An Adaptive Bayesian Lasso Approach with Spike-and-Slab Priors to Identify Multiple Linear and Nonlinear Effects in Structural Equation Models

Supplemental material

Holger Brandt, Jenna Cambria, and Augustin Kelava

April 23, 2018

1 Supplement to Introduction

Empirical example

In this supplement, we provide details on the empirical example from the introduction section. Data were obtained from the Trends in International Mathematics and Science Study 2015 (TIMMS; Mullis & Martin, 2013) which is publicly available for research purposes under

`https://timssandpirls.bc.edu/timss2015/international-database/`.

For this example, we used the US subsample with $N = 10,163$ students from grade 8 who took part in a mathematical skill test and provided information about their expectancies and values with regard to mathematics. For the prediction of the math skills, measures of students' expectancies of success ($BSBM19$), their intrinsic motivation ($BSBM17$), their utility and their attainment values ($BSBM20$; items 6, 8, and 9 measure attainment values) were selected. Math skills were operationalized using three plausible values for geometry, algebra, and number knowledge ($BSMGEO01$, $BSMALG01$, $BSMNUM01$).

For each of the 3 latent predictor variables, the items were aggregated to three indicator variables for each construct (item parcels). The clustering of the math skills due to the school membership was considered in the model by using a random intercept for the latent math skill factor. Only complete cases were included in the sampling process ($N = 9912$).

A cfa was conducted to test the appropriateness of the measurement models. Although the χ^2 test was large with $\chi^2 = 5803, df = 80$, other fit indices indicated good to acceptable fit (CFI=.96, TLI=.95, RMSEA=.09,SRMR=.05). Standardized factor loadings lay between .62 and .97, with lower values for attainment values (since only three items were available, no items parcels were constructed for attainment value) and higher values for math. Correlations between the predictor variables lay between .30 and .79.

From this original data set, 100 random samples of $N = 400$ were drawn and analyzed.

Model formulation For the measurement model, we used a factor loading matrix $\mathbf{\Lambda}_x$ that followed a simple structure, i.e., each item loaded only on one latent variable (cf. Eq. (24) in the main document). We specified three different structural models that included all 2, 3, or 4 latent predictor variables simultaneously as well as all 1, 3, or 6 (two-way) interactions and 2, 3, or 4 quadratic effects [cf. Eq. (23)].

For the analysis of the model we used the UPI, the LMS, and the aBSS-lasso approaches. All priors and implementation aspects (burn-in etc.) were identical to the specification provided in the simulation study.

Results Table 1 provides the information on average SD's reported in the introduction of the main text. The results for the regression coefficients are illustrated in Table 2. For each model, the median estimates, and their standard deviations are depicted. For LMS and aBSS lasso, all data sets converged. For UPI the model with four predictor variables resulted in only 87 converged data sets.

Table 1: Median, minimum and maximum standard deviations of parameter estimates across 100 random samples (the models included 3, 6, and 10 nonlinear effects, respectively).

	median			min			max		
no of predictors	2	3	4	2	3	4	2	3	4
UPI	0.23	0.24	0.86	0.10	0.16	0.33	0.24	0.33	2.12
LMS	0.19	0.17	0.25	0.06	0.10	0.10	0.20	0.25	0.50
aBSS	0.10	0.08	0.09	0.04	0.05	0.05	0.10	0.10	0.12

Table 2: Results of the small simulation on the TIMSS data set. Enumeration of the coefficients follows Eq. (24).

	UPI			LMS			aBSS		
no. of predictors	2	3	4	2	3	4	2	3	4
					Median				
$\gamma_{1,1}$	0.89	0.91	0.92	0.87	0.88	0.90	0.80	0.82	0.83
$\gamma_{1,2}$	-0.12	-0.13	-0.16	-0.10	-0.11	-0.13	-0.04	-0.04	-0.06
$\gamma_{1,3}$	–	-0.01	-0.21	–	-0.02	-0.12	–	-0.02	-0.04
$\gamma_{1,4}$	–	–	0.09	–	–	0.15	–	–	0.04
$\gamma_{2,1}$	0.00	-0.02	-0.04	-0.01	0.03	0.05	0.04	0.03	0.03
$\gamma_{2,2}$	0.29	0.21	0.20	0.24	0.18	0.14	0.09	0.07	0.06
$\gamma_{2,3}$	–	0.06	0.15	–	0.08	0.10	–	0.07	0.05
$\gamma_{2,4}$	–	–	-0.06	–	–	0.02	–	–	0.03
$\gamma_{2,5}$	-0.17	-0.15	-0.10	-0.11	-0.11	-0.08	-0.03	-0.03	-0.03
$\gamma_{2,6}$	–	0.08	0.05	–	0.03	0.05	–	-0.00	0.01
$\gamma_{2,7}$	–	–	0.03	–	–	-0.06	–	–	-0.00
$\gamma_{2,8}$	–	-0.15	-0.41	–	-0.07	-0.20	–	-0.04	-0.03
$\gamma_{2,9}$	–	–	0.76	–	–	0.43	–	–	0.05
$\gamma_{2,10}$	–	–	-0.48	–	–	-0.33	–	–	-0.11
					SD				
$\gamma_{1,1}$	0.13	0.13	0.19	0.12	0.12	0.12	0.11	0.11	0.11
$\gamma_{1,2}$	0.07	0.09	0.16	0.07	0.08	0.08	0.06	0.06	0.06
$\gamma_{1,3}$	–	0.09	0.36	–	0.08	0.15	–	0.05	0.08
$\gamma_{1,4}$	–	–	0.42	–	–	0.18	–	–	0.09
$\gamma_{2,1}$	0.24	0.27	0.38	0.19	0.21	0.23	0.10	0.10	0.10
$\gamma_{2,2}$	0.23	0.33	0.53	0.20	0.25	0.25	0.10	0.09	0.09
$\gamma_{2,3}$	–	0.26	1.15	–	0.19	0.31	–	0.10	0.09
$\gamma_{2,4}$	–	–	0.99	–	–	0.36	–	–	0.09
$\gamma_{2,5}$	0.10	0.16	0.33	0.06	0.10	0.10	0.04	0.05	0.05
$\gamma_{2,6}$	–	0.22	0.73	–	0.15	0.25	–	0.06	0.07
$\gamma_{2,7}$	–	–	0.70	–	–	0.26	–	–	0.08
$\gamma_{2,8}$	–	0.16	1.25	–	0.10	0.29	–	0.07	0.07
$\gamma_{2,9}$	–	–	2.12	–	–	0.50	–	–	0.12
$\gamma_{2,10}$	–	–	1.17	–	–	0.23	–	–	0.10

