Dynamic Spatial Autoregressive Models with Time-varying Spatial Weighting Matrices

Presentation · September 2019

Anna Gloria Billé
University of Padova

Leopoldo Catania
Aarhus University

Dynamic Spatial Autoregressive Models with Time–varying Spatial Weighting Matrices

Anna Gloria Billé[*], Francisco Blasques[◇] & Leopoldo Catania[†]

[*] Faculty of Economics and Management, Free University of Bozen–Bolzano, Italy;
[◇] VU University Amsterdam and Tinbergen Institute, Amsterdam, The Netherlands;
[†] Department of Economics and Business Economics and CREATES, Aarhus University, Aarhus, Denmark

September 10–13, 2019

Statistische Woche 2019
Trier University, Germany

Outline

1. **Introduction**
   - A brief literature review
   - Our motivation
2. **A Dynamic SAR model with a Time–varying Spatial Weighting matrix**
   - Model Specification
   - Different W_t parametrizations – the function $f(.,.)$
   - Normalization rules – the function $g(.)$
   - Filtering procedure for time–varying parameters
3. **The role of γ_t and a dynamic indicator of spatial association ϖ_t**
4. **Monte Carlo simulations**
   - Finite sample properties of MLE
   - Filtering properties
5. **Empirical applications**
   - Association between European countries' perceived risk
   - House price dynamics in the UK
6. **Conclusions and Future Developments**

Introduction: A brief literature review

Specification of the weighting matrix W within spatial autoregressive models

- In spatial econometrics [Ord, 1975], the weighting matrix W is a way to connect georeferenced random variables.
- **Exogenous** W. Several definitions: contiguity–based, *k*–neareast neighbors, distance-based criteria, etc.
- An *optimal* W is surely an unrealistic goal [Bavaud, 1998]: no obvious interpretation of the lag in a plane as in time [Whittle, 1954].
- *Are estimates and inferences sensitive to different* W *specifications?* School of thought ...
  1. Smoothing and Semiparametric approaches [McMillen, 2012]
  2. Robustness checks and Economic Theory [Corrado and Fingleton, 2012]
  3. Partial derivatives and summary measures [LeSage and Pace, 2014]
  4. Bayesian approach (prior knowledge)

Introduction: A brief literature review

Model Misspecification problems [Billé and Leorato, 2017]

(1) Estimation methods for potential **Endogenous** W [Qu and Lee [2015], Kelejian and Piras [2014]] (GMM, 2SLS, IV)

(2) Model selection [Seya et al., 2013]

(3) **Estimation of** W. Two main problems:

- W as a parameter $\Rightarrow \rho$ not identified.
- Even with panel data $T << N$, the estimation of $\frac{N\times(N-1)}{2}$ elements of W is intractable.

Current solutions:

- Matrix exponential spatial specification (MESS) [LeSage and Pace, 2007]
- Two-step residual regression estimator (spectral decomposition) [Bhattacharjee and Jensen-Butler, 2013]

Economic systems are essentially *dynamic by nature*, leading to potentially wrong conclusions even with economic W in a static spatial model.

Our motivation: Time–varying W_t matrices

(1) When $T << N$: **Spatial Dynamic Panel Data (SDPD) Models** [Lee and Yu [2012]; Wang and Yu [2015]]
(2) When $T >> N$: **Spatio-temporal model specifications** (exploit more the time information)

- LASSO approach [Otto and Steinert, 2018]
- Parameterizing W_t: $W(\gamma_t, \mathbf{d})$
  1. *Even if there are theoretical reasons indicating that spatial interaction effects are related to distance, it is often not clear from the theory the degree at which the spatial dependence diminishes as distance increases.* [Halleck Vega and Elhorst, 2015]
  2. Although geographical locations/regions are time–invariant, the strength of spatial dependence may also depend on economic variables that are time–varying.

(1) γ_t time–varying distance decay parameter $\Rightarrow$ managing the relative importance of geographical distance

(2) Score Driven (SD) models [[Harvey, 2013]; [Creal et al., 2013]] $\Rightarrow$ exploiting information contained in the score of the conditional distribution of the observables.

