

# Appendix to “Models for spatial structure in small-area estimation”

Corresponding author: Adam Howes ([ath19@ic.ac.uk](mailto:ath19@ic.ac.uk))

## Contents

<b>1</b>	<b>Implementation details for the Besag model</b>	<b>2</b>
<b>2</b>	<b>Comparison of AGHQ to NUTS</b>	<b>3</b>
<b>3</b>	<b>Simulation study</b>	<b>4</b>
<b>4</b>	<b>HIV study</b>	<b>19</b>
	<b>References</b>	<b>20</b>

# 1 Implementation details for the Besag model

Here we briefly review three best practices for using the Besag model, scaling, singletons, and constraints, as recommended by Freni-Stortino, Ventrucci, and Rue (2018):

1. *Scaling* The structure matrix  $R$  should be rescaled to have generalised variance, defined by the geometric mean of the diagonal elements of its generalised inverse

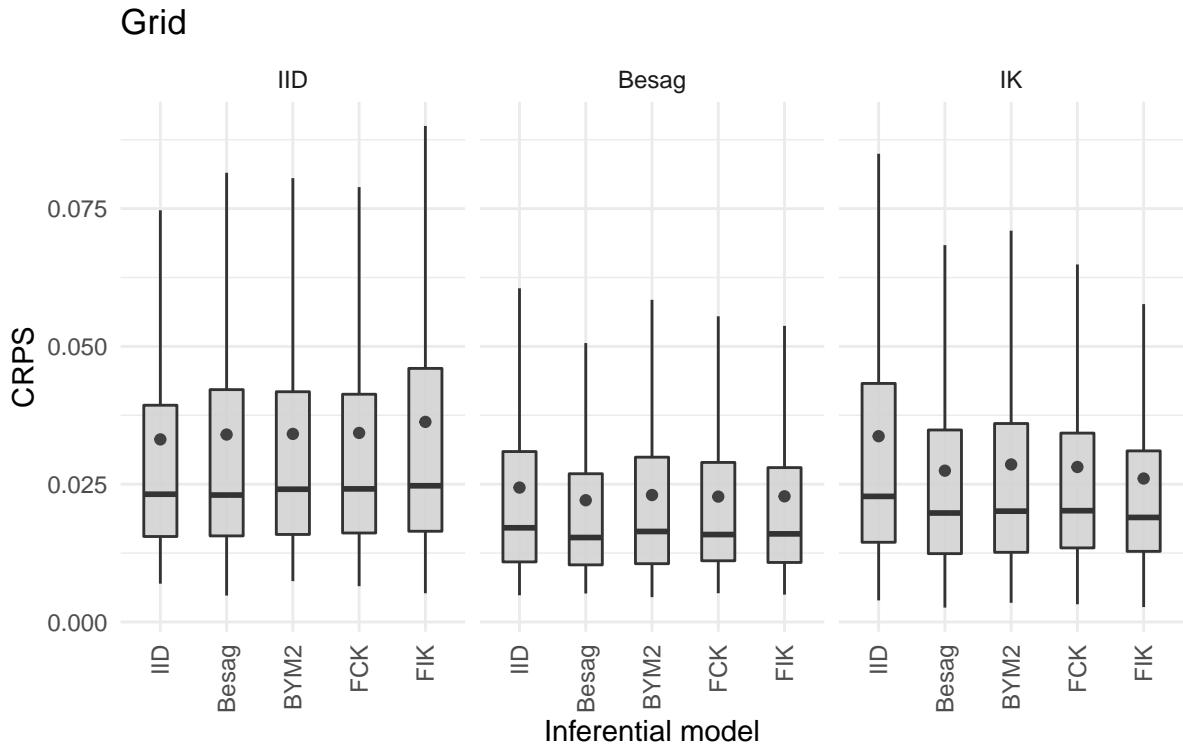
$$\sigma_{\text{GV}}^2(R) = \prod_{i=1}^n (R_{ii}^-)^{1/n} = \exp \left( \frac{1}{n} \sum_{i=1}^n \log(R_{ii}^-) \right), \quad (1)$$

equal to one, by replacing  $R$  with  $R^* = R/\sigma_{\text{GV}}^2(R)$ . As the diagonal elements  $R_{ii}^-$  correspond to marginal variances, the generalised variance gives a measure of the average marginal variance. However, this measure, introduced by Sørbye and Rue (2014), ignores off-diagonal entries and more broadly any measure of typical variance could be used. Scaling mitigates the influence of the adjacency graph on the variance of  $\phi$ . Allowing the variance to be controlled by  $\tau_\phi$  alone is important as it allows for consistent, interpretable prior selection. When the adjacency graph is disconnected it is not appropriate to scale the structure matrix  $R$  uniformly since for a given precision  $\tau_\phi$ , local smoothing operates on each connected component independently. As such, each connected component should be scaled independently to have generalised variance one giving  $R_I^* = R_I/\sigma_{\text{GV}}^2(R_I)$  where  $R_I$  is the sub-matrix of the structure matrix corresponding to index set  $I$ .

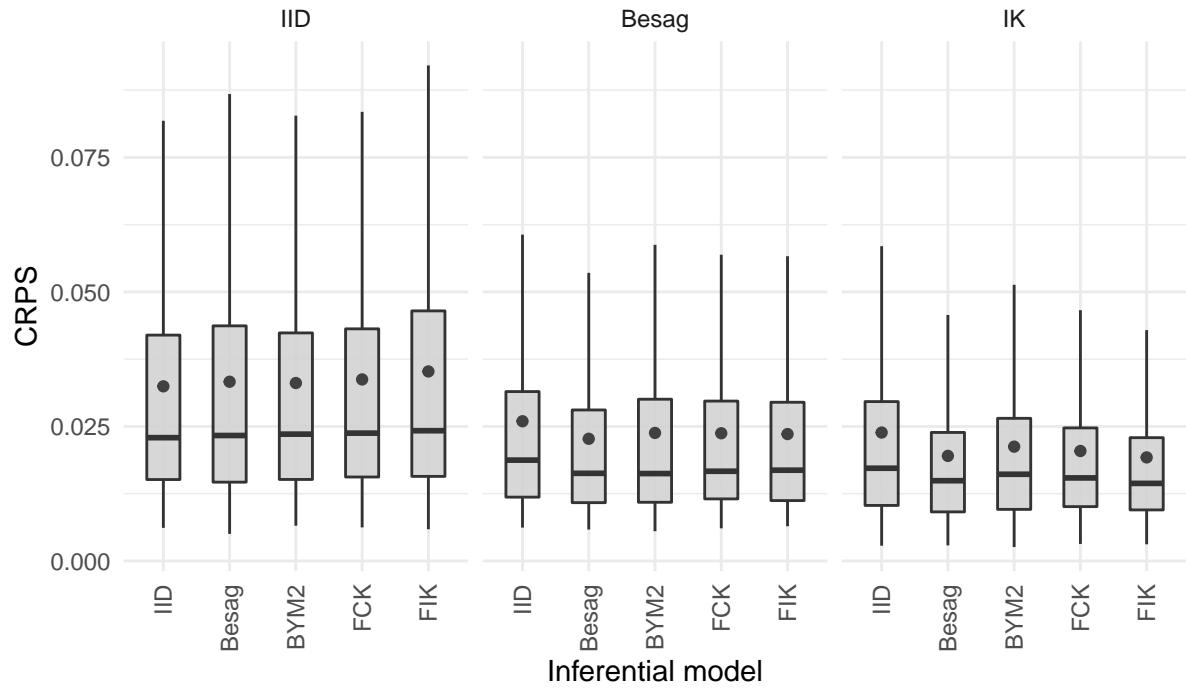
2. *Singletons* When one of the connected components is a single area, known either as a singleton or an island, the probability density  $\exp \left( -\frac{\tau_\phi}{2} \sum_{i \sim j} (\phi_i - \phi_j)^2 \right)$  has no dependence on  $\phi_i$ . This is equivalent to using an improper prior  $p(\phi_i) \propto 1$  and can be avoided by setting each singleton to have independent Gaussian noise  $p(\phi_i) \sim \mathcal{N}(0, 1)$ .
3. *Constraints* To avoid confounding of the spatial random effects with the intercept, it is recommended to place a sum-to-zero constraint on each non-singleton connected component. In other words, for each  $|I| > 1$  that  $\sum_{i \in I} \phi_i = 0$ .

