Supplementary Materials for *Sensitivity Analysis of the Bayesian Dirichlet Auto-Regressive Moving Average Model*

Harrison Katz* Liz Medina Robert E. Weiss

Catalogue of Items

Table S1: Parameter estimation—Simulation Study 4 (correct specification);
Table S2: Parameter estimation—Simulation Study 5 (overfitting);
Table S3: Parameter estimation—Simulation Study 6 (underfitting);
Table S4: Forecast performance summary;
Table S5: Forecast–error ratios (across scenarios);
Table S6: Forecast–error ratios (within scenarios).

*harrisonekatz@gmail.com

Supplementary Material: Additional Simulation Studies

Here we present three additional simulation studies conducted under the same general framework as in the main text. The only difference is the use of alternative coefficient matrices for the data-generating process (DGP). As before, each study is based on 50 simulated datasets, and we examine the same set of priors: Informative Normal, Horseshoe, Laplace, Spike-and-Slab, and Hierarchical. We focus on the same three scenarios: (1) correct model specification, (2) intentional overfitting, and (3) intentional underfitting.

Data-Generating Process

In these supplementary simulations, we use a DARMA(2,1) model with the following coefficient matrices:

$$\mathbf{A}_1 = \begin{bmatrix} 0.80 & -0.08 & -0.08 & -0.09 & -0.08 \\ -0.06 & 0.70 & -0.08 & 0.06 & 0.06 \\ -0.06 & 0.07 & 0.90 & 0.05 & -0.09 \\ 0.07 & 0.09 & 0.07 & 0.85 & 0.08 \\ 0.05 & -0.09 & -0.07 & -0.09 & 0.75 \end{bmatrix} \quad \mathbf{A}_2 = \begin{bmatrix} -0.30 & -0.07 & -0.06 & -0.08 & 0.09 \\ 0.06 & -0.20 & -0.07 & -0.07 & -0.09 \\ 0.05 & -0.07 & -0.25 & 0.07 & -0.05 \\ -0.06 & 0.08 & -0.09 & -0.15 & 0.07 \\ -0.05 & 0.06 & 0.10 & 0.07 & -0.20 \end{bmatrix},$$

$$\mathbf{B}_1 = \begin{bmatrix} 0.50 & -0.06 & -0.06 & 0.05 & 0.09 \\ 0.07 & 0.40 & 0.07 & 0.08 & -0.06 \\ -0.10 & -0.05 & 0.45 & 0.08 & 0.10 \\ 0.09 & 0.08 & -0.09 & 0.35 & 0.10 \\ 0.09 & -0.08 & 0.06 & -0.09 & 0.40 \end{bmatrix}.$$

As in the main study, we set $\phi = 500$ and use a 5-component vector for the covariate effect $\boldsymbol{\beta}^\top = (0.1, -0.05, 0.03, -0.02, 0.04)$. The initial two observations of the compositional series $\mathbf{y}_t$ are generated from a Dirichlet($\mathbf{1}$) distribution to start the recursion.

Study S1: Correct Model Specification (DARMA(2,1))

Parameter Estimation

When fitting the correctly specified DARMA(2,1) model under these new coefficient matrices, the patterns are similar to those in the main text. Horseshoe and Laplace priors successfully induce shrinkage on small coefficients while preserving large signals, resulting in relatively low RMSE and near-nominal coverage for both VAR and VMA parameters. Hierarchical priors again strike a balance, offering moderate shrinkage without overly broad credible intervals. Spike-and-Slab priors tend to more aggressively zero out parameters, sometimes overshrinking moderate-size coefficients. Informative Normal priors perform adequately but do not penalize unnecessary complexity as strongly.

Forecasting Performance

Forecasting accuracy remains high for the correctly specified model. Forecast RMSE values are similar across priors, with Horseshoe and Hierarchical priors again offering slight improvements in predictive accuracy. The differences are subtle, as an appropriate model structure provides a strong baseline. Shrinkage priors mainly help by stabilizing parameter estimates and reducing uncertainty.

Study S2: Intentional Overfitting (B-DARMA(4,2))

Parameter Estimation

With overfitting (fitting B-DARMA(4,2) instead of DARMA(2,1)), the key challenge is the presence of additional VAR and VMA terms that are truly zero. As in the main simulations, Horseshoe priors excel at shrinking these extraneous coefficients to near zero, yielding low bias and RMSE and appropriate coverage for the truly null parameters. Spike-and-Slab also mitigates overfitting but may overly penalize some coefficients on the borderline between zero and small nonzero values. Informative Normal priors, lacking strong shrinkage,

more frequently attribute nonzero mass to spurious terms, resulting in inflated RMSE and wider intervals. Laplace and Hierarchical priors offer intermediate levels of shrinkage, improving upon non-informative priors but not matching the Horseshoe's consistently strong performance.