Table 3: Articles with simulation studies on latent interaction and quadratic effects in SEM since 2000.

Authors	# effects	Approaches	Other
Brandt et al. (2014)	3	UPI, LMS, QML, 2SMM, MM	Nonnormality
Cham et al. (2012)	1	UPI, GAPI, LMS	Nonnormality
Cham et al. (2017)	1	UPI, LMS	Missing data
Chen and Cheng (2014)	1	PI	Complexer linear model
Foldnes and Hagtvet (2014)	1	PI	No of PI's
Gerhard, Klein, et al. (2014)	2	LMS	Model difference test, nonnormality
Gerhard, Büchner, et al. (2014)	2	LMS	Model test, heteroskedasticity
Harring et al. (2012)	1	PI, LMS, Bayes	
Jackman et al. (2011)	1	UPI	No of PI's
Kelava et al. (2008)	3	LMS, UPI, Ping	
Kelava et al. (2014)	3	LMS, UPI, NSEMM	Nonnormality
Kelava and Nagengast (2012)	3	LMS, Bayes	Nonnormality
Klein et al. (2009)	3	LMS	Misspecification
Klein and Moosbrugger (2000)	1	LMS, PI, 2SLS	
Klein and Muthén (2007)	1	LMS, QML, PI	Nonnormality
Lee (2007)	3	Bayes	Prior sensitivity
Leite and Zuo (2011)	1	PI	MLM
Li et al. (2000)	2	PI	LGM
Little et al. (2006)	1	PI	Orthogonalization
Lyhagen (2007)	1	MM	Moment method
Marsh et al. (2004)	1	PI, QML	Nonnormality
Marsh et al. (2007)	1	PI, QML	Nonnormality
Maslowsky et al. (2015)	1	LMS	
Mooijaart and Bentler (2010)	1	MM, LMS	Moment based approach
Moulder and Algina (2002)	1	PI, 2SLS/2SML	
Sun and Willson (2009)	1	LMS	LGM
Wall and Amemiya (2000)	1	2SMM, PI	polynomial effects
Wall and Amemiya (2001)	1	PI	Moment based approach
Wall and Amemiya (2000)	1	2SMM, PI	nonnormality

PI – Product Indicator Approaches based on Kenny and Judd (1984); GAPI – Generalized Appended PI (Wall & Amemiya, 2001); UPI – Unconstrained PI (Marsh et al., 2004); LMS – latent moderated structures (Klein & Moosbrugger, 2000); QML – Quasi Maximum Likelihood estimator (Klein & Muthén, 2007); 2SMM – 2 stage method of moments (Wall & Amemiya, 2003); Ping – Ping (1998)'s approach; Bayes – Bayesian approaches; MM – method of moments (Mooijaart & Bentler, 2010; Lyhagen, 2007)

2 Supplement to Model Section

Bayesian model specification

The Bayesian hierarchical model specification of the lasso is given by

$$\mathbf{x}|\boldsymbol{\tau}_x, \boldsymbol{\Lambda}_x, \boldsymbol{\xi} \sim MVN(\boldsymbol{\tau}_x + \boldsymbol{\Lambda}_x\boldsymbol{\xi}, D_{\sigma^2_\delta}) \tag{1}$$

$$\mathbf{y}|\boldsymbol{\tau}_y, \boldsymbol{\Lambda}_y, \boldsymbol{\eta} \sim MVN(\boldsymbol{\tau}_y + \boldsymbol{\Lambda}_y\boldsymbol{\eta}, D_{\sigma^2_\epsilon}) \tag{2}$$

$$\boldsymbol{\eta}|\boldsymbol{\alpha}, \boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, \boldsymbol{\xi} \sim MVN(\boldsymbol{\alpha} + \boldsymbol{\Gamma}_1\boldsymbol{\xi} + \boldsymbol{\Gamma}_2 vech(\boldsymbol{\xi}\boldsymbol{\xi}'), \Psi) \tag{3}$$

$$\boldsymbol{\xi}|\boldsymbol{\kappa} \sim MVN(\boldsymbol{\kappa}, \Phi) \tag{4}$$

where $MVN(\boldsymbol{\mu}, \Sigma)$ are multivariate normal distributions with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ. $D_{\sigma^2_\delta}$ and $D_{\sigma^2_\epsilon}$ are diagonal matrices with the residual variances in their diagonals. For the construction of Ψ in situations with off-diagonal elements (residual covariances) see Merkle and Rosseel (in press). The prior distributions (and hyperpriors) are given by

$$\tau_x^k \sim N(0, \sigma_{\delta kk}) \qquad \tau_y^j \sim N(0, \sigma_{\epsilon jj}) \tag{5}$$

