Testing the Correct Specification of a Spatial Dependence Panel Model for Stock Returns

Tim Kutzker* and Dominik Wied†

May 8, 2020

Abstract

This paper provides specification tests for the m-dimensional spatial autoregressive (SAR) panel model by deriving the limiting distribution of the specification test statistics and examines size and power in finite sample simulations. In the empirical application we analyzed the Euro Stoxx 50 returns. Regarding this, a 3-dimensional SAR panel model incorporating global dependencies, dependencies inside industrial branches and local dependencies is assumed. The investigation shows the tests' ability to detect inaccurate Value-at-Risk forecasts.

Keywords: heteroscedasticity, method of moments, spatial dependence, stock returns, Value-at-Risk

JEL Classification Numbers: C13, C51, G12.

*Corresponding Author, Institute of Econometrics and Statistics, University of Cologne, Address: Albertus Magnus Platz, D- 50923 Cologne, Germany, telephone: +49 221 470 6561, e-mail: tim.kutzker@uni-koeln.de.

†Institute of Econometrics and Statistics, University of Cologne, 50923 Cologne, Germany, email: dwied@uni-koeln.de

1 Introduction and Summary

In recent years the literature in economics and finance has found some interest in the connection between spatial dependence and stock returns. For example, Asgharian et al. (2013) use techniques from spatial econometrics in order to investigate in which way stock market co-movements are determined by countries' economic and geographical relations. One result shows that trade is the most important factor. Tam (2014) analyzes equity market linkages in East Asia with the result, among others, that Japan is a dominant driver. Selan and Kalatzis (2017) analyze peer effects in Brazil and find a positive spatial dependence between stock returns from peer companies, but a negative feedback effect from fundamental characteristics. A seminal methodological contribution is given by Blasques et al. (2016) who extend the spatial Durbin model by a time-varying spatial dependence parameter.

Furthermore, Arnold et al. (2013) propose a spatial autoregressive (SAR) panel model for stock returns in order to capture local dependencies and dependencies within industrial branches. Wied (2013) considers structural breaks in these models and Schmitt et al. (2016) combine the approach with local normalization techniques. Gong and Weng (2016) use the model for value at risk forecasts in the Chinese stock market. Catania and Billé (2017) generalize the SAR model with autoregressive and heteroscedastic disturbances by including methods from score-driven models. Moreover, Zhang et al. (2017) propose a dynamic spatial panel with generalized autoregressive conditional heteroscedastic model (DSP-GJR-GARCH). Lu (2017) considers a spatial panel data model, that models three effects jointly. Various empirical analyses in the aforementioned papers show that the SAR panel model is generally suitable for Value-at-Risk (VaR) forecasts and outperforms, e.g. the one-factor model.

One aspect which is often missing in recent literature is the question how good the model fits the data. In general, people tend to look at Moran's I (Moran, 1950 and Li et al., 2007) to analyze if there is spatial dependence in a given data set.

However, this measure is not connected to a specific model. One could apply it to somehow obtained model residuals, but even then, the question would remain in which way we can use this for a test. Born and Breitung (2011) and Su and Qu (2017) propose specification tests for SAR models, but they do not consider a panel context. Kelejian and Piras (2016) propose a J-test procedure for testing a null model against non-nested alternatives for a fixed effects spatial panel data framework. A crucial prerequisite of this test is to formulate what they call G alternative models under H_1.

In this paper, we revisit the SAR panel model from Arnold et al. (2013) and propose two methods on how to check the model fit. The basic idea stems from the model assumption that spatial weighting matrices capture all spatial dependence and that the remaining error terms are spatially uncorrelated. Therefore, we consider the model residuals such that the tests keep the null hypothesis of model fit if the covariance matrix of the residuals is basically diagonal, i.e. its off-diagonal elements are close to zero. We derive the asymptotic distribution of our test statistics and show in simulations, that the tests have reasonable power properties against sparse error term covariance matrices. An empirical application on stock data shows that the tests can potentially also be used as backtests for Value-at-Risk forecasts.

This paper is organized as follows: Section 2 describes the classical spatial autoregressive model, discusses the assumptions for a GMM estimation procedure and derives the specification tests. Section 3 provides an extensive Monte Carlo Simulation and Section 4 an empirical application. Finally, Section 5 concludes.

2 A Cross Sectional Correlation Based Specification Test for SAR(m) Panel Models

In this section, we introduce the general SAR(m), $m \in I\!N$ panel model and discuss briefly the slightly modified assumptions for the two step GMM estimator given in

(Arnold et al., 2013) which turn out to also hold for the m-dimensional case.

2.1 The Model

The SAR(m) panel model assumes that the dependent variable is correlated in the cross-sectional dimension n and that the spatial dependence can be separated into m different parts. The number m[1] and the specific form of the spatial matrices depend on the practitioner[2]. Thus, the spatial matrices W_i, $i = 1, ..., m$ are pre-specified and fixed. In what follows, let $\boldsymbol{y}_t$ and $\boldsymbol{\varepsilon}_t$ be n-dimensional random vectors for $t = 1, \dots, T$. The m-dimensional SAR panel model without any explanatory variables is given by

$$\boldsymbol{y}_t = \sum_{i=1}^{m} \rho_i W_i \, \boldsymbol{y}_t + \boldsymbol{\varepsilon}_t, \ t = 1, \dots, T \tag{1}$$

where $\rho_i \in \mathbb{R}$ for $i = 1, ..., m$. For asymptotic results, n n is fixed and T is sent to infinity. To derive limit theorems we impose the following assumptions:

Assumption 1.

1. *The sequence of random vectors $\{\boldsymbol{y}_t\}_{t \in I\!N}$ has zero mean, is stationary and ergodic.*

2. *For $i \in \{1, ..., m\}$, $r = 1, \dots, n$, $s = 1, \dots, n$, $W_{i,rs} \geq 0$, $W_{i,rr} = 0$.*

3. *For $i \in \{1, ..., m\}$ and $r = 1, \dots, n$, $\sum_{s=1}^{n} W_{i,rs} = 1$.*

4. *The parameter space S is defined as $S := \{\boldsymbol{\rho} \in \mathbb{R}^m : ||\boldsymbol{\rho}||_1 < 1\}$ where $|| \cdot ||_1$ defines the L_1-norm.*

5. *For $t \in \mathbb{Z}$, $\mathsf{Cov}(\boldsymbol{\varepsilon}_t) = diag\{\sigma_1^2, \dots, \sigma_n^2\} =: \Sigma \in \mathbb{R}^n$.*

[1]In the application later on, we will introduce three different spatial matrices which are assumed to capture the structure of daily stock returns. The first part covers a general dependence which affects all subjects equally. The second part captures dependencies among industrial branches and national effects are included with the help of the third dependency structure.

[2]An overview of commonly used matrices is given in J.P. Elhorst (2012).

6. *Each element of the vector* $\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\right)_{i<j}$ *meets the assumption of a central limit theorem and the corresponding long-term covariances* $\sum_{s,t\in I\!N}\mathsf{Cov}\left[\varepsilon_{i1}\varepsilon_{jt},\varepsilon_{ks}\varepsilon_{ls}\right]$ *are finite for every* $i<j$ *and* $k<l$*.*

In the context of daily stock returns, the zero mean and stationarity Assumption 1.1 is plausible (see also Aue et al. (2009)). Assumption 1.2 excludes "self influence" since the elements on the leading diagonal are zero and postulates that all elements are non-negative, which is usually the case in empirical applications. Assumption 1.3 ensures that the matrices are bounded and standardized. For the GMM estimator based on Arnold et al. (2013) we assume row-standardized weighting matrices. Depending on the underlying GM-estimation technique this assumption could be relaxed (cf. Kelejian and Prucha (2010), Breitung and Wigger (2017) et al.). Assumption 1.4 restricts the parameter space such that the sum of the absolute values of the elements of $\boldsymbol{\rho}\in\mathbb{R}^m$ is smaller than 1. Even though the assumption could be slightly generalized (cf. J.P. Elhorst (2012)) we follow the notation of Arnold et al. (2013) as it guarantees that the matrix $(I_n-\sum_{i=1}^{m}\rho_i W_i)$ is non-singular[3]. Hence, Assumption 1.1-1.4 ensure the model 1 to be well defined.

The crucial assumption, on which we will base our specification test, is that the covariance matrix of the error terms $\boldsymbol{\varepsilon}_t$ is diagonal. Consequently, all cross-sectional dependence is captured by the spatial terms, which corresponds to Assumption 1.5, although, heteroscedasticity is not excluded. Assumption 1.6. guarantees that the limiting distribution of our suggested test statistic is not degenerated, i.e. the dependence structure of the error vector $\boldsymbol{\varepsilon}_t$ meets certain regularity conditions, such that the serial dependence structure is bounded.

For the estimation, a two step GMM procedure is considered. First, we estimate the correlation parameters by the method of moments along the lines of Kelejian and Prucha (1999) or Kapoor et al. (2007). This step does not depend on the parameters of variance. Secondly, we estimate the variance parameters. Under some

[3]The matrix $(I_n-\sum_{i=1}^{m}\rho_i W_i)$ is strictly diagonally dominant.

regularity assumptions the GMM estimator $\hat{\boldsymbol{\rho}}$ is consistent and as asymptotically normal. While this is worked out in Arnold et al. (2013) for the special case of $m = 3$, a detailed derivation for the GMM estimator in the general case is presented in the Appendix A.

2.2 The Specification Test

We outline the test for the case of Assumption 1, noting that simulation results in section 3.1 indicate that the test also works, if we replace the error terms by GARCH residuals. So subsequently, the word data set can be regarded either as the original or the GARCH adjusted data.

Following the discussion given in the previous subsection, what remains is to check whether Assumptions 1.5 holds. Even if the course of action seems technical, the idea behind the test statistic is straightforward: we do not reject the null hypothesis if the covariance matrix of the errors is basically a diagonal matrix, i.e. its off-diagonal elements deviate not too far from zero. Let $\hat{H} \in \mathbb{R}^{n\times n}$ denote the empirical covariance matrix of the residuals times the square root of the time horizon, i.e. $\hat{H} := \sqrt{T}\hat{\mathsf{Cov}}\,[\hat{\boldsymbol{\varepsilon}}_t]$ and $\hat{H}_{ij}$ with $i,j \in \{1,2,...,n\}$ its elements. Let σ_{ij}^2 denote the (i,j)-th element of the theoretical counterpart Σ, i.e. the error covariance matrix. Since $\hat{H}$ and Σ are symmetric, it is sufficient to consider only the elements of the upper triangle of the matrix Σ. Hence, the null hypothesis is given by

$$H_0 : \sigma_{ij}^2 = 0 \;\text{ for all }\; i<j \qquad \text{vs.} \qquad H_1 : \exists s,t \text{ with } s<t : \sigma_{st}^2 \neq 0. \tag{2}$$

We opt to use χ^2-type tests for this testing problem. Instead of considering each element or the maximum of the absolute value of all off-diagonals, we take the sum of each element squared into account. Thus, the naive test statistic is given by

$$S := \sum_{i<j,i,j=1,...,n} (\hat{H}_{ij})^2. \tag{3}$$

The following theorem identifies the limiting distribution of the empirical covariance matrix times $\sqrt{T}$.

