Spillover Dynamics for Systemic Risk Measurement Using Spatial Financial Time Series Models

Francisco Blasques[a], Siem Jan Koopman[a,b], Andre Lucas[a], Julia Schaumburg[a]

[a] Vrije Universiteit Amsterdam and Tinbergen Institute, The Netherlands

[b] CREATES, Aarhus University, Denmark

Abstract

We extend the well-known static spatial Durbin model by introducing a time-varying spatial dependence parameter. The updating steps for this model are functions of past data and have information theoretic optimality properties. The static parameters are conveniently estimated by maximum likelihood. We establish the theoretical properties of the model and show that the maximum likelihood estimators of the static parameters are consistent and asymptotically normal. Using spatial weights based on cross-border lending data and European sovereign CDS spread data over the period 2009–2014, we find evidence of contagion in terms of high, time-varying spatial spillovers in the perceived credit riskiness of European sovereigns during the sovereign debt crisis. We find a particular downturn in spatial dependence in the second half of 2012 after the outright monetary transactions policy measures taken by the European Central Bank. Earlier non-standard monetary operations by the ECB did not induce such changes. The findings are robust to a wide range of alternative model specifications.

Keywords: Spatial correlation, time-varying parameters, systemic risk, European debt crisis, generalized autoregressive scores.

JEL classification: C58, C23, G15

1 Introduction

We propose a new parsimonious model to measure the time-varying cross-sectional dependence in European sovereign credit spread changes in order to investigate the effectiveness of non-standard monetary operations by the ECB in reducing contagion concerns during the European sovereign debt crisis. The model builds on the well-known spatial Durbin model for panel data. The strength of contemporaneous spillover effects is summarized in a single time-varying parameter: the spatial dependence parameter. We argue that this parameter may be interpreted as a measure of sovereign systemic risk that relates to the connectedness of the system in a similar way as the unconditional correlations of Forbes and Rigobon (2002). The changes in the dependence parameter can thus be labeled as contagion in the technical sense of Forbes and Rigobon (2002).

Our paper contributes to two strands of literature. First, we contribute to the applied spatial econometrics literature. Spatial models have been widely used in applied geographic and regional science studies, and have recently also been applied in empirical finance; see Fernandez (2011) for a CAPM model augmented by spatial dependencies, Wied (2013), Arnold et al. (2013), Kelly et al. (2013), and Asgharian et al. (2013) for analyses of spatial dependencies in stock markets, Denbee et al. (2014) for a network approach to assess interbank liquidity, and Saldias (2013) for a spatial error model to identify sector risk determinants. Keiler and Eder (2015) and Tonzer (2015) both use spatial lag models, to model CDS spreads of financial institutions and banking sector risks, respectively.

The above models, however, treat the spatial dependence parameter as static. To the best of our knowledge, explicitly endowing the spatial dependence parameter in the spatial Durbin model with time series dynamics is a new development. Allowing for such dynamics may be important empirically; see for example our financial systemic stability application in Section 5. We model the dynamics using the score-driven framework proposed by Creal et al. (2011, 2013) and Harvey (2013). Given the nonlinear impact of the time-varying parameter in the model, the theoretical properties of this model and the asymptotic properties of the maximum likelihood estimator (MLE) for the remaining static parameters are challenging and have not been established so far. We show under what conditions the filtered spatial dependence parameters are well behaved, such that the model is invertible. Invertibility is a key property for establishing consistency and asymptotic normality of the MLE; see for example Wintenberger (2013). We derive new conditions for the asymptotic properties of the MLE compared to Blasques et al. (2014), allowing for exogenous regressors to be part of the specification. We also discuss the information theoretic optimality of the model and illustrate in a simulation study that the model is able to track a range of different patterns for the time-varying spatial dependence parameter.

Second, we contribute to the literature that studies the dynamics of financial systemic risk in the context of a network of sovereigns or financial firms. Since the beginning of the European sovereing debt crisis in 2009, the sharp increases and comovements of sovereign credit spreads have been the subject of a growing number of empirical studies in finance. For instance, by employing an asset pricing model, Ang and Longstaff (2013) investigate the differences between U.S. and European credit default swap (CDS) spreads as a reflection of systemic risk. Lucas et al. (2014) and Kalbaska and Gatkowski (2012) use multivariate time series models to model comovements in European sovereign CDS spreads. Ait-Sahalia et al. (2014) model sovereign credit default intensities using multivariate jump processes. De Santis (2012) and Arezki et al. (2011) study credit risk spillover effects that are induced by rating events, such as downgrades of Greek government bonds. Leschinski and Bertram (2013) find contagion effects in European sovereign bond spreads using the simultaneous equations approach of Pesaran and Pick (2007). Caporin et al. (2013), on the other hand, employ Bayesian quantile regressions, and conclude that comovements in European credit spreads during the debt crisis are only due to increased volatities, but not contagion.

Our approach differs from the studies above since we introduce cross-sectional correlation not only through contemporaneous error correlations, but also through spillovers induced by shocks to the regressors, such as stock market crashes or interbank lending rates. Furthermore, we explicitly offer financial sector linkages as the source of sovereign credit risk comovements. This view is supported by the results of Korte and Steffen (2015), Kallestrup et al. (2016), Gorea and Radev (2014), and Beetsma et al. (2012), in which cross-border exposures between international financial sectors are relevant drivers of sovereign credit spreads. By exploiting these debt interconnections as economic distances between sovereigns in our spatial model, we obtain a scalar time-varying (spatial) dependence coefficient. We interpret this parameter in the systemic context as the overall tendency for shock spillovers. Such changes to spillovers are directly linked to contagion as defined in the technical sense of Forbes and Rigobon (2002). As such, the spatial dependence coefficient provides a measure of changes in systemic risk and the market's perception of contagion within the euro area.

We organize the remainder of this paper as follows. Section 2 introduces our spatial score model with time-varying parameters, formulates the information theoretic optimality properties of the steps, and establishes the consistency and asymptotic normality of the maximum likelihood estimator. In Section 3, we provide Monte Carlo evidence of the model's ability to track different dynamic patterns in spatial dependence over time. Section 4 describes the data for our study on European sovereign CDS spread dynamics. Section 5 provides the results for our main model, its extensions and some alternative specifications. Section 6 concludes.

2 Spatial models with dynamic spatial dependence

2.1 Static spatial model for panel data

The Spatial Durbin Model (SDM) for panel data is given by

$$y_t = \rho W y_t + \beta_1 \mathbf{1}_n + A_t\beta_2 + W\, A_t\beta_3 + e_t, \qquad e_t \sim p_e(e_t; \Sigma, \lambda), \qquad t = 1, \ldots, T, \tag{1}$$

where $y_t = (y_{1t}, \ldots, y_{nt})'$ denotes a vector of n cross-sectional observations at time t, ρ is the spatial dependence coefficient, W is an $n \times n$ matrix of exogenous spatial weights, β_1 is an unknown scalar intercept, $\mathbf{1}_n$ is an $n \times 1$-vector of ones, A_t is an $n \times k$ matrix of exogenous regressors, β_2 and β_3 are $k \times 1$ vectors of unknown coefficients, respectively,[1] and e_t is an $n \times 1$ disturbance vector with multivariate density $p_e(e_t, \Sigma; \lambda)$, mean zero, unknown $k \times k$ covariance (or scale) matrix Σ, and other parameters describing the shape of the distribution are collected in the parameter vector λ. For example, if p_e is a Student's t distribution, λ contains the degrees of freedom parameter.

Model (4) implies that each entry y_{it}, for $i = 1, \ldots, n$, of the vector y_t depends on the other entries y_{jt}, for $j \neq i$. For a moderately large n, we cannot estimate such a system of contemporaneous dependencies without imposing further restrictions. The idea of a spatial dependence model is to specify the spatial weight matrix W as a function of geographic or economic distances, and in this way exogenously define a neighborhood structure between the cross-sectional units. It is stan-

[1] Here, we assume that A_t only contains individual-specific regressors. In our empirical application, we also consider regressors that are common to all units. In this case, to avoid multicollinearity due to the row-normalization of W, $W\, A_t$ only comprises the subset of individual-specific regressors.

dard practice to use a row-normalized weight matrix W such that $\sum_{j=1}^{n} w_{ij} = 1$ for $i = 1, \ldots, n$, where w_{ij} is the (i, j)th element of W. The impact of the (spatially weighted) contemporaneous dependent variables Wy_t on y_t is captured by a scalar spatial dependence parameter ρ. For shocks to die out over space, we require $\rho \in (1/\omega_{min}, 1)$ where ω_{min} is the smallest eigenvalue of W; see for example Lee (2004).

In addition to the spatial lag of the dependent variable, the Spatial Durbin Model (1) features spatial lags of the individual-specific regressors. This implies that each panel unit's dependent variable may react to shocks to the regressor(s) of its neighboring units. The model formulation not only nests the widely used Spatial Lag Model (SLM) for $\beta_2 = 0$, it is also the reduced form of a model with spatial dependence in the error term, the so-called Spatial Error Model (SEM). The SEM has the form

$$y_t = \gamma_1 \mathbf{1}_n + A_t\gamma_2 + u_t, \quad u_t = \delta W u_t + e_t. \tag{2}$$

where γ_1 and δ are unknown scalars, γ_2 is an unknown coefficient vector and e_t is defined as above. The model can be rewritten as

$$y_t = \delta W y_t + \widetilde{\gamma}_1 \mathbf{1}_n + (\mathrm{I}_n - \delta W) A_t \gamma_2 + e_t \tag{3}$$

with $\widetilde{\gamma}_1 = \gamma_1(\mathrm{I}_n - \delta W)$, which is a SDM model with $\beta_2 = \gamma_2$ and parameter restriction $\beta_3 = -\delta\gamma_2$, see also LeSage and Pace (2008).

In the following, we write the SDM as

$$y_t = \rho W y_t + X_t\beta + e_t \tag{4}$$

with $X_t := (\mathbf{1}_n : A_t : WA_t)$ and $\beta := (\beta_1, \beta_2', \beta_3')'$. It can be shown that this basic form can capture nonlinear feedback effects across units by rewriting it as

$$y_t = Z X_t \beta + Z e_t, \tag{5}$$

where we assume that the inverse matrix $Z = (\mathrm{I}_n - \rho W)^{-1}$ exists, with I_n denoting the $n \times n$ identity matrix. Using an infinite power series expansion as in LeSage and Pace (2008), we obtain

$$y_t = X_t\beta + \rho W X_t\beta + \rho^2 W^2 X_t\beta + \cdots + e_t + \rho W e_t + \rho^2 W^2 e_t + \cdots . \qquad (6)$$

Equation (6) reveals that e_{it} and $x'_{it}\beta$ for unit i spill over to other units $j \neq i$. The extent of spillover depends on the relative proximity of j to i via the weight matrix W and the spatial dependence parameter ρ. At the same time, there are possible feedback effects back to unit i itself, for example if w_{ij} and w_{ji} are both non-zero, such that i and j are mutual neighbors, and i is a 'second-order neighbor' to itself.

