WIFO

ÖSTERREICHISCHES INSTITUT
FÜR WIRTSCHAFTSFORSCHUNG

WORKING PAPERS

A Bayesian Spatial Autoregressive Logit Model With An Empirical Application to European Regional FDI Flows

Tamás Krisztin, Philipp Piribauer

586/2019

A Bayesian Spatial Autoregressive Logit Model With An Empirical Application to European Regional FDI Flows

Tamás Krisztin, Philipp Piribauer

WIFO Working Papers, No. 586

August 2019

Abstract

In this paper we propose a Bayesian estimation approach for a spatial autoregressive logit specification. Our approach relies on recent advances in Bayesian computing, making use of Pólya-Gamma sampling for Bayesian Markov-chain Monte Carlo algorithms. The proposed specification assumes that the involved log-odds of the model follow a spatial autoregressive process. Pólya-Gamma sampling involves a computationally efficient treatment of the spatial autoregressive logit model, allowing for extensions to the existing baseline specification in an elegant and straightforward way. In a Monte Carlo study we demonstrate that our proposed approach significantly outperforms existing spatial autoregressive probit specifications both in terms of parameter precision and computational time. The paper moreover illustrates the performance of the proposed spatial autoregressive logit specification using pan-European regional data on foreign direct investments. Our empirical results highlight the importance of accounting for spatial dependence when modelling European regional FDI flows.

E-mail address: philipp.piribauer@wifo.ac.at
2019/999/W/0

Medieninhaber (Verleger), Hersteller: Österreichisches Institut für Wirtschaftsforschung • 1030 Wien, Arsenal, Objekt 20 • Tel. (43 1) 798 26 01-0 • Fax (43 1) 798 93 86 • http://www.wifo.ac.at/ • Verlags- und Herstellungsort: Wien

A Bayesian spatial autoregressive logit model with an empirical application to European regional FDI flows

Tamás Krisztin[1] and Philipp Piribauer*[2]

[1]International Institute for Applied Systems Analysis (IIASA)
[2]Austrian Institute of Economic Research (WIFO)

Abstract

In this paper we propose a Bayesian estimation approach for a spatial autoregressive logit specification. Our approach relies on recent advances in Bayesian computing, making use of Pólya-Gamma sampling for Bayesian Markov-chain Monte Carlo algorithms. The proposed specification assumes that the involved log-odds of the model follow a spatial autoregressive process. Pólya-Gamma sampling involves a computationally efficient treatment of the spatial autoregressive logit model, allowing for extensions to the existing baseline specification in an elegant and straightforward way. In a Monte Carlo study we demonstrate that our proposed approach significantly outperforms existing spatial autoregressive probit specifications both in terms of parameter precision and computational time. The paper moreover illustrates the performance of the proposed spatial autoregressive logit specification using pan-European regional data on foreign direct investments. Our empirical results highlight the importance of accounting for spatial dependence when modelling European regional FDI flows.

Keywords: Spatial autoregressive logit, Bayesian MCMC estimation, FDI flows, European regions

JEL Codes: C11, C21, C25, F23, R11, R30

*Corresponding author: Philipp Piribauer, Austrian Institute of Economic Research (WIFO), Arsenal 20, 1030 Vienna, Austria. *E-mail*: philipp.piribauer@wifo.ac.at. The research carried out in this paper was supported by funds of the Oesterreichische Nationalbank (Jubilaeumsfond project number: 18116).

1 Introduction

In the regional economic literature, the logit model represents a popular choice for capturing discrete choice behaviours, for example, in explaining foreign direct investment (FDI) decisions of multinational enterprises (Ascani et al., 2016; Crescenzi et al., 2013). These specifications moreover allow for individual-level heterogeneity by relying on fatter distributional tails as compared to competing probit specifications (see, for example, Cameron and Trivedi, 2005), and assume that individual observations are independent. When modelling spatial data, however, the latter assumption are often violated, since regional random choice observations frequently exhibit spatial dependence (Calabrese and Elkink, 2014). Spatial dependence in choice outcomes results in a situation where observed choices at one location are similar to choices made at nearby locations. Particularly for regional economic applications, spatial dependence thus appears generally more likely than spatial independence (LeSage and Pace, 2009). However, it is worth noting that neglecting spatial dependence among the observations might result in severely biased and inconsistent estimates (Anselin, 2002).

The importance of controlling for spatial dependence when modelling spatial data in the random choice framework has been well documented using the probit model (see, for example, Anselin, 2002; McMillen, 1992; Smith and LeSage, 2004). Proposed estimation techniques rely both on methods of general moments (GMM) (McMillen, 1992; Anselin, 2002), as well as Bayesian Markov-chain Monte Carlo (MCMC) estimation (Smith and LeSage, 2004). The Bayesian estimation technique exploits the (non-spatial) approach of Albert and Chib (1993), who treat the observed binary outcome variables as indicators relating to a latent (unobserved) level of utility, which is directly in line with the random choice literature. This latent utility approach allows for an implementation in MCMC algorithms using draws from a truncated normal distribution. Spatial autoregressive variants by LeSage (2000) and Smith and LeSage (2004) instead rely on draws from truncated multivariate normal distributions. An overview on spatial discrete choice models is given by Smirnov (2010). In contrast, using the same procedure for the spatial autoregressive logit model have been held back by the difficulty of the logit log-likelihood incorporating multiple intractable integrals. Two notables exceptions (Brasington et al., 2016; Klier and McMillen, 2008) rely on a GMM estimation approach. This, however, suffers from the usual shortcoming of GMM-related approaches, i.e. for the estimation it requires instrumental variables and the parameter space of the spatial autocorrelation coefficient is not bounded.

The current paper builds on recent advances in Bayesian estimation of binomial-type model specifications (Polson et al., 2013), which rely on introducing a latent Pólya-Gamma distributed variable to facilitate Bayesian Markov-chain Monte Carlo estimation. Our version of the logit model allows for spatial dependence in the log-odds of the logit process. A particular advantage of this approach is that conditional on the latent Pólya-Gamma variable, the resulting conditional posterior distribution of the slope parameters is Gaussian, hence standard sampling procedures can be employed. Additionally, sampling for both the latent variable as well as for the slope parameter is computationally efficient. These virtues of the proposed spatial logit specification appear particularly useful, since recent advances in the spatial econometric literature rely on more flexible model specifications, which can be incorporated in such a modelling framework in a straightforward and computationally efficient way. Examples of such flexible potential extensions include explicitly allowing for non-linearity in the parameters (Cornwall and Parent, 2017; LeSage and Chih, 2018; Koch and Krisztin, 2011; Krisztin, 2017; 2018; Piribauer, 2016), multivariate spatial econometric frameworks (Crespo Cuaresma et al. 2018), shrinkage approaches for big data applications (Pfarrhofer and Piribauer, 2019; Piribauer and Crespo Cuaresma, 2016), uncertainty about the nature of spatial spillovers (Vega and Elhorst, 2013; LeSage and Fischer, 2008), or allowing for continuous spatial effects (Laurini 2017).

Due to the resulting normally distributed conditional posterior distribution of the slope parameters, the proposed spatial autoregressive logit specification might be extended to allow for more flexible and hierarchical model specifications in a straightforward way. We demonstrate the appealing virtues of the proposed approach – both in terms of parameter precision and computing time – in a series of Monte Carlo studies. Alternative benchmark specifications include a standard (non-spatial) logit model, a classic linear spatial autoregressive model estimated directly on the binary data, the spatial autoregressive specification of Klier and McMillen (2008), as well as a spatial autoregressive probit model specification (LeSage et al., 2011; LeSage and Pace, 2009).

We moreover use European regional data on FDI activities to illustrate the performance of the proposed Bayesian spatial autoregressive logit specification. In the empirical application we utilise information on FDI press announcements from the *fDi Markets* database, which is maintained by a specialist division of Financial Times Ltd. Specifically, we aim at modelling the occurrence of greenfield FDI investments in host regions in Europe along different stages of the value chain by explicitly accounting for spatial autoregressive log-odds and regional neighborhood characteristics. To illustrate the flexibility of the Bayesian estimation approach we allow for uncertainty over the number of nearest neighbors in the spatial specification.

