

## 0.1 Simulation Controlling Underlying Data Generation Process

The fully simulated data simulation allows for insight into the behavior of the MLLT under correct and incorrect model specification. This simulation study also highlights shortcomings of independently fit LLT models when correlation exists in the observation or state equations. The data is simulated under three scenarios: 1.) correlation only exists in the observation equation (O), 2.) correlation only exists in the state equation (S), and 3.) correlation exists in both the observation and state equation (OS). The independent LLT, the O MLLT, S MLLT, and OS MLLT are all fit to each of the three data generation processes.

The covariates are randomly generated to mirror the predictors of interest (section \_\_\_\_\_). There is a time effect, a binary group effect, and a linear continuity point effect. For each simulation, 100 subjects are generated with between 2-12 observations. The “true” linear effects  $\beta$ , observation error covariance ( $\Sigma_\epsilon$ ), and underlying state process covariance ( $\Sigma_\eta$ ) are denoted in equation (\_\_\_\_\_).

The underlying data generation for each of the three simulation scenarios are as follows,

$$\begin{aligned} \begin{bmatrix} y_{ij1} \\ y_{ij2} \\ y_{ij3} \end{bmatrix} &= \begin{bmatrix} \alpha_{ij1} \\ \alpha_{ij2} \\ \alpha_{ij3} \end{bmatrix} + \begin{bmatrix} \mathbf{x}_{ij}\beta_1 \\ \mathbf{x}_{ij}\beta_2 \\ \mathbf{x}_{ij}\beta_3 \end{bmatrix} + \begin{bmatrix} \epsilon_{ij1} \\ \epsilon_{ij2} \\ \epsilon_{ij3} \end{bmatrix}, \quad \begin{bmatrix} \epsilon_{ij1} \\ \epsilon_{ij2} \\ \epsilon_{ij3} \end{bmatrix} \sim N(0, \Sigma_\epsilon) \\ \begin{bmatrix} \alpha_{ij1} \\ \alpha_{ij2} \\ \alpha_{ij3} \end{bmatrix} &= \begin{bmatrix} \alpha_{i(j-1)1} \\ \alpha_{i(j-1)2} \\ \alpha_{i(j-1)3} \end{bmatrix} + \begin{bmatrix} \eta_{ij1} \\ \eta_{ij2} \\ \eta_{ij3} \end{bmatrix}, \quad \begin{bmatrix} \eta_{ij1} \\ \eta_{ij2} \\ \eta_{ij3} \end{bmatrix} \sim N(0, \delta_{ij}\Sigma_\eta) \\ \beta &= [\beta_1 \quad \beta_2 \quad \beta_3] = \begin{bmatrix} 4 & -3 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \end{aligned}$$

for  $i \in \{1, 2, \dots, 100\}$  and  $j \in \{2, 3, \dots, 12\}$ . The parameters in  $\beta$  were chosen to have no, small, medium, large, and negative effects. For the three different data generation scenarios we utilize differing values of  $\Sigma_\epsilon$  and  $\Sigma_\eta$ .

### 0.1.0.1 O Model

$$\Sigma_\epsilon = \begin{bmatrix} 15 & 2.4 & 1 \\ 2.4 & 15 & 1 \\ 1 & 1 & 10 \end{bmatrix}, \quad \Sigma_\eta = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

### 0.1.0.2 S Model

$$\Sigma_\epsilon = \begin{bmatrix} 15 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 10 \end{bmatrix}, \quad \Sigma_\eta = \begin{bmatrix} 5 & 3.7 & 0 \\ 3.7 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

### 0.1.0.3 OS Model

$$\Sigma_\epsilon = \begin{bmatrix} 15 & 2.4 & 1 \\ 2.4 & 15 & 1 \\ 1 & 1 & 10 \end{bmatrix}, \quad \Sigma_\eta = \begin{bmatrix} 5 & 3.7 & 0 \\ 3.7 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

The covariance matrices  $\Sigma_\epsilon$  and  $\Sigma_\eta$  were chosen to approximate values observed when the MLLT models were fit to real data.