The simultaneous equations structure of (4) leads to an endogeneity problem and causes the least squares estimator in (4) to be inconsistent, As an alternative solution, we can estimate the parameters by the method of Maximum Likelihood (ML) or Quasi-ML (QML) where the latter is typically based on the normal distribution.[2] The ML Estimator (MLE) for spatial models with static dependence parameter was first studied in Ord (1975) in the context of cross-sectional data sets. Lee (2004) derives asymptotic properties of the QML Estimator (QMLE) for $n \to \infty$, and Hillier and Martellosio (2013) investigate its finite sample distribution. Large n and large T asymptotics for the QMLE of the spatial model with static dependence parameter are studied in Yu et al. (2008). For further textbook treatments, we refer to Anselin (1988) and LeSage and Pace (2008). For a survey on the panel data spatial lag model and parameter estimation, see Lee and Yu (2010).

2.2 Score dynamics for the spatial dependence parameter

We can interpret the spatial dependence parameter ρ in (4) as a measure of the strength of cross-sectional spillovers. In many empirical applications involving panel data, it is unrealistic to assume that ρ is constant over the entire sample period. We therefore introduce a time-varying spatial dependence parameter ρ_t in the model, that is

$$y_t = \rho_t W y_t + X_t\beta + e_t, \qquad e_t \sim p_e(e_t; \Sigma, \lambda), \qquad t = 1, \ldots, T, \qquad (7)$$

where $\rho_t = h(f_t)$ is a monotonic transformation of a time-varying parameter f_t. Time-variation in ρ_t has been at the core of attention in financial economics. In particular, the line of literature starting with Forbes and Rigobon (2002) states that changes in spillovers (as for example picked

[2] Alternatively, we can use GMM as in, for example, Kelejian and Prucha (2010).

up by changes in ρ_t) are much better measures of financial contagion than are pairwise correlations. We choose the link function h such that $\rho_t \in (-1, 1)$. To describe the dynamics of f_t, we adopt the autoregressive score framework of Creal et al. (2011, 2013) and Harvey (2013). The score framework for time-varying parameters has been adopted successfully in a range of different empirical settings, including the multivariate volatility model of Creal et al. (2011), the systemic risk model of Oh and Patton (2016) and Lucas et al. (2014), the credit risk dynamic factor model of Creal et al. (2014), and the location and scale models with fat tails of Harvey and Luati (2014).[3]

The score framework centers around the use of the scaled score of the conditional density p_e to drive the time-variation in f_t. The updating equation for f_t is given by

$$f_{t+1} = \omega + A s_t + B f_t, \tag{8}$$

where ω, A, and B are fixed unknown parameters, and $s_t = S_t \nabla_t$ is the scaled score function. The scaled score function is defined as the first derivative of the predictive log-likelihood function at time t with respect to f_t, possibly multiplied by some local scaling factor S_t. In our case, the score function is given by $\nabla_t = (\partial \ell_t / \partial \rho_t) \cdot (\partial h(f_t)/\partial f_t)$ with $\rho_t = h(f_t)$, where

$$\ell_t = \log p_e \left(y_t - \rho_t W y_t - X_t \beta, \Sigma; \lambda \right) + \log \left| (\mathrm{I}_n - \rho_t W) \right|. \tag{9}$$

Throughout this paper, we use unit scaling, that is $S_t \equiv 1$ such that $s_t = \nabla_t$. Other scaling choices are also feasible; see Creal et al. (2013).[4] Equation (9) differs from the likelihood of a simple linear regression model by the term $\log |(\mathrm{I}_n - h(f_t)W)|$. This term accounts for the nonlinearity of the model in ρ_t as shown in equation (5). We define the vector of static parameters $\boldsymbol{\theta} = (\omega, A, B, \beta, \lambda)'$ and estimate $\boldsymbol{\theta}$ via the numerical maximization of the likelihood function

$$\mathcal{L}_T = \sum_{t=1}^{T} \ell_t. \tag{10}$$

We consider two specifications for the disturbance density p_e, namely the multivariate normal distribution and the multivariate Student's t distribution. The latter is particularly relevant for our empirical study because changes in credit default swap (CDS) spreads may be fat-tailed. Also,

[3] See `www.gasmodel.com` for a more complete compilation of papers.

[4] In a simulation (not reported here) we show that different choices of scaling, such as scaling by the inverse information matrix scaling or by its square root, did not have much impact on our empirical results.

Creal et al. (2011) and Harvey and Luati (2014) argue that the Student's t distribution can render the dynamics more robust to incidental influential observations and outliers.

Using the standard expression for the multivariate normal density, we obtain the time t contribution to the log-likelihood function as

$$\begin{aligned}\ell_t &= \log|\mathrm{I} - h(f_t)W| - \frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma| \\ &\quad -\frac{1}{2}(y_t - h(f_t)Wy_t - X_t\beta)'\Sigma^{-1}(y_t - h(f_t)Wy_t - X_t\beta),\end{aligned}$$

and the resulting score

$$\nabla_t = \big(y_t'W'\Sigma^{-1}(y_t - h(f_t)Wy_t - X_t\beta) - \mathrm{tr}(Z(f_t)W)\big)\cdot \dot{h}(f_t), \tag{11}$$

where $\mathrm{tr}(\cdot)$ is the trace operator, $Z(f_t) = (\mathrm{I}_n - h(f_t)W)^{-1}$, and $\dot{h}(f_t)$ is the first derivative of the transformation function h with respect to f_t. For instance, if $h(f_t) = \gamma\tanh(f_t)$ with $\gamma \in (0,1)$, then $\dot{h}(f_t) = \gamma(1 - \tanh^2(f_t))$. When the density of the disturbance vector e_t is a multivariate Student's t distribution with λ degrees of freedom, we obtain

$$\begin{aligned}\ell_t &= \log|Z(f_t)^{-1}| + \log\left(\frac{\Gamma\left(\frac{\lambda+n}{2}\right)}{|\Sigma|^{1/2}(\lambda\pi)^{n/2}\Gamma\left(\frac{\lambda}{2}\right)}\right) \\ &\quad - \left(\frac{\lambda+n}{2}\right)\log\left(1 + \frac{(y_t - h(f_t)Wy_t - X_t\beta)'\Sigma^{-1}(y_t - h(f_t)Wy_t - X_t\beta)}{\lambda}\right),\end{aligned}$$

with the corresponding score function

$$\begin{aligned}\nabla_t &= \big(\tilde{w}_t\cdot y_t'W'\Sigma^{-1}(y_t - h(f_t)Wy_t - X_t\beta) - \mathrm{tr}(Z(f_t)W)\big)\cdot\dot{h}(f_t), \qquad (12)\\ \tilde{w}_t &= (1+\lambda^{-1}n)/\big(1+\lambda^{-1}(y_t - h(f_t)Wy_t - X_t\beta)'\Sigma^{-1}(y_t - h(f_t)Wy_t - X_t\beta)\big).\end{aligned}$$

Is is easy to verify that for $\lambda \to \infty$ we obtain $\tilde{w}_t \to 1$. The score expression in (12) in that case collapses to the one in (11). The weight $\tilde{w}_t$ is small if the residuals $y_t - h(f_t)Wy_t - X_t\beta$ are 'large' in a multivariate sense. The implication of a small weight $\tilde{w}_t$ is that the observation has a smaller impact on the updates of f_t. This provides a robustness feature to the dynamics of f_t if we assume a fat-tailed distribution such as the Student's t; see also the discussion in Creal et al. (2011, 2013) and Harvey (2013). The intuition is straightforward: a large residual may be attributable to the fat-tailedness of the Student's t distribution rather than to a recent increase in the spatial correlation parameter $\rho_t = h(f_t)$.

The score expressions in (11) and (12) also depart from the expressions for the standard linear regression model. In particular, the additional correction term $-\mathrm{tr}(Z(f_t)W)$ accounts for the simultaneity bias in the standard least squares estimator and follows from the presence of the term $\log|Z(f_t)^{-1}|$ in the likelihood at time t. Economically, this term accounts for the fact that there may be feedback effects from unit i to unit j and then back to unit i. Hence the spatial autoregressive score model integrates time-varying direct and indirect effects; both are used to determine the appropriate transition dynamics for ρ_t.

2.3 Optimality of score updating in the time-varying spatial model

The score-driven framework may provide an intuitively and statistically appealing way to update the time-varying spatial dependence parameter ρ_t. But possibly more importantly, the score based updates have also optimal properties in an information theoretic sense under very mild regularity conditions. This was proven in a generic setting by Blasques et al. (2015). To understand the issue for our particular time-varying spatial dependence model, we repeat the main argument of Blasques et al. (2015) for our specific setting.

Let $p_t := p(\,\cdot\,|X_t)$ denote the true unknown conditional density of y_t. Similarly, let $\tilde{p}_t := \tilde{p}(\,\cdot\,|\tilde{f}_t, X_t)$ denote the conditional density implied by the score model given the filtered time-varying parameter $\tilde{f}_t$, the regressors X_t, the postulated innovation density p_e, and the static parameter vector $\boldsymbol{\theta}$. Ideally, whenever a new observation y_t becomes available, we want the filtered value $\tilde{f}_{t+1}$ to be such that the new conditional density implied by the model $\tilde{p}_{t+1} := \tilde{p}(\cdot|\tilde{f}_{t+1}, X_t)$ is as close as possible to the true unknown conditional density p_t from which y_t was drawn.

Following Blasques et al. (2015), we focus on the notion of Kullback-Leibler divergence to measure the distance between the two densities

$$\mathcal{D}_{\mathrm{KL}}\big(p_t\,,\,\tilde{p}_{t+1}\big) = \int_Y p(y|X_t)\log\frac{p(y|X_t)}{\tilde{p}(y|\tilde{f}_{t+1}, X_t;\boldsymbol{\theta})}\,\mathrm{d}y, \tag{13}$$

where $Y \subseteq \mathbb{R}$ is the set over which the divergence is evaluated locally. In particular, we would like an update $\tilde{f}_{t+1}$ for which the divergence $\mathcal{D}_{\mathrm{KL}}(p(\,\cdot\,|f_t, X_t)\,,\,\tilde{p}(\,\cdot\,|\tilde{f}_{t+1}, X_t))$ is smaller than the previous divergence $\mathcal{D}_{\mathrm{KL}}(p(\,\cdot\,|f_t, X_t)\,,\,\tilde{p}(\,\cdot\,|\tilde{f}_t, X_t))$, implying that the update from $\tilde{f}_t$ to $\tilde{f}_{t+1}$

reduces the Kullback-Leibler divergence to the true unknown conditional density.

We can show that only score updates are special in the following sense.

PROPOSITION 2.1 (Proposition 2 in Blasques et al. (2015)). *A smooth observation-driven update from* $\tilde{f}_t$ *to* $\tilde{f}_{t+1}$ *is optimal in the sense of* $\mathcal{D}_{\mathrm{KL}}(p_t, \tilde{p}_{t+1}) < \mathcal{D}_{\mathrm{KL}}(p_t, \tilde{p}_t)$ *for every* $(y_t, \tilde{f}_t, \bar{f}_t)$ *if and only if the update is score equivalent.*

It follows that only score (equivalent) updates have the property that they always locally reduce the Kullback-Leibler divergence and thus provide a local improvement to the statistical model given the data. In particular, the spatial model structure and Student's t specification in Section 2.2 are sufficiently smooth for all local optimality results to apply. Moreover, the score-driven time-varying spatial correlation model is sufficiently regular to also obtain non-local regions where the score steps ensure Kullback-Leibler improvements. We refer to Blasques et al. (2015) for more details, optimality results, and proofs.

