
Munich Personal RePEc Archive

Some correlation properties of spatial autoregressions

Martellosio, Federico

University of Reading

September 2009

Online at https://mpra.ub.uni-muenchen.de/17254/
MPRA Paper No. 17254, posted 11 Sep 2009 09:52 UTC

Some Correlation Properties of Spatial Autoregressions

Federico Martellosio*
University of Reading
10 September, 2009

Abstract

This paper investigates how the correlations implied by a first-order simultaneous autoregressive (SAR(1)) process are affected by the weights matrix and the autocorrelation parameter. An interpretation of the covariance structure of the process is provided, based on the walks connecting the spatial units. The interpretation serves to explain a number of correlation properties of SAR(1) processes, and clarifies why in practical applications it is difficult, or even impossible, to use SAR(1) processes to impose some desired correlation properties on a given data set.

Keywords: simultaneous autoregressions; spatial autocorrelation; spatial weights matrices; walks in graphs.

JEL Classification: C21.

Mailing address: School of Economics, University of Reading, Whiteknights, Reading RG6 6AW, UK. Tel: +44 (0) 118 378 6033.

E-mail: f.martellosio@reading.ac.uk

*I am grateful to Grant Hillier, Paolo Paruolo, Peter Robinson, Tony Smith, Kees Jan van Garderen for discussions and encouragement.

1 Introduction

First-order simultaneous autoregressive (SAR(1)) processes are widely used to analyze variables that interact over some geographic, social, or economic space (e.g., Cliff and Ord, 1981; Anselin, 1988).[1] Such processes postulate a simple relationship between each of its two main components—the weights matrix $\boldsymbol{W}$ and the correlation parameter ρ—and the inverse of their variance matrix. Due to the matrix inversion, however, the way in which $\boldsymbol{W}$ and ρ affect the correlations implied by the model may be difficult to understand. Wall (2004) discusses some unexpected correlation properties of SAR(1) models defined on irregularly spaced lattices, and concludes that the models exhibit nonintuitive behavior. On the other hand, the correlation structure of SAR(1) models is well-understood when the observational units form a regular lattice (e.g., Whittle, 1954; Besag, 1972), but this is rarely the case in economics.

The main objective of this paper is to shed some light on how the correlation structure of SAR(1) processes depends on $\boldsymbol{W}$ and on ρ. We provide a simple interpretation of the covariances between two variables observed at two spatial units in terms of a particular type of walks connecting the two units. Such an interpretation clarifies why, in the case of irregularly spaced lattices, it is difficult to relate properties of $\boldsymbol{W}$ to the correlation properties of SAR(1) models. The problem is that the way different walks contribute to the covariances depends on ρ. Indeed, when $|\rho|$ is small the correlation structure is largely determined by short walks, but, as $|\rho|$ increases, the importance of longer walks increases. Because of this reason, there are a number of correlation properties that hold for sufficiently small $|\rho|$, but are not guaranteed to hold for all values of ρ in the parameter space of the models. Since ρ is unknown, and can be estimated only after $\boldsymbol{W}$ has been chosen, it follows that the user of a SAR(1) model cannot control the correlation properties by specifying $\boldsymbol{W}$.

The rest of the paper is organized as follows. SAR(1) models are presented in Section 2. Section 3 introduces some graph theoretic terminology, which is used in Section 4 to formulate the interpretation of the covariance structure of SAR(1) models. Section 5 analyzes the correlation properties of SAR(1) models. Besides focusing on properties that hold for small $|\rho|$, we discuss what happens when $|\rho|$ is large, and we study explicitly the behavior of the correlations at the extremes of the parameter space. Throughout the analysis several consequences of the common practice of row-standardizing $\boldsymbol{W}$ are pointed out. Section 6 concludes. Proofs are collected in the Appendix.

2 Specification of the SAR(1) Model

Consider a *fixed* and *finite* set of n observational units. The units are labelled by the first n integers in some arbitrary way, and the i-th unit has a random variable y_i associated with it.

[1] For empirical applications in economics, see, e.g., Case (1991), Bell and Bockstael (2000), Millimet and Collier (2008), and Calvó-Armengol et al. (2009).

Letting $\boldsymbol{y} = (y_1, ..., y_n)'$, a SAR(1) model is specified through the equation

$$\boldsymbol{y} = \rho \boldsymbol{W} \boldsymbol{y} + \boldsymbol{\varepsilon}, \tag{1}$$

where ρ is a real unknown parameter, $\boldsymbol{W}$ is an $n \times n$ weights matrix, $\boldsymbol{\varepsilon}$ is a vector of error terms such that $\mathrm{E}(\boldsymbol{\varepsilon}) = \boldsymbol{0}$ and $\mathrm{var}(\boldsymbol{\varepsilon}) = \sigma^2 \boldsymbol{V}$, with $\sigma^2 > 0$ and $\boldsymbol{V}$ a known $n \times n$ diagonal matrix with positive diagonal entries. We do not make any other assumptions on the distribution of $\boldsymbol{\varepsilon}$, but it should be kept in mind that if such a distribution is far from being Gaussian, then correlations may be poor indicators of dependence. Also, because in this paper we are interested only in correlations, we do not consider extensions of (1) to models with $\mathrm{E}(\boldsymbol{y}) \neq \boldsymbol{0}$.[2]

The matrix $\boldsymbol{W}$ is non-stochastic and known. It is chosen to reflect a priori information on relations among the observational units (see, e.g., Anselin, 1988). For example, the entries $(\boldsymbol{W})_{ij}$ may be taken to be a certain function of some physical or economic distance between the i-th and the j-th observational units. Throughout the paper we assume that, as it is virtually always the case in applications, $\boldsymbol{W}$ has zero diagonal entries and nonnegative off-diagonal entries.

Provided that ρ is different from the reciprocal of the nonzero real eigenvalues of $\boldsymbol{W}$, so that $\boldsymbol{I} - \rho \boldsymbol{W}$ is invertible, model (1) implies the positive definite variance matrix

$$\mathrm{var}(\boldsymbol{y}) = \sigma^2 (\boldsymbol{I} - \rho \boldsymbol{W})^{-1} \boldsymbol{V} (\boldsymbol{I} - \rho \boldsymbol{W}')^{-1}. \tag{2}$$

For the purpose of studying the correlation structure of SAR(1) models, there is no loss of generality in setting $\sigma^2 = 1$ and $\boldsymbol{V} = \boldsymbol{I}$.[3] Accordingly, from now on and unless otherwise specified, we take[4]

$$\mathrm{var}(\boldsymbol{y}) = \left[\left(\boldsymbol{I} - \rho \boldsymbol{W}' \right) \left(\boldsymbol{I} - \rho \boldsymbol{W} \right) \right]^{-1}. \tag{3}$$

Let $\lambda_{\max}$ denote the spectral radius (i.e., the largest modulus of the eigenvalues) of $\boldsymbol{W}$. Since $\boldsymbol{W}$ is nonnegative, $\lambda_{\max}$ is an eigenvalue of $\boldsymbol{W}$ (Horn and Johnson, 1985, Theorem 8.3.1). Also, if $\boldsymbol{W}$ has at least one (real) negative eigenvalue, let $\lambda_{\min}$ denote the smallest negative eigenvalue. Then, the largest connected interval of values of ρ around the origin where $\mathrm{var}(\boldsymbol{y})$ exists is $(\rho_{\min}, \rho_{\max})$, with[5]

$$\rho_{\min} := \begin{cases} \lambda_{\min}^{-1} & \text{if } \boldsymbol{W} \text{ has at least one negative eigenvalue} \\ -\infty & \text{otherwise,} \end{cases}$$

[2] Model (1) can be extended to accomodate a nonzero mean in two different ways, leading to the so-called spatial lag and spatial error models (e.g., Anselin, 1988). Both models have variance matrix (2).

[3] We can take $\boldsymbol{V} = \boldsymbol{I}$ because the correlation matrix of a random vector $\boldsymbol{y}$ is invariant to transformations $\boldsymbol{y} \to \boldsymbol{T}\boldsymbol{y}$, for any diagonal matrix $\boldsymbol{T}$ with positive diagonal entries (if $\boldsymbol{V} \neq \boldsymbol{I}$, just take $\boldsymbol{T} = \boldsymbol{V}^{-1/2}$).

[4] The matrix $\boldsymbol{W}$ in (3) is related to that in (2) through the transformation $\boldsymbol{W} \to \boldsymbol{V}^{-1/2} \boldsymbol{W} \boldsymbol{V}^{1/2}$. Hence, the results to be given in this paper continue to hold for SAR(1) models with $\boldsymbol{V} \neq \boldsymbol{I}$, provided that $\boldsymbol{W}$ is replaced with $\boldsymbol{V}^{-1/2} \boldsymbol{W} \boldsymbol{V}^{1/2}$. Note that the eigenvalues of $\boldsymbol{W}$ are invariant to the above (similarity) transformation.

[5] The case $\rho_{\min} = -\infty$ is relevant when $\boldsymbol{W}$ is nilpotent (Horn and Johnson, 1985, p. 139), or when all eigenvalues of $\boldsymbol{W}$ are either positive or complex. On the other hand, $\lambda_{\max} = 0$ (and hence $\rho_{\max} = \infty$) occurs if and only if $\boldsymbol{W}$ is nilpotent.

and

$$\rho_{\max} := \begin{cases} \lambda_{\max}^{-1} & \text{if } \lambda_{\max} > 0 \\ \infty & \text{if } \lambda_{\max} = 0. \end{cases}$$

One could take $(\rho_{\min}, \rho_{\max})$ or its subset $(-\rho_{\max}, \rho_{\max})$ as the set of admissible values for ρ. In Section 4 we shall see that the correlation structure of SAR(1) models admits a simple interpretation when $\rho \in (-\rho_{\max}, \rho_{\max})$.

3 Some Graph Theoretic Notions

For our analysis, it is useful to consider the graph underlying $\boldsymbol{W}$, that is, the graph having as *vertices* the n observational units, and as *arcs* the ordered pairs (i, j) such that $(\boldsymbol{W})_{ij} > 0$. For a general introduction to graph theory see, e.g., Harary (1969).