For each scenario, the simulations are carried out 1000 times. The most effective model is one that is unbiased, maintains 95% coverage, and small parameter variance (indicated by small confidence interval length) for the parameters in  $\beta$ ,  $\Sigma_\epsilon$ , and  $\Sigma_\eta$ . Maintaining proper 95% coverage indicates proper type I error at the level of 0.05. If we can fix type I error, the next step is to minimize type II error which then leads to greater power to detect significant differences. Minimizing type II error will occur by minimizing the

parameter variance if the estimates are unbiased and have proper 95% coverage. As we are using a Bayesian Gibb’s sampling approach, the confidence interval is used to judge parameter variance.

The Bayesian Gibb’s sampler is repeated 5000 times with a burn-in of 2000 for each model. This means samples 2001-5000 are used for parameter inference.

### 0.1.1 Fully Simulated Results

The LLT, O, S, and OS models all show unbiasedness and near 95% coverage in the linear effect parameters in all three simulation scenarios. This occurs despite the LLT, O, and S models having a level of misspecification in the covariance parameters in the different scenarios. In the linear effects there is not a large difference in confidence interval length between the different methods. The O, S, and OS models maintain performance when compared to the LLT with respect to linear effects. However, the primary motivation of the MLLT models are to gain inference on correlation between cognitive tests.

The parameters in  $\Sigma_\epsilon$  and  $\Sigma_\eta$  vary much more across the data generation processes. When generating from the O model the O and OS estimation maintains unbiased estimates and 95% coverage. Unsurprisingly, the misspecified S estimation provides erroneous estimates and poor coverage for the covariance matrices. A similar pattern arises when generating from the S model. The correctly specified S and OS models show near unbiasedness and 95% coverage. The coverage does lack in some, most likely due to identifiability issues, but the estimates are still in a close neighborhood of the correct values. As the primary interest is identifying underlying constructs, and not hypothesis testing, these results are promising.

For the OS data generation processes, the LLT and the S estimation models assume there is not any covariance in the observation error, therefore, do not seek to estimate the non-diagonal values of  $\Sigma_\epsilon$ . Even so, the LLT and S models are fairly accurate in estimating the observation variances. The O estimation model, which does assume observation error covariance, greatly over estimates the covariance parameters. This occurs because the O model assumes no covariance in the state equation, therefore, any correlation in the state equation ends up being allocated to the observation error. Additionally, we see the same principle for the S model in estimating  $\Sigma_\eta$ . The covariance parameters of  $\Sigma_\eta$  are inflated as model S is assuming correlation truly occurring in  $\Sigma_\epsilon$  is actually occurring in the underlying cognitive process. The OS model, which is correctly specified, unsurprisingly accurately estimates the  $\Sigma_\epsilon$  and  $\Sigma_\eta$ .

The OS model estimates  $K(K - 1)/2$  more parameters than the O and S models. Even with the extra parameters, the OS model estimation performed as well as the correctly specified O and S models respectively. However, if O or S are misspecified it can lead to miscalculation in the covariance matrices. The OS model is much more robust in terms of handling different observation error and cognitive process correlation.

Shortcomings are also blatantly evident as the LLT provides no observation error or cognitive process correlation estimation. The O and S models do provide some insight into the inter-relatedness between cognition tests, but the OS provides the most descriptive form of the correlations. With the added covariance estimation benefits of the OS MLLT, it perform just as well as the LLT in modeling linear effects.

#### 0.1.1.1 O Model

Table 1: Linear Effect Coverage

Test	Variable	Beta	LLT	O	S	OS
Y1	X1	4	0.964	0.965	0.969	0.963
Y1	X2	2	0.965	0.964	0.958	0.957
Y1	X3	1	0.953	0.953	0.955	0.954
Y2	X1	-3	0.940	0.948	0.945	0.943
Y2	X2	0	0.955	0.954	0.952	0.962
Y2	X3	1	0.937	0.932	0.942	0.938
Y3	X1	0	0.957	0.950	0.954	0.952