2.4 Statistical properties of the model

In this section, we establish the existence, strong consistency and asymptotic normality of the MLE of the static parameters $\boldsymbol{\theta}$ that define the stochastic properties of the spatial score model from Section 2. The results for this specific model hold in a much more general context, and we use this more general framework for the formal proves in the web appendix to this paper. In fact, we extend the results in Blasques et al. (2014) to allow for the presence of exogenous regressors.

The observation-driven structure of the time-varying spatial Durbin model allows us to perform maximum likelihood (ML) estimation in a straightforward way. Following equation (10), we define the ML estimator (MLE) of the static parameter vector $\boldsymbol{\theta}$ as an element of the $\arg\max$ set of the sample log likelihood function $\mathcal{L}_T(\boldsymbol{\theta}, \bar{f}_1)$,

$$\hat{\boldsymbol{\theta}}_T(\bar{f}_1) \in \arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \mathcal{L}_T(\boldsymbol{\theta}, \bar{f}_1), \tag{14}$$

where

$$
\begin{aligned}
\mathcal{L}_T(\boldsymbol{\theta}, \bar{f}_1) &= \frac{1}{T}\sum_{t=1}^{T} \ell_t(\boldsymbol{\theta}, \bar{f}_1) \\
&= \frac{1}{T}\sum_{t=1}^{T} \log p_e\Big(y_t - h\big((\tilde{f}_t(\boldsymbol{\theta}, \bar{f}_1)\big)Wy_t - X_t\beta\ ;\ \lambda\Big) - \log|Z\big(\tilde{f}_t(\boldsymbol{\theta}, \bar{f}_1)\big)|.
\end{aligned}
$$

with $Z(f_t)$ defined below (11).

It is interesting to highlight the main complications in the proof of consistency and asymptotic normality of the MLE. Apart from the usual complications of showing the existence of the appropriate number of moments, a major effort in the proof is proving invertibility of the model. For the crucial importance of proving model invertibility, see for example Wintenberger (2013). As seen above, the likelihood function holds in terms of the data y_t and the filter $\tilde{f}_t$. For the appropriate laws of large numbers and central limit theorems to apply, we therefore need stationarity and ergodicity of both y_t and $\tilde{f}_t$. The former can be established by studying the properties of the model as a data generating process at the true parameter. The latter can be established by studying the properties of the model as a filter for $\tilde{f}_t$ for given data at arbitrary values of the parameter vector. In particular, we prove that for stationary and ergodic data sequences $\{y_t\}$ the filter converges almost surely and pathwise for any starting value $\bar{f}_1$ to a stationary ergodic sequence $\{\tilde{f}_t\}$. Both of the result for y_t and $\tilde{f}_t$ hinge on the contraction properties of quite different stochastic recurrence equations. Given the non-linear structure of the model, studying the properties of these equations is substantially more complicated than in the GARCH case. We refer to the web appendix for more details.

We state the result for the model in (7) with Student's t distributed innovations with $\lambda > 0$ degrees of freedom. Consider a transformation function h that is (a.s.) bounded away from minus one and one with uniformly bounded derivatives $h^{(i)}$,

$$
-1 \ <\ \underline{\rho}\ \leq\ \rho_t = h(f_t)\ \leq\ \bar{\rho}\ <\ 1 \quad \text{a.s.}; \qquad \sup_{f\in\mathcal{F}} |h^{(i)}(f)| < \infty\ , \quad i = 1, 2. \tag{15}
$$

For example, to restrict the correlation to the interval $(-\bar{\rho}, \bar{\rho})$, we can take $h(f_t) = \bar{\rho}\,\tanh(f_t)$, where $\bar{\rho}$ can be arbitrarily close to one. We have the following result.

THEOREM 1. *Consider the spatial score model with link function* (15). *If* $\{y_t\}_{t\in\mathbb{Z}}$ *and* $\{X_t\}_{t\in\mathbb{Z}}$

are SE with $\mathbb{E}|y_t| < \infty$ and $\mathbb{E}|X_t| < \infty$, then there exists a compact parameter space $\boldsymbol{\Theta}$ with $|B| < 1 \; \forall \, \boldsymbol{\theta} \in \boldsymbol{\Theta}$, such that the MLE exists (a.s.) and is strongly consistent for any initialization $\bar{f}_1 \in \mathcal{F}$. If $\mathbb{E}|y_t|^{2+\epsilon} < \infty$ and $\mathbb{E}|X_t|^{2+\epsilon} < \infty$ for some $\epsilon > 0$, then the MLE is asymptotically normal with covariance matrix $\mathcal{I}(\boldsymbol{\theta}_0)^{-1}$ where $\mathcal{I}(\boldsymbol{\theta}_0) := -\mathbb{E}\tilde{\ell}_t''(\boldsymbol{\theta}_0)$ is the Fisher information matrix.

Theorem 1 establishes that we can use the MLE both for estimation and inference.

3 Monte Carlo study

To study the performance of the time-varying spatial score model in filtering out different dynamic patterns for the spatial dependence parameter, we conduct a simulation study. In this study, we also investigate whether the MLE is well-behaved and approximately normally distributed in larger samples as claimed in Theorem 1.

To limit the complexity of the experiment, we consider a spatial lag model without regressors. We set the sample size to realistic values given the empirical application in Section 5. The data generating process is

$$y_t = Z(f_t)e_t, \quad e_t \overset{i.i.d.}{\sim} \text{Student's } t(0, \mathrm{I}_n; 5), \tag{16}$$

where $Z(f_t) = (\mathrm{I}_n - \tanh(f_t)W)^{-1}$, $t = 1, ..., T$ and with cross-sectional dimension $n = 9$. The spatial weight matrix W is specified similar to the row-normalized cross-border exposures of the financial sectors of European countries as used in our empirical application. We simulate 250 data sets according to (16) using five processes with different dynamic patterns for the spatial dependence parameter. These patterns are similar to the ones in Engle (2002).[5]

Figure 1 shows that the filtered spatial dependence parameters are able to capture the patterns of the simulated processes quite accurately. At the low extremes of each path for ρ_t there is some over-smoothing compared to the high extremes, but this is intuitively plausible: the signal present in strongly cross-sectionally correlated data y_t is much more apparent than that in weakly correlated data.

[5] In particular, we consider a constant ($\rho_t = 0.9$); sine ($\rho_t = 0.5 + 0.4\cos(2\pi t/200)$); fast sine ($\rho_t = 0.5 + 0.4\cos(2\pi t/20)$); step function ($\rho_t = 0.9 - 0.5 * I(t > T/2)$); and ramp ($\rho_t = \mod(t/200)$).

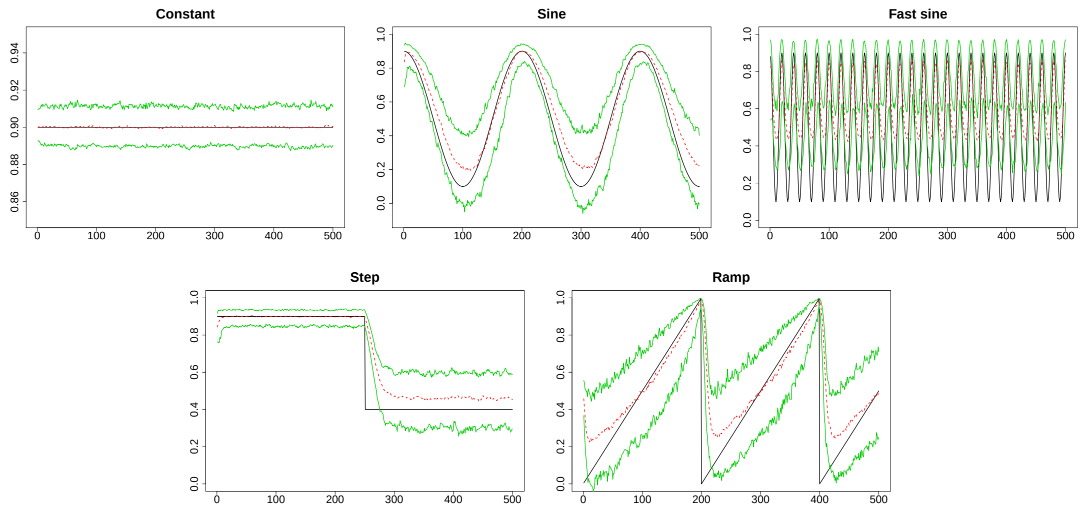


Figure 1: Simulated true spatial dependence process (black line), median filtered parameter (dashed red line) and 2.5% and 97.5% (green lines) quantiles of the filtered parameters. The figures are based on 250 replications.

In our second simulation study, we again use $n = 9$ cross-sectional units. We assume that the disturbances are normally distributed with common variance σ^2, and we include one regressor variable $X_t \sim N(0, I_9)$. The data-generating process is the Gaussian spatial score model laid out in Section 2. In contrast to our previous experiment, the model is now correctly specified. We simulate 500 paths y_t using the parameters $\omega = 0.05$, $A = 0.05$, $B = 0.8$, $\beta = 1.5$, and $\sigma^2 = 2$. We plot the kernel density estimates of the distribution of the MLE for three different sample sizes, $T = \{\ 500, 1000, 2000\}$, in Figure 2.

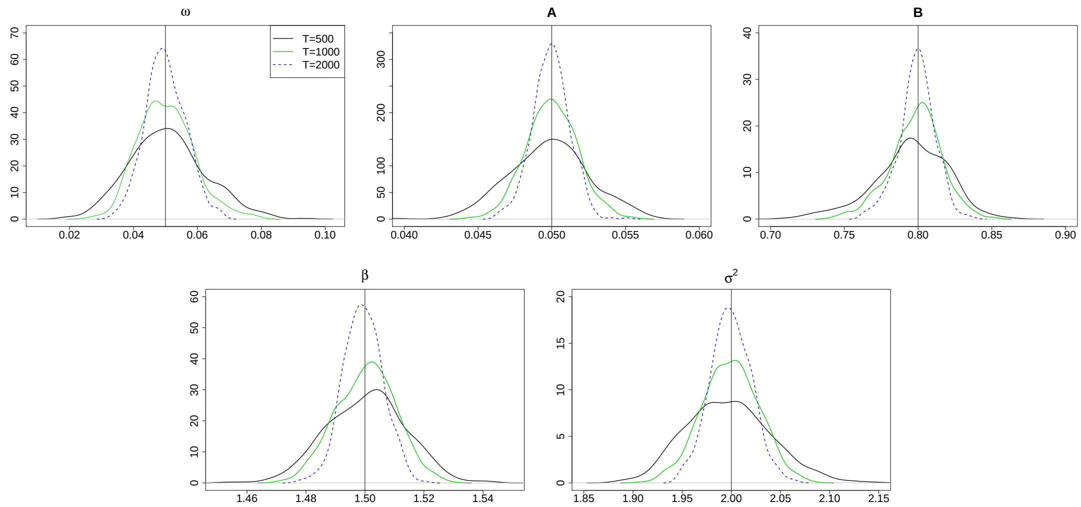


Figure 2: Kernel density estimates of estimated parameters from 500 simulations for 3 sample sizes ($T = 500, 1000, 2000$), vertical lines indicating the true parameter value

The figure clearly shows that for smaller sample sizes of around $T = 500$, the estimators are still not perfectly normal. For larger sample sizes, however, we see a clear convergence to the limiting result. In particular, for empirically relevant sample sizes of around $T = 2,000$ given our empirical application in the next section, all distributions look close to a normal centered around the true parameter values. We therefore apply the MLE and its associated standard errors in our empirical application below.