Forecasting Performance

Forecasting under intentional overfitting highlights the value of shrinkage priors. Horseshoe priors minimize the performance degradation caused by unnecessary lags, showing lower RMSE relative to Informative Normal priors. Laplace and Hierarchical priors also help restrain overfitting, improving forecasts compared to non-informative priors. Spike-and-Slab's aggressive thresholding can stabilize forecasts but may slightly reduce predictive performance if it eliminates coefficients too aggressively. Overall, priors that induce sparsity lead to more stable, interpretable, and accurate forecasts under overfitting conditions.

Study S3: Intentional Underfitting (B-DARMA(1,0))

Parameter Estimation

Underfitting by fitting a B-DARMA(1,0) model to data generated from DARMA(2,1) again demonstrates that no prior can compensate for missing VAR(2) and VMA(1) terms. All priors show higher bias and RMSE for the parameters that cannot be properly identified due to structural omissions in the model. Coverage rates are reduced, and credible intervals often fail to capture the true values adequately. This pattern is consistent with the main text results and confirms that when the model is fundamentally misspecified, prior choice plays a limited role in improving inference.

Forecasting Performance

In the underfitting scenario, forecasting accuracy deteriorates uniformly across priors. The lack of necessary model terms impairs the capture of true temporal dynamics, leading to systematically higher RMSE and suboptimal predictive log-likelihoods. None of the considered priors meaningfully improves forecasts when the underlying process complexity is not well represented by the chosen model order.

Summary of Supplementary Simulations

These three supplementary studies corroborate and reinforce the findings from the main simulations. Under correct model specification, all priors perform reasonably well, with Horseshoe, Laplace, and Hierarchical priors offering slight advantages in estimation and prediction. Under overfitting, Horseshoe and other shrinkage-oriented priors effectively suppress extraneous terms, preventing inflated uncertainty and degrading forecasts. Underfitting remains impervious to prior-based corrections, with all priors exhibiting similarly poor performance when key dynamics are omitted.

In combination, the supplementary results confirm that the main conclusions are robust to changes in the underlying coefficient matrices. Priors that induce shrinkage or sparsity consistently help navigate overfitting challenges, while none can overcome the fundamental limitations imposed by an underspecified model. As a result, these supplementary studies bolster our recommendation to match model complexity to the data-generating process and to employ shrinkage priors when structural uncertainty or risk of overfitting is present.

Table S1: **Parameter Estimation Summary for Simulation Study 4 (Correct Specification).** We show mean bias, RMSE, average credible interval length, and coverage for each prior, summarizing over all parameters in each of β, A_1, A_2, and B_1. Lower RMSE and shorter intervals typically indicate more effective shrinkage, while coverage near the nominal 0.95 is desirable.

Coefficient	Prior	Mean Bias	Mean RMSE	Mean CI Length	Coverage
β	Informative	0.000	0.042	0.203	0.972
	Horseshoe	-0.002	0.042	0.184	0.956
	Laplace	0.003	0.052	0.249	0.964
	Spike-Slab	0.011	0.087	0.302	0.812
	Hierarchical	-0.000	0.044	0.185	0.952
A_1	Informative	-0.038	0.170	0.616	0.897
	Horseshoe	-0.008	0.092	0.344	0.949
	Laplace	-0.011	0.105	0.520	0.987
	Spike-Slab	-0.010	0.190	0.589	0.818
	Hierarchical	-0.030	0.123	0.507	0.958
A_2	Informative	0.043	0.156	0.533	0.885
	Horseshoe	0.009	0.079	0.286	0.938
	Laplace	0.006	0.088	0.436	0.989
	Spike-Slab	0.005	0.163	0.495	0.800
	Hierarchical	0.016	0.102	0.435	0.971
B_1	Informative	0.017	0.162	0.677	0.952
	Horseshoe	-0.018	0.106	0.426	0.960
	Laplace	-0.008	0.121	0.613	0.984
	Spike-Slab	-0.003	0.227	0.665	0.795
	Hierarchical	-0.006	0.108	0.584	0.990

Table S2: Parameter Estimation Summary for Simulation Study 5 (Overfitting). This table reflects the setting where the fitted model (B-DARMA(4,2)) exceeds the true DARMA(2,1) order. We report mean bias, RMSE, credible interval length, and coverage for each prior and parameter.

