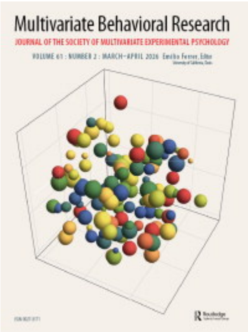

Routledge
Taylor & Francis Group

Multivariate Behavioral Research

ISSN: 0027-3171 (Print) 1532-7906 (Online) Journal homepage: www.tandfonline.com/journals/hmbr20

Bayesian Variable Selection via Shrinkage Priors in Growth Mixture Models

Ihnwhi Heo, Fan Jia & Sarah Depaoli

MULTIVARIATE BEHAVIORAL RESEARCH
https://doi.org/10.1080/00273171.2026.2673279

ABSTRACT

Bayesian Variable Selection via Shrinkage Priors in Growth Mixture Models

Ihnwhi Heo, Fan Jia, and Sarah Depaoli

Department of Psychological Sciences, University of California, Merced, CA, USA

Growth mixture models (GMMs; Muthén & Shedden, 1999) capture unobserved heterogeneity in longitudinal processes by modeling latent classes that represent qualitatively distinct patterns of change. Evaluating antecedents of heterogeneous developmental trajectories enhances the practical utility of GMMs through

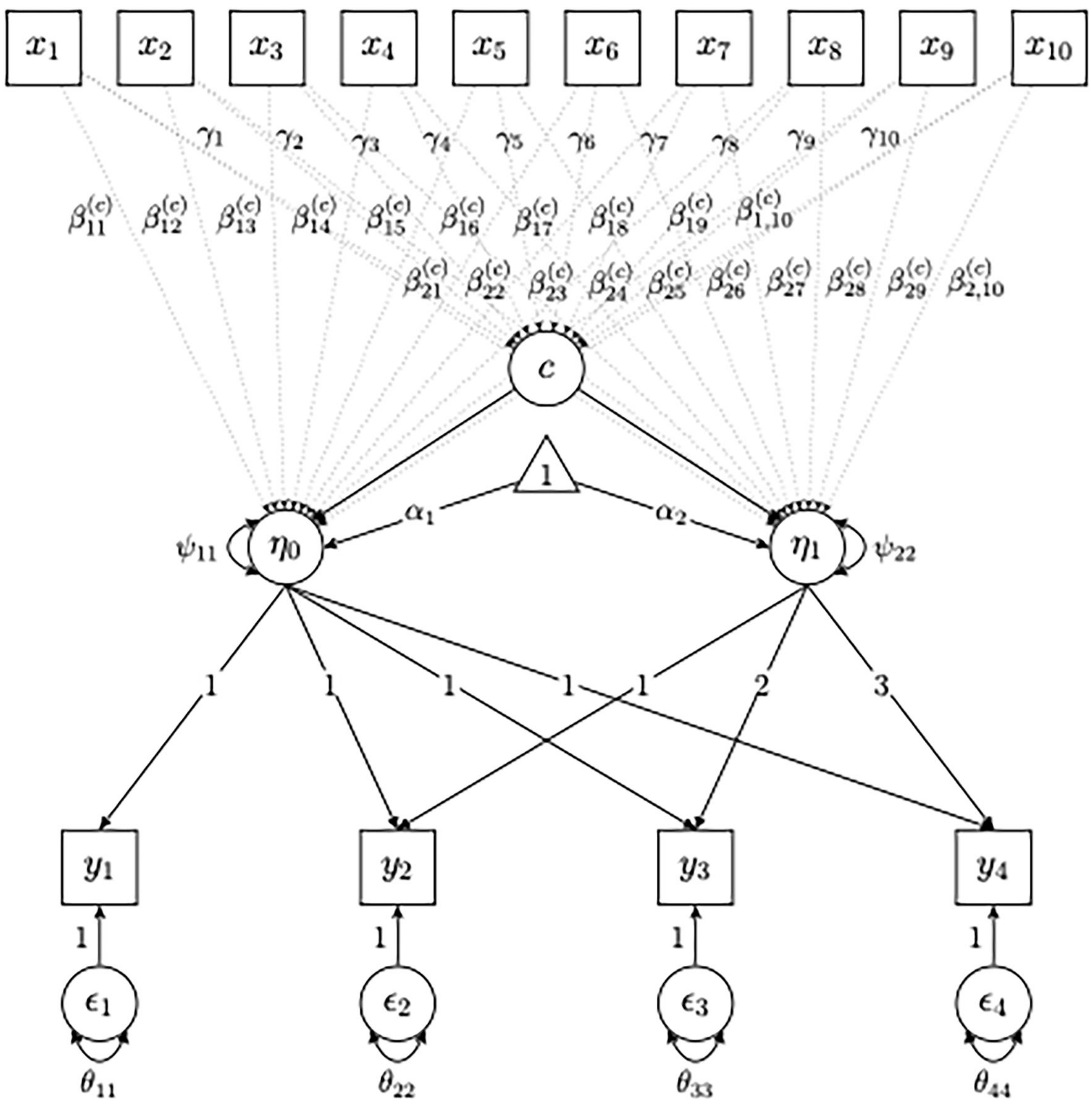

Figure 1. Path diagram representing the data-generating model. The latent class variable is denoted by c. The latent intercept and slope are denoted by η_0 and η_1, respectively. These three latent variables are predicted by ten covariates ($x_1 - x_{10}$). Four repeated measures ($y_1 - y_4$) are the observed indicators of the growth process.

CONTACT Ihnwhi Heo, ihnwhi.heo@gmail.com Department of Psychological Sciences, University of California, Merced, 5200 N. Lake Road, Merced, CA 95343, USA.

the inclusion of covariates that predict class membership and growth factors (Petras & Masyn, 2010). This naturally raises a looming question: How can we select important covariates in GMMs? Including irrelevant or sample-dependent predictors risks overfitting, poor generalizability, and misclassification. Traditional selection strategies (e.g., forward selection, backward elimination) are ill-suited for mixture modeling and can undermine the credibility of results. Long recognized as "the most important problem in statistics" (Hesterberg et al., 2008), variable selection has, paradoxically, received little methodological attention in GMM research.