Y3	X2	0	0.948	0.949	0.950	0.956
Y3	X3	0	0.950	0.950	0.949	0.947

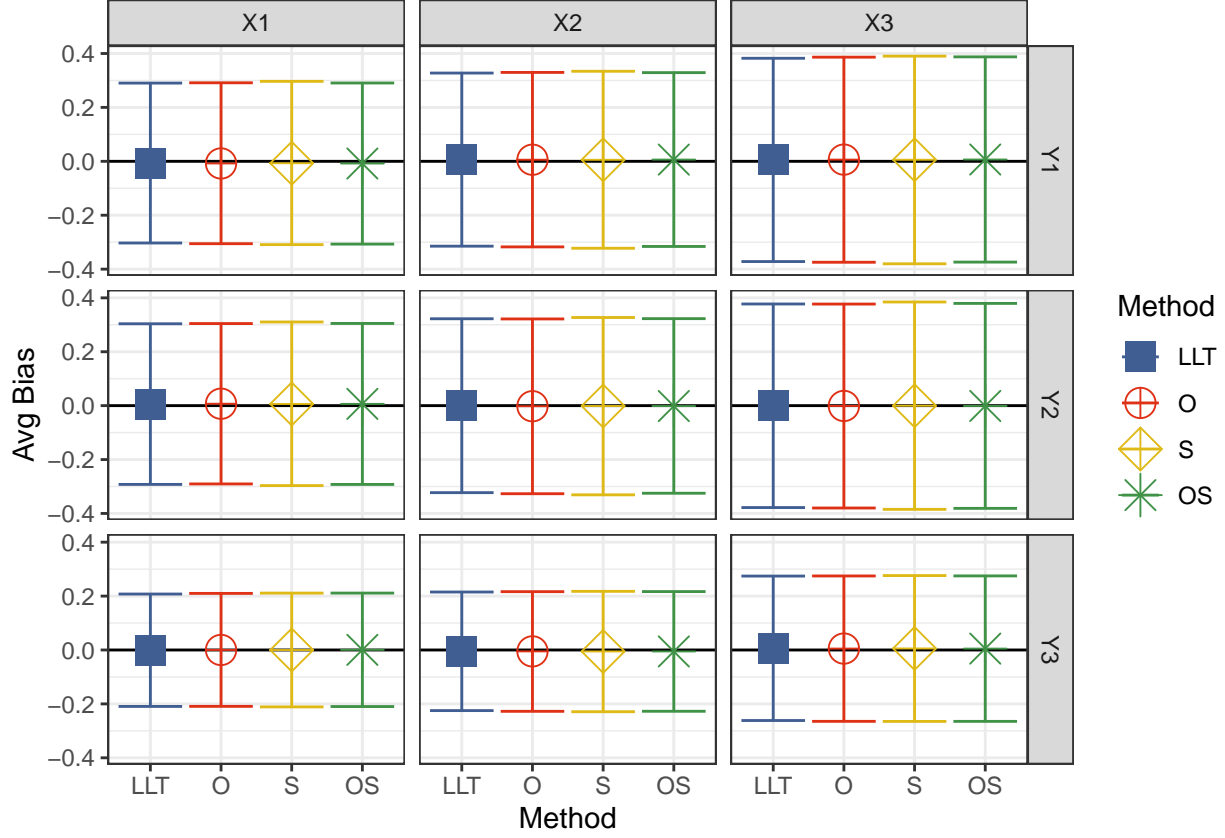


Figure 1: Parameter Bias and Variability

Table 2: Observation Covariance Estimates

param	true	LLT	O	S	OS
1,1	15.0	14.909	14.919	14.688	14.950
1,2	2.4	-	2.371	-	2.425
2,2	15.0	14.892	14.903	14.668	14.927
1,3	1.0	-	0.990	-	1.027
2,3	1.0	-	0.988	-	1.008
3,3	10.0	10.056	10.063	10.021	10.093

Table 3: Observation Covariance Coverage

param	LLT	O	S	OS
1,1	0.95	0.951	0.928	0.948
1,2	-	0.953	-	0.941
2,2	0.944	0.949	0.934	0.946
1,3	-	0.956	-	0.956