$$\boldsymbol{\Lambda}_x^k \sim MVN(\mathbf{0}, D^2_{\sigma_{\delta kk}}) \qquad \boldsymbol{\Lambda}_y^j \sim MVN(\mathbf{0}, D_{\sigma^2_{\epsilon jj}}) \tag{6}$$

$$\alpha_v \sim N(0, \sigma_{\zeta vv}) \qquad \kappa_u \sim N(0, 1) \tag{7}$$

$$\sigma_{\delta kk} \sim C^+(\alpha, \beta) \qquad \sigma_{\epsilon jj} \sim C^+(\alpha, \beta) \tag{8}$$

$$\sigma_{\zeta vv} \sim C^+(\alpha, \beta) \qquad \Phi \sim LK_j(\nu) \tag{9}$$

where $k = 1 \ldots p$, $j = 1 \ldots q$, $v = 1 \ldots n$, and $u = 1 \ldots m$ indicates the respective element or row from the parameter matrices (cf. specification of priors in Guo, Zhu, Chow, & Ibrahim, 2012). These priors can be viewed as standard priors (see Bayesian SEM specifications in, e.g., Feng, Wang, Wang, & Song, 2015; Feng, Wu, & Song, 2017; Guo et al., 2012; Song, Li, Cai, & Ip, 2013; Wang, Feng, & Song, 2016). In line with Gelman (2006) and Polson and Scott (2012), we used Half Cauchy priors for the standard deviations instead of inverse Gamma priors for the variances. Further, instead of an inverse Wishart distribution for the covariance matrix of the latent predictor variables, we suggest using an (m dimensional) LK_j prior for the generation of correlation matrices (Gelman, Lee, & Guo, 2015; Lewandowski, Kurowicka, & Joe, 2009); the LK_j distribution is a multivariate generalization of the Beta distribution (with shape parameter $\nu > 0$).

3 Supplement to the Simulation Study

Bayesian model specification for simulation study

For both implementations of the adaptive Bayesian lasso (aB-lasso and aBSS-lasso), we used the following weakly informative priors for the measurement model parameters:

$$\lambda_x^k \sim N(1, \sigma_{\delta kk}) \qquad \sigma_{\delta kk} \sim C^+(0, 2.5) \tag{10}$$

$$\lambda_y^j \sim N(1, \sigma_{\epsilon jj}) \qquad \sigma_{\epsilon jj} \sim C^+(0, 2.5) \tag{11}$$

where the factor loadings λ_x^k, λ_y^j refer to the k-the indicator variable x_k ($k = 1 \ldots 12$) or j-th indicator variable y_j ($j = 1 \ldots 3$). $\sigma_{\delta kk}, \sigma_{\epsilon jj}$ are the respective standard deviations of the residual variables. For the residual standard deviations, Half Cauchy distributions with mean zero and a scaling factor of 2.5 were selected (in line with recommendations by Gelman, 2006; Polson & Scott, 2012). For the latent intercept (α), the standard deviation of the latent residual (ζ) of η, and the correlation matrix of the latent factors (Φ), the following priors were chosen:

$$\alpha \sim N(0, \sigma_\zeta) \qquad \sigma_\zeta \sim C^+(0, 2.5) \qquad \Phi \sim Lk_j(2). \tag{12}$$

Latent factors were generated using a Cholesky decomposition to increase the speed (and stability) of the computation.

Table 4: Parameter bias for different conditions of multicollinearity (ρ), sample size (N) and reliability. Note that the statistics are averaged for the two linear and the two nonlinear effects that are unequal to zero in the population ($\gamma_1 \neq 0, \gamma_2 \neq 0$), and for the effects that are zero in the population ($\gamma_1 = 0, \gamma_2 = 0$).

		Low Reliability				*High Reliability*			
ρ		$\gamma_1 \neq 0$	$\gamma_2 \neq 0$	$\gamma_1 = 0$	$\gamma_2 = 0$	$\gamma_1 \neq 0$	$\gamma_2 \neq 0$	$\gamma_1 = 0$	$\gamma_2 = 0$
					$N = 200$				
.2	aBSS	-0.02	-0.06	0.01	0.01	-0.01	-0.03	0.00	0.00
	aB	0.00	-0.02	0.01	0.01	0.00	-0.00	0.00	0.00
	lms	0.00	-0.00	-0.00	0.00	0.00	-0.00	-0.00	-0.00
	upi	0.05	0.16	-0.01	-0.02	0.01	0.01	-0.00	-0.00
.4	aBSS	-0.02	-0.07	0.01	0.02	-0.01	-0.03	0.01	0.01
	aB	0.01	-0.03	0.01	0.01	0.00	-0.00	0.01	0.00
	lms	0.01	-0.01	-0.01	0.00	0.00	0.00	0.00	-0.00
	upi	0.03	0.04	-0.02	-0.00	0.00	0.02	0.00	-0.00
.6	aBSS	-0.04	-0.09	0.04	0.02	-0.01	-0.06	0.02	0.01
	aB	-0.01	-0.05	0.04	0.02	0.00	-0.02	0.01	0.01
	lms	-0.00	-0.01	0.00	0.00	0.00	-0.00	0.00	0.00
	upi	0.02	-0.18	-0.03	0.06	0.00	0.02	-0.00	-0.01
.8	aBSS	-0.05	-0.13	0.06	0.03	-0.02	-0.10	0.02	0.03
	aB	-0.02	-0.10	0.05	0.03	0.01	-0.05	0.01	0.01
	lms	0.03	-0.03	-0.03	0.01	0.02	0.00	-0.01	-0.00
	upi	-0.09	-0.02	0.10	0.01	0.04	0.13	-0.04	-0.03
					$N = 400$				
.2	aBSS	-0.01	-0.03	0.00	0.01	0.00	-0.00	0.00	0.00
	aB	0.00	-0.01	0.00	0.00	0.01	0.00	0.00	0.00
	lms	0.00	-0.01	-0.00	0.00	0.00	0.00	-0.00	0.00
	upi	0.01	0.04	-0.00	-0.01	0.00	0.01	-0.00	0.00
.4	aBSS	-0.00	-0.04	0.01	0.01	-0.00	-0.02	0.00	0.00
	aB	0.01	-0.01	0.01	0.01	0.00	-0.00	0.00	0.00
	lms	0.01	0.00	-0.00	-0.00	0.00	-0.00	0.00	0.00
	upi	0.04	0.12	-0.03	-0.03	0.00	0.00	-0.00	-0.00
.6	aBSS	-0.00	-0.07	0.01	0.02	-0.01	-0.03	0.01	0.01
	aB	0.01	-0.03	0.00	0.01	-0.00	-0.00	0.01	0.00
	lms	0.01	-0.01	-0.01	0.00	-0.00	0.01	0.00	-0.00
	upi	0.03	-0.06	-0.03	0.02	-0.00	0.02	0.00	-0.00
.8	aBSS	-0.03	-0.11	0.04	0.03	-0.02	-0.09	0.02	0.02
	aB	-0.01	-0.06	0.02	0.02	-0.01	-0.05	0.01	0.01
	lms	0.01	0.02	-0.01	-0.00	-0.00	-0.02	0.00	0.00
	upi	0.02	-0.13	-0.02	0.03	0.01	0.02	-0.01	-0.01
					$N = 800$				
.2	aBSS	0.00	-0.01	0.00	0.00	0.00	-0.00	0.00	0.00
	aB	0.01	0.00	-0.00	0.00	0.00	0.00	0.00	0.00
	lms	0.00	0.01	-0.00	-0.00	0.00	0.00	0.00	0.00
	upi	0.01	0.03	-0.00	-0.00	0.00	0.00	0.00	-0.00
.4	aBSS	0.00	-0.02	0.00	0.01	0.00	-0.01	0.00	0.00