This study leverages the strengths of the Bayesian framework for handling GMMs (Depaoli, 2013; Depaoli et al., 2017; Heo et al., 2024) and achieves variable selection through shrinkage priors. Shrinkage priors induce sparsity by specifying hyperpriors that control the degree of shrinkage applied to estimated effects, shrinking small coefficients toward zero while maintaining large coefficients (Kaplan, 2023). We evaluate the performance of shrinkage priors in variable selection within GMMs through simulation.

We manipulate six design factors: sample size (150, 300, 900), class separation (Mahalanobis distance: 3.0, 4.0), class proportion (equal, unequal), covariate effect pattern (homogeneous, heterogeneous), covariate correlation (0.2, 0.8), and shrinkage prior (ridge, lasso, hyperlasso, elastic net, horseshoe, regularized horseshoe, spike-and-slab). The data-generating model (Figure 1) is a two-class linear GMM with four time points and ten predictors, whose effect sizes vary by condition (one large, two medium, two small, and five null). Covariates predict latent class variable *via* multinomial logistic regression ($\gamma_1 - \gamma_{10}$) and within-class growth factors *via* linear regression ($\beta_{11}^{(c)} - \beta_{1,10}^{(c)}$ and $\beta_{21}^{(c)} - \beta_{2,10}^{(c)}$). The outcome measure is variable selection accuracy, defined as the proportion of variables with different effect sizes that are selected.

Figure 2 displays variable selection accuracy based on posterior medians for the smallest sample size condition, a class separation of 3.0, and unequal class proportions. For Class 1 (larger class) with a covariate correlation of 0.2, ridge and lasso exhibited weaker shrinkage and tended to select more null effects, with ridge producing the highest false selections. In contrast, horseshoe, regularized horseshoe, and spike-and-slab induced stronger sparsity, with spike-and-slab producing the strongest shrinkage and the lowest false selections. These results were similar across covariate effect patterns. For Class 2 (smaller class), selection rates were lower overall, and the results resembled those from Class 1 under the homogeneous pattern. However, under the heterogeneous pattern, false selections increased for horseshoe, regularized horseshoe, and spike-and-slab. When the covariate correlation