4 Data

In our empirical study we evaluate the evolution of perceived sovereign credit risk over a period that includes the Eurozone sovereign debt crisis. In particular, we investigate the time-varying features of the spatial dependence structure between the changes in sovereign credit default swap (CDS) spreads, particularly in relation to a number of the policy responses by regulators. Our spatial structure is directly linked to the bank sectors' cross-exposures to other sovereigns and financial sectors within the European Union.

4.1 Credit default spread data

Since EU countries have been affected by the crisis to different degrees, sovereign credit spreads in Europe are strongly cross-sectionally dependent. Figure 3 shows the credit default swap spreads from February 2, 2009, until May 12, 2014 (1375 daily observations) for the eight euro area countries in our sample: Belgium, France, Germany, Ireland, Italy, the Netherlands, Portugal, and Spain. As in Acharya et al. (2014), we use relative changes (log differences multiplied by 100) of U.S. Dollar-denominated sovereign CDS spreads for each of these countries using data obtained from Bloomberg.

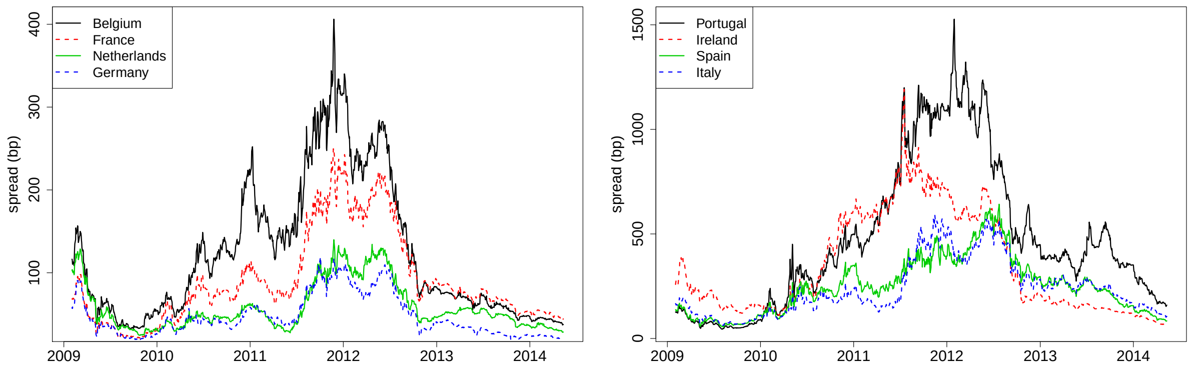


Figure 3: Credit default swap spreads of eight European sovereigns, Feb 2, 2009 – May 12, 2014. The different countries are split in two groups.

The time series reveal clear common patterns, particularly among the non-stressed Eurozone countries (Germany, France, Netherlands, Belgium, and to a lesser extend Spain and Italy). At the same time, there are temporary dissimilarities: for example, the evolution of the Ireland credit spread appears to be roughly in line with that of the other countries before mid 2010 and after mid 2012, but departing during the height of the European sovereign debt crisis. The combination of commonalities with possible temporary changes in commonality warrants the use of the time-varying spatial score model proposed in this paper.

4.2 Other explanatory variables

Our empirical model contains three regressors that capture the state of European financial markets; see also Caporin et al. (2013). The first variable is the change in the volatility index VStoxx. The VStoxx is measured using the implied volatility of the EuroStoxx 50 and captures changes in risk appetite. Our second variable is the difference between the three month Euribor and the overnight rate EONIA. This measure captures financial sector stress and the perceived counterparty credit risk between banks. The third variable is the change in the three month Euribor as a proxy for the monetary policy rate.

We also incorporate two country-specific regressors, namely the (log) returns of the main stock index in each of the respective countries, and absolute changes in the interest rate spreads between government bonds with one year and ten year maturities. We list the local stock indices in Table 1. Local stock market returns are a measure of the well-being of the local economy and in this way an indirect measure of the ability of governments to pay off debt in the long run through tax collection. We expect a negative relation with credit spread changes. The term spreads reflect the difference between long-term and short-term borrowing costs of governments, and we expect a positive relation with sovereign credit default swap spreads.

Table 1: List of country-specific stock indices included in the time-varying spatial score model as regressor variables.

Belgium	BEL 20 Price Index	France	CAC 40 Price Index
Germany	DAX 30 Price Index	Ireland	ISEQ 20 Price Index
Italy	FTSE MIB Price Index	Netherlands	AEX Price Index
Portugal	PSI 20 Price Index	Spain	IBEX 35 Price Index

All variables are included in the model with a lag of one period. The data are obtained from Datastream, except for the short-term government bond yields for France, Germany, Ireland, Italy, and Portugal, which are obtained from Bloomberg. Augmented Dickey-Fuller unit root test statistics indicate that all time series are stationary. Table 2 presents the summary statistics.

Table 2: Data summary. Stock index log returns are calculated from closing prices. All stock indices are quoted in domestic currency (Euro).

	mean	min.	25% quant.	median	75% quant.	max.
CDS spread changes (log changes*100)						
Belgium	-0.08	-19.34	-1.9	-0.07	1.78	17.04
France	-0.03	-19.44	-1.84	-0.07	1.56	19.82
Germany	-0.07	-26.71	-1.89	0	1.56	25.43
Ireland	-0.11	-32.69	-1.57	-0.03	1.32	26.81
Italy	-0.03	-43.73	-2.09	-0.1	1.76	20.27
Netherlands	-0.09	-22.2	-1.66	-0.03	1.39	14.92
Portugal	0.02	-47.38	-1.8	0	1.66	20.54
Spain	-0.04	-37.04	-2.02	0	1.99	25.17
local stock index returns (log returns*100)						
Belgium	0.04	-5.49	-0.59	0.03	0.69	8.96
France	0.03	-5.63	-0.68	0.02	0.80	9.22
Germany	0.06	-5.99	-0.57	0.07	0.75	5.90
Ireland	0.06	-6.79	-0.62	0.02	0.83	6.95
Italy	0.01	-7.04	-0.88	0.04	1.03	10.68
Netherlands	0.04	-5.34	-0.58	0.04	0.71	7.07
Portugal	0.01	-5.51	-0.69	0.02	0.77	10.20
Spain	0.02	-6.87	-0.82	0.01	0.87	13.48
local term spreads (changes)						
Belgium	0	-1.15	-0.03	0	0.03	0.45
France	0	-0.18	-0.02	0	0.02	0.2
Germany	0	-0.17	-0.02	0	0.02	0.24
Ireland	0	-3.89	-0.04	0	0.05	3.76
Italy	0	-1.55	-0.03	0	0.03	1.23
Netherlands	0	-1.02	-0.03	0	0.02	1.1
Portugal	0	-3.94	-0.07	0	0.06	12.79
Spain	0	-1.17	-0.04	0	0.05	1.01
Eurozone-wide variables						
VStoxx change	-0.02	-10.94	-0.86	-0.11	0.67	12.79
term spread	0.35	-0.37	0.14	0.34	0.52	1
Euribor change	-0.13	-9.2	-0.3	0	0.1	6.4

4.3 Spatial weights matrix

The choice of the spatial weights matrix is a key ingredient of the spatial model, as it determines the structure of the 'economic distance' between the sovereign CDS spread changes and defines the channel for cross-sectional spillovers. Recently, domestic banks' cross-border exposures have been identified as relevant pricing factors for sovereign credit spreads, see for example Kallestrup et al. (2016), Korte and Steffen (2015), and Beetsma et al. (2012). A possible reason for this connection is outlined in Korte and Steffen (2015). They argue that until recently, risk management rules for banks implied a so-called 'zero risk weight channel': European banks were not required to hold capital buffers against EU member states' debt. This led to regulatory arbitrage incentives for banks to hold more government debt; see also Acharya and Steffen (2015). At the same time and due to the banks' willingness to take on government debt, governments were able to issue large amounts of debt, thus creating a potentially problematic feedback loop: if sovereign credit risk materialized, banks could become stressed, and due to possible bail-outs, governments in turn might become stressed as well.

To account for this type of possible feedback loop, we use a weight matrix that is constructed from cross-border debt data provided by the Bank for International Settlements (BIS).[6] The data are published on a quarterly basis. Therefore, our weights matrix is updated quarterly as in Denbee et al. (2014). To avoid endogeneity and to account for the time gap in data availability, we lag the matrices by two quarters. For all quarters the raw exposure matrix, which we denote by W_{raw}, is row-normalized to form proper weights that sum up to one. In her spatial model for banking sector interconnections, Tonzer (2015) uses a similar data set, and averages the entries in W_{raw} over her sample period. Another alternative would be to normalize the exposure data by the GDP of the country. We investigate this and other alternatives for constructing the weight matrix in our robustness checks in Section 5.2.

[6]The data can be found at http://www.bis.org/statistics/consstats.htm, Table 9B: International bank claims, consolidated - immediate borrower basis. Last accessed on March 20, 2014.

5 Results

5.1 Main results

Table 3 contains the estimation results for both the static spatial lag model and the time-varying spatial score model for normally and Student's t-distributed disturbances. For the benchmark models, we have a common, time-invariant variance. We relax this assumption in Section 5.2.

For the static model, we find strong evidence for spatial dependence, indicated by the high estimate and small standard error for ρ. Given that CDS spread changes are fat-tailed, it is not surprising to find that the model fit improves substantially for the Student's t vis-à-vis the normal distribution. The likelihood value increases by more than 1800 points upon adding a single parameter to the model, thus decreasing the AICc.

Table 3: Estimated parameters and their robust (sandwich) standard errors in parentheses, for the static Spatial Durbin Model and the time-varying spatial model, based on normally (N) and Student's t (t_λ) distributed disturbances. The maximized log-likelihood value (logL) and the Akaike information criterion, corrected for finite numbers of observations, (AICc) are also reported. Estimation period is February 2, 2009 – May 12, 2014.