Model Specification

A Dynamic SAR model with Time–varying Spatial Weighting matrix (DSWM model)

Let $\mathbf{y}_t$ be a N–dimensional stochastic vector of spatial variables located on a possibly unevenly spaced lattice $Z \subseteq \Re^N$ at time t. We assume that $\mathbf{y}_t$ is generated according to a spatial autoregressive–regressive (SAR) model

$$\mathbf{y}_t = \rho \mathrm{W}\left(\gamma_t, \mathbf{d}\right)\mathbf{y}_t + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \overset{iid}{\sim} \mathcal{D}\left(\mathbf{0}, \boldsymbol{\Sigma}_t, \psi\right), \tag{1}$$

where ρ is a static (time–invariant) spatial autocorrelation coefficient, $\mathbf{X}_t = \left(\mathbf{x}_{j,t}; j = 1, \ldots, K\right)$ is a $N \times K$ matrix of exogenous covariates with associated static vector of coefficients $\boldsymbol{\beta} = \left(\beta_j; j = 1, \ldots, K\right)'$, $\boldsymbol{\varepsilon}_t$ is the N–stochastic vector of innovations at each time with diagonal covariance matrix $\Sigma_t = \mathrm{diag}\left(\sigma^2_{i,t}; i = 1, \ldots, N\right)'$, ψ is the shape parameter (Normal and Student's t distributions)

Note: *Cross-sectional heteroskedasticity* if $\sigma^2_i \neq \sigma^2_j$ for at least one $i \neq j$, $i, j = 1, \ldots, N$. *Time-varying shock volatility* through the score driven methodology.

Model Specification

Definition (2.1)

$\mathrm{W}(\gamma_t, \mathbf{d}) = \mathrm{W}_t = \{\omega_{ij,t}\}_{i,j=1}^{N}$.
(a) $\omega_{ij,t} = \omega_{ji,t}, \omega_{ij,t} > 0$, $\omega_{ii,t} = 0 \quad \forall i,j = 1,\ldots,N$ and $\forall t$,
(b) $\mathbf{d} = (d_{ij}; i,j = 1,\ldots,N)$, $d_{ii} = 0, \quad \forall i = 1,\ldots,N$ and $d_{ij} = d_{ji}$
$d_{ij} > 0$ for all $i,j = 1,\ldots,N, \quad i \neq j$,
(c) $\gamma_t > 0$.

Def (2.1)(a): W_t is a sequence of symmetric $N \times N$ matrices of real positive entries with zero on the main diagonal. **Def (2.1)(b)**: $\mathbf{d}$ is a $N \times N$ matrix of strictly exogenous non–stochastic variables such that $\mathbf{d} \in \Re_{+}^{N \times N}$.

Assumption (1)

The rows and the columns of $\mathrm{W}_t, \forall t$ *are uniformly bounded as N goes to infinity.*

Different W_t parameterizations: the function $f(.,.)$

$\mathcal{F}_t = \sigma(\mathbf{y}_{t-s}, \mathbf{X}_{t-s+1}, s \geq 0)$ is the filtration generated by the process of $\mathbf{y}_t$ defined in equation (1).

Definition (2.2)

$\omega_{ij,t} = f(\gamma_t, d_{ij})$, where $f : \Re_+ \to \Re_+$ is an $\mathcal{F}_t$–measurable differentiable distance decay function among the spatial units i and j:
(a) $f(\gamma_t, d_{ij}) > 0$ for $i \neq j$,
(b) $f(\gamma_t, d_{ij}) = 0$ for $i = j$,
(c) $f(\gamma_t, d_{ij}) < f(\gamma_t, d_{hk}) \iff d_{ij} > d_{hk}$,
(d) $\lim_{d_{ij}\to\infty} f(\gamma_t, d_{ij}) = 0$, $\lim_{d_{ij}\to 0} f(\gamma_t, d_{ij}) = 1$,
$\lim_{\gamma_t\to\infty} f(\gamma_t, d_{ij}) = 0$, $\lim_{\gamma_t\to 0} f(\gamma_t, d_{ij}) = 1$.

Any **continuous monotonically decreasing function** $\to$ lower weights as distances increase. In our paper we consider the following $f(.,.)$:

(1) *inverse distance*, $f(\gamma_t, d_{ij}) = \frac{1}{d_{ij}^{\gamma_t}}$

(2) *negative exponential*, $f(\gamma_t, d_{ij}) = \exp(-\gamma_t d_{ij})$

Normalization rules: the function $g(.)$

In order to ensure *stable spatio–temporal processes* ...