The paper is structured as follows. Section 2 outlines the characteristics of our Bayesian spatial autoregressive logit model specification. Section 3 describes the Bayesian Markov-chain Monte Carlo estimation strategy employed to estimate the model. To demonstrate the efficiency both in terms of parameter precision as well as computational aspects, Section 4 presents Monte Carlo studies demonstrating the efficiency of the proposed approach, by benchmarking the results of our proposed approach to several other well-known specifications. Section 5 contains an applied illustration using regional NUTS-2 data on FDI investments in Europe. The final section concludes.

2 Model specification

Let y_i (for $i = 1, \ldots, N$) denote a binary outcome for a region i between two, mutually exclusive events. We aim to model the probability of making choice 1, which we denote as $p(y_i = 1)$. Based on the logit framework, we may model this probability as a function of the log-odds μ_i:

$$p(y_i = 1) = \frac{\exp \mu_i}{1 + \exp \mu_i}. \tag{2.1}$$

In a standard framework, μ_i is typically specified as a linear combination of a matrix of explanatory variables and a corresponding vector of unobserved slope parameters.

Particularly for (small-scale) regional data, the assumption of spatial dependence of observations might be more reasonable as compared to spatial independence (LeSage and Pace, 2009). In order to introduce spatial dependence among the regions we may assume that the log-odds μ_i in Eq. (2.1) do not solely depend on the characteristics of own region i, but also on other regions' characteristics. In the spatial econometric literature, such regional interdependencies are usually incorporated using non-negative and row-stochastic spatial weight matrices. A spatial weight matrix $\boldsymbol{W}$ contains information on the $N \times N$ spatial linkages between the regions in the sample and is usually treated as being given. Specifically, $w_{ij} > 0$ for $i \neq j$ if region i and j are assumed as being neighbors, otherwise $w_{ij} = 0$. Since no region is assumed as being a neighbor to itself, $w_{ii} = 0$ for all i.

The core part of the spatial autoregressive (SAR) logit model is given by:

$$\begin{aligned}\boldsymbol{\mu} &= \rho \boldsymbol{W} \boldsymbol{\mu} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\mu} &= \boldsymbol{A}^{-1}(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}),\end{aligned} \tag{2.2}$$

with $\boldsymbol{A}^{-1} = (\boldsymbol{I}_N - \rho \boldsymbol{W})^{-1}$ where $\boldsymbol{I}_N$ denotes an $N \times N$ identity matrix. The $N \times K$ matrix $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K]$ collects the K vectors of explanatory variables and $\boldsymbol{\beta}$ denote the respective $K \times 1$ vector of slope parameters. The $N \times 1$ vector $\boldsymbol{\varepsilon}$ contains independently and identically Gaussian distributed disturbance terms, with zero mean and σ^2 variance. The (scalar) parameter ρ measures the strength of spatial autocorrelation with sufficient stability condition $\rho \in (-1, 1)$, where positive (negative) values of ρ indicate positive (negative) spatial autocorrelation. In the absence of spatial autocorrelation ($\rho = 0$), the model framework collapses to a classical linear setup.

In such a spatial autoregressive (SAR) model specification, the $N \times 1$ vector of log-odds $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_N]'$ thus also depends on the characteristics of other regions in the sample. Spatial dependence is introduced by the spatial multiplier $\boldsymbol{A}^{-1} = (\boldsymbol{I}_N - \rho \boldsymbol{W})^{-1} = \sum_{r=0}^{\infty} \rho^r \boldsymbol{W}^r$.[1]

Note, that the residuals of the SAR logit model consist of two disinct components: first, the heteroscedastic errors arising from the logistic model in Eq. (2.1) and second, a normally distributed error term $\boldsymbol{\varepsilon}$ with σ^2 variance. Spatial dependence in the errors is captured through the latter. Similar as in the spatial variant of the probit model (see LeSage et al., 2011), the variance σ^2 is restricted to unity, in order to correctly identify the logistic errors.

The likelihood of the parameter vector $\boldsymbol{\beta}$, given an $N \times 1$ vector of observed choices $\boldsymbol{y}$ can be written as:

$$\mathcal{L}(\boldsymbol{\beta}|\boldsymbol{y}) = \prod_{i=1}^{N} \frac{(\exp \mu_i)^{y_i}}{1 + \exp \mu_i}, \tag{2.3}$$

[1]It is worth noting that the standard SAR model can be extended to more flexible spatial econometric model specifications in a straightforward way. Specifically, one may additionally include spatially lagged explanatory variables, resulting in a so-called spatial Durbin model (SDM) specification (see, for example LeSage and Pace 2009). A similar extension is presented in the empirical exercise.

where $\mathcal{L}(\cdot)$ denotes the likelihood function. However, it is worth noting that the likelihood contribution of observation i relies on the product of the probabilities $p(y_i = 1)$ and $p(y_i = 0)$. This results in the well-known problem that the log-likelihood is not linear, which may severely complicate parameter estimation and inference.

3 Estimation strategy

We propose estimation of the spatial autoregressive logit specification stated above using a Bayesian Markov-chain Monte Carlo approach. In order to efficiently deal with the logit framework, we follow an estimation strategy proposed in Polson et al. (2013) by introducing a latent Pólya-Gamma variable. The introduction of this latent quantity allows to recast the conditional posterior distribution of the slope parameters in the logit framework to a Gaussian distribution. Conditional on this latent Pólya-Gamma variable we may thus use standard Markov-chain Monte Carlo algorithms for spatial autoregressive methods in a straightforward and flexible way (see, for example, LeSage 1997, or LeSage and Pace 2009).

Polson et al. (2013) show that the introduction of a Pólya-Gamma variable can severely facilitate Bayesian estimation of binomial-type model specifications. A particularly useful result of the work by Polson et al. (2013) is given by the following identity:

$$\frac{(\exp \mu_i)^a}{(1+\exp \mu_i)^b} = 2^{-b} \exp(\kappa_i \mu_i) \int_0^\infty \left(\exp \frac{-\omega_i \mu_i^2}{2} \right) p(\omega_i) d\omega_i \qquad a, \mu_i \in \mathbb{R},\ b \in \mathbb{R}^+, \tag{3.1}$$

where $\kappa_i = a - b/2$ and ω_i is a Pólya-Gamma distributed random variable with scale b and location parameter zero, $p(\omega_i) \sim \mathcal{PG}(b, 0)$. Note that the above integral identity does not rely on numerical approximation of the binomial term, and can be estimated by sampling from the conditional posterior Pólya-Gamma distribution. Using the results in Polson et al. (2013) and Windle et al. (2014), the conditional posterior for $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_N]'$ also takes the form of a Pólya-Gamma distribution:

$$p(\boldsymbol{\omega}|\boldsymbol{\beta}, \rho, \boldsymbol{y}) = \mathcal{PG}\left(1, \boldsymbol{A}^{-1}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{A}^{-1}\boldsymbol{\varepsilon}\right). \tag{3.2}$$

Computationally efficient sampling algorithms for the Pólya-Gamma conditional posterior in Eq. (3.2) are provided by Polson et al. (2013) and Windle et al. (2014) and are implemented in the R package **BayesLogit**.

Combining Eq. (3.1) with Eq. (2.3) and setting $a = y_i$ and $b = 1$ yields the contribution to the likelihood of observation i in our specific econometric setting:

$$\frac{\exp(\mu_i)^{y_i}}{1+\exp(\mu_i)} \propto \exp(\kappa_i \mu_i) \int_0^\infty \left(\exp \frac{-\omega_i \mu_i^2}{2} \right) p(\omega_i) d\omega_i. \tag{3.3}$$