Theorem 2.1. *Under the null hypothesis $H_0 : \sigma_{ij}^2 = 0$ for all $i < j$, the assumptions of Theorem A.3, the following holds for $1 \leq i,j \leq n$*

$$\underset{T\to\infty}{dlim} \sqrt{T}\ \hat{\mathsf{Cov}}\ [\hat{\boldsymbol{\varepsilon}}_t] = A + B + B^{'} \in \mathbb{R}^{n\times n} \tag{4}$$

with $(A)_{ii} = \underset{T\to\infty}{lim} \sqrt{T} \sum_{t=1}^{T} \sigma_{it}^2 = \infty$ and the components of A are jointly normally distributed for $i \neq j$ with $(A)_{ij} \sim N(0, \lim_{T\to\infty} \mathsf{Var}\ [\frac{1}{\sqrt{T}} \sum_{t=1}^{\infty} \boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{jt}])$ and $Cov((A)_{ij},(A)_{kl}) = 0$ for $i \neq j$ and $k \neq l$ with $(i,j) \neq (k,l)$. Moreover, $B \overset{d}{=} (\sum_{g=1}^{m} X_g W_g)(I_n - \sum_{g=1}^{m} \rho_g W_g)^{-1}\Sigma$, where $\boldsymbol{X} := (X_1,\ldots,X_m) \sim N(0, \boldsymbol{d}^{-1} S_W (\boldsymbol{d}^{-1})^{'}) \in \mathbb{R}^{1\times m}$ with $S_W = \sum_{t=-\infty}^{\infty} \mathsf{E}[f(\boldsymbol{y}_t,\boldsymbol{\rho})f(\boldsymbol{y}_t,\boldsymbol{\rho})']$ for $f(\boldsymbol{y}_t,\boldsymbol{\rho}) = (\boldsymbol{\varepsilon}_t' W_1 \boldsymbol{\varepsilon}_t, \cdots, \boldsymbol{\varepsilon}_t' W_m \boldsymbol{\varepsilon}_t)'$ and $\boldsymbol{d}$ defined in Assumption (2).

Here and in the following $\underset{T\to\infty}{\text{dlim}}$ denotes limit in distribution and $\overset{d}{=}$ equality in distribution. Three remarks about Theorem 2.1 are in order. First, the leading elements of matrix A diverge to infinity. However, the tests considers only the off-diagonal elements ($i \neq j, i,j = 1,...,n$), which are finite by Assumption 1.6. This in turn ensures, that the test is well defined. Second, since $\left(I_n - \sum_{g=1}^{m} \rho_g W_g\right)$ is strictly diagonally dominant, the inverse exists. Third, we note that the matrices B and its transposed appear in the limit. This is due to the effect of estimating $\boldsymbol{\rho}$ instead of using the unknown population quantity. The analysis of such a residual effect (compare Demetrescu and Wied, 2018+) is somewhat complicated, since the additional terms need different standardizing factors in the proof[4]. However, all terms in the limiting distribution are based on the same error terms, thus, the convergence is jointly and the limiting distribution in (4) is multivariate normal. If we additionally assume serially independence in the error vector, the variance of the elements in the limiting matrix A simplifies to a product, shown in the following remark.

Remark 2.2. *Suppose the assumptions of Theorem 2.1 hold. If $\{\boldsymbol{\varepsilon}_t\}_{t\in\{1,\ldots,T\}}$ is*

[4]For a detailed analysis of the convergence rate we refer to Lemma B.1 in the Appendix.

serially independent, then

$$(A)_{ij} \sim N(0,\sigma_i^2\sigma_j^2) \quad \textit{for} \quad i \neq j. \tag{5}$$

In accordance with our test statistic (3), we can reformulate the test in vectorial notation, i.e.

$$S = \hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\alpha}}, \tag{6}$$

where $\hat{\boldsymbol{\alpha}}$ represents the vector of the upper triangle of the empirical covariance matrix of the residuals times $\sqrt{T}$. Since the empirical covariance matrix consists of n^2 elements, the upper triangle matrix vector (i.e. stacking every element above the leading diagonal, but excluding elements from the leading diagonal) consists of $n(n-1)/2$ elements and has the following form:

$$\begin{aligned}
\hat{\boldsymbol{\alpha}} \ &:= \ \underset{T\to\infty}{\mathrm{dlim}} \left(\sqrt{T}\hat{\mathsf{Cov}}\ [\hat{\boldsymbol{\varepsilon}}_t]\right)_{i<j,\ i,j=1,\ldots,n} \\
&= \ \underset{T\to\infty}{\mathrm{dlim}} \left(\frac{1}{\sqrt{T}}\sum \hat{\boldsymbol{\varepsilon}}_t\hat{\boldsymbol{\varepsilon}}_t'\right)_{i<j,\ i,j=1,\ldots,n} \ = \ \underset{T\to\infty}{\mathrm{dlim}} \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\hat{\boldsymbol{d}}_t \in \mathbb{R}^{\frac{n(n-1)}{2}} \\
\text{with } \hat{\boldsymbol{d}}_t &:= \left(\hat{\varepsilon}_{1t}\hat{\varepsilon}_{2t},\ldots,\hat{\varepsilon}_{1t}\hat{\varepsilon}_{nt},\hat{\varepsilon}_{2t}\hat{\varepsilon}_{3t},\ldots,\hat{\varepsilon}_{2t}\hat{\varepsilon}_{nt},\ldots,\hat{\varepsilon}_{(n-1)t}\hat{\varepsilon}_{nt}\right)'.
\end{aligned}$$

By means of Slutsky's theorem we define the theoretical counterpart

$$\begin{aligned}
\boldsymbol{\alpha} \ &:= \ (A)_{i<j,\ i,j=1,\ldots,n} \\
&= \ \underset{T\to\infty}{\mathrm{dlim}} \left(\frac{1}{\sqrt{T}}\sum \boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\right)_{i<j,\ i,j=1,\ldots,n} \ = \ \underset{T\to\infty}{\mathrm{dlim}} \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{d}_t \in \mathbb{R}^{\frac{n(n-1)}{2}} \\
\text{with } \boldsymbol{d}_t &:= \left(\varepsilon_{1t}\varepsilon_{2t},\ldots,\varepsilon_{1t}\varepsilon_{nt},\varepsilon_{2t}\varepsilon_{3t},\ldots,\varepsilon_{2t}\varepsilon_{nt},\ldots,\varepsilon_{(n-1)t}\varepsilon_{nt}\right)'
\end{aligned}$$

which stacks the upper triangular matrix of the covariance matrix of the errors times $\sqrt{T}$ in a vector. Analogously, $\boldsymbol{\beta}$ defines the vector of the stacked upper triangular matrix of B and $\boldsymbol{\beta}^*$ of B', respectively, i.e. for $Z_W := \underset{T\to\infty}{\mathrm{dlim}} \sum_{g=1}^{m} \sqrt{T}(\rho_g - \hat{\rho}_g)W_g$

$$\begin{aligned}
\boldsymbol{\beta} \ &:= \ (B)_{i<j,i,j=1,\ldots,n} \ = \ \left(Z_W(I_n - \sum_{g=1}^{m}\rho_g W_g)^{-1}\Sigma\right)_{i<j,\ i,j=1,\ldots,n} \in \mathbb{R}^{\frac{n(n-1)}{2}}, \\
\boldsymbol{\beta}^* \ &:= \ \left(B'\right)_{i<j,i,j=1,\ldots,n} \ = \ \left(\Sigma'(I_n - \sum_{g=1}^{m}\rho_g W_g')^{-1}Z_W'\right)_{i<j,\ i,j=1,\ldots,n} \in \mathbb{R}^{\frac{n(n-1)}{2}}.
\end{aligned}$$

The vectors $\boldsymbol{\beta}$ and $\boldsymbol{\beta}^*$ are well defined, since B is not necessarily symmetric.

Lemma 2.3. *$\boldsymbol{\beta}$ represents the vector of the upper triangle and $\boldsymbol{\beta}^*$ the vector of the lower triangle of the matrix $Z_W(I_n - \sum_{g=1}^{m} \rho_g W_g)^{-1}\Sigma$, i.e. for $i,j \in \{1,...,n\}$*

$$\boldsymbol{\beta}^* \; = \; \left(Z_W(I_n - \sum_{g=1}^{m} \rho_g W_g)^{-1}\Sigma \right)_{i>j,\; i,j=1,...,n} \in \mathbb{R}^{\frac{n(n-1)}{2}}. \tag{7}$$

The next Lemma provides the limit distribution of our test statistic S (6).

Lemma 2.4. *Suppose the assumptions of Theorem 2.1 hold. Then the test statistic S (3) is asymptotically distributed as*

$$S \; = \; \hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\alpha}} \xrightarrow[T\to\infty]{d} (\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\beta}^*)'(\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\beta}^*),$$

where the covariance matrix for $\boldsymbol{\alpha}$ is given by

$$\mathsf{Cov}\;[\boldsymbol{\alpha}] \;=\; \begin{pmatrix} \lim\limits_{T\to\infty} \mathsf{Var}\;[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \varepsilon_{1t}\varepsilon_{2t}] & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lim\limits_{T\to\infty} \mathsf{Var}\;[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \varepsilon_{(n-1)t}\varepsilon_{nt}] \end{pmatrix}.$$

Consequently, the critical value for the test statistic S (3) can be derived by drawing independently from the limiting distribution given in Lemma 2.4 and computing the corresponding quantile[5]. The test takes care of the size demands and has good power properties as shown in Section 3. The next subsection presents another specification test, which has greater size and consequently better power properties.