ρ		*Low Reliability* $\gamma_1 \neq 0$	$\gamma_2 \neq 0$	$\gamma_1 = 0$	$\gamma_2 = 0$	*High Reliability* $\gamma_1 \neq 0$	$\gamma_2 \neq 0$	$\gamma_1 = 0$	$\gamma_2 = 0$
	aB	0.01	-0.00	0.00	0.00	0.01	-0.00	0.00	0.00
	lms	0.00	0.00	-0.00	0.00	0.00	-0.00	-0.00	0.00
	upi	0.01	0.03	-0.01	-0.01	0.00	-0.00	-0.00	0.00
.6	aBSS	0.00	-0.05	0.01	0.01	-0.00	-0.02	0.00	0.01
	aB	0.01	-0.02	0.01	0.01	-0.00	-0.01	0.00	0.00
	lms	0.00	-0.00	-0.00	0.00	-0.00	-0.00	0.00	0.00
	upi	0.03	-0.02	-0.03	0.00	-0.00	0.00	-0.00	-0.00
.8	aBSS	-0.02	-0.10	0.03	0.03	-0.00	-0.07	0.01	0.02
	aB	-0.00	-0.06	0.02	0.02	0.00	-0.03	0.00	0.01
	lms	-0.00	-0.02	0.00	0.00	0.00	-0.01	-0.00	0.00
	upi	-0.04	-0.15	0.04	0.04	0.01	0.01	-0.01	-0.00

aBSS – lasso with spike-and-slab priors; aB – standard lasso; lms – LMS approach; upi – unconstrained product indicator approach.

Table 5: 95% Coverage rates for different conditions of multicollinearity (ρ), sample size (N) and reliability. Note that the statistics are averaged for the two linear and the two nonlinear effects that are unequal to zero in the population ($\gamma_1 \neq 0, \gamma_2 \neq 0$), and for the effects that are zero in the population ($\gamma_1 = 0, \gamma_2 = 0$).