## 2 Comparison of AGHQ to NUTS

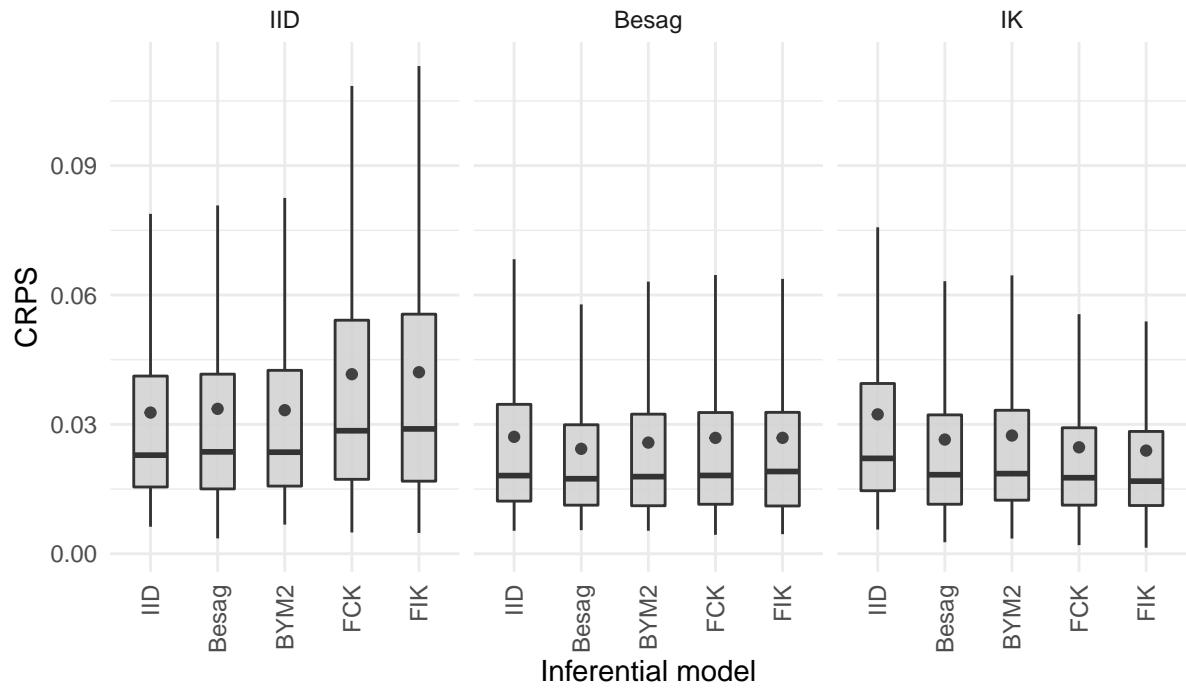
### 3 Simulation study



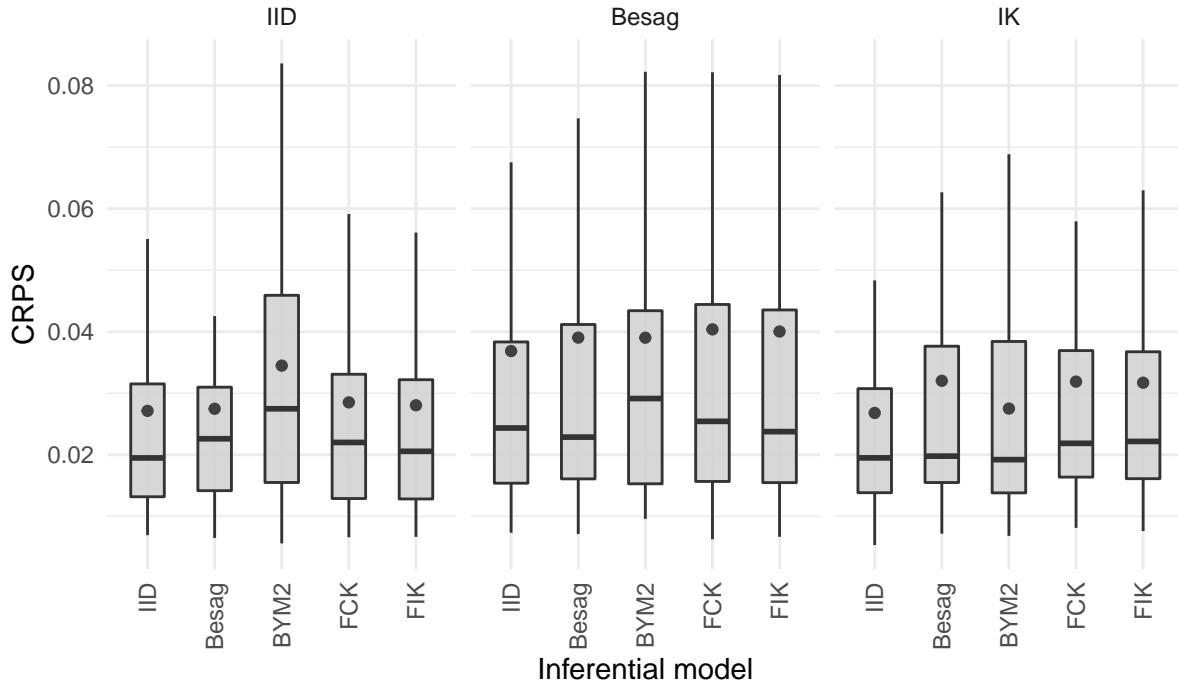
## Cote d'Ivoire



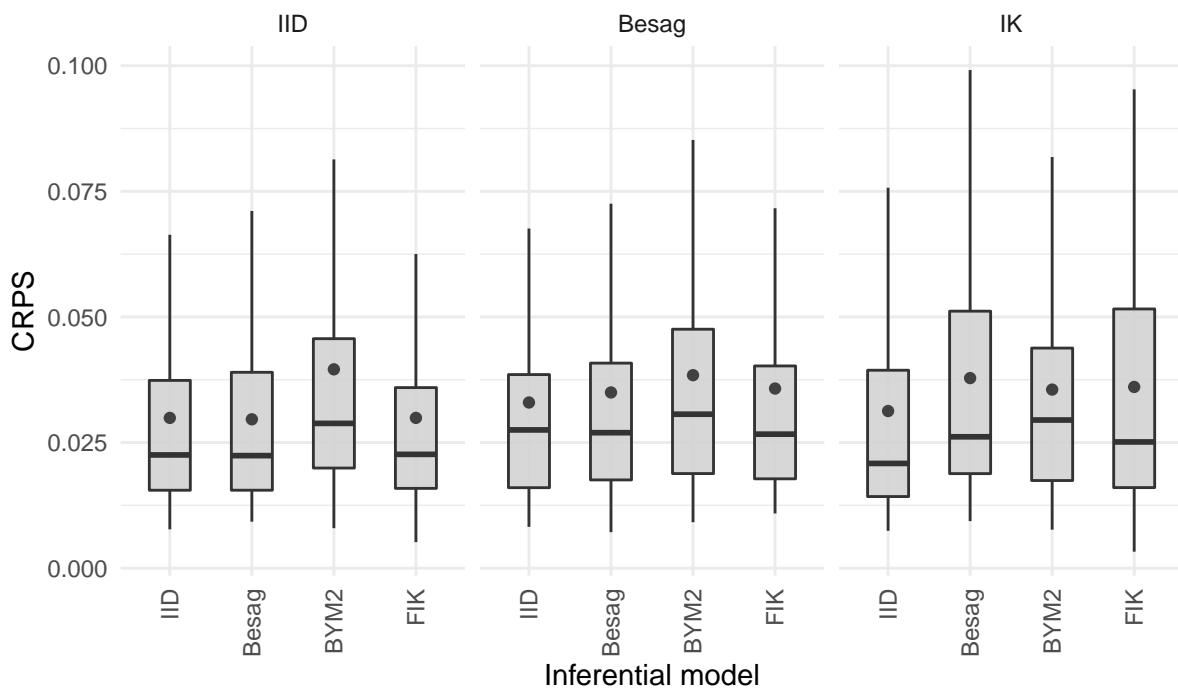