2.3 A More Powerful Test

In Theorem 2.1 we have shown, that the elements of the limiting distribution follow a multivariate normal distribution. Thus, if we standardize the test statistic S (6) by its covariance matrix we receive a new test statistic S^*_χ which is χ^2-distributed,

[5]The complex dependence structure of α and β can be simulated with the help of the Taylor series approximation.

i.e.

$$S_\chi^* \ := \ \hat{\boldsymbol{\alpha}}'(\mathsf{Cov}\ [\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\beta}^*])^{-1}\hat{\boldsymbol{\alpha}} \sim \chi^2_{\frac{n(n-1)}{2}}. \tag{8}$$

The terms $\boldsymbol{\beta}$ and $\boldsymbol{\beta}^*$ can be regarded as additional noise which comes from the estimation procedure. This additional noise can be extracted by decomposing the covariance matrix given in (8) into two parts. Thus, we have

$$\mathsf{Cov}\ [\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\beta}^*] \ = \ \mathsf{Cov}\ [\boldsymbol{\alpha}]+\Psi \tag{9}$$

with $\Psi \ := \ \mathsf{Cov}\ [\boldsymbol{\beta}+\boldsymbol{\beta}^*]+\mathsf{Cov}\ [\boldsymbol{\alpha},\boldsymbol{\beta}+\boldsymbol{\beta}^*]+\mathsf{Cov}\ [\boldsymbol{\alpha},\boldsymbol{\beta}+\boldsymbol{\beta}^*]'$. The first part $\mathsf{Cov}\ [\boldsymbol{\alpha}]$ covers the underlying variance structure while the second part Ψ can be considered as a noise term[6]

If either $||(\mathsf{Cov}\ [\boldsymbol{\alpha}])^{-1}\Psi|| < 1$ or $||\Psi(\mathsf{Cov}\ [\boldsymbol{\alpha}])^{-1}|| < 1$ hold[7], we can estimate the inverse of covariance matrix (9) with the help of the Taylor series approximation and the telescoping sum[8]. It yields

$$\begin{aligned}(\mathsf{Cov}\ [\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\beta}^*])^{-1} \ &= \ (\mathsf{Cov}\ [\boldsymbol{\alpha}])^{-1}-(\mathsf{Cov}\ [\boldsymbol{\alpha}])^{-1}\Psi(\mathsf{Cov}\ [\boldsymbol{\alpha}])^{-1}\\ &\quad +(\mathsf{Cov}\ [\boldsymbol{\alpha}])^{-1}\Psi(\mathsf{Cov}\ [\boldsymbol{\alpha}])^{-1}\Psi(\mathsf{Cov}\ [\boldsymbol{\alpha}])^{-1}-\ldots\\ &\leq \ (\mathsf{Cov}\ [\boldsymbol{\alpha}])^{-1}.\end{aligned}$$

Thus, $(\mathsf{Cov}\ [\boldsymbol{\alpha}])^{-1}$ is an upper bound for the inverse of the covariance matrix (9). Hence,

$$S_\chi \ := \ \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\hat{\boldsymbol{d}}_t{}'(\mathsf{Cov}\ [\boldsymbol{\alpha}])^{-1}\sum_{t=1}^{T}\hat{\boldsymbol{d}}_t \tag{10}$$

provides a more powerful test, since $S_\chi \geq S_\chi^* \sim \chi^2_{\frac{n(n-1)}{2}}$. In order to study the behavior of S and S_χ in finite samples we perform an extensive Monte Carlo Simulation which can be found in the next section.

[6]If we additionally assume serial independence the covariance matrix of $\boldsymbol{\alpha}$ can easily be implemented, since only the variances need to be estimated, cf. Lemma B.3. Otherwise, the covariance matrix of $\boldsymbol{\alpha}$ is given in Lemma 2.4.

[7]In our Monte Carlo simulation we observed that this is usually the case whenever the variance of $\boldsymbol{\varepsilon}_{it}$ is greater than 1 for all $i = 1,\ldots,n$.

[8]The sum and product of two symmetric positive semidefinite (psd) matrices is still psd.

3 Monte Carlo Simulation

The Monte Carlo (MC) Simulation consists of three major simulations. While the first two simulations assume serial independence, the third simulation examines the behavior of the test in the case of GARCH(1,1) driven errors. The first less comprehensive simulation depicts a 3-dimensional SAR model, i.e.

$$\boldsymbol{y}_t = \rho_1 W_1 \boldsymbol{y}_t + \rho_2 W_2 \boldsymbol{y}_t + \rho_3 W_3 \boldsymbol{y}_t + \boldsymbol{\varepsilon}_t,\ t = 1, \ldots, T$$

with $(W_1)_{ij} = \frac{1}{n-1}$ for all $i \neq j$ and $(W_1)_{ii} = 0$. The spatial matrices W_2 and W_3 are defined as

$$(W_2)_{ij} = \begin{cases} 1,\ \text{if } j \text{ even and } i \neq j \\ 1,\ \text{if } j-1 = i \\ 0,\ \text{otherwise} \end{cases} \qquad (W_3)_{ij} = \begin{cases} 1/(n/2-1),\ \text{if } i,j \gtrless \frac{n}{2} \\ 0,\ \text{otherwise,} \end{cases}$$

where additionally the matrix W_2 is row standardized by its row sum $\sum_j (W_2)_{ij}$. The expression $i,j \gtrless \frac{n}{2}$ in the definition of W_3 indicates that both i and j are either smaller or equal or greater than $\frac{n}{2}$. In terms of interpretation the matrix W_1 can be regarded as a weighting matrix, where each firm has the same weight with respect to a portfolio. Thus, the matrix W_1 captures a general effect, e.g. global crisis, market performance in the past etc.[9]. The spatial matrix W_2 can be considered as industry affiliation. W_3 may be regarded as the dichotomous component of the market which divides the market into two different fields (e.g. the beneficiaries of a given change, e.g. fiscal reform, aid payments, etc) and those who are not affected.

In the first part, the vector of observation $\boldsymbol{y}_t$ is generated by a multivariate normal error vector $\boldsymbol{\varepsilon}_t$ with zero mean and covariance matrix $\Sigma := \sigma^2 I_n$, where I_n represents the n-dimensional identity matrix. The parameter of spatial dependence is given by $\boldsymbol{\rho} = (0.45, 0.3, 0.15)$ and the homoscedastic variance equals $\sigma^2 = 2$. For calculating

[9] Even if W_1 is equally weighted ρ_1 cannot be considered as a fixed affect which affects market participant equally, since fixed affects are time independent. SAR models try to capture this time dependence structure with fixed weighting matrices.

the power of our tests we use the following misspecification: If we consider a market with n participants then there are $n(n-1)/2$ possible pairs (e.g. participants that are correlated with each other). The parameter ζ describes the portion of how many pairs we wish to consider[10], the parameter κ^2 describes their correlation. E.g. if we consider a market that consists of $n = 20$ actors then there are $n(n-1)/2 = 190$ different pairs. If $\zeta = 0.1$ and $\kappa = 0.2$ we presume that there are 19 pairs that have a correlation coefficient that is equal to 0.04. No further assumptions are made about the structure of the correlation. It is possible, for instance, that the actor number of 20 has the correlation 0.04 with every other participant, i.e. Σ is a diagonal matrix with 0 in the off diagonals. Only the last column and row of Σ is non-zero. However, in general, the correlation structure is completely random[11]. To determine the size and the power of the first test (2.1) we draw $B = 400$ times from the asymptotic limit distribution given in Lemma (2.4). The dependence between α and β is modelled by the Taylor series approximation. The overall number of MC repetitions is equal to 701.

Table 1 about here

We begin by studying the size of the first test for $n = 20, 50$ and $T = 50, 100, 200, 500$. Results are presented in Table (1). Collectively, the test has good size. Similar properties are derived for the power analysis of the test. Whenever the ratio of T over n is small and the dependence structure in the error term is more or less negligible (cf. $\kappa = \zeta = 0.05$) the power of the test is low. However, if there are sufficient observations (i.e. $\frac{T}{n} > 10$) and if the dependence structure in the data set is not negligible ($\kappa, \zeta \geq 0.1$) then the test provides good power properties. All in all we observe an increasing power whenever the dependence structure (κ or ζ) or the number of observations (n or T) increases.

[10] In case that $\zeta \cdot n(n-1)/2$ is not even we round down.

[11] This procedure of misspecfication ensures that the moment conditions (12) are violated, thus, the GMM estimator is biased, cf. Hansen (1982).

Table 2 about here

Similar results are obtained for the second test S_χ (3) which can be found in Table (2)[12]. In small sample studies S_χ performs worse than test S in terms of size power. This is due to the fact that the we used the empirical approximation for the inverse covariance matrix that is employed in S_χ, which is biased in small samples. Consequently, as T tends to infinity the size of the test S_χ converges clearly to the desired nominal level of 5% and the power increases as the level of misspecification rises.

However, additional simulations show that the tests' power decreases in the case of too large ζ, i.e. in the case of a highly non-sparse covariance matrix. Here, the population moment conditions from (12) are severely violated so that the model is misspecified and the behavior of the model estimators $\hat{\rho}$ is unclear, compare also Fleming (2004).

To summarize, both tests show good size and power properties whenever the ratio T over n is greater or equal to 10. Based on the simple limiting distribution of S_χ^*, the test S_χ is also very easy to implement since the approximation test S_χ requires only the empirical covariance matrix of the residuals.

The second MC simulation extends the investigations. Here, we consider a SAR(4) model

$$\boldsymbol{y}_t = \rho_1 W_1 \boldsymbol{y}_t + \rho_2 W_2 \boldsymbol{y}_t + \rho_3 W_3 \boldsymbol{y}_t + \rho_4 W_4 \boldsymbol{y}_t + \boldsymbol{\varepsilon}_t,\ t = 1, \ldots, T$$

where W_1 is a group interaction matrix of the first two-thirds, W_2 is a group interaction matrix of the last one-third, W_3 a binary contiguity matrix of the third-order neighbors only (the observations $1,...,n$ are assumed to be in a circle, i.e. 2 is a

[12] The second test is applicable since we observed in every study and simulation we conducted that either $||(\mathsf{Cov}\ [\boldsymbol{\alpha}])^{-1}\Psi|| < 1$ or $||\Psi(\mathsf{Cov}\ [\boldsymbol{\alpha}])^{-1}|| < 1$ hold

neighbor of $n-1, n, 1, 3, 4, 5$ and[13]

$$(W_4)_{ij} = \begin{cases} \frac{1}{2 \cdot \lfloor n-1 \rfloor}, \text{ if } i \text{ is even and } j \text{ odd or vice versa} \\ 0, \text{ otherwise.} \end{cases}$$

The weighting vector ρ is given by $\boldsymbol{\rho} = (-0.2\ 0.05\ 0.1\ 0.5)$. Moreover, we presuppose heteroscedastic normal error terms, i.e. $\sigma_i \sim N(0,1)$ for $i = 1, ..., n$. In order to analyze the power in case of misspecification, we choose ζ and κ likewise to the first MC simulation. To determine the size and power we follow the recommendations given in MacKinnon (2002) and draw $B = 400$ times from the asymptotic limit distribution given in Lemma (2.4). The overall number of MC repetitions is equal to 701. The results of the tests can be found in Table 3.

Table 3 about here

Even if the results of the second analysis are not one-to-one comparable with those from the first simulation[14] it is clearly observable that the tests hold the size level. The power increases if either the correlation structure (κ or ζ) or the amount of observation grows (n or T). Thus, the results presented in the second, more complex study are in line with those given in the first simulation.

In summary, the MC study has shown that the test is also applicable in case of small samples as long as the vector of observations is sufficiently large compared to the cross sectional dimension n. The next section shows that the test even holds size and power demands if the error terms follow a GARCH process.

3.1 GARCH(1,1)

One of the many problems researchers and practitioners face when analyzing data series in financial markets is their structure. Thus, volatility of financial assets has

[13]Matrices W_1, W_2, W_3 are the counterparts to the matrices $G1, G2, BC3$ given in J.P. Elhorst (2012).