		Low Reliability				*High Reliability*			
ρ		$\gamma_1 \neq 0$	$\gamma_2 \neq 0$	$\gamma_1 = 0$	$\gamma_2 = 0$	$\gamma_1 \neq 0$	$\gamma_2 \neq 0$	$\gamma_1 = 0$	$\gamma_2 = 0$
					$N = 200$				
.2	aBSS	0.95	0.93	0.99	1.00	0.94	0.93	0.98	0.99
	aB	0.98	0.98	0.97	0.99	0.95	0.96	0.95	0.98
	lms	0.96	0.93	0.93	0.94	0.93	0.92	0.93	0.93
	upi	0.99	0.99	0.99	1.00	0.93	0.93	0.90	0.94
.4	aBSS	0.94	0.95	1.00	1.00	0.94	0.96	0.99	0.99
	aB	0.98	0.98	0.98	0.99	0.96	0.97	0.97	0.98
	lms	0.94	0.95	0.95	0.94	0.95	0.93	0.93	0.94
	upi	1.00	1.00	1.00	1.00	0.95	0.95	0.93	0.93
.6	aBSS	0.94	0.96	0.99	1.00	0.95	0.95	1.00	1.00
	aB	0.97	0.98	0.99	0.99	0.96	0.98	0.97	0.98
	lms	0.94	0.93	0.98	0.95	0.95	0.94	0.95	0.92
	upi	1.00	1.00	1.00	1.00	0.94	0.94	0.95	0.94
.8	aBSS	0.95	0.99	0.99	1.00	0.93	0.97	0.99	1.00
	aB	0.98	1.00	0.99	1.00	0.96	0.98	0.97	0.99
	lms	0.97	0.92	0.98	0.96	0.94	0.95	0.95	0.93
	upi	1.00	1.00	1.00	1.00	0.99	1.00	0.98	0.99
					$N = 400$				
.2	aBSS	0.95	0.92	1.00	0.99	0.92	0.93	0.99	0.99
	aB	0.96	0.95	0.98	0.97	0.92	0.95	0.97	0.97
	lms	0.94	0.92	0.96	0.94	0.92	0.94	0.96	0.94
	upi	0.96	0.97	0.97	0.99	0.92	0.94	0.94	0.93
.4	aBSS	0.96	0.94	0.99	0.99	0.96	0.95	0.99	1.00
	aB	0.96	0.96	0.98	0.98	0.97	0.96	0.96	0.97
	lms	0.96	0.93	0.97	0.94	0.96	0.93	0.94	0.94
	upi	0.99	0.99	1.00	1.00	0.95	0.94	0.93	0.94
.6	aBSS	0.94	0.94	1.00	1.00	0.94	0.95	0.98	0.99
	aB	0.97	0.97	0.97	0.99	0.95	0.94	0.96	0.97
	lms	0.96	0.94	0.94	0.94	0.94	0.92	0.95	0.93
	upi	1.00	1.00	1.00	1.00	0.95	0.93	0.96	0.93
.8	aBSS	0.96	0.98	0.98	1.00	0.93	0.95	0.99	1.00
	aB	0.98	0.99	0.97	1.00	0.96	0.98	0.97	0.99
	lms	0.97	0.94	0.96	0.94	0.95	0.94	0.96	0.94
	upi	1.00	1.00	1.00	1.00	0.96	0.98	0.98	0.97
					$N = 800$				
.2	aBSS	0.95	0.95	1.00	0.99	0.96	0.95	0.99	0.99
	aB	0.95	0.96	0.96	0.97	0.95	0.95	0.97	0.96
	lms	0.95	0.94	0.95	0.93	0.94	0.94	0.96	0.94
	upi	0.96	0.96	0.96	0.97	0.94	0.93	0.95	0.94
.4	aBSS	0.95	0.97	0.99	1.00	0.95	0.96	0.98	0.99

ρ		*Low Reliability* $\gamma_1 \neq 0$	$\gamma_2 \neq 0$	$\gamma_1 = 0$	$\gamma_2 = 0$	*High Reliability* $\gamma_1 \neq 0$	$\gamma_2 \neq 0$	$\gamma_1 = 0$	$\gamma_2 = 0$
	aB	0.96	0.96	0.96	0.98	0.94	0.96	0.95	0.95
	lms	0.96	0.95	0.96	0.94	0.96	0.96	0.94	0.93
	upi	0.97	0.96	0.97	0.99	0.96	0.95	0.94	0.93
.6	aBSS	0.96	0.94	0.99	1.00	0.96	0.95	1.00	0.99
	aB	0.95	0.95	0.98	0.98	0.96	0.96	0.98	0.97
	lms	0.95	0.93	0.97	0.96	0.95	0.96	0.96	0.95
	upi	0.99	0.99	0.99	1.00	0.95	0.96	0.96	0.95
.8	aBSS	0.95	0.97	0.99	1.00	0.94	0.94	0.99	0.99
	aB	0.97	0.99	0.97	0.99	0.94	0.96	0.97	0.97
	lms	0.97	0.95	0.96	0.94	0.94	0.94	0.96	0.93
	upi	1.00	1.00	1.00	1.00	0.94	0.95	0.95	0.95

aBSS – lasso with uninformative spike-and-slab priors; aB – standard lasso; lms – LMS approach; upi – unconstrained product indicator approach.

Table 6: Root mean square error (rmse) for different conditions of multicollinearity (ρ), sample size (N) and reliability. Note that the statistics are averaged for the two linear and the two nonlinear effects that are unequal to zero in the population ($\gamma_1 \neq 0, \gamma_2 \neq 0$), and for the effects that are zero in the population ($\gamma_1 = 0, \gamma_2 = 0$).