## Texas



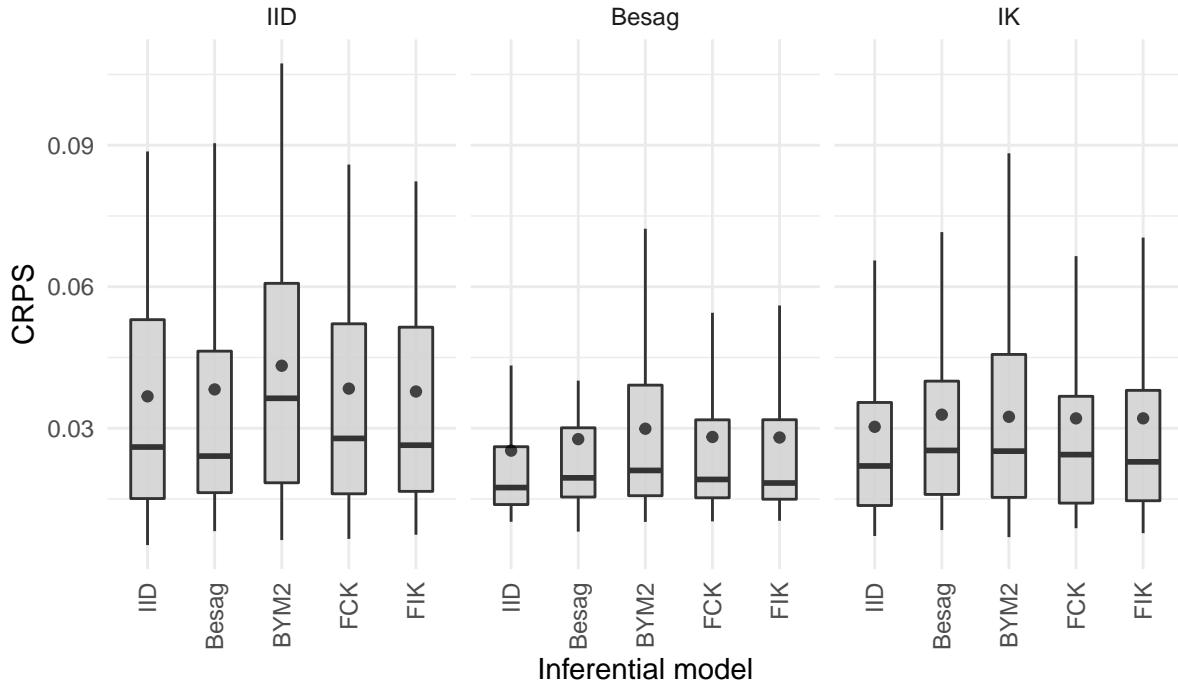
1



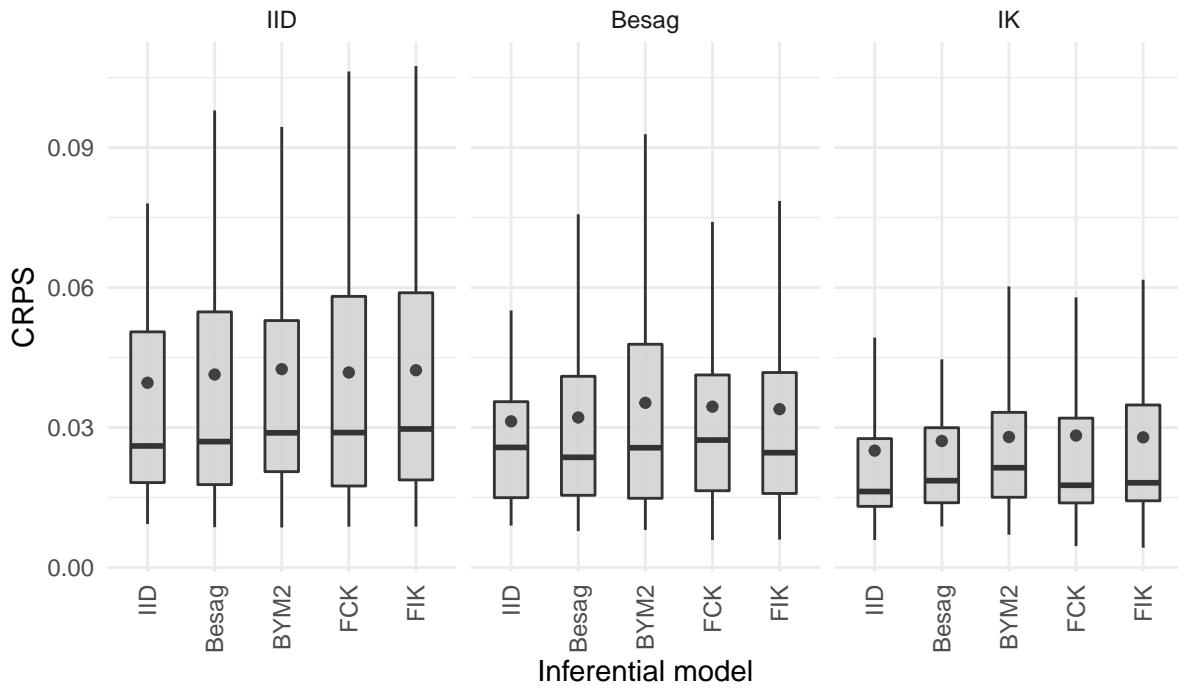
2



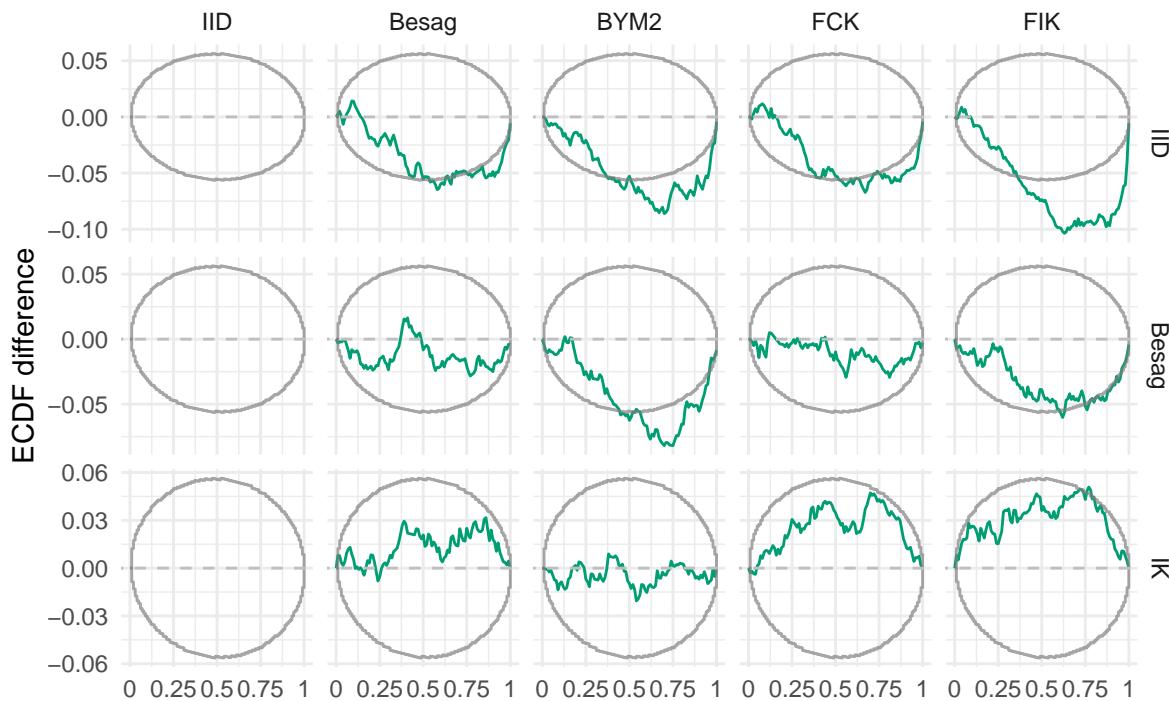
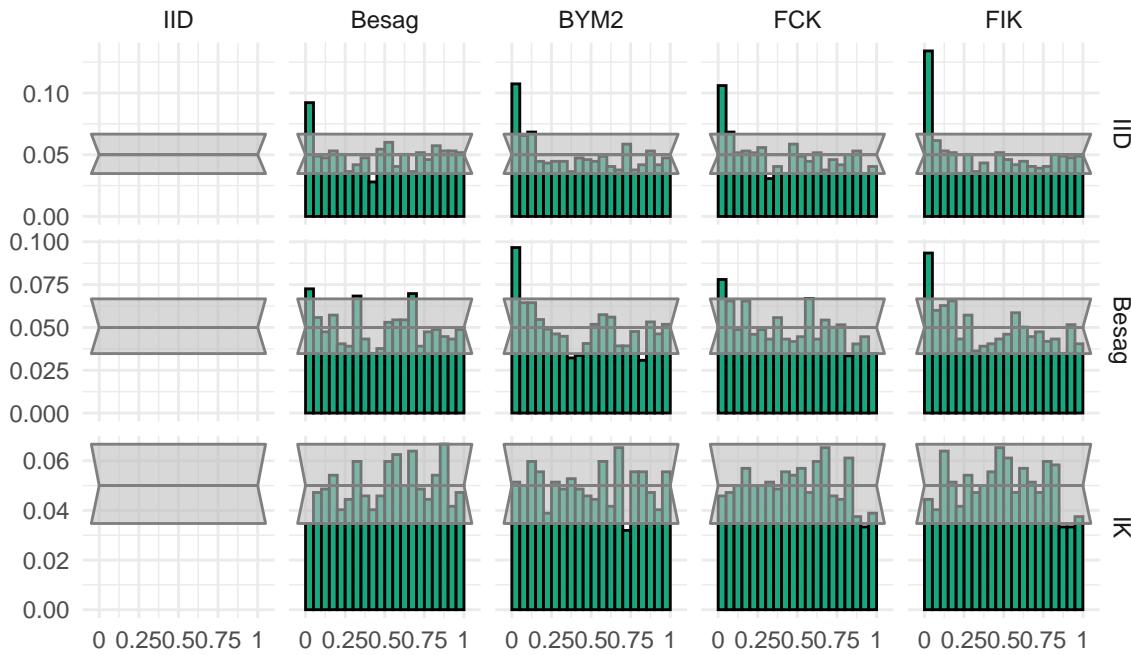
3



