Bayesian regularized SEM

Current capabilities and constraints

Sara van Erp
Utrecht University
s.j.vanerp@uu.nl

Preprint: *https://psyarxiv.com/92vh8*

What do we want to do?

Finding an accurate description of the phenomenon under investigation that is as parsimonious as possible.

Classical regularized SEM

Idea: add a penalty to the parameters to regularize

$$F_{regsem}\big(S, \Sigma(\theta)\big) = F\big(S, \Sigma(\theta)\big) + \boxed{\lambda P\big(\theta_{reg}\big)}$$

e.g. lasso, ridge, or elastic net.

The penalty shrinks small coefficients to zero, thereby identifying the model.

Regularized SEM: possible applications

- Cross-loadings in CFA (Jacobucci et al., 2016)
- Loadings in EFA (Jin et al., 2018; Trendafilov et al., 2017; Hirose and Yamamoto, 2014)
- Selection of covariates in MIMIC models (Jacobucci et al., 2016, 2019)
- Selection of mediators in mediation analysis (Serang et al., 2017; Zhao and Luo, 2016; van Kesteren and Oberski, 2019)
- More stable estimation of latent class models in small samples (Chen et al., 2017; Robitzsch, 2020)
- Detection of violations of MI (Magis et al., 2015; Tutz and Schauberger, 2013)
- ...

Any model with (too) many parameters where some are assumed to be zero.

Bayesian regularized SEM

Instead of a penalty function, we can use a Bayesian *shrinkage prior:*

$$p(\theta|y) \propto p(y|\theta)\boldsymbol{p}(\boldsymbol{\theta})$$

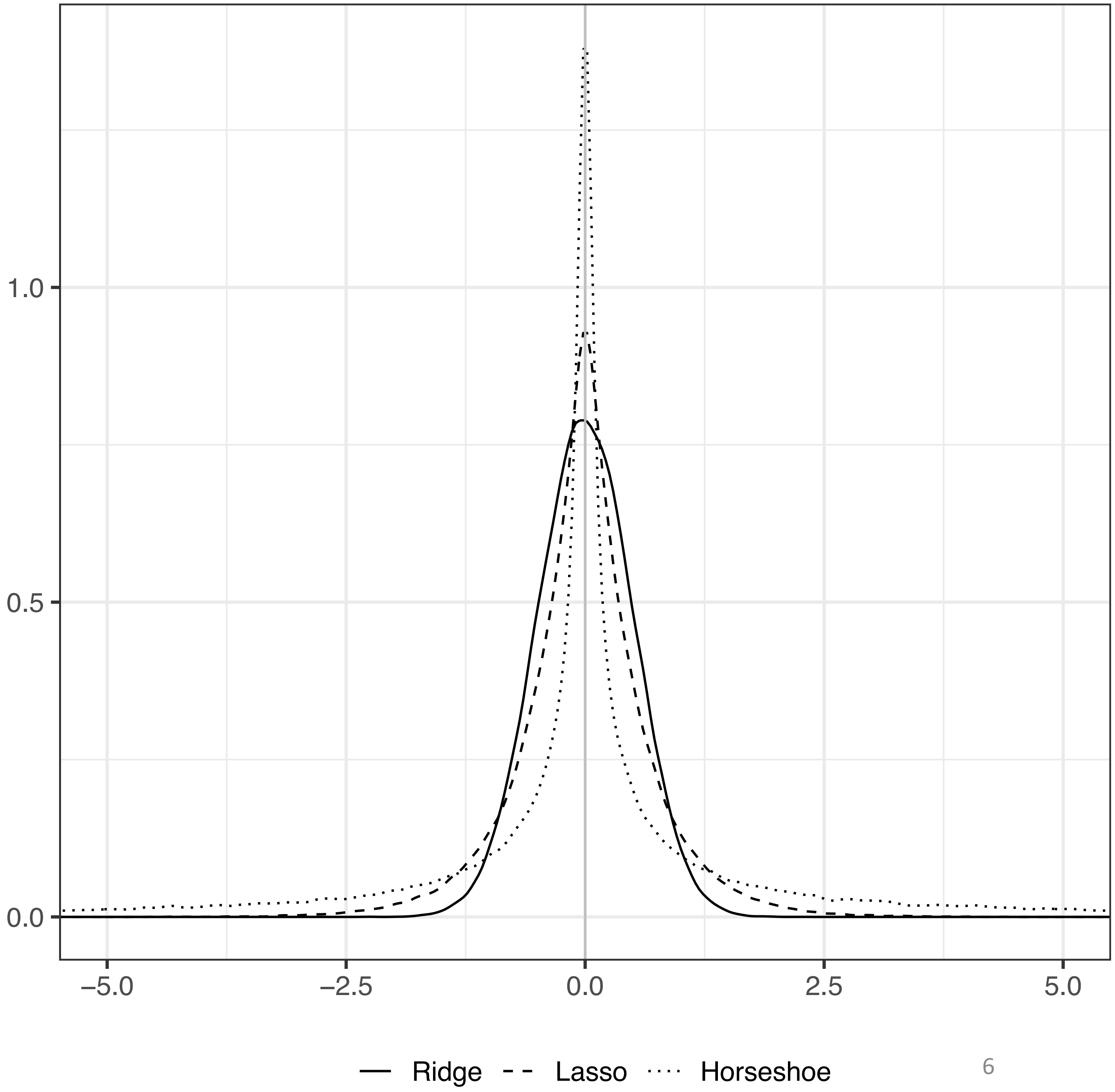
Shrinkage priors pull small coefficients towards zero, some priors in the same way as classical penalties.