[14]The model presupposes heteroscedasticity and the spatial structure is completely different. From this it follows, that the violation of the moment condition (12) is not one-to-one comparable.

been extensively studied in the last twenty years. An important aspect is volatility clustering, i.e. conditional heteroskedasticity, which leads to an increase in the probability of rare events, that is often modelled with GARCH errors. Since the SAR(m) model is a powerful instrument in modelling financial data[15], the third and final Monte Carlo simulation for the suggested test (6) assumes that the errors of the data generating process (DGP) are driven by a GARCH(1,1) model, i.e. for $t = 1,...,T$ and $i = 1,...n$

$$y_{it} = \sigma_{it}(I_n - \rho_1 W_1 - \rho_2 W_2 - \rho_3 W_3)^{-1}\epsilon_{it},$$
$$\sigma_{it}^2 = 0.33 + 0.33\sigma_{i(t-1)}^2 + 0.075 y_{i(t-1)}^2,$$
$$\epsilon_{it} \overset{i.i.d}{\sim} N(0,1).$$

To receive comparable results, the weighting matrices W_1, W_2, W_3 are similar to those of the first MC simulation of Section 3. The size and power results are presented in Table 4. At first, it should be noted that the amount of observation of a GARCH adjusted data set needs to be significantly higher compared to a data set with no GARCH adjustment, since for the case of a GARCH adjustment an initial estimate needs to be conducted. Thus, a primarily high error of estimation distorts the stationarity assumption. However, with a sufficiently large set of observations, the test S (6) performs also well with reference to size and power.

Table 4 about here

4 Empirical Analysis

We analyze the spatial dependencies in the daily stock returns of the Euro Stoxx 50 members in the composition of January 2010 for the period from 2003 until 2009, using adjusted stock prices from Datastream which we transfer to log returns. Our

[15]The empirical analysis in Section 4 shows, that a SAR(3) seems reasonable in times of no economic crisis.

basic model for the stock returns on day $t = 1, ..., T$, is

$$\boldsymbol{y}_t = \rho_g W_g \boldsymbol{y}_t + \rho_b W_b \boldsymbol{y}_t + \rho_l W_l \boldsymbol{y}_t + \boldsymbol{\varepsilon}_t \tag{11}$$

where $\boldsymbol{y}_t$ is the vector of stock returns on day t while the weighting matrices W_g, W_b, W_l capture general dependencies[16], dependencies insides branches and local dependencies[17] . In all weighting matrices, market capitalization is taken into account, i.e. the share of the respective firm is written into the non-zero entries of the rows, before the matrices are row-standardized to 1.

Table 5 about here

Thus, none of the spatial matrices W_g, W_b, W_l are symmetric. The unknown parameters ρ_g, ρ_b and ρ_l represent the corresponding factors. Our main interest is to provide statistical evidence whenever the spatial model (11) is applicable. Therefore, we conducted an extensive empirical analysis for the transferred initial data set to log returns with and without a GARCH(1,1) adjustment.

Figure 1 about here

Figure 1 shows a rolling window parameter estimation for $\boldsymbol{\rho}$ for a window of size $T = 100$ in a data set of size 1861 of the Euro Stoxx 50 from 2003 until 2009. The blue line equals the ratio of the 95%-quantile of the limit distribution over the value of the test statistic of S (6).[18] Thus, the null hypothesis is rejected whenever the value of the blue line is smaller than 1. Figure 1 illustrates that in periods of economic crisis the spatial model (11) is not applicable. This is consistent with the observation that in times of bear markets the correlation among market participants

[16]The elements of this matrix are non-zero outside the diagonal and all these entries in a single row have the same value, so that it captures impacts which affect all stocks in a similar way like prior performance of stock markets.

[17]For the partitioning of the Euro Stoxx 50 members into branches and countries we refer to Table 5 in the appendix C.

[18]The results for the second test statistic S_{χ^2} are similar, so we omit them.

rises dramatically. The resulting extensive dependency structure cannot be captured by the simple spatial model (11). Accordingly, the findings of our test give evidence that the effects of the dot-com bubble crisis around 2000 last until summer 2004, since the test declines to apply model (11). In the two following years, Figure 1 depicts evidence to apply the model, since the blue line is often greater than 1. However, roughly speaking, from the beginning of 2006 until the end of the observation period the test indicates that a spatial model is inappropriate. This in accordance with the financial crisis, that started in summer of 2006.

We continue our empirical analysis by looking at Value-at-Risk (VaR) forecasts to see if our new specification test could also be used as a backtest in the spirit of Ziggel et al. (2014) among others. Figure 4 depicts the VaR forecast with standard normally distributed errors for the mean-variance optimal portfolio (see Arnold et al., 2013) based on a rolling historical window of $T = 50$. We observe that in times of moderate economic peaks (2004-2006), where the test provides statistical evidence for a spatial model, the VaR forecasts also seem to be accurate. In times of crises the test (6) rejects the null, such that both the spatial model (11) and VaR forecasts seem to be inappropriate instruments to describe the prevailing market situation. The facts that in time of a crisis both the SAR(3) is rejected and a superproportional number of VaR-exceedances are observed, are two consequences of high market volatility and cross-sectional dispersions in the stock market, such that the proposed testing procedures can also be applied as a VaR backtest.

Figure 2 about here

Figure 2 is the analogon to Figure 1 under a GARCH(1,1) adjustment and it depicts that the overall structure is in accordance with those from Figure 1. Beyond that, the GARCH(1,1) filter seems to point out the typical scope of application of spatial models, that in times of economic crisis classical SAR(m) models seem to be too restrictive and not complex enough. This is consistent with the results given in Figure 3, which shows, that VaR-forecasts are less violated in moderate economic times compared to an economic depression (cf. roughly summer 2006 until end of the data set).

Figure 3 about here

Furthermore, the amount of clusters regarding VaR violations decreases from Figure 4 compared to Figure 3, where there is less clustering of VaR violations.

Figure 4 about here

However, in bear markets clustering is still clearly observable which is in accordance with the findings that the extensive structure could not be fully captured by the spatial model (11). Overall, our empirical investigation shows the tests' ability to detect misspecifications for classical SAR models for both the initial and for a GARCH adjusted data set.

5 Conclusion

We propose two specification tests for spatial models and analyze the size and power of these tests. The proposed tests show good size and power properties in finite samples for both initial data and GARCH adjusted data. An empirical analysis of the Euro Stoxx 50 between 2003 and 2009 substantiates that bull markets provide statistical evidence to apply a SAR(3) model. However, in bear markets a simple spatial model does not capture the extensive structure of relations and dependencies in the market. Accordingly, the test provided statistical evidence for the empirical observation that both, the time after the dot-com bubble and the time around the Lehman Brothers bankruptcy could not be captured correctly by a spatial model which models only a general, branches and national dependence. For that reason it seems to be useful to introduce a test which provides statistical evidence if a given data set fulfils the assumptions of a classical SAR(m) model. To the best of the authors knowledge this is the first specification test for a classical SAR(m).

An interesting task for further research would be to see if the new specification test can be reasonably combined with the test for structural changes proposed in Wied (2013). Maybe, structural changes are a key reason for misspecification. Also, one could think about extending the ideas in this paper to extensions of the SAR model including additional exogenous regressors.

References

M. Arnold, S. Stahlberg, and D. Wied. Modelling different kinds of spatial dependence in stock returns. *Empirical Economics*, 44(2):761–774, 2013.

H. Asgharian, W. Hess, and L. Liu. A spatial analysis of international stock market linkages. *Journal of Banking and Finance*, 37(12):4738–4754, 2013.

A. Aue, S. Hörmann, L. Horvath, and M. Reimherr. Break detection in the covariance structure of multivariate time series models. *Annals of Statistics, (37(6B), 4046-4087*, 2009.

F. Blasques, S.J. Koopman, A. Lucas, and J. Schaumburg. Spillover dynamics for systemic risk measurement using spatial financial time series models. *Journal of Econometrics*, 195(2):211–223, 2016.

B. Born and J. Breitung. Simple regression based tests for spatial dependence. *Econometrics Journal*, 14(2):330–342, 2011.

J. Breitung and C. Wigger. Alternative gmm estimators for spatial regression models. *Spatial Economic Analysis*, 13:2:148–170, 2017.

L. Catania and A.G. Billé. Dynamic spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Applied Econometrics*, 32(6): 1178–1196, 2017.

M. Demetrescu and D. Wied. Testing for constant correlation of filtered series under structural change. *The Econometrics Journal*, 2018+.

M.M. Fleming. Techniques for estimating spatially dependent discrete choice models. In L. Anselin, R. Florax, and S.J. Rey, editors, *Advances in Spatial Econometrics: Methodology, Tools and Applications.* Springer, Berlin, Heidelberg, 2004.

P. Gong and Y. Weng. Value-at-Risk Forecasts by a Spatiotemporal Model in Chinese Stock Market. *Physica A: Statistical Mechanics and its Applications*, 441(C): 173–191, 2016.

L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50:1029–1054, 1982.

G. Piras J.P. Elhorst, D.J. Lacombe. On model specification and parameter space definitions in higher order spatial econometric models. *Regional Science and Urban Economics*, 42:211–220, 2012.

M. Kapoor, H. H. Kelejian, and I.R. Panel data models with spatially correlated error components. *Journal of Econometrics, 140(1):97-130*, 2007.

H.H. Kelejian and G. Piras. An extension of the j - test to a spatial panel data frameworkwork. *Journal of Applied Econometrics*, 31(2):387–402, 3 2016. ISSN 1099-1255. doi: 10.1002/jae.2425. URL `http:https://doi.org/10.1002/jae.2425`.

H.H. Kelejian and I.R. Prucha. A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40:509–533, 1999.

H.H. Kelejian and I.R. Prucha. Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157:53–67, 2010.

H. Li, C.A. Calder, and N. Cressie. Beyond Moran's I: Testing for Spatial Dependence Based on the Spatial Autoregressive Model. *Geographical Analysis*, 39(4): 357–375, 2007.

Lina Lu. Simultaneous spatial panel data models with common shocks. *FRB Boston Risk and Policy Analysis Unit Paper No. RPA*, pages 17–3, 2017.

J. G. MacKinnon. Bootstrap inference in econometrics. *The Canadian Journal of Economicss*, 35:615–645, 2002.

P.A.P. Moran. Notes on continuous stochastic phenomena. *Biometrika*, 37(1):17–23, 1950.

T. Schmitt, R. Schäfer, D. Wied, and T. Guhr. Spatial Dependence in Stock Returns - Local Normalization and VaR Forecasts. *Empirical Economics*, 50(3):1091–1109, 2016.

B. Selan and A.E.G. Kalatzis. Peer effects of stock returns and financial characteristics: Spatial approach for an emerging market. *working paper*, 2017.

L. Su and X. Qu. Specification tests for spatial autoregressive models. *Journal of Business and Economic Statistics*, 35(4):572–584, 2017.

P.S. Tam. A spatial-temporal analysis of east asian equity market linkages. *Journal of Comparative Economics*, 42(2):304–327, 2014.

D. Wied. CUSUM-type Testing for Changing Parameters in a Spatial Autoregressive Model for Stock Returns. *Journal of Time Series Analysis*, 34(1):211–229, 2013.

W.-G. Zhang, G.-L. Mo, F. Liu, and Y.-J. Liu. Value-at-risk forecasts by dynamic spatial panel gjr-garch model for international stock indices portfolio. *Soft Computing, forthcoming*, 2017.

D. Ziggel, T. Berens, G. Weiß, and D. Wied. A new set of improved value-at-risk backtests. *Journal of Banking and Finance*, 48:29–41, 2014.

