Political Analysis (2025), **33**, 171–177
doi:10.1017/pan.2024.30

PA

LETTER

Adding Regularized Horseshoes to the Dynamics of Latent Variable Models

Garret Binding[1,2] and Piotr Koc[3,4]

[1]Institute of Political Science, University of Basel, Basel, Switzerland; [2]Department of Political Science, University of Zurich, Zurich, Switzerland; [3]GESIS – Leibniz Institute for the Social Sciences, Mannheim, Germany; [4]Institute of Philosophy and Sociology of the Polish Academy of Sciences, Warsaw, Poland

Corresponding author: Garret Binding; Email: garret.binding@unibas.ch

(Received 27 February 2024; revised 5 December 2024; accepted 6 December 2024; published online 17 January 2025)

Abstract

Dynamic latent variable models generally link units' positions on a latent dimension over time *via* random walks. Theoretically, these trajectories are often expected to resemble a mixture of periods of stability interrupted by moments of change. In these cases, a prior distribution such as the regularized horseshoe—that allows for both stasis and change—can prove a better theoretical and empirical fit for the underlying construct than other priors. Replicating Reuning, Kenwick, and Fariss (2019), we find that the regularized horseshoe performs better than the standard normal and the Student's t-distribution when modeling dynamic latent variable models. Overall, the use of the regularized horseshoe results in more accurate and precise estimates. More broadly, the regularized horseshoe is a promising prior for many similar applications.

Keywords: Bayesian inference; item response theory; latent variable models

Edited by: Jeff Gill

1. Introduction

When modeling dynamic trends in latent variable models, researchers typically use random walks to link observations from one time-period to the next: tomorrow's position on a latent dimension is today's position plus a random effect (see, e.g., Caughey and Warshaw 2015; König, Marbach, and Osnabrügge 2013; Martin and Quinn 2002; Schnakenberg and Fariss 2014). Often, the expected trajectory of the political constructs modeled in such a manner (such as countries' democracy scores) is a mixture of stasis and change: periods of stability are interrupted by moments of rupture (see also Reuning *et al.* 2019, 504).

In this letter, we suggest the use of the regularized horseshoe (RHS; Piironen and Vehtari 2017) as a prior distribution for dynamic latent variable models. Stemming from the Bayesian literature on sparsity-inducing priors (Carvalho, Polson, and Scott 2010; Polson and Scott 2011), draws from the regularized horseshoe distribution are shrunk towards 0 (stasis) or allow for deviations from 0 (change). The correspondence between the prior's properties and expectations about the construct makes the RHS an ideal candidate for these applications.

Below, we introduce the regularized horseshoe as a useful addition to the distributions often employed in political science[1] and highlight its utility in the domain of latent variable models. We

[1]To the best of our knowledge, the RHS has so far not been discussed or used in political science. It's predecessor—the horseshoe—was employed by Ratkovic and Tingley (2017) and considered by Park and Yamauchi (2023).

© The Author(s), 2025. Published by Cambridge University Press on behalf of The Society for Political Methodology.
This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/pan.2024.30 Published online by Cambridge University Press

build on Reuning *et al.* (2019, from hereon: RKF), who assessed the performance of the standard normal and the Student's t-distribution (the two distributions generally used in these applications) in a similar endeavor. We show that using the RHS leads to better results than using either of the other two distributions. Model specifications using the RHS produce more accurate and precise estimates, and better model fit than those using the other distributions.

2. Motivation

The standard normal distribution as well as the Student's t-distribution are sensible starting points for the random walks of dynamic models. Both concentrate their probability around 0 (stasis), while also allowing for deviations from 0 (change). Their empirical density functions are shown in Figure 1. The key difference between the tails of these two distributions is best visible in panels (b) and (c). As the tails of the t-distribution are determined *via* the degrees of freedom (ν), their decay can be more gradual than those of a standard normal (here, we show draws from $StudentT(4,0,1)$ in line with RKF, where the parameters of the t-distribution are ν, location, and scale).