In graph theory, a (directed) *walk* from a vertex i_0 to a vertex i_r is an alternating sequence $(i_0, a_1, i_1, ..., a_r, i_r)$, $r \geq 0$, of vertices and arcs in which each arc a_t is (i_{t-1}, i_t). For the purposes of this paper, we need to modify the definition of a walk.

Definition 3.1 *A SAR-walk from i_0 to i_r is an alternating sequence $(i_0, a_1, i_1, ..., a_r, i_r)$ of vertices and arcs in which, for some $k = 0, ..., r$, the first k arcs a_t are (i_{t-1}, i_t) and the remaining $r - k$ arcs are $a_t = (i_t, i_{t-1})$.*

Note that a SAR-walk is a walk only for $k = r$. For $k < r$, the first k steps of a SAR-walk are in the direction of the sequence $(i_0, a_1, i_1, ..., a_r, i_r)$ and the remaining ones are in the opposite direction. In Section 4 we shall see that Definition 3.1 is the appropriate one to analyze SAR(1) models, because all SAR-walks (not only those that are walks) from one unit to another unit affect the correlation implied by a SAR(1) model between those units.

Next, we adapt a number of standard graph theoretic notions to our definition of a SAR-walk. We say that a graph is *connected* if there is a SAR-walk from i to j, for all $i, j = 1, ..., n$. The *length* of a SAR-walk is the number of arcs in it (r in the case of the SAR-walk in Definition 3.1). The trivial sequence (i_0) is considered a SAR-walk of length 0 from i_0 to i_0.

The definition of a SAR-walk imposes the following distance on the set of observational units.

Definition 3.2 *The distance $d(i, j)$ between any two units i and j is the length of the shortest SAR-walk from i to j, if any; otherwise, $d(i, j) = \infty$.*

It is immediately clear that: $d(i, j) = 0$ if and only if $i = j$; $d(i, j) = d(j, i)$ (which is why we speak of the distance between i and j, rather than from i to j); $d(i, j) = 1$ if and only if $(\boldsymbol{W})_{ij} + (\boldsymbol{W})_{ji} > 0$.[6]

[6] It is also easily checked that, if the graph is connected, $d(\cdot, \cdot)$ satisfies the triangle inequality, and hence is a metric.

SAR-walks and the induced distance $d(\cdot,\cdot)$ are determined by which entries of $\boldsymbol{W}$ are zero and which are nonzero. It is also convenient to have a measure of the importance of different SAR-walks, based on the magnitude of the nonzero entries of $\boldsymbol{W}$. To each arc (i,j) we assign the weight $(\boldsymbol{W})_{ij}$. Then, based on a standard graph theoretic notion (e.g., Godsil, 1993, p. 56), we define the weight of a SAR-walk as follows.

Definition 3.3 *The weight of a SAR-walk is the product of the weights of its arcs.*

The weight of a SAR-walk of length 0, which does not have any arcs in it, is taken to be 1.

We shall also use the following terminology. If $(\boldsymbol{W})_{ij} > 0$, j is said to be a *neighbor* of i, which we indicate in symbols by $i \to j$. When j is a neighbor of i, and i of j, we say that i and j are neighbors. A graph such that $(\boldsymbol{W})_{ij} = 0$ if and only if $(\boldsymbol{W})_{ji} = 0$, for all $i,j = 1,...,n$, is said to be *undirected*; otherwise, it is said to be *directed*.

The following example illustrates the above graph theoretic notions.

Example 3.4 Consider a spatial configuration consisting of a central unit (e.g., a big city) surrounded by 8 units on a circle (e.g., neighboring smaller cities). For such a configuration, we consider the two graphs G_u and G_d displayed in Figure 1. The indexes u and d on G stand for undirected and directed, respectively. A line with an arrow represents an arc, and a line joining two vertices i and j without arrows indicates that there is both an arc from i to j and an arc from j to i. The weight matrix corresponding to G_u is

$$\boldsymbol{W}_u = \begin{bmatrix} \cdot & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & 1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & 1 & \cdot & 1 & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & 1 & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & 1 \\ 1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \end{bmatrix}, \tag{4}$$

where the dots stand for zeros. The weights matrix $\boldsymbol{W}_d$ corresponding to G_d is obtained from $\boldsymbol{W}_u$ by replacing all the ones in the first column with zeros. Both $\boldsymbol{W}_u$ and $\boldsymbol{W}_d$ specify that the central unit directly influences the surrounding units. The choice between $\boldsymbol{W}_u$ and $\boldsymbol{W}_d$ should be based on whether the surrounding units are believed to directly influence the central unit; a lack of influence from the surrounding units to the central unit could be due, for instance, to the absence of knowledge spillover in that direction, or to transportation costs being too high in that direction. To illustrate the concepts of SAR-walks and distance, let us focus on the pairs of units $(1,2)$ and $(2,6)$. On both G_u and G_d, $d(1,2) = 1$. However, while on G_u there are two SAR-walks of length 1 from 1 to 2 ($1 \to 2$ and $2 \leftarrow 1$), on G_d there is only one ($1 \to 2$). As for the units 2 and 6, they are at different distance in the two

graphs. Namely, $d(2,6) = 2$ on G_u ($2 \to 1 \to 6$ is a shortest SAR-walk from 2 to 6), whereas $d(2,6) = 4$ on G_d ($2 \to 3 \to 4 \to 5 \to 6$ is a shortest SAR-walk from 2 to 6).

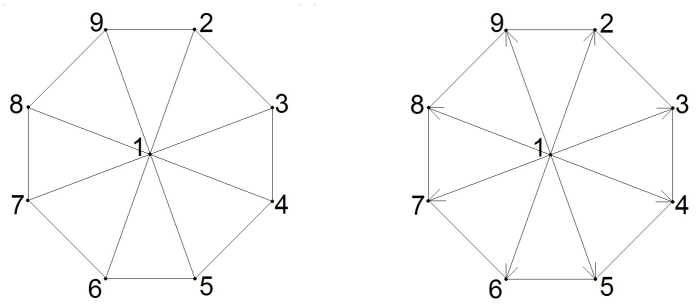


Figure 1: The graphs G_u (left) and G_d (right)

The next example serves to clarify the difference between walks and SAR-walks.

Example 3.5 Figure 2 displays the graph underlying the weights matrix

$$\boldsymbol{W} = \begin{bmatrix} \cdot & 1 & \cdot & \cdot & \cdot \\ 1 & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & 1 \\ \cdot & \cdot & \cdot & 1 & \cdot \end{bmatrix}. \tag{5}$$

Observe that there are no walks from unit 2 to unit 4. However, there are SAR-walks joining 2 and 4: for instance $2 \to 3 \leftarrow 4$ is a SAR-walk of length 2 from 2 to 4. For a plot of $\text{corr}(y_2, y_4)$ implied by a SAR(1) model with weights matrix (5) see Figure 6. It will become clear in the next section that in order for two variables y_i and y_j to be correlated it is not necessary that i and j are joined by a walk, but that they are joined by a SAR-walk.

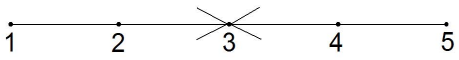


Figure 2: The graph underlying the weights matrix (5).

4 Interpretation of the Covariance Structure of a SAR(1) Model

The graph theoretic notions introduced in the previous section allow us to formulate an interpretation of SAR(1) models. Let $\text{len}(\omega)$ and $\text{wei}(\omega)$ denote, respectively, the length and the weight of a SAR-walk ω, and let $\mathcal{K}_{ij}$ be the set of all SAR-walks from i to j. Our interpretation rests on the following result.

Theorem 4.1 *In a SAR(1) model, when* $|\rho|\,\lambda_{\max} < 1$, *and for any* $i, j = 1, ..., n$,

$$\operatorname{cov}(y_i, y_j) = \sum_{\omega \in \mathcal{K}_{ij}} \operatorname{wei}(\omega)\rho^{\operatorname{len}(\omega)}. \tag{6}$$

Theorem 4.1 relates the covariance structure of a SAR(1) model to the walk structure of the graph underlying $\boldsymbol{W}$. More specifically, it asserts that $\operatorname{cov}(y_i, y_j)$ can be interpreted as the sum of the contributions of all SAR-walks from i to j, the contribution of a SAR-walk ω being $\operatorname{wei}(\omega)\rho^{\operatorname{len}(\omega)}$. Such an interpretation holds for $|\rho|\,\lambda_{\max} < 1$, or, equivalently, for $\rho \in (-\rho_{\max}, \rho_{\max})$.

Example 4.2 Consider the very simple weights matrix

$$\boldsymbol{W} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{7}$$

The underlying graph consists of the vertices 1 and 2, with an arc from 1 to 2. There is only one SAR-walk from 1 to 2, with length and weight both equal to 1. Hence, by Theorem 4.1, $\operatorname{cov}(y_1, y_2) = \rho$, for any ρ ($\lambda_{\max} = 0$ for matrix (7)). Similarly, Theorem 4.1 yields $\operatorname{var}(y_1) = 1 + \rho^2$ and $\operatorname{var}(y_2) = 1$, where the two terms equal to 1 are due to the two walks of length 0, and the term ρ^2 is due to the walk $1 \rightarrow 2 \leftarrow 1$. An alternative to the weights matrix (7) is

$$\boldsymbol{W} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{8}$$

which indicates symmetric interaction between units 1 and 2, and has $\lambda_{\max} = 1$. The extra arc from 2 to 1 implies that there are SAR-walks of any odd length contributing to $\operatorname{cov}(y_1, y_2)$. More specifically, for any odd r, there are $r + 1$ SAR-walks of length r and weight 1 from 1 to 2 (for $r = 1$, they are: $1 \rightarrow 2$ and $1 \leftarrow 2$; for $r = itive3$, they are: $1 \rightarrow 2 \rightarrow 1 \rightarrow 2$, $1 \rightarrow 2 \rightarrow 1 \leftarrow 2$, $1 \rightarrow 2 \leftarrow 1 \leftarrow 2$, and $1 \leftarrow 2 \leftarrow 1 \leftarrow 2$; and so on). By Theorem 4.1, for $|\rho| < 1$, we can then write $\operatorname{cov}(y_1, y_2) = \sum_{r=0;r\ odd}^{\infty}(r+1)\rho^r$, which is easily seen to converge to $2\rho(1-\rho^2)^{-2}$.

