WIFO ■ WORKING PAPERS
722/2026

estimateW: An R Package for Bayesian Estimation of Weight Matrices in Spatial Econometric Panels

Tamás Krisztin
Philipp Piribauer

estimateW: An R Package for Bayesian Estimation of Weight Matrices in Spatial Econometric Panels

Tamás Krisztin (IIASA), Philipp Piribauer (WIFO)

WIFO Working Papers 722/2026
March 2026

Abstract

This document introduces the R library estimateW to estimate spatial weight matrices for Bayesian spatial econometric panel models. The approach focuses on spatial weights that are binary prior to row-standardization. However, unlike recent literature our approach requires no strong a priori assumptions on (socio-)economic distances between the spatial units. The estimation approach relies on efficient Bayesian Gibbs sampling techniques and the library supports a variety of the most common spatial econometric panel specifications. estimateW moreover supports to elicit flexible shrinkage priors, which allow to estimate spatial spillovers even in settings where the number of time period is small relative to number of cross-sectional units. An empirical illustration for European NUTS-1 regions demonstrates that the method recovers plausible spatial dependence patterns, interpretable spillover effects, and meaningful clustering in the estimated network structure.

E-Mail: philipp.piribauer@wifo.ac.at

2026/1/W/0

Media owner (publisher), producer: Austrian Institute of Economic Research
1030 Vienna, Arsenal, Objekt 20 | Tel. (43 1) 798 26 01 0 | https://www.wifo.ac.at
Place of publishing and production: Vienna

Free download: https://www.wifo.ac.at/publication/pid/67747780

estimateW: An R package for Bayesian estimation of weight matrices in spatial econometric panels

Tamás Krisztin

International Institute for Applied Systems Analysis, Laxenburg, Austria

Philipp Piribauer

Austrian Institute for Economic Research (WIFO), Vienna, Austria

Abstract

This document introduces the R library **estimateW** to estimate spatial weight matrices for Bayesian spatial econometric panel models. The approach focuses on spatial weights that are binary prior to row-standardization. However, unlike recent literature our approach requires no strong a priori assumptions on (socio-)economic distances between the spatial units. The estimation approach relies on efficient Bayesian Gibbs sampling techniques and the library supports a variety of the most common spatial econometric panel specifications. **estimateW** moreover supports to elicit flexible shrinkage priors, which allow to estimate spatial spillovers even in settings where the number of time period is small relative to number of cross-sectional units. An empirical illustration for European NUTS-1 regions demonstrates that the method recovers plausible spatial dependence patterns, interpretable spillover effects, and meaningful clustering in the estimated network structure.

Keywords: Bayesian spatial econometrics, spatial weight matrix estimation, regional economic growth, R.

JEL Classification: C11, C23, C31, R11.

Corresponding author: Philipp Piribauer, Austrian Institute for Economic Research (WIFO), Arsenal 20, 1030 Vienna, Austria. E-mail: philipp.piribauer@wifo.ac.at. URL: `https://www.wifo.ac.at/philipp.piribauer/`.

1 Introduction

Spatial econometric panel models are central in empirical research on regional interdependencies, policy diffusion, and network externalities across space and time. These models rely on an $n \times n$ spatial weight matrix $\boldsymbol{W}$ and a spatial autoregressive parameter ρ to encode cross-sectional linkages and propagate shocks via the spatial multiplier $(\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1}$. In most applications, $\boldsymbol{W}$ is treated as exogenous and row-stochastic, commonly derived from contiguity, distance bands, or nearest neighbors, with nonnegative entries and zeros on the diagonal to ensure a stable parameter space. While this practice facilitates estimation, it risks misspecifying the underlying network structure and may mask heterogeneity in spatial spillovers that is substantively relevant for policy analysis and regional economics.

This paper introduces **estimateW**, an R package for Bayesian estimation of spatial weight matrices within spatial econometric panels. **estimateW** (Krisztin and Piribauer, 2026) is available from the Comprehensive R Archive Network (CRAN) at `https://CRAN.R-project.org/package=estimateW`. The package targets binary adjacency structures prior to row-standardization and treats off-diagonal entries of the spatial weight matrix as unknown parameters, aligning with common spatial assumptions while relaxing strong a priori contiguity rules. Estimation proceeds via a Gibbs sampling framework with element-wise Bernoulli updates for the adjacency matrix, semi-conjugate updates for slope and variance parameters, and a common griddy Gibbs (Ritter and Tanner, 1992) (or optionally a Metropolis–Hastings step) for the spatial autoregressive parameter. Computational efficiency is achieved through exact, low-cost log-determinant and inverse updates that exploit Sherman–Morrison type identities, enabling estimation with large cross-sections and limited time periods.[1]

A key design feature is a two-part prior for the spatial weights: a spatially structured component that encodes prior spatial knowledge and a sparsity component that regularizes the number of neighbors per unit. The framework preserves established impact reporting by delivering posterior summaries of average direct, indirect, and total effects (LeSage and Pace, 2009), while also providing full posterior draws of $\boldsymbol{W}$ for heterogeneity analyses beyond conventional aggregate effects. The package covers the standard taxonomy of spatial econometric panel specifications used in applied research, including Spatial Durbin Models (SDM), Spatial Error Models (SEM), Spatial Autoregressive Models (SAR), Spatial Durbin Error Models (SDEM), and SLX panels.

Allowing the spatial weight matrix to be unknown raises nontrivial identification challenges, since the autoregressive parameter and network structure jointly shape the spatial transmission of shocks. Building on recent identification results for network formation in

[1] Simulation evidence in Krisztin and Piribauer (2023) and the regional growth application in Piribauer et al. (2023) further indicate that estimating the spatial weight matrix can perform well even when the number of time periods T is relatively small.

spatial panels by De Paula et al. (2025), the paper discusses sufficient assumptions that ensure global identification of the parameters.

An empirical illustration for European NUTS-1 regions demonstrates the practical value of the approach. The results indicate substantial spatial dependence, parameter signs consistent with growth theory, and economically meaningful clustering in the estimated network and multipliers. The example also highlights the role of sparsity priors when $n \gg T$, which enhance stability and yield interpretable adjacency structures for inference on regional spillovers.

The remainder of the article is structured as follows. Section 2 discusses the model architecture and estimation approach implemented in **estimateW**, along with key aspects of parameter identification. Section 3 presents the options for eliciting priors for the spatial weight matrix and the remaining model parameters. Section 4 provides details on the Markov-chain Monte Carlo (MCMC) estimation approach. Section 5 offers an empirical illustration using subnational economic data for European NUTS-1 regions. Section 6 concludes.

2 Spatial autoregressive panels with unknown spatial weights

Consider a panel spatial autoregressive framework for n cross-sectional observations and $t = 1, ..., T$ time observations of the form:

$$\boldsymbol{y}_t = \rho \boldsymbol{W} \boldsymbol{y}_t + \boldsymbol{X}_t \boldsymbol{\beta}_1 + \boldsymbol{W} \boldsymbol{X}_t \boldsymbol{\beta}_2 + \boldsymbol{Z}_t \boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}_t, \tag{1}$$

where $\boldsymbol{y}_t$ is an n by 1 matrix of the dependent variable measured in time t. $\boldsymbol{W}\boldsymbol{y}_t$ is a spatial lag in the dependent variable where $\boldsymbol{W}$ denotes a $n \times n$ row-stochastic and non-negative spatial weight matrix. ρ is a corresponding (scalar) spatial autoregressive parameter with sufficient stability condition $|\rho| < 1$.

Matrix $\boldsymbol{Z}_t$ $(n \times q_2)$ contains explanatory variables that are not spatially lagged and may include a constant or fixed effects. The second matrix of regressors $\boldsymbol{X}_t$ $(n \times q_1)$ enters both with and without a spatial lag. $\boldsymbol{\varepsilon}_t$ is a vector of normally distributed errors, $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_n)$.

Equation (1) corresponds to the panel Spatial Durbin Model (SDM), which is used as a general starting point because it nests the most common spatial panel specifications (see LeSage and Pace, 2009, or Elhorst, 2010). By imposing restrictions on the spatial lags in the dependent variable, regressors, or error process, the SDM reduces to SAR, SEM, SDEM, or SLX-type models, each of which is supported by the package described below.