Lemma (2.1)

Let $\tau(\mathrm{W}_t)$ denote the **time–varying spectral radius** *of the square N–dimensional W_t matrix at time t:*
$\tau(\mathrm{W}_t) = max\{|e_{1t}|, ..., |e_{Nt}|\}$, where $e_{1t}, ..., e_{Nt}$ are the eigenvalues of W_t at time t. Then, $(\mathbb{I}_N - \rho \mathrm{W}_t)^{-1}$ is non–singular for all the values of ρ in the interval $(-1/\tau(\mathrm{W}_t), 1/\tau(\mathrm{W}_t))$ at each time t.

Assumption (2)

$\rho \in \left(-\frac{1}{\tau(\mathrm{W}_t)}, \frac{1}{\tau(\mathrm{W}_t)}\right)_{\setminus\{0\}}, \forall t.$

Two important issues ...

(1) The **parameter space of** ρ in model (1), with no further assumptions, is not fixed, i.e. $\rho \in \left(-\frac{1}{\tau(\mathrm{W}_t)}, \frac{1}{\tau(\mathrm{W}_t)}\right)_{\setminus\{0\}}$

(2) The spectral–normalization rule [Kelejian and Prucha, 2010] is not sufficient in our case since the autoregressive coefficient of the equivalent normalized model will be $\rho^* = \rho \times \tau(\mathrm{W}_t)$

Normalization rules: the function $g(.)$

Definition (2.3)

$g(\mathrm{W}_t) = \mathrm{W}_t^*$, where $g(\cdot)$ is a $\mathcal{F}_t$–measurable differentiable normalizing function such that $\tau(\mathrm{W}_t^*) = 1, \ \forall t$.

Then, we finally consider the following model

$$\mathbf{y}_t = \rho^* \mathrm{W}_t^* \mathbf{y}_t + \mathbf{X}_t \beta + \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} \mathcal{D}(\mathbf{0}, \mathbf{\Sigma}_t, \psi), \tag{2}$$

where $\mathrm{W}_t^* = \{\omega_{ij,t}^*\}_{i,j=1}^N$ and $\omega_{ij,t}^* = g_{ij}(\gamma_t, \mathrm{W}_t)$, with $\tau(\mathrm{W}_t^*) = 1$. The inverse matrix $(\mathbb{I} - \rho^* \mathrm{W}_t^*)^{-1}$ exists for all values of ρ^* in the interval $(-1, 1)_{\setminus\{0\}}$. Then, the reduced form of model (2) is

$$\mathbf{y}_t = \mathbf{A}_t^{-1} \mathbf{X}_t \beta + \mathbf{A}_t^{-1} \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} \mathcal{D}(\mathbf{0}, \mathbf{\Sigma}_t, \psi) \tag{3}$$

and the conditional distributions of $\mathbf{y}_t$ and ε_t are

$$\mathbf{y}_t | \mathcal{F}_{t-1} \sim \mathcal{D}\left(\mathbf{y}_t; \mathbf{A}_t^{-1} \mathbf{X}_t \beta, \mathbf{A}_t^{-1} \mathbf{\Sigma}_t \mathbf{A}_t^{-1\prime}, \psi\right) \tag{4}$$

and

$$\varepsilon_t | \mathcal{F}_{t-1} \sim \mathcal{D}(\mathbf{0}; \mathbf{\Sigma}_t, \psi), \tag{5}$$

respectively.

Filtering procedure for time–varying parameters

γ_t is updated through a filter based on the *scaled score* of the conditional density in (4). Let $\gamma_t = \exp(\tilde{\gamma}_t)$. The filter is

$$\tilde{\gamma}_{t+1} = (1 - \xi)\,\kappa + \alpha \tilde{s}_t + \xi \tilde{\gamma}_t, \tag{6}$$

where the condition $|\xi| < 1$ is imposed to ensure stationarity of the process and

$$\tilde{s}_t = \tilde{\mathcal{I}}_t\left(\tilde{\gamma}_t, \boldsymbol{\eta}, \mathbf{X}_t\right)^{-1/2} \tilde{\nabla}_t\left(\mathbf{y}_t, \tilde{\gamma}_t, \boldsymbol{\eta}, \mathbf{X}_t\right), \tag{7}$$

where $\boldsymbol{\eta} = \left(\rho^*, \mathrm{diag}\left(\boldsymbol{\Sigma}\right)', \boldsymbol{\beta}'\right)'$. *Time–varying shock volatilities*: given the conditional density in (5) we introduce time–variation in the $\sigma^2_{i,t}$ by letting $\sigma^2_{i,t} = \exp(\tilde{\sigma}^2_{i,t})$ and

$$\tilde{\sigma}^2_{i,t+1} = (1 - \xi_\sigma)\kappa_{\sigma,i} + \alpha_\sigma s_{i,t} + \xi_\sigma \tilde{\sigma}^2_{i,t}, \tag{8}$$

where $s_{i,t}$ is the scaled score of the conditional distribution $\mathbf{y}_t|\mathcal{F}_{t-1}$ evaluated with respect to $\tilde{\sigma}^2_{i,t}$.

