Study 2: Comparative Performance of Normal–Gaussian and Exponential–Gaussian Copula VAR Models Under Exponential Innovations

Table of contents

0. Summary

0.1 Computational Stability versus Statistical Inference

Both models exhibit excellent computational stability across all simulation conditions. The Normal–Gaussian (NG) model shows no post-warmup divergent transitions and max $\hat{R} \leq 1.01$ in all replications. The Exponential–Gaussian (EG) model is equally well-behaved, with zero divergent transitions and good convergence diagnostics. This is an improvement over the SG model under extremeCHI in Study 1, where divergences were frequent.

0.2 Model Performance Under Exponential Innovations

Exponential–Gaussian (EG): EG is (i) approximately unbiased for the VAR dynamics Φ and intercepts μ, and (ii) well calibrated for dependence when ρ is interpreted on the correct scale in mixed-direction cells (see Section 1.1 for the sign convention under mirroring). Empirical coverage is close to 0.95 in most conditions

Normal–Gaussian (NG): NG is computationally stable but exhibits attenuation in $|\rho|$ and corresponding under-coverage for ρ under Exponential margins. This attenuation arises from PIT distortion (see Study 1, Section 6.1).

0.3 Insights

PIT distortion induced by marginal CDF misspecification is the dominant failure mechanism, consistent with findings from Study 1. When NG assumes Gaussian margins but the true innovations are Exponential, the probability integral transform yields non-uniform PITs, which attenuate the effective dependence seen by the Gaussian copula.

1. Introduction

This simulation study extends the analysis from Study 1 by comparing two Bayesian VAR(1) models under exponential innovations: a Normal–Gaussian (NG) model and an Exponential–Gaussian (EG) model. While Study 1 examined skew-normal and chi-squared innovations, this study focuses on exponential margins to assess whether correctly specifying the marginal distribution recovers the models parameters without bias.

1.1. Data Generating Process (DGP)

The DGP follows the same structure as Study 1:

$$Y_t = \mu + \Phi Y_{t-1} + \varepsilon_t, \qquad t = 2, \ldots, T,$$

with $\mu = \mathbf{0}$. Innovations $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})$ have standardized Exponential margins (mean 0, variance 1), with optional mirroring to induce left-skewness, and are coupled through a Gaussian copula with correlation parameter ρ.

---

i Standardization of Exponential Innovations

A standard Exponential(1) random variable X has $\mathbb{E}[X] = 1$ and $\text{Var}(X) = 1$. To standardize:

$$Z = X - 1$$

yields $\mathbb{E}[Z] = 0$ and $\text{Var}(Z) = 1$. Mirroring $(-Z)$ produces left-skewed innovations. For the joint density via Gaussian copula, see Study 1, Section 1.1.

! Copula Sign Under Mirroring

Left-skew Exponential margins are produced by mirroring standardized Exponential innovations (multiplying by -1). For a Gaussian copula, mirroring exactly one margin corresponds to the transformation $u \mapsto 1-u$ on the PIT scale. Since $\Phi^{-1}(1-u) = -\Phi^{-1}(u)$,

---

this flips the sign of the latent Gaussian score and therefore maps

$$\rho \mapsto -\rho.$$

Accordingly, in mixed-direction settings (e.g., $+-$), the effective copula correlation associated with the observed marginals is

$$\rho_{\text{eff}} = s_1 s_2 \rho, \qquad s_j \in \{+1, -1\},$$

where $s_j = -1$ indicates a mirrored (left-skew) margin.

1.2. Simulation Design

The study employs a factorial design crossing four factors.

Table 1: Summary of the Simulation Design Factors.

Factor	Levels
DGP Level	Standardized Exponential
Time Series Length (T)	50, 100, 200
Copula Correlation (ρ)	0.30, 0.50 (reported as input; evaluation uses ρ_{eff} in mixed-direction cells)
VAR Parameters (Φ)	**Set A**: $\begin{pmatrix} 0.40 & 0.10 \\ 0.10 & 0.40 \end{pmatrix}$
	Set B: $\begin{pmatrix} 0.55 & 0.10 \\ 0.10 & 0.25 \end{pmatrix}$
Skewness Direction	++ (both right), -- (both left), +- (mixed)

🔥 Mixed-Direction Symmetry

Only one mixed-direction case $(+-)$ is included. For the asymmetric VAR set B, $+-$ and $-+$ are not equivalent under variable relabeling. See Study 1, Section 1.2 for discussion of this design choice.

1.3. True Parameter Values

Table 2: True Parameter Values Used in the Data Generating Process.
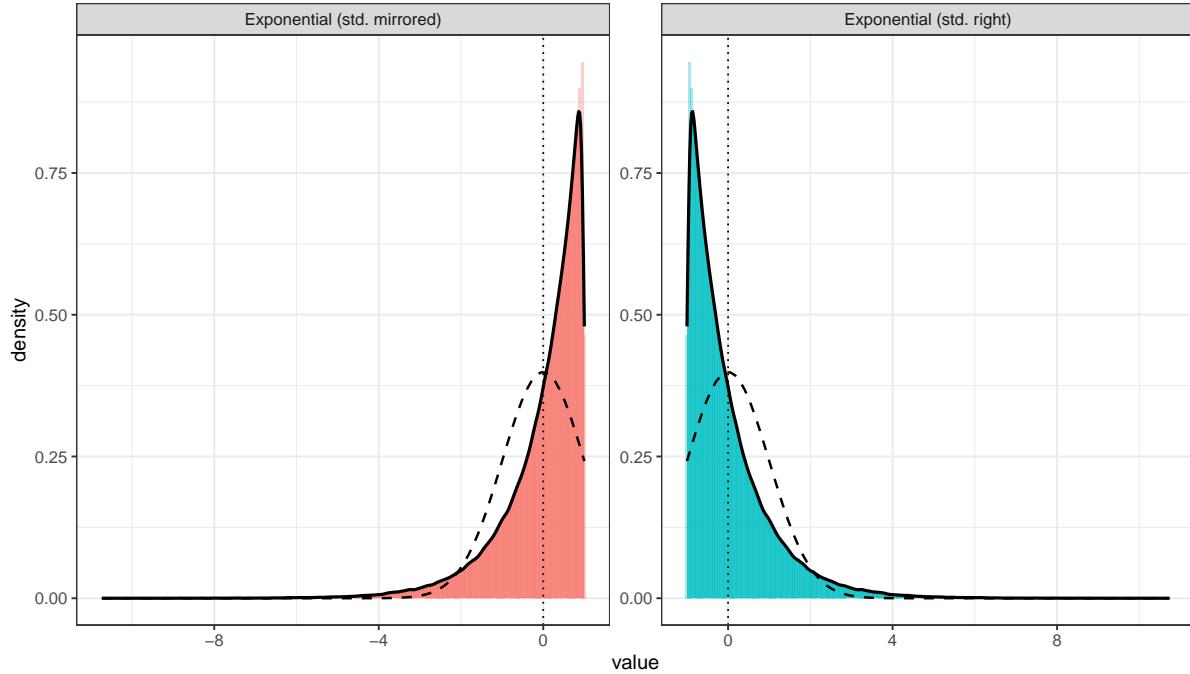
Parameter	True Value	Notes
μ_1, μ_2	0, 0	Innovations are mean-zero
ϕ_{11} (Set A / Set B)	0.40 / 0.55	Diagonal AR coefficients
$\phi_{12} = \phi_{21}$	0.10	Cross-effects (symmetric)
ϕ_{22} (Set A / Set B)	0.40 / 0.25	Diagonal AR coefficients
ρ	0.30 or 0.50	Copula correlation
σ_1, σ_2 (NG model)	1.0, 1.0	Innovations are unit-variance
$\sigma_{\exp,1}, \sigma_{\exp,2}$ (EG model)	1.0, 1.0	Scale parameter for standardized Exponential