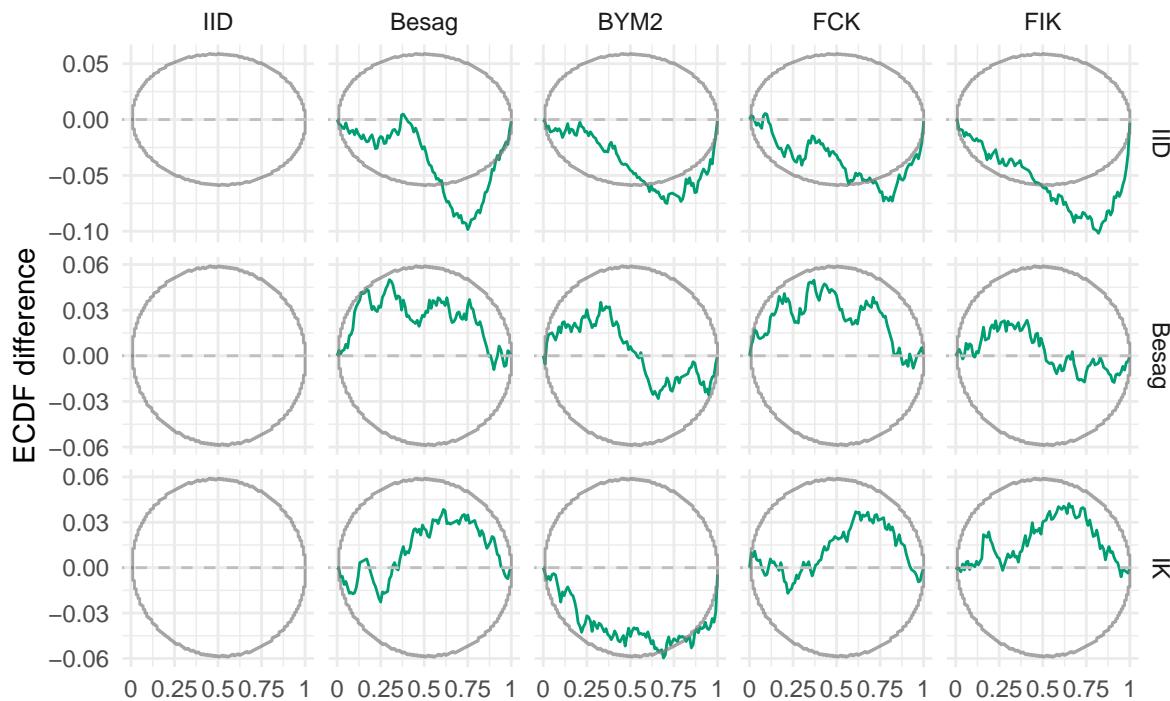
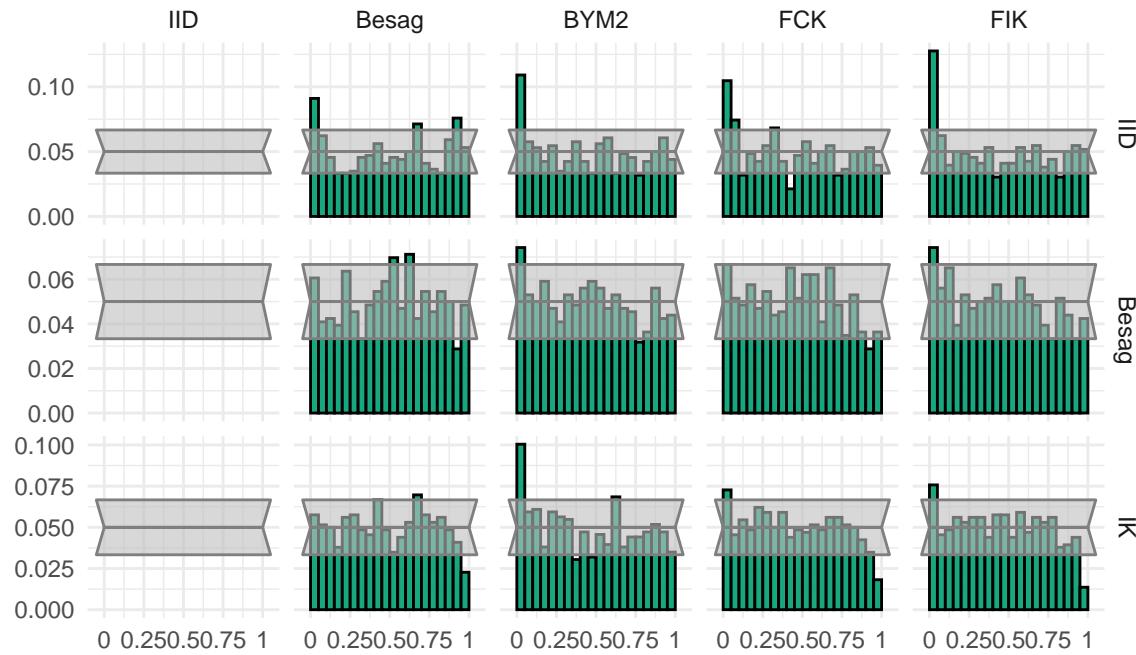
4