Filtering procedure for time–varying parameters

Given the conditional distributions in equations (4) and (5), the log–likelihood contribution (Normal case) of $\mathbf{y}_t$ is proportional to

$$\log p\left(\mathbf{y}_t|\gamma_t,\boldsymbol{\eta},\mathbf{X}_t\right) \propto -\frac{1}{2}\log|\mathbf{\Sigma}_t| + \log|\mathbf{A}_t| - (1/2)\,\boldsymbol{\nu}_t'\boldsymbol{\nu}_t, \qquad (9)$$

where

$$\boldsymbol{\nu}_t'\boldsymbol{\nu}_t = \left(\mathbf{A}_t\mathbf{y}_t - \mathbf{X}_t\beta\right)'\mathbf{\Sigma}_t^{-1}\left(\mathbf{A}_t\mathbf{y}_t - \mathbf{X}_t\beta\right) \qquad (10)$$

Given a spatial sample of time dimension T, model parameters can be easily estimated through Maximum Likelihood (ML), i.e.

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}\in\mathbf{\Theta}} \sum_{t=1}^{T} \log p\left(\mathbf{y}_t|\boldsymbol{\theta},\mathbf{X}_t\right), \qquad (11)$$

where $\boldsymbol{\theta} = \left(\boldsymbol{\eta}',\kappa,\alpha,\xi,\kappa_{\sigma,i},\alpha_\sigma,\xi_\sigma\right)'$ are time–invariant parameters to be estimated, and $\mathbf{\Theta}$ is a compact set where model parameters take values [[Harvey, 2013]; [Blasques et al., 2014]].

The role of γ_t

The goal: consistently estimate γ_t

- $\gamma_t \in \Re_+$ is the *time–varying distance decay parameter*, measuring the relative importance of higher–order neighbourhoods in space and over time t.
- Interactions may continue even over a first or second–order neighbourhood, suggesting that a *sparse* matrix might be non appropriate most of the times.
- Useful to selecting appropriate weighting schemes and knowing the optimal level at which the correlation rapidly decreases over time, without any a priori knowledge of the spatial process $\Rightarrow$ avoinding **model misspecification** due to prespecified W_t.
- Interpretation: $\gamma_t \rightarrow 0$, the higher the role played by the spatial units at greater distances, i.e. *higher–order neighbourhoods* $\Rightarrow$ *dense* matrices are maybe more appropriate.

Dynamic indicators of spatial association ϖ_t

Problem: the domain of γ_t is unbounded

Dynamic indicator

$$\varpi_t = 1 - \frac{\lambda(R_t(\infty))}{\lambda(R_t(\gamma_t))} \tag{12}$$

such that $\varpi_t \in (0, 1)$, and where $R_t(\infty)$ (no spatial association) and $\lambda(.)$ (highest eigenvalue). When γ_t diverges, the spatial units become uncorrelated and $\lambda(R_t(\infty)) = 1$.

- $\gamma_t \to 0 \Rightarrow \varpi_t \to 1$, stronger connections among the spatial units (dense W_t more appropriate),
- $\gamma_t \to \infty \Rightarrow \varpi_t \to 0$, viceversa.

Monte Carlo simulations: FSP of MLE

(1) **Finite sample properties of MLE**

- Number of series $B = 1000$ of pseudo observations from the true model (eq. (4) and (6))
- No time–varying shock volatilities
- $\rho^* = \{-0.7, -0.3, 0.3, 0.7\}$, $N = \{50, 100\}$, $\sigma = 1.0$ (cross-sectional homoskedasticity)
- $\alpha = 0.02$, $\xi = 0.97$, $\kappa = \log 2 \approx 0.6931$ for the γ_t recursion

Note: The true value of the $\kappa = 0.6931$ parameter implies an unconditional level of the γ_t dynamic of 2, i.e. $\mathbb{E}[\gamma_t] = 2$ for all the considered specifications.