> **i** Bias Metric for Intercepts (μ)
>
> Because $\mu = 0$ in the DGP, "relative bias" is undefined. We report absolute bias for μ (i.e., $\hat{\mu} - 0$), using the convention rel_bias = bias when |truth| is near 0.

1.4 Visual Check: Standardized Marginal Innovations (DGP)

Standardized Exponential innovations used in Study 2
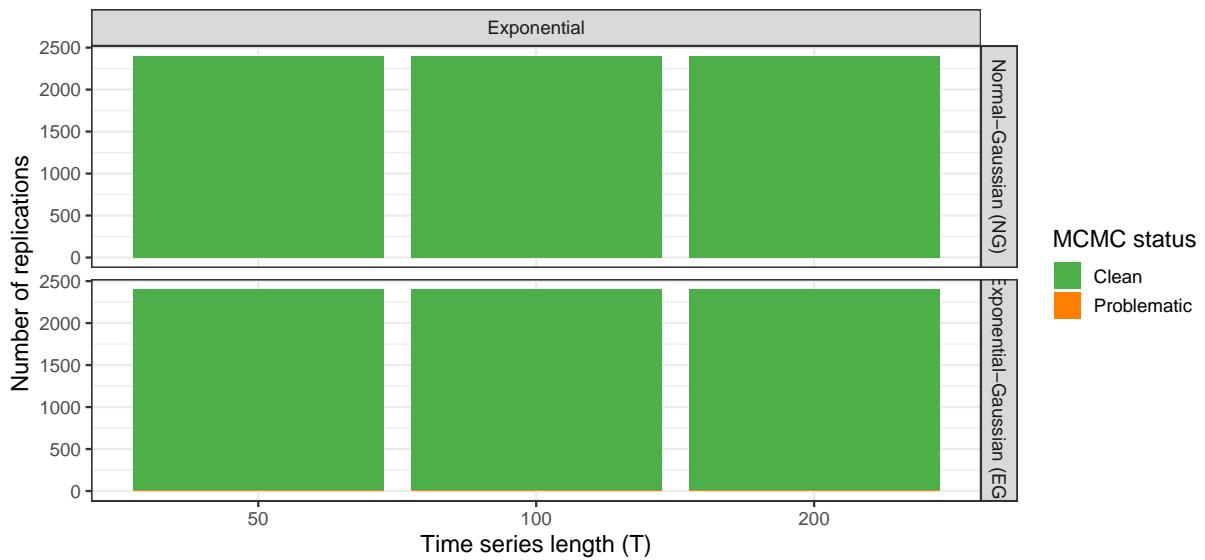

2. Data Loading and Preparation

2.1. MCMC Classification and Overview

We classify runs based on MCMC diagnostics ($\hat{R}$ and divergent transitions) using the same criteria as Study 1:
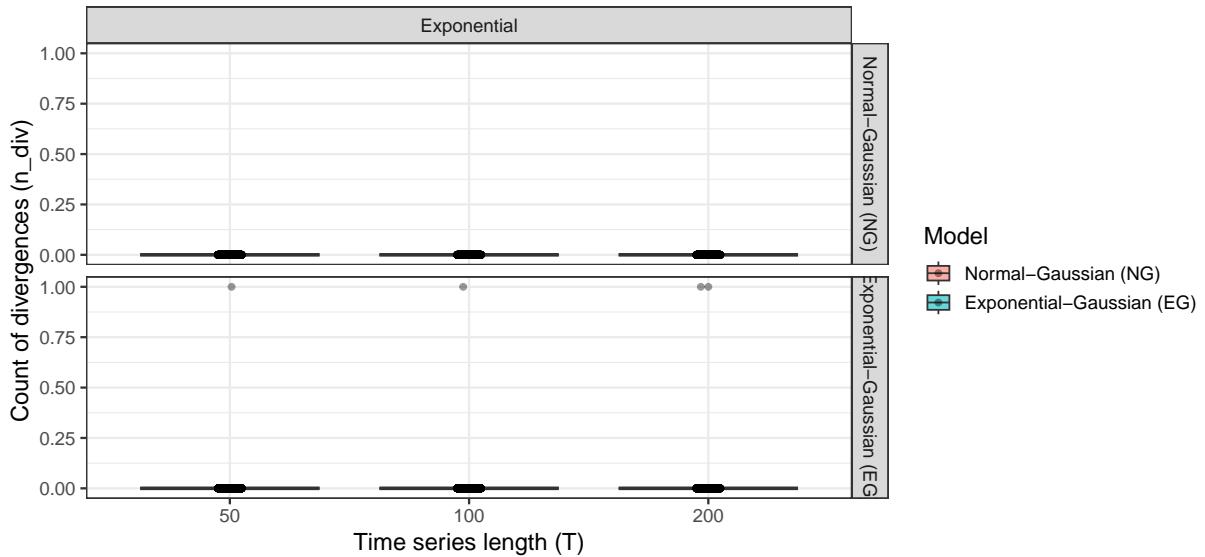
- **Clean:** $\hat{R} \leq 1.01$ and no post-warmup divergences.
- **Problematic:** $\hat{R} > 1.01$ or at least one divergence.
- **Failed/Error:** Non-OK status or missing diagnostics.