By conditioning on $\boldsymbol{\omega}$ and the spatial autoregressive parameter ρ, the conditional posterior for $\boldsymbol{\beta}$ is thus given by:

$$p(\boldsymbol{\beta}|\rho, \boldsymbol{\omega}, \boldsymbol{y}) \propto p(\boldsymbol{\beta}) \prod_{i=1}^{N} \exp(\kappa_i \mu_i - \omega_i \mu_i^2/2) \tag{3.4}$$

$$\propto p(\boldsymbol{\beta}) \exp\left\{-\frac{1}{2}(\boldsymbol{A}\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta})' \boldsymbol{\Omega} (\boldsymbol{A}\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta})\right\}, \tag{3.5}$$

with $N \times N$ matrix $\boldsymbol{\Omega} = diag(\omega_1, ..., \omega_N)$ and $\boldsymbol{z} = [\kappa_1/\omega_1, ..., \kappa_N/\omega_n]'$. By eliciting a normally distributed prior density for $\boldsymbol{\beta}$ with prior mean $\underline{\boldsymbol{\mu}}_\beta$ and variance $\underline{\boldsymbol{\Sigma}}_\beta$, $p(\boldsymbol{\beta}) \sim \mathcal{N}(\underline{\boldsymbol{\mu}}_\beta, \underline{\boldsymbol{\Sigma}}_\beta)$, the conditional posterior distribution for the slope parameters $\boldsymbol{\beta}$ is therefore Gaussian:

$$p(\boldsymbol{\beta}|\rho, \boldsymbol{\omega}, \boldsymbol{y}) = \mathcal{N}(\overline{\boldsymbol{\mu}}_\beta, \overline{\boldsymbol{\Sigma}}_\beta), \tag{3.6}$$

with posterior quantities $\overline{\boldsymbol{\mu}}_\beta = \boldsymbol{\Sigma}_\beta \left(\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{A}\boldsymbol{z} + \underline{\boldsymbol{\Sigma}}_\beta^{-1}\underline{\boldsymbol{\mu}}_\beta\right)$ and $\overline{\boldsymbol{\Sigma}}_\beta = \left(\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{X} + \underline{\boldsymbol{\Sigma}}_\beta^{-1}\right)^{-1}$. Equation (3.6) shows the particular advantage of the introduction of the latent Pólya-Gamma variable. Due to the fact that given $\boldsymbol{\omega}$ the conditional posterior distribution for $\boldsymbol{\beta}$ is normally distributed, the proposed framework may be extended to hierarchical or more flexible specifications in a straightforward way. This appears to be appealing for extending binomial-type model specifications to allow for recent advances in spatial econometric literature, such as but not limited to, including uncertainty in spatial structure (LeSage and Fischer 2008, LeSage and Pace 2007), uncertainty about the explanatory variables (Pfarrhofer and Piribauer 2019, Crespo Cuaresma et al. 2018, Piribauer and Crespo Cuaresma 2016), or flexible specifications for parameter heterogeneity (LeSage and Chih 2018, Cornwall and Parent 2017, Piribauer 2016).

Conditional on $\boldsymbol{\omega}$ using Theorem 1 from Polson et al. (2013), the conditional posterior distribution for the spatial autoregressive parameter ρ is given by:

$$p(\rho|\boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{y}) \propto |\boldsymbol{A}| \exp\left\{-\frac{1}{2}(\boldsymbol{A}\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta})' \boldsymbol{\Omega} (\boldsymbol{A}\boldsymbol{z} - \boldsymbol{X}\boldsymbol{\beta})\right\} p(\rho), \tag{3.7}$$

where $p(\rho)$ denotes the prior density of ρ. Standard prior choices for ρ involve a uniform or a beta distribution (see, for example, LeSage and Pace 2009). However, the conditional posterior for ρ is not reducible to a well-known distribution which can easily be sampled from. We therefore use a griddy Gibbs step (Ritter and Tanner 1992) in order to sample from the conditional posterior for ρ. This can be easily achieved using the numerical integration procedure as in LeSage and Pace (2009).[2]

[2]An alternative, however, computationally more intensive approach also frequently used in the spatial econometric literature involves a Metropolis-Hastings step for the spatial autoregressive parameter (see, for example, LeSage and Pace 2009).

Markov-chain Monte Carlo sampling procedure

Given the conditional posterior distributions stated above, Markov-chain Monte Carlo algorithms can be employed by sequentially sampling from the conditional posteriors. With suitable starting values for $\boldsymbol{\beta}$ and ρ, our sampler involves the following steps:

I. Update $\boldsymbol{\omega}$ by drawing from $p(\boldsymbol{\omega}|\boldsymbol{\beta}, \rho, \boldsymbol{y})$ using Eq. (3.2),

II. Update $\boldsymbol{\beta}$ by drawing from $p(\boldsymbol{\beta}|\rho, \boldsymbol{\omega}, \boldsymbol{y})$ using Eq. (3.6),

III. Update ρ using a griddy Gibbs step from $p(\rho|\boldsymbol{\beta}, \boldsymbol{\omega}, \boldsymbol{y})$ based on Eq. (3.7).

The Markov-chain Monte Carlo algorithm cycles through steps I. to III. B times by excluding the first B_0 draws as burn-ins. Inference on the parameters is conducted using the $B - B_0$ remaining draws.[3]

4 Simulation study

To assess the accuracy of our model, we benchmark its performance with regards to several competing model specifications on a set of artificially generated data in a series of Monte Carlo studies[4]. Our benchmark data generating process is a spatial autoregressive logit model with a constant and containing two randomly generated explanatory variables:

$$\begin{aligned}\tilde{y}_i &= Binom\left(1, \frac{\exp \tilde{\mu}_i}{1 + \exp \tilde{\mu}_i}\right),\\ \tilde{\boldsymbol{\mu}} &= \tilde{\boldsymbol{A}}^{-1}\left(\tilde{\beta}_0 + \tilde{\boldsymbol{x}}_1\tilde{\beta}_1 + \tilde{\boldsymbol{x}}_2\tilde{\beta}_2 + \tilde{\boldsymbol{\varepsilon}}\right),\end{aligned}$$

where $\tilde{\boldsymbol{A}} = \left(\boldsymbol{I}_N - \tilde{\rho}\tilde{\boldsymbol{W}}\right)$ and $Binom(\cdot, \cdot)$ denotes the binomial distribution. To maintain succinct notation we label the simulated true values in the Monte Carlo application with a tilde. The explanatory variables $\tilde{\boldsymbol{x}}_1$ and $\tilde{\boldsymbol{x}}_2$ are both normally distributed, with zero mean and unity variance. For each Monte Carlo iteration we randomly generate $\tilde{\beta}_0$, $\tilde{\beta}_1$, and $\tilde{\beta}_2$ from a normal distribution with standard deviation 0.05 and means of 0.5, 1, and -1, respectively. The vector of residuals $\tilde{\boldsymbol{\varepsilon}}$ is generated from a normal distribution, with zero mean and unity variance. The row-stochastic spatial weight matrix $\tilde{\boldsymbol{W}}$ is constructed using 5-nearest neighbors based on a randomly generated spatial location pattern, stemming from a normal distribution with zero mean and unity variance. We vary the strength of spatial dependence $\tilde{\rho} \in [0, 0.5, 0.8]$. Additionally, to evaluate the performance of the samplers with regard to sample size, we let $N \in [400, 1000]$.

For the Monte Carlo simulation study, we compare the following five model specifications:

1. *SAR Logit*: The Bayesian spatial autoregressive logit specification sketched above.

[3] Convergence of the MCMC algorithm was checked using the convergence diagnostics proposed by Geweke (1991) and Raftery and Lewis (1992). Convergence diagnostics have been calculated using the R package **coda**.

[4] Detailed R-codes are available from the authors upon request.

2. *GMM SAR Logit*:[5] The GMM spatial autoregressive estimation approach introduced in Klier and McMillen (2008).

3. *Linearized GMM SAR Logit*:[6] The linearized version of the GMM spatial autoregressive estimation approach, which is advocated as a resonable approximation of the computationally expensive *GMM SAR Logit* by Klier and McMillen (2008).

4. *SAR Probit*: A Bayesian spatial autoregressive probit specification put forward by LeSage et al. (2011).

5. *SAR*: A Bayesian version of a standard spatial autoregressive model specification (LeSage and Pace 2009), assuming normally distributed errors with σ^2 variance, without imposing restrictions to the binary nature of the dependent variable: $\boldsymbol{y} = \boldsymbol{A}^{-1}(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})$, with $\boldsymbol{\varepsilon} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_N)$.

6. *Logit*: A standard logit specification assuming no spatial autocorrelation ($\rho = 0$).