Monte Carlo simulations: FSP of MLE

$N = 100$	T	DGP1					DGP2				
		ρ =-0.7	α =0.02	β =0.97	κ =0.6931	σ =1	ρ =-0.3	α =0.02	β =0.97	κ =0.6931	σ =1
Median	500	-0.6998	0.0201	0.9671	0.7024	1.0000	-0.2998	0.0193	0.9567	0.7008	0.9999
	1000	-0.7000	0.0200	0.9672	0.7113	1.0000	-0.3003	0.0203	0.9652	0.6913	1.0000
	2000	-0.6999	0.0200	0.9686	0.7013	0.9999	-0.2997	0.0200	0.9676	0.6948	1.0000
SD	500	0.0028	0.0011	0.0129	0.1241	0.0049	0.0059	0.0095	0.0809	0.0905	0.0040
	1000	0.0019	0.0008	0.0090	0.1040	0.0033	0.0039	0.0058	0.0496	0.0649	0.0031
	2000	0.0014	0.0005	0.0058	0.0825	0.0021	0.0034	0.0047	0.0368	0.0772	0.0023
MSE	500	8.0E-06	1.2E-06	0.0001	0.0154	2.4E-05	3.4E-05	9.0E-05	0.0076	0.0082	1.6E-05
	1000	3.8E-06	6.4E-07	9.8E-05	0.0110	1.1E-05	1.5E-05	3.4E-05	0.0028	0.0042	9.7E-06
	2000	2.1E-06	3.0E-07	3.7E-05	0.0068	4.4E-06	1.1E-05	2.2E-05	0.0014	0.0060	5.5E-06
		DGP3					DGP4				
		ρ =0.3	α =0.02	β =0.97	κ =0.6931	σ =1	ρ =0.7	α =0.02	β =0.97	κ =0.6931	σ =1
Median	500	0.3004	0.0194	0.9561	0.6879	1.0000	0.6999	0.0199	0.9667	0.7011	0.9999
	1000	0.3002	0.0202	0.9638	0.7000	0.9999	0.6999	0.0199	0.9680	0.7016	1.0000
	2000	0.3000	0.0199	0.9663	0.6960	1.0000	0.6999	0.0200	0.9689	0.6952	1.0000
SD	500	0.0056	0.0081	0.0953	0.0863	0.0044	0.0026	0.0011	0.0136	0.1200	0.0043
	1000	0.0040	0.0065	0.0455	0.0745	0.0032	0.0020	0.0008	0.0089	0.1021	0.0030
	2000	0.0029	0.0044	0.0456	0.0551	0.0021	0.0014	0.0005	0.0062	0.0826	0.0019
MSE	500	3.2E-05	6.6E-05	0.0105	0.0074	2.0E-05	7.2E-06	1.3E-06	0.0002	0.0144	1.8E-05
	1000	1.6E-05	4.2E-05	0.0024	0.0056	1.0E-05	4.3E-06	6.5E-07	9.0E-05	0.0104	9.5E-06
	2000	8.8E-06	2.0E-05	0.0022	0.0030	4.5E-06	2.1E-06	3.2E-07	4.4E-05	0.0068	3.7E-06

Monte Carlo simulations: FSP of MLE

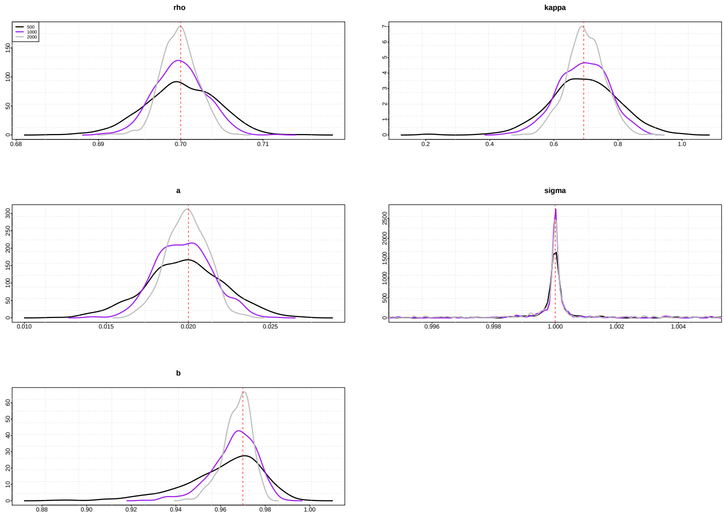


Figure: Finite sample properties with a spatial dimension N=50, $T = \{500, 1000, 2000\}$ and a positive autocorrelation coefficient ρ^*=0.7.