MCMC convergence status by model (Study 2)

Interpretation: Both NG and EG fits are uniformly Clean across the entire design, with zero post-warmup divergent transitions. This excellent computational behavior contrasts with the SG model under extremeCHI in Study 1, which exhibited frequent divergences. The absence of sampling difficulties here demonstrates that correctly specifying the marginal distribution (EG with Exponential margins) avoids the posterior geometry challenges that arose from model misspecification.
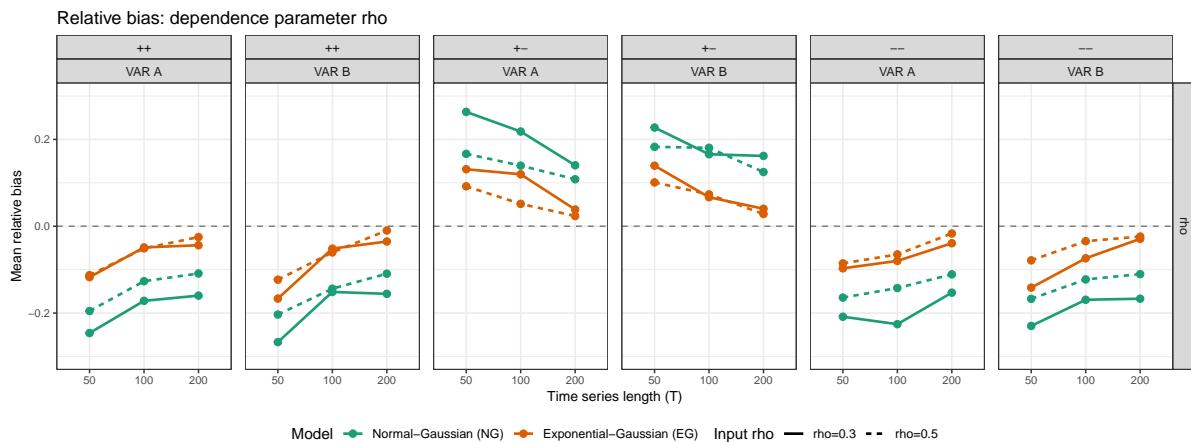
Distribution of post−warmup divergent transitions
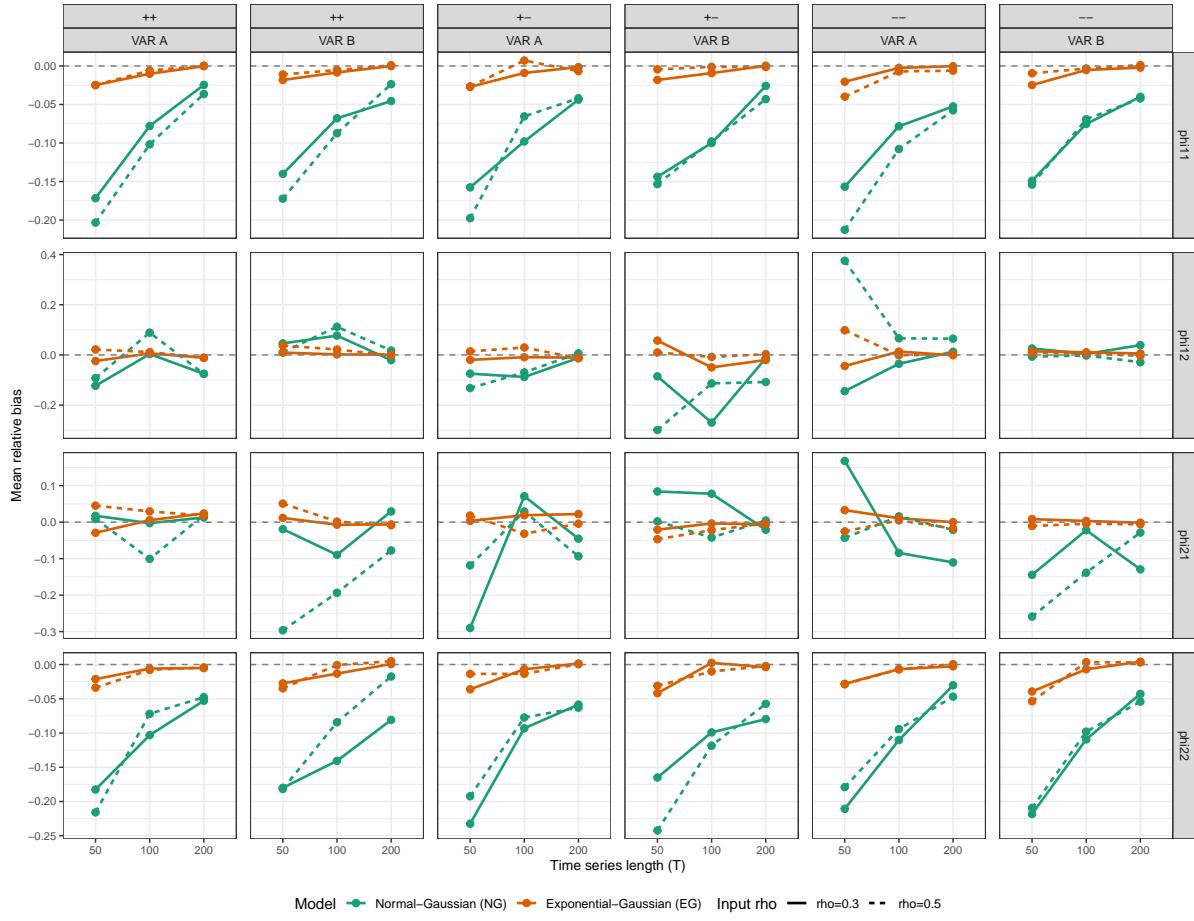
3. Helpers