See also van Erp et al. (2019)

Shrinkage priors

Properties:

1. Very peaked around zero
2. Heavy tails



Advantages Bayesian regularized SEM

- Intuitive interpretation
- Automatic uncertainty estimates
- Incorporation of prior information
- Flexibility in terms of priors considered
- Automatic estimation penalty parameter

	Ridge	*Lasso*	*Lasso extensions*	*Spike-and-slab*	*Other*
EFA				**Carvalho et al. (2008); Chen (2021);** West (2003)	Bhattacharya & Dunson (2011); Conti et al. (2014); Legramanti et al. (2020)
CFA	**Liang (2020)**; Lu et al, (2016); **Muthén & Asparouhov (2012)**; Vamvourellis et al. (2021)	**Chen, Guo et al. (2021)**; Pan et al. (2017)	**Chen, Guo et al. (2021)**	Lu et al, (2016)	
Neural drift diffusion		Kang et al. (2022)			
Item response models	Vamvourellis et al. (2021)	**Chen (2020)**			
Multiple group factor models	**Shi et al. (2017)**		Chen, Bauer et al. (2021)	Chen, Bauer et al. (2021)	
Non-linear SEM		Guo et al. (2012)	Brandt et al. (2018); Feng, Wang et al. (2015)	Brandt et al. (2018)	
General SEM		Feng, Wu et al. (2015); Feng, Wu et al. (2017)	Feng, Wu et al. (2015); Feng, Wu et al. (2017)		
Quantile SEM		Feng, Wang et al. (2017)	Feng, Wang et al. (2017)		
Latent growth curve models			Jacobucci & Grimm (2018)		

Software for Bayesian regularized SEM

User-friendly

- `Mplus`: only ridge available
- R-package `blcfa`: covariance lasso for CFA (requires `Mplus`)
- R-package `blavaan`: only ridge available
- R-package `LAWLB`: adaptive and covariance lasso for CFA
- R-package `infinitefactor`: aimed at high-dimensional EFA

Not so user-friendly:

- Description MCMC algorithm or (R) implementation online
- Code it yourself (Stan or JAGS)

Added difficulty: prior should be informative enough to identify the model

An illustration

Aim: Comparing various methods on *realistic* data.

Data (N = 1495) was generated based on an observed covariance matrix of 65 items from the ADaptive Ability Performance Test (ADAPT; Jonker et al., 2022)

EFA resulted in an eight factor solution, based on which the first three factors and 36 items were used for further analysis.

