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 $\boldsymbol{\beta}$, observation error covariance (Σ_ε), and underlying state process covariance (Σ_η) are denoted in equation (________________).

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

$$\begin{bmatrix} y_{ij1} \\ y_{ij2} \\ y_{ij3} \end{bmatrix} = \begin{bmatrix} \alpha_{ij1} \\ \alpha_{ij2} \\ \alpha_{ij3} \end{bmatrix} + \begin{bmatrix} \boldsymbol{x_{ij}\beta_1} \\ \boldsymbol{x_{ij}\beta_2} \\ \boldsymbol{x_{ij}\beta_3} \end{bmatrix} + \begin{bmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \\ \varepsilon_{ij3} \end{bmatrix}, \quad \begin{bmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \\ \varepsilon_{ij3} \end{bmatrix} \sim N(0, \Sigma_\varepsilon)$$

$$\begin{bmatrix} \alpha_{ij1} \\ \alpha_{ij2} \\ \alpha_{ij3} \end{bmatrix} = \begin{bmatrix} \alpha_{i(j-1)1} \\ \alpha_{i(j-1)2} \\ \alpha_{i(j-1)3} \end{bmatrix} + \begin{bmatrix} \eta_{ij1} \\ \eta_{ij2} \\ \eta_{ij3} \end{bmatrix}, \quad \begin{bmatrix} \eta_{ij1} \\ \eta_{ij2} \\ \eta_{ij3} \end{bmatrix} \sim N(0, \delta_{ij}\Sigma_\eta)$$

$$\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta_1} & \boldsymbol{\beta_2} & \boldsymbol{\beta_3} \end{bmatrix} = \begin{bmatrix} 4 & -3 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

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

0.1.0.1 O Model

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

0.1.0.2 S Model

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

0.1.0.3 OS Model

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

The covariance matrices $\boldsymbol{\Sigma_\varepsilon}$ and $\boldsymbol{\Sigma_\eta}$ 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 $\boldsymbol{\beta}$, $\boldsymbol{\Sigma_\varepsilon}$, and $\boldsymbol{\Sigma_\eta}$. 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 occurs 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. This occurs despite the LLT, O, and S models having a level of misspecification in the covariance parameters. In the linear effects there is not a large difference in confidence interval length between the different methods. The O, S, and OS models measure up well to the LLT.

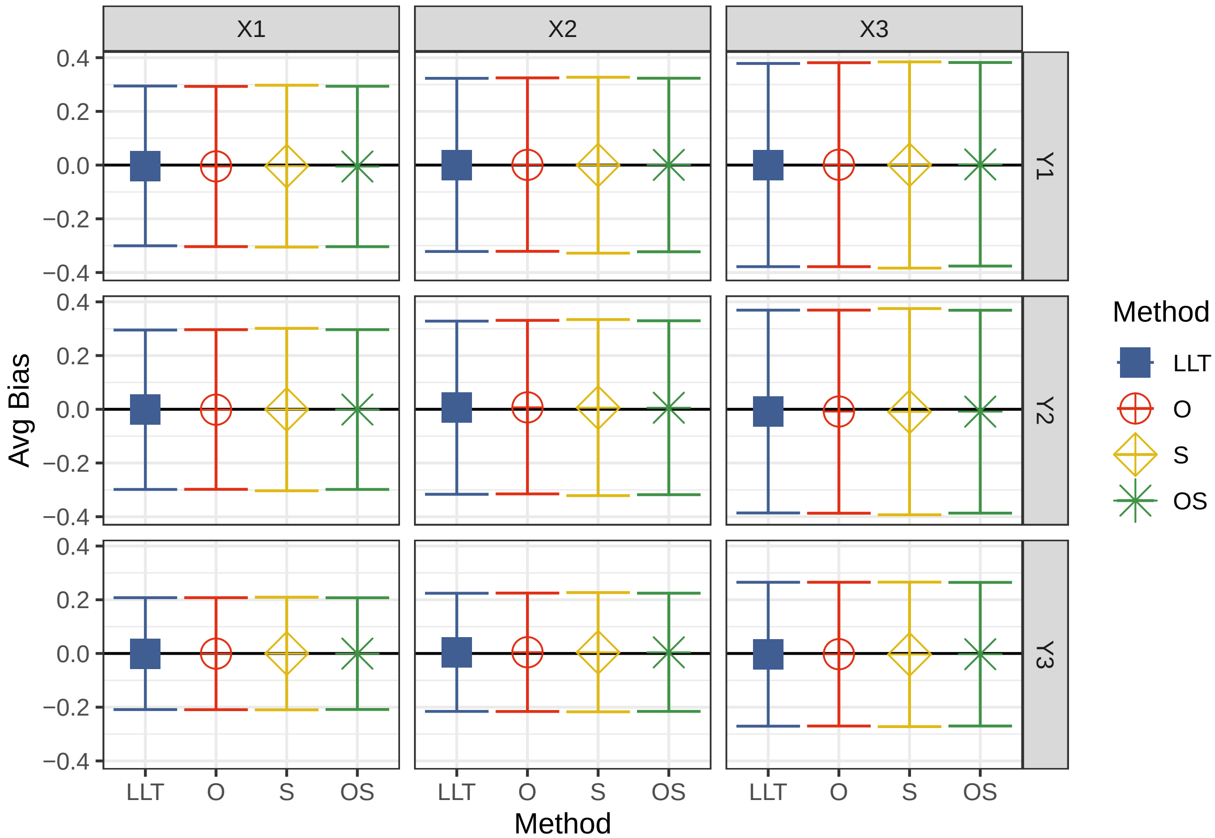
The parameters in Σ_ε vary much more. The LLT and the S models assume there is not any covariance in the observation error, therefore, do not seek to estimate the non-diagonal values of Σ_ε. Even so, the LLT and S models are fairly accurate in estimating the observation variances. The O model, which does assume observation error covariance, greatly over estimates the covariance parameters. This is 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 Σ_η. The covariance parameters of Σ_η are inflated as model S is assuming correlation truly occurring in Σ_ε is actually occurring in the underlying cognitive process.