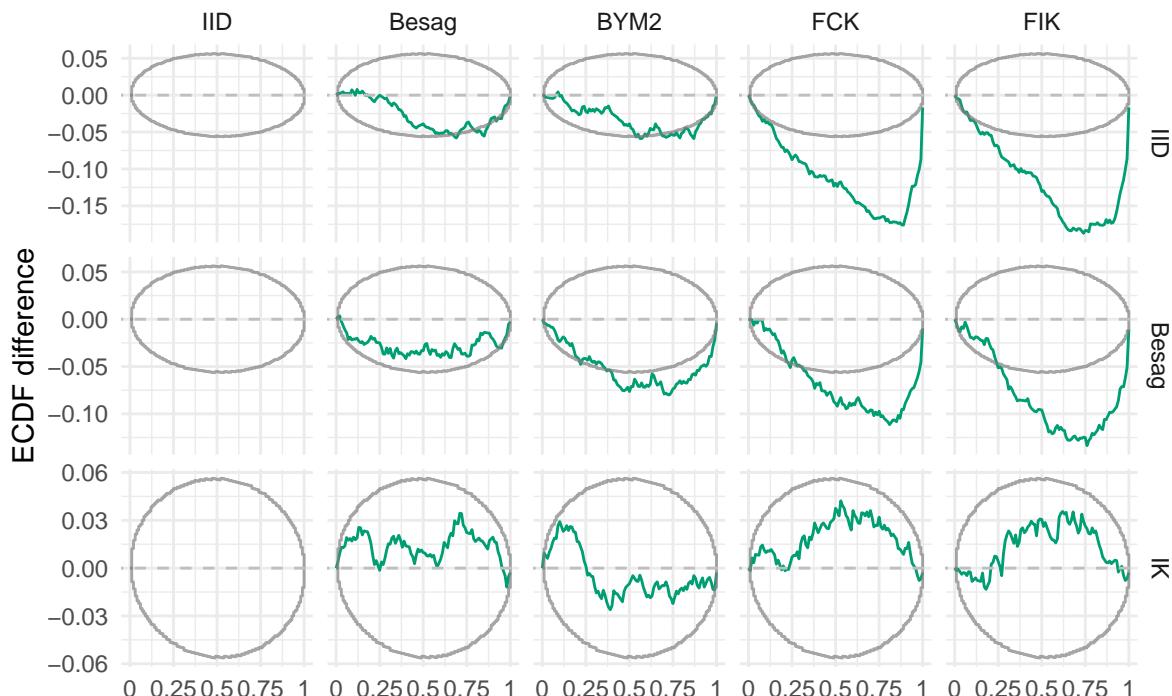
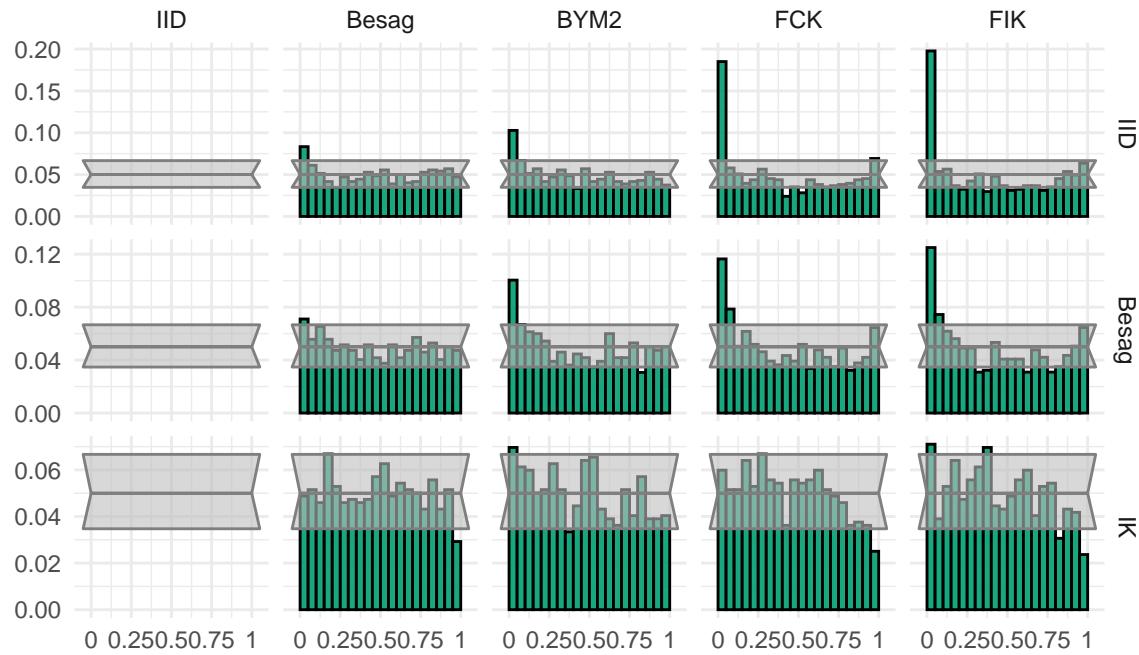
## Grid



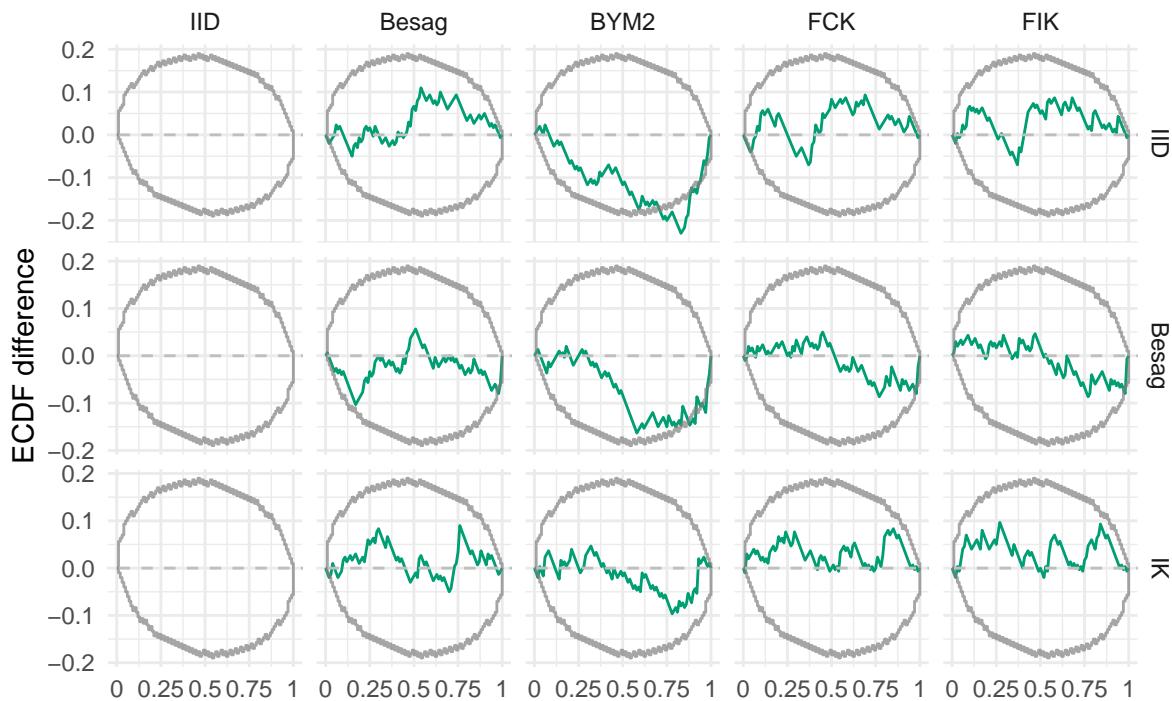
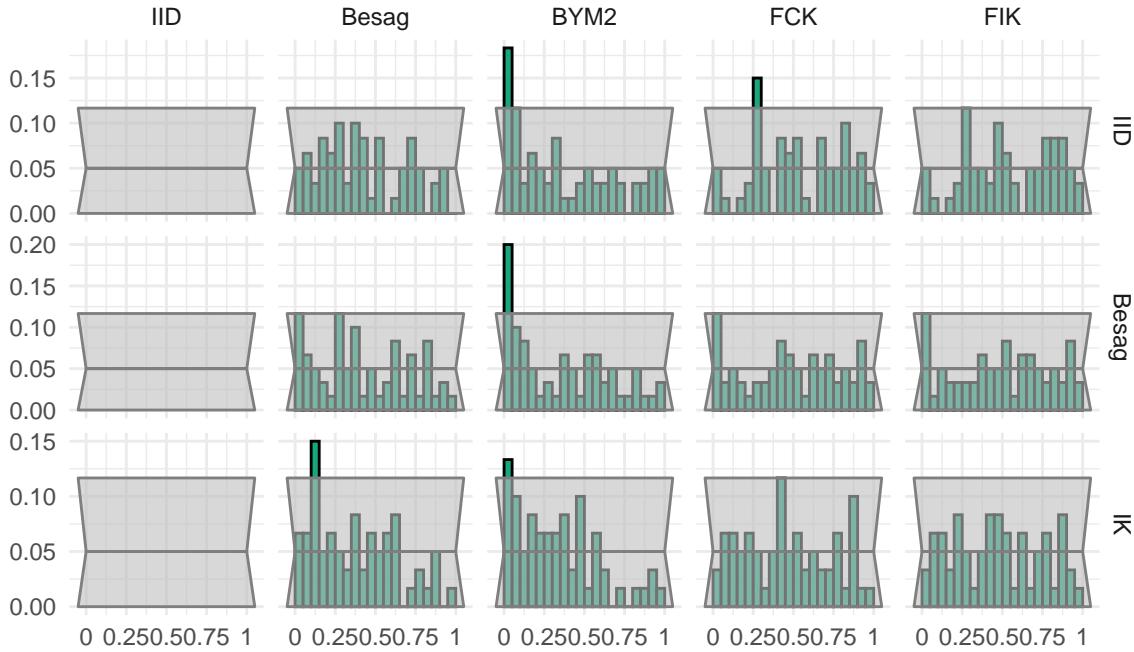
## Cote d'Ivoire