Illustration: Methods

I compared the following methods:

-`lslx`

mcp, lasso, and elastic net penalties

-`blavaan`

ridge prior with fixed sd

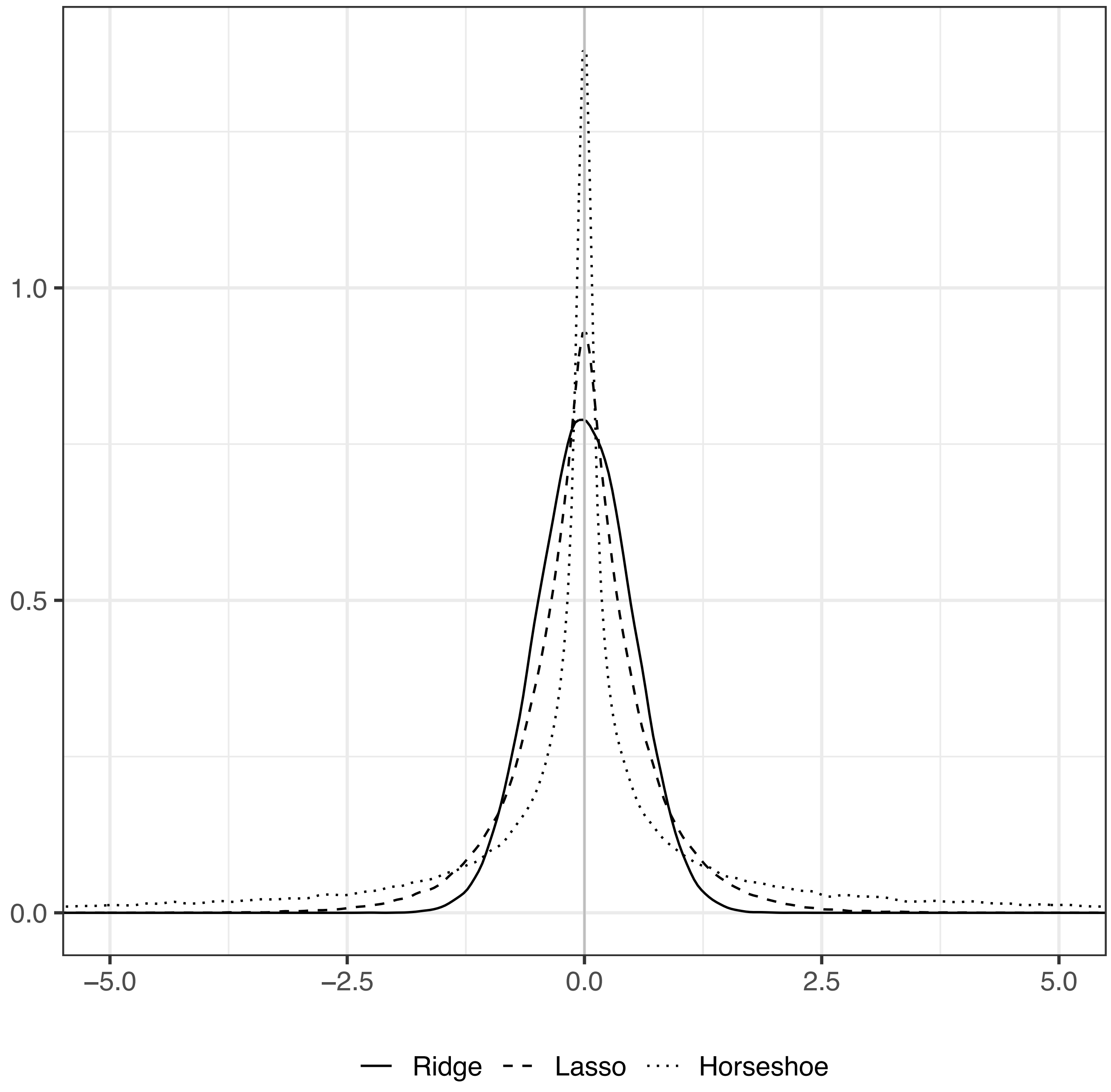
-`lawlb`

lasso and adaptive lasso

-`Stan`

ridge prior with estimated sd and regularized horseshoe prior

Shrinkage priors compared (1)



Shrinkage priors compared (2)