The OS model, which is correctly specified, unsurprisingly accurately estimates the Σ_ε and Σ_η. Similar studies were conducted, except O or S were correctly specified. When either O or S were correctly specified they slightly outperformed the OS model. This is because the OS model estimates $K(K-1)/2$ more parameters than the other models. But, as can be seen from this simulation analysis, 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. Even with the added benefits of the MLLT, it perform just as well as the LLT in modeling linear effects.

0.1.1.1 O Model

Test	Variable	Beta	LLT	O	S	OS
Y1	X1	4	0.948	0.943	0.956	0.947
Y1	X2	2	0.943	0.943	0.946	0.949
Y1	X3	1	0.954	0.953	0.955	0.956
Y2	X1	-3	0.963	0.956	0.961	0.956
Y2	X2	0	0.944	0.945	0.951	0.949
Y2	X3	1	0.960	0.953	0.960	0.956
Y3	X1	0	0.966	0.960	0.956	0.959
Y3	X2	0	0.937	0.936	0.933	0.935
Y3	X3	0	0.949	0.944	0.944	0.949



param	true	LLT	O	S	OS
1,1	15.0	14.894	14.918	14.676	14.936
1,2	2.4	-	2.380	-	2.416
2,2	15.0	14.881	14.903	14.666	14.926
1,3	1.0	-	0.997	-	1.019
2,3	1.0	-	1.013	-	1.038
3,3	10.0	10.013	10.030	9.987	10.068

param	true	LLT sq diff	O sq diff	S sq diff	OS sq diff
1,1	15.0	0.011	0.007	0.105	0.004
1,2	2.4	5.760	0.000	5.760	0.000
2,2	15.0	0.014	0.009	0.112	0.005
1,3	1.0	1.000	0.000	1.000	0.000
2,3	1.0	1.000	0.000	1.000	0.001
3,3	10.0	0.000	0.001	0.000	0.005

param	LLT	O	S	OS
1,1	0.959	0.958	0.941	0.957
1,2	-	0.944	-	0.946
2,2	0.94	0.940	0.921	0.938
1,3	-	0.961	-	0.956
2,3	-	0.941	-	0.933
3,3	0.945	0.949	0.947	0.951