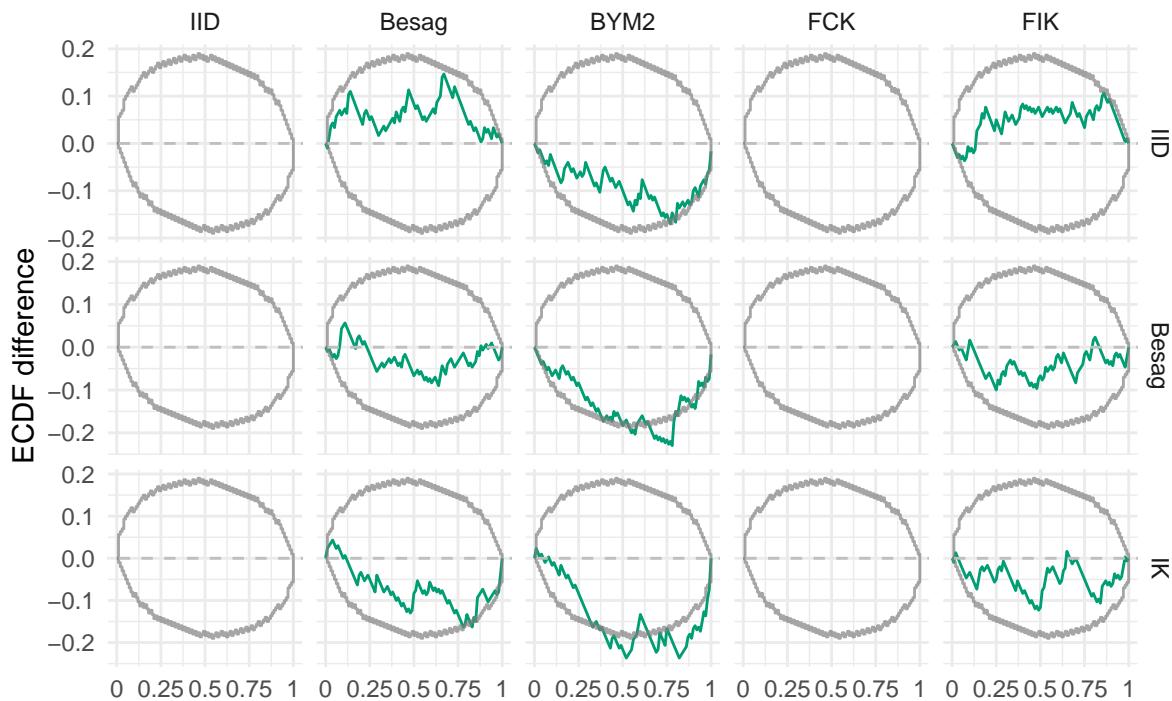
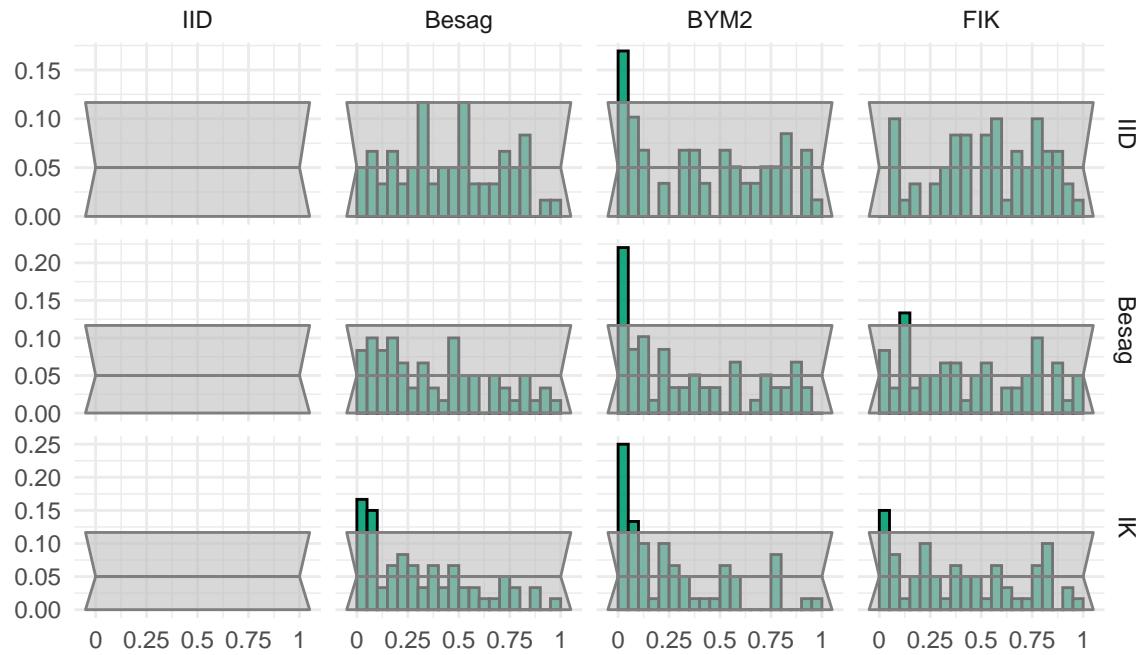
## Texas