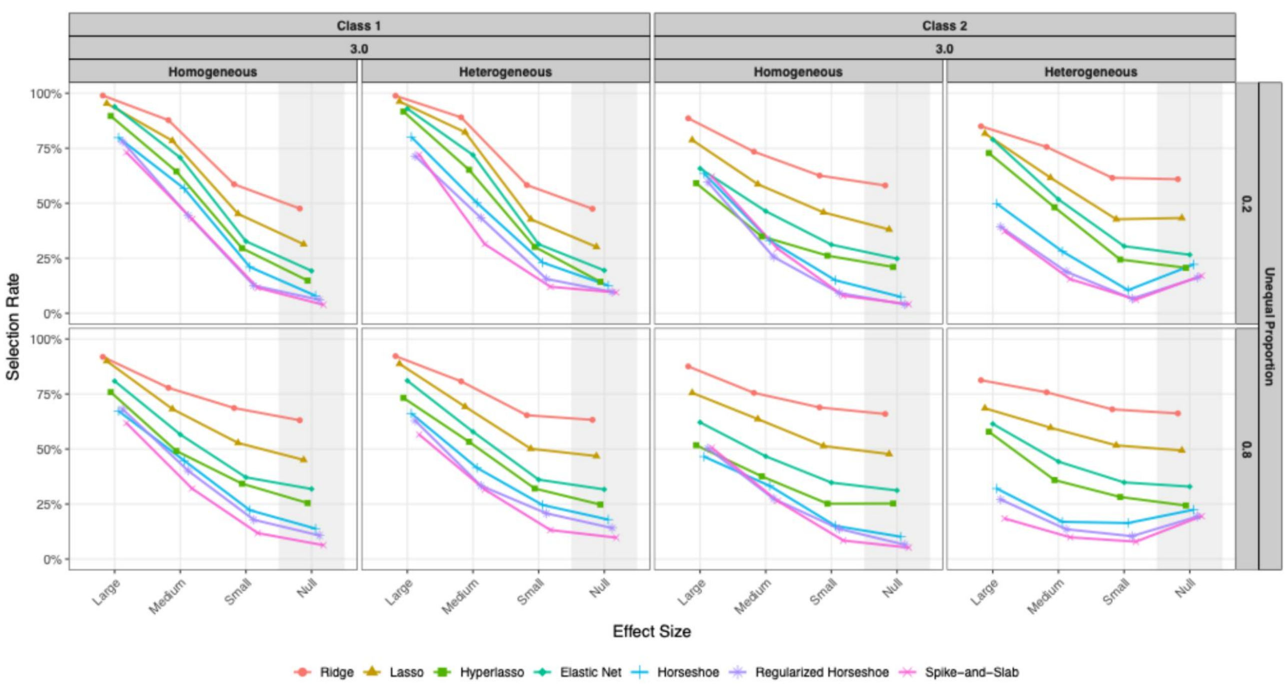

Figure 2. Selection rates for covariates predicting the latent slope when sample size = 150, class separation = 3.0, and class proportions are unequal. Rows correspond to covariate correlations, and columns represent covariate effect patterns across the two latent classes. The x-axis represents effect sizes, the y-axis indicates selection rates, and each line corresponds to a shrinkage prior. Selection rates for variables with null effects are shaded in gray to denote false selections.

was 0.8, selection rates decreased in both classes, particularly for variables with large and medium effects.

The contributions of this study are evident. Methodologically, we provide a systematic evaluation of Bayesian shrinkage priors in mixture modeling and fill a research gap in variable selection for GMMs. For applied researchers, the findings offer guidance for choosing priors and highlight caveats, improving the implementation of GMMs.

ORCID

Ihnwhi Heo http://orcid.org/0000-0002-6123-3639
Fan Jia http://orcid.org/0000-0003-3855-532X
Sarah Depaoli http://orcid.org/0000-0002-1277-0462

References

Depaoli, S. (2013). Mixture class recovery in GMM under varying degrees of class separation: Frequentist versus Bayesian estimation. *Psychological Methods*, *18*(2), 186–219. https://doi.org/10.1037/a0031609

Depaoli, S., Yang, Y., & Felt, J. (2017). Using Bayesian statistics to model uncertainty in mixture models: A sensitivity analysis of priors. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(2), 198–215. https://doi.org/10.1080/10705511.2016.1250640

Heo, I., Depaoli, S., Jia, F., & Liu, H. (2024). Bayesian approach to piecewise growth mixture modeling: Issues and applications in school psychology. *Journal of School Psychology*, *107*, 101366. https://doi.org/10.1016/j.jsp.2024.101366

Hesterberg, T., Choi, N. H., Meier, L., & Fraley, C. (2008). Least angle and $\ell 1$ penalized regression: A review. *Statistics Surveys*, *2*(none), 61–93. https://doi.org/10.1214/08-SS035

Kaplan, D. (2023). *Bayesian statistics for the social sciences.* (2nd ed.) Guilford Press.

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, *55*(2), 463–469. https://doi.org/10.1111/j.0006-341X.1999.00463.x

Petras, H., & Masyn, K. (2010). General growth mixture analysis with antecedents and consequences of change. In Piquero, A. R. & Weisburd, D. (Eds.), *Handbook of quantitative criminology.* (1st ed., pp. 69–100) Springer.