A A Two Step GMM Estimation Procedure for SAR(m) Models

Given the assumptions given in 2nd section hold true. The covariance matrix of $\boldsymbol{y}_t = (I_n - \sum_{i=1}^{m} \rho_i W_i)^{-1}\boldsymbol{\varepsilon}_t$ is given by

$$\mathsf{Cov}[\boldsymbol{y}_t] \quad = \quad \left(I_n - \sum_{i=1}^{m} \rho_i W_i\right)^{-1} \Sigma \left(I_n - \sum_{i=1}^{m} \rho_i W_i'\right)^{-1} =: V.$$

For the estimation, a two step procedure is considered. First, we estimate the correlation parameters by the method of moments which does not depend on the parameters of variance. Second, we estimate the variance parameters.

The moment estimator for the correlation parameters uses the following m-moment conditions:

$$\mathsf{E}\left[\boldsymbol{\varepsilon}_t' W_i \boldsymbol{\varepsilon}_t\right] \quad = \quad \mathsf{tr}(W_i\Sigma) = 0 \qquad \text{for} \qquad i = 1, ..., m. \tag{12}$$

Clearly, the variance parameters σ_i^2 for $i = 1, ..., m$ do not enter the moment conditions. Replacing $\boldsymbol{\varepsilon}_t$ by

$$\boldsymbol{\varepsilon}_t = \left(I_n - \sum_{i=1}^{m} \rho_i W_i\right) \boldsymbol{y}_t$$

and averaging over t gives the theoretical system of equations

$$\Gamma\boldsymbol{\lambda} + \boldsymbol{\gamma} = 0,$$

where $\boldsymbol{\lambda} := \boldsymbol{\lambda}(\boldsymbol{\rho})$ is a functional vector of $\boldsymbol{\rho} := (\rho_1, ...\rho_m)$ of dimension $M := \binom{m}{1} + \binom{m+2-1}{2}$, $\binom{\cdot}{\cdot}$ denoting the binomial coefficient, such that

$$\lambda_i \quad = \rho_i \qquad \text{for} \qquad i = 1, ..., m \tag{13}$$

$$\lambda_{m+i} \quad = \rho_i^2 \qquad \text{for} \qquad i = 1, ..., m \tag{14}$$

$$\lambda_{2m+\#\{ij\,|\,i<j, i<l, j\leq k\}} \quad = \rho_l \rho_k \qquad \text{for} \qquad l, k = 1, ..., m, \tag{15}$$

where $\#\{ij\,|\,i < j, i < l, j \leq k\}$ represents the number of integer pairs ij such that the conditions $i < j, i < l$ and $j \leq k$ are fulfilled for $l, k = 1, ..., m$. The elements of

$\Gamma \in \mathbb{R}^{m\times M}$ and $\gamma \in \mathbb{R}^m$ are defined by for $i,j = 1,...,m$,

$$\Gamma_{i,j} = \mathsf{E}\left[-\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{y}_t'\left(W_i + W_i'\right)W_j\boldsymbol{y}_t\right], \tag{16}$$

$$\Gamma_{i,m+j} = \mathsf{E}\left[\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{y}_t'W_j'W_iW_j\boldsymbol{y}_t\right], \tag{17}$$

$$\Gamma_{i,2m+\#\{ij\,|\,i<j,i<l,j\leq k\}} = \mathsf{E}\left[\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{y}_t'W_l'\left(W_i + W_i'\right)W_k\boldsymbol{y}_t\right], \tag{18}$$

$$\gamma_i = \mathsf{E}\left[\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{y}_t'W_i\boldsymbol{y}_t\right].$$

Let G and $\boldsymbol{g}$ be the empirical counterparts of Γ and $\boldsymbol{\gamma}$, i.e. the expectation operator is left out. The moment estimator for $\boldsymbol{\rho} = (\rho_1,...,\rho_m)'$ is defined as

$$\hat{\boldsymbol{\rho}} := (\hat{\rho}_1,...,\hat{\rho}_m)' := \arg\min_{\boldsymbol{\rho}\in S} ||G\boldsymbol{\lambda} + \boldsymbol{g}||$$

where $||\cdot||$ represents the euclidean norm.

Remark A.1. *For $k,l \in \{1,...,m\}$, the entries of $\mathsf{E}[G] = \Gamma$ given in (16)-(18) can be calculated as*

$$\Gamma_{k,l} = \mathsf{tr}\left(\left(W_k + W_k'\right)W_lV\right),$$

$$\Gamma_{k,m+l} = \mathsf{tr}\left(W_l'W_kW_lV\right),$$

$$\Gamma_{i,2m+\#\{ij\,|\,i<j,i<l,j\leq k\}} = \mathsf{tr}\left(W_l'\left(W_i + W_i'\right)W_kV\right).$$

The following remark illustrates the results for the SAR(3) model.

Remark A.2. *For the case $m = 3$ we have to estimate the spatial vector $\boldsymbol{\rho} := (\rho_1,\rho_2,\rho_3)$. The corresponding theoretical system of equations is given by $\Gamma\boldsymbol{\lambda} + \boldsymbol{\gamma} = 0$ with $\Gamma := \left(\Gamma_{(1)},\Gamma_{(2)},\Gamma_{(3)}\right) \in \mathbb{R}^{m\times M}$, $\boldsymbol{\lambda} \in \mathbb{R}^{M\times 1}$ and $\boldsymbol{\gamma} \in \mathbb{R}^{m\times 1}$ with $M = 3 + \binom{4}{2} = 9$ which are defined as*

$$\Gamma := \mathsf{E}\left[\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{y}_t'\left(\Gamma_{(1)},\Gamma_{(2)},\Gamma_{(3)}\right)\boldsymbol{y}_t\right]$$

with

$$\Gamma_{(1)} := \begin{pmatrix} (W_1+W_1')W_1 & (W_1+W_1')W_2 & (W_1+W_1')W_3 \\ (W_2+W_2')W_1 & (W_2+W_2')W_2 & (W_2+W_2')W_3 \\ (W_3+W_3')W_1 & (W_3+W_3')W_2 & (W_3+W_3')W_3 \end{pmatrix},$$

$$\Gamma_{(2)} := \begin{pmatrix} W_1'W_1W_1 & W_2'W_1W_2 & W_3'W_1W_3 \\ W_1'W_2W_1 & W_2'W_2W_2 & W_3'W_2W_3 \\ W_1'W_3W_1 & W_2'W_3W_2 & W_3'W_3W_3 \end{pmatrix},$$

$$\Gamma_{(3)} := \begin{pmatrix} W_1'(W_1+W_1')W_2 & W_1'(W_1+W_1')W_3 & W_2'(W_1+W_1')W_3 \\ W_1'(W_2+W_2')W_2 & W_1'(W_2+W_2')W_3 & W_2'(W_2+W_2')W_3 \\ W_1'(W_3+W_3')W_2 & W_1'(W_3+W_3')W_3 & W_2'(W_3+W_3')W_3 \end{pmatrix},$$

$$\boldsymbol{\lambda} := \left(\rho_1,\rho_2,\rho_3,\rho_1^2,\rho_2^2,\rho_3^2,\rho_1\rho_2,\rho_1\rho_3,\rho_2\rho_3\right)$$

and

$$\boldsymbol{\gamma} := \mathsf{E}\left[\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{y}_t'\left(W_1',W_2',W_3'\right)'\boldsymbol{y}_t\right].$$

Since the theoretical term $\Gamma\boldsymbol{\lambda}+\boldsymbol{\gamma}$ is equal to zero for the true parameter values, the moment estimator for $\hat{\boldsymbol{\rho}}$ minimizes the corresponding empirical system $G\boldsymbol{\lambda}+\boldsymbol{g}$. Arnold et al. (2013) prove consistency and asymptotic normality of the moment estimator (cf. Theorem A.3) for $T\to\infty$, for which an additional assumption is needed.

Assumption 2.

1. *The true parameter $\boldsymbol{\rho}\in S$ is the unique solution of the theoretical system of equations, i.e.*

$$\Gamma\boldsymbol{\lambda}+\boldsymbol{\gamma}=0 \Leftrightarrow \hat{\boldsymbol{\rho}}=\boldsymbol{\rho}.$$

2. *The matrix $\mathsf{E}\left(\frac{\partial(G\boldsymbol{\lambda}+\boldsymbol{g})}{\partial\hat{\boldsymbol{\rho}}}(\boldsymbol{y}_t,\boldsymbol{\rho})\right) =: \boldsymbol{d} = \Gamma\boldsymbol{\lambda}^{(1)}$ exists, is finite and has full rank with $\boldsymbol{\lambda}^{(1)}$ a $(M\times m)$ dimensional matrix defined as*

$$\boldsymbol{\lambda}^{(1)}(l,l)=1, \qquad \boldsymbol{\lambda}^{(1)}(2m+\#\{ij\,|\,i<j,i<l,j\le k\},l)=\rho_k$$

$$\boldsymbol{\lambda}^{(1)}(m+l,l)=2\rho_l, \qquad \boldsymbol{\lambda}^{(1)}(2m+\#\{ij\,|\,i<j,i<l,j\le k\},k)=\rho_l$$

for all $l,k = 1,...,m$.

3. *For*

$$f(\boldsymbol{y}_t, \boldsymbol{\rho}) = \begin{pmatrix} \boldsymbol{\varepsilon}_t' W_1 \boldsymbol{\varepsilon}_t \\ \vdots \\ \boldsymbol{\varepsilon}_t' W_m \boldsymbol{\varepsilon}_t \end{pmatrix},$$

it holds that, for $j \to \infty$, $\mathsf{E}[f(\boldsymbol{y}_t, \boldsymbol{\rho}) \,|\, f(\boldsymbol{y}_{t-j}, \boldsymbol{\rho}), f(\boldsymbol{y}_{t-j-1}, \boldsymbol{\rho}), \ldots]$ converges in mean square to zero and that, for

$\boldsymbol{v}_j := \mathsf{E}\,[f(\boldsymbol{y}_t, \rho) \,|\, f(\boldsymbol{y}_{t-j}, \rho), f(\boldsymbol{y}_{t-j-1}, \rho), ...) - \mathsf{E}\,[f(\boldsymbol{y}_t, \rho) \,|\, f(\boldsymbol{y}_{t-j-1}, \rho), f(\boldsymbol{y}_{t-j-2}, \rho), ...]$

the infinite sum $\sum_{t=-\infty}^{\infty} \mathsf{E}[(\boldsymbol{v}_j \boldsymbol{v}_j)^{\frac{1}{2}}]$ is finite.

Under the Assumptions 1 and 2 the GMM estimator $\hat{\boldsymbol{\rho}}$ is consistent and asymptotic normal as the following theorem shows:

Theorem A.3. *Let Assumption 1 and 2 hold. Then, for $S_W = \sum_{t=-\infty}^{\infty} \mathsf{E}[f(\boldsymbol{y}_t, \boldsymbol{\rho}) f(\boldsymbol{y}_t, \boldsymbol{\rho})']$ and $T \to \infty$ it holds:*

1. $\hat{\boldsymbol{\rho}} \overset{p}{\to} \boldsymbol{\rho}$

2. $\sqrt{T}(\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}) \overset{d}{\to} N(0, \boldsymbol{d}^{-1} S_W (\boldsymbol{d}^{-1})')$.