Our prior set-up is as follows. We use a Gaussian prior for $\boldsymbol{\beta}$ with zero mean and variance 10^8 and for ρ we use a four-parameter Beta prior specification as outlined in LeSage and Pace (2009), with bounds set equal to -1 and 1 and shape parameters 1.01. For the *SAR* specification we use an inverse-gamma prior for σ^2 with rate and shape parameters 0.1. This corresponds to weakly informative prior distributions for all of our parameters. In the Monte Carlo study we benchmark the root mean squared error of the estimated spatial autoregressive parameter, as well as the summary spatial impact metrics to their true values. Since the data generating process in the simulation study involves a spatial lag, we follow LeSage and Pace (2018) and report predictive performances with respect to spatial impact metrics in the spirit of LeSage and Pace (2009) rather than focussing on slope parameters. In addition to the magnitude of spatial dependence ρ, we moreover report average direct and indirect (spillover) impacts. Both metrics are given by a non-

[5] Estimates for GMM SAR Logit have been produced using the R package **McSpatial**. However, for high spatial autocorrelation $\tilde{\rho} = 0.8$, GMM SAR Logit appeared to have severe problems to produce any estimates in more than 99% of all simulation runs. We have therefore omitted GMM SAR Logit in the simulation study for $\tilde{\rho} = 0.8$.

[6] Estimates for Linearized GMM SAR Logit have been produced using the R package **McSpatial**. In this setting, however, it is worth noting that estimates of ρ may exceed unity. In these cases we have restricted the estimate for ρ to 0.99 for calculation of spatial impact metrics.

linear combination of the slope parameters, ρ, and $\boldsymbol{W}$. Details on the calculation of the impact metrics can be found in the appendix.

Table 1: Root mean squared error measures for the Monte Carlo runs

N	Model	$\tilde{\rho} = 0.0$			$\tilde{\rho} = 0.5$			$\tilde{\rho} = 0.8$		
		direct	indirect	ρ	direct	indirect	ρ	direct	indirect	ρ
400	SAR Logit	0.097	0.033	0.081	0.169	0.323	0.344	**0.359**	**1.424**	**0.243**
	GMM SAR Logit	0.118	0.101	0.258	**0.136**	**0.237**	**0.234**			
	Linearized GMM SAR Logit	0.105	0.096	0.232	0.222	12.915	0.334	0.539	43.039	0.897
	SAR Probit	0.182	0.029	0.128	0.191	0.364	0.369	0.509	1.529	0.482
	SAR	0.317	0.027	0.076	0.335	0.393	0.420	0.416	1.580	0.493
	Logit	**0.079**	**0.000**	**0.000**	0.247	0.417	0.500	0.960	1.667	0.800
1,000	SAR Logit	0.056	0.022	0.052	**0.088**	0.284	0.261	0.383	**1.435**	**0.221**
	GMM SAR Logit	0.063	0.066	0.152	0.103	**0.199**	**0.194**			
	Linearized GMM SAR Logit	0.059	0.070	0.147	0.187	9.270	0.285	0.575	44.066	0.886
	SAR Probit	0.184	0.018	0.078	0.188	0.346	0.302	**0.305**	1.506	0.431
	SAR	0.313	0.019	0.052	0.331	0.384	0.376	0.416	1.570	0.476
	Logit	**0.050**	**0.000**	**0.000**	0.229	0.417	0.500	0.871	1.657	0.800

Notes: Results are based on 1,000 Monte Carlo iterations. For each Monte Carlo iteration, the corresponding sampling algorithms are run using 1,000 draws, where the initial 700 draws were discarded as burn-in. The columns *direct* and *indirect* correspond to summary marginal effects (for details, see the Appendix). The values given for *direct*, *indirect*, and ρ correspond to the average RMSE($\cdot$) over all Monte Carlo iterations. Bold values denote the lowest average RMSE scores.

The results of the Monte Carlo study are presented in Table 1. For each scenario and model specification, the reported metrics in the table correspond to 1,000 Monte Carlo runs. Information on the sample size N and the alternative model specifications used are shown in the first and second columns. The table reports average root mean squared error point estimates for the average direct and indirect impacts as well as ρ corresponding to $\tilde{\rho} = [0.0, 0.5, 0.8]$.

In the case of no spatial autocorrelation ($\tilde{\rho} = 0$), Table 1 shows particular appealing results for the standard logit specification (*Logit*). This result holds true for both moderate and large sample sizes. Since *Logit* resembles the true data generating process for $\tilde{\rho} = 0$, these results appear little surprising. However, for both sample sizes considered, our spatial autoregressive logit specification (*SAR Logit*) appears to produce the second-best solution for predicting the summary direct effects by closely tracking the reported root mean squared errors for the standard logit case. However, for indirect effects and for the spatial autoregressive parameter ρ both the *SAR* and *SAR Probit* specifications also appears to slightly outperform *SAR Logit*.

In terms of predicting the spatial dependence parameter in the absence of spatial autocorrelation in the data generating process ($\tilde{\rho} = 0$), the results in the table moreover reveal that both reported GMM-based approaches (*GMM SAR Logit* and *Linearized GMM SAR Logit*) perform rather poorly as compared to the alternative specifications. This outperformance of the alternative approaches can be seen for both moderate ($N = 400$) and large sample sizes ($N = 1,000$).

In terms of point predictions for moderate spatial autocorrelation ($\tilde{\rho} = 0.5$), results for direct and indirect effects significantly change. For moderate sample sizes ($N = 400$) the *GMM SAR Logit* produces the most precise predictions, closely tracked by *SAR Logit*. In the case of large sample sizes ($N = 1,000$) the results appear more mixed, with *GMM SAR Logit* producing the most precise predictions for indirect effects, while the *SAR Logit* outperforms all other competing specifications in predicting direct effects. In terms of predicting the spatial autoregressive parameter ρ the *GMM SAR Logit* exhibits the lowest bias in terms of RMSE for both sample sizes considered.

Note, that the *Linearized GMM SAR Logit* performs particularly poorly in the case of moderate to high spatial autocorrelation when predicting indirect effects. This is due to the fact that estimate for ρ produced by the linearized GMM specifications is not bounded. While we impose a bound for the indirect effect calculations an estimate for ρ close to one still results in severely biased indirect effects, as evidenced by the results.

In settings with high spatial autocorrelation ($\tilde{\rho} = 0.8$), the outperformance of *SAR Logit* compared to the competing specifications appears particularly pronounced. As an only exception the *SAR Probit* specification slightly outperforms the *SAR Logit* in terms of predicting direct effects in the case of large sample sizes ($N = 1,0000$).

To sum up, Table 1 demonstrates that the proposed *SAR Logit* specification produces particularly appealing results in the presence of spatial autocorrelation. The standard (non-spatial) logit model specification is preferable only in the absence of spatial autocorrelation. *GMM SAR Logit* outperforms *SAR Logit* only in the presence of moderate spatial autocorrelation. It is moreover worth noting, that in the presence of high spatial autocorrelation the *GMM SAR Logit* algorithm failed to converge in the vast majority of the draws.

5 An empirical illustration to modelling European subnational FDI

This section aims at empirically illustrating our proposed estimation approach for pan-European subnational foreign direct investment data. More specifically we aim at explaining the occurrence of FDI activity in a certain destination region along different stages of the value chain. The importance of attracting foreign direct investment for the purposes of fostering economic growth is empirically well documented (see, for example, Huber et al., 2017; Blonigen and Piger, 2014; Eicher et al., 2012). The determinants that appear particularly crucial in this aspect are trade-openness (Balasubramanyam et al., 1996), a well-developed human capital stock (Borensztein et al. 1998) or the degree of embeddedness of foreign firms in local economies (Markusen and Venables 2000; Rodriguez-Clare 1996) as well as the general business environment (Blomstrom and Kokko, 2003; Xu, 2000). Multiple studies emphasize the importance of spatial issues as well (see Baltagi et al., 2007; Ekholm et al., 2007, among others), emphasizing the role of global, as well as local spillovers.