We now briefly discuss some implications of Theorem 4.1. According to representation (6), *all* SAR-walks between two units i and j contribute to $\operatorname{cov}(y_i, y_j)$. Thus, a SAR(1) model implicitly assumes that one unit can influence simultaneously many other units, and that, unless the graph is directed, one unit can at the same time influence and be influenced by another unit. In particular, SAR-walks that contain repetitions of arcs also contribute to the covariances. This may generate a complicated influence process. In economic applications, the influence between units may be due, for example, to neighborhood spillover effects, strategic interaction, spatial externalities, or social interaction. Equation (6) represents the equilibrium of the influence process.

Theorem 4.1 clarifies the role of ρ in shaping the correlation structure of a SAR(1) model. Since the contribution of a SAR-walk ω depends on ρ through $\rho^{\operatorname{len}(\omega)}$, it follows that: (i)

when $|\rho|$ is small, the covariance structure of a SAR(1) model is mostly determined by short SAR-walks; (ii) as $|\rho|$ increases, longer walks yield increasingly more important contributions, relative to shorter walks. As we shall see below, the fact that ρ controls the relative importance of SAR-walks of different lengths implies that some properties of SAR(1) models that hold for sufficiently small $|\rho|$ do not necessarily hold for larger $|\rho|$.

According to representation (6), the contribution of a SAR-walk to the covariance structure of a SAR(1) model depends not only on its length but also on its weight. More specifically, the form $\text{wei}(\omega)\rho^{\text{len}(\omega)}$ of the contributions implies that: (i) for a fixed length, SAR-walks with bigger weight give a larger contribution to $\text{cov}(y_i, y_j)$; (ii) for a fixed weight, and for $|\rho| < 1$, the contribution of a SAR-walk is decreasing with its length.[7]

Another implication of Theorem 4.1 regards the sign of the correlations. When $\rho \in (0, \rho_{\max})$, $\text{corr}(y_i, y_j) > 0$ provided that there is at least one SAR-walk from i to j, and $\text{corr}(y_i, y_j) = 0$ otherwise. Conversely, when $\rho \in (-\rho_{\max}, 0)$ the sign of $\text{corr}(y_i, y_j)$, $i \neq j$, may depend on ρ. To see why this is the case, first observe that, since the lowest power of ρ in (6) must be $d(i, j)$ (by the definition of $d(\cdot, \cdot)$), it follows that, as $\rho \uparrow 0$, $\text{corr}(y_i, y_j) > 0$ if $d(i, j)$ is even, $\text{corr}(y_i, y_j) < 0$ if $d(i, j)$ is odd. For any fixed $\rho \in (-\rho_{\max}, 0)$, however, SAR-walks of even length yield a positive contributions whereas SAR-walks of odd length yield a negative contributions. Hence, whether $\text{corr}(y_i, y_j)$ is positive or negative depends on whether the total contribution of the SAR-walks of even length is larger or smaller than (minus) the total contribution of the SAR-walks of odd length. Which of the two total contributions is larger may depend on ρ, because, as we have seen above, the relative importance of SAR-walks of different lengths depends on ρ.

Representation (6) provides an explanation of how correlations are formed in SAR(1) processes. Such an explanation contrasts, at least to some extent, with the conclusion in Wall (2004) that the correlations implied by a SAR(1) model do not follow an intuitive scheme. Indeed, all properties that are deemed to be counterintuitive in Wall (2004) can be explained using representation (6). Nevertheless, we do agree with Wall (2004) that it is impossible to understand the correlation structure of a SAR(1) model just by looking at $\boldsymbol{W}$. This point will be discussed in detail in the next section.

5 The Correlations

In this section we investigate how the correlation properties of SAR(1) models depend on $\boldsymbol{W}$ and on ρ. Section 5.1 gives an expansion of the correlations, which is used in Section 5.2 to analyze the effect of row-standardizing $\boldsymbol{W}$, and in Section 5.3 to study the dependence of the correlations on distance. Section 5.4 is concerned with the behavior of the correlations at the extremes of the parameter space.

[7] As long as $\lambda_{\max} > 0$, the inequality $|\rho|\,\lambda_{\max} < 1$ (required by Theorem 4.1) can be rewritten as $|\rho| < 1$ without any loss of generality, because $\lambda_{\max}$ can be taken to be 1 (by reparametrizing the model so that the weights matrix is $\lambda_{\max}^{-1}\boldsymbol{W}$). On the other hand, if $\lambda_{\max} = 0$ (see footnote 5), then the inequality $|\rho| < 1$ identifies a subset of the set of values where Theorem 4.1 holds.

5.1 An Expansion of the Correlations

For $r = 0, ..., \infty$, let $\boldsymbol{Z}_r := \sum_{k=0}^{r} \boldsymbol{W}^k (\boldsymbol{W}')^{r-k}$, where, as usual, $\boldsymbol{W}^0 = \boldsymbol{I}$. The entries of the matrices $\boldsymbol{Z}_r$ admit the following interpretation.

Lemma 5.1 *$(\boldsymbol{Z}_r)_{ij}$ equals the total weight of the SAR-walks of length r from i to j.*

The matrices $\boldsymbol{Z}_r$ appear in an expansion of $\text{corr}(y_i, y_j)$ in powers of ρ. The next result gives the two terms that lead the expansion as $\rho \to 0$. For two functions $f(x)$ and $g(x)$, we write $f(x) = O(g(x))$ to indicate that $|f(x)/g(x)|$ is bounded from above in some neighborhood of $x = 0$.

Theorem 5.2 *In a SAR(1) model,*

$$\text{corr}(y_i, y_j) = (\boldsymbol{Z}_{d(i,j)})_{ij} \rho^{d(i,j)} + (\boldsymbol{Z}_{d(i,j)+1})_{ij} \rho^{d(i,j)+1} + O(\rho^{d(i,j)+2}), \tag{9}$$

for any $i, j = 1, ..., n$.

Theorem 5.2 is not aimed at approximating $\text{corr}(y_i, y_j)$. Indeed, the correlations can be either computed exactly (if n is not too large) or they can be approximated efficiently by, e.g., numerical simulation. We shall see in the rest of the paper that Theorem 5.2 is helpful to study some correlation properties of SAR(1) models.

Expression (9) specifies the low-order derivatives with respect to ρ of the correlations at $\rho = 0$.[8] Namely, it says that $\frac{\mathrm{d}^r}{\mathrm{d}\rho^r} \text{corr}(y_i, y_j)|_{\rho=0} = r!(\boldsymbol{Z}_r)_{ij}$, for any $r = 1, ..., d(i, j) + 1$. In particular, it is worth emphasizing that

$$\frac{\mathrm{d}}{\mathrm{d}\rho} \text{corr}(y_i, y_j)|_{\rho=0} = (\boldsymbol{Z}_1)_{ij} = (\boldsymbol{W})_{ij} + (\boldsymbol{W})_{ji}, \tag{10}$$

that is, the first derivative of $\text{corr}(y_i, y_j)$ at $\rho = 0$ is completely determined by the entries $(\boldsymbol{W})_{ij}$ and $(\boldsymbol{W})_{ji}$. Eq. (10) shows that $\text{corr}(y_i, y_j)$ starts from the origin with zero derivative if and only if i is not a neighbor of j and j is not a neighbor of i.

Example 5.3 Consider the graphs G_u and G_d of Figure 1, and suppose that the corresponding weights matrices have been row-standardized (hence $\lambda_{\max} = 1$). Figure 3(b) compares $\text{corr}(y_1, y_2)$ on G_u and G_d, for $\rho \in (0, 1)$. Here, as in the other figures of this paper, correlations have been obtained by symbolic inversion of $(\boldsymbol{I} - \rho\boldsymbol{W}')(\boldsymbol{I} - \rho\boldsymbol{W})$. The quantity $(\boldsymbol{W})_{ij} + (\boldsymbol{W})_{ji}$ equals $1/8$ for the pair $(1, 2)$ (as the only SAR-walk of length 1 from 1 to 2 is $1 \to 2$, with weight $1/8$). Due to the extra arc $2 \to 1$, $(\boldsymbol{W})_{ij} + (\boldsymbol{W})_{ji}$ is larger on G_u, equal to $1/8 + 1/3 = 11/24$ (besides the SAR-walk $1 \to 2$, which has weight $1/8$, on G_u there is also $1 \leftarrow 2$, which has weight $1/3$). Accordingly, the first derivative of $\text{corr}(y_1, y_2)$ at $\rho = 0$ is larger on G_u than on G_d.

[8] This is because (9) corresponds to a full power series expansion, uniformly convergent in a neighborhood of $\rho = 0$. The full expansion is not needed explicitly and is confined to the proof of Theorem 5.2.

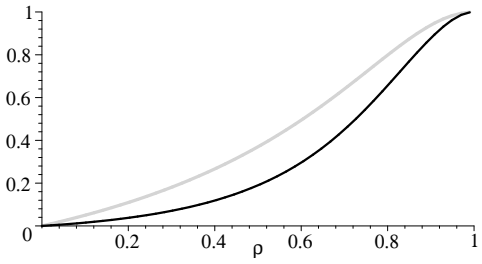


Figure 3: $\text{corr}(y_1, y_2)$ implied by a SAR(1) model on G_d (dark line) and on G_u (light line), when the weights matrices are row-standardized, and for $\rho \in (0, 1)$.