In most spatial econometric specifications, the $n \times n$ spatial weight matrix $\boldsymbol{W}$ is assumed exogenously known, which implies a number of unknown parameters (the slope parameters, ρ and σ^2) of the model in (1) of $2 + 2q_1 + q_2$. In contrast, the main focus of the package is to treat the $n^2 - n$ off-diagonal elements of $\boldsymbol{W}$ as unknown parameters (the diagonal being restricted to zero). This dramatically increases the dimensionality of the model: the total number of unknown parameters becomes $2 + (n^2 - n) + 2q_1 + q_2$. When the number of time periods T is modest relative to n, this leads to a high-dimensional inference problem in which the number of parameters may far exceed the effective sample size nT. Such settings require regularization, making prior information essential for ensuring stable estimation and meaningful inference.

In practice, the computational burden of our approach grows quickly with n. For this reason, the current implementation of **estimateW** is primarily intended for spatial panels with a moderate number of cross-sectional (regional) units, and full estimation of $\boldsymbol{W}$ is recommended for applications with roughly $n \lesssim 300$. For larger systems, users may consider fixing parts of the spatial structure ex ante or relying on exogenous weight matrices.

2.1 Estimation of spatial weight matrices

Consider an $n \times n$ binary adjacency matrix $\boldsymbol{\Omega}$ with element ω_{ij}. The spatial weight matrix is obtained through row-standardization, $\boldsymbol{W} = \text{rs}(\boldsymbol{\Omega})$, where $\text{rs}(\cdot)$ denotes the row-standardization operator. A typical element of the spatial weight matrix w_{ij} is then given by

$$w_{ij} = \begin{cases} \omega_{ij} / \sum_{j=1}^{n} \omega_{ij} & \text{if } \sum_{j=1}^{n} \omega_{ij} > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Constructing $\boldsymbol{W}$ from a binary adjacency matrix is standard in spatial econometrics, with common choices including contiguity, k-nearest neighbours, or distance-band definitions. In these cases, however, $\boldsymbol{\Omega}$ is typically treated as exogenously known. In contrast, **estimateW** treats all off-diagonal elements of the adjacency matrix (and thus of the spatial weight matrix) as unknown and estimates them jointly with the remaining model parameters.

Following Krisztin and Piribauer (2023), the package employs a Bayesian Markov Chain Monte Carlo (MCMC) algorithm to estimate the unknown parameters. Conditional posteriors are derived for the blocks σ^2, the slope coefficients $\boldsymbol{\beta}$, the spatial autoregressive parameter ρ, and the adjacency matrix $\boldsymbol{\Omega}$. Semi-conjugate normal and inverse-gamma priors are used for $\boldsymbol{\beta}$ and σ^2, while ρ follows a flexible four-parameter beta prior following LeSage and Pace (2009).

Estimation of the elements of the adjacency matrix is based on recent work by Krisztin and Piribauer (2023). The Bayesian MCMC estimation approach combines likelihood with prior information and uses conditional posterior distributions to produce estimates and inference of the unknown parameters.

For the SDM-type panel model in Eq. (1), the likelihood is

$$p(\mathcal{D}|\bullet) = \frac{1}{(2\pi\sigma^2)^{nT/2}}|\boldsymbol{S}| \exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{S}\boldsymbol{Y} - \boldsymbol{U}\boldsymbol{\beta})'(\boldsymbol{S}\boldsymbol{Y} - \boldsymbol{U}\boldsymbol{\beta})\right\}, \tag{3}$$

where $\mathcal{D}$ denotes the data, and $\bullet$ summarizes the model-specific unknown parameters. To simplify notation, the remaining quantities collect the respective variables, $\boldsymbol{S} = \boldsymbol{I}_T \otimes (\boldsymbol{I}_n - \rho \boldsymbol{W})$, $\boldsymbol{Y} = [\boldsymbol{y}_1', ..., \boldsymbol{y}_T']'$, $\boldsymbol{\beta} = [\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', \boldsymbol{\beta}_3']'$, and matrix $\boldsymbol{U}$ collects the respective $Q = 2q_1 + q_2$ explanatory variables.

The conditional posterior of ω_{ij}, $p(\omega_{ij}|\boldsymbol{\Omega}_{-ij}, \bullet, \mathcal{D})$ is given by:

$$p(\omega_{ij}|\boldsymbol{\Omega}_{-ij}, \bullet, \mathcal{D}) \propto p(\mathcal{D}|\bullet)p(\omega_{ij}), \tag{4}$$

The posterior distribution for ω_{ij}, conditional on all the other entries $\boldsymbol{\Omega}_{-ij}$ and the other unknown parameters can be written as:

$$p(\omega_{ij}|\boldsymbol{\Omega}_{-ij}, \bullet, \mathcal{D}) \sim \text{Bernoulli}\left(\frac{\overline{p}_{ij}^{(1)}}{\overline{p}_{ij}^{(0)} + \overline{p}_{ij}^{(1)}}\right). \tag{5}$$

Using the law of total probability, the binary nature of the individual entries in $\boldsymbol{\Omega}$ results in a Bernoulli-distributed conditional posterior distribution for any proper prior choice of $p(\omega_{ij})$. As a result, standard Gibbs sampling can be used for each element in $\boldsymbol{\Omega}$.

$\overline{p}_{ij}^{(0)} = p(\omega_{ij} = 0 \mid \boldsymbol{\Omega}_{-ij}, \bullet, \mathcal{D})$ and $\overline{p}_{ij}^{(1)} = p(\omega_{ij} = 1 \mid \boldsymbol{\Omega}_{-ij}, \bullet, \mathcal{D})$ are obtained by evaluating (4) at $\omega_{ij} = 0$ and $\omega_{ij} = 1$. As shown in Krisztin and Piribauer (2023), the resulting Gibbs sampler mixes well and converges rapidly even in cases with a limited number of observations.

2.2 Impact assessment

The model in equation (1) is written in structural form, where $\rho \boldsymbol{W} y_t$ captures the spatial lag of the endogenous variable. Its reduced form is given by:

$$\boldsymbol{y}_t = (\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1}(\boldsymbol{X}_t\boldsymbol{\beta}_1 + \boldsymbol{W}\boldsymbol{X}_t\boldsymbol{\beta}_2 + \boldsymbol{Z}_t\boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}_t), \tag{6}$$

where $(\boldsymbol{I}_n - \rho \boldsymbol{W})$ is the spatial filter and its inverse the spatial multiplier. Under non-negative, row-stochastic $\boldsymbol{W}$ with $\|\boldsymbol{W}\|_\infty \leq 1$ and $|\rho| < 1$, the spatial multiplier admits the Neumann expansion:

$$(\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1} = \sum_{r=0}^{\infty} \rho^r \boldsymbol{W}^r = \boldsymbol{I}_n + \rho \boldsymbol{W} + \rho^2 \boldsymbol{W}^2 + \dots \tag{7}$$

Equations (6)–(7) illustrate that the slope coefficients cannot be interpreted as marginal effects, since shocks at one spatial unit may propagate throughout the entire system. For the lth exogenous variable ($l = 1, \ldots, q_1$) we follow De Paula et al. (2025) and link the relationship between structural and reduced-form parameters:

$$\boldsymbol{\Pi}_l = (\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1}(\boldsymbol{I}_n \beta_{1,l} + \boldsymbol{W}\beta_{2,l}). \tag{8}$$

Empirical studies typically summarize the effects in (8) using the metrics introduced by LeSage and Pace (2009). The diagonal elements of $\boldsymbol{\Pi}_l$ represent direct effects, and the average direct impact is given by $\text{trace}(\boldsymbol{\Pi}_l)/n$. Off-diagonal elements capture spillover (indirect) effects. The average indirect impact is $\mathbf{1}_n'(\boldsymbol{\Pi}_l - \text{diag}(\boldsymbol{\Pi}_l))\mathbf{1}_n/n$, where $\mathbf{1}_n$ is an $(n \times 1)$ vector of ones, and the average total impact is the sum of the average direct and indirect impacts. **estimateW** automatically reports average direct, indirect, and total impacts for all exogenous variables, and provides full MCMC output for each effect, enabling the computation of credible intervals and additional summaries.