		Low Reliability				*High Reliability*			
ρ		$\gamma_1 \neq 0$	$\gamma_2 \neq 0$	$\gamma_1 = 0$	$\gamma_2 = 0$	$\gamma_1 \neq 0$	$\gamma_2 \neq 0$	$\gamma_1 = 0$	$\gamma_2 = 0$
					$N = 200$				
.2	aBSS	0.11	0.11	0.07	0.06	0.08	0.08	0.05	0.05
	aB	0.10	0.10	0.09	0.10	0.08	0.08	0.07	0.07
	lms	0.11	0.13	0.11	0.13	0.08	0.07	0.08	0.07
	upi	0.29	0.82	0.28	0.65	0.08	0.08	0.09	0.09
.4	aBSS	0.13	0.12	0.07	0.07	0.08	0.08	0.05	0.05
	aB	0.12	0.12	0.10	0.12	0.08	0.08	0.07	0.08
	lms	0.12	0.15	0.12	0.16	0.08	0.09	0.08	0.09
	upi	0.39	0.90	0.44	0.96	0.08	0.11	0.09	0.11
.6	aBSS	0.15	0.13	0.09	0.09	0.10	0.11	0.06	0.07
	aB	0.13	0.15	0.12	0.14	0.10	0.11	0.08	0.11
	lms	0.16	0.26	0.15	0.27	0.10	0.12	0.09	0.13
	upi	0.49	1.02	0.56	1.07	0.11	0.19	0.11	0.20
.8	aBSS	0.19	0.16	0.13	0.10	0.13	0.14	0.08	0.09
	aB	0.19	0.21	0.17	0.18	0.13	0.17	0.11	0.17
	lms	0.33	0.69	0.29	0.83	0.13	0.27	0.13	0.29
	upi	0.85	1.02	0.79	1.18	0.27	1.17	0.29	1.07
					$N = 400$				
.2	aBSS	0.07	0.08	0.04	0.05	0.06	0.05	0.03	0.03
	aB	0.07	0.08	0.06	0.07	0.06	0.05	0.05	0.05
	lms	0.08	0.08	0.07	0.08	0.05	0.05	0.05	0.05
	upi	0.10	0.17	0.09	0.18	0.06	0.06	0.05	0.05
.4	aBSS	0.08	0.09	0.05	0.06	0.05	0.06	0.04	0.04
	aB	0.08	0.10	0.07	0.09	0.06	0.06	0.05	0.06
	lms	0.08	0.11	0.08	0.11	0.06	0.06	0.06	0.06
	upi	0.22	0.61	0.25	0.63	0.06	0.07	0.06	0.07
.6	aBSS	0.10	0.12	0.06	0.07	0.07	0.08	0.04	0.05
	aB	0.10	0.13	0.09	0.12	0.07	0.09	0.06	0.08
	lms	0.10	0.16	0.10	0.17	0.07	0.10	0.07	0.10
	upi	0.34	0.78	0.37	0.89	0.07	0.12	0.07	0.11
.8	aBSS	0.14	0.15	0.11	0.09	0.11	0.13	0.06	0.08
	aB	0.14	0.20	0.14	0.18	0.10	0.15	0.08	0.15
	lms	0.18	0.47	0.18	0.47	0.10	0.20	0.09	0.21
	upi	0.62	1.10	0.69	1.13	0.12	0.37	0.12	0.40
					$N = 800$				
.2	aBSS	0.05	0.05	0.03	0.03	0.04	0.03	0.02	0.02
	aB	0.05	0.05	0.05	0.05	0.04	0.04	0.03	0.03
	lms	0.05	0.05	0.05	0.05	0.04	0.04	0.03	0.04
	upi	0.06	0.09	0.06	0.08	0.04	0.04	0.03	0.04
.4	aBSS	0.05	0.07	0.04	0.04	0.04	0.04	0.03	0.03

ρ		Low Reliability $\gamma_1 \neq 0$	$\gamma_2 \neq 0$	$\gamma_1 = 0$	$\gamma_2 = 0$	High Reliability $\gamma_1 \neq 0$	$\gamma_2 \neq 0$	$\gamma_1 = 0$	$\gamma_2 = 0$
	aB	0.06	0.07	0.05	0.06	0.04	0.04	0.04	0.04
	lms	0.06	0.07	0.06	0.07	0.04	0.04	0.04	0.04
	upi	0.07	0.14	0.07	0.14	0.04	0.04	0.04	0.05
.6	aBSS	0.07	0.10	0.04	0.06	0.04	0.06	0.03	0.04
	aB	0.07	0.10	0.06	0.09	0.05	0.06	0.04	0.06
	lms	0.07	0.12	0.07	0.11	0.05	0.06	0.05	0.06
	upi	0.20	0.54	0.20	0.69	0.05	0.07	0.05	0.07
.8	aBSS	0.12	0.13	0.08	0.09	0.07	0.11	0.04	0.07
	aB	0.11	0.17	0.10	0.17	0.07	0.12	0.06	0.11
	lms	0.11	0.29	0.12	0.31	0.07	0.15	0.06	0.15
	upi	0.35	1.01	0.37	0.94	0.07	0.19	0.07	0.19

aBSSu – lasso with uninformative spike-and-slab priors; aB – standard lasso; lms – LMS approach; upi – unconstrained product indicator approach.

Table 7: Relative accuracy (rmse(·)/rmse(aBSS)) for different conditions of multicollinearity (ρ), sample size (N) and reliability. Note that the statistics are averaged for the two linear and the two nonlinear effects that are unequal to zero in the population ($\gamma_1 \neq 0, \gamma_2 \neq 0$), and for the effects that are zero in the population ($\gamma_1 = 0, \gamma_2 = 0$).

		Low Reliability				*High Reliability*			
ρ		$\gamma_1 \neq 0$	$\gamma_2 \neq 0$	$\gamma_1 = 0$	$\gamma_2 = 0$	$\gamma_1 \neq 0$	$\gamma_2 \neq 0$	$\gamma_1 = 0$	$\gamma_2 = 0$
					$N = 200$				
.2	aB	0.91	0.93	1.39	1.48	0.97	0.93	1.39	1.49
	lms	0.97	1.15	1.63	1.93	0.94	0.92	1.55	1.66
	upi	2.61	7.52	4.17	10.05	0.98	1.04	1.64	1.92
.4	aB	0.94	1.02	1.39	1.54	0.95	0.98	1.40	1.57
	lms	0.99	1.31	1.66	2.18	0.93	1.04	1.55	1.80
	upi	3.11	7.63	6.15	12.82	0.98	1.28	1.64	2.18
.6	aB	0.90	1.12	1.34	1.63	0.97	1.05	1.39	1.64
	lms	1.06	1.90	1.77	3.17	0.96	1.17	1.59	2.07
	upi	3.36	7.53	6.42	12.62	1.06	1.77	1.82	3.05
.8	aB	1.02	1.32	1.26	1.83	0.95	1.25	1.34	1.95
	lms	1.77	4.27	2.17	8.28	1.01	1.99	1.62	3.30
	upi	4.58	6.36	5.94	11.69	2.03	8.60	3.48	11.97
					$N = 400$				
.2	aB	0.99	0.93	1.45	1.47	1.01	0.99	1.42	1.50
	lms	1.05	0.97	1.57	1.65	0.96	0.97	1.53	1.61
	upi	1.37	2.01	2.03	3.71	0.98	1.09	1.55	1.71
.4	aB	1.00	1.01	1.46	1.55	1.02	1.00	1.40	1.56
	lms	1.03	1.13	1.66	1.92	1.02	0.99	1.51	1.73
	upi	2.63	6.44	5.36	11.40	1.04	1.12	1.57	1.88
.6	aB	0.95	1.10	1.42	1.65	0.98	1.08	1.41	1.56
	lms	0.99	1.39	1.67	2.33	0.96	1.16	1.53	1.82
	upi	3.28	6.76	6.03	12.43	1.00	1.40	1.58	2.14
.8	aB	0.96	1.36	1.25	2.00	0.92	1.19	1.36	1.81
	lms	1.25	3.22	1.61	5.20	0.93	1.54	1.56	2.62
	upi	4.25	7.51	6.21	12.62	1.11	2.83	1.90	5.02
					$N = 800$				
.2	aB	1.01	0.99	1.44	1.50	1.02	1.03	1.45	1.48
	lms	1.05	1.02	1.50	1.64	1.01	1.03	1.54	1.57
	upi	1.11	1.70	1.80	2.36	1.03	1.11	1.55	1.64
.4	aB	1.05	1.00	1.44	1.56	1.02	1.07	1.40	1.53
	lms	1.05	1.03	1.55	1.79	0.99	1.06	1.48	1.65
	upi	1.28	2.11	1.85	3.38	1.00	1.14	1.49	1.76
.6	aB	1.04	1.05	1.45	1.58	1.04	1.04	1.48	1.61
	lms	1.06	1.21	1.58	2.00	1.03	1.04	1.59	1.82
	upi	2.99	5.52	4.78	12.37	1.04	1.16	1.64	1.99
.8	aB	0.93	1.27	1.30	1.97	1.01	1.14	1.44	1.63
	lms	0.99	2.18	1.51	3.53	1.02	1.35	1.60	2.11
	upi	3.06	7.52	4.77	10.69	1.10	1.75	1.75	2.70