Coefficient	**Prior**	**Mean Bias**	**Mean RMSE**	**Mean CI Length**	**Coverage**
	Informative	-0.012	0.043	0.207	0.956
	Horseshoe	-0.009	0.043	0.176	0.940
β	Laplace	-0.009	0.054	0.223	0.932
	Spike-Slab	-0.004	0.100	0.310	0.748
	Hierarchical	-0.010	0.042	0.194	0.960
	Informative	-0.060	0.224	0.754	0.813
	Horseshoe	-0.055	0.160	0.308	0.914
A_1	Laplace	-0.046	0.178	0.661	0.960
	Spike-Slab	-0.020	0.215	0.680	0.804
	Hierarchical	0.073	0.158	0.157	0.645
	Informative	0.033	0.177	0.761	0.925
	Horseshoe	0.032	0.092	0.222	0.895
A_2	Laplace	0.019	0.128	0.613	0.962
	Spike-Slab	0.002	0.203	0.686	0.810
	Hierarchical	-0.079	0.182	0.197	0.682
	Informative	0.017	0.143	0.674	0.968
	Horseshoe	-0.004	0.022	0.166	1.000
A_3	Laplace	0.001	0.100	0.517	0.987
	Spike-Slab	0.005	0.184	0.598	0.814
	Hierarchical	0.035	0.093	0.172	0.962
	Informative	0.003	0.113	0.475	0.950
	Horseshoe	-0.002	0.017	0.137	0.999
A_4	Laplace	-0.004	0.081	0.366	0.975
	Spike-Slab	-0.006	0.134	0.409	0.808
	Hierarchical	-0.008	0.035	0.118	0.990
	Informative	0.051	0.232	0.849	0.856
	Horseshoe	0.029	0.145	0.419	0.946
B_1	Laplace	0.040	0.203	0.770	0.951
	Spike-Slab	0.010	0.231	0.748	0.815
	Hierarchical	-0.096	0.187	0.237	0.756
	Informative	0.054	0.230	0.867	0.864
	Horseshoe	0.014	0.073	0.349	0.996
B_2	Laplace	0.026	0.173	0.758	0.971
	Spike-Slab	0.016	0.217	0.753	0.846
	Hierarchical	-0.005	0.023	0.236	1.000

Table S3: **Parameter Estimation Summary for Simulation Study 6 (Underfitting).** This table reports key metrics (mean bias, RMSE, average CI length, coverage) when crucial VAR(2) and VMA(1) terms are omitted. All priors suffer from higher errors and coverage shortfalls, indicating that structural misspecification is the dominant source of inaccuracy.

Coefficient	Prior	Mean Bias	Mean RMSE	Mean CI Length	Coverage
β	Informative	-0.006	0.044	0.228	0.972
	Horseshoe	-0.009	0.047	0.221	0.940
	Laplace	-0.006	0.054	0.333	0.984
	Spike-Slab	-0.001	0.068	0.719	0.996
	Hierarchical	-0.009	0.047	0.216	0.944
A_1	Informative	-0.035	0.210	0.342	0.609
	Horseshoe	-0.037	0.198	0.307	0.619
	Laplace	-0.034	0.206	0.332	0.628
	Spike-Slab	-0.029	0.223	0.372	0.616
	Hierarchical	-0.040	0.210	0.339	0.602

Table S4: Forecast Performance Summary Across Supplementary Simulations. M-RMSE is the mean (across simulations) of the root mean squared error on the test set, and SD-RMSE is its standard deviation.

Prior	Sim 4: True DGP		Sim 5: Overfitting		Sim 6: Underfitting	
	M-RMSE	SD RMSE	M-RMSE	SD RMSE	M-RMSE	SD RMSE
Informative	0.0378	0.0060	0.0414	0.0079	0.0394	0.0085
Horseshoe	0.0379	0.0063	0.0381	0.0057	0.0394	0.0085
Laplace	0.0381	0.0063	0.0420	0.0073	0.0399	0.0086
Spike-Slab	0.0396	0.0064	0.0412	0.0072	0.0408	0.0085
Hierarchical	0.0376	0.0063	0.0385	0.0062	0.0394	0.0083

Table S5: Forecasting Performance Ratios for Mean RMSE (M-RMSE) and its Standard Deviation (SD-RMSE). "S5" = Overfitting scenario, "S6" = Underfitting, "S4" = Correct DGP. Columns show how each simulation compares in terms of M-RMSE and SD-RMSE. Ratios > 1 indicate worse performance (higher RMSE) relative to the denominator; < 1 indicates better (lower RMSE).

Prior	M-RMSE Ratio				SD-RMSE Ratio			
	5/4	6/4	6/5	5/6	5/4	6/4	6/5	5/6
Informative	1.091	1.030	0.945	1.058	1.385	1.051	0.758	1.319
Horseshoe	1.016	1.042	1.026	0.974	1.278	1.139	0.892	1.121
Laplace	1.060	1.042	0.983	1.017	1.385	1.103	0.796	1.256
Spike-Slab	1.042	1.025	0.984	1.016	1.227	1.091	0.889	1.125
Hierarchical	1.022	1.032	1.010	0.990	1.211	1.079	0.891	1.122

Table S6: Forecasting Performance Ratios Within Simulations (Best Model as Denominator). M-RMSE is the mean root mean squared error and SD-RMSE is the standard deviation of the root mean squared errors on the test set across simulations.

Prior	Sim 4: True DGP		Sim 5: Overfitting		Sim 6: Underfitting	
	M-RMSE	SD RMSE	M-RMSE	SD RMSE	M-RMSE	SD RMSE
Informative	1.01	1.08	1.08	1.17	1.00	1.00
Horseshoe	1.00	1.00	1.00	1.00	1.00	1.00
Laplace	1.01	1.08	1.05	1.17	1.01	1.05
Spike-Slab	1.04	1.22	1.07	1.17	1.03	1.17
Hierarchical	1.01	1.06	1.01	1.00	1.00	1.00