While the two distributions differ in their tails, they are similar in their center: a gradual decline from a peak at 0 assigns a lot of probability to small deviations from the center. Within the Bayesian literature, distributions such as the horseshoe have been suggested that exhibit a pronounced peak at 0 to induce sparsity in the posterior distributions of parameters (Carvalho *et al.* 2010; Polson and Scott 2011). Essentially, the horseshoe prior is a continuous mixture of two components with one component (potentially) shrinking values towards 0 (τ), and another component allowing for large deviations from 0 (λ).[2]

More recently, Piironen and Vehtari 2017 proposed a regularized version of the horseshoe (RHS) to address the heavy tails of the horseshoe prior. By adding a hyperparameter c^2 to the horseshoe's parameterization, the extent of deviations from 0 can be regularized. Formally, the regularized horseshoe distribution for a parameter γ is constructed in the following manner:

$$\gamma_i \sim N(0, \tau^2 \tilde{\lambda}_i^2), \quad \tilde{\lambda}_i^2 = \frac{c^2 \lambda_i^2}{c^2 + \tau^2 \lambda_i^2},$$
$$\tau \sim StudentT^{+}(1,0,1),$$
$$\lambda_i \sim StudentT^{+}(1,0,1) \quad \forall i = 1, \ldots, I,$$

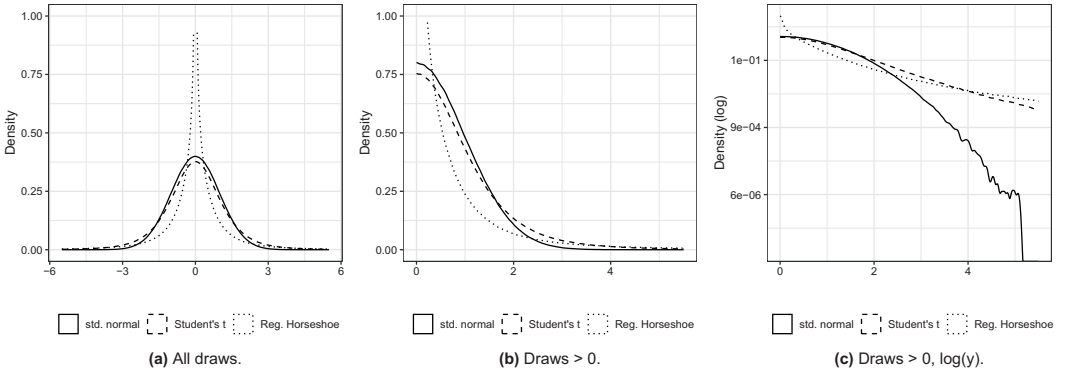


Figure 1. Empirical density functions.

[2]Here, our discussion focuses on the RHS. In the Supplementary Information (SI), we also discuss and assess the original horseshoe and discrete spike-and-slab priors (George and McCulloch 1993; Mitchell and Beauchamp 1988).

https://doi.org/10.1017/pan.2024.30 Published online by Cambridge University Press

where i indexes elements of the vector γ. As discussed by Piironen and Vehtari 2017, c^2 is given a prior too:

$$c^2 \sim InvGamma(\alpha, \beta), \quad \alpha = \nu/2, \quad \beta = \nu s^2/2,$$

which corresponds to a $StudentT^+(\nu, 0, s^2)$ prior. We set $\nu = s = 1$, so that $c^2 \sim InvGamma(0.5, 0.5)$.[3] If c^2 is dropped and $\gamma_i \sim N(0, \tau^2\lambda_i^2)$, the prior corresponds to the original horseshoe.

The distribution of the RHS is also shown in Figure 1. More probability is assigned to values at or around 0, with a far steeper decline from the spike at the center than the more gradual descent of the other distributions. Simultaneously, the tails of the RHS are fatter, exhibiting similar, and then lower decay rates than those of the t-distribution (visible in panel (c)). These two attributes of this distribution—a spike at 0 combined with fat tails—make the RHS an ideal *a priori* candidate to model the theoretically expected mixture of stasis and change in latent constructs over time. Its use should result in a posterior distribution that is both more accurate (as it better reflects the expected trajectory) and precise (as its probability is more concentrated) than the other two distributions. Both aspects are important in applications, especially if estimates are used in subsequent models that propagate measurement uncertainty (rather than focusing on point estimates only; see Tai, Hu, and Solt 2024).