4. Exponential DGP: NG vs EG Performance
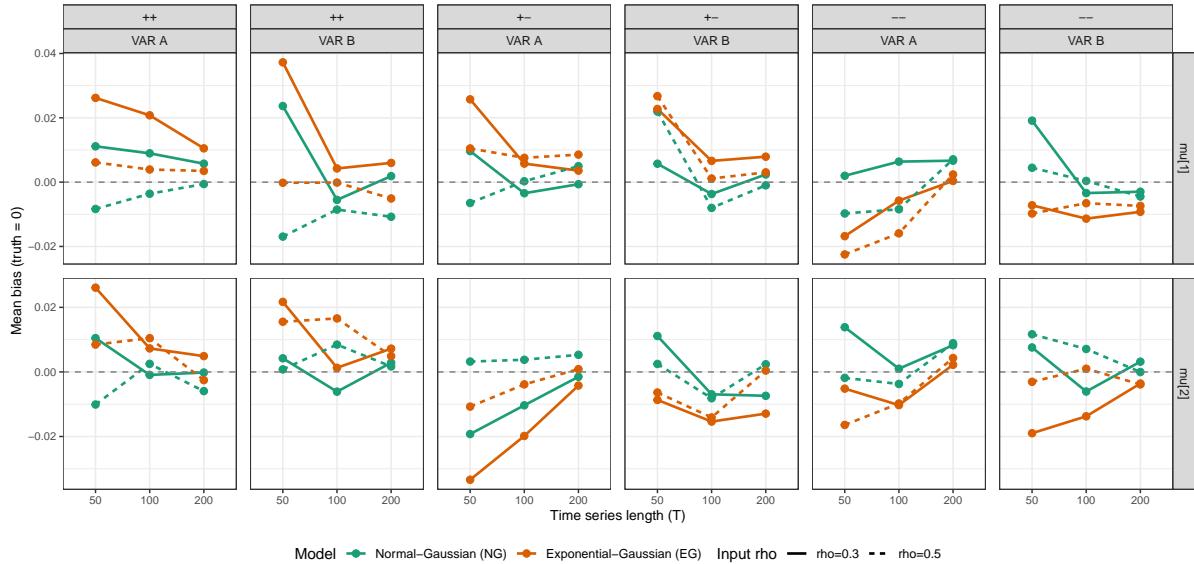
4.1. Relative Bias

Relative bias: dependence parameter rho

Relative bias: VAR coefficients (Phi elements)
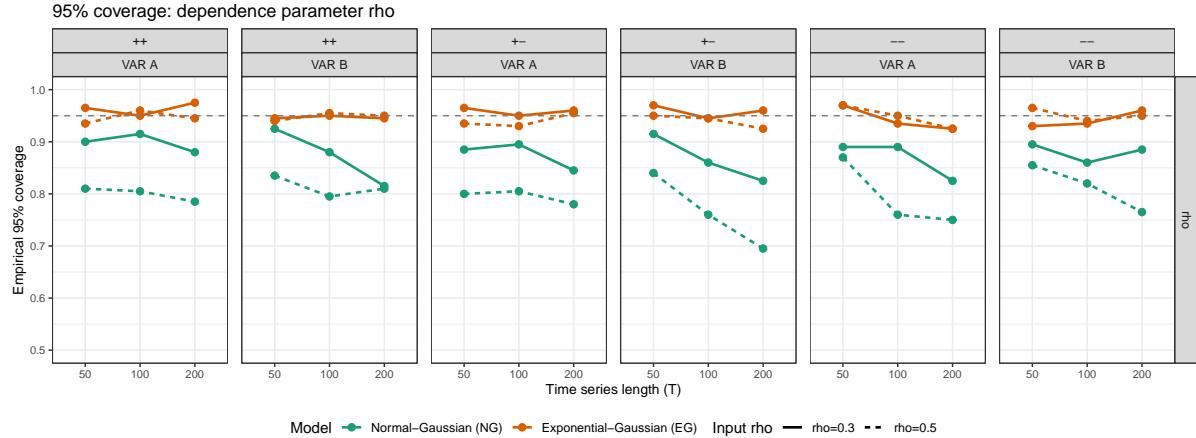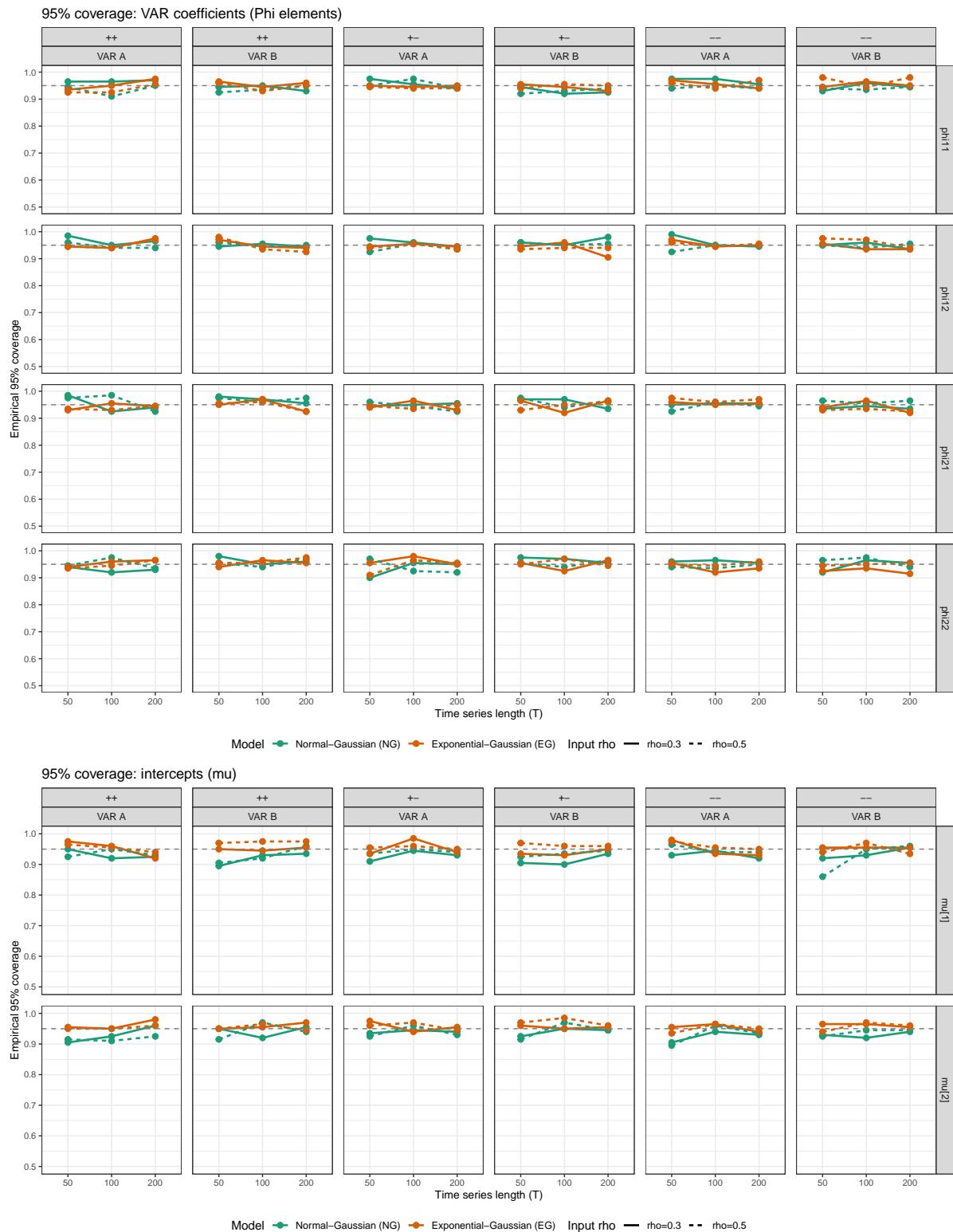
Bias: intercepts (mu)