Using these average impact measures is appropriate when $\boldsymbol{W}$ is exogenous, since spatial heterogeneity is then driven solely by ρ. When $\boldsymbol{W}$ is estimated from the data, however, average effects may obscure meaningful heterogeneity in spillovers, which may itself be of substantive interest. Because **estimateW** retains all MCMC draws of $\boldsymbol{W}$ and the corresponding impact matrices, more detailed analyses are straightforward. For example, Piribauer et al. (2023) decompose spatial spillovers into spill-in and spill-out effects to study the complex patterns of regional labour productivity growth in Europe.

2.3 Parameter identification

Treating the elements of the spatial weight matrix $\boldsymbol{W}$ as unknown introduces additional identification challenges that go beyond those associated with the spatial and slope parameters. The nature of these challenges depends on the specific model specification and does not hinge on whether the number of observations exceeds the number of parameters to be estimated.[2] A thorough discussion of identification in general spatial autoregressive models is provided by De Paula et al. (2025). Under six sufficient assumptions, global identification of the parameters of interest can be established (see Corollary 3 in De Paula et al., 2025).[3]

I. $[\boldsymbol{W}]_{ii} = 0 \; \forall i$: This assumption rules out self-links and is automatically satisfied when using `estimateW`, where the diagonal of $\boldsymbol{\Omega}$ (and hence $\boldsymbol{W}$) is set to zero by construction.

[2] In situations with a limited number of observations, we suggest applying shrinkage priors as outlined in the following section.

[3] We maintain the standard assumption that the matrix of explanatory variables is of full rank.

II. $\sum_{j=1}^{n} |\rho[\boldsymbol{W}]_{ij}| < 1 \; \forall i$, $\rho < 1$, and $\|\boldsymbol{W}\|_{\infty} < C$ for some positive $C \in \mathbb{N}$. These conditions ensure that feedback effects decay as they propagate through the network and guarantee non-singularity of $(\boldsymbol{I}_n - \rho \boldsymbol{W})$. Under the default settings, **estimateW** enforces these restrictions during estimation.

III. $\rho\beta_{1,q} + \beta_{2,q} \neq 0$ for $q = 1, \ldots, q_1$. This standard regularity condition for SDM specifications excludes cases where endogenous and exogenous spatial effects exactly offset each other. It is not imposed mechanically, but can be checked ex post using the MCMC output.

IV. There is at least one i such that $\sum_{j=1}^{n} [\boldsymbol{W}]_{ij} = 1$: It is required that at least one row of $\boldsymbol{W}$ adds to a fixed and known number. In **estimateW**, the spatial weight matrix is constructed as row-stochastic by default (i.e. its rows sum to unity; or zero in the case of no neighbors). Additional priors can also be chosen to ensure a minimum (and also a maximum) number of neighbours (e.g. to ensure that each spatial observation has at least one neighbour).[4]

V. The diagonal elements of $\boldsymbol{W}^2$ are not constant (network asymmetry condition; see also Bramoullé et al., 2009). This condition ensures sufficient heterogeneity in second-order neighbourhood structures and can be verified in a post-estimation step.

VI. $\rho > 0$ and $[\boldsymbol{W}]_{ij} \geq 0$. Non-negativity of $\boldsymbol{W}$ is satisfied by construction. The sign restriction on ρ is the only non-standard requirement relative to spatial models with exogenous weights. Since most empirical applications suggest positive spatial interactions, the default prior in `estimateW` restricts ρ to $(0, 1)$, although the common range $(-1, 1)$ can be used instead. If $\boldsymbol{W}$ is not irreducible, no sign restriction on ρ is required (see De Paula et al., 2025). Users may therefore either rely on the positive-sign restriction for ρ or allow unrestricted ρ while incorporating stronger priors on the spatial structure of $\boldsymbol{W}$.[5]

2.4 Other spatial panel models supported by the package

In addition to the SDM-type model sketched above, the package supports a broad set of spatial econometric panel specifications. Following the taxonomy in Elhorst (2010), Table 1 illustrates the alternative panel models implemented in the package. These classes differ with respect to which model components include spatial lags (endogenous variable,

[4]Note that in the presence of common shocks this assumption can be adjusted by assuming $\sum_{j=1}^{n} [\boldsymbol{W}]_{ij} = 1 \; \forall i$. This adapted assumption ensures that the parameters of the model can be identified even in the presence of common shocks, and can be easily enforced by setting a minimum number of neighbours (for details, see Proposition 1 in De Paula et al., 2025).

[5]Further discussion of identification through spatial priors is available in Krisztin and Piribauer (2023).

exogenous variables, or error term). For example, SDM specifications include a spatial lag of the dependent variable and (at least some) spatially lagged covariates. The table also lists the corresponding function names for each model class.

For every specification, the package provides two function variants: one with an "w" suffix (e.g. `sdmw()`) and one without (e.g. `sdm()`). The "w" functions constitute the core functionality of the package, where (at least parts of) the spatial weight matrix are treated as unknown. The corresponding functions without the suffix estimate the same models under an exogenously specified spatial weight matrix. The package therefore also supports estimation of the most common spatial econometric panel models under predefined $\boldsymbol{W}$.

Beyond the models shown in Table 1, the package allows users to construct more customized spatial panel specifications. It provides modular objects that generate draws from the respective conditional posterior distributions for all parameter blocks. These building blocks can be used to assemble MCMC schemes for alternative spatial panel models, such as dynamic spatial panels (Debarsy et al., 2012), models with instrumental variables (Kelejian and Prucha, 1998), time-varying spatial structures (Iacopini et al., 2025), spatial parameter shrinkage approaches (Pfarrhofer and Piribauer, 2019), and others.

Table 1: Spatial econometric panel specifications in **estimateW**

Model name	**Model specification**	**R Functions**
Spatial Durbin (SDM)	$\boldsymbol{y}_t = \rho \boldsymbol{W} \boldsymbol{y}_t + \boldsymbol{X}_t \boldsymbol{\beta}_1 + \boldsymbol{W} \boldsymbol{X}_t \boldsymbol{\beta}_2 + \boldsymbol{Z}_t \boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}_t$	`sdm()`, `sdmw()`
Spatial Durbin Error (SDEM)	$\boldsymbol{y}_t = \boldsymbol{X}_t \boldsymbol{\beta}_1 + \boldsymbol{W} \boldsymbol{X}_t \boldsymbol{\beta}_2 + \boldsymbol{Z}_t \boldsymbol{\beta}_3 + (\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1} \boldsymbol{\varepsilon}_t$	`sdem()`, `sdemw()`
Spatial Autoregressive (SAR)	$\boldsymbol{y}_t = \rho \boldsymbol{W} \boldsymbol{y}_t + \boldsymbol{Z}_t \boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_t$	`sar()`, `sarw()`
Spatial Error (SEM)	$\boldsymbol{y}_t = \boldsymbol{Z}_t \boldsymbol{\beta}_1 + (\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1} \boldsymbol{\varepsilon}_t$	`sem()`, `semw()`
Spatial Lag X (SLX)	$\boldsymbol{y}_t = \boldsymbol{X}_t \boldsymbol{\beta}_1 + \boldsymbol{W} \boldsymbol{X}_t \boldsymbol{\beta}_2 + \boldsymbol{Z}_t \boldsymbol{\beta}_3 + \boldsymbol{\varepsilon}_t$	`slx()`, `slxw()`

3 Prior specification for spatial weights and model parameters

Given the high dimensionality of the parameter space and the structural role of the spatial weight matrix, prior specification is central for stability, identification, and interpretability. This section describes how priors for the unknown parameters can be specified. Priors for the spatial weight matrix are defined via the function `W_priors()`. The slope coefficients and the error variance parameter σ^2 are handled by `beta_priors` and `sigma_priors()`, respectively. Priors for the spatial autoregressive parameter ρ can be specified using `rho_priors()`.