B Proofs

Theorem 2.1 is proved by means of the following Lemmas.

Lemma B.1. *Let I_n denote the n-dimensional identity matrix and W the m-dimensional stack of spatial matrices, i.e. $W' = (W_1', \dots, W_m')$ with $W_i \in R^{n \times n}$ for $i = 1, ..., m$. Under Assumption 1 and given that $\{\boldsymbol{\varepsilon}_t\}_{t \in \{1,...,T\}}$ is serially independent the following holds for $\boldsymbol{\rho} := (\rho_1, \dots, \rho_m)$ and $\hat{\boldsymbol{\rho}} = (\hat{\rho}_1, \dots, \hat{\rho}_m)$*

$$\sqrt{T}\ \hat{\mathsf{Cov}}\ [\hat{\boldsymbol{\varepsilon}}_t] = \frac{1}{\sqrt{T}} \sum \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' + \frac{1}{T} \sum \Delta_T \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' + \frac{1}{T} \sum \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \Delta_T' + \frac{1}{T} \sum \Delta_T \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \frac{\Delta_T'}{\sqrt{T}}$$

with $\Delta_T := \sqrt{T}((\boldsymbol{\rho} - \hat{\boldsymbol{\rho}}) \otimes I_n) W (I_n - (\boldsymbol{\rho} \otimes I_n) W)^{-1}$, where $\otimes$ represents the Kronecker product.

Proof. It holds:

$$\begin{aligned}
&\sqrt{T} \hat{\mathsf{Cov}}\ [\hat{\boldsymbol{\varepsilon}}_t] = \sqrt{T} \hat{\mathsf{E}}\ [\hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t'] \\
&= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t' = \frac{1}{\sqrt{T}} \sum (I_n - (\hat{\boldsymbol{\rho}} \otimes I_n) W) \boldsymbol{y}_t \boldsymbol{y}_t' (I_n - (\hat{\boldsymbol{\rho}} \otimes I_n) W)' \\
&= \frac{1}{\sqrt{T}} \sum (I_n - (\hat{\boldsymbol{\rho}} \otimes I_n) W)(I_n - (\boldsymbol{\rho} \otimes I_n) W)^{-1} \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' [(I_n - (\hat{\boldsymbol{\rho}} \otimes I_n) W)(I_n - (\boldsymbol{\rho} \otimes I_n) W)^{-1}]' \\
&= \frac{1}{\sqrt{T}} \sum (I_n - (\boldsymbol{\rho} \otimes I_n) W + (\boldsymbol{\rho} \otimes I_n) W - (\hat{\boldsymbol{\rho}} \otimes I_n) W)(I_n - (\boldsymbol{\rho} \otimes I_n) W)^{-1} \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \\
&\qquad [(I_n - (\boldsymbol{\rho} \otimes I_n) W + (\boldsymbol{\rho} \otimes I_n) W - (\hat{\boldsymbol{\rho}} \otimes I_n) W)(I_n - (\boldsymbol{\rho} \otimes I_n) W)^{-1}]' \\
&= \frac{1}{T} \sum [\sqrt{T} I_n + \sqrt{T}((\boldsymbol{\rho} - \hat{\boldsymbol{\rho}}) \otimes I_n) W (I_n - (\boldsymbol{\rho} \otimes I_n) W)^{-1}] \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \\
&\qquad [\sqrt{T} I_n + \sqrt{T}((\boldsymbol{\rho} - \hat{\boldsymbol{\rho}}) \otimes I_n) W (I_n - (\boldsymbol{\rho} \otimes I_n) W)^{-1}]' \\
&= \frac{1}{T} \sum [\sqrt{T} I_n + \Delta_T] \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' [I_n + \frac{\Delta_T}{\sqrt{T}}]' \\
&= \frac{1}{T} \sum [\sqrt{T} \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' + \Delta_T \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'][\frac{\Delta_T'}{\sqrt{T}} + I_n] \\
&= \frac{1}{T} \sum [\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \Delta_T' + \Delta_T \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \frac{\Delta_T'}{\sqrt{T}} + \sqrt{T} \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' + \Delta_T \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'] \\
&= \underbrace{\frac{1}{\sqrt{T}} \sum \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'}_{\lim_{T \to \infty} = A} + \underbrace{\frac{1}{T} \sum \Delta_T \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'}_{\lim_{T \to \infty} = B} + \underbrace{\frac{1}{T} \sum \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \Delta_T'}_{\lim_{T \to \infty} = B'} + \underbrace{\frac{1}{T} \sum \Delta_T \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \frac{\Delta_T'}{\sqrt{T}}}_{o_p(\sqrt{T})}
\end{aligned}$$

□

The claim in Theorem 2.1 is achieved by standard arguements and an adjustment of Theorem 2.1. in Arnold et al. (2013).

Lemma B.2. *If presume the same Assumptions as in Lemma B.1 , then* $\boldsymbol{\alpha} = (A)_{i<j,i\neq j} = \left(\frac{1}{\sqrt{T}}\sum \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'\right)_{i<j,i\neq j}$ *has expectation zero and the following covariance matrix*

$$\mathsf{Cov}\,[\boldsymbol{\alpha}] = \begin{pmatrix} \lim\limits_{T\to\infty} \mathsf{Var}\,[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_{1t}\boldsymbol{\varepsilon}_{2t}] & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & \lim\limits_{T\to\infty} \mathsf{Var}\,[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_{(n-1)t}\boldsymbol{\varepsilon}_{nt}] \end{pmatrix}.$$

Proof. The zero mean statement follows directly from the cross-sectional uncorrelatedness for every $t = 1,...,T$. Furthermore, we observe

$$\begin{aligned} \mathsf{Cov}\,[\boldsymbol{\alpha}] &= \lim_{T\to\infty} \begin{pmatrix} \mathsf{Var}\,[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_{1t}\boldsymbol{\varepsilon}_{2t}] & \cdots & \frac{1}{T}\mathsf{Cov}\,[\sum \epsilon_{1t}\epsilon_{2t}\sum \epsilon_{1s}\epsilon_{ns}] \\ \vdots & \cdots & \vdots \\ \frac{1}{T}\mathsf{Cov}\,[\sum \epsilon_{1t}\epsilon_{nt}\sum \epsilon_{1s}\epsilon_{2s}] & \cdots & \mathsf{Var}\,[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_{(n-1)t}\boldsymbol{\varepsilon}_{nt}] \end{pmatrix} \\ &= \begin{pmatrix} \lim\limits_{T\to\infty} \mathsf{Var}\,[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_{1t}\boldsymbol{\varepsilon}_{2t}] & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & \lim\limits_{T\to\infty} \mathsf{Var}\,[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_{(n-1)t}\boldsymbol{\varepsilon}_{nt}] \end{pmatrix} \\ &\in \mathbb{R}^{\frac{n(n-1)}{2}\times\frac{n(n-1)}{2}}. \end{aligned}$$

$\square$

Lemma B.3. *If presume the same Assumptions as in Lemma B.1 and* $\{\boldsymbol{\varepsilon}_t\}_{t\in\{1,...,T\}}$ *being serially independent, then* $\boldsymbol{\alpha} = (A)_{i<j,i\neq j} = \left(\frac{1}{\sqrt{T}}\sum \boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\right)_{i<j,i\neq j}$ *is multivariate normally distributed with expectation zero and*

$$\mathsf{Cov}\left[\left(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\right)_{i<j,i\neq j}\right] = \mathsf{Cov}\left[\boldsymbol{\alpha}\right] = diag\left(\sigma_1^2\sigma_2^2,...,\sigma_{n-1}^2\sigma_n^2\right).$$

Proof. The vector α can be rewritten as $\sqrt{T}\left(\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\right)_{i<j,i\neq j}$. By Assumption 1.5 and the multivariate central limit theorem we obtain that $\boldsymbol{\alpha}$ is normally distributed with expectation zero. Since we assume uncorrelatedness in the cross-section for every $t = 1,...,T$, we have for $i \neq j \neq k \neq i$

$$\mathsf{Cov}\left[\varepsilon_{it}\varepsilon_{jt},\varepsilon_{it}\varepsilon_{jt}\right] = \mathsf{E}\left[\varepsilon_{it}^2\varepsilon_{jt}^2\right] - 0 = \sigma_i^2\sigma_j^2, \tag{19}$$

$$\mathsf{Cov}\left[\varepsilon_{it}\varepsilon_{jt},\varepsilon_{it}\varepsilon_{kt}\right] = \mathsf{E}\left[\varepsilon_{it}^2\varepsilon_{jt}\varepsilon_{kt}\right] - 0 = \mathsf{E}\left[\varepsilon_{it}^2\right]\mathsf{E}\left[\varepsilon_{jt}\varepsilon_{kt}\right] = 0. \tag{20}$$

Thus, the covariance matrix for the limiting normal distribution is given by

$$\begin{aligned}
\mathsf{Cov}\left[\boldsymbol{\alpha}\right] &= \begin{pmatrix} \mathsf{Cov}\left[\varepsilon_{1t}\varepsilon_{2t},\varepsilon_{1t}\varepsilon_{2t}\right] & \cdots & \mathsf{Cov}\left[\varepsilon_{1t}\varepsilon_{2t},\varepsilon_{(n-1)t}\varepsilon_{nt}\right] \\ \mathsf{Cov}\left[\varepsilon_{1t}\varepsilon_{3t},\varepsilon_{1t}\varepsilon_{2t}\right] & \cdots & \mathsf{Cov}\left[\varepsilon_{1t}\varepsilon_{3t},\varepsilon_{(n-1)t}\varepsilon_{nt}\right] \\ \vdots & \cdots & \vdots \\ \mathsf{Cov}\left[\varepsilon_{(n-1)t}\varepsilon_{nt},\varepsilon_{1t}\varepsilon_{2t}\right] & \cdots & \mathsf{Cov}\left[\varepsilon_{(n-1)t}\varepsilon_{nt},\varepsilon_{(n-1)t}\varepsilon_{nt}\right] \end{pmatrix} \\
&= \begin{pmatrix} \sigma_1^2\sigma_2^2 & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & \sigma_{(n-1)}^2\sigma_n^2 \end{pmatrix} \in \mathbb{R}^{\frac{n(n-1)}{2}\times\frac{n(n-1)}{2}}.
\end{aligned}$$

$\square$

C Tables and Figures

List of Tables and Figures

Table 1: Size and Power of S for SAR(3)