aB – standard lasso; lms – LMS approach; upi – unconstrained product indicator approach.

References

Brandt, H., Kelava, A., & Klein, A. G. (2014). A simulation study comparing recent approaches for the estimation of nonlinear effects in SEM under the condition of non-normality. *Structural Equation Modeling*, *21*, 181–195. doi: 10.1080/10705511.2014.882660

Cham, H., Reshetnyak, E., Rosenfeld, B., & Breitbart, W. (2017). Full information maximum likelihood estimation for latent variable interactions with incomplete indicators. *Multivariate Behavioral Research*, *52*, 12–30. doi: 10.1080/00273171.2016.1245600

Cham, H., West, S. G., Ma, Y., & Aiken, L. S. (2012). Estimating latent variable interactions with nonnormal observed data: A comparison of four approaches. *Multivariate Behavioral Research*, *47*, 840–876. doi: 10.1080/00273171.2012.732901

Chen, S.-P., & Cheng, C.-P. (2014). Model specification for latent interactive and quadratic effects in matrix form. *Structural Equation Modeling*, *21*, 94–101. doi: 10.1080/10705511.2014.859509

Feng, X.-N., Wang, G.-C., Wang, Y.-F., & Song, X.-Y. (2015). Structure detection of semiparametric structural equation models with Bayesian adaptive group lasso. *Statistics in Medicine*, *34*, 1527–1547. doi: 10.1002/sim.6410

Feng, X.-N., Wu, H.-T., & Song, X.-Y. (2017). Bayesian regularized multivariate generalized latent variable models. *Structural Equation Modeling*. doi: 10.1080/10705511.2016.1257353

Foldnes, N., & Hagtvet, K. A. (2014). The choice of product indicators in latent variable interaction models: Post hoc analyses. *Psychological Methods*, *19*, 444–457. doi: 10.1037/a0035728

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*, *1*, 515–534. doi: 10.1214/06-BA117A

Gelman, A., Lee, D., & Guo, J. (2015). Stan: A probabilistic programming language for Bayesian inference and optimization. *Journal of Educational and Behavioral Statistics*, *40*(5), 540–543. doi: 10.3102/1076998615606113

Gerhard, C., Büchner, R. D., Klein, A. G., & Schermelleh-Engel, K. (2014). A fit index to assess model fit and detect omitted terms in nonlinear sem. *Structural Equation Modeling*, *24*, 414–27. doi: 10.1080/10705511.2016.12689239

Gerhard, C., Klein, A. G., Schermelleh-Engel, K., Moosbrugger, H., Gäde, J., & Brandt, H. (2014). On the performance of likelihood-based difference tests in nonlinear structural equation models. *Structural Equation Modeling*, *22*, 276–287. doi: 10.1080/10705511.2014.935752

Guo, R., Zhu, H., Chow, S.-M., & Ibrahim, J. G. (2012). Bayesian lasso for semiparametric structural equation models. *Biometrics*, *68*, 567–577. doi: 10.1111/j.1541-0420.2012.01751.x

Harring, J. R., Weiss, B. A., & Hsu, J.-C. (2012). A comparison of methods for estimating quadratic effects in nonlinear structural equation models. *Psychological Methods*, *17*, 193–214. doi: 10.1037/a0027539

Jackman, M. G. A., Leite, W. L., & Cochrane, D. J. (2011). Estimating latent variable interactions with the unconstrained approach: A comparison of methods to form product indicators for large, unequal numbers of items. *Structural Equation Modeling*, *18*, 274–288. doi: 10.1080/10705511.2011.557342

Kelava, A., Moosbrugger, H., Dimitruk, P., & Schermelleh-Engel, K. (2008). Multicollinearity and missing constraints: A comparison of three approaches for the analysis of latent nonlinear effects. *Methodology*, *4*, 51–66. doi: 10.1027/1614-2241.4.2.51