Example 5.3 is concerned with the application of expression (10) to two different weights matrices. Let us now consider a fixed $\boldsymbol{W}$. Expression (10) says that the magnitude of $(\boldsymbol{W})_{ij} + (\boldsymbol{W})_{ji}$ provides a ranking of all pairs (i, j) such that $(\boldsymbol{W})_{ij} + (\boldsymbol{W})_{ji} > 0$ (i.e., i is a neighbor of j or j is a neighbor of i) in terms of their degree of correlation as $\rho \to 0$. It is important to stress that, for an arbitrary $\boldsymbol{W}$, the ranking of the correlations established by $(\boldsymbol{W})_{ij} + (\boldsymbol{W})_{ji}$ is guaranteed to hold only as $\rho \to 0$. Indeed, setting $(\boldsymbol{W})_{ij} + (\boldsymbol{W})_{ji} > (\boldsymbol{W})_{lm} + (\boldsymbol{W})_{ml}$ for some two pairs of units (i, j) and (l, m) does not necessarily imply that $|\text{corr}(y_i, y_j)| > |\text{corr}(y_l, y_m)|$ for some fixed value of ρ (for example, see Figure 4 below, where it is clear that neighbor correlations can intersect, or Figure 5). What is more, the width of the interval of values of ρ around 0 where $(\boldsymbol{W})_{ij} + (\boldsymbol{W})_{ji} > (\boldsymbol{W})_{lm} + (\boldsymbol{W})_{ml}$ does imply $|\text{corr}(y_i, y_j)| > |\text{corr}(y_l, y_m)|$ depends in a complicated way on i, j, l, m, and $\boldsymbol{W}$. The practical implication of such observations is that the user of a SAR(1) model is unable to control the ranking of pairs of variables according to their correlations by specifying $\boldsymbol{W}$, because such a ranking may depend on the unknown parameter ρ.

5.2 The Effect of Row-Standardization

In economic applications of SAR(1) models, it is common practice to standardize $\boldsymbol{W}$ so that all its row sums are 1. As recently pointed out by Kelejian and Prucha (2009), some consequences of this practice are not completely clear. In this section we study the effect of row-standardization on the correlation structure of SAR(1) models.[9] For simplicity, we focus on the correlations between units at distance 1, and on the case when the original weights matrix is a $(0, 1)$ matrix (that is, a matrix containing only zeros and ones) and is symmetric. Extensions of the results in this section can be obtained easily for other specifications of $\boldsymbol{W}$.

The following corollary of Theorem 5.2 requires some new notation. We define $N_i := \{j : (\boldsymbol{W})_{ij} > 0\}$, $n_i := |N_i|$, $n_{ij} := |N_i \cap N_j|$, and $s_{ij} := \sum_{l \in N_i \cap N_j} n_l^{-1}$. In words, N_i is the set of neighbors of i, n_i is the number of neighbors of i, n_{ij} is the number of common neighbors of i and j, and s_{ij} is the sum of n_l^{-1} over the common neighbors l of i and j.

[9] Note that row-standardization is achievable only if $\boldsymbol{W}$ does not have any zero rows, that is, every unit has at least one neighbor. In the case of an undirected graph, this condition simply rules out isolated units.

Corollary 5.4 *Consider a SAR(1) model with weights matrix* $\boldsymbol{W}$*, and let* i *and* j *be any two units such that* $d(i,j) = 1$*.*

(*a*) *If* $\boldsymbol{W}$ *is a symmetric* $(0,1)$ *matrix, then*

$$\text{corr}(y_i, y_j) = 2\rho + 3n_{ij}\rho^2 + O(\rho^3); \tag{11}$$

(*b*) *if* $\boldsymbol{W}$ *is a row-standardized version of a symmetric* $(0,1)$ *matrix, then*

$$\text{corr}(y_i, y_j) = \left(\frac{1}{n_i} + \frac{1}{n_j}\right)\rho + \left[\frac{n_{ij}}{n_i n_j} + \left(\frac{1}{n_i} + \frac{1}{n_j}\right)s_{ij}\right]\rho^2 + O(\rho^3). \tag{12}$$

Despite their approximate nature, expressions (11) and (12) are helpful to understand some consequences of row-standardizing $(0,1)$ matrices. Two such consequences are described next, and are then illustrated by means of a representative example.

Firstly, observe that the leading term in (11) is the same for each pair of neighbors, whereas that in (12) can be very different for different pairs of neighbors. Thus, when ρ is small (more precisely, up to order $O(\rho^2)$), different pairs of neighbors tend to be much more similarly correlated when $\boldsymbol{W}$ is a symmetric $(0,1)$ matrix than when it is row-standardized.

A second consequence of row-standardization regards the approximate ranking of neighbors correlations implied by (11) and (12). In typical applications, the number n_i of neighbors is large for vertices i in the central part of a graph, and small for vertices close to the borders of the graph. Accordingly, the number n_{ij} of common neighbors is generally large for pairs of neighbors (i,j) in the central part of a graph, and small for pairs of neighbors close to the borders of the graph, whereas the reverse holds for the quantity $n_i^{-1} + n_j^{-1}$. Thus, by expression (11), when $\boldsymbol{W}$ is a symmetric $(0,1)$ matrix and up to order $O(\rho^3)$, neighbors correlations are larger in the central part of the graph than at the borders of the graph. After row-standardization, the situation is reversed: according to expression (12), neighbors correlations are larger at the borders of the graph than in the central part, up to order $O(\rho^2)$. The intervals of values of ρ around 0 where such implications hold depend on $\boldsymbol{W}$ in a complicated manner, but it is clear that the inequality $|\text{corr}(y_i, y_j)| \geq |\text{corr}(y_l, y_m)|$ is satisfied over a large interval if $n_{ij} - n_{lm}$ is large (for a $(0,1)$ weights matrix) or if $n_i^{-1} + n_j^{-1} - (n_l^{-1} + n_m^{-1})$ is large (for a row-standardized weights matrix).

Example 5.5 Consider a random vector $\boldsymbol{y}$ observed over the map of the 48 continental US states. Suppose that two states are taken to be neighbors if and only if they share a common boundary or a common corner. Figure 4 displays all correlations implied by a SAR(1) model between pairs of variables y_i and y_j such that the states i and j are neighbors. The correlations are plotted for $\rho \in [0, \lambda_{\max}^{-1})$, $\lambda_{\max}^{-1}$ being about 0.185 for a $(0,1)$ $\boldsymbol{W}$ (panel (b)), 1 for a row-standardized $\boldsymbol{W}$ (panel (c)). Out of the 107 correlations, we have emphasized those between Missouri and Tennessee (crosses) and Maine and New Hampshire (dark solid line). The number n_{ij} of common neighbors is 2 for Missouri and Tennessee, 0 for Maine and New Hampshire. Thus, by Corollary 5.4(a), Missouri and Tennessee are more correlated than

Maine and New Hampshire, when $\boldsymbol{W}$ is a $(0,1)$ matrix and for sufficiently small ρ. Indeed, panel (b) shows that this remains true for all $\rho \in (0, \lambda_{\max}^{-1})$. A similar observation applies to the case of a row-standardized $\boldsymbol{W}$. Observe that $n_i^{-1} + n_j^{-1}$ is larger for Maine and New Hampshire ($n_i^{-1} + n_j^{-1} = 4/3$) than for Missouri and Tennessee ($n_i^{-1} + n_j^{-1} = 1/4$). Corollary 5.4(b) then implies that, for sufficiently small ρ, Maine and New Hampshire are more correlated than Missouri and Tennessee; in fact, this remains true unless ρ is very close to 1.

Following the usual rules for division of power series, from Corollary 5.4 it is possible to derive expansions for ratios of neighbors correlations. This may be useful to quickly compare the degrees of correlation of different pairs of neighbors. For instance, in the case of a symmetric $(0,1)$ weights matrix, expression (11) yields

$$\frac{\operatorname{corr}(y_i, y_j)}{\operatorname{corr}(y_l, y_m)} = 1 + \frac{3}{2}\,(n_{ij} - n_{lm})\,\rho + O(\rho^2). \tag{13}$$

In the case of Example 5.5, application of (13) yields that Missouri and Tennessee ($n_{ij} = 2$) are, up to order $O(\rho^2)$, $1 + 3\rho$ times more correlated than Maine and New Hampshire ($n_{lm} = 0$).

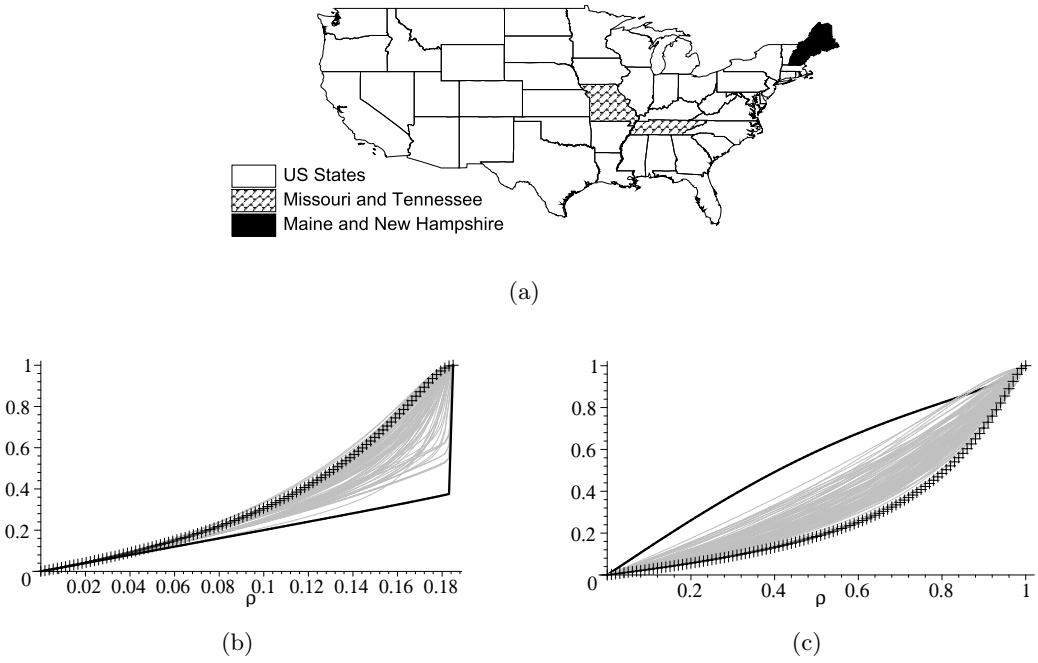


Figure 4: The correlations, as a function of ρ, implied by a SAR(1) model with $(0,1)$ weights matrix (panel (b)) and row-standardized weights matrix (panel (c)), for the 107 pairs of contiguous US states, with emphasis on Missouri and Tennessee (crosses) and Maine and New Hampshire (dark solid line).