3. Model Specification

All subsequent empirical analyses are based on the model specifications and priors described in this section (unless mentioned otherwise). In line with RKF, we use Bayesian inference implemented in Stan (Carpenter *et al.* 2017) to approximate the posterior distribution of our model.

Following RKF, the likelihood for a unidimensional binary item response theory (IRT) model for units $i = 1, \ldots, I$, items $k = 1, \ldots, K$, and time-points $t = 1, \ldots, T$ is

$$\mathcal{L} = \prod_{i,t=1}^{I,T} \prod_{k=1}^{K} \Lambda(\alpha_k - \beta_k \theta_{i,t})^{y_{ikt}} (1 - \Lambda(\alpha_k - \beta_k \theta_{i,t})^{1-y_{ikt}}),$$

where $\theta_{i,t}$ is a unit's ideal point on a latent dimension at a given time-point, α_k and β_k are an item's difficulty and discrimination parameters, and Λ is the logistic function. y_{ikt} is an observed binary outcome for a unit × item × time-point combination. The latent model is identified by constraining the discrimination parameters to be positive (rotational invariance), and by fixing the mean and standard deviation of the ideal points to 0 and 1 (local and scalar invariance; see Bafumi *et al.* 2005). We assign standard (half) normal distributions to the vectors $\vec{\alpha}$ and $\vec{\beta}$ of length K, and to the vector $\vec{\theta}_{t=1}$ of length I: $\vec{\alpha} \sim N(0,1)$, $\vec{\beta} \sim N^+(0,1)$, and $\vec{\theta}_{t=1} \sim N(0,1)$.

The dynamic nature of the model is ensured by specifying yesterday's $\theta_{i,t-1}$ as the expectation of today's $\theta_{i,t}$, so that:

$$\theta_{i,t} = \theta_{i,t-1} + \gamma_{i,t}\sigma \quad \forall i = 1, \ldots, I \quad \& \quad \forall t = 2, \ldots, T,$$

where $\gamma_{i,t}$ are the random effects over time and σ is a smoothing hyperparameter that governs how much the past informs the present with a standard half normal prior: $\sigma \sim N^+(0,1)$.[4]

The primary difference in the model comparisons below are the priors used for $\gamma_{i,t}$.[5] We compare the results from models using three different prior distributions for $\gamma_{i,t}$: (i) a specification using the standard

[3] Both ν and s could be set to other values. We chose these values so that τ, λ_i, and c^2 have the same Student's t prior (equivalent to a Cauchy distribution), and to thereby allow for large deviations from 0 across all three parameters *via* their prior. Future research could explore the results' sensitivity to these choices.

[4] RKF assign (half) normal distributions with a standard deviation of 3 to α, β, and σ, while we assign a standard deviation of 1. This difference does not matter within this comparison.

[5] We omit σ when using the RHS because τ performs a similar function in the RHS prior and its inclusion would lead to identification issues.

https://doi.org/10.1017/pan.2024.30 Published online by Cambridge University Press

normal distribution ($N(0,1)$), (ii) one using the Student's t-distribution (with $\nu = 4$; $StudentT(4,0,1)$), and (iii) our suggested specification with the regularized horseshoe (as detailed in Section 2 above).

4. Results

In this section, we replicate the simulation study of RKF and their re-analysis of Pemstein, Meserve, and Melton 2010, who estimated the trajectories of countries' democracy scores—a construct likely to exhibit stasis and change across time.[6] We limit ourselves to a selection of quantities of interest here (focusing on the accuracy and precision of estimates as well as model fit). We present more background and quantities of interest in the SI (Section B focuses on the simulation study, Section C on the application). The results there corroborate the results here.

4.1. Simulation Study

RKF simulate data according to the model specification in Section 3. $\theta_{i,t}$ is either $\theta_{i,t-1} + \gamma_{i,t}\sigma$ (and $\gamma_{i,t} \sim N(0,1)$) or $\theta_{i,t}$ is a new draw from $N(0,1)$. This induces breaks in the otherwise more gradual trajectory of θ over time. RKF vary the extent of smoothing over time (*via*σ) as well as the probability of a break, set $I = T = 50$, and $K = 5$, and generate 25 datasets per condition. This results in a total of 225 datasets.