$\mathbf{n=20}$	$\zeta=0.05$				
	$\kappa=0.00$	$\kappa=0.05$	$\kappa=0.1$	$\kappa=0.15$	$\kappa=0.25$
$T=50$	0.031	0.039	0.060	0.108	0.332
$T=100$	0.039	0.056	0.956	0.213	0.742
$T=200$	0.042	0.069	0.220	0.532	0.973
$T=500$	0.034	0.114	0.576	0.943	1.00
$T=1000$	0.045	0.257	0.929	0.999	1.00
$\mathbf{n=20}$	$\zeta=0.1$				
	$\kappa=0.00$	$\kappa=0.05$	$\kappa=0.1$	$\kappa=0.15$	$\kappa=0.25$
$T=50$	0.034	0.044	0.083	0.210	0.739
$T=100$	0.038	0.067	0.173	0.526	0.984
$T=200$	0.041	0.097	0.444	0.917	1.00
$T=500$	0.034	0.219	0.944	1.00	1.00
$T=1000$	0.045	1.00	1.00	1.00	1.00
$\mathbf{n=20}$	$\zeta=0.2$				
	$\kappa=0.00$	$\kappa=0.05$	$\kappa=0.1$	$\kappa=0.15$	$\kappa=0.25$
$T=50$	0.033	0.063	0.172	0.459	0.984
$T=100$	0.036	0.089	0.415	0.888	1.00
$T=200$	0.029	0.166	0.830	1.00	1.00
$T=500$	0.037	0.508	100	1.00	1.00
$T=1000$	0.045	0.926	1.00	1.00	1.00
$\mathbf{n=50}$	$\zeta=0.05$				
	$\kappa=0.00$	$\kappa=0.05$	$\kappa=0.1$	$\kappa=0.15$	$\kappa=0.25$
$T=50$	0.048	0.056	0.148	0.366	0.949
$T=100$	0.041	0.079	0.264	0.763	1.00
$T=200$	0.051	0.141	0.716	1.00	1.00
$T=500$	0.059	0.383	1.00	1.00	1.00
$T=1000$	0.046	0.862	1.00	1.00	1.00
$\mathbf{n=50}$	$\zeta=0.1$				
	$\kappa=0.00$	$\kappa=0.05$	$\kappa=0.1$	$\kappa=0.15$	$\kappa=0.25$
$T=50$	0.048	0.075	0.281	0.758	1.00
$T=100$	0.041	0.125	0.677	0.993	1.00
$T=200$	0.051	0.304	0.990	1.00	1.00
$T=500$	0.059	0.810	1.00	1.00	1.00
$T=1000$	0.045	1.00	1.00	1.00	1.00
$\mathbf{n=50}$	$\zeta=0.2$				
	$\kappa=0.00$	$\kappa=0.05$	$\kappa=0.1$	$\kappa=0.15$	$\kappa=0.25$
$T=50$	0.049	0.104	0.608	0.988	1.00
$T=100$	0.041	0.244	0.974	1.00	1.00
$T=200$	0.051	0.602	1.00	1.00	1.00
$T=500$	0.059	0.997	1.00	1.00	1.00
$T=1000$	0.046	1.00	1.00	1.00	1.00

Power and Size Analysis of the Test (3) with $\rho=(0.45,0.3,0.15)\in\mathbb{R}^3$. The DGP follows a multivariate normal distribution where ζ describes the expected portion of pairs that are correlated with each other with correlation κ^2 and variance $\sigma_i^2=2$ for all $i\in\{1,...,n\}$. The amount of draws from the limit distribution is $B=400$ by 701 Monte Carlo repetitions.

Table 2: Size and Power of S_χ for SAR(3)

n = 20	$\zeta = 0.05$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.028	0.025	0.040	0.057	0.220
$T = 100$	0.042	0.030	0.072	0.175	0.679
$T = 200$	0.038	0.057	0.158	0.503	0.965
$T = 500$	0.047	0.121	0.567	0.988	1.00
$T = 1000$	0.055	0.233	0.922	0.998	1.00
n = 20	$\zeta = 0.1$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.028	0.037	0.045	0.092	0.561
$T = 100$	0.042	0.057	0.133	0.414	0.987
$T = 200$	0.038	0.060	0.384	0.912	1.00
$T = 500$	0.047	0.238	0.937	1.00	1.00
$T = 1000$	0.055	0.591	0.998	1.00	1.00
n = 20	$\zeta = 0.2$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.028	0.033	0.060	0.253	0.954
$T = 100$	0.043	0.063	0.346	0.844	1.00
$T = 200$	0.039	0.113	0.831	0.997	1.00
$T = 500$	0.047	0.483	1.00	1.00	1.00
$T = 1000$	0.055	0.894	1.00	1.00	1.00
n = 50	$\zeta = 0.05$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.005	0.014	0.016	0.065	0.709
$T = 100$	0.018	0.027	0.156	0.601	1.00
$T = 200$	0.030	0.771	0.617	0.991	1.00
$T = 500$	0.033	0.369	0.998	1.00	1.00
$T = 1000$	0.047	0.829	1.00	1.00	1.00
n = 50	$\zeta = 0.1$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.005	0.010	0.0411	0.330	0.995
$T = 100$	0.018	0.047	0.045	0.989	1.00
$T = 200$	0.023	0.164	0.982	1.00	1.00
$T = 500$	0.033	0.773	1.00	1.00	1.00
$T = 1000$	0.047	0.999	1.00	1.00	1.00
n = 50	$\zeta = 0.2$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.005	0.013	0.203	0.900	0.997
$T = 100$	0.018	0.106	0.927	1.00	1.00
$T = 200$	0.030	0.435	1.00	1.00	1.00
$T = 500$	0.033	0.997	1.00	1.00	1.00
$T = 1000$	0.047	1.00	1.00	1.00	1.00

Power and Size Analysis of the Test (8) S_χ where ζ describes the expected portion of pairs that are correlated with each other with correlation κ^2 and variance $\sigma_i^2 = 2$ for all $i \in \{1, ..., n\}$.

Table 3: Size and Power of S for SAR(4)

$\mathbf{n=60}$	$\zeta=0.05$				
	$\kappa=0.00$	$\kappa=0.05$	$\kappa=0.1$	$\kappa=0.15$	$\kappa=0.25$
$T=50$	0.036	0.049	0.073	0.134	0.438
$T=100$	0.043	0.050	0.114	0.263	0.805
$T=200$	0.036	0.069	0.227	0.640	0.989
$T=500$	0.035	0.117	0.681	0.979	1.00
$T=1000$	0.050	0.329	0.960	0.997	1.00
$\mathbf{n=60}$	$\zeta=0.1$				
	$\kappa=0.00$	$\kappa=0.05$	$\kappa=0.1$	$\kappa=0.15$	$\kappa=0.25$
$T=50$	0.036	0.061	0.096	0.251	0.782
$T=100$	0.043	0.084	0.208	0.629	0.991
$T=200$	0.036	0.097	0.509	0.939	1.00
$T=500$	0.036	0.270	0.976	1.00	1.00
$T=1000$	0.050	0.684	1.00	1.00	1.00
$\mathbf{n=60}$	$\zeta=0.2$				
	$\kappa=0.00$	$\kappa=0.05$	$\kappa=0.1$	$\kappa=0.15$	$\kappa=0.25$
$T=50$	0.036	0.074	0.191	0.535	0.988
$T=100$	0.043	0.103	0.479	0.949	1.00
$T=200$	0.036	0.193	0.919	1.00	1.00
$T=500$	0.036	0.604	1.00	1.00	1.00
$T=1000$	0.050	0.962	1.00	1.00	1.00
$\mathbf{n=90}$	$\zeta=0.05$				
	$\kappa=0.00$	$\kappa=0.05$	$\kappa=0.1$	$\kappa=0.15$	$\kappa=0.25$
$T=50$	0.029	0.059	0.080	0.191	0.685
$T=100$	0.054	0.064	0.176	0.461	0.972
$T=200$	0.044	0.101	0.398	0.893	1.00
$T=500$	0.046	0.214	0.930	1.00	1.00
$T=1000$	0.047	0.551	1.00	1.00	1.00
$\mathbf{n=90}$	$\zeta=0.1$				
	$\kappa=0.00$	$\kappa=0.05$	$\kappa=0.1$	$\kappa=0.15$	$\kappa=0.25$
$T=50$	0.029	0.074	0.161	0.431	0.942
$T=100$	0.054	0.080	0.382	0.853	1.00
$T=200$	0.044	0.164	0.810	1.00	1.00
$T=500$	0.046	0.503	0.997	1.00	1.00
$T=1000$	0.047	0.930	1.00	1.00	1.00
$\mathbf{n=90}$	$\zeta=0.2$				
	$\kappa=0.00$	$\kappa=0.05$	$\kappa=0.1$	$\kappa=0.15$	$\kappa=0.25$
$T=50$	0.029	0.094	0.365	0.806	0.999
$T=100$	0.054	0.127	0.752	0.999	1.00
$T=200$	0.044	0.338	0.997	1.00	1.00
$T=500$	0.046	0.888	1.00	1.00	1.00
$T=1000$	0.047	0.999	1.00	1.00	1.00

Power and Size Analysis of the test S with $\boldsymbol{\rho} = (-0.2\ 0.05\ 0.1\ 0.5)$. The errors are heteroscedastic, i.e. $\sigma_i \sim N(0,1)$, $i = 1,...n$. The parameter ζ describes the portion of expected pairs of firms that are correlated to each other with correlation intensity κ^2.

Table 4: Size and Power of S under GARCH model for SAR(3)

$\mathbf{n = 50}$	$\zeta = 0.02$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 1000$	0.086	0.118	0.211	0.490	0.960
$T = 1500$	0.078	0.128	0.331	0.719	0.996
$T = 2000$	0.062	0.140	0.459	0.906	1.00
$T = 2500$	0.068	0.156	0.565	0.960	1.00
$T = 3000$	0.044	0.114	0.673	0.988	1.00
$\mathbf{n = 50}$	$\zeta = 0.04$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 1000$	0.086	0.126	0.440	0.881	1.00
$T = 1500$	0.078	0.178	0.699	0.986	1.00
$T = 2000$	0.062	0.156	0.872	1.00	1.00
$T = 2500$	0.068	0.250	0.936	1.00	1.00
$T = 3000$	0.044	0.315	0.972	1.00	1.00
$\mathbf{n = 50}$	$\zeta = 0.1$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 1000$	0.086	0.253	0.932	1.00	1.00
$T = 1500$	0.078	0.425	0.994	1.00	1.00
$T = 2000$	0.062	0.520	1.00	1.00	1.00
$T = 2500$	0.068	0.711	1.00	1.00	1.00
$T = 3000$	0.044	0.792	1.00	1.00	1.00
$\mathbf{n = 80}$	$\zeta = 0.02$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 1000$	0.073	0.129	0.334	0.810	1.00
$T = 1500$	0.070	0.126	0.518	0.964	1.00
$T = 2000$	0.043	0.143	0.771	0.997	1.00
$T = 2500$	0.060	0.206	0.877	1.00	1.00
$T = 3000$	0.050	0.229	0.954	1.00	1.00
$\mathbf{n = 80}$	$\zeta = 0.04$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 1000$	0.073	0.128	0.709	0.998	1.00
$T = 1500$	0.070	0.202	0.954	1.00	1.00
$T = 2000$	0.043	0.291	0.990	1.00	1.00
$T = 2500$	0.060	0.409	1.00	1.00	1.00
$T = 3000$	0.050	0.517	1.00	1.00	1.00
$\mathbf{n = 80}$	$\zeta = 0.1$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 1000$	0.073	0.416	0.998	1.00	1.00
$T = 1500$	0.070	0.607	1.00	1.00	1.00
$T = 2000$	0.043	0.826	1.00	1.00	1.00
$T = 2500$	0.060	0.972	1.00	1.00	1.00
$T = 3000$	0.050	0.988	1.00	1.00	1.00