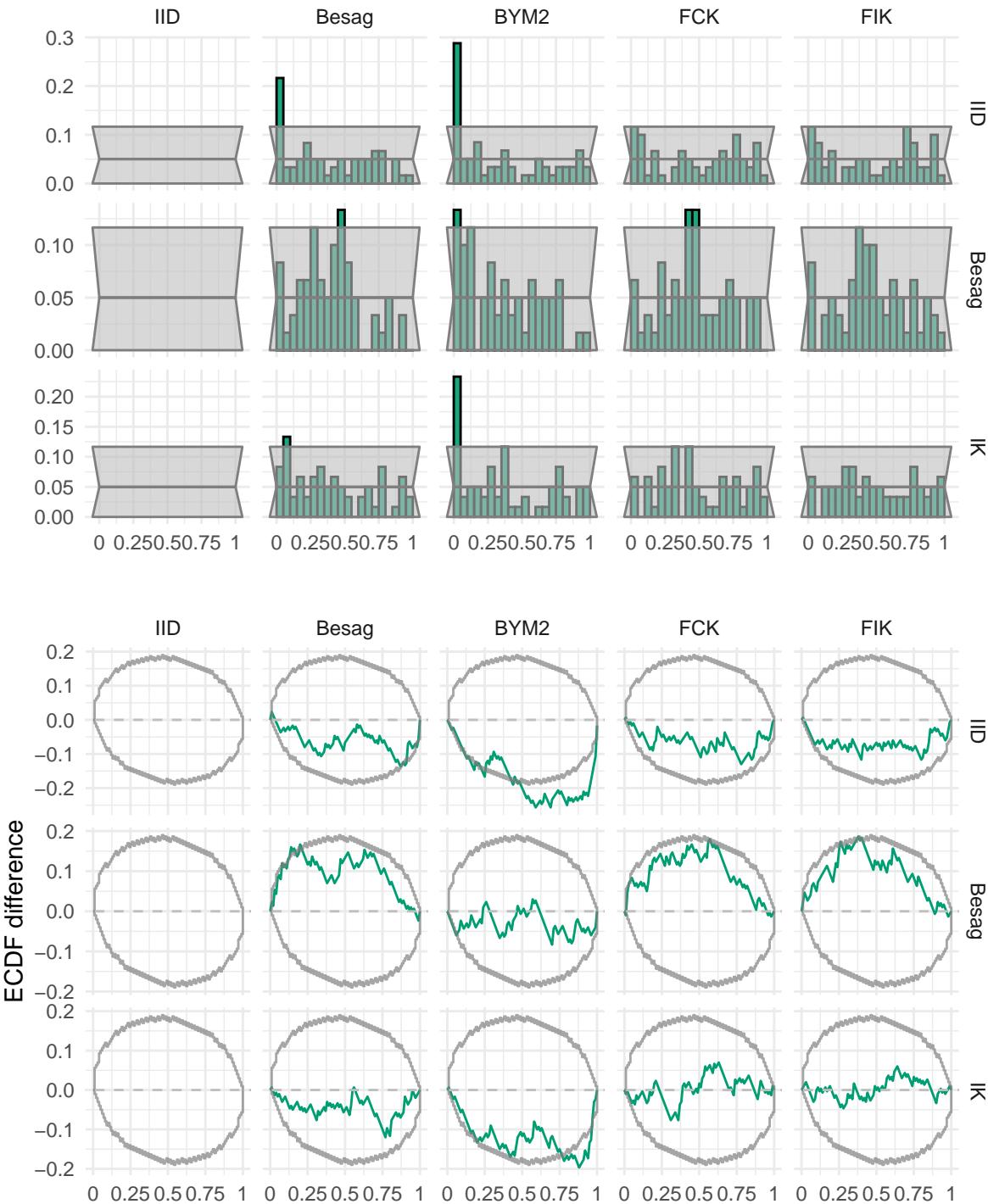
1



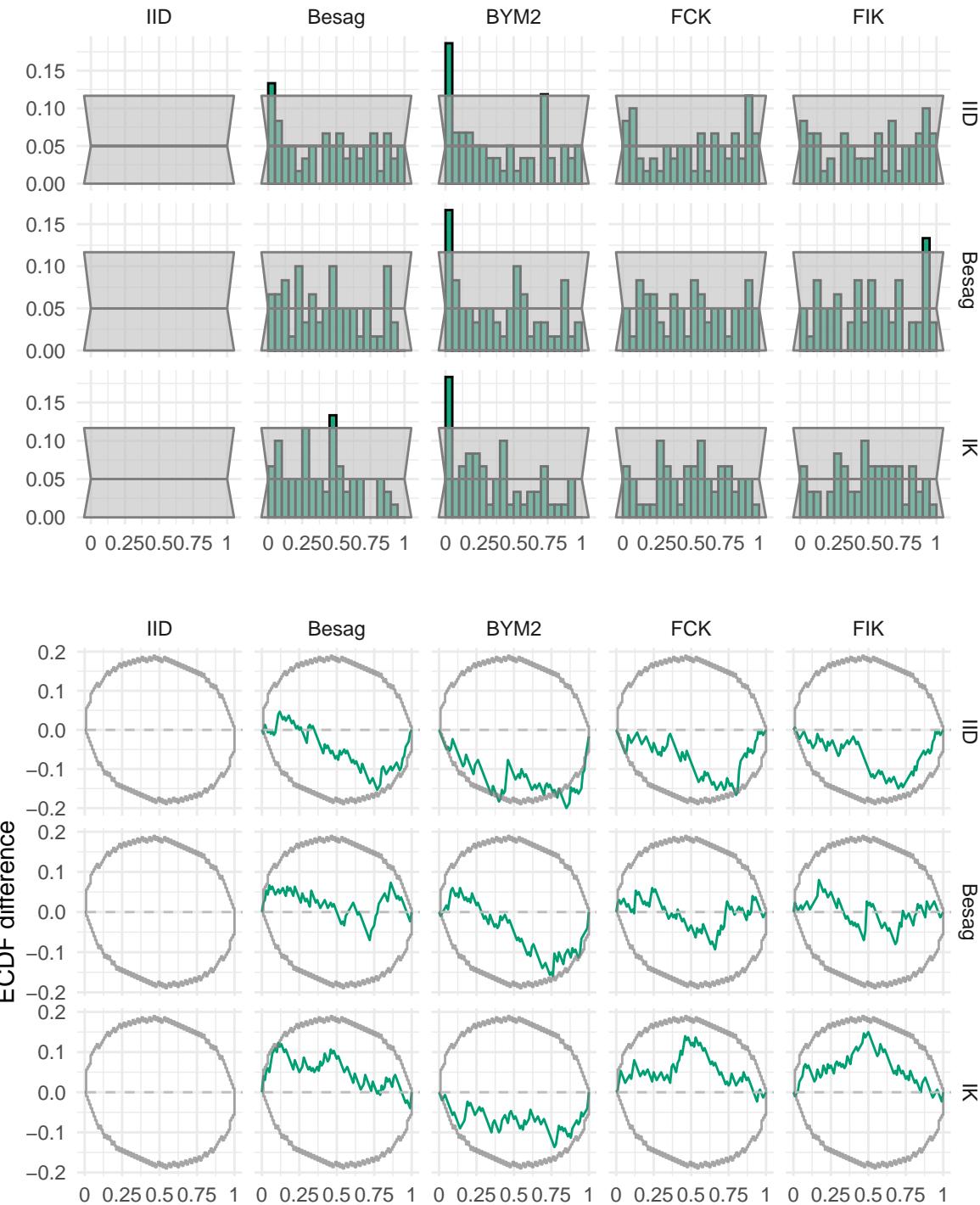
2

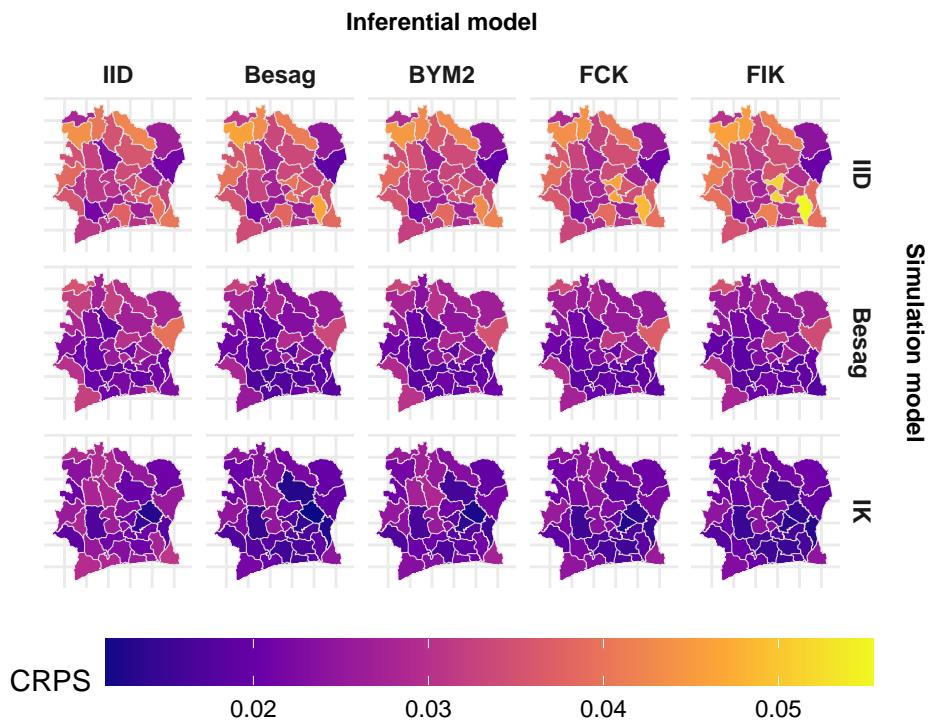


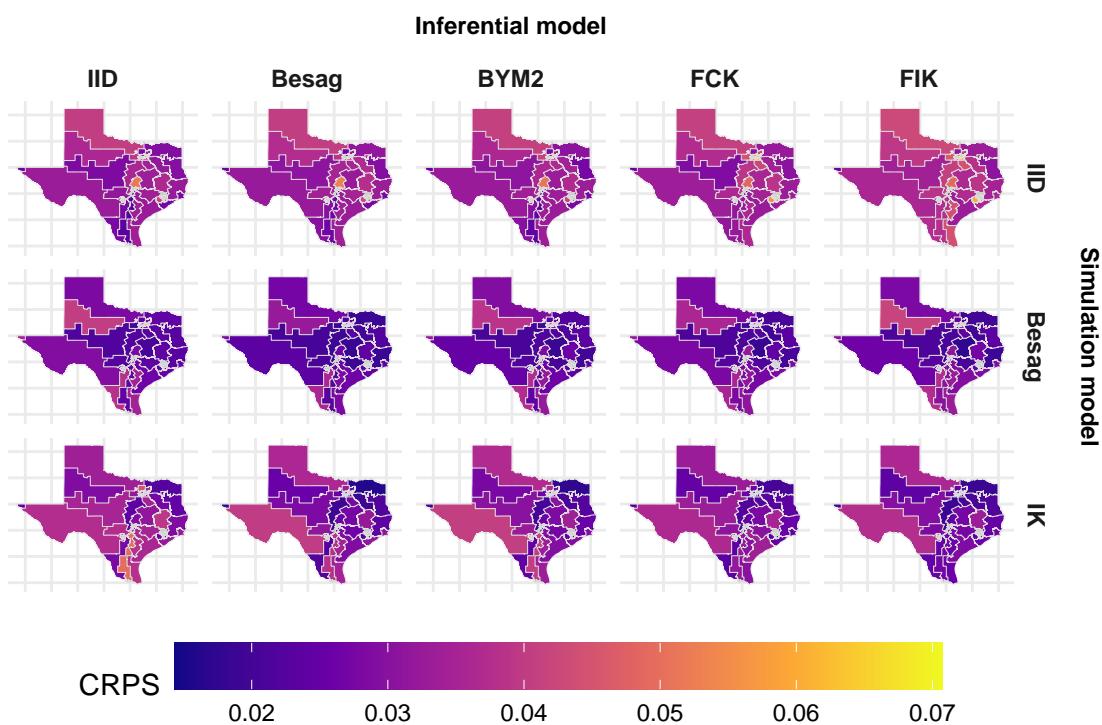
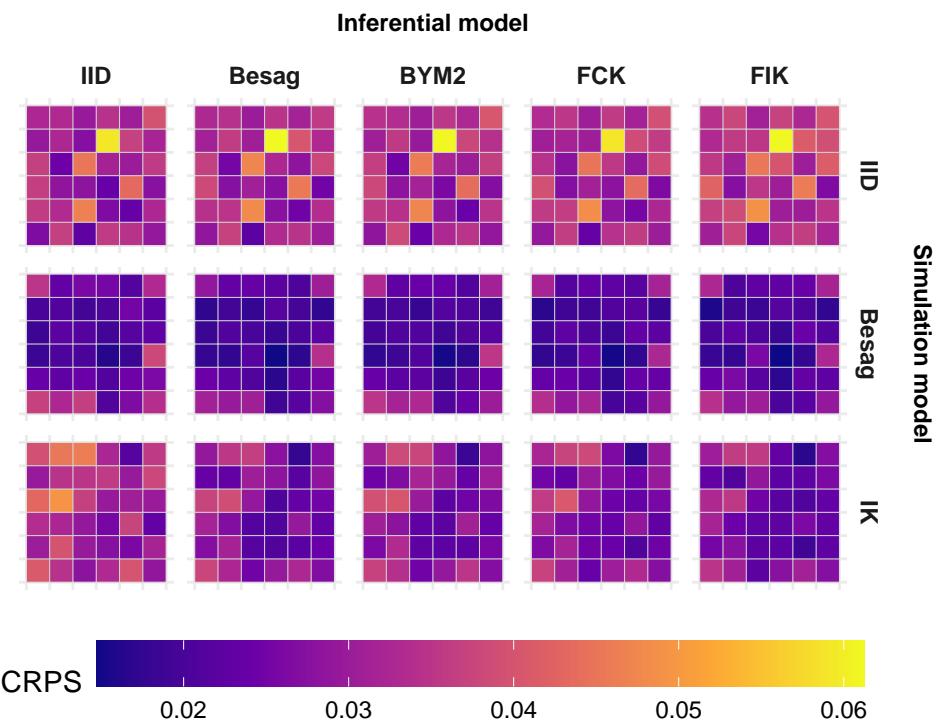
3