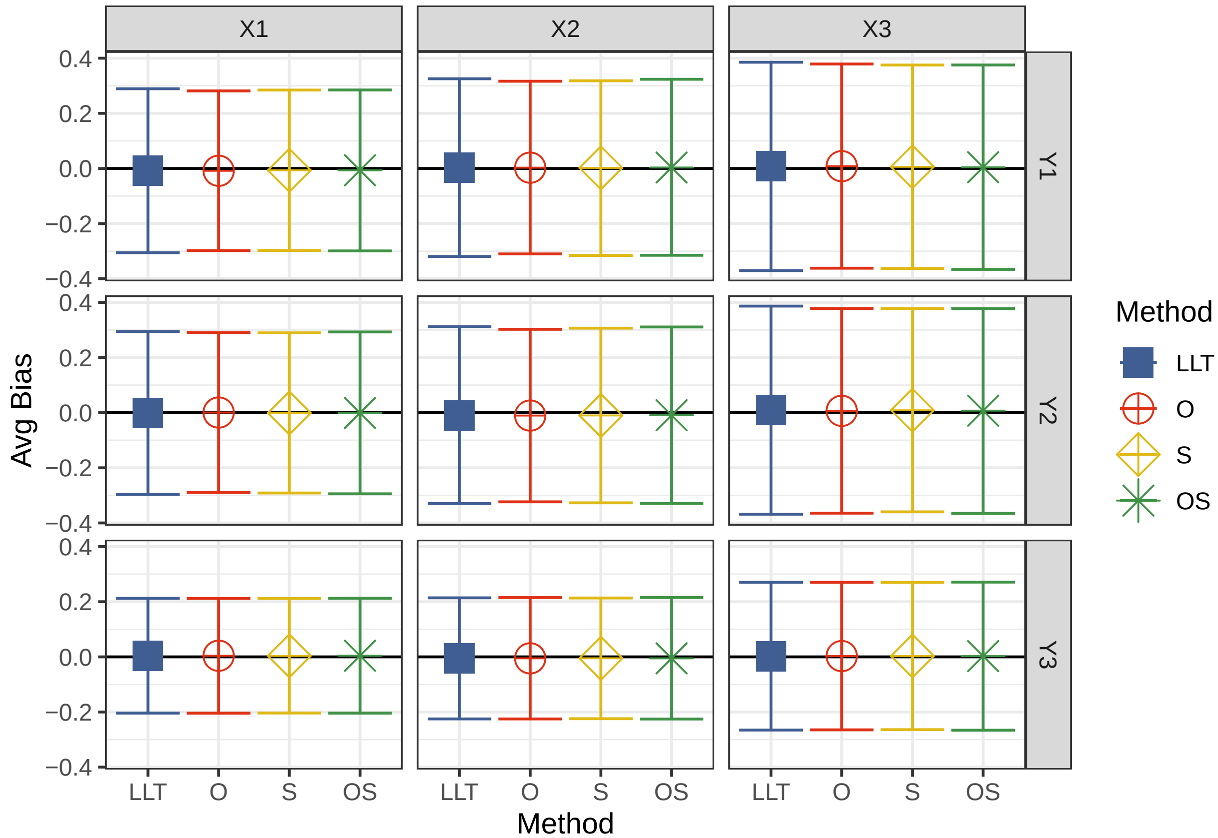
param	true	LLT	O	S	OS
1,1	5	4.91	4.932	5.181	4.960
1,2	0	-	-	1.012	-0.043
2,2	5	4.887	4.909	5.153	4.934
1,3	0	-	-	0.391	-0.022
2,3	0	-	-	0.391	-0.027
3,3	2	1.887	1.895	1.936	1.889

param	true	LLT sq diff	O sq diff	S sq diff	OS sq diff
1,1	5	0.008	0.005	0.033	0.002
1,2	0	0.000	0.000	1.024	0.002
2,2	5	0.013	0.008	0.023	0.004
1,3	0	0.000	0.000	0.153	0.000
2,3	0	0.000	0.000	0.153	0.001
3,3	2	0.013	0.011	0.004	0.012

param	LLT	O	S	OS
1,1	0.949	0.945	0.938	0.949
1,2	-	-	0.479	0.939
2,2	0.928	0.931	0.927	0.929
1,3	-	-	0.794	0.955
2,3	-	-	0.791	0.936
3,3	0.921	0.923	0.927	0.922