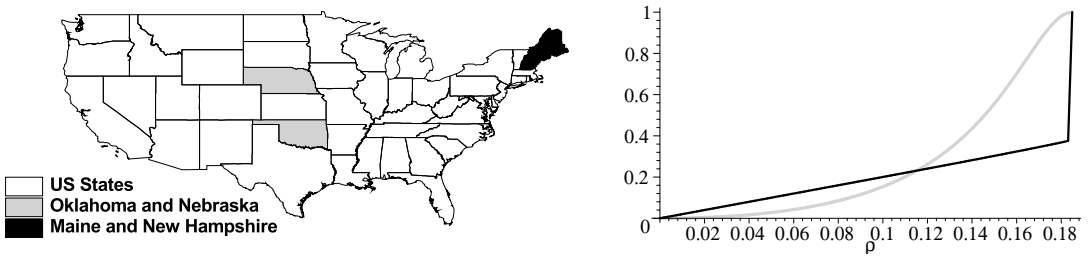
5.3 Correlations as the Distance Changes

We now turn to analyze the dependence of the correlations implied by a SAR(1) model on the graph distance $d(\cdot,\cdot)$. The following corollary of Theorem 5.2 establishes that, when $|\rho|$ is sufficiently small, the absolute value of $\operatorname{corr}(y_i, y_j)$ is inversely related to $d(i,j)$ for any $\boldsymbol{W}$.

Corollary 5.6 *As $\rho \to 0$ in a SAR(1) model, $|\text{corr}(y_i, y_j)| > |\text{corr}(y_l, y_m)|$ if $d(i,j) < d(l,m)$, for any $i, j, l, m = 1, ..., n$.*

Similarly to other correlation properties studied above, the ordering established by Corollary 5.6 does not need to hold over the whole parameter space of a SAR(1) model, and the interval where it holds depends in a complicated way on $\boldsymbol{W}$ and on i, j, l, m. An example is given next.

Example 5.7 Figure 5 displays $\text{corr}(y_i, y_j)$ when $\boldsymbol{W}$ is a $(0,1)$ matrix and i and j are Maine and New Hampshire (darker line; $d(i,j) = 1$) and Oklahoma and Nebraska (lighter line; $d(i,j) = 2$), for $0 < \rho < \lambda_{\max}^{-1}$. For large values of ρ the correlation between the units at distance 2 is much larger than the correlation between the units at distance 1.



(a) (b)

Figure 5: The correlations, as a function of ρ, implied by a SAR(1) model with $(0,1)$ weights matrix, for a pair of neighbors (darker line) and for a pair of non-neighbors (lighter line).

Example 5.7 shows that, for large enough ρ, the correlations implied by a SAR(1) model are not guaranteed to be inversely related to the distance $d(\cdot,\cdot)$. This is a consequence of the fact, pointed out in Section 4, that the importance of the contributions to the covariance structure of long SAR-walks increases with ρ. In some applications, one may wish to impose that units that are close together according to $d(\cdot,\cdot)$ cannot be much less correlated than units that are further away from each other. Recall that the contribution of a SAR-walk depends not only on its length, but also on its weight. It follows that to counterbalance the effect of an increasing ρ (longer walks become more important), one can adopt a specification of $\boldsymbol{W}$ such that the weights of SAR-walks decay quickly with the length. This is precisely what happens with a row-standardized $\boldsymbol{W}$, since in that case the weight of a SAR-walk ω is inversely related to n_i, for each $i \in \omega$. Thus, row-standardization of $\boldsymbol{W}$ helps to attenuate the possible nonmonotonicity of $\text{corr}(y_i, y_j)$ in $d(i,j)$. Indeed, when the weights matrix of Example 5.7 is row-standardized, it is never the case that a correlation between non-neighbors is much larger than a correlation between neighbors (contrary to the case of Figure 5).

5.4 Correlations at the Extremes of the Parameter Space

So far, we have studied properties of SAR(1) models that are guaranteed to hold for small $|\rho|$, but not necessarily for larger $|\rho|$. This section analyzes the behavior of the correlations as ρ approaches the extremes of the largest connected interval around the origin where $\text{var}(\boldsymbol{y})$ exists (see Section 2). For this purpose, we assume that $\lambda_{\max} > 0$ and that $\boldsymbol{W}$ has at least one negative eigenvalue, so that such extremes are $\lambda_{\min}^{-1}$ and $\lambda_{\max}^{-1}$.[10] Some numerical investigation of the behavior of the correlations close to the extremes of the parameter space can be found in Kelejian and Robinson (1995).

It is natural to wonder whether two variables y_i and y_j achieve perfect correlation as ρ approaches $\lambda_{\min}^{-1}$ and $\lambda_{\max}^{-1}$. To answer this question, let us denote by $g_{\min}$ (resp. $g_{\max}$) the geometric multiplicity of $\lambda_{\min}$ (resp. $\lambda_{\max}$),[11] and by $\boldsymbol{q}_{\min}$ (resp. $\boldsymbol{q}_{\max}$) an eigenvector of $\boldsymbol{W}$ associated to $\lambda_{\min}$ (resp. $\lambda_{\max}$).

Theorem 5.8 *In a SAR(1) model, for any $\boldsymbol{W}$ and for any $i, j = 1, ..., n$,*

(a) *as $\rho \to \lambda_{\min}^{-1}$,*

$$\text{corr}(y_i, y_j) \to \begin{cases} +1 & \text{if } g_{\min} = 1 \text{ and } (\boldsymbol{q}_{\min})_i (\boldsymbol{q}_{\min})_j > 0 \\ -1 & \text{if } g_{\min} = 1 \text{ and } (\boldsymbol{q}_{\min})_i (\boldsymbol{q}_{\min})_j < 0; \end{cases}$$

(b) *as $\rho \to \lambda_{\max}^{-1}$, $\text{corr}(y_i, y_j) \to +1$ if $g_{\max} = 1$ and $(\boldsymbol{q}_{\max})_i (\boldsymbol{q}_{\max})_j \neq 0$.*

Theorem 5.8 specifies the cases when the behavior of the correlations at the extremes of the parameter space does not depend on $\boldsymbol{W}$. In all other cases, the limiting behavior depends on $\boldsymbol{W}$.[12]

In the rest of this section we focus on the right extreme $\lambda_{\max}^{-1}$, because positive autocorrelation is much more common in applications than negative autocorrelation. In addition, a SAR(1) model with ρ close to $\lambda_{\max}^{-1}$ has an intrinsic interest due to the analogy with the near unit root case in an AR(1) model; see, e.g., Fingleton (1999); Lee and Yu (2008). The following corollary gives a sufficient condition for all pairs of variables to achieve perfect correlation as $\rho \to \lambda_{\max}^{-1}$. The condition is given in terms of walks (not SAR-walks).

Corollary 5.9 *If there is a walk from each unit to every other unit, then all correlations implied by a SAR(1) model tend to 1 as $\rho \to \lambda_{\max}^{-1}$.*

[10] When $\lambda_{\max} = 0$, the right extreme of the parameter space is ∞ (see Section 2). Similarly, when $\boldsymbol{W}$ does not have any negative eigenvalues, the left extreme is $-\infty$. Clearly, the behavior of the correlations as $\rho \to \pm\infty$ is not very interesting from a practical point of view.

[11] In most applications to irregular lattices, $g_{\min}$ and $g_{\max}$ are 1, since values of $g_{\min}$ and $g_{\max}$ larger than 1 generally require $\boldsymbol{W}$ to satisfy some symmetries (e.g., Biggs, 1993, Ch. 15). One example of a weights matrix that statisfies several symmetries and is sometimes used in social network analysis is a block diagonal matrix having r blocks equal to the matrix $\boldsymbol{E}_n$ defined in Example 5.11. For such a matrix, $g_{\min} = n - r$ and $g_{\max} = r$.

[12] Namely, when $(\boldsymbol{q}_{\min})_i (\boldsymbol{q}_{\min})_j = 0$ or $g_{\min} > 1$, $\text{corr}(y_i, y_j)$ can approach any value in $[-1, 1]$ as $\rho \to \lambda_{\min}^{-1}$, depending on $\boldsymbol{W}$. Similarly, when $(\boldsymbol{q}_{\max})_i (\boldsymbol{q}_{\max})_j = 0$ or $g_{\max} > 1$, $\text{corr}(y_i, y_j)$ can approach any value in $[0, 1]$ as $\rho \to \lambda_{\max}^{-1}$, depending on $\boldsymbol{W}$.

Clearly, the condition in Corollary 5.9 is satisfied as long as there is no ordering of the units such that $\boldsymbol{W}$ is block-triangular. An example of a SAR(1) model with block-triangular $\boldsymbol{W}$ is given next.

Example 5.10 Figure 6 displays $\text{corr}(y_2, y_4)$ implied by a SAR(1) model with the (block-triangular) weights matrix of Example 3.5. The condition in Corollary 5.9 is not satisfied, because, for instance, there are no walks from 2 to 4 (although there are SAR-walks from 2 to 4). Observe that $\text{corr}(y_2, y_4)$ does not approach 1 as $\rho \to \lambda_{\max}^{-1} = 1$.

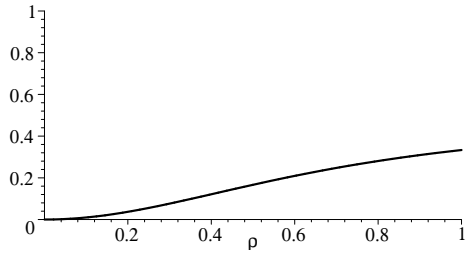


Figure 6: $\text{corr}(y_2, y_4)$ implied by a SAR(1) model with weights matrix (5).