3.1 Eliciting priors for the spatial weights

The package allows users to specify Bernoulli priors for the elements of the spatial adjacency matrix $\mathbf{\Omega}$, $p(\omega_{ij}) \sim \text{Bernoulli}(\underline{p}_{ij})$, based on two components:

$$\underline{p}_{ij} \propto \underline{\omega}_{ij}\, \underline{m}(k_i). \tag{9}$$

The first component, $\underline{\omega}_{ij} \in [0, 1]$, is a scalar hyperparameter controlling the *spatial structure* of the prior. The second component, $\underline{m}(k_i)$, regularizes the number of neighbours of unit i and induces sparsity in the adjacency matrix $\mathbf{\Omega}$ (and hence in $\boldsymbol{W}$).

The main arguments of the `W_priors()` function are: the scalar `n`, specifying the number of cross-sectional units; `W_prior`, an $n \times n$ matrix containing the hyperparameters $\underline{\omega}_{ij}$; and `nr_neighbors_prior`, an n-dimensional vector encoding the hyperparameters $\underline{m}(k_i)$.

Additional options can be specified. The argument `row_standardized_prior = TRUE` allows estimation of spatial weight matrices without row-standardization.[6] The option `symmetric_prior = TRUE` enforces a symmetric adjacency matrix $\mathbf{\Omega}$. Users may also specify minimum and maximum numbers of neighbours for each spatial unit.

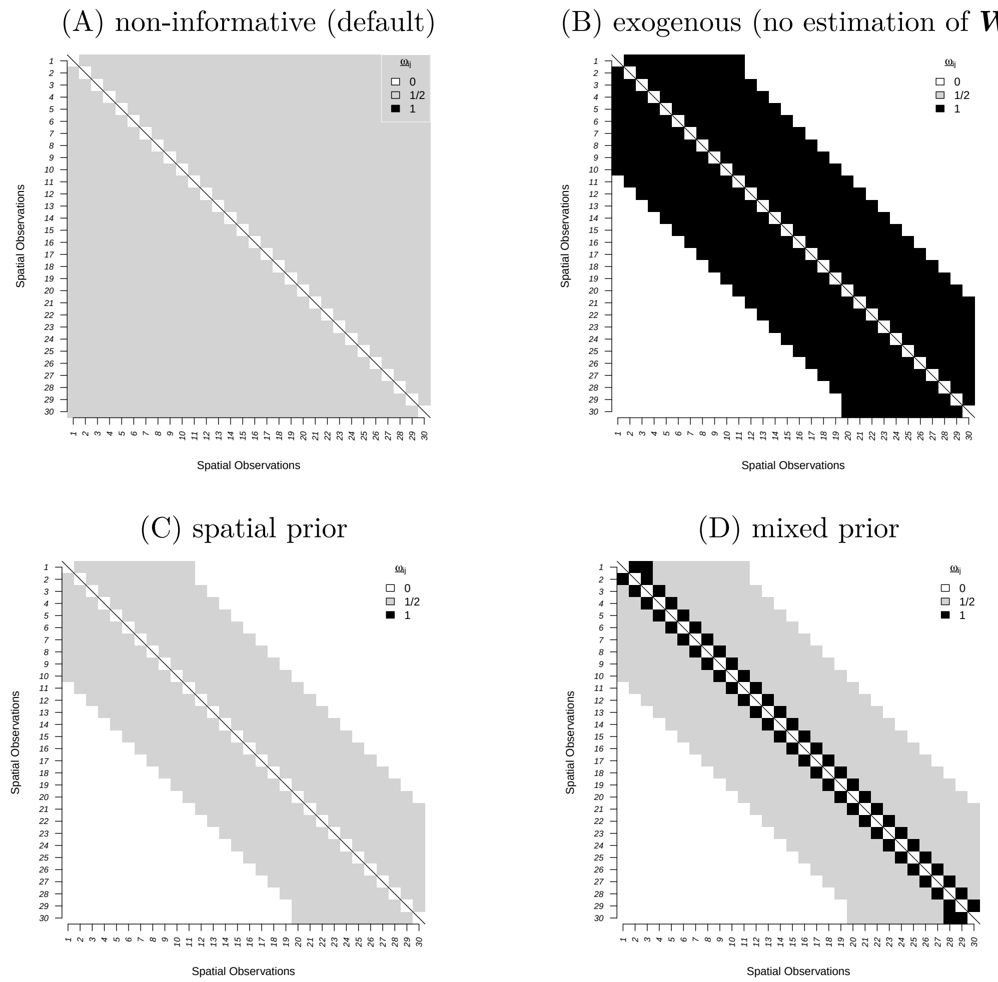
Spatially structured priors, $\underline{\omega}_{ij}$:
The default specification for the spatially structured component is $\underline{\omega}_{ij} = 0.5$ for all $i \neq j$, reflecting the absence of explicit prior information about the spatial structure. Users may, however, define alternative spatial prior configurations in a flexible and transparent way. Larger (smaller) values of $\underline{\omega}_{ij}$ correspond to higher (lower) prior probabilities of inclusion. Spatial prior information can also be encoded through the extreme values 0 and 1: setting $\underline{\omega}_{ij} = 0$ excludes the corresponding link from the adjacency matrix, while setting $\underline{\omega}_{ij} = 1$ forces inclusion. In particular, specifying only 0/1 values in $\underline{\omega}_{ij}$ via the argument `W_prior` reduces the model to one with an exogenous spatial weight matrix.[7]

Figure 1 illustrates four stylised examples of spatially structured priors using only the hyperparameters $\underline{\omega}_{ij}$. The example assumes a *linear city* with 30 regions, where region 1 neighbours region 2, region 2 neighbours regions 1 and 3, and so forth.

[6] However, note that a model without row-standardization may suffer from identification problems, and the admissible parameter space for ρ becomes unclear; see LeSage and Pace, 2009.

[7] The package internally fixes the diagonal of $\underline{\mathbf{\Omega}}$ at zero. Moreover, setting some non-diagonal entries of $\underline{\mathbf{\Omega}}$ to 0 or 1 can substantially reduce computation time, as these elements are no longer sampled.

Figure 1: Some stylized examples for spatially structured priors $\underline{\omega}_{ij}$



Case (A) shows the default specification, where all elements of the adjacency matrix $\boldsymbol{\Omega}$ are treated as unknown ($\underline{\omega}_{ij} = 1/2$ for $i \neq j$). Case (B) represents the standard situation in spatial econometrics, where the spatial weight matrix is treated as known. This is implemented by setting $\underline{\omega}_{ij} = 1$ for neighbouring units and 0 otherwise; in the example, a 10-nearest neighbour structure is used.

Case (C) illustrates a spatial prior that incorporates partial prior knowledge: potential neighbours within a specified distance band are treated as unknown, while units outside that band are excluded. Besides enabling the use of prior spatial information, this approach reduces the number of parameters to be estimated, which can be advantageous in samples with limited observations. Finally, case (D) combines the logic of (B) and (C) by assigning some elements of $\underline{\Omega}$ the values 0, 1, or 1/2.

Sparsity priors on the spatial neighbors, $\underline{m}(k_i)$:
The second prior component in equation (10), $\underline{m}(k_i)$, governs the degree of sparsity in the spatial weight matrix by controlling the prior on the number of neighbors of unit i. Let $k_i = \sum_{j=1}^{n} \omega_{ij}$ denote the number of nonzero entries in the i-th row of the adja-

cency matrix $\boldsymbol{\Omega}$. The function `W_priors()` allows users to specify an n-dimensional vector of prior weights over the possible values $k_i \in \{0, 1, \ldots, n-1\}$ via the argument `nr_neighbors_prior`.

Table 2 illustrates three shrinkage priors for $\underline{m}(k_i)$ for a stylised sample of $n = 30$ spatial units under a non-informative spatial structure (case (A) in Figure 1). For each prior, the table shows its analytic form and corresponding R implementation, the shape of $\underline{m}(k_i)$, and the implied prior distribution of k_i.