Interpretation. Across T, EG is approximately unbiased for Φ and μ. For ρ, the target in mixed-direction cells is ρ_{eff} (callout above): EG is close to unbiased, while NG exhibits attenuation in $|\rho|$, most pronounced at small T.

4.2. 95% Coverage

95% coverage: dependence parameter rho

95% coverage: VAR coefficients (Phi elements)
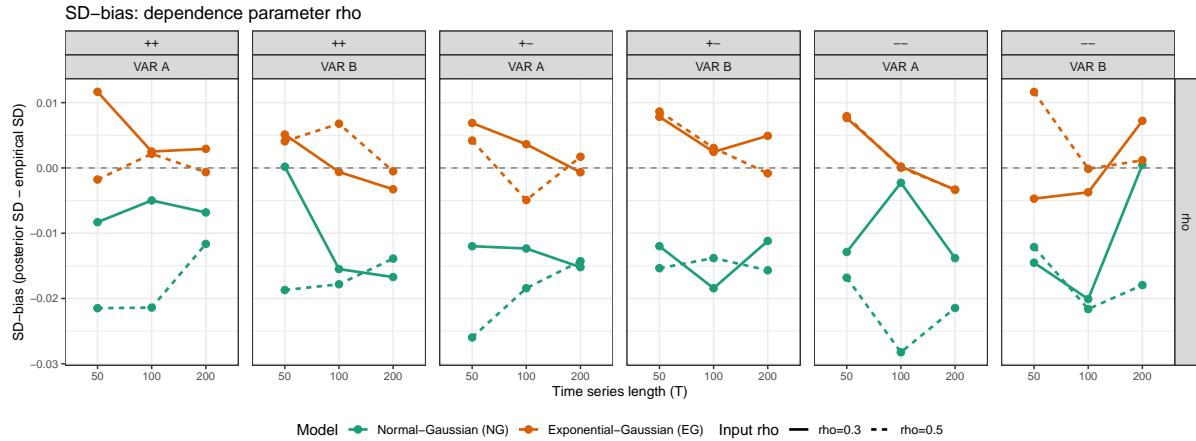
95% coverage: intercepts (mu)
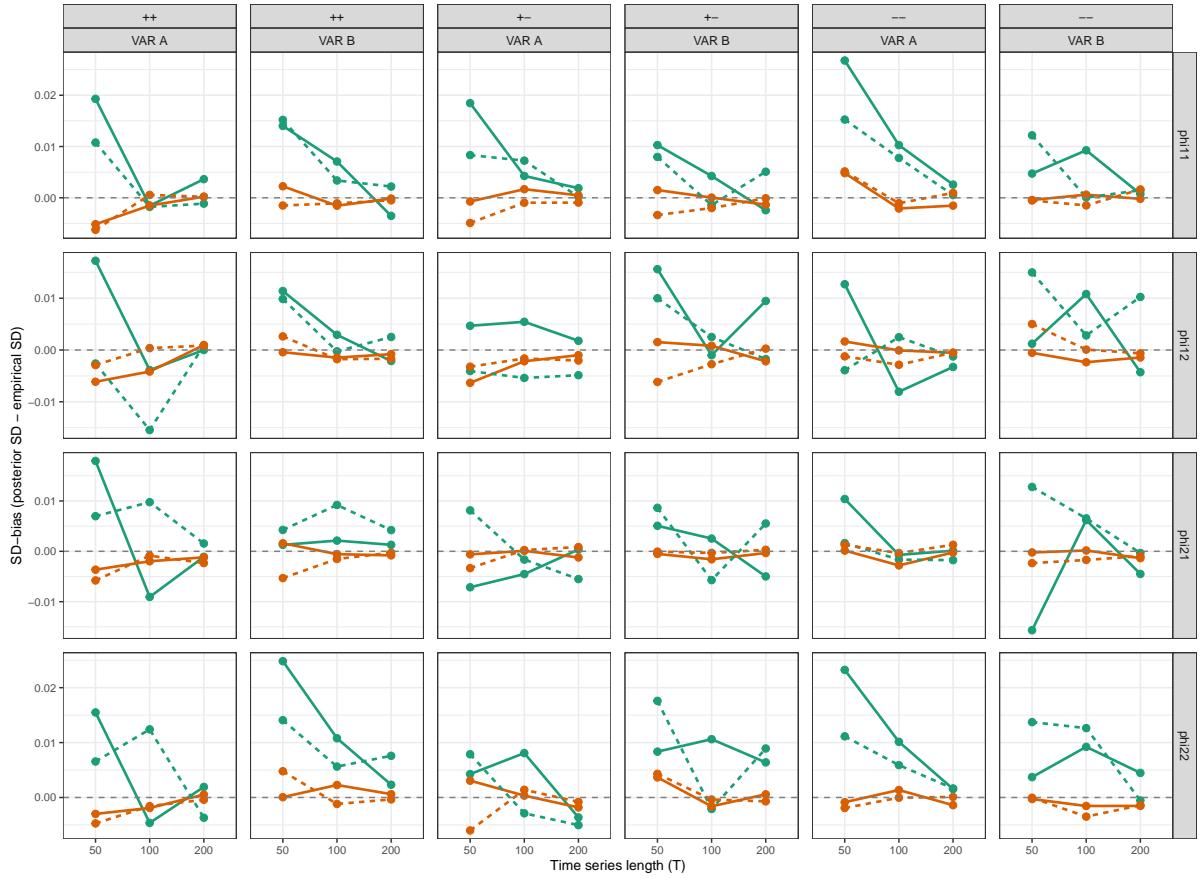
Interpretation. EG coverage is typically close to nominal for the core parameters. NG coverage shortfalls are concentrated in ρ, consistent with attenuation bias under marginal misspecification.