Monte Carlo simulations: Filtering Properties

(2) **Flexibility of the score updating mechanism assumed for the γ_t recursion**

We assume the following artificial deterministic and stochastic patterns

- Constant: $\gamma_t = 2$
- Sine: $\gamma_t = 3 + 2\sin(2\pi t/200)$
- Fast Sine: $\gamma_t = 3 + 2\sin(2\pi t/20)$
- Step: $\gamma_t = 4 - 3\,(t > 500)$
- Ramp: $\gamma_t = \mod(t/400)/100 + 1$
- AR: $\gamma_t = \exp(\tilde{\gamma}_t)$, $\tilde{\gamma}_t = 0.015 + 0.98\tilde{\gamma}_{t-1} + 0.1\eta_t, \eta_t \sim \mathcal{N}(0,1)$.

We simulate $B = 1000$ series of length $T = 1000$ and $N = 25$, whereas $\sigma = 1$ and $\rho^* = 0.7$ are fixed and are not estimated.

Monte Carlo simulations: Filtering Properties

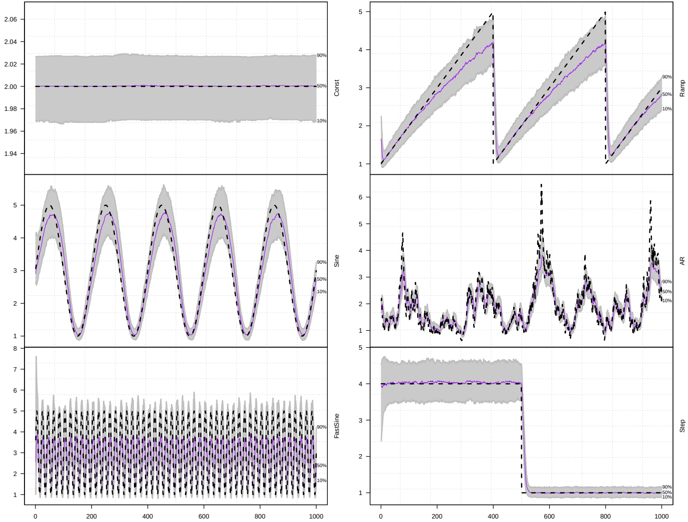


Figure: Filtering properties of the model.

(2) House price dynamics in the UK

Recent literature on house prices [Shi and Lee [2017]; Billé et al. [2017]; Bailey et al. [2016]; Holly et al. [2010]]

(1) Quality adjusted (to price variations) regional house price series from Nationwide Building Society website [Holly et al., 2011].

(2) Period from Q1 1974 to Q2 2018 (178 quarterly observations).

(3) Starting from NUTS data from the UK Office of National Statistics website, Nationwide regions are obtained by aggregating different NUTS regions.

(4) Explanatory Variables: population level, unemployment index, etc. [Brady, 2011].

(5) **d**: geographical distances based on Euclidean distances between centroids of the regions.

Aim: *identifying periods of time in which the evaluated time–series are more inter–connected, suggesting a common behaviour among them*

(2) House price dynamics in the UK

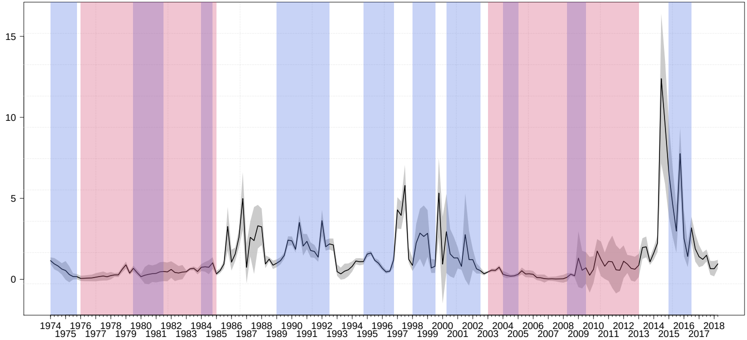


Figure: Filtered γ_t for the **Student's t model** for UK house price dynamics and $(10\% - 90\%)$ in–sample simulation–based confidence bands (gray bands). The blue shaded bars indicate periods of UK recession according to the OECD based Recession Indicators for the United Kingdom from the Peak through the Trough (GBRRECDM), see Federal Reserve Bank of St. Louis [2018].