	Static model		Time-varying model	
	N	t_λ	N	t_λ
ρ	0.6979	0.6888		
	(0.0083)	(0.0071)		
ω			0.0106	0.0126
			(0.0057)	(0.0177)
A			0.0108	0.0139
			(0.0024)	(0.0093)
B			0.9867	0.9848
			(0.0073)	(0.0215)
$\log(\sigma^2)$	1.8636	0.8708	1.8519	0.8701
	(0.0512)	(0.0455)	(0.0506)	(0.0456)
VStoxx	-0.1193	-0.0403	-0.0946	-0.0389
	(0.0589)	(0.0203)	(0.0398)	(0.0195)
term spread	0.1373	0.0985	0.2008	0.1459
	(0.1228)	(0.0754)	(0.1138)	(0.0735)
Euribor change	0.1119	0.066	0.0886	0.062
	(0.0404)	(0.0305)	(0.0343)	(0.0258)
local stocks	-0.1985	-0.1038	-0.1884	-0.1047
	(0.0479)	(0.028)	(0.0471)	(0.0277)
local term spread	0.2244	0.1288	0.2353	0.1337
	(0.1138)	(0.0796)	(0.1149)	(0.0809)
w.local stocks	-0.0668	-0.0468	-0.0376	-0.0407
	(0.0582)	(0.0333)	(0.0514)	(0.0318)
w.local term spread	0.3517	0.3484	0.3705	0.3906
	(0.3462)	(0.2556)	(0.3286)	(0.2641)
const	-0.0447	-0.0542	-0.0934	-0.0778
	(0.044)	(0.0245)	(0.0451)	(0.0279)
λ_0		2.4708		2.512
		(0.1229)		(0.1266)
logLik	-26632.2	-24780.9	-26464.9	-24692.2
AICc	53284.7	49584.2	52954.1	49410.7

The dynamic spatial score model based on the normal distribution increases the likelihood by almost 170 points compared to the static Gaussian model at the cost of adding two model parameters. The dynamics of the spatial dependence parameter are highly persistent with a value of B close to unity. The unconditional mean of f_t equals $\omega/(1-B) \approx 0.797$ with $\tanh(0.797) \approx 0.6624$. Accounting for the fact that the expected value of $\tanh(f_t)$ is slightly larger than this due to Jensen's inequality, we see that the unconditional level for the Gaussian spatial score model is close to the static estimate of 0.6979. Similarly, the dynamic Student's t model increases the likelihood by approximately 88 points compared to its static counterpart. The unconditional level of $\tanh(f_t)$ again lies close to its static estimate.

On the basis of the reported AICc values, the data clearly favors time variation in the spatial dependence parameter ρ_t using the Student's t distribution for both the disturbance e_t and the transition dynamics of ρ_t. The estimated degrees of freedom parameter λ for the Student's t models is around 2.5. Hence there is a substantial degree of fat-tailedness. A part of the unconditional fat-tailedness may also be due to the presence of volatility clustering. We discuss these robustness issues in more detail in Section 5.2.

The coefficients for the included regressors have the same signs throughout the four model specifications. Although the regression estimates vary somewhat, particularly between the normal and Student's t based models, the overall picture remains the same. A higher implied stock volatility (VStoxx) correlates with lower CDS spreads. This is consistent with the phenomenon of 'flight to quality' from stocks to bonds when the price of risk increases in stock markets. A higher term spread on the interbank credit market implies a higher tendency to borrow overnight. This is correlated with higher CDS spread changes and may be a sign of a perceived bank-sovereign feedback loop: problems in the functioning of the interbank lending market may induce a fear of possible future bailouts and subsequent sovereign debt problems. An increase in the Euribor, which is a measure of the monetary policy rate, may signal that it becomes more costly for banks to obtain liquidity from the central bank, which may induce refinancing problems and impair the functioning of the financial sector. As expected, local stock market upturns have a dampening effect on sovereign credit spreads. The same is true, though to a lesser extent, for neighboring stock markets. Increases in local and neighboring term spreads point to relatively higher borrowing costs for government bonds with longer maturities compared to government bonds with shorter maturities, and have a positive relation with sovereign CDS spread changes.

Figure 4 presents the evolution of the filtered spatial dependence parameter. We observe that the path of the spatial dependence coefficient corresponding to the Student's t spatial score model is more robust to outliers than its normal counterpart. This phenomenon is a common finding in the volatility literature; see for example Creal et al. (2013) and Harvey (2013). Comparing the score expressions in equations (11) and (12), it is clear that the time-varying spatial score model shares this feature. While the normal score is unbounded in the dependent variable and the regressors, the Student's t score contains a compensating effect in the denominator that leads to a down-weighting of large positive or negative observations; see the factor $\tilde{w}_t$ in (12). This leads to a different pattern between the two filtered spatial dependence series for the two distributions, particularly during mid 2010, the first half of 2012, and late 2013.

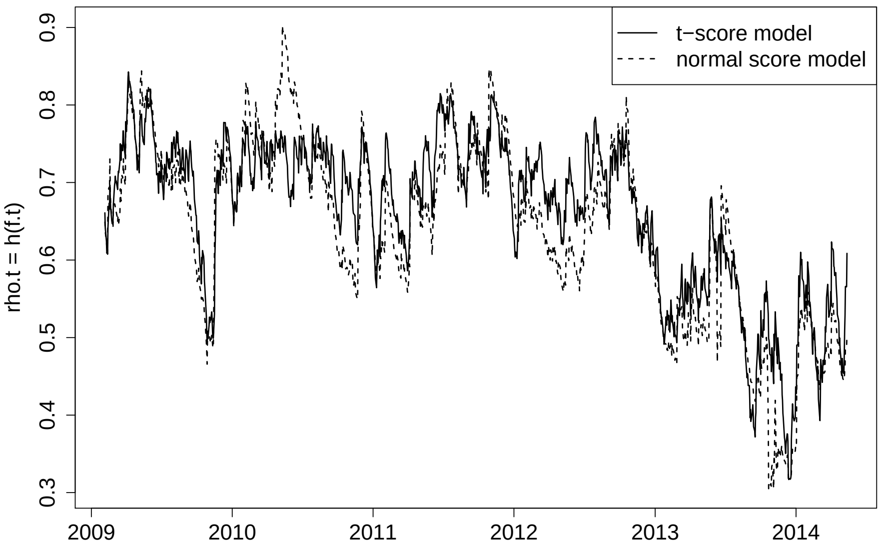


Figure 4: Filtered spatial dependence parameters obtained by imposing normally (dashed line) and Student's t (solid line) distributed disturbances.

Throughout the sample period, systemic risk as captured by the spatial dependence coefficient is high, fluctuating around a value of 0.75 until the end of 2012. At that time, the level starts to decline towards a lower level of around 0.5 to 0.6. Using the Forbes and Rigobon (2002) terminology, only from 2013 onwards markets perceive contagion concerns to be mitigated as the propagation strength (measured by ρ_t) falls. The pattern can be related to a number of important policy events during the European sovereign debt crisis, in particular a number of non-standard monetary operations by the ECB.[7] Some events have a high visible impact. For example, the first Long Term Refinancing Operation (LTRO) at the end of 2012 causes a sudden and sharp drop in the spatial dependence parameter. The effect, however, is short-lived and the value of ρ_t bounces back soon after to similar levels as before. The second LTRO hardly has any visible effect on the spatial dependence parameter. It is not until Mario Draghi's speech at the Global Investment Conference in London in July 2012[8] and the subsequent announcements and implementation of the Outright Monetary Transactions (OMT) and the European Stability Mechanism (ESM) in the months thereafter, that the contagion concerns appear to break down with ρ_t decreasing more permanently to a lower level.

5.2 Extensions

In this section, we extend the time-varying spatial score model in different directions. First, we allow for sovereign-specific volatility clustering. Second, we let the parameters corresponding to the regressors vary over time.

Unobserved time-varying volatility factors

Given the patterns in the data, it is clearly unrealistic to assume a common, time-invariant variance for all sovereign CDS spread changes. We therefore extend the baseline model by adding a time-varying diagonal covariance matrix Σ_t for the disturbances in the spatial model,

$$y_t = h(f_t)Wy_t + X_t\beta + e_t \qquad e_t \sim p_e(0, \Sigma_t), \quad \text{with} \tag{17}$$

$$\Sigma_t := \Sigma(\boldsymbol{f}_t^\sigma) = \text{diag}\left(\sigma_1^2(f_{1,t}^\sigma), \ldots, \sigma_n^2(f_{n,t}^\sigma)\right) = \text{diag}\left(\exp(f_{1,t}^\sigma), ..., \exp(f_{n,t}^\sigma)\right), \tag{18}$$

[7] A list of events can be found in Figure **??** in the supplemental appendix. See also Table **??** with a list of sources.

[8] Quote: "Within our mandate, the ECB is ready to do whatever it takes to preserve the euro. And believe me, it will be enough." Source: see web appendix, Table **??**.

where $\boldsymbol{f}_t^\sigma = (f_{1,t}^\sigma, ..., f_{n,t}^\sigma)'$ is a vector of sovereign-specific variance factors. As before, we endow the factors $\boldsymbol{f}_{j,t}^\sigma$ with score updating dynamics. To enforce parsimony, we allow for sovereign-specific intercepts in the score updating equations for $\boldsymbol{f}_t^\sigma$, but impose common score sensitivity and persistence parameters A^σ and B^σ, so $f_{j,t+1}^\sigma = \omega_j^\sigma + A^\sigma\, s_{j,t}^\sigma + B^\sigma\, f_{j,t}^\sigma$; see Appendix A for further details. Although the covariance matrix of the disturbance vector Σ_t is diagonal, the reduced form covariance matrix of y_t is still a full (time-varying) matrix

$$\mathrm{Cov}(y_t) = Z(f_t)\Sigma_t Z(f_t)'. \tag{19}$$

Time-varying coefficients

It is easy to also accommodate time-variation in a subset or all of the coefficients corresponding to the regressors X_t in our model. The model becomes

$$y_t = h(f_t)Wy_t + X_{t1}\theta_t + X_{t2}\beta + e_t \qquad e_t \sim t_{\lambda_0}(0, \Sigma_t). \tag{20}$$

Here, X_{t1} may contain parts of A_t and/or (WA_t) (see equation (1)). The score dynamics for θ_t are easy to derive. As our interest is in studying the potential time-variation in spillovers between financial markets in the Eurozone, we consider the special case of X_{t1} being the spatial lags of our two individual-specific regressors, local stock market returns, and the changes in local term spreads.