0.1.1.2 S Model

Test	Variable	Beta	LLT	O	S	OS
Y1	X1	4	0.949	0.947	0.948	0.941
Y1	X2	2	0.948	0.941	0.945	0.944
Y1	X3	1	0.958	0.955	0.954	0.960
Y2	X1	-3	0.954	0.950	0.955	0.949
Y2	X2	0	0.938	0.938	0.936	0.941
Y2	X3	1	0.952	0.949	0.948	0.950
Y3	X1	0	0.955	0.962	0.951	0.957
Y3	X2	0	0.942	0.943	0.940	0.948
Y3	X3	0	0.966	0.961	0.962	0.966



param	true	LLT	O	S	OS
1,1	15	14.88	15.277	15.24	15.170
1,2	0	-	2.466	-	-0.402
2,2	15	14.921	15.315	15.271	15.191
1,3	0	-	0.007	-	-0.007
2,3	0	-	0.000	-	-0.005
3,3	10	10.047	10.069	10.098	10.115

param	true	LLT sq diff	O sq diff	S sq diff	OS sq diff
1,1	15	0.014	0.077	0.058	0.029
1,2	0	0.000	6.081	0.000	0.162
2,2	15	0.006	0.099	0.073	0.036
1,3	0	0.000	0.000	0.000	0.000
2,3	0	0.000	0.000	0.000	0.000
3,3	10	0.002	0.005	0.010	0.013

param	LLT	O	S	OS
1,1	0.929	0.931	0.911	0.916
1,2	-	0.024	-	0.893
2,2	0.951	0.933	0.901	0.923
1,3	-	0.937	-	0.931
2,3	-	0.959	-	0.931
3,3	0.958	0.959	0.959	0.951

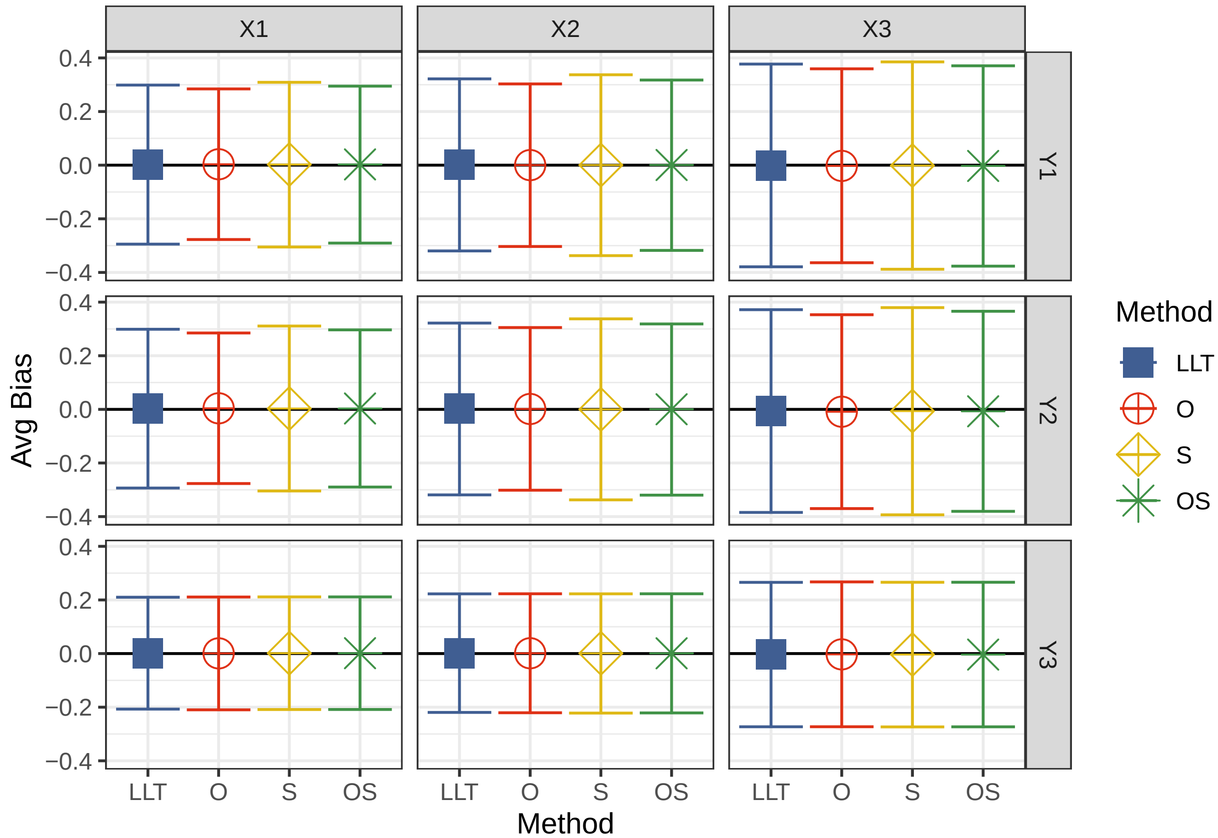
param	true	LLT	O	S	OS
1,1	5.000	4.866	4.521	4.702	4.839
1,2	3.714	-	-	3.758	3.956
2,2	5.000	4.869	4.528	4.712	4.857
1,3	0.000	-	-	0.012	0.019
2,3	0.000	-	-	0.004	0.008
3,3	2.000	1.876	1.878	1.858	1.877