(2) House price dynamics in the UK

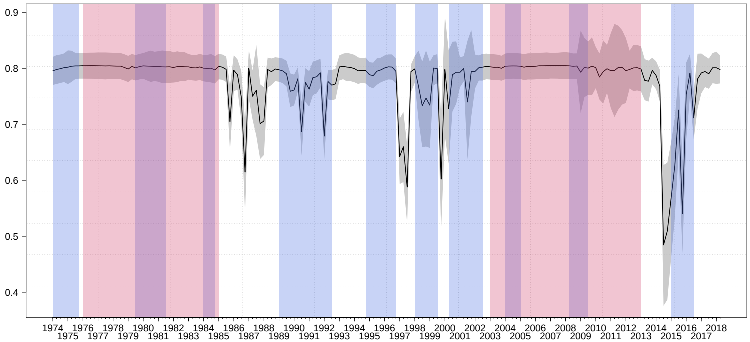
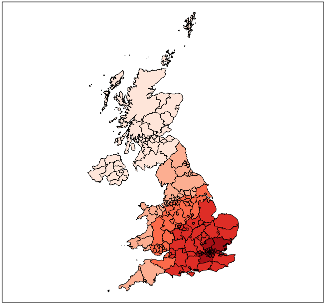
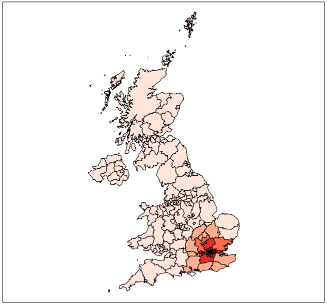


Figure: The evolution of the indicator of spatial association as in Equation (12) for house price dynamics and (10% − 90%) in–sample simulation–based confidence bands (gray bands). The blue shaded bars indicate periods of UK recession according to the OECD based Recession Indicators for the United Kingdom from the Peak through the Trough (GBRRECDM), see Federal Reserve Bank of St. Louis [2018].

(2) House price dynamics in the UK



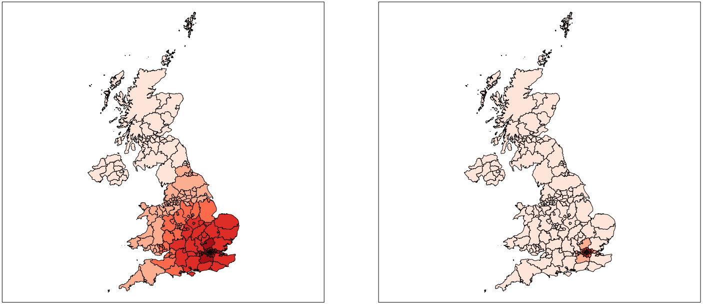
(a) Period 1976-1985



(b) Period 1986-2003

Figure: Higher–order effects of W_t with respect to London for 4 different periods of time. The intra–periods average values of γ_t are: (a) $\gamma_1 = \mathbf{0.415}$, (b) $\gamma_2 = \mathbf{1.734}$.

(2) House price dynamics in the UK



(a) Period 2004-2012 (b) Period 2013-2016

Figure: Higher–order effects of W_t with respect to London for 4 different periods of time. The intra–periods average values of γ_t are: (a) $\gamma_3 = \mathbf{0.492}$, (b) $\gamma_4 = \mathbf{4.054}$.

Conclusions

- A Dynamic Spatial Weighting Matrix (DSWM) model.
- Parametrization of W_t: any continuous monotonically decreasing function (distance decay functions).
- Feasible even with large N.
- Score Driven models, exploiting information from the conditional distributions.
- Finite sample as well as asymptotic properties of the ML estimator. Filtering properties. Filter invertibility.
- UK house prices:
  (1) Periods with substantially different spatial processes (different weighting schemes).
  (2) The magnitude of London's dominance over the other regions as described by Holly et al. [2011] is remarkably different across time.

Future Developments

(1) Including lagged exogenous regressors (i.e. W$\mathbf{X}$), leading to a time–varying spatial Durbin specification of the model, or spatially autocorrelated shocks (i.e. Wε_t), leading to a time–varying spatial autoregressive model with autoregressive (and eventually heteroskedastic) disturbances, can be addressed at the cost of additional calculation starting from our specification.