Therefore, in the empirical specification we include spatially lagged explanatory variables as well, resulting in a so-called spatial Durbin log-odds specification. In the spatial econometric literature, the choice of the spatial weight matrix is often seen as being crucial. Since empirical results may be strongly affected by the choice of the spatial weight matrix, we therefore extend the model specification by allowing for ξ alternative spatial weight matrix specification. Our spatial Durbin specification can be written as:

$$\boldsymbol{\mu}_t = (\boldsymbol{I}_N - \rho \boldsymbol{W}(\xi)_t)^{-1}(\alpha_t + \boldsymbol{Z}_t\boldsymbol{\gamma} + \boldsymbol{W}_t(\xi)\boldsymbol{Z}_t\boldsymbol{\vartheta} + \boldsymbol{\varepsilon}), \tag{5.1}$$

where the (scalar) spatial autoregressive parameter ρ, and the $N \times N$ spatial weight matrix $\boldsymbol{W}_t(\xi)$, conditional on the parameter ξ, are defined as before. The positive discrete scalar parameter $\xi \in \{1, w\}$ acts as an indicator to choose between w alternative spatial weight matrices. Conditional on ξ, the entries in $\boldsymbol{W}_t(\xi)$ are assumed as being given. $\boldsymbol{Z}_t$ denotes a $N \times K$ matrix of explanatory variables at time period $t = 1, \ldots, T$ with corresponding $K \times 1$ vector of slope parameters $\boldsymbol{\gamma}$. α_t denotes a time period fixed effect and the term $\boldsymbol{W}_t(\xi)\boldsymbol{Z}_t$ is a spatial lag of the matrix of covariates with associated vector of parameters $\boldsymbol{\vartheta}$ and explicitly includes the regions' characteristics of their neighbors. Note that the spatial Durbin version of the logit specification given in Eq. (5.1) can be easily specified as a standard spatial autoregressive logit specification by collecting all the explanatory variables into a matrix $\boldsymbol{X}_t$. By vertically stacking $\boldsymbol{X}_t$ for all t along with the time fixed effects into a matrix $\boldsymbol{X}$ and defining $\boldsymbol{W}(\xi) = diag(\boldsymbol{W}_1(\xi), \ldots, \boldsymbol{W}_T(\xi))$, the equations for the Markov-chain Monte Carlo algorithm can be applied (conditional on ξ) as discussed before.

However, since the additional parameter ξ is treated as unknown, an additional sampling step for ξ has to be employed in the MCMC sampling algorithm. Similar to ρ, the conditional posterior distribution for ξ is also not of well-known form and is given by:

$$p(\xi|\boldsymbol{\beta}, \boldsymbol{\omega}, \rho, \boldsymbol{y}) \propto |\boldsymbol{A}(\xi)| \exp\left\{-\frac{1}{2}\left(\boldsymbol{A}(\xi)z - \boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{\varepsilon}\right)' \boldsymbol{\Omega}\left(\boldsymbol{A}(\xi)z - \boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{\varepsilon}\right)\right\} p(\xi), \tag{5.2}$$

with $\boldsymbol{A}(\xi) = \boldsymbol{I}_N - \rho\boldsymbol{W}(\xi)$. Sampling for the conditional posterior distribution is straightforward by for example employing an additional Metropolis-Hastings step using independent uniform proposals for ξ. Very similar sampling and estimation strategies to account for uncertainty among $\boldsymbol{W}$ have also been employed in Piribauer and Crespo Cuaresma (2016) or LeSage and Pace (2009).

5.1 Regions, data, and spatial weights

Our empirical application makes use of the *fDi Markets* database, which is maintained by fDi Intelligence – a specialist division of the Financial Times Ltd. Our data set provides information on regional FDI activities from 2003 to 2011. *fDi Markets* relies on media sources and company data, and reports information (among others) on the host and source (country, region and city) of the FDI flows, sector classifications, the amount of the capital investment, as well as the amount of created jobs. The FDI flows in the database contain all cross-border greenfield investments, and the inclusion of investments in the database is conditional on the FDI flow generating new employment or capital investments in the host region. The investments reported in the *fDi Markets* database have been subject to several robustness checks and comparisons with other data sources on FDI activities. The reported cross-checks support the reliability of the *fDi Markets* data set on the spatial distribution of FDI (Crescenzi et al. 2013).

The dependent variable comprises information on global FDI in-flows to 266 European NUTS-2 host regions in the period 2003 to 2011. The complete list of regions is presented in Table A2 in the Appendix. The binary dependent variable measures annually whether one or more FDI investment occurred for a panel of NUTS-2 host regions, which yields a total of 2,394 observations.

Observations on FDI flows contain information on 17 distinct business activities, which are associated with various stages of the value chain, from product research and development, through testing and manufacturing, up to to delivery and recycling. Investing companies at different stages of the value chain exhibit heterogeneous preference structure for regional endowments. This heterogeneity stems from unobserved characteristics of the investing firms (i.e. at their utility structure related to their relative position along the value chain) and provides valuable insights on investment decision processes (Ascani et al., 2016). To account for this source of heterogeneity, we differentiate between FDI flows based on their position along the value chain. This is achieved by grouping FDI inflows with respect to their associated business activities into three distinct categories: upstream, downstream, and production. The upstream category comprises conceptual product development, including design and testing, as well as management and business administration activities. The downstream category encompasses activities linked to the consumer, such as sales, product delivery or support. Finally, the production category contains activities associated with the actual product creation – such as manufacturing and extraction – as well as related activities – such as recycling. This classification is based on the general value stage chain classification by Sturgeon (2008) and are partly reminiscent to those employed by Crescenzi et al. (2013) and Ascani et al. (2016). Table A1 in the Appendix shows the complete list of investment activities and categories.

Our set of covariates comprises the main variables used in the prevalent literature on regional FDI activities and regional economic growth (see, for example, Pintar et al., 2016; Crescenzi et al., 2013; Crespo Cuaresma et al., 2014). Since a higher demand for goods should promote FDI activities which are launched in order to expand the local production of final goods, recent empirical studies predominantly focus on the market size of the host region. A large market size of the host region is therefore assumed to increase the probability to attract FDI investments. In national and regional studies of FDI activities, the corporate tax rate is considered as another key driver for attracting FDI flows (Bellak and Leibrecht, 2009). A higher corporate income tax is expected to deter investment, while lower tax rates are used as a national measure to attract FDI investments.

Additionally, typical supply-side quantities include measures of regional human, and knowledge capital endowments. To approximate regional human capital endowments we use two variables focussing on regions' educational attainment. One variable measures the share of population with tertiary education attainment, and a second is given by the share with low education. As a measure for regional knowledge capital endowments we use data on patent counts to proxy regional knowledge production. The use of regional data on patent activities appears reasonable since patents can be seen as a direct outcome of R&D investments (see, for example, LeSage and Fischer 2012). To construct regional knowledge capital stocks we make use of the perpetual inventory method. We therefore follow work by Fischer and LeSage (2015) and LeSage and Fischer (2012) by assuming $K_{i,t} = K_{i,t-1}(1 - r_K) + P_{i,t}$, where $K_{i,t}$ and $P_{i,t}$ denotes the knowledge capital stock and patent counts in region i and time period t, respectively. $r_K = 0.10$ denotes a constant

Table 2: Variables used in the empirical illustration

Variable	**Description**
Upstream	1 denotes that FDI in-flows associated with upstream business activities have been reported in the region, 0 otherwise. *Source*: *fDi Markets*
Downstream	1 denotes that FDI in-flows associated with downstream business activities have been reported in the region, 0 otherwise. *Source*: *fDi Markets*
Production	1 denotes that FDI in-flows associated with production business activities have been reported in the region, 0 otherwise. *Source*: *fDi Markets*
Market size	Proxied by means of regional gross value added, in log terms. *Source*: Cambridge Econometrics
Population density	Population per square km, in log terms. *Source*: Cambridge Econometrics
Corporate income tax	Country-specific top statutory corporate income tax rates (including surcharges) *Source*: Eurostat
Employment in industry	Share of NACE B to F (industry and construction) in total employment. *Source*: Cambridge Econometrics
Employment in services	Share of NACE G to U (services) in total employment. *Source*: Cambridge Econometrics
Lower education workers	Share of population (aged 25 and over) with lower education (ISCED levels 0-2). *Source*: Eurostat
Higher education workers	Share of population (aged 25 and over) with higher education (ISCED levels 6+). *Source*: Eurostat
Regional knowledge capital	Knowledge stock formation measured in terms of patent accumulation, in log terms. *Source*: Eurostat

Notes: ISCED and NACE refer to the international standard classification of education and the second revision of the statistical classification of economic activities in the European community, respectively.

depreciation rate of regional knowledge capital.[7] In addition to information on host regional market size, human and knowledge capital stocks, we moreover include several other covariates to control for the regional industry mix, population density, as well as other socio-economic characteristics typically found in the empirical regional economic literature (Crespo Cuaresma et al. 2014).