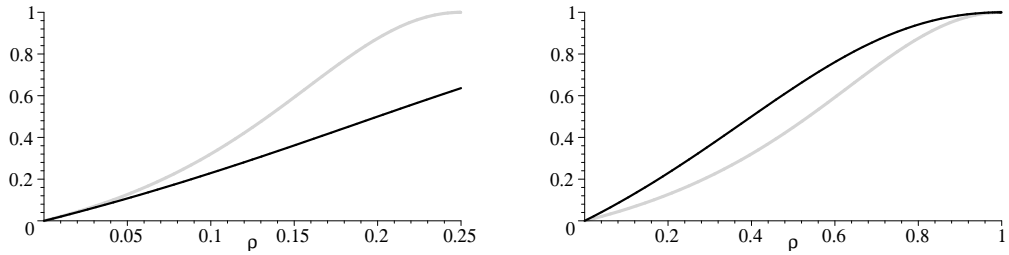
Note that on an undirected graph (i.e., when the neighborhood relation is symmetric), a block-triangular $\boldsymbol{W}$ must be block-diagonal. Thus, by Corollary 5.9, all correlations implied by a SAR(1) model on an undirected graph tend to 1 as $\rho \to \lambda_{\max}^{-1}$ provided that there is no ordering of the units such that $\boldsymbol{W}$ is block-diagonal. Block-diagonal weights matrices are important in several spatial econometric applications; see, e.g., Case (1991), Kelejian et al. (2006), Baltagi (2006) and Lee (2007).

Example 5.11 Let $\boldsymbol{E}_n$ be the $n \times n$ "equal weights" matrix $\boldsymbol{J}_n - \boldsymbol{I}_n$, where $\boldsymbol{J}_n$ denotes the $n \times n$ matrix of all ones. Consider a SAR(1) model with block-diagonal weights matrix

$$\boldsymbol{W} = \text{diag}(\boldsymbol{E}_3, \boldsymbol{E}_5). \tag{14}$$

For this matrix, $g_{\max} = 1$ and $\boldsymbol{q}_{\max} = (0, 0, 0, 1, 1, 1, 1, 1)'$. Figure 7(a) displays $\text{corr}(y_1, y_2)$ and $\text{corr}(y_4, y_5)$. Observe that while y_4 and y_5 tend to be perfectly correlated as $\rho \to \lambda_{\max}^{-1} = 0.25$, the same is not true of y_1 and y_2. This is because the SAR(1) model with weights matrix (14) can be decomposed into the product of two independent models, one for $(y_1, y_2, y_3)'$ and one for $(y_1, ..., y_5)'$. The correlations implied by the first submodel tend to 1 as ρ approaches the spectral radius of $\boldsymbol{E}_3$ (i.e., 1/2), whereas those implied by the second submodel tend to 1 as ρ approaches the spectral radius of $\boldsymbol{E}_5$ (i.e., 1/4). Next, consider the row-standardized version of (14), $\boldsymbol{W} = \text{diag}(\frac{1}{2}\boldsymbol{E}_3, \frac{1}{4}\boldsymbol{E}_5)$. In this case, the two diagonal blocks have both spectral radius 1. Figure 7(b) shows that both $\text{corr}(y_1, y_2)$ and $\text{corr}(y_4, y_5)$ tend to 1 as $\rho \to 1$.

The fact that both correlations in Figure 7(b) approach 1 as $\rho \to 1$ can be generalized as follows.



(a) (b)

Figure 7: $\text{corr}(y_1, y_2)$ (darker line) and $\text{corr}(y_4, y_5)$ (lighter line) for the weights matrix (14) (panel (a)) and its row-standardized version (panel (b)).

Corollary 5.12 *Let $\boldsymbol{W}$ be a block-diagonal weights matrix such that there is a walk from each unit to every other unit in the same block. If $\boldsymbol{W}$ is row-standardized, then all correlations implied by a SAR(1) model tend to 1 as $\rho \to 1$.*

We conclude this section by pointing out an interesting connection between Theorem 5.8 and a particular concept of centrality in a network. Theorem 5.8 establishes that, as long as $g_{\max} = 1$ and the i-th and j-th entries of $\boldsymbol{q}_{\max}$ are nonzero, $\text{corr}(y_i, y_j) \to 1$ as $\rho \to \lambda_{\max}^{-1}$. However, if the i-th and j-th entries of $\boldsymbol{q}_{\max}$ are nonzero but very close to zero (compared to all other entries of $\boldsymbol{q}_{\max}$), then Theorem 5.8 may not be very informative on the behavior of $\text{corr}(y_i, y_j)$ in a neighborhood of $\lambda_{\max}^{-1}$, because in that case $\text{corr}(y_i, y_j)$ may approach 1 very slowly. For example, the correlation between Maine and New Hampshire in Figure 4(b) does approach 1 as $\rho \to \lambda_{\max}^{-1}$, but it remains less than 0.4 unless ρ is very close to $\lambda_{\max}^{-1}$.[13] Now, it is well-known that $(\boldsymbol{q}_{\max})_i$ is a measure of the "centrality" of unit i, i.e., of the involvement of unit i in the the walk structure of the graph underlying $\boldsymbol{W}$; see, e.g., Bonacich (1972), Cvetković et al. (1980), p. 104, and Straffin (1980). A small $(\boldsymbol{q}_{\max})_i$ indicates low centrality. Hence, $\text{corr}(y_i, y_j)$ approaches 1 very slowly as $\rho \to \lambda_{\max}^{-1}$ for pairs of units that are far from the centre of the graph underlying $\boldsymbol{W}$. Note that such a nonsmooth behavior of $\text{corr}(y_i, y_j)$ close to $\lambda_{\max}^{-1}$ does not occur when $\boldsymbol{W}$ is row-standardized, because in that case $\boldsymbol{q}_{\max}$ is a vector of identical entries.

6 Discussion

This paper has analyzed how the correlation structure of SAR(1) processes depends on the weights matrix $\boldsymbol{W}$ and on the autoregressive parameter ρ. We have shown that the correlations implied by a SAR(1) model can be fully understood in terms of a particular type of walks connecting the spatial units. Each walk between two units i and j gives a contribution to the covariance between the random variables y_i and y_j, and the parameter ρ controls the relative importance of the contributions coming from walks of different lengths. This role

[13] When $\boldsymbol{q}_{\max}$ is normalized to have length 1, the entry of $\boldsymbol{q}_{\max}$ corresponding to Maine is $1.39 \cdot 10^{-7}$ and that corresponding to New Hampshire is $4.08 \cdot 10^{-6}$.

of ρ implies that the correlation properties of a SAR(1) model cannot be predicted from $\boldsymbol{W}$ alone. Indeed, we have studied a number of correlation properties of SAR(1) model that, for a fixed $\boldsymbol{W}$, hold in some interval of values of ρ around the origin, but are not guaranteed to hold for larger values of ρ. A researcher wishing to impose certain correlation properties on a set of spatial data should consider using other models, such as geostatistical models (see, e.g., Cressie, 1993, Chapter 2).

A class of spatial processes that share several similarities with SAR(1) models is that of first-order conditional autoregressive (CAR(1)) processes (e.g., Besag, 1974; Cressie, 1993). CAR(1) processes are very popular in many fields, for instance disease mapping and image analysis, but not in economics. All the results given in this paper for SAR(1) processes can be easily extended to CAR(1) processes. Indeed, the analysis of the correlation structure of CAR(1) processes is simpler. This is because the weights matrix of a CAR(1) model can be taken to be symmetric without loss of generality,[14] so that walks and their weights can be defined without regards to the direction of their steps.

Appendix A Proofs

We first give two auxiliary lemmata, and then prove all results stated in the main text.

Lemma A.1 *In a SAR(1) model, when* $|\rho|\,\lambda_{\max} < 1$*, and for any* $i, j = 1, ..., n$*,*

$$\operatorname{cov}(y_i, y_j) = \sum_{r=d(i,j)}^{\infty} (\boldsymbol{Z}_r)_{ij}\rho^r. \tag{15}$$

Proof. When $|\rho|\,\lambda_{\max} < 1$, $(\boldsymbol{I} - \rho\boldsymbol{W})^{-1} = \sum_{r=0}^{\infty}(\rho\boldsymbol{W})^r$, where, as usual, $\boldsymbol{W}^0 = \boldsymbol{I}$ (e.g., Horn and Johnson, 1985, p. 301), and hence $\operatorname{var}(\boldsymbol{y}) = (\boldsymbol{I} - \rho\boldsymbol{W})^{-1}(\boldsymbol{I} - \rho\boldsymbol{W}')^{-1} = \sum_{r,s=0}^{\infty}\rho^{r+s}\boldsymbol{W}^r(\boldsymbol{W}')^s$. Transforming the double summation into a single summation yields $\operatorname{var}(\boldsymbol{y}) = \sum_{r=0}^{\infty}\rho^r\boldsymbol{Z}_r$. The proof is completed on observing that $(\boldsymbol{Z}_r)_{ij} = 0$ if $r < d(i,j)$, by Lemma 5.1 and the definition of $d(\cdot,\cdot)$. ∎

Lemma A.2 *Let* λ *be an eigenvalue of an* $n \times n$ *matrix* $\boldsymbol{M}$*, with geometric multiplicity* g*. Then,* $\operatorname{rk}((\boldsymbol{I} - \lambda^{-1}\boldsymbol{M}')(\boldsymbol{I} - \lambda^{-1}\boldsymbol{M})) = n - g$*.*

Proof. For any matrix $\boldsymbol{B}$, $\operatorname{rk}(\boldsymbol{B}'\boldsymbol{B}) = \operatorname{rk}(\boldsymbol{B})$ and $\operatorname{rk}(\boldsymbol{B}) + \dim(\mathcal{N}(\boldsymbol{B})) = n$, where $\mathcal{N}(\boldsymbol{B})$ denotes the nullspace of $\boldsymbol{B}$ (e.g., Horn and Johnson, 1985). The result follows, because, by definition, $g = \dim(\mathcal{N}(\boldsymbol{M} - \lambda\boldsymbol{I}))$. ∎

[14] Let $\boldsymbol{L}$ be a known $n \times n$ diagonal matrix with positive diagonal entries. Provided that $(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\boldsymbol{L}$ is symmetric and positive definite, the joint distribution implied by a CAR(1) model is $\mathrm{N}\left(\boldsymbol{0}, \sigma^2(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\boldsymbol{L}\right)$. By the same argument as in footnote 3, in order to study correlations, there is no loss of generality in setting $\boldsymbol{L} = \boldsymbol{I}$, which implies that $\boldsymbol{W}$ can be assumed to be symmetric.