(2) **Forecasting** spatial-temporal structures

Thank you for your attention!

References I

Natalia Bailey, Sean Holly, and M Hashem Pesaran. A two-stage approach to spatio-temporal analysis with strong and weak cross-sectional dependence. *Journal of Applied Econometrics*, 31(1): 249–280, 2016.

Francois Bavaud. Models for spatial weights: a systematic look. *Geographical analysis*, 30(2):153–171, 1998.

Arnab Bhattacharjee and Chris Jensen-Butler. Estimation of the spatial weights matrix under structural constraints. *Regional Science and Urban Economics*, 43(4):617–634, 2013.

Anna Gloria Billé and Samantha Leorato. Quasi-ml estimation, marginal effects and asymptotics for spatial autoregressive nonlinear models. Technical report, Faculty of Economics and Management at the Free University of Bozen, 2017.

Anna Gloria Billé, Roberto Benedetti, and Paolo Postiglione. A two-step approach to account for unobserved spatial heterogeneity. *Spatial Economic Analysis*, 12(4):452–471, 2017.

References II

Francisco Blasques, Siem Jan Koopman, and Andre Lucas. Maximum likelihood estimation for generalized autoregressive score models. Technical report, Tinbergen Institute, 2014.

Ryan R Brady. Measuring the diffusion of housing prices across space and over time. *Journal of Applied Econometrics*, 26(2):213–231, 2011.

Luisa Corrado and Bernard Fingleton. Where is the economics in spatial econometrics? *Journal of Regional Science*, 52(2):210–239, 2012.

Drew Creal, Siem Jan Koopman, and André Lucas. Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5):777–795, 2013.

Federal Reserve Bank of St. Louis. Oecd based recession indicators for the united kingdom from the peak through the trough [gbrrecdm], retrieved from fred, federal reserve bank of st. louis; `https://fred.stlouisfed.org/series/GBRRECDM`. Technical report, 2018.

Solmaria Halleck Vega and J Paul Elhorst. The slx model. *Journal of Regional Science*, 55(3):339–363, 2015.

References III

Andrew C Harvey. *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Cambridge University Press, 2013.

Sean Holly, M Hashem Pesaran, and Takashi Yamagata. A spatio-temporal model of house prices in the usa. *Journal of Econometrics*, 158(1):160–173, 2010.

Sean Holly, M Hashem Pesaran, and Takashi Yamagata. The spatial and temporal diffusion of house prices in the uk. *Journal of urban economics*, 69(1):2–23, 2011.

Harry H Kelejian and Gianfranco Piras. Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes. *Regional Science and Urban Economics*, 46:140–149, 2014.

Harry H Kelejian and Ingmar R Prucha. Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157(1):53–67, 2010.

References IV

Lung-fei Lee and Jihai Yu. Qml estimation of spatial dynamic panel data models with time varying spatial weights matrices. *Spatial Economic Analysis*, 7(1):31–74, 2012.

James P LeSage and R Kelley Pace. A matrix exponential spatial specification. *Journal of Econometrics*, 140(1):190–214, 2007.

James P LeSage and R Kelley Pace. The biggest myth in spatial econometrics. *Econometrics*, 2(4):217–249, 2014.

Daniel P McMillen. Perspectives on spatial econometrics: linear smoothing with structured models. *Journal of Regional Science*, 52(2): 192–209, 2012.

Keith Ord. Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70(349):120–126, 1975.

Philipp Otto and Rick Steinert. Estimation of the spatial weighting matrix for spatiotemporal data under the presence of structural breaks. *arXiv preprint arXiv:1810.06940*, 2018.

Xi Qu and Lung-fei Lee. Estimating a spatial autoregressive model with an endogenous spatial weight matrix. *Journal of Econometrics*, 184(2): 209–232, 2015.

References V

Hajime Seya, Yoshiki Yamagata, and Morito Tsutsumi. Automatic selection of a spatial weight matrix in spatial econometrics: Application to a spatial hedonic approach. *Regional Science and Urban Economics*, 43(3):429–444, 2013.

Wei Shi and Lung-fei Lee. A spatial panel data model with time varying endogenous weights matrices and common factors. *Regional Science and Urban Economics (to appear)*, 2017.

Wei Wang and Jihai Yu. Estimation of spatial panel data models with time varying spatial weights matrices. *Economics Letters*, 128:95–99, 2015.

Peter Whittle. On stationary processes in the plane. *Biometrika*, pages 434–449, 1954.