For the specification of the spatial weight matrix $\boldsymbol{W}$ we used a range of nearest neighbor type specification. Specifically, we use nearest neighbor spatial weight matrices ranging from one to 20 (i. e. $\xi \in \{1, 20\}$), by using a discrete uniform prior for ξ. Data on the explanatory variables used stem from both the *Cambridge Econometrics* regional database as well as the *Eurostat* regional database. Detailed information on the set of dependent and explanatory variables used in our spatial autoregressive logit framework is provided in Table 2.

5.2 Empirical results

In this subsection we present the Markov-chain Monte Carlo estimation results by using 20,000 posterior draws for our Bayesian spatial Durbin logit specification, from which 15,000 are discarded as burn-in. Table 3 presents empirical summary metrics for spatial models in the form of average direct and indirect (spillover) effects (LeSage and Pace 2009) for the explanatory variables under scrutiny. The presented summary impact measures are evaluated at the mean of the explanatory variables. Details for the calculation of average direct and indirect (spillover) effects in the SDM logit specifications are presented in the Appendix.

Average direct effects are reminiscent to non-spatial impact assessments by representing the average impacts to a region's dependent variable due to a marginal change of an explanatory variable in the same region. Average indirect (or spillover) effects represent the impact due to a marginal change in all other regions. In addition, one may also interpret average total effects which are given by the sum of direct and indirect effects and embody the average regional impact resulting from a marginal change in all regions in the sample. Estimation results are presented for host regional FDI flows separated for the value chain related upstream, downstream, and production activities in Table 3. In addition to the spatial impact metrics, the table also reports posterior quantities for the degree of spatial dependence ρ, the parameter ξ for the spatial neighborhood, and McFadden's pseudo R^2, a measure for the goodness of fit in logit specifications (McFadden, 1974) using posterior mean quantities.[8]

Columns (i) and (ii) in Table 3 show the impact metrics for upstream investment activities. Similar to the other investment activities, upstream FDI activities appear to exhibit positive and significant spatial autocorrelation. However, for upstream FDI spatial autocorrelation is somewhat weaker as compared to the other two activities. As expected, the overall market size appears to have a highly significant and positive direct impact on the probability for upstream FDI investments occuring in the host region. The estimated direct impact for the market size variable turns out

[7]In the empirical application, we have also used alternative values for r_K. The results, however, appeared rather robust for the choice of r_K.

[8]McFadden's pseudo R^2 is defined as $1 - \frac{\mathcal{L}_1}{\mathcal{L}_0}$, where $\mathcal{L}_1$ denotes the posterior log-likelihood of the fitted model and $\mathcal{L}_0$ the log-likelihood of a null mode containing only an intercept. Based on McFadden (1974) values between 0.2 and 0.4 are considered to be an excellent fit.

to be particularly large among the three value chain categories under scrutiny. This corroborates the findings of Henderson and Ono (2008), Defever (2006), or Duranton and Puga (2005), namely that such activities are driven more by functional rather than by sectoral aspects and tend to be located particularly in urban agglomerations. The relative importance of regional proximity to large metropolitan areas can also be seen by a positive and significant direct and indirect impact of the population density variable. The importance of proximity to urban agglomerations becomes even more accentuated due to the negative and highly significantly estimated indirect effect of the market size variable. This result suggests that a marginal increase in the market size solely in all other regions decreases the probability to attract upstream FDI activities in the own region.

Concerning the host region's industry mix, the table also shows positive direct impacts for both the industry as well as service sector shares. Interestingly, the average indirect effect for the employment share in the services sector has a negative sign. This means that an increase in service intensity in foreign regions decreases the probability of a region's FDI inflow. This finding is in line with Strauss-Kahn and Vives (2009), who highlight the importance of agglomerations in the same sector as well as a high level of business service activities for location decisions of headquarters. Strauss-Kahn and Vives (2009) also find that low corporate tax rates appear to be of particular importance. Table 3 also shows a negative direct impact estimate for the corporate tax rate. Interestingly, the estimated spillover impacts for the corporate tax rates is positive which corroborates the finding of tax competition (Bellak and Leibrecht 2009). An increase in tax rates only in foreign regions therefore has a positive impact on attracting upstream FDI activities.

While a low educated working age population appears to have an insignificant impact on attracting upstream FDI inflows, the impact of regional tertiary education attainment is positive and highly significant. However, spatial spillover impacts for both human capital variables are insignificant. Knowledge capital endowments also seem to be of importance for regions to attract FDI inflows. Interestingly, we find no positive direct effect for own regional knowledge capital endowments, a result also found by Dimitropoulou et al. (2013), who study the locational determinants of FDI in UK regions. However, our spatial autoregressive framework shows significant and positive average spillover impacts for regional knowledge capital. While the own regional knowledge capital endowments appears insignificant, spatial proximity to knowledge capital thus still plays a key role.

Concerning the number of nearest neighbors ξ, Table 3 reports rather precise posterior estimates for all three industry classifications. For upstream FDI, the posterior mean of ξ is with 10.037 higher and more precisely estimated than for downstream and production (posterior means of 6.062 and 7.597, respectively).

Columns (iii) and (iv) of Table 3 report posterior mean estimates for regional FDI for downstream activities. For this classification of host FDI flows, the posterior mean estimates for the spatial autoregressive parameter is positive and highly significant. Similar to upstream FDI, the direct impact of the market size variable is positive and the spatial spillover metrics also appear to be negative. Both impact measures, however, are a bit less pronounced compared to upstream activities. The somewhat less accentuated importance of regional proximity to urban agglomerations of downstream activities as compared to upstream FDI can also be seen by an insignificant direct effect of the population density variable. Albeit the own regional population density appears less important, the indirect impacts show a positive (and significant) posterior mean. In order to attract downstream FDI flows, spatial proximity to urban agglomerations thus also plays a major role. The direct impact of the corporate tax rate also exhibits a negative sign with positive spatial spillover effects. The spatial summary metrics for the regional sectoral structure for downstream activities are very similar to those for upstream FDI. Employment shares for both industry and service sectors are positively related to the probability of attracting FDI. Moreover, the table again indicates negative spatial spillover effects of the share of employment in services.

Table 3: Summary impact measures for value chain classifications

	Variable	Upstream		Downstream		Production	
		Mean	SD	Mean	SD	Mean	SD
		(i)	(ii)	(iii)	(iv)	(v)	(vi)
Direct	Market size	**1.497**	0.147	**1.460**	0.148	**1.255**	0.137
	Population density	**0.118**	0.075	0.060	0.057	**-0.046**	0.013
	Corporate tax rate	**-0.081**	0.039	**-0.152**	0.078	**-0.202**	0.096
	Employment in industry	**2.561**	1.319	**3.612**	1.566	**2.749**	1.523
	Employment in services	**4.036**	1.397	**3.940**	1.569	1.084	1.146
	Lower education workers	1.190	0.857	0.447	0.664	-0.457	0.378
	Tertiary education workers	**5.096**	1.436	**2.572**	1.168	-0.733	0.366
	Regional knowledge capital	-0.025	0.024	-0.004	0.042	-0.002	0.037
Indirect	Market size	**-0.101**	0.028	**-0.066**	0.027	**-0.132**	0.050
	Population density	**0.147**	0.086	**0.254**	0.095	**0.268**	0.103
	Corporate tax rate	**6.438**	1.989	**3.205**	2.060	**7.894**	2.886
	Employment in industry	0.216	1.621	-0.723	1.497	1.972	2.561
	Employment in services	**-3.463**	0.728	**-2.572**	0.763	-0.598	1.661
	Lower education workers	-0.415	0.738	0.362	0.668	0.568	0.832
	Tertiary education workers	0.753	1.182	-0.148	1.045	**2.342**	1.417
	Regional knowledge capital	**0.173**	0.095	**0.231**	0.100	-0.008	0.105
	ρ	**0.163**	0.061	**0.195**	0.054	**0.381**	0.084
	ξ	**10.037**	1.403	**6.062**	1.891	**7.597**	2.292
	McFadden's R^2	0.319		0.254		0.190	
	Observations with positive inflows	1,027		1,465		1,328	
	Total number of observations	2,394		2,394		2,394	

Notes: All models include fixed effects for time and a constant. Results based on 20,000 Markov-chain Monte Carlo iterations, where the first 15,000 were discarded as burn-in. Estimates in bold are statistically significant under a 90% confidence interval.