Unobserved time-varying mean factor

To distinguish commonalities from spatial spillovers, we also extend the model with an additional unobserved time-varying mean factor. This factor is independent of the spatial lag structure,

$$y_t = h(f_t)Wy_t + X_t\beta + Z(f_t)^{-1}\boldsymbol{\lambda} f_t^\lambda + e_t, \qquad e_t \sim t_{\lambda_0}(0, \Sigma_t) \tag{21}$$

where λ_0 is the degrees of freedom parameter of the Student's t distribution, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)'$ is an $(n \times 1)$-vector of factor loadings, and $f_t^\lambda \in \mathbb{R}$ is an additional time-varying parameter endowed with score updating. Explicit formulas for the dynamics are given in Appendix A. Rewriting equation (21) in reduced form, we obtain

$$y_t = \boldsymbol{\lambda} f_t^{\lambda} + Z(f_t) X_t \beta + Z(f_t) e_t, \tag{22}$$

which allows for a direct comparison with the benchmark model without spatial lag structure,

$$y_t = X_t \beta + \boldsymbol{\lambda} f_t^{\lambda} + e_t. \tag{23}$$

Goodness of fit comparison

Table 4 compares the goodness of fit of the eight empirical model specifications we consider in our analysis. Almost each extension improves the performance of the model. The exception is the model that allows for time-variation in the coefficients corresponding to the spatially lagged regressors θ. The model without any spatial structure performs worst, despite featuring an unobserved time-varying mean and time-varying volatilities. We therefore conclude that explicitly accounting for dynamic contemporaneous spillovers of shocks, as is done by the time-varying spatial score model, is an important feature when analyzing the dynamics of sovereign credit spread data.

Table 4: Goodness of fit comparison of all empirical specifications considered. The largest log-likelihood value (logL) and smallest Akaike Information Criterion (AICc) are bolded.

	Static spatial		Time-varying spatial	
$e_t \sim$	$N(0, \sigma^2\mathrm{I}_n)$	$t_\lambda(0, \sigma^2\mathrm{I}_n)$	$N(0, \sigma^2\mathrm{I}_n)$	$t_\lambda(0, \sigma^2\mathrm{I}_n)$
logL	-26632.25	-24781.00	-26464.94	-24692.20
AICc	53284.66	49584.19	52954.12	49410.66

	Time-varying spatial-t			Benchmark-t
	(+tv. volas)	(+tv.θ+tv.volas)	(+mean f.+tv.volas)	(+mean f.+tv.volas)
logL	-24365.43	-24364.06	**-24343.39**	-26934.08
AICc	48775.61	48781.16	**48754.46**	53927.45

The parameter estimates from the model with spatial score updating and score-driven, time-varying variances is given in Table 5. In contrast to the spatial factor, the variance factors are less persistent, which is seen by the value of B^{σ}. This is off-set by a larger impact of the scores in the transition equation; see the value of A^{σ}.[9]

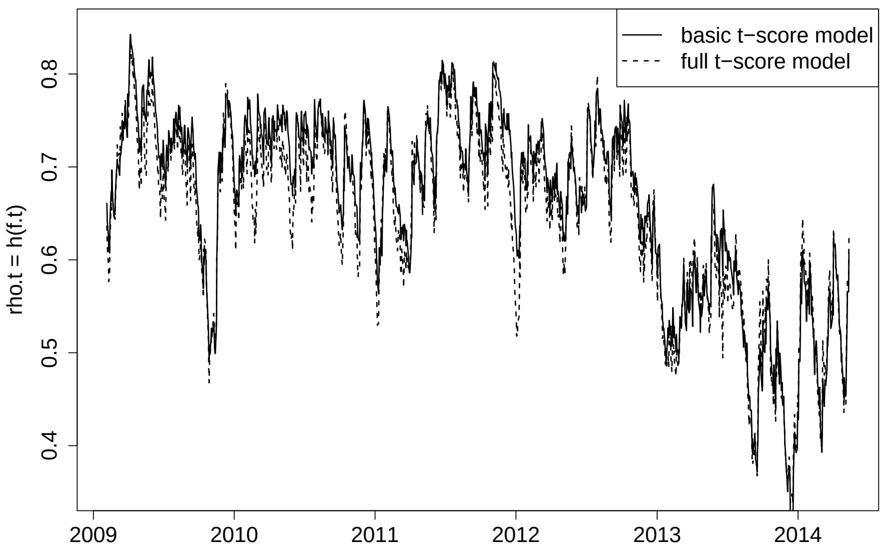
[9]In case of the model with unobserved mean factor, none of the parameters λ_i, $i = 1, \ldots, n$, corresponding to the mean factor are individually significantly different from zero. Jointly, these parameters slightly improve the model fit, as is indicated by the AICc in Table 4. Also, the loading estimates have an economic interpretation: the non-stressed Eurozone countries have a negative coefficient λ_i, while the most stressed countries during part of the European sovereign debt crisis (Portugal, Ireland, Spain) have positive loadings.

Table 5: Estimated parameters and their numerically approximated (sandwich-)standard errors in parentheses, for the full model featuring spatial score updating, time-varying sovereign-specific variances, an unobserved mean factor, and t-distributed disturbances. The maximized log-likelihood value (logL) and the Akaike information criterion (AICc) are also reported. Estimation period is February 2, 2009 - May 12, 2014.

const.	-0.081	ω_1^σ	Belgium	0.0457	ω	0.0241
	(0.0244)			(0.0133)		(0.0188)
VStoxx	-0.0345	ω_2^σ	France	0.0473	A	0.0163
	(0.0198)			(0.015)		(0.0064)
term spread	0.14	ω_3^σ	Germany	0.0582	A^σ	0.189
	(0.0677)			(0.0159)		(0.0233)
Euribor	0.0664	ω_4^σ	Ireland	0.0354	B	0.9697
	(0.0253)			(0.0115)		(0.0237)
stocks	-0.0917	ω_5^σ	Italy	0.0651	B^σ	0.9437
	(0.0262)			(0.0178)		(0.0147)
loc. term sp.	0.1969	ω_6^σ	Netherlands	0.0447	λ_0	3.004
	(0.094)			(0.0136)		(0.1849)
w.stocks	-0.0354	ω_7^σ	Portugal	0.0562		
	(0.0303)			(0.0165)		
w.loc. term sp.	0.4309	ω_8^σ	Spain	0.065	logLik	-24365.4
	(0.2605)			(0.0175)	AICc	48775.6

With respect to the dynamic spatial dependence, the qualitative implications of all models with t-errors are very similar. This is shown in Figure 5. Omitting the additional variance and mean dynamics leads to a slight upward adjustment in the filtered spatial dependence parameter, but the overall pattern does not change.

Figure 5: Filtered spatial dependence parameters obtained from the basic time-varying spatial score model with t-distributed disturbances (solid) as well as with sovereign-specific, dynamic variances and an unobserved mean factor (dashed).



Results from standard residual diagnostic tests are given in Table 6. The model with dynamic spatial dependence and time-varying variances substantially reduces auto-correlations and ARCH effects for most individual series. Furthermore, cross-correlations are, on average, much lower for the model residuals than for the raw data. The full correlation matrices are provided in the web appendix. The web appendix also contains further robustness results using absolute instead of relative CDS spread changes as a dependent variable. Apart from an overall lower level of spatial dependence and a more clearly visible impact of the financial crisis at the beginning of the sample, the pattern for the spatial dependence parameter is similar to that obtained using log changes.

Table 6: Diagnostic tests for the residuals of the full model featuring a spatial updating factor and volatilities, all driven by dynamic score updating, compared to the raw CDS spread changes. LB refers to the Ljung-Box test for residual serial correlation, ARCH LM refers to the test for remaining auto-correlation in the squared residuals. The right-hand panel contains averages of pairwise cross-correlations.

sovereign	LB test stat.		ARCH LM test stat.		average cross-corr.	
	raw	residuals	raw	residuals	raw	residuals
Belgium	108.64	19.82	169.91	25.32	0.70	0.10
France	49.48	25.70	160.44	39.14	0.66	0.00
Germany	62.61	19.67	142.70	61.77	0.63	-0.10
Ireland	129.89	38.13	302.23	89.51	0.64	-0.02
Italy	99.02	40.71	102.13	117.98	0.71	0.09
Netherlands	55.69	46.50	124.41	19.66	0.64	-0.01
Portugal	167.91	52.30	189.35	52.25	0.65	0.08
Spain	105.81	39.12	253.68	131.42	0.69	0.08

5.3 Comparison with other models

In this section, we investigate the robustness of our empirical results. We compare the outcomes of our score-driven Spatial Durbin Model (SDM) with the outcomes of alternative spatial models as well as the DCC model of Engle (2002). Furthermore, to check the sensitivity of the results with respect to the spatial weights matrix, we re-estimate the dynamic SDM for several alternative choices for W_t.

Alternative dynamic spatial specifications

The static Spatial Durbin Model (1) nests the Spatial Lag Model (SLM) and the Spatial Error Model (SEM), which are both frequently used in the literature. In this section, we compare the performance of dynamic versions of these models with our benchmark model, in order to obtain the most parsimonious spatial model for our data. The dynamic SLM is defined as

$$y_t = \rho_t W y_t + \beta_1 \mathbf{1}_n + A_t \beta_2 + e_t, \quad e_t \sim t_{\lambda_0}(0, \Sigma_t)$$

where the model components are defined as in equation (1) except that, as before, $\rho_t = h(f_t)$. The SLM is therefore a restricted version of the SDM with $\beta_3 = 0$. The dynamic SEM is defined as

$$y_t = \gamma_1 \mathbf{1}_n + A_t \gamma_2 + u_t, \quad u_t = \delta_t W u_t + e_t, \quad e_t \sim t_{\lambda_0}(0, \Sigma_t). \tag{24}$$

with $\delta_t = h(f_t^\delta)$. The dynamics for the score-driven SLM model are identical to the dynamics of the score-driven SDM, with $\beta_3 = 0$. We give explicit formulas for the dynamics of the score-driven SEM in Appendix A. We compare the fits of the three models in Table 7. The difference between the fits of SLM and SDM is very small, which is not surprising given the individual insignificance of the two spatially lagged regressors in Table 5. However, the AICc still is slightly smaller for the SDM. The SEM performs worse than the other two candidates, suggesting that its implied parameter restrictions are not supported by the data. We conclude that the dynamic SDM is the most adequate dynamic spatial model to describe our data, closely followed by the SLM.

Table 7: Comparison of likelihood values and AICc for three dynamic spatial model specifications: the Spatial Durbin Model, the Spatial Lag Model and the Spatial Error Model.

	SDM	SEM	SLM
logL	-24365.43	-24404.44	-24368.07
AICc	48775.61	48849.50	48776.76

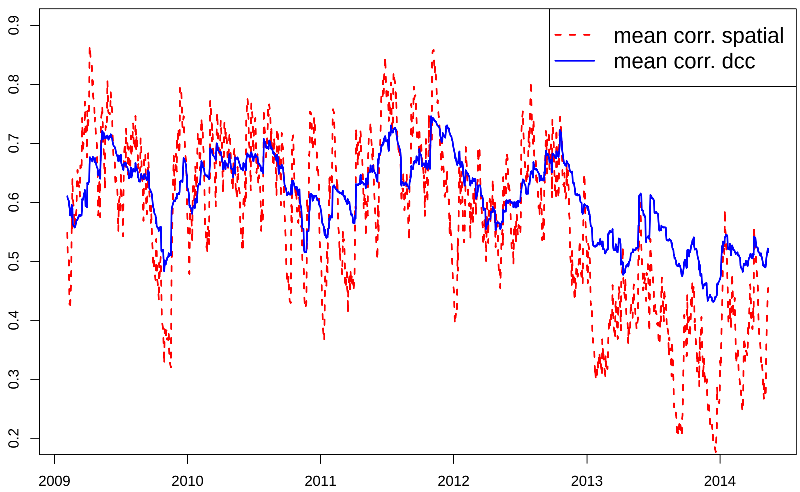
Dynamic conditional correlation model

The Dynamic Conditional Correlation (DCC) model of Engle (2002) and Engle and Sheppard (2001) is a widely used parsimonious model for the time evolution of the correlations of a panel of time series. In contrast to our score-driven spatial model, it does not produce a time-varying scalar measure of spillover strength, but instead a matrix of conditional correlations at each time point t (for details, see Engle (2002)). In fact, Forbes and Rigobon (2002) clearly argue that the use of pairwise conditional correlations may overstate contagion effects by picking up increased spillover where there is none. Following their argument, the spatial correlation parameter is a more structural parameter and better suited to pick up whether spillover strength has changed over time.