Kelava, A., & Nagengast, B. (2012). A Bayesian model for the estimation of latent interaction and quadratic effects when latent variables are non-normally distributed. *Multivariate Behavioral Research*, *47*, 717–742. doi: 10.1080/00273171.2012.715560

Kelava, A., Nagengast, B., & Brandt, H. (2014). A nonlinear structural equation mixture modeling approach for non-normally distributed latent predictor variables. *Structural Equation*

Modeling, *21*, 468–481. doi: 10.1080/10705511.2014.915379

Kenny, D., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, *96*, 201–210. doi: 10.1037/0033-2909.96.1.201

Klein, A. G., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, *65*, 457–474. doi: 10.1007/BF02296338

Klein, A. G., & Muthén, B. O. (2007). Quasi maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects. *Multivariate Behavioral Research*, *42*, 647–674. doi: 10.1080/00273170701710205

Klein, A. G., Schermelleh-Engel, K., Moosbrugger, H., & Kelava, A. (2009). Assessing spurious interaction effects. In T. Teo & M. Khine (Eds.), *Structural equation modeling in educational research: Concepts and applications* (pp. 13–28). Rotterdam, NL: Sense Publishers.

Lee, S.-Y. (2007). *Structural equation modeling: A Bayesian approach.* Chichester, England: Wiley.

Leite, W., & Zuo, Y. (2011). Modeling latent interactions at level 2 in multilevel structural equation models: An evaluation of mean-centered and residual-centered approaches. *Structural Equation Modeling*, *18*, 449–464. doi: 10.1080/10705511.2011.582400

Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, *100*, 1989–2001. doi: 10.1016/j.jmva.2009.04.008

Li, F., Duncan, T. E., & Acock, A. (2000). Modeling interaction effects in latent growth curve models. *Structural Equation Modeling*, *7*, 497–533. doi: 10.1207/S15328007SEM0704_1

Little, T. D., Bovaird, J. A., & Widaman, K. F. (2006). On the merits of orthogonalizing powered and interaction terms: Implications for modeling interactions among latent variables. *Structural Equation Modeling*, *13*, 497–519. doi: 10.1207/s15328007sem1304_1

Lyhagen, J. (2007). Estimating nonlinear structural models: EMM and the Kenny-Judd model. *Structural Equation Modeling*, *14*, 391–403. doi: 10.1080/10705510701301487

Marsh, H. W., Wen, Z., & Hau, K.-T. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods*, *9*, 275–300. doi: 10.1037/1082-989X.9.3.275

Marsh, H. W., Wen, Z., Hau, K.-T., Little, T. D., Bovaird, J. A., & Widaman, K. F. (2007). Unconstrained structural equation models of latent interactions: Contrasting residual- and mean-centered approaches. *Structural Equation Modeling*, *14*, 570–580. doi: 10.1080/10705510701303921

Maslowsky, J., Jager, J., & Hemken, D. (2015). Estimating and interpreting latent variable interactions: A tutorial for applying the latent moderated structural equations method. *International Journal of Behavioral Development*, *39*, 87–96. doi: 10.1177/0165025414552301

Merkle, E. C., & Rosseel, Y. (in press). blavaan: Bayesian structural equation models via parameter expansion. *Journal of Statistical Software*. Retrieved from `https://arxiv.org/abs/1511.05604`

Mooijaart, A., & Bentler, P. M. (2010). An alternative approach for nonlinear latent variable models. *Structural Equation Modeling*, *17*, 357–573. doi: 10.1080/10705511.2010.488997

Moulder, B. C., & Algina, J. (2002). Comparison of methods for estimating and testing latent variable interactions. *Structural Equation Modeling*, *9*, 1–19. doi: 10.1207/S15328007SEM0901_1

Mullis, I., & Martin, M. E. (2013). *TIMSS 2015 Assessment Frameworks.* TIMSS & PIRLS International Study Center. Retrieved from `http://timssandpirls.bc.edu/timss2015/frameworks.html`

Ping, R. A. (1998). EQS and LISREL Examples: Using Survey Data. In R. E. Schumacker & G. A. Marcoulides (Eds.), *Interaction and nonlinear effects in structural equation modeling* (pp. 63–100). Mahwah, NJ: Lawrence Erlbaum Associates.

Polson, N. G., & Scott, J. G. (2012). On the half-Cauchy prior for a global scale parameter. *Bayesian Analysis*, *7*(4), 887–902. doi: 10.1214/12-BA730

Song, X.-Y., Li, Z.-H., Cai, J.-H., & Ip, E. H.-S. (2013). A Bayesian approach for generalized semi-parametric structural equation models. *Psychometrika*, *78*, 624–647. doi: 10.1007/s11336-013-9323-7

Sun, R., & Willson, V. L. (2009). Evaluating intercept-slope interactions in latent growth modeling. *Structural Equation Modeling*, *16*, 226–244. doi: 10.1080/10705510902751051

Wall, M. M., & Amemiya, Y. (2000). Estimation for polynomial structural equation models. *Journal of the Statistical American Association*, *95*, 929–940. doi: 10.1080/01621459.2000.10474283

Wall, M. M., & Amemiya, Y. (2001). Generalized appended product indicator procedure for nonlinear structural equation analysis. *Journal of Educational and Behavioral Statistics*, *26*, 1–29. doi: 10.3102/10769986026001001

Wall, M. M., & Amemiya, Y. (2003). A method of moments technique for fitting interaction effects in structural equation models. *British Journal of Mathematical and Statistical Psychology*, *56*, 47–63. doi: 10.1348/000711003321645331

Wang, Y., Feng, X.-N., & Song, X.-Y. (2016). Bayesian quantile structural equation models. *Structural Equation Modeling*, *23*(2), 246–258. doi: 10.1080/10705511.2015.1033057