The results for the importance of the market size, population density, as well as the industry mix show that relatively urbanized regions with a particularly distinctive sectoral service structure seemingly have the most promising prerequisites to attract FDI for both up- and downstream activities (see also Burger et al., 2012). Similarities in the determinants for attracting FDI between downstream and upstream activities can also be seen by the posterior estimates for the human capital endowments. Specifically, the table shows a positive and precisely estimated direct effect of the tertiary education attainment. Indirect effects for both human capital variables, however, are insignificant.

Spatial impact metrics for production-related FDI are depicted in columns (v) and (vi) of Table 3. Not surprisingly, for this category of FDI inflows both the market size variable and the corporate tax rates appear important. Inspection of the results, however, reveals that the determinants of attracting production-related FDI appears to be different as compared to up- and downstream activities. These findings are in line with work by Fallon and Cook (2014) and Crescenzi et al. (2013), who also find (albeit using different functional classifications) different locational determinants for production-related FDI as compared to business service oriented activities. Specifically, the direct impact for the employment share in the industry sector appears to be of particularly relevance, while the employment share in services is insignificant. This corroborates the findings of Defever (2006), who argue that locational manufacturing decisions of multinational enterprises is particularly affected by sectoral aspects. The table moreover reports negative direct impacts for the population density variable with positive and significant indirect effects. These resemble the stylized fact that regions with a relatively high specialization in manufacturing tend to be spatially clustered and are typically located not directly but in the closer proximity to urban agglomerations.

Results for the human capital variable for production FDI also differ from the outcomes of the two alternative classifications. Interestingly, the direct impact for regional tertiary education attainment is negative. However, its indirect impact measure is positive and much larger in magnitude. Similar to Crescenzi et al. (2013), we find no significant regional pull-factors for knowledge capital endowments for production FDI.

6 Concluding remarks

In this paper we propose a Bayesian approach for estimating a logit specification which exibits spatial autoregressive log-odds. Our proposed approach builds on recent advances in Bayesian econometric modelling by introducing a latent Pólya-Gamma variable (see Polson et al. 2013). Due to the introduction of the latent variable, the conditional posterior distribution of the slope parameters in our spatial autoregressive logit specification takes a Gaussian form, which renders Bayesian Markov-chain Monte Carlo estimation particularly efficient. Moreover, the resulting Gaussian conditional posterior distribution allows to extend the proposed baseline spatial autoregressive logit model to more flexible specifications in a straightforward way. In a simulation study the paper highlights the advantages of our proposed model specification as compared to existing spatial autoregressive Probit specifications (Smith and LeSage 2004) both in terms of parameter precision and computational time.

To illustrate our spatial logit model on a real economic data set, the paper moreover studies pan-European regional foreign direct investment flows by focussing on FDI inflows in different stages of the value chain. Specifically, in our empirical application we differentiate between upstream, downstream, and production activities. European regional FDI data are constructed using the *fDi Markets* database, which contains detailed information on regional FDI activities using media sources and company data. Our empirical results suggest that controlling for spatial autocorrelation when studying European regional FDI inflows is of fundamental importance. This particular holds true for downstream- and production-related investment activities. For all stages in the value chain, we find the market size, their spatial spillover effects, as well as the corporate tax rates to be of particular importance. Our results moreover corroborate the findings in the recent literature on regional FDI flows, that service-related activities tend to be attracted by large metropolitan areas while manufacturing activities are typically attracted by smaller regions in the closer proximity to urban agglomerations (see, for example, Defever 2006 and Duranton and Puga 2005). We additionally find marked differences in the locational investment decisions between manufacturing and service-oriented FDI inflows (Fallon and Cook 2014, Crescenzi et al. 2013). Specifically, our results are in line with work by Strauss-Kahn and Vives (2009) and Defever (2006), who argue that industry-related investment decisions of multinational enterprises are particularly affected by sectoral aspects, while service-related decisions are more determined by functional aspects. The results moreover highlight human and knowledge capital variables as important drivers to attract FDI.

References

Albert JH and Chib S (1993) Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association* 88(422), 669–679

Anselin L (2002) Under the hood: Issues in the specification and interpretation of spatial regression models. *Agricultural Economics* 27(2), 247–267

Ascani A, Crescenzi R and Iammarino S (2016) Economic institutions and the location strategies of European multinationals in their geographic neighborhood. *Economic Geography* 92(4), 401–429

Balasubramanyam VN, Salisu M and Sapsford D (1996) Foreign direct investment and growth in EP and IS countries. *The Economic Journal* 106(434), 92–105

Baltagi BH, Egger P and Pfaffermayr M (2007) Estimating models of complex FDI: Are there third-country effects? *Journal of Econometrics* 140(1), 260–281

Bellak C and Leibrecht M (2009) Do low corporate income tax rates attract FDI? – Evidence from Central- and East European countries. *Applied Economics* 41(21), 2691–2703

Blomstrom M and Kokko A (2003) The economics of foreign direct investment incentives. Technical report, National Bureau of Economic Research

Blonigen BA and Piger J (2014) Determinants of foreign direct investment. *Canadian Journal of Economics* 47(3), 775–812

Borensztein E, De Gregorio J and Lee JW (1998) How does foreign direct investment affect economic growth? *Journal of International Economics* 45(1), 115–135

Brasington D, Flores-Lagunes A and Guci L (2016) A spatial model of school district open enrollment choice. *Regional Science and Urban Economics* 56, 1–18

Burger MJ, van der Knaap B and Wall RS (2012) Revealed competition for greenfield investments between European regions. *Journal of Economic Geography* 13(4), 619–648

Calabrese R and Elkink JA (2014) Estimators of binary spatial autoregressive models: A Monte Carlo study. *Journal of Regional Science* 54(4), 664–687

Cameron AC and Trivedi KP (2005) *Microeconometrics*. Cambridge University Press, New York, 2nd edition

Cornwall GJ and Parent O (2017) Embracing heterogeneity: The spatial autoregressive mixture model. *Regional Science and Urban Economics* 64, 148–161

Crescenzi R, Pietrobelli C and Rabellotti R (2013) Innovation drivers, value chains and the geography of multinational corporations in Europe. *Journal of Economic Geography* 14(6), 1053–1086

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

Crespo Cuaresma J, Doppelhofer G, Huber F and Piribauer P (2018) Human capital accumulation and long-term income growth projections for European regions. *Journal of Regional Science* 58(1), 81–99

Defever F (2006) Functional fragmentation and the location of multinational firms in the enlarged Europe. *Regional Science and Urban Economics* 36(5), 658–677

Dimitropoulou D, McCann P and Burke SP (2013) The determinants of the location of foreign direct investment in UK regions. *Applied Economics* 45(27), 3853–3862

Duranton G and Puga D (2005) From sectoral to functional urban specialisation. *Journal of Urban Economics* 57(2), 343–370

Eicher TS, Helfman L and Lenkoski A (2012) Robust FDI determinants: Bayesian model averaging in the presence of selection bias. *Journal of Macroeconomics* 34(3), 637–651

Ekholm K, Forslid R and Markusen JR (2007) Export-platform foreign direct investment. *Journal of the European Economic Association* 5(4), 776–795

Fallon G and Cook M (2014) Explaining manufacturing and non-manufacturing inbound FDI location in five UK regions. *Tijdschrift voor Economische en Sociale Geografie* 105(3), 331–348

Fischer MM and LeSage JP (2015) A Bayesian space-time approach to identifying and interpreting regional convergence clubs in Europe. *Papers in Regional Science* 94(4), 677–702

Geweke J (1991) *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*, volume 196. Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA

Henderson JV and Ono Y (2008) Where do manufacturing firms locate their headquarters? *Journal of Urban Economics* 63(2), 431–450

Huber F, Fischer MM and Piribauer P (2017) The role of US based FDI flows for global output dynamics. *Macroeconomic Dynamics* , forthcoming

Klier T and McMillen DP (2008) Clustering of auto supplier plants in the United States: Generalized method of moments spatial logit for large samples. *Journal of Business and Economic Statistics* 26(4), 460–471

Koch M and Krisztin T (2011) Applications for asynchronous multi-agent teams in nonlinear applied spatial econometrics. *Journal of Internet Technology* 12(6), 1007–1014

Krisztin T (2017) The determinants of regional freight transport: A spatial, semiparametric approach. *Geographical Analysis* 49(3), 268–308