The main results are shown in Figure 2. We focus on four quantities of interest: the correlation between $\hat{\theta}$ and θ (using Pearson's ρ), the root mean squared error of $\hat{\theta}$, the size of the standard errors around $\hat{\theta}$, and the expected log pointwise predictive density (ELPD) based on leave-one-out cross-validation (LOO-CV) as a measure of model fit (Vehtari *et al.* 2017). We set the RHS as the reference distribution, subtract the values of the other two specifications (indicated along the vertical axis), and calculate the median relative improvement of the RHS (shown alongside the median line). Correlation and model fit using the RHS are larger, while RMSE and standard errors are smaller. The relative improvements in RMSE and SE are larger (between 2.3%–7.3%) than in correlation or model fit (between 0.2%–2.3%). The improvements are larger when the probability of a shock is larger and σ is smaller (see Figures 8 and 9 in the SI). The RHS outperforms both the standard normal and the Student's t-distribution in the simulation study.

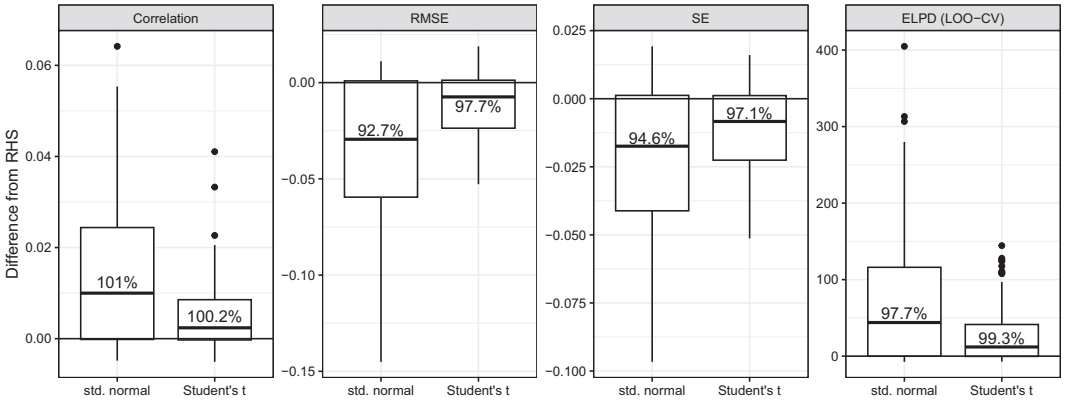


Figure 2. Simulation results. Differences shown along the vertical axis, relative improvement alongside the median lines. For example, the correlation estimated *via* the RHS is 1% larger than *via* the standard normal. Note that the difference for the ELPD is positive although the relative improvement is smaller than 100% because ELPD is measured on a negative scale.

[6] We omit RKF's replication of Martin and Quinn 2002 for brevity and because the lack of shared items from one time-period to the next (as Supreme Court Justices are faced with new cases every session) necessitates a series of rigid constraints on parameters.

https://doi.org/10.1017/pan.2024.30 Published online by Cambridge University Press

4.2. Revisiting Pemstein et al. 2010

RKF estimate a dynamic model of countries' democracy scores based on the static model of Pemstein *et al.* 2010 (*via* an ordered logit specification). Here, we first focus on model fit *via* the ELPD (based on LOO-CV). The differences between specifications are large and statistically significant. The specification using the RHS performs better than the one using the t-distribution by 1781 units (SE: 58), and better than the one using the normal distribution by 6023 (SE: 125). Using the RHS results in a better model fit than using either of the other two distributions. Second, we turn to predictive accuracy: how often do predictions equal the observed values of democracy indicators? Here, too, the RHS performs better than the other specifications. Overall, the predictive accuracy using the RHS is 62.9% (95% credible interval: 62.6%–63.2%). It is 61.7% (95% ci: 61.4%–62%) for the t-distribution and 58.5% (95% ci: 58.2%–58.8%) for the standard normal distribution. A breakdown by indicator is provided in Figure 11 in the SI. Using the RHS results in a higher predictive accuracy than using the other two priors.