2,3	-	0.952	-	0.947
3,3	0.951	0.945	0.947	0.951

Table 4: State Covariance Estimates

param	true	LLT	O	S	OS
1,1	5	4.89	4.969	5.206	4.985
1,2	0	-	-	0.987	-0.071
2,2	5	4.887	4.964	5.205	4.988
1,3	0	-	-	0.373	-0.042
2,3	0	-	-	0.386	-0.021
3,3	2	1.893	1.932	1.973	1.932

Table 5: State Covariance Coverage

param	LLT	O	S	OS
1,1	0.935	0.948	0.938	0.945
1,2	-	-	0.500	0.939
2,2	0.948	0.952	0.955	0.951
1,3	-	-	0.815	0.942
2,3	-	-	0.783	0.933
3,3	0.92	0.936	0.930	0.934

#### 0.1.1.2 S Model

Table 6: Linear Effect Coverage

Test	Variable	Beta	LLT	O	S	OS
Y1	X1	4	0.954	0.956	0.957	0.959
Y1	X2	2	0.946	0.951	0.950	0.956
Y1	X3	1	0.959	0.957	0.958	0.957
Y2	X1	-3	0.957	0.965	0.955	0.961
Y2	X2	0	0.941	0.943	0.948	0.951
Y2	X3	1	0.953	0.957	0.955	0.957
Y3	X1	0	0.945	0.948	0.946	0.945
Y3	X2	0	0.950	0.954	0.949	0.958
Y3	X3	0	0.938	0.942	0.939	0.938

Table 7: Observation Covariance Estimates

param	true	LLT	O	S	OS
1,1	15	14.866	15.248	15.147	15.065
1,2	0	-	2.436	-	-0.387
2,2	15	14.908	15.293	15.18	15.092
1,3	0	-	0.012	-	0.006
2,3	0	-	-0.020	-	-0.033
3,3	10	10.06	10.064	10.087	10.102