To check whether our model's implications are similar to the implications of a DCC model, we compare the cross-sectional averages of the time-varying correlation matrix derived from our model (the standardized version of the covariance matrix given in (19)) with the corresponding cross-sectional averages of the DCC correlation matrices. Figure 6 shows a plot of the two averages over time. The mean correlation implied by the spatial model resemble the plots of ρ_t in Figure 5, but the two are not identical.

Both the average correlations from the spatial model and the DCC model are qualitatively similar, but the spatial model seems to be more responsive to shocks. The DCC-correlations evolve gradually and appear too smooth: they show no immediate reaction to major policy events, such as for instance the first Long Term Refinancing Operation (LTRO) in December 2011. This was a nonstandard monetary policy measure carried out by the ECB to provide banks with liquidity (see also Table **??** and Figure **??** in the Supplemental Appendix). The policy resulted in a temporary break in perceived contagion as picked up by ρ_t and the implied pairwise correlations from the spatial model. No effect is seen, however, for the pairwise average DCC correlations. Furthermore, the decline of the perceived spillovers in CDS spreads after the implementation of the European Stability Mechanism (ESM) and the Outright Monetary Transactions (OMT) program in the second half of 2012 is more pronounced in the time series of correlations implied by the spatial model. We concluding that both models are able to capture salient features of the data, but following Forbes and Rigobon (2002) we prefer the time-varying spatial dependence parameter as a measure of systemic (contagion) risk. By its construction as a scalar measure it summarizes perceived spillover tendencies, and has a structural interpretation due to its ability to incorporate shocks from the regressors; see also equation (6).

Figure 6: Averages of correlation matrix entries (excluding main diagonals) implied by the score-driven SDM with time-varying volatilities (dashed red) and the DCC model (solid blue).



Alternative spatial weights matrices

So far, all results reported were obtained using the spatial weights matrix W_{raw} described in Section 4.3. As a final robustness check, we re-estimate the model using different choices for W: a matrix containing the raw exposure data at the beginning of the sample, i.e. a matrix that is not updated quarterly (W_{const}), a binary matrix indicating the geographical neighborhood of the countries in our sample (W_{geo}), and a (time-varying) weights matrix in which we weight the raw exposures of the financial markets by the countries' respective quarterly GDP[10]. As the different models all have the same number of parameters, we can simply compare the likelihood values at the optimum.

Table 8 shows that the goodness of fit is quite different between the different specifications. The model with a time-varying raw weights matrix provides the best fit. Despite the differences in fit, however, the parameter estimates and the dynamics of the spatial dependence parameter are robust towards the specification of W, and none of the qualitative implications of our model change.

[10]source: OECD statistics

Table 8: Comparison of likelihood values for the time-varying spatial score model with Student's t disturbances using different spatial weights matrices.

	W_{raw}	W_{const}	W_{gdp}	W_{geo}
logL	-24692.2	-24776.95	-24865.78	-25586.64

It is particularly interesting to see that the weight matrices based on economic distances as measured through financial cross-exposures (W_{raw}, W_{const}, and W_{gdp}) provide a much better fit than a matrix based on geographic distances (W_{geo}). However, as mentioned before, scaling the exposures by the size of the economy (as measured by GDP) does not provide an improvement in terms of model fit.

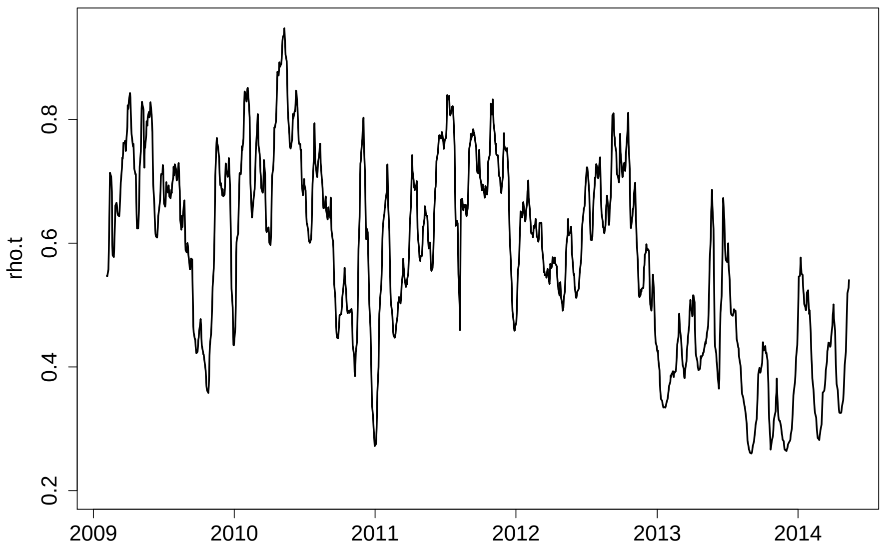
Spatial correlation as a dynamic latent variable

To enable comparisons of the estimates from our spatial score-driven model with the estimates from a complete, parameter-driven specification of the model, we also implement a state-space version of the spatial model with normally distributed disturbances. In particular, the time-varying spatial correlation parameter f_t for the state-space version of the model is specified as a dynamic latent variable via the stochastic equation

$$f_{t+1} = \omega + Bf_t + \eta_{t+1}, \qquad \eta_t \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\eta^2),$$

which is the state equation of the state-space model. The observation equation (7) remains the same in the state-space model. We estimate the parameters of this model and extract the time-varying correlation parameter f_t using the numerically accelerated importance sampling (NAIS) method of Koopman et al. (2015). The combination of the complexities in our empirical data set and the highly non-linear impact of the dynamic parameter f_t on the likelihood function of the state-space model creates a challenge for parameter estimation and signal extraction. It is outside the scope of this paper to provide a full account of the estimation process for this state-space model, but all results are available upon request from the authors. The estimated path of the time-varying correlation parameter f_t from the state-space model is presented in Figure 7 and we observe that the estimated path is somewhat more noisier and less persistent compared to the estimated f_t from the spatial score model. It is, however, comforting that these results leave our main conclusion unaltered: only after late 2012 the systemic risk link appears to be broken. This analysis provides some evidence that our key finding does not hinge on whether we adopt an observation- or a parameter-driven approach in modelling the time-varying correlation parameter.

Figure 7: The estimated spatial dependence parameter from the state-space model with normally distributed disturbances.



6 Conclusion

In this paper, we propose a new model for time-varying spatial dependence in panel data sets. The model extends the widely used spatial lag model to a time-varying parameter framework by endowing the spacial dependence parameter with generalized autoregressive score dynamics and fat tails. Allowing for time-variation is particularly useful if we apply spatial models over longer time periods, where we can no longer be sure that the spatial dependence parameter is constant. The fat-tailed feature of our model is useful in a setting where we apply the model to financial data, which typically exhibit fatter tails than the normal.

We established the theoretical properties of our new model: the dynamics of the model are optimal in the sense that with each update step they locally reduce the Kullback-Leibler distance of the statistical model to the true unknown conditional density. Moreover, we established conditions for model invertibility and for consistency and asymptotically normality of the maximum likelihood estimator in this model.

In our empirical study based on our time-varying spatial score model, we showed that European sovereign CDS spread changes exhibit a strong, time-varying degree of spatial dependence. Cross-border debt linkages appear as a suitable transmission channel for the spatial spillovers. In our final model, we incorporated a time-varying common mean factor as well as time-varying volatilities into the specification. Using the filtered time-varying parameters of this final model, we found evidence for a break in spatial dependence (contagion) towards the end of 2012, i.e., after the implementation of the outright monetary transactions (OMT) by the ECB. Earlier non-standard monetary policies by the ECB, such as the long term refinancing operations (LTROs) only resulted in temporary, short-lived breaks in perceived contagion in the technical sense of Forbes and Rigobon (2002). This illustrates that policies by regulators have at least been partly effective in breaking perceived contagion concerns during the height of the European sovereign debt crisis, but that such actions must be chosen with care and credibly implemented, as otherwise their effect might still be only temporary.

Acknowledgements

We thank Rutger Lit for his excellent help and valuable insights in implementing the importance sampling methods. We also thank conference participants at 7th Annual SoFiE Conference in Toronto 2014, International Association of Applied Econometrics conference in London 2014, Econometric Society European Meeting in Toulouse 2014, Verein fuer Socialpolitik Annual Conference in Hamburg 2014, 2nd workshop on "Models driven by the score of predictive likelihoods" in Tenerife 2014, and SYRTO workshop at the Bundesbank 2014, for helpful comments. Lucas and Blasques thank the Dutch Science Foundation (NWO, grant VICI453-09-005) for financial support. Schaumburg thanks the Dutch Science Foundation (NWO, grant VENI451-15-022) for financial support. Koopman, Lucas, and Schaumburg thank the European Union Seventh Framework Programme (FP7-SSH/2007-2013, grant agreement 320270 - SYRTO) for financial support. Koopman acknowledges support from CREATES, Center for Research in Econometric Analysis of Time Series (DNRF78) at Aarhus University, Denmark, and funded by the Danish National Research Foundation. Research support by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk" is gratefully acknowledged.

References

Acharya, V., Drechsler, I., and Schnabl, P. (2014). A pyrrhic victory? bank bailouts and sovereign credit risk. *Journal of Finance*, 69:2689–2739.

Acharya, V. V. and Steffen, S. (2015). The "greatest" carry trade ever? understanding eurozone bank risks. *Journal of Financial Economics*, 115:215–236.

Ait-Sahalia, Y., Laeven, R. J. A., and Pelizzon, L. (2014). Mutual excitation in european sovereign cds. *Journal of Econometrics*, 183:151–167.

Ang, A. and Longstaff, F. A. (2013). Systemic sovereign risk: Lessons from the u.s. and europe. *Journal of Monetary Economics*, 60:493–510.

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

Arezki, R., Candelon, B., and Sy, A. N. R. (2011). Sovereign rating news and financial markets spillovers: Evidence from the european debt crisis. *IMF Working Paper WP/11/68*.

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

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

Beetsma, R., Giuliodori, M., de Jong, F., and Widijanto, D. (2012). Spread the news: The impact of news on the european sovereign bond markets during the crisis. *Journal of International Money and Finance*, 34:83–101.