Third, we revisit the two countries discussed by RKF: the Philippines and Afghanistan (Figure 3). The countries exhibit longer periods of stasis (e.g., after the Marcos regime in the Philippines) as well as rapid change (e.g., the Saur Revolution in Afghanistan). Both the RHS and the t-distribution more aptly recover abrupt large changes in latent democracy scores than the normal distribution does. In addition, the RHS appears to have narrower credible intervals than the t-distribution does in some periods of stability (e.g., before and after the Marcos regime).

What if we consider the results of all country-year observations rather than this selection? We turn to this in Figure 4 in which we compare the magnitude of year-over-year changes in countries' mean democracy scores and the credible intervals of their scores between models. We subtract the values

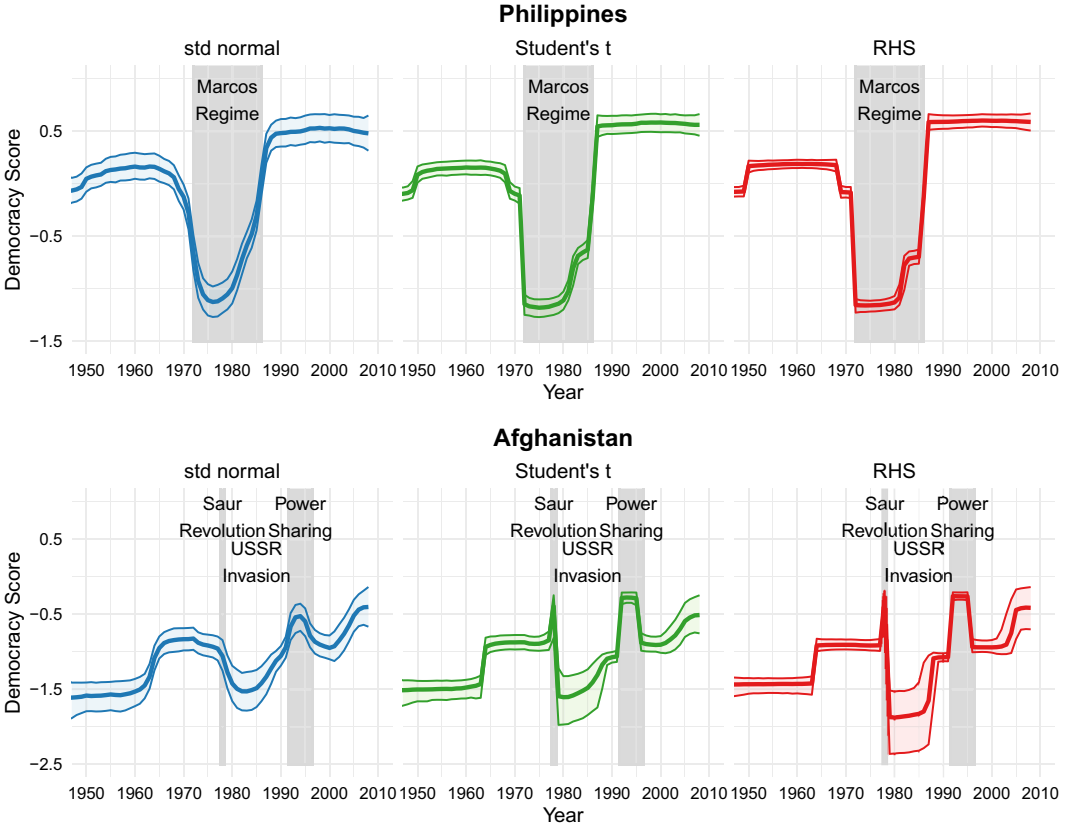
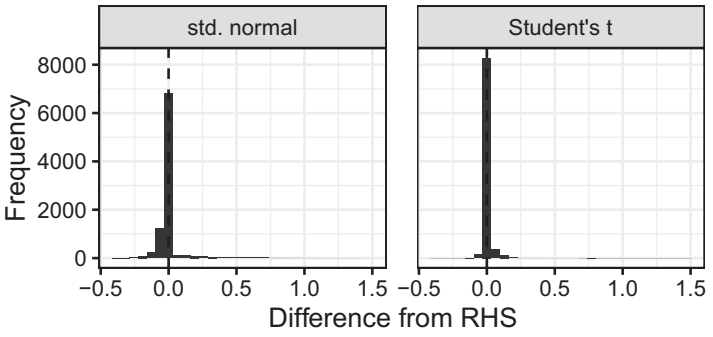
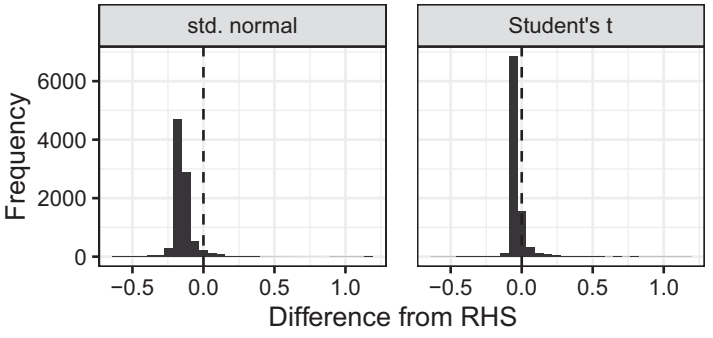