4.3. SD-Bias

SD–bias: dependence parameter rho
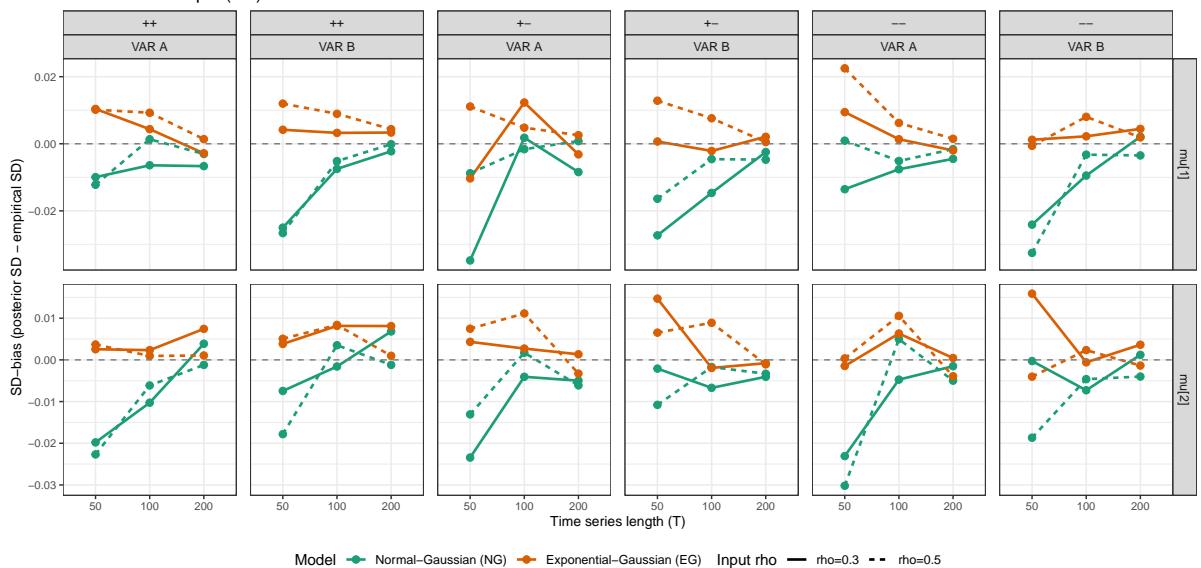
SD–bias: VAR coefficients (Phi elements)
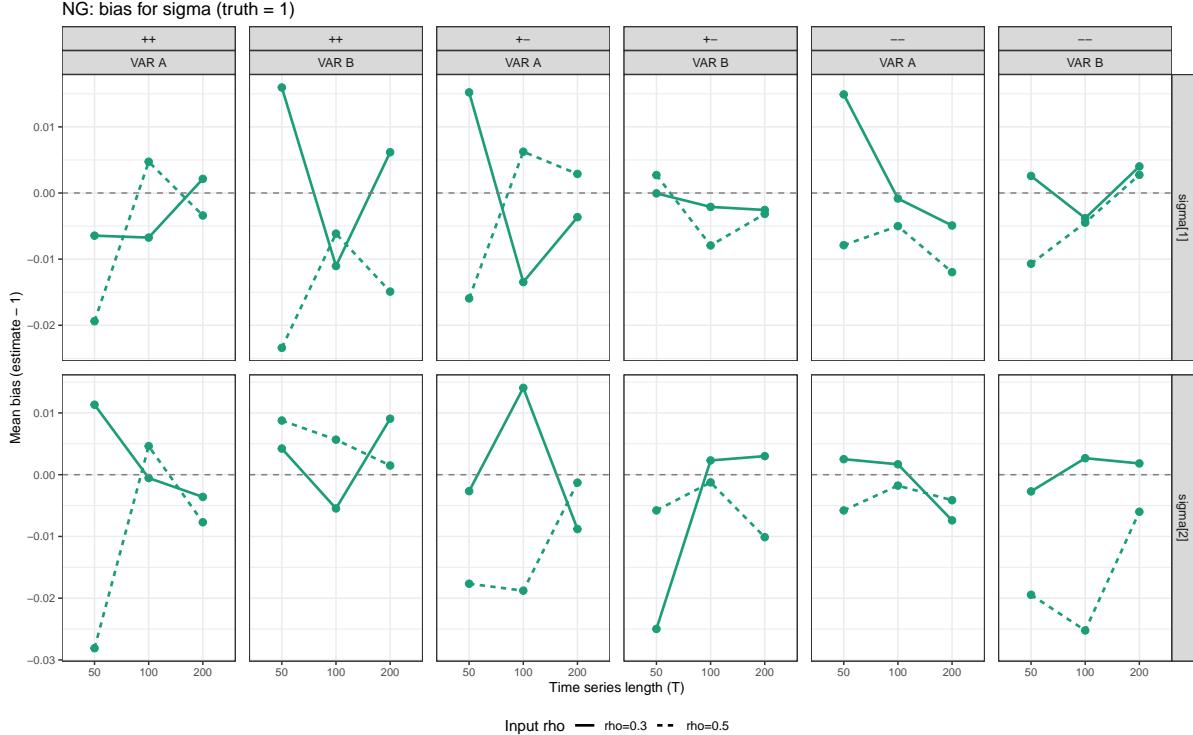
SD–bias: intercepts (mu)

Interpretation. EG is generally well calibrated (SD-bias near 0). NG frequently shows negative SD-bias for ρ, indicating overconfident posterior uncertainty for dependence under misspecified margins.