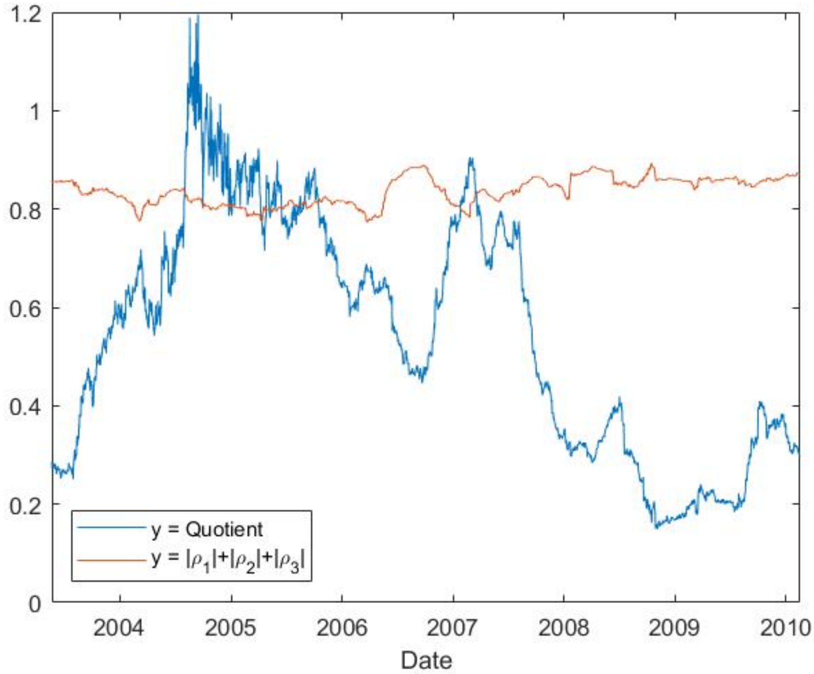
Power and Size Analysis of the Test (3) S^* with $\rho = (0.45, 0.3, 0.15)$ under a GARCH model. The data generating process is GARCH(1,1) with constant and GARCH parameter equal to 0.33 and ARCH parameter equal to 0.075 with standard normal errors. ζ describes the expected portion of pairs that are correlated with each other with correlation κ^2 and variance $\sigma_i^2 = 2$ for all $i \in \{1, ..., n\}$.

Table 5: Partitioning of Euro Stoxx 50 members into branches and countries.

Finance	Aegon, Allianz, AXA, Banco Bilbao, Banco Santander, BNP, Crédit Agricole, Deutsche Bank, Deutsche Börse, Generali, ING, Intesa, Münchener Rück, Société Générale, Unicredit
Automobil	Daimler, Renault, VW
Energy	Alstom, E.ON, ENEL, ENI, Iberdrola, Repsol, RWE, SUEZ, Total
Telecom and Media	Deutsche Telekom, France Telecom, Telecom Italia, Telefonica, Vivendi
Pharma and Chemicals	Air Liquide, BASF, Bayer, Sanofi
Consumer Electronics	Nokia, Philips, SAP, Siemens, Schneider
Consumer retail	Anheuser Busch, Carrefour, Danone, L'Oreal, LVMH, Unilever
Basic Industry	Arcelor Mittal, CRH, Saint Gobain, Vinci
Benelux	Aegon, Anheuser Busch, Arcelor, ING, Philips, Unilever
France	Air Liquide, Alstom, AXA, BNP, Carrefour, Crédit Agricole, France Telecom, Danone, L'Oreal, LVMH, Saint Gobain, Sanofi, Schneider, Société Générale, SUEZ, Total, Vinci, Vivendi
Germany	Allianz, BASF, Bayer, Daimler, Deutsche Bank, Deutsche Börse, Deutsche Telekom, E.ON, Münchner Rück, RWE, SAP, Siemens, VW
Italy	Generali, ENEL, ENI, Intesa, Telecom Italia, Unicredito
Spain	Banco Bilbao, Banco Santander, Iberdrola, Repsol, Telefonica
Others	CRH, Nokia

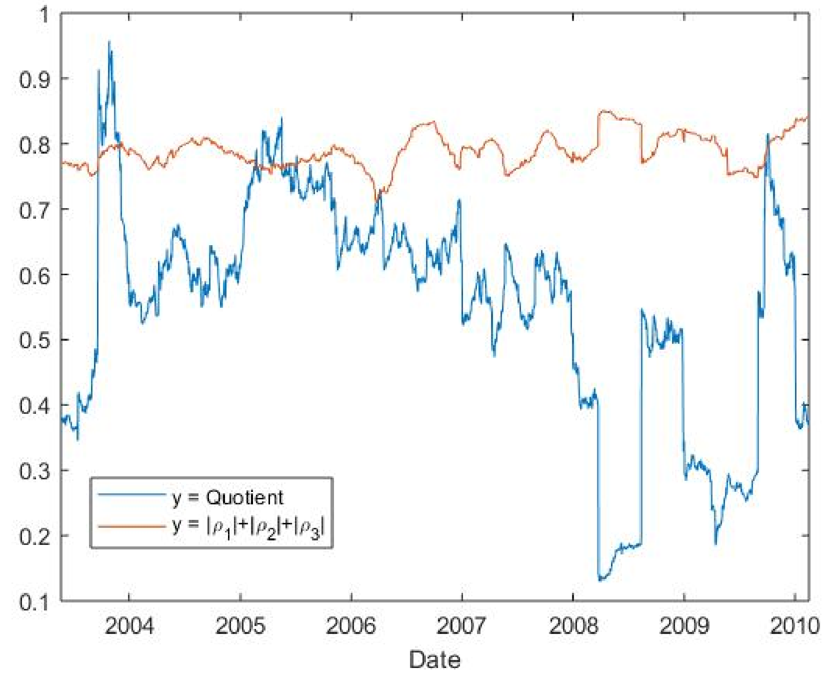
Both matrices are constructed in the following way: The off-diagonal elements are nonzero if the corresponding stocks belong to the same branch (W_b) or country (W_l). In each row, the nonzero entries are identical and sum up to 1 (row-wise).

Figure 1: Rolling Window for $T = 100$



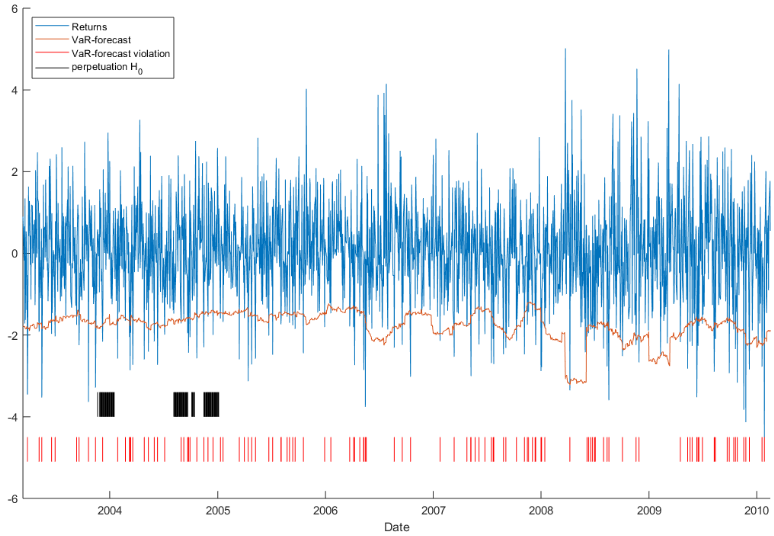
Rolling window parameter estimation for $\boldsymbol{\rho}$ for a window of size $T = 100$ in a data set of size $T = 1861$ and dimension $n = 50$. The number of Bootstrap repetitions is equal to 400. The blue line depicts the ratio of the 95%-quantile of the limit distribution given in Lemma (2.4) over the test statistic S from (3). The orange line is the accumulated spatial dependence parameter $\boldsymbol{\rho}$ within the L_1-norm.

Figure 2: Rolling Window for $T = 100$ under GARCH(1,1) adjustment



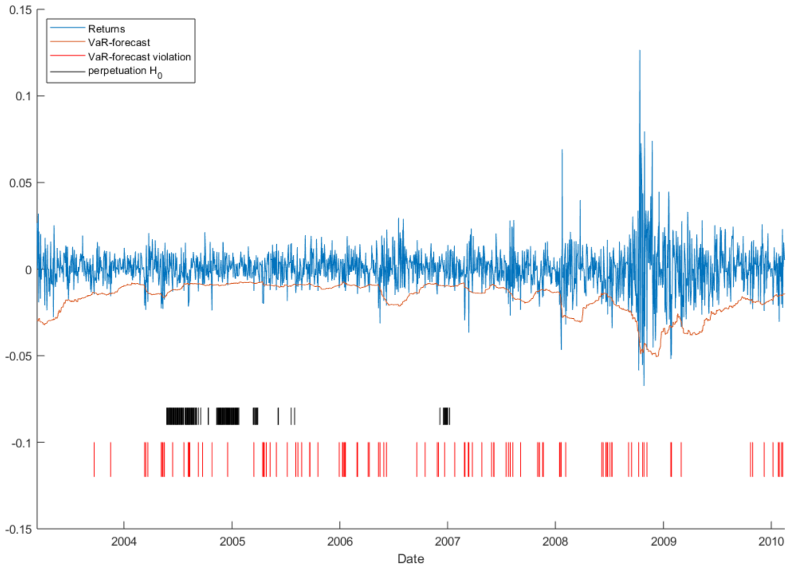
Rolling window parameter estimation for $\boldsymbol{\rho}$ for a window of size $T = 100$ in a data set of size $T = 1861$ and dimension $n = 50$ under GARCH(1,1) adjustment. The number of Bootstrap repetitions is equal to 400. The blue line depicts the ration of the 95%-quantile of the limit distribution given in Lemma (2.4) over the test statistic S from (3). The orange line is the accumulated spatial dependence parameter $\boldsymbol{\rho}$ within the L_1-norm.

Figure 3: VaR-Forecasts for $T = 50$ under GARCH(1,1) adjustment



The figure depicts a Value-at-Risk forecast with standard normal distributed errors based on the Euro Stoxx 50 from 2003 to 2009 in a rolling window of size 50 under a GARCH(1,1) adjustment. The orange line represents the VaR-forecast based on the data. The blue line represents the returns. A VaR violation is reported with a red dashed line at the bottom of that figure. The black line indicates statistical significance to apply the spatial model (1). The VaR-level is chosen at 0.05 and the number of Bootstrap repetitions is equal to 400.

Figure 4: VaR-Forecasts for $T = 50$



The figure depicts a Value-at-Risk forecast with standard normal distributed errors based on the Euro Stoxx 50 from 2003 to 2009 in a rolling window of size 50. The orange line represents the VaR-forecast based on the data. The blue line represents the returns. A VaR violation is reported with a red dashed line at the bottom of that figure. The black line indicates statistical significance to apply the spatial model (1). The VaR-level is chosen at 0.05 and the number of Bootstrap repetitions is equal to 400.