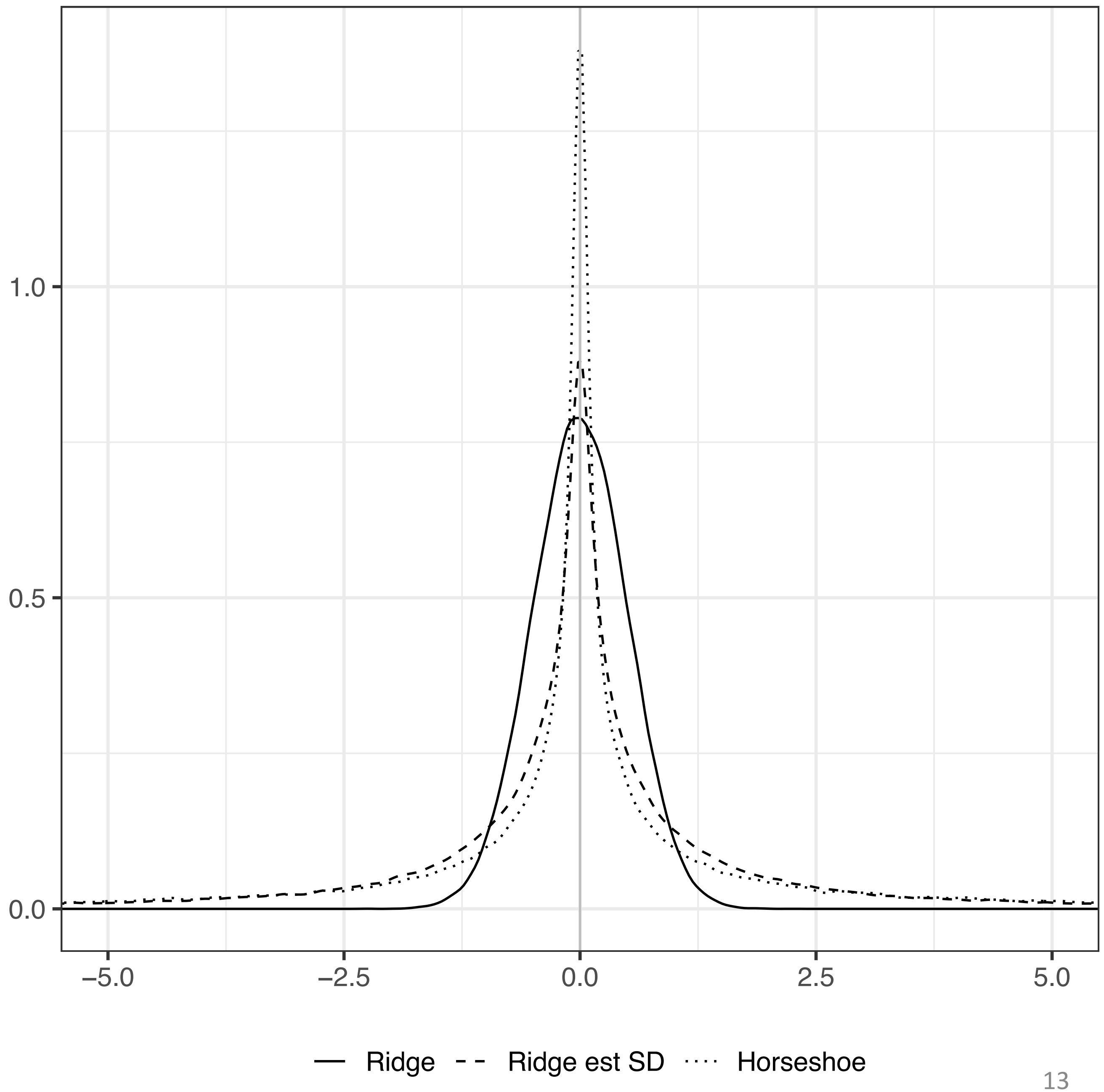


Illustration: General observations

- Model specification can be a hassle (especially for large models)
- Some packages allow simultaneous regularization of cross-loadings and residual covariances. This was not tested here.
- Some packages provide information on fit indices but these can be impacted by prior choice (see Edwards and Konold, 2023)

- Variational Bayes seems way off, especially for the ridge
- Practically no differences between the lasso and adaptive lasso
- Providing a prior guess for the number of substantial cross-loadings in the regularized horseshoe does not matter for the results