Case a) no shrinkage. A flat prior $\underline{m}(k_i) \propto 1$ is implemented by `nr_neighbors_prior = rep(1, n)`.[8] This corresponds to independent Bernoulli priors with $\underline{p}_{ij} = 1/2$ for all $i \neq j$ without using the shrinkage component ($\underline{m}(k_i) \propto 1$), so that

$$p(k_i) = \text{Binomial}(n-1, 0.5),$$

with a priori expected number of neighbors of $(n-1)/2$. Even though this setup appears non-informative, it places substantial prior mass on dense networks, as seen in the right panel of case a) in the table. For large n and limited T, such implicit informativeness may lead to prior dominance and weak posterior identification. This flat setup of $\underline{m}(k_i)$ is shown in the second column of the table and can be specified with the argument `nr_neighbors_prior=rep(1,n)` in the function `W_priors()`.[9]

Case b) default setup. The default setup follows Krisztin and Piribauer (2023) by treating the prior inclusion probabilities $\underline{p}_{ij}$ as random (rather than being fixed) using a Beta prior with hyperparameters $(\underline{a}_\omega, \underline{b}_\omega)$. The resulting prior on k_i is Beta–Binomial (see Ley and Steel, 2009):

$$p(k_i) \propto \binom{n-1}{k_i} \Gamma(\underline{a}_\omega + k_i)\, \Gamma(n - 1 + \underline{b}_\omega - k_i),$$

where $\Gamma(\cdot)$ denotes the Gamma function. However, the user has to define $\underline{m}(k_i)$, which is given by:

$$\underline{m}(k_i) = \Gamma(\underline{a}_\omega + k_i)\, \Gamma(n - 1 + \underline{b}_\omega - k_i). \tag{10}$$

The default choice $\underline{a}_\omega = \underline{b}_\omega = 1$ yields a uniform prior on k_i. This hierarchical prior can be implemented via `bbinompdf()` for the argument `nr_neighbors_prior` in `W_priors()`.

To impose stronger sparsity in high-dimensional settings ($n \gg T$), one may fix $\underline{a}_\omega = 1$ and anchor $\underline{b}_\omega$ to a desired expected number of neighbors $\underline{k}$:

$$\underline{b}_\omega = \frac{(n-1) - \underline{k}}{\underline{k}}. \tag{11}$$

[8] The n-dimensional vector supplied in `nr_neighbors_prior` does not need to integrate to unity; normalization is done internally.

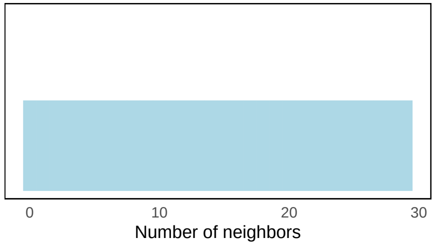
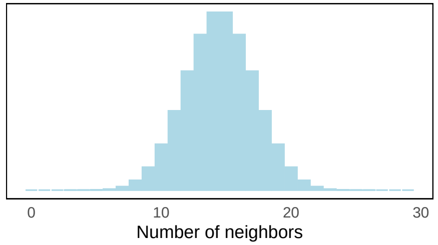
[9] Note that the package automatically ensures that the prior integrates to unity, such that any positive constant may be chosen for the n-dimensional vector in `nr_neighbors_prior`.

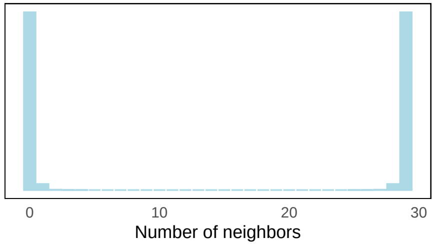
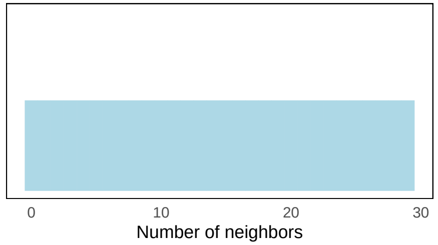
In the default setup with $\underline{a}_\omega = \underline{b}_\omega = 1$, $\underline{k} = (n-1)/2$.

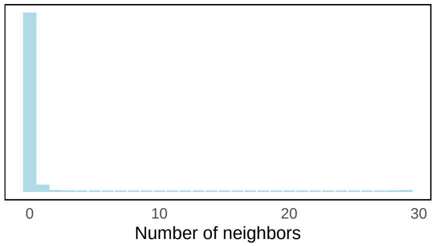
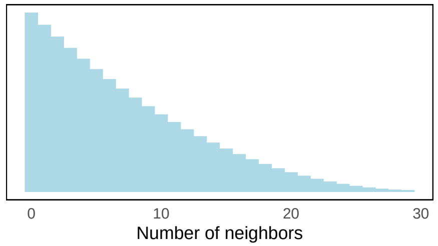
Case c) strong hierarchical shrinkage. For samples with limited temporal information, a more restrictive choice of $\underline{b}_\omega$ (keeping $\underline{a}_\omega = 1$) by controlling the prior expected number of neighbors $\underline{k}$ induces stronger shrinkage toward sparse adjacency structures. The implied prior on k_i is shown in the last column of Table 2. This setup can substantially stabilize estimation and improve identification in high-dimensional spatial panels by concentrating prior mass on a smaller, more plausible number of neighbors per unit.

Table 2: Sparsity priors on the spatial neighbors

Prior setup and implementation	$\underline{m}(k_i)$	**Implied** $p(k_i)$

a) no shrinkage
$\underline{m}(k_i) \propto 1$
 	
`R> n = 30;`
`R> Wpr_a = W_priors(n = n,`
`    +  W_prior = matrix(1/2,n,n),`
`    +  nr_neighbors_prior = rep(1,n))`
		

b) default setup
$\underline{m}(k_i) = \Gamma(\underline{a}_\omega + k_i) + \Gamma(n - 1 + \underline{b}_\omega + k_i)$,
$\underline{a}_\omega = 1,\ \underline{b}_\omega = 1$
 	
`R> n = 30;`
`R> Wpr_a = W_priors(n = n,`
`    + W_prior = matrix(1/2,n,n),`
`    + nr_neighbors_prior =`
`    + bbinompdf(0:(n-1), nsize = n-1,a = 1,b = 1))`
		

c) strong shrinkage
$\underline{m}(k_i) = \Gamma(\underline{a}_\omega + k_i) + \Gamma(n - 1 + \underline{b}_\omega + k_i)$
$\underline{a}_\omega = 1,\ \underline{b}_\omega = \frac{(n-1)-\underline{k}}{\underline{k}},\ \underline{k} = 7$
 	
`R> n = 30; kbar=7; bbar <- (n - 1 - kbar) / kbar;`
`R> Wpr_a = W_priors(n = n,`
`    + W_prior = matrix(1/2,n,n),`
`    + nr_neighbors_prior =`
`    + bbinompdf(0:(n-1), nsize = n-1,a = 1,b = bbar))`
		

3.2 Priors for the remaining parameters

The functions `beta_priors()`, `sigma_priors()`, and `rho_priors()` allow users to specify prior distributions for the remaining parameter blocks $\boldsymbol{\beta}$, σ^2, and ρ, respectively.