Figure 3. Trajectories of the Philippines and Afghanistan (1946–2008; mean and 95% ci).

https://doi.org/10.1017/pan.2024.30 Published online by Cambridge University Press



(**a**) Year-over-year changes.



(**b**) Credible intervals.

Figure 4. Year-over-year changes and credible intervals by model.

from the standard normal and t-distribution specifications from those based on the RHS. There is little systematic difference in the year-over-year changes across models (panel (a)). In contrast, the credible intervals of the RHS are systematically smaller than those of either other distribution (panel (b)). On average, the intervals of the RHS are 44% the size of those using the standard normal distribution, and 73% of those based on the Student's t-distribution. If researchers use estimates from similar models in further explanatory models and take measurement uncertainty into account (as suggested by Tai *et al.* 2024), then the RHS will produce less variance and more precision to begin with.

5. Conclusion

In this letter, we introduce the regularized horseshoe as a prior to political science, and discuss its theoretical and empirical utility in the domain of dynamic latent variable models where periods of stasis can be interrupted by moments of rapid change. In our replication of RKF, the RHS prior performed better than either the standard normal or the Student's t-distribution. The use of the RHS leads to more accurate and precise estimates as well as better model fit.

Our substantive focus has been on modeling countries' democracy scores over time. It is likely that the RHS prior would perform similarly well in other areas where dynamic processes are characterized by a punctuated equilibrium such as, e.g., the trajectories of party positions, public opinion, or budgetary allocations. In addition, due to its sparsity-inducing nature, the RHS could be a candidate for applications where researchers wish to find a balance between a model's flexibility and rigidity *via* a prior distribution, such as in the case of differential item functioning (see, e.g., Binding, Koedam, and Steenbergen 2024).

https://doi.org/10.1017/pan.2024.30 Published online by Cambridge University Press

Acknowledgments. We thank Marco R. Steenbergen and Artur Pokropek for their insightful feedback, as well as the reviewers and editor.

Competing Interest. None.

Funding Statement. The work of the first author was supported by the Swiss National Science Foundation (Grant No. 212396). The work of the second author was supported by the Polish National Science Centre (Sonata Bis-10: UMO-2020/38/E/HS6/00302).

Data Availability Statement. Replication code for this article is available in Binding and Koc (2024) at https://doi.org/10.7910/DVN/G2VRQH.

Supplementary Material. For supplementary material accompanying this paper, please visit https://doi.org/10.1017/pan.2024.30.

References

Bafumi, J., A. Gelman, D. K. Park, and N. Kaplan. 2005. "Practical Issues in Implementing and Understanding Bayesian Ideal Point Estimation." *Political Analysis* 13 (2): 171–187.