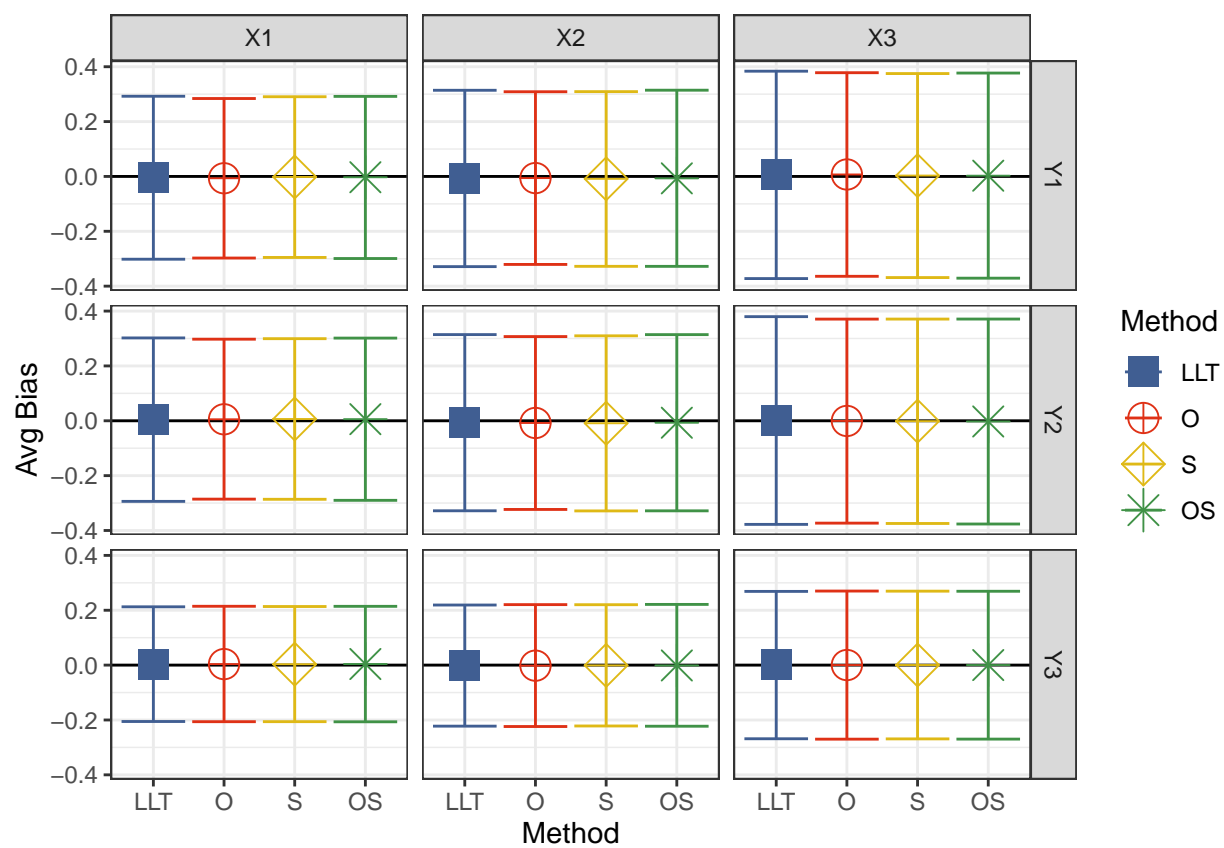


Figure 2: Parameter Bias and Variability

Table 8: Observation Covariance Coverage

param	LLT	O	S	OS
1,1	0.949	0.949	0.938	0.935
1,2	-	0.029	-	0.904
2,2	0.95	0.935	0.936	0.935
1,3	-	0.959	-	0.955
2,3	-	0.952	-	0.946
3,3	0.945	0.948	0.948	0.947

Table 9: State Covariance Estimates

param	true	LLT	O	S	OS
1,1	5.000	4.879	4.59	4.803	4.940
1,2	3.714	-	-	3.786	3.980
2,2	5.000	4.899	4.606	4.838	4.977
1,3	0.000	-	-	0.017	0.015
2,3	0.000	-	-	0.003	0.015
3,3	2.000	1.89	1.932	1.916	1.928

Table 10: State Covariance Coverage

param	LLT	O	S	OS
1,1	0.932	0.908	0.922	0.930
1,2	-	-	0.946	0.918
2,2	0.937	0.902	0.913	0.911
1,3	-	-	0.951	0.935
2,3	-	-	0.950	0.933
3,3	0.927	0.937	0.922	0.927

### 0.1.1.3 OS Model

Table 11: Linear Effect Coverage

Test	Variable	Beta	LLT	O	S	OS
Y1	X1	4	0.962	0.948	0.968	0.961
Y1	X2	2	0.946	0.934	0.955	0.945
Y1	X3	1	0.952	0.941	0.958	0.950
Y2	X1	-3	0.950	0.944	0.958	0.957
Y2	X2	0	0.945	0.931	0.959	0.952
Y2	X3	1	0.949	0.940	0.968	0.964
Y3	X1	0	0.957	0.962	0.960	0.957
Y3	X2	0	0.950	0.947	0.949	0.952
Y3	X3	0	0.945	0.948	0.949	0.949

Table 12: Observation Covariance Estimates

param	true	LLT	O	S	OS
1,1	15.0	14.964	15.752	14.273	15.075
1,2	2.4	-	4.976	-	2.140

2,2	15.0	14.948	15.733	14.248	15.066
1,3	1.0	-	0.997	-	0.979
2,3	1.0	-	1.008	-	0.997
3,3	10.0	10.021	10.021	9.988	10.056

Table 13: Observation Covariance Coverage

param	LLT	O	S	OS
1,1	0.944	0.882	0.867	0.943
1,2	-	0.024	-	0.917
2,2	0.941	0.879	0.851	0.938
1,3	-	0.946	-	0.942
2,3	-	0.948	-	0.926
3,3	0.937	0.935	0.937	0.931