5. Marginal Parameters

We report mean bias for the Exponential scale parameters in EG ($\sigma_{\exp}$) and the innovation standard deviations in NG (σ). Under the standardized DGP, the reference value is 1.

NG: bias for sigma (truth = 1)

Interpretation. The σ estimates under NG show small but systematic deviations from the true value of 1. As discussed in Study 1 (Section 7.1), these deviations arise from marginal misspecification but are not the mechanism driving ρ attenuation—both phenomena stem from the same root cause (marginal misspecification) but through different pathways.

6. Details

This section provides technical details on the implementation of the simulation study, including prior specifications, MCMC settings, and mathematical derivations specific to the Exponential–Gaussian (EG) model.

6.1. Prior Specifications

Both models use weakly informative priors. The EG model employs a novel reparameterization for the scale parameter to handle the support constraint of exponential marginals.

Table: Prior Specifications for the Normal–Gaussian (NG) Model

Parameter	Prior	Support	Rationale
μ_1, μ_2	Normal$(0, 1)$	$\mathbb{R}$	Weakly informative; centered at true value (0)
$\phi_{11}, \phi_{12}, \phi_{21}, \phi_{22}$	Normal$(0, 0.5)$	$(-1, 1)$	Regularizes toward stationarity; truncated by bounds
σ_1, σ_2	Half-Normal$(0, 1)$	$(0, \infty)$	Weakly informative scale prior
ρ	Normal$(0, 0.5)$	$(-1, 1)$	Regularizes toward independence; truncated by bounds

Table: Prior Specifications for the Exponential–Gaussian (EG) Model

Parameter	Prior	Support	Rationale
μ_1, μ_2	Normal$(0, 1)$	$\mathbb{R}$	Weakly informative; centered at true value (0)
$\phi_{11}, \phi_{12}, \phi_{21}, \phi_{22}$	Normal$(0, 0.5)$	$(-1, 1)$	Regularizes toward stationarity; truncated by bounds
$\sigma_{\exp,1}, \sigma_{\exp,2}$	Log-Normal$(0, 0.5)$	(b_i, ∞)	Induced prior via reparameterization (see Section 6.4)
ρ	Normal$(0, 0.5)$	$(-1, 1)$	Regularizes toward independence; truncated by bounds

> **ℹ EG Scale Parameter Prior**
>
> The EG model does **not** place a direct prior on $\sigma_{\exp}$. Instead, it estimates an unconstrained parameter $\eta \in \mathbb{R}$ and applies a change-of-variables transformation:

$$\sigma_{\text{exp}} = b + \exp(\eta)$$

where b is a data-dependent feasibility bound (Section 6.4). The induced prior on σ_{exp} is Log-Normal$(0, 0.5)$ with support shifted to (b, ∞).

6.2. MCMC Settings

All models were fitted using the No-U-Turn Sampler (NUTS) implemented in Stan via the `rstan` package.

Table: MCMC Sampling Configuration

Setting	Value	Description
Chains	4	Number of independent Markov chains
Total iterations	4,000	Iterations per chain (including warmup)
Warmup iterations	2,000	Discarded adaptation period
Post-warmup draws	2,000	Retained samples per chain
`adapt_delta`	0.995	Target acceptance probability
`max_treedepth`	15	Maximum tree depth for NUTS
Parallelization	Outer loop	Replications parallelized; chains run sequentially

> **i** Higher adapt_delta Than Study 1
>
> Study 2 uses `adapt_delta = 0.995` compared to 0.95 in Study 1. This higher value was chosen because the EG model's feasibility-bound reparameterization creates regions where the posterior density changes rapidly near the boundary. The elevated target acceptance rate reduces step sizes in these regions, preventing divergent transitions. Despite this challenging geometry, **zero divergent transitions** were observed across all 14,400 fits (7,200 per model), demonstrating that the reparameterization strategy (Section 6.4) successfully regularizes the posterior.

6.3. Gaussian Copula Log-Density

The Gaussian copula density implementation is identical to Study 1. For uniform marginals $(u, v) \in (0, 1)^2$ with correlation parameter $\rho \in (-1, 1)$:

Let $z_1 = \Phi^{-1}(u)$ and $z_2 = \Phi^{-1}(v)$ denote the standard normal quantile transforms. The copula log-density is:

$$\log c(u, v; \rho) = -\frac{1}{2} \log(1 - \rho^2) - \frac{1}{2(1 - \rho^2)} \left(z_1^2 - 2\rho z_1 z_2 + z_2^2 \right) + \frac{1}{2} \left(z_1^2 + z_2^2 \right)$$

> **i Boundary Clamping**
>
> As in Study 1, the implementation applies boundary clamping with $\varepsilon = 10^{-9}$:
>
> $$u_{\text{clamped}} = \max(\varepsilon, \min(1 - \varepsilon, u))$$
>
> This prevents numerical overflow when $\Phi^{-1}(u)$ is evaluated near 0 or 1.

6.4. EG Model: Feasibility-Bound Reparameterization

The Exponential–Gaussian model requires that the shifted residuals have positive support. This section documents the reparameterization strategy that ensures feasibility while maintaining good posterior geometry.

The Support Constraint Problem

For an Exponential marginal with scale σ_{exp}, the likelihood requires:

$$x_{\text{shifted},t} = \sigma_{\text{exp}} + s \cdot \text{res}_t > 0, \quad \forall t$$

where $s \in \{+1, -1\}$ is the skew direction ($+1$ for right-skewed, -1 for left-skewed/mirrored). Naively parameterizing $\sigma_{\text{exp}} > 0$ can lead to divergent transitions when the sampler proposes values that violate this constraint for some t.

Data-Dependent Feasibility Bound

Define the feasibility bound:

$$b_i = \max_t \left(-s_i \cdot \text{res}_{t,i} \right)$$

This is the minimum value of $\sigma_{\text{exp},i}$ such that $x_{\text{shifted},t,i} > 0$ for all observations. Any $\sigma_{\text{exp},i} > b_i$ is guaranteed to satisfy the support constraint.