- **Slope parameters $\boldsymbol{\beta}$.** A Gaussian prior of the form $p(\boldsymbol{\beta}) \sim \mathcal{N}(\underline{\boldsymbol{\mu}}_\beta, \underline{\boldsymbol{V}}_\beta)$ is used. The function `beta_priors()` allows users to set the prior mean vector $\underline{\boldsymbol{\mu}}_\beta$ $(Q \times 1)$ and the prior variance matrix $\underline{\boldsymbol{V}}_\beta$ $(Q \times Q)$. By default, the prior is centered at zero with a rather diffuse variance $100\boldsymbol{I}_Q$:[10]

```
R> beta_priors(Q,
   + beta_mean_prior = matrix(0, Q, 1),
   + beta_var_prior  = diag(Q) * 100)
```

- **Error variance parameter σ^2.** An inverse-Gamma prior $p(\sigma^2) \sim \mathcal{IG}(\underline{a}_{\sigma^2}, \underline{b}_{\sigma^2})$ is used, with standard default values $\underline{a}_{\sigma^2} = \underline{b}_{\sigma^2} = 0.001$:

```
R> sigma_priors(
    sigma_rate_prior  = 0.001,
    sigma_shape_prior = 0.001)
```

- **Spatial autoregressive parameter ρ.** The prior for ρ is taken from the four-parameter Beta family, $p(\rho) \sim \mathcal{B}_4(\underline{a}_\rho, \underline{b}_\rho, \underline{\rho}_{\min}, \underline{\rho}_{\max})$. The hyperparameters $\underline{\rho}_{\min}$ and $\underline{\rho}_{\max}$ determine the support of ρ, with default values $[0, 1]$.[11] The shape parameters $\underline{a}_\rho$ and $\underline{b}_\rho$ control the form of the distribution and are set to one by default, yielding a uniform prior:

```
R> rho_priors(
   + rho_a_prior = 1,
   + rho_b_prior = 1,
   + rho_min     = 0,
   + rho_max     = 1)
```

[10]Non–spatially lagged regressors ($\boldsymbol{Z}_t$) are always ordered last, i.e. $\boldsymbol{U}_t = [\boldsymbol{X}_t \ \boldsymbol{W}\boldsymbol{X}_t \ \boldsymbol{Z}_t]$ and $\boldsymbol{\beta} = [\boldsymbol{\beta}_1' \ \boldsymbol{\beta}_2' \ \boldsymbol{\beta}_3']'$.

[11]The lower bound may be set to -1 to allow $\rho \in (-1, 1)$. However, when both ρ and $\boldsymbol{W}$ are estimated without informative priors, identification problems may arise. To mitigate this, Krisztin and Piribauer (2023) recommend using informative spatial priors or restricting ρ to be positive. The default setup reflects the latter.

4 Markov Chain Monte Carlo sampling

Given the prior choices outlined above, the conditional posterior distributions for the slope parameters, the spatial autoregressive parameter, and the error variance take standard forms (for an introduction, see LeSage and Pace, 2009). For the specification in equation (1), the conditional posterior of the slope parameters $\boldsymbol{\beta}$ is Gaussian:

$$
\begin{aligned}
p(\boldsymbol{\beta} \mid \bullet, \mathcal{D}) &\sim \mathcal{N}(\overline{\boldsymbol{\mu}}_\beta, \overline{\boldsymbol{V}}_\beta), && (12)\\
\overline{\boldsymbol{\mu}}_\beta &= \sigma^{-2}\overline{\boldsymbol{V}}_\beta \boldsymbol{U}'\boldsymbol{S}\boldsymbol{Y}, && (13)\\
\overline{\boldsymbol{V}}_\beta &= \left(\sigma^{-2}\boldsymbol{U}'\boldsymbol{U} + \underline{\boldsymbol{V}}_\beta^{-1}\right)^{-1}, && (14)
\end{aligned}
$$

with $\boldsymbol{S} = \boldsymbol{I}_T \otimes (\boldsymbol{I}_n - \rho \boldsymbol{W})$. The conditional posterior of the error variance σ^2 follows an inverse-Gamma distribution:

$$
\begin{aligned}
p(\sigma^2 \mid \bullet, \mathcal{D}) &\sim \mathcal{IG}(\overline{a}_{\sigma^2}, \overline{b}_{\sigma^2}), && (15)\\
\overline{a}_{\sigma^2} &= \underline{a}_{\sigma^2} + nT/2, && (16)\\
\overline{b}_{\sigma^2} &= \underline{b}_{\sigma^2} + (\boldsymbol{S}\boldsymbol{Y} - \boldsymbol{U}\boldsymbol{\beta})'(\boldsymbol{S}\boldsymbol{Y} - \boldsymbol{U}\boldsymbol{\beta})/2. && (17)
\end{aligned}
$$

In contrast, the conditional posterior of the spatial autoregressive parameter ρ does not correspond to a closed-form distribution:

$$
p(\rho \mid \bullet, \mathcal{D}) \propto p(\rho)\,|\boldsymbol{S}|\,\exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{S}\boldsymbol{Y} - \boldsymbol{U}\boldsymbol{\beta})'(\boldsymbol{S}\boldsymbol{Y} - \boldsymbol{U}\boldsymbol{\beta})\right], \qquad (18)
$$

so alternative sampling methods must be employed.

The package implements an MCMC sampler that iteratively draws from the respective conditional posterior distributions. The sampling scheme proceeds as follows:

1) *Initialization.* A random configuration of $\boldsymbol{\Omega}$ is generated based on the specified priors, from which the initial spatial weight matrix $\boldsymbol{W}$ is constructed. Initial values for the remaining parameters are drawn from their respective prior distributions.

2) *Sequential parameter updating.* In each iteration, the following updates are performed:

   a) Sequentially update the elements of $\boldsymbol{\Omega}$ row-wise in random order.

   b) Draw $\boldsymbol{\beta}$ from its Gaussian conditional posterior.

   c) Draw σ^2 from its inverse-Gamma conditional posterior.

   d) Sample ρ using a griddy Gibbs step (Ritter and Tanner, 1992), following LeSage and Pace (2009), because its conditional posterior has no closed form. To reduce computational cost when evaluating $\log|\boldsymbol{S}|$ over a fine grid, the package uses the approximation of Barry and Pace (1999).[12] Alternatively, users may enable a Metropolis–Hastings step via `rho_priors()`.

Step 2) is repeated for B iterations after discarding the first B_0 draws as burn-in. Inference is then based on the remaining $(B - B_0)$ posterior samples.

4.1 Efficient evaluation of log-determinants

Sampling the individual elements ω_{ij} of the adjacency matrix can create substantial computational burdens in a Bayesian MCMC setting. Each update of ω_{ij} alters the spatial weight matrix $\boldsymbol{W}$ and thus the matrix $\boldsymbol{S} = \boldsymbol{I}_T \otimes (\boldsymbol{I}_n - \rho \boldsymbol{W})$, requiring repeated evaluation of the log-determinant $|\boldsymbol{S}|$. For large n, direct computation becomes infeasible, as evaluating $|\boldsymbol{S}|$ scales with $\mathcal{O}(n^3)$ and must be performed at every Gibbs iteration for each unknown entry in $\boldsymbol{\Omega}$.

To address this issue, the package follows the approach of Krisztin and Piribauer (2023), which exploits the Sherman–Morrison formula for log-determinant updates and the matrix inverse lemma within the Gibbs sampling scheme. Because a change in ω_{ij} corresponds to a rank-one modification of $\boldsymbol{W}$, this strategy enables exact and numerically inexpensive element-wise updates of $|\boldsymbol{S}|$ and $(\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1}$, largely avoiding costly direct determinant evaluations. A detailed derivation of this update mechanism is provided in Krisztin and Piribauer (2023).

5 Empirical example: Economic growth in European regions

This section provides an empirical illustration of the package using an in-built panel dataset on pan-European regional economic growth. The empirical design follows the conceptual framework of Piribauer et al. (2023), who analyse the spatial structure of regional economic growth across European NUTS-2 regions. Their study emphasises that economic growth spillovers need not follow simple geographic or distance-based decay patterns. Instead, empirically relevant spatial interactions often reflect complex economic, institutional, or socio-cultural linkages that may strongly deviate from exogenous contiguity structures.

In the spirit of this argument, the following illustration estimates the spatial interaction structure from data, rather than imposing a fixed spatial weights matrix. However, for computational tractability within a vignette setting, the in-built dataset aggregates observations to the NUTS-1 level. This reduces the cross-sectional dimension relative to

[12] As an alternative to this approximation, function `rho_priors()` allows to use a quadratic spline interpolation instead.

Piribauer et al. (2023), but still preserves the central empirical challenge of spatial growth regressions: n is moderately large relative to the available time periods T, which makes data-driven estimation of spatial networks non-trivial and calls for the use of sparsity-inducing priors as discussed in Section 3.