Proof of Theorem 4.1. Let $\mathcal{K}_{ij}^{(r)}$ denote the set of all SAR-walks of length r from i to j, so that, according to Lemma 5.1, $(\boldsymbol{Z}_r)_{ij} = \sum_{\omega \in \mathcal{K}_{ij}^{(r)}} \text{wei}(\omega)$. Then, by Lemma A.1, when $|\rho|\,\lambda_{\max} < 1$, $\text{cov}(y_i, y_j) = \sum_{r=d(i,j)}^{\infty} \sum_{\omega \in \mathcal{K}_{ij}^{(r)}} \text{wei}(\omega)\rho^r = \sum_{\omega \in \mathcal{K}_{ij}} \text{wei}(\omega)\rho^{\text{len}(\omega)}$. ∎

Proof of Lemma 5.1. On expanding the products of k matrices $\boldsymbol{W}$ and $r-k$ matrices $\boldsymbol{W}'$ that appears in the definition $\boldsymbol{Z}_r := \sum_{k=0}^{r} \boldsymbol{W}^k (\boldsymbol{W}')^{r-k}$, we obtain

$$\begin{aligned}(\boldsymbol{Z}_r)_{ij} &= \sum_{k=0}^{r} \sum_{l_1,\ldots,l_{r-1}=1}^{n} (\boldsymbol{W})_{il_1}(\boldsymbol{W})_{l_1l_2}\ldots(\boldsymbol{W})_{l_{k-1}l_k}(\boldsymbol{W}')_{l_kl_{k+1}}\ldots(\boldsymbol{W}')_{l_{r-2}l_{r-1}}(\boldsymbol{W}')_{l_{r-1}j} \\ &= \sum_{k=0}^{r} \sum_{l_1,\ldots,l_{r-1}=1}^{n} (\boldsymbol{W})_{il_1}(\boldsymbol{W})_{l_1l_2}\ldots(\boldsymbol{W})_{l_{k-1}l_k}(\boldsymbol{W})_{l_{k+1}l_k}\ldots(\boldsymbol{W})_{l_{r-1}l_{r-2}}(\boldsymbol{W})_{jl_{r-1}}.\end{aligned} \tag{16}$$

It is now clear that if, for a fixed k, $i \to l_1 \to \ldots \to l_k \leftarrow \ldots \leftarrow l_{r-1} \leftarrow j$ is a SAR-walk, then the product of entries of $\boldsymbol{W}$ in (16) is the weight of that SAR-walk. Note that, because of the summations over $k = 0, \ldots, r$ and over $l_1, \ldots, l_{r-1} = 1, \ldots, n$, expression (16) takes into account all possible SAR-walks of length r from i to j, which completes the proof. ∎

Proof of Theorem 5.2. Let $\sigma_{ij}(\rho)$ and $\sigma_{ij}^*(\rho)$ denote, respectively, $\text{cov}(y_i, y_j)$ and $\text{corr}(y_i, y_j)$ in a SAR(1) model. Let $D^r f(x)$ denote the r-th derivative of a function $f : \mathbb{R} \to \mathbb{R}$, and let $D^r f(0)$ denote $D^r f(x)$ evaluated at $x = 0$. In any neighborhood of $\rho = 0$ such that $|\rho|\,\lambda_{\max} < 1$, $\sigma_{ij}^*(\rho)$ is an infinitely differentiable function of ρ, and hence it admits the MacLaurin expansion $\sum_{r=0}^{\infty}(r!)^{-1} D^r \sigma_{ij}^*(0)\rho^r$. The bulk of the proof consists of expressing the coefficients $(r!)^{-1} D^r \sigma_{ij}^*(0)$ in terms of entries of the matrices $\boldsymbol{Z}_0, \ldots, \boldsymbol{Z}_r$. Write $\sigma_{ij}^*(\rho) = \sigma_{ij}(\rho)\eta(v(\rho))$, with $\eta(z) := z^{-1/2}$ and $v(\rho) := \sigma_{ii}(\rho)\sigma_{jj}(\rho)$. By Leibniz's formula,

$$\frac{1}{r!} D^r \sigma_{ij}^*(0) = \sum_{s=0}^{r} \frac{1}{s!(r-s)!} D^{r-s}\sigma_{ij}(0) D^s \eta(v(0)), \tag{17}$$

where, applying Faá di Bruno's formula (e.g., Albramowitz and Stegun, 1979),

$$D^s \eta(v(0)) = \sum_{\substack{k_1>0,\ldots,k_s>0: \\ k_1+2k_2+\ldots+sk_s=s}} \frac{s!}{k_1!\ldots k_s!} (D^k\eta)(v(0)) \prod_{t=1}^{s} \left(\frac{D^t v(0)}{t!}\right)^{k_t}, \tag{18}$$

with $k := \sum_{t=1}^{s} k_t$. Explicit expressions for the terms $(D^k\eta)(v(0))$ (the k-th derivative of $\eta(z)$ evaluated at $v(0)$) and $D^t v(0)/t!$ appearing in (18) can be derived as follows. Since $v(0) = 1$ and $D^k\eta(z) = \alpha_k z^{-1/2-k}$, with $\alpha_k := \prod_{l=0}^{k-1}(1/2 - l)$, we obtain $(D^k\eta)(v(0)) = \alpha_k$. In addition, from expression (15), we have that, for $|\rho|\,\lambda_{\max} < 1$, $v(\rho) = \sum_{r,s=0}^{\infty} \rho^{r+s} (\boldsymbol{Z}_r)_{ii}(\boldsymbol{Z}_s)_{jj}$. The last expression can be rewritten as $v(\rho) = \sum_{r=0}^{\infty} \rho^r \sum_{u=0}^{r} (\boldsymbol{Z}_{r-u})_{ii}(\boldsymbol{Z}_u)_{jj}$, showing that $D^t v(0)/t! = \sum_{u=0}^{t} (\boldsymbol{Z}_{t-u})_{ii}(\boldsymbol{Z}_u)_{jj}$. Thus, since $D^1 v(0) = 0$, (18) yields $D^s\eta(v(0)) = 0$ if $s = 1$, and

$$D^s \eta(v(0)) = \sum_{\substack{k_2>0,\ldots,k_s>0: \\ 2k_2+\ldots+sk_s=s}} \frac{s!\alpha_K}{k_2!\ldots k_s!} \prod_{t=2}^{s} \left(\sum_{u=0}^{t} (\boldsymbol{Z}_{t-u})_{ii}(\boldsymbol{Z}_u)_{jj}\right)^{k_t} \tag{19}$$

if $s > 1$, with $K := \sum_{t=2}^{s} k_t$. Substituting expressions (19) and $D^{r-s}\sigma_{ij}(0) = (r-s)!(\boldsymbol{Z}_{r-s})_{ij}$ into (17), we obtain $(r!)^{-1}D^r\sigma^*_{ij}(0) = (\boldsymbol{Z}_r)_{ij} - b_r(i,j)$, with

$$b_r(i,j) := \sum_{s=2}^{r}(\boldsymbol{Z}_{r-s})_{ij} \sum_{\substack{k_2>0,\ldots,k_s>0:\\ tk_2+\ldots+tk_s=s}} \frac{\alpha_K}{k_2!\ldots k_s!}\prod_{t=2}^{s}\left(\sum_{u=0}^{t}(\boldsymbol{Z}_{t-u})_{ii}(\boldsymbol{Z}_u)_{jj}\right)^{k_t}. \tag{20}$$

Observe that $b_r(i,j)$ is a sum of terms each of which contains one entry $(\boldsymbol{Z}_{r-s})_{ij}$, $s = 2,\ldots r$, as a factor. Since, by Lemma 5.1, $(\boldsymbol{Z}_r)_{ij} = 0$ if $r < d(i,j)$, it follows that $b_r(i,j) = 0$ if $(\boldsymbol{Z}_{r-2})_{ij} = 0$, or, equivalently, if $r \leq d(i,j)+1$. Thus, the first two terms in the MacLaurin expansion of $\sigma^*_{ij}(\rho)$ are $(\boldsymbol{Z}_{d(i,j)})_{ij}\rho^{d(i,j)}$ and $(\boldsymbol{Z}_{d(i,j)+1})_{ij}\rho^{d(i,j)+1}$, completing the proof. ∎

Proof of Corollary 5.4. Both parts (a) and (b) are obtained by substituting the relevant expression for $(\boldsymbol{Z}_1)_{ij}$ and $(\boldsymbol{Z}_2)_{ij}$ into Theorem 5.2. If $\boldsymbol{W}$ is a symmetric $(0,1)$ matrix, then, for any pair of neighbors (i,j), $(\boldsymbol{Z}_1)_{ij} = 2(\boldsymbol{W})_{ij} = 2$ and $(\boldsymbol{Z}_2)_{ij} = 3(\boldsymbol{W}^2)_{ij} = 3\sum_{l=1}^{n}\boldsymbol{W}_{il}\boldsymbol{W}_{lj} = 3n_{ij}$. Consider now a row-standardized weights matrix $\boldsymbol{W} = \boldsymbol{D}^{-1}\boldsymbol{A}$, where $\boldsymbol{A}$ is a symmetric $(0,1)$ matrix, and $\boldsymbol{D}$ is the diagonal matrix with $(\boldsymbol{D})_{ii} = \sum_{j=1}^{n}(\boldsymbol{A})_{ij}$, $i = 1,\ldots,n$. For any pair of neighbors (i,j), $(\boldsymbol{Z}_1)_{ij} = n_i^{-1}(\boldsymbol{A})_{ij} + n_j^{-1}(\boldsymbol{A})_{ji} = n_i^{-1} + n_j^{-1}$ and $(\boldsymbol{Z}_2)_{ij} = (\boldsymbol{W}^2)_{ji} + (\boldsymbol{W}\boldsymbol{W}')_{ij} + (\boldsymbol{W}^2)_{ij}$, where $(\boldsymbol{W}^2)_{ij} = (\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}\boldsymbol{A})_{ij} = n_i^{-1}\sum_{l=1}^{n}n_l^{-1}(\boldsymbol{A})_{il}(\boldsymbol{A})_{lj} = n_i^{-1}s_{ij}$, $(\boldsymbol{W}\boldsymbol{W}')_{ij} = (\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{A}\boldsymbol{D}^{-1})_{ij} = (n_in_j)^{-1}\sum_{l=1}^{n}\boldsymbol{A}_{il}\boldsymbol{A}_{lj} = (n_in_j)^{-1}n_{ij}$, and $(\boldsymbol{W}^2)_{ji} = n_j^{-1}s_{ij}$. ∎