Binding, G., and P. Koc. 2024. "Replication Data for: Adding Regularized Horseshoes to the Dynamics of Latent Variable Models." Version DRAFT VERSION. https://doi.org/10.7910/DVN/G2VRQH.

Binding, G., J. Koedam, and M. R. Steenbergen. 2024. "The Comparative Meaning of Political Space: A Comprehensive Modeling Approach." *Political Science Research and Methods* 12 (3): 643–51. https://doi.org/10.1017/psrm.2023.16.

Carpenter, B., et al. 2017. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software* 76 (1). https://doi.org/10.18637/jss.v076.i01.

Carvalho, C. M., N. G. Polson, and J. G. Scott. 2010. "The Horseshoe Estimator for Sparse Signals." *Biometrika* 97 (2): 465–480. https://doi.org/10.1093/biomet/asq017.

Caughey, D., and C. Warshaw. 2015. "Dynamic Estimation of Latent Opinion Using a Hierarchical Group-Level IRT Model." *Political Analysis* 23 (2): 197–211.

George, E. I., and R. E. McCulloch. 1993. "Variable Selection via Gibbs Sampling." *Journal of the American Statistical Association* 88 (423): 881–889.

König, T., M. Marbach, and M. Osnabrügge. 2013. "Estimating Party Positions across Countries and Time—A Dynamic Latent Variable Model for Manifesto Data." *Political Analysis* 21 (4): 468–491.

Martin, A. D., and K. M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999." *Political Analysis* 10 (2): 134–153.

Mitchell, T. J., and J. J. Beauchamp. 1988. "Bayesian Variable Selection in Linear Regression." *Journal of the American Statistical Association* 83 (404): 1023–1032.

Park, J. H., and S. Yamauchi. 2023. "Change-Point Detection and Regularization in Time Series Cross-Sectional Data Analysis." *Political Analysis* 31 (2): 257–277.

Pemstein, D., S. A. Meserve, and J. Melton. 2010. "Democratic Compromise: A Latent Variable Analysis of Ten Measures of Regime Type." *Political Analysis* 18 (4): 426–449.

Piironen, J., and A. Vehtari. 2017. "Sparsity Information and Regularization in the Horseshoe and Other Shrinkage Priors." *Electronic Journal of Statistics* 11 (2): 5018–5051. https://doi.org/10.1214/17-EJS1337SI.

Polson, N. G., and J. G. Scott. 2011. "Shrink Globally, Act Locally: Sparse Bayesian Regularization and Prediction." In *Bayesian Statistics 9*, 501–538. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199694587.003.0017. With discussions by Bertrand Clark, C. Severinski, Merlise A. Clyde, Robert L. Wolpert, Jim e. Griffin, Philiip J. Brown, Chris Hans, Luis R. Pericchi, Christian P. Robert and Julyan Arbel. . .

Ratkovic, M., and D. Tingley. 2017. "Sparse Estimation and Uncertainty with Application to Subgroup Analysis." *Political Analysis* 25 (1): 1–40.

Reuning, K., M. R. Kenwick, and C. J. Fariss. 2019. "Exploring the Dynamics of Latent Variable Models." *Political Analysis* 27 (4): 503–517.

Schnakenberg, K. E., and C. J. Fariss. 2014. "Dynamic Patterns of Human Rights Practices." *Political Science Research and Methods* 2 (1): 1–31.

Tai, Y. C., Y. Hu, and F. Solt. 2024. "Democracy, Public Support, and Measurement Uncertainty." *American Political Science Review* 118 (1): 512–518. https://doi.org/10.1017/S0003055422000429. https://www.cambridge.org/core/product/identifier/S0003055422000429/type/journal_article.

Vehtari, A., A. Gelman, and J. Gabry. 2017. "Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC." *Statistics and Computing* 27: 1413–1432.

Cite this article: Binding, G. and Koc, P. (2025). Adding Regularized Horseshoes to the Dynamics of Latent Variable Models. *Political Analysis*, 171–177. https://doi.org/10.1017/pan.2024.30

https://doi.org/10.1017/pan.2024.30 Published online by Cambridge University Press