As a benchmark, we estimate a standard spatial autoregressive (SAR) panel growth specification of the form:[13]

$$\boldsymbol{y}_t = \rho \boldsymbol{W} \boldsymbol{y}_t + \boldsymbol{Z}_{t-1}\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t, \tag{19}$$

where $\boldsymbol{y}_t = (\boldsymbol{g}_t - \boldsymbol{g}_{t-1})$ denotes labour productivity growth between years $t-1$ and t, with $\boldsymbol{g}_t$ capturing the (log-)level of real gross value added (GVA) per worker. The matrix $\boldsymbol{Z}_{t-1}$ contains lagged initial conditions—most prominently the initial level of labour productivity—and two indicators of regional educational attainment, following the spatially-augmented growth literature (LeSage and Fischer, 2009, Crespo Cuaresma and Feldkircher, 2013, Crespo Cuaresma et al., 2014, Piribauer, 2016). For clarity of exposition, we include only a common intercept in the baseline specification.[14]

The dataset comprises annual growth rates for $n = 90$ European NUTS-1 regions for the period 2001–2019 ($T = 19$), with respective one-year lags of the explanatory variables. This structure closely mirrors that of Piribauer et al. (2023), albeit at a more aggregated level. Because n remains relatively large compared to T, the use of hierarchical sparsity priors for the spatial weight matrix is particularly advantageous, ensuring stable estimation while allowing the data to inform the strength and pattern of spatial linkages.

We begin by initialising the data objects and loading the relevant package environment. We also set a seed for reproducibility of the results:

```
R> library(estimateW)
R> set.seed(571)

R> Y <- as.matrix(nuts1growth$growth_gdp_pw)
R> Z <- cbind(1,nuts1growth$init_gdp_pw,
+         nuts1growth$loweduc,
+         nuts1growth$higheduc)
R> n <- 90
R> tt <- 19
```

Object `Y` (the stacked dependent variable $\boldsymbol{Y}$) is a $nT \times 1$ matrix of annual growth rates of GVA per worker (labor productivity) for $n = 90$ European NUTS-1 regions for

[13] Such a spatial growth regression can be derived from theoretical spatially augmented growth models (see López-Bazo et al., 2004 or Fischer, 2011).

[14] Region- or country-fixed effects may easily be added by including the respective dummies as regressors.

the period 2001-2019 ($T = 19$).[15] In addition to the dependent variable, the dataset also includes the key explanatory variables for regional growth regressions, such as the natural logarithm of initial gross value added (GVA) per worker, and the share of the working age population with low and high levels of education based on the International Standard Classification of Education (ISCED). These explanatory variables are collected in `X`, an $nT \times 4$ matrix (including the intercept, this object corresponds to the stacked matrix of explanatory variables $\boldsymbol{U}$).

Given the moderate cross-sectional dimension relative to the time dimension, $n \gg T$, we follow the hierarchical shrinkage approach described in Section 3, targeting a prior expected number of $\underline{k} = 7$ neighbours per region. We impose no prior spatial structure other than fixing the diagonal to zero, letting the data inform the adjacency pattern:

```
R> kbar=7
R> Wprior <- matrix(0.5,ncol=n,nrow=n)
R> diag(Wprior) <- 0
R> a_pr = 1; b_pr = ((n-1)-kbar)/(kbar)
R> priorW_hierarchical <- bbinompdf(
   +         0:(n-1),
   +         nsize = (n-1), a = a_pr, b = b_pr)
R> AA = W_priors(
   +         n=n,
   +         W_prior = Wprior,
   +         nr_neighbors_prior = priorW_hierarchical)
```

To keep runtime minimal for illustrative purposes, the SAR model with unknown spatial weights is estimated using a small number of MCMC draws. We therefore run the function only using 100 retained MCMC draws after discarding the first batch of 100 draws:

```
R> res_sarw = sarw(
   +     Y=Y, tt=tt, Z=Z,
   +     niter = 200, nretain = 100,
   +     W_prior = AA)
```

The object `res_sarw` is a list that inter alia contains all the retained posterior MCMC draws for the slope parameters (`postb`), their corresponding average direct, indirect (spillover) and total impacts (`post.direct`, `post.indirect`, `post.total`), spatial parameter ρ (`postr`), σ^2 (`posts`), or the elements of the spatial weight matrix $\boldsymbol{W}$ (`postw`).

The posterior means and standard deviations of the key model parameters are summarised in Table 3. A quick overview may also be obtained via the command `summary(res_sarw)`.

[15]The countries covered are: Austria, Belgium, Bulgaria, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Hungary, Ireland, Italy, Lithuania, Luxembourg, Latvia, Malta, Netherlands, Poland, Portugal, Romania, Sweden, Slovenia, and Slovakia. The NUTS-1 codes identify the first-level administrative regions within these countries.

Table 3: Results for key quantities

Variable/Impact	**Post. Mean**	**Post. SD**
intercept	0.17651	0.01090
log initial GVA per worker	-0.01692	0.00117
share low education	0.00004	0.00005
share high education	0.00044	0.00011
ρ	0.71322	0.01574
σ^2	0.00053	0.00002
av. direct initial GVA per worker	-0.01880	0.00125
av. direct share low education	0.00004	0.00006
av. direct share high education	0.00049	0.00012
av. indirect initial GVA per worker	-0.03972	0.00386
av. indirect share low education	0.00008	0.00012
av. indirect share high education	0.00104	0.00026

Notes: The table presents estimates of the slope parameters and other key quantities for the benchmark specification by reporting posterior mean and standard deviations.

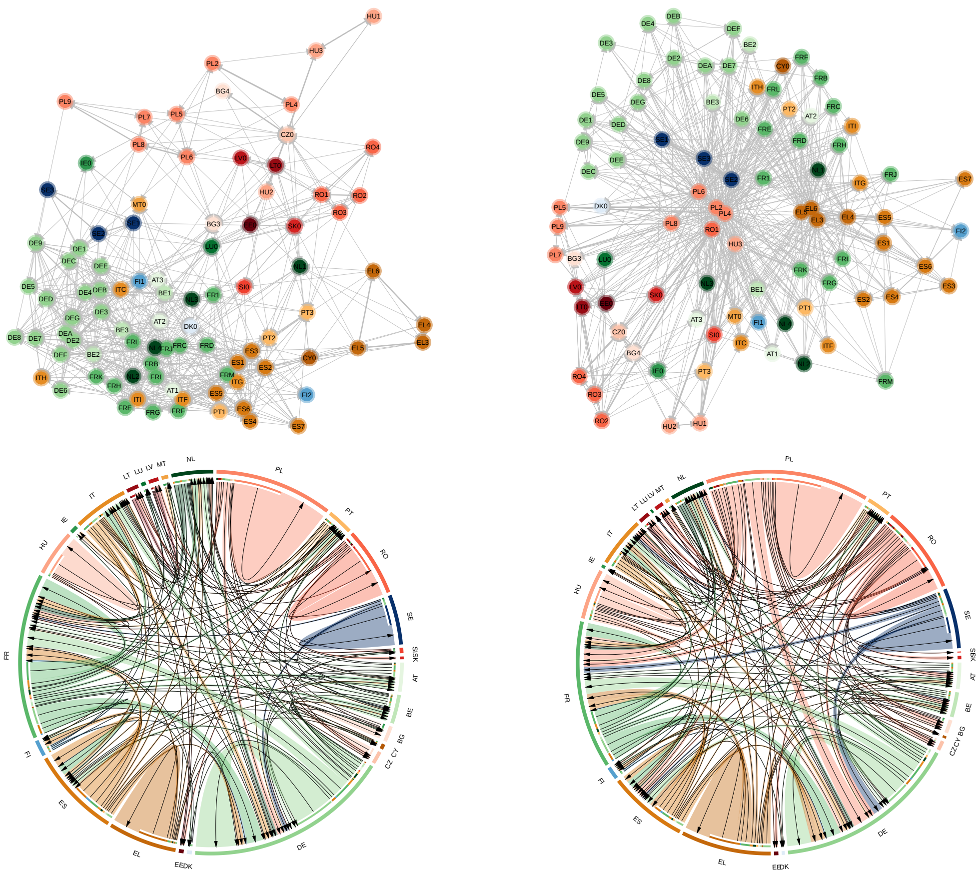
Table 3 shows a rather strong and precisely estimated degree of spatial autocorrelation ρ. The parameters corresponding to the explanatory variables show the expected signs, with a negative impact of the initial income variable (suggesting conditional convergence) and a positive impact of tertiary education attainment (share high education) on economic growth. While the table also shows a positive sign on the share of low educated in the working age population (relative to the benchmark of a medium educated population), the parameter is estimated rather imprecisely. The results presented in Table 3 are thus mainly in line with a similar specification in Piribauer et al. (2023) for the NUTS-2 level.