Proof of Corollary 5.6. By Lemma 5.1 and the definition of $d(\cdot,\cdot)$, $(\boldsymbol{Z}_{d(i,j)})_{ij} > 0$. Hence, Theorem 5.2 implies that the larger $d(i,j)$ is, the faster $\mathrm{corr}(y_i,y_j)$ goes to zero as $\rho \to 0$. ∎

Proof of Theorem 5.8. Let $\boldsymbol{\Sigma}(\rho)$ denote the variance matrix $(\boldsymbol{I}-\rho\boldsymbol{W})^{-1}(\boldsymbol{I}-\rho\boldsymbol{W}')^{-1}$ of a SAR(1) model. For any fixed $\rho \in (\lambda_{\min}^{-1}, \lambda_{\max}^{-1})$, let $\eta_1(\rho) \leq \eta_2(\rho) \leq \ldots \leq \eta_n(\rho)$ denote the eigenvalues of $\boldsymbol{\Sigma}(\rho)$, and let $\boldsymbol{v}_1(\rho),\ldots,\boldsymbol{v}_n(\rho)$ denote n orthonormal eigenvectors of $\boldsymbol{\Sigma}(\rho)$, with $\boldsymbol{v}_l(\rho)$ associated to $\eta_l(\rho)$, $l = 1,\ldots,n$. By Lemma A.2, $\mathrm{rk}(\boldsymbol{\Sigma}^{-1}(\lambda_{\min}^{-1})) = n - g_{\min}$. Since, in addition, all eigenvalues $\eta_l(\rho)$, $l = 1,\ldots,n$, are continuous in ρ, it follows that, if $g_{\min} = 1$, then $\eta_n(\rho)$ is the only eigenvalue of $\boldsymbol{\Sigma}(\rho)$ that does not have a finite limit as $\rho \to \lambda_{\min}^{-1}$. Next, observe that $\boldsymbol{\Sigma}^{-1}(\lambda_{\min}^{-1})\boldsymbol{q}_{\min} = (\boldsymbol{I}-\lambda_{\min}^{-1}\boldsymbol{W}')(\boldsymbol{I}-\lambda_{\min}^{-1}\boldsymbol{W})\boldsymbol{q}_{\min} = 0$, that is, $\boldsymbol{\Sigma}^{-1}(\lambda_{\min}^{-1})$ has an eigenvector $\boldsymbol{q}_{\min}$ associated to its smallest eigenvalue 0. Thus, if $g_{\min} = 1$, $\boldsymbol{v}_1(\rho) \to \boldsymbol{q}_{\min}$ as $\rho \to \lambda_{\min}^{-1}$. Consider now the spectral decomposition $\boldsymbol{\Sigma}(\rho) = \sum_{l=1}^{n}\eta_l(\rho)\boldsymbol{v}_l(\rho)\boldsymbol{v}_l'(\rho)$. Dividing by $\eta_n(\rho)$ and taking the (i,j)th entry, we obtain

$$\begin{aligned}\frac{1}{\eta_n(\rho)}\mathrm{cov}(y_i,y_j) &= (\boldsymbol{v}_1(\rho)\boldsymbol{v}_1'(\rho))_{ij} + \sum_{l=2}^{n}\frac{\eta_l(\rho)}{\eta_n(\rho)}\left(\boldsymbol{v}_l(\rho)\boldsymbol{v}_l'(\rho)\right)_{ij} \\ &= (\boldsymbol{v}_1(\rho))_i\left(\boldsymbol{v}_1'(\rho)\right)_j + \sum_{l=2}^{n}\frac{\eta_l(\rho)}{\eta_n(\rho)}\left(\boldsymbol{v}_l(\rho)\right)_i\left(\boldsymbol{v}_l'(\rho)\right)_j.\end{aligned}$$

Since, as $\rho \to \lambda_{\min}^{-1}$, $\boldsymbol{v}_1(\rho) \to \boldsymbol{q}_{\min}$ and $\eta_l(\rho)/\eta_n(\rho) \to 0$ for any $l = 2,\ldots,n$, it follows that

$$\frac{1}{\eta_n(\rho)}\mathrm{cov}(y_i,y_j) \to (\boldsymbol{q}_{\min})_i(\boldsymbol{q}_{\min})_j.$$

Thus, provided that $(\boldsymbol{q}_{\min})_i(\boldsymbol{q}_{\min})_j \neq 0$, $\text{corr}(y_i, y_j) \to (\boldsymbol{q}_{\min})_i(\boldsymbol{q}_{\min})_j / |(\boldsymbol{q}_{\min})_i(\boldsymbol{q}_{\min})_j| = \text{sgn}((\boldsymbol{q}_{\min})_i(\boldsymbol{q}_{\min})_j)$, which completes the proof of part (a) of the theorem. The proof of part (b) is essentially identical, with $\boldsymbol{v}_1(\rho)$ replaced by $\boldsymbol{v}_n(\rho)$ and all the indices min replaced by max. The only substantial difference is that, by Theorem 8.3.1 of Horn and Johnson (1985), if $g_{\max} = 1$, then $\boldsymbol{q}_{\max}$ is nonnegative or nonpositive, and hence $(\boldsymbol{q}_{\max})_i(\boldsymbol{q}_{\max})_j$ cannot be negative. ∎

Proof of Corollary 5.9. If there is a walk from each unit to every other unit, then $\boldsymbol{W}$ is irreducible. By the Perron-Frobenius Theorem (e.g., Horn and Johnson, 1985), a nonnegative irreducible matrix has $g_{\max} = 1$ and an entrywise positive $\boldsymbol{q}_{\max}$, so the result follows from Theorem 5.8(b). ∎

Proof of Corollary 5.12. A SAR(1) model with block-diagonal weights matrix can be decomposed into the product of a number of submodels equal to the number of diagonal blocks. Since there is a walk from each unit to every other unit in the same block, $\boldsymbol{W}$ does not have any zero rows, and can therefore be row-standardized. By Corollary 5.9, all correlations implied by each submodel tend to 1 as $\rho \to \lambda_{\max}^{-1}$, with $\lambda_{\max} = 1$ for any row-standardized $\boldsymbol{W}$. ∎

References

Albramowitz, A. and S. Stegun (1979). *Handbook of Mathematical Functions.* Nauka, Moskow.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models.* Kluwer, Dordrecht.

Baltagi, B. H. (2006). Random effects and spatial autocorrelation with equal weights. *Econometric Theory 22*, 973-984.

Bell, K. P. and N. E. Bockstael (2000). Applying the generalized-moments estimation approach to spatial problems involving microlevel data. *The Review of Economics and Statistics 82*, 72-82.

Besag, J. E. (1972). On the correlation structure of some two-dimensional stationary processes. *Biometrika 59*, 43-48.

Besag, J. E. (1974). Spatial interaction and the statistical analysis of lattice data. *Journal of the Royal Statistical Society B 36*, 192-236.

Biggs, N. L. (1993). *Algebraic Graph Theory*, second edition. Cambridge University Press, Cambridge.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology 2*, 113–120.

Calvó-Armengol, A., E. Patacchini and Y. Zenou (2009). Peer effects and social networks in education. *The Review of Economic Studies*, forthcoming.

Case, A. (1991). Spatial patterns in household demand. *Econometrica 59*, 953-966.

Cliff, A. D. and J. K. Ord (1981). *Spatial Processes: Models and Applications*. Pion, London.

Cressie, N. (1993). *Statistics for Spatial Data*. Wiley, New York.

Cvetković, D. M., M. Doob and H. Sachs (1980). *Spectra of Graphs*. Academic Press, New York.

Fingleton, B. (1999). Spurious spatial regression: some Monte Carlo results with a spatial unit root and spatial cointegration. *Journal of Regional Science 39*, 1-19.

Godsil, C. (1993). *Algebraic Combinatorics*. Chapman and Hall, New York.

Harary, F. H. (1969). *Graph Theory*. Addison-Wesley, Reading.

Horn, R. and C. R. Johnson (1985). *Matrix Analysis*. Cambridge University Press, Cambridge.

Kelejian, H. H. and D. P. Robinson (1995). Spatial correlation: a suggested alternative to the autoregressive model, in: L. Anselin, R. Florax (Eds.), *New Directions in Spatial Econometrics*. Springer Verlag, Berlin.

Kelejian, H. H. and I. R. Prucha (2009). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, forthcoming.

Kelejian, H. H., I. R. Prucha and Y. Yuzefovich (2006). Estimation problems in models with spatial weighting matrices which have blocks of equal elements. *Journal of Regional Science 46*, 507-515.

Lee, L. F. (2007). Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics 140*, 333-374.

Lee, L. F. and J. Yu (2008). Near unit root in the spatial autoregressive model. Manuscript, The Ohio State University.

Millimet, D. L. and T. Collier (2008). Efficiency in public schools: Does competition matter? *Journal of Econometrics 145*, 134-157.

Straffin, P. D. (1980). Linear algebra in geography: eigenvectors of networks. *Mathematics Magazine 53*, 269-276.

Wall, M. M. (2004). A close look at the spatial structure implied by the CAR and SAR models. *Journal of Statistical Planning and Inference 121*, 311-324.

Whittle, P. (1954). On stationary processes in the plane. *Biometrika 41*, 434-449.