param	true	LLT sq diff	O sq diff	S sq diff	OS sq diff
1,1	5.000	0.018	0.229	0.089	0.026
1,2	3.714	13.794	13.794	0.002	0.059
2,2	5.000	0.017	0.223	0.083	0.020
1,3	0.000	0.000	0.000	0.000	0.000
2,3	0.000	0.000	0.000	0.000	0.000
3,3	2.000	0.015	0.015	0.020	0.015

param	LLT	O	S	OS
1,1	0.932	0.884	0.888	0.898
1,2	-	-	0.933	0.908
2,2	0.945	0.9	0.879	0.901
1,3	-	-	0.930	0.931
2,3	-	-	0.934	0.941
3,3	0.924	0.924	0.920	0.913

0.1.1.3 OS Model

Test	Variable	Beta	LLT	O	S	OS
Y1	X1	4	0.957	0.939	0.960	0.953
Y1	X2	2	0.943	0.922	0.951	0.944
Y1	X3	1	0.957	0.948	0.968	0.962
Y2	X1	-3	0.961	0.948	0.968	0.963
Y2	X2	0	0.943	0.923	0.953	0.944
Y2	X3	1	0.949	0.944	0.963	0.955
Y3	X1	0	0.952	0.952	0.951	0.958
Y3	X2	0	0.948	0.946	0.944	0.942
Y3	X3	0	0.944	0.948	0.950	0.949



param	true	LLT	O	S	OS
1,1	15.0	14.924	15.728	14.32	15.095
1,2	2.4	-	5.022	-	2.174
2,2	15.0	14.913	15.709	14.304	15.065
1,3	1.0	-	1.022	-	1.016
2,3	1.0	-	1.009	-	1.002
3,3	10.0	9.985	10.001	9.974	10.042

param	true	LLT sq diff	O sq diff	S sq diff	OS sq diff
1,1	15.0	0.006	0.530	0.462	0.009
1,2	2.4	5.760	6.875	5.760	0.051
2,2	15.0	0.008	0.503	0.484	0.004
1,3	1.0	1.000	0.000	1.000	0.000
2,3	1.0	1.000	0.000	1.000	0.000
3,3	10.0	0.000	0.000	0.001	0.002

param	LLT	O	S	OS
1,1	0.936	0.890	0.85	0.923
1,2	-	0.015	-	0.934
2,2	0.95	0.891	0.846	0.933
1,3	-	0.956	-	0.945
2,3	-	0.953	-	0.951
3,3	0.945	0.948	0.947	0.940

param	true	LLT	O	S	OS
1,1	5.000	4.87	4.146	5.715	4.892
1,2	3.714	-	-	4.984	3.855
2,2	5.000	4.852	4.137	5.713	4.896
1,3	0.000	-	-	0.465	-0.023
2,3	0.000	-	-	0.464	-0.027
3,3	2.000	1.922	1.934	1.961	1.928

param	true	LLT sq diff	O sq diff	S sq diff	OS sq diff
1,1	5.000	0.017	0.729	0.511	0.012
1,2	3.714	13.794	13.794	1.613	0.020
2,2	5.000	0.022	0.745	0.508	0.011
1,3	0.000	0.000	0.000	0.216	0.001
2,3	0.000	0.000	0.000	0.215	0.001
3,3	2.000	0.006	0.004	0.002	0.005

param	LLT	O	S	OS
1,1	0.931	0.769	0.823	0.927
1,2	-	-	0.435	0.929
2,2	0.931	0.758	0.809	0.915
1,3	-	-	0.730	0.934
2,3	-	-	0.736	0.929
3,3	0.943	0.943	0.939	0.935