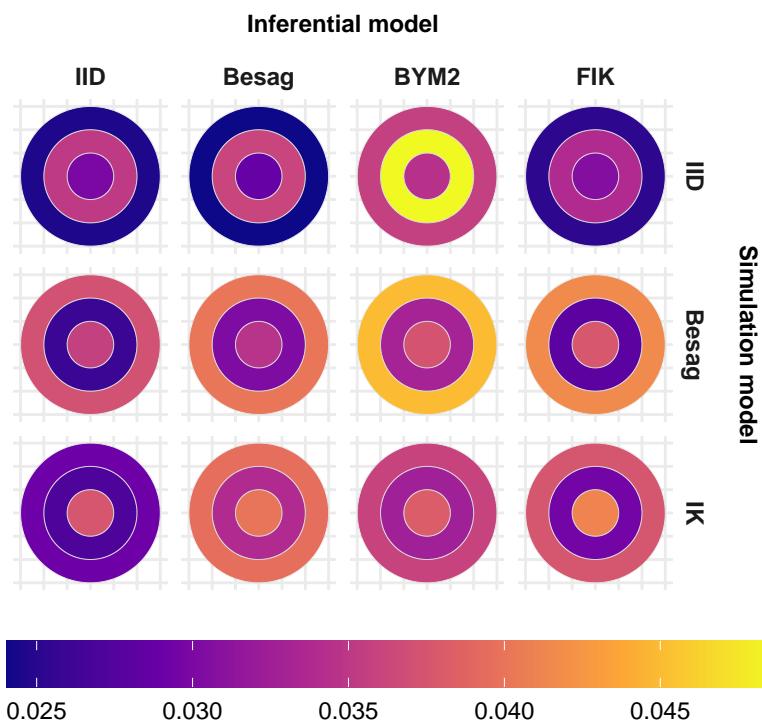
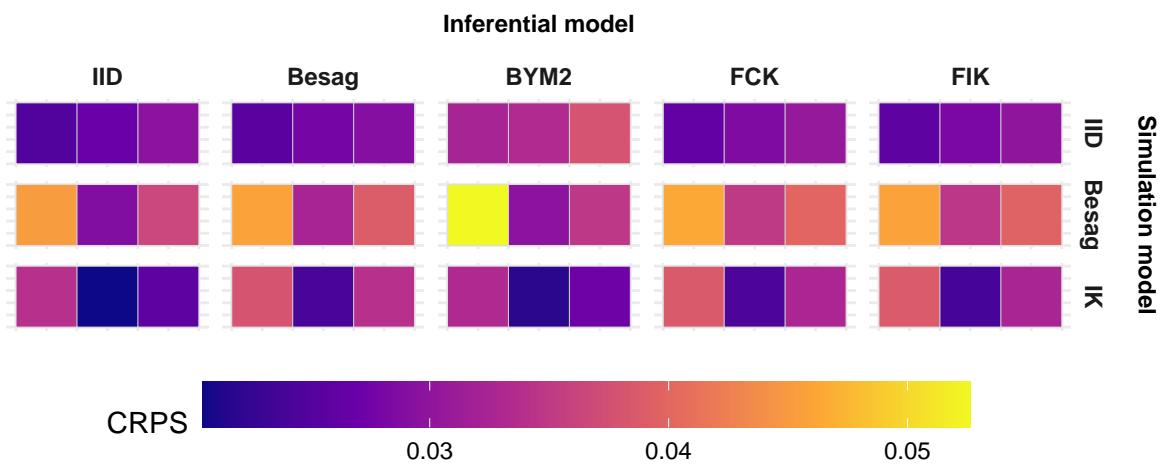


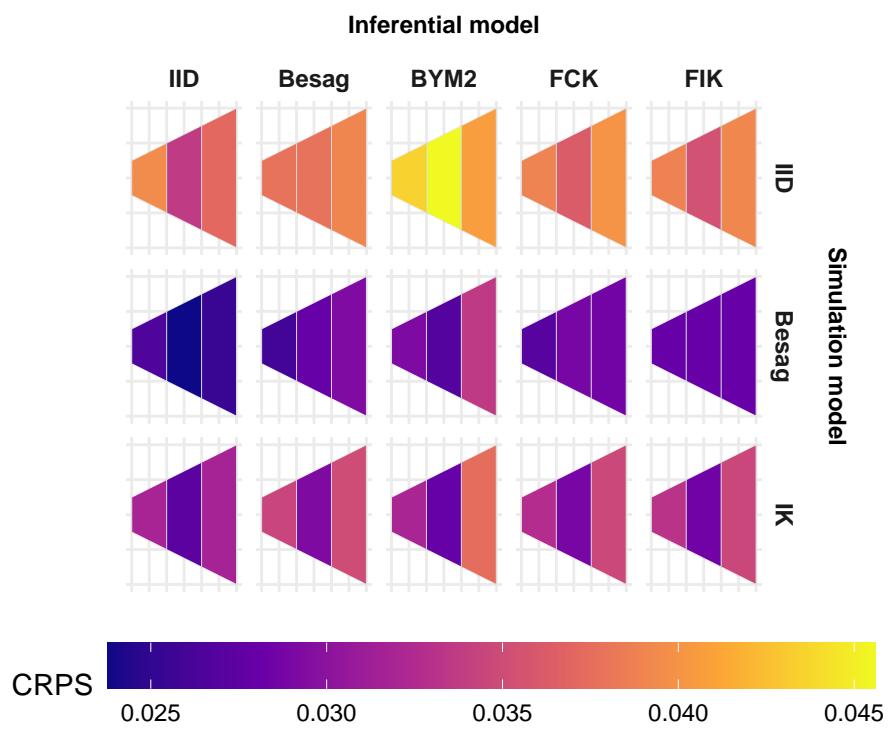
4











## **4 HIV study**

## References

- Freni-Stortino, Anna, Massimo Ventrucci, and Håvard Rue. 2018. “A Note on Intrinsic Conditional Autoregressive Models for Disconnected Graphs.” *Spatial and Spatio-Temporal Epidemiology* 26: 25–34.
- Sørbye, Sigrunn Holbek, and Håvard Rue. 2014. “Scaling Intrinsic Gaussian Markov Random Field Priors in Spatial Modelling.” *Spatial Statistics* 8: 39–51.