The η Reparameterization

Rather than estimating σ_{exp} directly, the EG model estimates an unconstrained parameter $\eta \in \mathbb{R}$ and transforms:

$$\sigma_{\text{exp}} = b + \exp(\eta)$$

where $\exp(\eta) > 0$ represents the "slack" above the feasibility bound.

Key benefits:

1. **Automatic feasibility**: $\sigma_{\text{exp}} > b$ by construction for any $\eta \in \mathbb{R}$
2. **Smooth geometry**: No hard boundaries for the sampler to navigate
3. **Well-defined gradients**: The transformation is smooth and invertible

Induced Prior via Change-of-Variables

To place a Log-Normal$(0, 0.5)$ prior on σ_{exp}, we apply the change-of-variables formula. The Jacobian of the transformation $\sigma_{\text{exp}} = b + \exp(\eta)$ is:

$$\frac{d\sigma_{\text{exp}}}{d\eta} = \exp(\eta)$$

The log-posterior contribution is:

$$\log p(\sigma_{\text{exp}}) + \log \left| \frac{d\sigma_{\text{exp}}}{d\eta} \right| = \text{lognormal\_lpdf}(\sigma_{\text{exp}} \mid 0, 0.5) + \eta$$

> **!** No Additional Prior on η
>
> The Stan implementation explicitly avoids placing an independent prior on η. Doing so would double-regularize the scale parameter and conflict with the intended Log-Normal$(0, 0.5)$ prior on σ_{exp}, potentially worsening posterior geometry and introducing bias.

Stan Implementation

```
model {
  // Compute feasibility bounds: b_i = max_t(-s_i * res_{t,i})
  for (i in 1:2) {
    real m = -skew_direction[i] * residuals[1, i];
    for (t in 2:(T-1)) {
      m = fmax(m, -skew_direction[i] * residuals[t, i]);
    }
    b[i] = m;
  }

  // Reparameterized sigma: always feasible
  for (i in 1:2) {
    sigma_exp[i] = b[i] + exp(eta[i]);
  }
  rate_exp = 1.0 ./ sigma_exp;

  // Induced prior on sigma_exp via change-of-variables
  for (i in 1:2) {
    target += lognormal_lpdf(sigma_exp[i] | 0, 0.5) + eta[i];
  }

  // Likelihood contributions...
}
```

6.5. CDF Mirroring for Left-Skewed Exponential Margins

For left-skewed (mirrored) Exponential margins, the model must correctly transform the CDF to preserve the copula correlation structure.

The Mirroring Problem

Let $X \sim \text{Exponential}(\lambda)$ with CDF $F_X(x) = 1 - e^{-\lambda x}$. For a mirrored variable $Y = -X$:

$$F_Y(y) = P(Y \leq y) = P(-X \leq y) = P(X \geq -y) = 1 - F_X(-y)$$

On the uniform (PIT) scale, if $U = F_X(X)$, then the PIT of $Y = -X$ is $V = 1 - U$.

Copula Correlation Sign Flip

For a Gaussian copula, the latent normal scores are $Z = \Phi^{-1}(U)$. Under the $U \mapsto 1 - U$ transformation:

$$\Phi^{-1}(1 - U) = -\Phi^{-1}(U) = -Z$$

This negates the latent score, flipping the sign of the copula correlation. If the DGP has $\rho > 0$ and exactly one margin is mirrored, the effective correlation on the observed scale is $\rho_{\text{eff}} = -\rho$.

Stan Implementation

The EG model applies the CDF flip after computing the exponential CDF:

```
for (i in 1:2) {
  // Shift residual to positive support
  x_shifted[i] = sigma_exp[i] + skew_direction[i] * res[i];
  target += exponential_lpdf(x_shifted[i] | rate_exp[i]);

  // Compute PIT
  u_vec[i] = exponential_cdf(x_shifted[i], rate_exp[i]);

  // Flip CDF for left-skewed (mirrored) margins
  if (skew_direction[i] < 0) u_vec[i] = 1.0 - u_vec[i];
}

target += gaussian_copula_ld(u_vec[1], u_vec[2], rho);
```

> **i** Effective Correlation in Evaluation
>
> The analysis script computes $\rho_{\text{eff}} = s_1 s_2 \rho$ where $s_j = -1$ if margin j is mirrored. Bias and coverage metrics for ρ are evaluated against ρ_{eff}, not the input ρ, ensuring correct interpretation in mixed-direction cells.

6.6. Reproducibility Strategy

The simulation study employs deterministic seeding to ensure full reproducibility.

Data Generation Seeds

Each simulated dataset receives a unique seed:

$$\text{seed}_{\text{data}} = \text{seed\_base} + \text{condition\_id} \times 100000 + \text{rep\_id}$$

with seed_base = 2026 for Study 2. The larger multiplier (100,000 vs 10,000 in Study 1) provides wider separation between condition seed ranges.

MCMC Initialization Seeds

The EG model uses data-driven initialization based on OLS pre-fits:

1. Fit OLS: $y_{1,t} \sim y_{1,t-1} + y_{2,t-1}$ and $y_{2,t} \sim y_{1,t-1} + y_{2,t-1}$
2. Extract coefficients as initial values for μ and Φ
3. Compute feasibility bounds b from OLS residuals
4. Set $\eta = \log(0.2)$ (slack 0.2) as default

For NG, standard random initialization is used with seed:

$$\text{seed}_{\text{init}} = \text{seed\_base}_R + \text{seed\_key}(\text{cid}, \text{rid}, \text{chain})$$

where $\text{seed\_key}(\text{cid}, \text{rid}, \text{chain}) = \text{cid} \times 100000 + \text{rid} \times 10 + \text{chain}$.

Stan Sampler Seeds

$$\text{seed}_{\text{Stan}} = \text{seed\_base}_{\text{Stan}} + \text{seed\_key}(\text{cid}, \text{rid}, 0)$$

with $\text{seed\_base}_{\text{Stan}} = 5 \times 10^6$.

> **ℹ Resume Capability**
>
> The pipeline supports resumption via environment variables `START_COND` and `START_REP`. Because seeds are deterministic functions of (condition, replication), resuming from any point produces identical results to running the full pipeline.

Initialization Sidecar Files

For each fitted model, the pipeline optionally saves a "sidecar" file containing:

- All random seeds used (data, init, Stan)
- Exact initial values for each chain
- Timestamp and metadata

This enables exact reproduction of any individual fit.