Illustration: Estimated cross-loadings

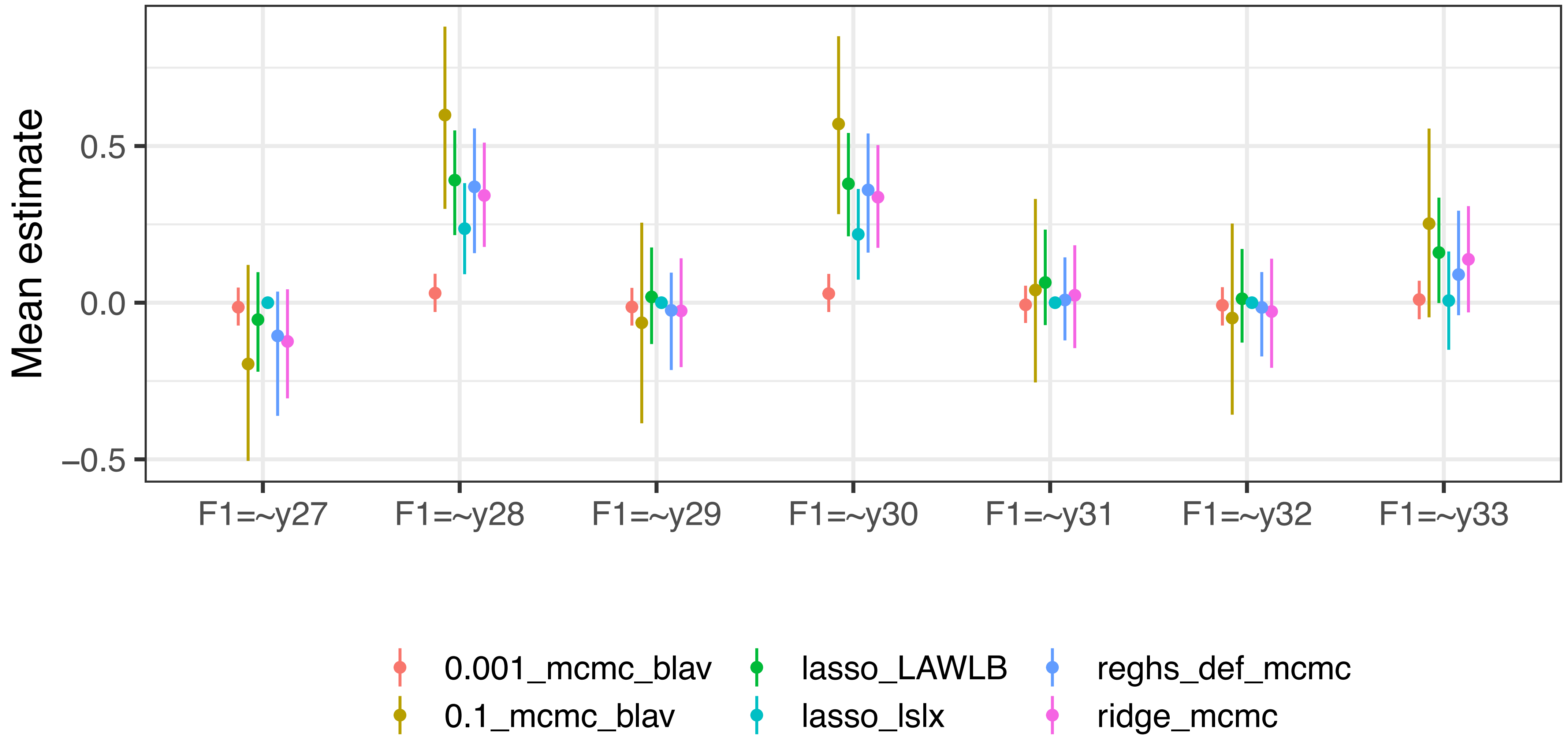