Table 14: State Covariance Estimates

param	true	LLT	O	S	OS
1,1	5.000	4.835	4.168	5.776	4.941
1,2	3.714	-	-	5.007	3.891
2,2	5.000	4.849	4.183	5.800	4.952
1,3	0.000	-	-	0.470	-0.009
2,3	0.000	-	-	0.470	-0.015
3,3	2.000	1.894	1.943	1.977	1.939

Table 15: State Covariance Coverage

param	LLT	O	S	OS
1,1	0.94	0.78	0.828	0.942
1,2	-	-	0.447	0.937
2,2	0.924	0.789	0.820	0.927
1,3	-	-	0.719	0.938
2,3	-	-	0.738	0.939
3,3	0.926	0.938	0.930	0.930

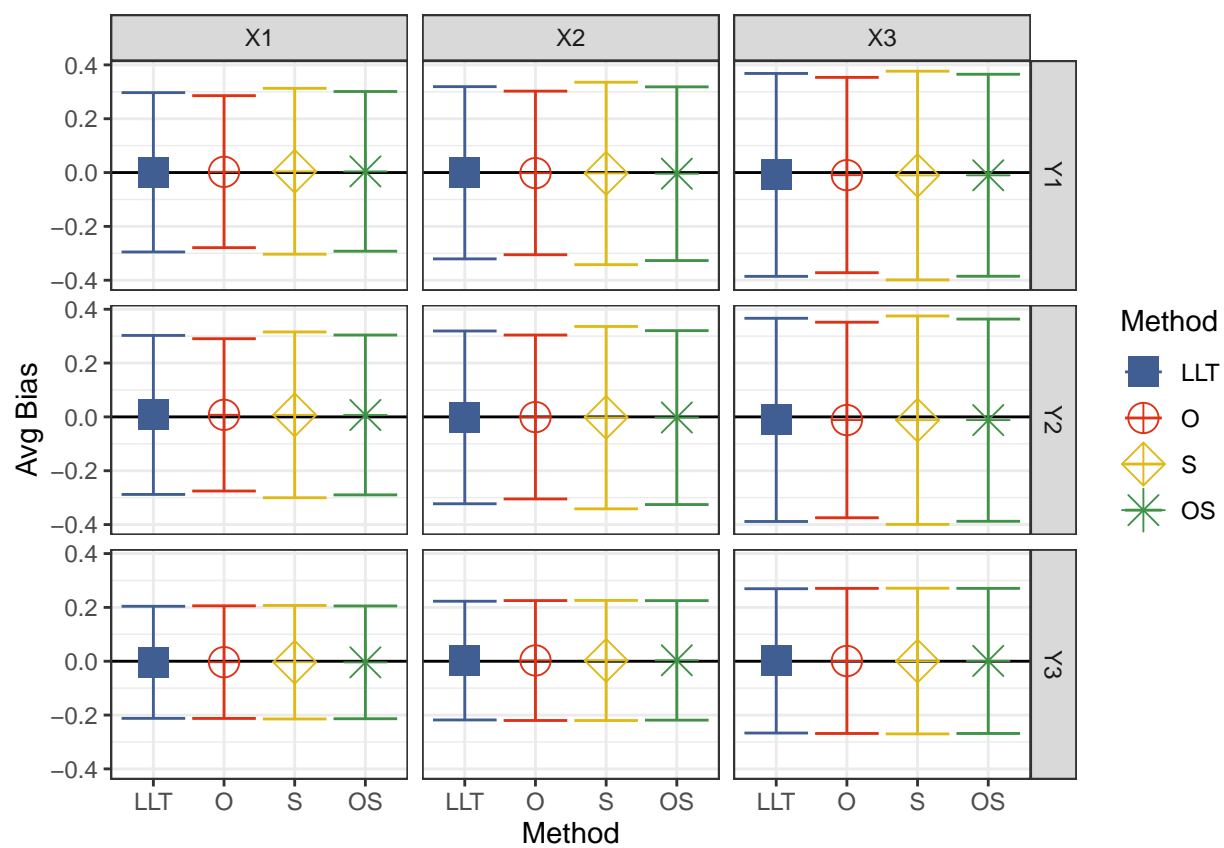


Figure 3: Parameter Bias and Variability