Note that the approach involves estimation of the $n \times n$ matrix linkages between the regions. The posterior draws of $\boldsymbol{W}$ are stored in the array `postw`. We follow Piribauer et al. (2023) and visualize the most important estimated linkages of $\boldsymbol{W}$ and the spatial multipliers $(\boldsymbol{I}_n - \rho\boldsymbol{W})^{-1}$ as a network as well as a chord diagram using the `circlize` package (Gu, 2024, Gu et al., 2014). To increase visualization, the regions (nodes) are colored according to the countries. We use the same colors as in Piribauer et al. (2023), which are based on a supranational country classification, covering covers of 26 countries of the European Union. Specifically, Southern European countries (Cyprus, Greece, Italy, Malta, Spain and Portugal) are shown in different shades of orange and brown, Northern Europe (Denmark, Finland, Sweden) in blue, Western European countries (Austria, Belgium, Germany, France, Ireland, Luxembourg and the Netherlands) in green and Central and Eastern Europe (Bulgaria, the Czech Republic, Hungary, Poland, Slovakia, Slovenia and Romania) in different tones of light red. The three Baltic countries (Estonia, Latvia and Lithuania) are rendered in dark red. The left subplots of Figure 2 show the estimated spatial weight matrix $\boldsymbol{W}$ and the right subplots the corresponding spatial spillover multipliers $(\boldsymbol{I}_n - \rho\boldsymbol{W})^{-1}$. In both cases, the most important links are presented as a

Figure 2: Estimated links of $\boldsymbol{W}$ and spatial spillovers by region and country

(a) Links of $\boldsymbol{W}$

(b) Spillovers of spatial multipliers $(\boldsymbol{I}_n - \rho\boldsymbol{W})^{-1}$



Notes: The figures depict the estimated spatial weight matrix $\boldsymbol{W}$ (left subplots a) and the implied spatial spillover multipliers (i.e. the off-diagonal elements of $(\boldsymbol{I}_n - \rho\boldsymbol{W})^{-1}$) (right subplots b). The upper subplots depict the most important links as a graph and using country colors. For better visualization, the lower subplots show the respective chord diagrams by aggregating by country.

network for regions (top subplots) and as an aggregated country-level chord plot (bottom subplots).[16] Examining these estimated links corroborates the findings in Piribauer et al. (2023) that regions belonging to the same country are more strongly interconnected. This is particularly apparent in the chord diagrams, where the most notable inflows by country typically come from the same country. Furthermore, regions belonging to the same (supranational) country group are also more strongly connected.

[16] To improve visual clarity and prevent 'hairballing', only the strongest links of the respective posterior means are shown in the graphs. Specifically, the region-by-region graphs show only the strongest 10 per cent of the posterior means and the country-aggregated chord diagrams show only the largest 20 per cent.

6 Concluding remarks

This paper has introduced the **estimateW** package for R, which provides a flexible and transparent framework for estimating spatial weight matrices within hierarchical Bayesian spatial panel models. The package enables researchers to move beyond the conventional assumption of a fixed and exogenous spatial structure by estimating the strength and pattern of spatial dependencies directly from the data. A modular prior architecture separates spatial structure from sparsity, allowing users to encode domain knowledge while retaining a high degree of data-driven flexibility.

Methodologically, **estimateW** builds on a Gibbs-sampling scheme that exploits efficient matrix determinant and inverse updates. This ensures numerically stable inference even in high-dimensional settings where the number of spatial units is large relative to the time dimension. By combining structured priors, sparsity control, and computational efficiency, the package facilitates credible Bayesian estimation of spatial network structures and associated spillover effects.

The package offers a user-friendly interface for the most common spatial econometric panel models, including the spatial Durbin model (SDM), spatial error model (SEM), spatial autoregressive model (SAR), spatial Durbin error model (SDEM), and SLX specifications, in versions that allow for either exogenous or estimated spatial weight matrices. An empirical illustration using European NUTS-1 regions demonstrates that the approach uncovers plausible spatial dependence patterns, economically meaningful parameter estimates, and intuitively interpretable clustering in the estimated network and multiplier effects.

It is worth noting that the computational burden of the estimation approach of the spatial weight matrix grows quickly with the number of cross-sectional (regional) units n. As a rule of thumb, the fully data-driven estimation of $\boldsymbol{W}$ implemented in **estimateW** is designed for applications with cross-sectional dimensions up to roughly 300 units.

References

Barry RP and Pace RK (1999) Monte Carlo estimates of the log determinant of large sparse matrices. *Linear Algebra and its Applications* 289(1-3), 41–54

Bramoullé Y, Djebbari H and Fortin B (2009) Identification of peer effects through social networks. *Journal of Econometrics* 150(1), 41–55

Crespo Cuaresma J, Doppelhofer G and Feldkircher M (2014) The determinants of economic growth in European regions. *Regional Studies* 48(1), 44–67

Crespo Cuaresma J and Feldkircher M (2013) Spatial filtering, model uncertainty and the speed of income convergence in Europe. *Journal of Applied Econometrics* 28(4), 720–741

De Paula A, Rasul I and Souza PC (2025) Identifying network ties from panel data: Theory and an application to tax competition. *Review of Economic Studies* 92(4), 2691–2729

Debarsy N, Ertur C and LeSage JP (2012) Interpreting dynamic space–time panel data models. *Statistical Methodology* 9(1), 158–171

Elhorst JP (2010) Applied spatial econometrics: Raising the bar. *Spatial Economic Analysis* 5(1), 9–28

Fischer MM (2011) A spatial Mankiw-Romer-Weil model: Theory and evidence. *The Annals of Regional Science* 47, 419–436

Gu Z (2024) *circlize: Circular Visualization*. R package version 0.4.16, https://CRAN.R-project.org/package=circlize

Gu Z, Gu L, Eils R, Schlesner M and Brors B (2014) Circlize implements and enhances circular visualization in R. *Bioinformatics* 30(19), 2811–2812

Iacopini M, Krisztin T and Piribauer P (2025) Analysing dynamic cross-price dependencies with a Markov-Switching spatial autoregressive model. *The Annals of Applied Statistics* , 1–23

Kelejian HH and Prucha IR (1998) A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics* 17(1), 99–121

Krisztin T and Piribauer P (2023) A Bayesian approach for the estimation of weight matrices in spatial autoregressive models. *Spatial Economic Analysis* 18(1), 44–63

Krisztin T and Piribauer P (2026) *estimateW: Estimation of Spatial Weight Matrices.* R package version 0.2.0, https://CRAN.R-project.org/package=estimateW

LeSage JP and Fischer MM (2009) Spatial growth regressions: Model specification, estimation and interpretation. *Spatial Economic Analysis* 3(3), 275–304

LeSage JP and Pace RK (2009) *Introduction to Spatial Econometrics.* CRC Press, Boca Raton London New York

Ley E and Steel MF (2009) On the effect of prior assumptions in Bayesian model averaging with applications to growth regression. *Journal of Applied Econometrics* 24(4), 651–674

López-Bazo E, Vayá E and Artis M (2004) Regional externalities and growth: evidence from European regions. *Journal of Regional Science* 44(1), 43–73

Pfarrhofer M and Piribauer P (2019) Flexible shrinkage in high-dimensional Bayesian spatial autoregressive models. *Spatial Statistics* 29, 109–128

Piribauer P (2016) Heterogeneity in spatial growth clusters. *Empirical Economics* 51(2), 659–680

Piribauer P, Glocker C and Krisztin T (2023) Beyond distance: The spatial relationships of European regional economic growth. *Journal of Economic Dynamics and Control* 155, 104735

Ritter C and Tanner MA (1992) Facilitating the Gibbs sampler: The Gibbs stopper and the Griddy–Gibbs sampler. *Journal of the American Statistical Association* 87(419), 861–868