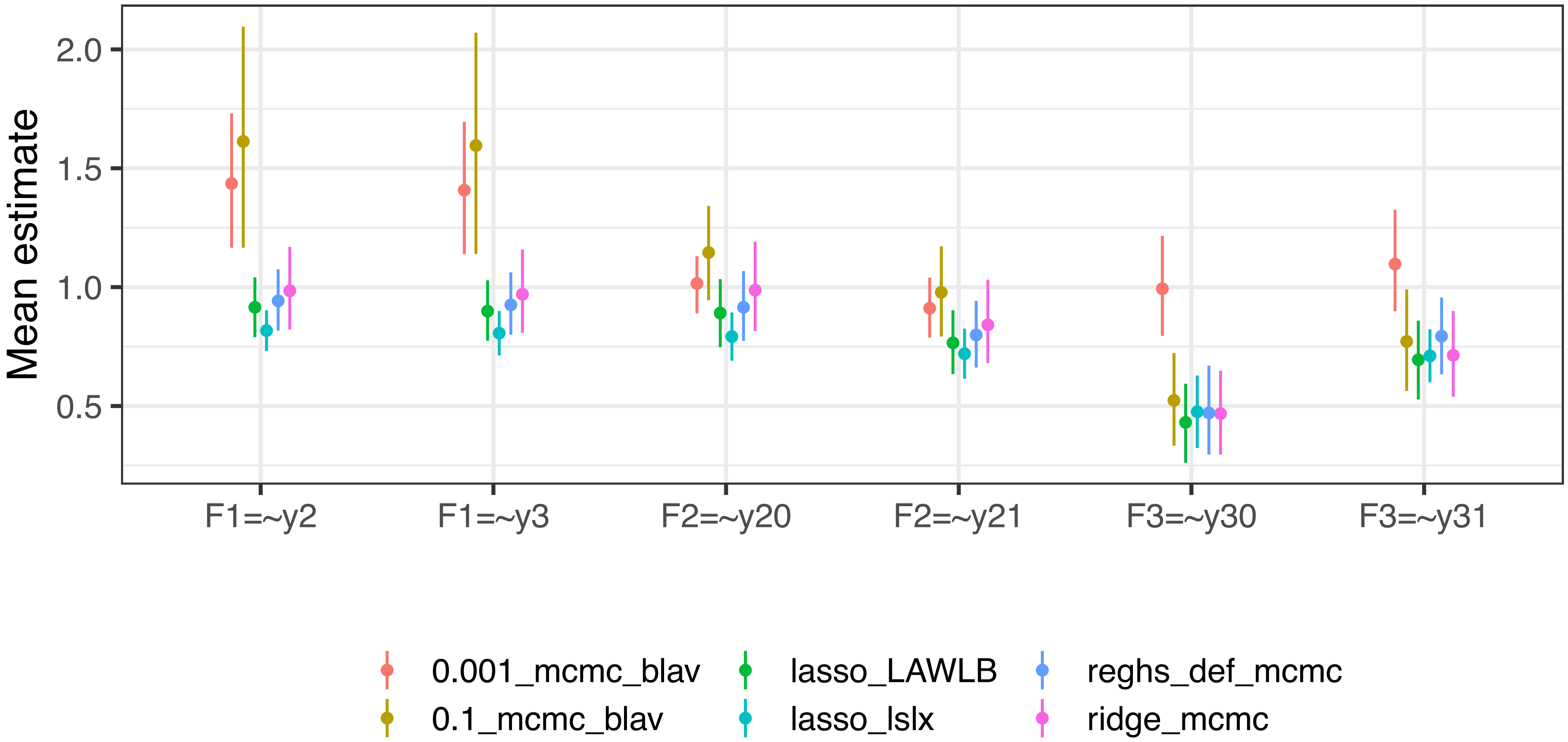