Blasques, F., Koopman, S. J., and Lucas, A. (2014). Maximum likelihood estimation for generalized autoregressive score models. *Tinbergen Institute Discussion Papers 14-029/III*.

Blasques, F., Koopman, S. J., and Lucas, A. (2015). Information-theoretic optimality of observation-driven time series models for continuous responses. *Biometrika*, 102(2):325–343.

Caporin, M., Pelizzon, L., Ravazzolo, F., and Rigobon, R. (2013). Measuring souvereign contagion in europe. *NBER Working Paper No. 18741*.

Creal, D., Koopman, S. J., and Lucas, A. (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business and Economic Statistics*, 29(4):552–563.

Creal, D., Koopman, S. J., and Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28:777–795.

Creal, D., Schwaab, B., Koopman, S. J., and Lucas, A. (2014). Observation driven mixed-measurement dynamic factor models. *Review of Economics and Statistics*, 96:898–915.

De Santis, R. A. (2012). The euro area sovereign debt crisis - safe haven, credit rating agencies and the spread of the fever from greece, ireland and portugal. *ECB Working Paper No. 1419*.

Denbee, E., Julliard, C., Li, Y., and Yuan, K. (2014). Network risk and key players: A structural analysis of interbank liquidity. *LSE Financial Markets Group (FMG) Discussion Paper 734*.

Engle, R. (2002). Dynamic conditional correlation. *Journal of Business & Economic Statistics*, 20:339–350.

Engle, R. and Sheppard, K. (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate garch. *NYU Working Paper No. S-DRP-01-10. Available at SSRN: http://ssrn.com/abstract=1296441*.

Fernandez, V. (2011). Spatial linkages in international financial markets. *Quantitative Finance*, 11:237–245.

Forbes, K. J. and Rigobon, R. (2002). No contagion, only interdependence: measuring stock market comovements. *Journal of Finance*, 57(5):2223–2261.

Gorea, D. and Radev, D. (2014). The euro area sovereign debt crisis: Can contagion spread from the periphery to the core? *International Review of Economics & Finance*, 30:78–100.

Harvey, A. C. (2013). *Dynamic Models for Volatility and Heavy Tails*. Econometric Society Monographs, Cambridge University Press.

Harvey, A. C. and Luati, A. (2014). Filtering with heavy tails. *Journal of the American Statistical Association*, 109:1112–1122.

Hillier, G. and Martellosio, F. (2013). Properties of the maximum likelihood estimator in spatial autoregressive models. *cemmap working paper CWP44/13*.

Kalbaska, A. and Gatkowski, M. (2012). Eurozone sovereign contagion: Evidence from the cds market (20052010). *Journal of Economic Behavior & Organization*, 83:657–673.

Kallestrup, R., Lando, D., and Murgoci, A. (2016). Financial sector linkages and the dynamics of bank and sovereign credit spreads. *Journal of Empirical Finance*, page forthcoming.

Keiler, S. and Eder, A. (2015). Cds spreads and contagion amongst systemically important financial institutions - a spatial econometric approach. *International Journal of Finance & Economics*, 20:291–309.

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

Kelly, B., Lustig, H., and Van Nieuwerburgh, S. (2013). Firm volatility in granular networks. *NBER Working Paper No. 19466*.

Koopman, S. J., Lucas, A., and Scharth, M. (2015). Numerically accelerated importance sampling for nonlinear non-Gaussian state-space models. *Journal of Business & Economic Statistics*, 33(1):114–127.

Korte, J. and Steffen, S. (2015). Zero risk contagion - banks' sovereign exposure and sovereign risk spillovers. *Working Paper, Available at SSRN: http://ssrn.com/abstract=2395097*.

Lee, L. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica*, 72:1899–1925.

Lee, L. and Yu, J. (2010). Some recent developments in spatial panel data models. *Regional Science and Urban Economics*, 40:255–271.

LeSage, J. P. and Pace, R. K. (2008). *Introduction to Spatial Econometrics*. CRC Press.

Leschinski, C. and Bertram, P. (2013). Pure contagion dynamics in emu government bond spreads. *Working Paper*.

Lucas, A., Schwaab, B., and Zhang, X. (2014). Conditional euro area sovereign default risk. *Journal of Business and Economic Statistics*, 32(2):271–284.

Oh, D. H. and Patton, A. (2016). Time-varying systemic risk: Evidence from a dynamic copula model of cds spreads. *Journal of Business & Economic Statistics*, page forthcoming.

Ord, K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70:120–126.

Pesaran, M. H. and Pick, A. (2007). Econometric issues in the analysis of contagion. *Journal of Economic Dynamics & Control*, 31:1245–1277.

Saldias, M. (2013). A market-based approach to sector risk determinants and transmission in the euro area. *Journal of Banking & Finance*, 37:4534–4555.

Tonzer, L. (2015). Cross-border interbank networks, banking risk and contagion. *Journal of Financial Stability*, 18:19–32.

Wied, D. (2013). Cusum-type testing for changing parameters in a spatial autoregressive model for stock returns. *Journal of Time Series Analysis*, 34:221–229.

Wintenberger, O. (2013). Continuous invertibility and stable QML estimation of the EGARCH(1, 1) model. *Scandinavian Journal of Statistics*, 40(4):846–867.

Yu, J., de Jong, R., and Lee, L.-f. (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and t are large. *Journal of Econometrics*, 146(118-134).

Appendix A Model extensions and alternatives

We restrict the model extensions to the case of Student's t distributed disturbances. We obtain the equations for the Gaussian case as a special case by letting $\lambda_0 \to \infty$.

We assume that the vector of variance factors $\boldsymbol{f}_t^\sigma$ in (18) follows an n-dimensional score process as given by

$$\boldsymbol{f}_{t+1}^\sigma = \boldsymbol{\omega}^\sigma + A^\sigma \nabla_t^\sigma + B^\sigma \boldsymbol{f}_t^\sigma$$

with $\boldsymbol{\omega} = (\omega_1^\sigma, \ldots, \omega_n^\sigma)'$, and $A^\sigma, B^\sigma \in \mathbb{R}$. We thus allow for sovereign-specific intercepts in the variance score update, but restrict the dynamic parameters A^σ and B^σ to be common across all countries. This results in a parsimonious, yet flexible model. The score of the spatial dependence factor f_t is given in (12), with Σ replaced by Σ_t. For the variance factors, the score vector is

$$\nabla_t^\sigma = \frac{\partial \ell_t}{\partial \boldsymbol{f}_t^\sigma} = \frac{1}{2} \begin{pmatrix} \frac{(1+\lambda^{-1}n)\exp(-f_{1,t}^\sigma)\cdot\left(y_{1,t}-h(f_t)\sum_{j=1}^n w_{1j}y_{j,t}-x_{1,t}'\beta\right)^2}{1+\lambda^{-1}(y_t-h(f_t)Wy_t-X_t\beta)'\Sigma(\boldsymbol{f}_t^\sigma)^{-1}(y_t-h(f_t)Wy_t-X_t\beta)} - 1 \\ \vdots \\ \frac{(1+\lambda^{-1}n)\exp(-f_{n,t}^\sigma)\cdot\left(y_{n,t}-h(f_t)\sum_{j=1}^n w_{nj}y_{j,t}-x_{n,t}'\beta\right)^2}{1+\lambda^{-1}(y_t-h(f_t)Wy_t-X_t\beta)'\Sigma(\boldsymbol{f}_t^\sigma)^{-1}(y_t-h(f_t)Wy_t-X_t\beta)} - 1 \end{pmatrix},$$

with $X_t' = (x_{1,t}, \ldots, x_{n,t})$, and $x_{i,t} \in \mathbb{R}^{k\times 1}$.

In the presence of an additional mean factor f_t^λ as in (22), the score update for f_t changes from (12) to

$$\begin{aligned} \nabla_t &= \left[\tilde{w}_t \cdot \left(Wy_t - W\lambda f_t^\lambda\right)' \Sigma^{-1}\left(y_t - h(f_t)Wy_t - X_t\beta - Z(f_t)^{-1}\boldsymbol{\lambda} f_t^\lambda\right) - \mathrm{tr}(Z(f_t)W)\right] \cdot \dot{h}(f_t), \\ \tilde{w}_t &= \frac{(1+\lambda^{-1}n)}{1+\lambda^{-1}(y_t-h(f_t)Wy_t-X_t\beta-Z(f_t)^{-1}\boldsymbol{\lambda} f_t^\lambda)'\Sigma^{-1}(y_t-h(f_t)Wy_t-X_t\beta-Z(f_t)^{-1}\boldsymbol{\lambda} f_t^\lambda)}. \end{aligned} \tag{A.1}$$

The updating equation for f_t^λ is given by

$$f_{t+1}^\lambda = \omega^\lambda + A^\lambda \nabla_t^\lambda + B^\lambda f_t^\lambda,$$

with score

$$\nabla_t^\lambda = \tilde{w}_t \cdot (Z(f_t)^{-1}\lambda)'\Sigma^{-1}(y_t - h(f_t)Wy_t - X_t\beta - Z(f_t)^{-1}\boldsymbol{\lambda} f_t^\lambda). \tag{A.2}$$

Finally, in the benchmark model (23), the score expression equals that in (A.2) with $W = 0$ and

$Z(f_t) \equiv \mathrm{I}_n$.

The dynamic Spatial Error model is given in equation (24) and can be re-written as

$$y_t = \gamma_1 \mathbf{1}_n + A_t\gamma_2 + (\mathrm{I}_n - h(f_t^\delta)W)^{-1}e_t..$$

The factor f_t^δ is updated according to

$$f_{t+1}^\delta = \omega^\delta + A^\delta s_t^\delta + B^\delta f_t^\delta.$$

Define

$$\begin{aligned}\bar{e} &= (\mathrm{I}_n - h(f_t^\delta)W)(y_t - \gamma_1\mathbf{1}_n - A_t\gamma_2) \\ &= y_t - h(f_t^\delta)Wy_t - \gamma_1\mathbf{1}_n + h(f_t^\delta)W\gamma_1\mathbf{1}_n - A_t\gamma_2 + h(f_t^\delta)WA_t\gamma_2\end{aligned}$$

The t-likelihood is

$$\ell_t = \log|Z(f_t^\delta)^{-1}| + \log\left(\frac{\Gamma\left(\frac{\lambda+n}{2}\right)}{|\Sigma|^{1/2}(\lambda\pi)^{n/2}\Gamma\left(\frac{\lambda}{2}\right)}\right) - \left(\frac{\lambda+n}{2}\right)\log\left(1 + \frac{\bar{e}_t'\Sigma^{-1}\bar{e}_t}{\lambda}\right).$$

and the score functions is

$$s_t^\delta = \Big(\bar{w}_t \cdot (y_t' + \gamma_1\mathbf{1}_n' + \gamma_2'A_t')W'\Sigma^{-1}\bar{e}_t - \mathrm{tr}(Z(f_t^\delta)W)\Big) \cdot \dot{h}(f_t^\delta),$$

with

$$\bar{w}_t = (1 + \lambda^{-1}n)/(1 + \lambda^{-1}\bar{e}_t'\Sigma^{-1}\bar{e}_t).$$