Illustration: Estimated main loadings



Current capabilities and constraints

- There are quite some possibilities for simple Bayesian regularized SEM (e.g., ridge or lasso).
- More advanced priors are possible, but not yet implemented in user-friendly software.
- Different priors can lead to different estimates, which in turn influence other parameter estimates in the model.
- Existing methods for parameter selection are arbitrary (See also Zhang, Pan & Ip, 2021)
- In addition, general guidelines based on simulation studies are lacking. Long run times and convergence issues can make simulation studies difficult.

Thank you!

Check out the preprint: *https://psyarxiv.com/92vh8*

References

Biswas, N., Mackey, L., & Meng, X. L. (2022). Scalable Spike-and-Slab. *arXiv preprint arXiv:2204.01668*.

Chen, J. (2021). A Generalized Partially Confirmatory Factor Analysis Framework with Mixed Bayesian Lasso Methods. Multivariate Behavioral Research, 1-16.

Chen, Y., Li, X., Liu, J., and Ying, Z. (2017). Regularized latent class analysis with application in cognitive diagnosis. Psychometrika, 82(3):660–692.

Edwards, K. D. & Konold, T. R. (2023). Impact of Informative Priors on Model Fit Indices in Bayesian Confirmatory Factor Analysis, Structural Equation Modeling: A Multidisciplinary Journal, 30:2, 272-283

Geminiani, E., Marra, G., & Moustaki, I. (2021). Single- and Multiple-Group Penalized Factor Analysis: A Trust-Region Algorithm Approach with Integrated Automatic Multiple Tuning Parameter Selection. Psychometrika, 86(1), 65-95.

Hirose, K. and Yamamoto, M. (2014). Sparse estimation via nonconcave penalized likelihood in factor analysis model. Statistics and Computing, 25(5):863–875.

Huang, P. (2020). lslx: Semi-Confirmatory Structural Equation Modeling via Penalized Likelihood. Journal of Statistical Software_ 93(7), 1-37. doi:10.18637/jss.v093.i07

Jacobucci, R., Brandmaier, A. M., and Kievit, R. A. (2019). A practical guide to variable selection in structural equation modeling by using regularized multiple-indicators, multiple-causes models. Advances in Methods and Practices in Psychological Science, 2(1):55–76.

Jacobucci, R., Grimm, K. J., and McArdle, J. J. (2016). Regularized structural equation modeling. Structural Equation Modeling: A Multidisciplinary Journal, 23(4):555–566.

Jin, S., Moustaki, I., and Yang-Wallentin, F. (2018). Approximated penalized maximum likelihood for exploratory factor analysis: An orthogonal case. Psychometrika, 83(3):628–649.

Jonker, F., de Looff, P., van Erp, S., Nijman, H., & Didden, R. (2023). The adaptive ability performance test (ADAPT): A factor analytic study in clients with intellectual disabilities. Journal of Applied Research in Intellectual Disabilities, 36(1), 3–12.

MacCallum, R. C., Roznowski, M., and Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. Psychological Bulletin, 111(3):490–504.

Magis, D., Tuerlinckx, F., and Boeck, P. D. (2015). Detection of differential item functioning using the lasso approach. Journal of Educational and Behavioral Statistics, 40(2):111–135.

Piironen, J., & Vehtari, A. (2016). Comparison of Bayesian predictive methods for model selection. Statistics and Computing, 27(3), 711–735.

Robitzsch, A. (2020). Regularized latent class analysis for polytomous item responses: An application to SPM-LS data. Journal of Intelligence, 8(3):30.

Serang, S., Jacobucci, R., Brimhall, K. C., and Grimm, K. J. (2017). Exploratory mediation analysis via regularization. Structural Equation Modeling: A Multidisciplinary Journal, 24(5):733–744.

Trendafilov, N. T., Fontanella, S., and Adachi, K. (2017). Sparse exploratory factor analysis. Psychometrika, 82(3):778–794.

Tutz, G. and Schauberger, G. (2013). A penalty approach to differential item functioning in rasch models. Psychometrika, 80(1):21–43.

van Erp, S., Oberski, D. L., and Mulder, J. (2019). Shrinkage priors for Bayesian penalized regression. Journal of Mathematical Psychology, 89:31–50.

van Kesteren, E.-J. and Oberski, D. L. (2019). Exploratory mediation analysis with many potential mediators. Structural Equation Modeling: A Multidisciplinary Journal, 26(5):710–723.

Zhang, L., Pan, J., Dubé, L., & Ip, E. H. (2021). blcfa: An R package for Bayesian model modification in confirmatory factor analysis. Structural Equation Modeling: A Multidisciplinary Journal, 28(4), 649-658.

Zhang, L., Pan, J., & Ip, E. H. (2021). Criteria for Parameter Identification in Bayesian Lasso Methods for Covariance Analysis: Comparing Rules for Thresholding, p-value, and Credible Interval. Structural Equation Modeling: A Multidisciplinary Journal, 28(6), 941-950.

Zhao, Y. and Luo, X. (2016). Pathway lasso: estimate and select sparse mediation pathways with high dimensional mediators. arXiv preprint arXiv:1603.07749.