Krisztin T (2018) Semi-parametric spatial autoregressive models in freight generation modeling. *Transportation Research Part E: Logistics and Transportation Review* 114(September 2017),

121–143
Laurini MP (2017) A spatial error model with continuous random effects and an application to growth convergence. *Journal of Geographical Systems* 19(4), 371–398
LeSage JP (1997) Bayesian estimation of spatial autoregressive models. *International Regional Science Review* 20(1), 113–129
LeSage JP (2000) Bayesian estimation of limited dependent variable spatial autoregressive models. *Geographical Analysis* 32(1), 19–35
LeSage JP and Chih YY (2018) A Bayesian spatial panel model with heterogeneous coefficients. *Regional Science and Urban Economics* 72, 58–73
LeSage JP and Fischer MM (2008) Spatial growth regressions: Model specification, estimation and interpretation. *Spatial Economic Analysis* 3(3), 275–304
LeSage JP and Fischer MM (2012) Estimates of the impact of static and dynamic knowledge spillovers on regional factor productivity. *International Regional Science Review* 35(1), 103–127
LeSage JP, Kelley Pace R, Lam N, Campanella R and Liu X (2011) New Orleans business recovery in the aftermath of Hurricane Katrina. *Journal of the Royal Statistical Society. Series A: Statistics in Society* 174(4), 1007–1027
LeSage JP and Pace RK (2007) A matrix exponential spatial specification. *Journal of Econometrics* 140, 190–214
LeSage JP and Pace RK (2009) *Introduction to Spatial Econometrics*. CRC Press, Boca Raton London New York
LeSage JP and Pace RK (2018) Spatial econometric Monte Carlo studies: Raising the bar. *Empirical Economics* 55(1), 17–34
Markusen JR and Venables AJ (2000) The theory of endowment, intra-industry and multi-national trade. *Journal of International Economics* 52(2), 209–234
McFadden D (1974) Conditional logit analysis of qualitative choice behavior. In Zarembka P, ed., *Frontiers in Econometrics*. Academic Press, New York, 105–142
McMillen DP (1992) Probit with spatial autocorrelation. *Journal of Regional Science* 32(3), 335–348
Pfarrhofer M and Piribauer P (2019) Flexible shrinkage in high-dimensional Bayesian spatial autoregressive models. *Spatial Statistics* 29(1), 109–128
Pintar N, Sargant B and Fischer MM (2016) Austrian outbound foreign direct investment in Europe: A spatial econometric study. *Romanian Journal of Regional Science* 10(1), 1–22
Piribauer P (2016) Heterogeneity in spatial growth clusters. *Empirical Economics* 51(2), 659–680
Piribauer P and Crespo Cuaresma J (2016) Bayesian variable selection in spatial autoregressive models. *Spatial Economic Analysis* 11(4), 457–479
Polson NG, Scott JG and Windle J (2013) Bayesian inference for logistic models using Polya-Gamma latent variables. *Journal of the American Statistical Association* 108(504), 1339–1349
Raftery AE and Lewis S (1992) How many iterations in the Gibbs sampler? In Bernardo JM, Berger J, Dawid AP and Smith AFM, eds., *Bayesian Statistics 4*. Oxford University Press, Oxford, 763–773

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

Rodriguez-Clare A (1996) Multinationals, linkages, and economic development. *The American Economic Review* 86(4), 852–873

Smirnov OA (2010) Modeling spatial discrete choice. *Regional Science and Urban Economics* 40(5), 292–298

Smith TE and LeSage JP (2004) A Bayesian probit model with spatial dependencies. In Lesage JP and Pace RK, eds., *Spatial and Spatiotemporal Econometrics*. Emerald Group Publishing Limited, Bingley, 127–160

Strauss-Kahn V and Vives X (2009) Why and where do headquarters move? *Regional Science and Urban Economics* 39(2), 168–186

Sturgeon TJ (2008) Mapping integrative trade: Conceptualising and measuring global value chains. *International Journal of Technological Learning, Innovation and Development* 1(3), 237–257

Vega SH and Elhorst JP (2013) On spatial econometric models, spillover effects, and W. In *53rd ERSA Congress, Palermo, Italy*

Windle J, Polson NG and Scott JG (2014) Sampling Polya-Gamma random variates: Alternate and approximate techniques. *arXiv preprint arXiv:1405.0506*

Xu B (2000) Multinational enterprises, technology diffusion, and host country productivity growth. *Journal of Development Economics* 62(2), 477–493

Appendix

Table A1: Classification of *fDi Markets* business functions

Classification	Business activities	% of classification
Upstream	Business Services	64.0
	Design, Development and Testing	10.8
	Education and Training	2.5
	Headquarters	12.1
	Information and Communication Technology and Internet Infrastructure	4.3
	Research and Development	6.3
Downstream	Customer Contact Centre	4.2
	Logistics, Distribution and Transportation	26.9
	Maintenance and Servicing	3.4
	Sales, Marketing and Support	62.1
	Shared Services Centre	2.0
	Technical Support Centre	1.4
Production	Construction	21.0
	Electricity	5.3
	Extraction	0.3
	Manufacturing	72.1
	Recycling	1.3

Notes: The last column indicates the percent of industry activities per FDI classification. The values are based on the total observed FDI flows in the *fDi Markets* database targeting the selected NUTS-2 regions in the period 2003-2011.

Table A2: List of regions in the study.

Austria	**France [continued]**	**Hungary**	**Poland [continued]**	**UK**
Burgenland (AT)	Languedoc-Roussillon	Dél-Alföld	Lódzkie	Bedfordshire and Hertfordshire
Kärnten	Limousin	Dél-Dunántúl	Lubelskie	Berkshire, Buckinghamshire and
Niederösterreich	Lorraine	Észak-Alföld	Lubuskie	Oxfordshire
Oberösterreich	Midi-Pyrénées	Észak-Magyarország	Malopolskie	Cheshire
Salzburg	Nord - Pas-de-Calais	Közép-Dunántúl	Mazowieckie	Cornwall and Isles of Scilly
Steiermark	Pays de la Loire	Közép-Magyarország	Opolskie	Cumbria
Tirol	Picardie	Nyugat-Dunántúl	Podkarpackie	Derbyshire and Nottinghamshire
Vorarlberg	Poitou-Charentes	**Ireland**	Podlaskie	Devon
Wien	Provence-Alpes-Côte d'Azur	Border, Midland and Western	Pomorskie	Dorset and Somerset
Belgium	Rhône-Alpes	Southern and Eastern	Slaskie	East Anglia
Prov. Antwerpen	**Germany**	**Italy**	Swietokrzyskie	East Wales
Prov. Brabant Wallon	Arnsberg	Abruzzo	Warminsko-Mazurskie	East Yorkshire and
Prov. Hainaut	Berlin	Basilicata	Wielkopolskie	Northern Lincolnshire
Prov. Liège	Brandenburg	Calabria	Zachodniopomorskie	Eastern Scotland
Prov. Limburg (BE)	Braunschweig	Campania	**Portugal**	Essex
Prov. Luxembourg (BE)	Bremen	Emilia-Romagna	Alentejo	Gloucestershire, Wiltshire and
Prov. Namur	Chemnitz	Friuli-Venezia Giulia	Algarve	Bristol
Prov. Oost-Vlaanderen	Darmstadt	Lazio	Área Metropolitana de Lisboa	Greater Manchester
Prov. Vlaams-Brabant	Detmold	Liguria	Centro (PT)	Hampshire and Isle of Wight
Prov. West-Vlaanderen	Dresden	Lombardia	Norte	Herefordshire, Worcestershire
Région de Bruxelles-Capitale	Düsseldorf	Marche	**Romania**	and Warwickshire
Bulgaria	Freiburg	Molise	Bucuresti - Ilfov	Highlands and Islands
Severen tsentralen	Gießen	Piemonte	Centru	Inner London
Severoiztochen	Hamburg	Provincia Autonoma di Bolzano/	Nord-Est	Kent
Severozapaden	Hannover	Bozen	Nord-Vest	Lancashire
Yugoiztochen	Karlsruhe	Provincia Autonoma di Trento	Sud - Muntenia	Leicestershire, Rutland and
Yugozapaden	Kassel	Puglia	Sud-Est	Northamptonshire
Yuzhen tsentralen	Koblenz	Sardegna	Sud-Vest Oltenia	Lincolnshire
Czech Republic	Köln	Sicilia	Vest	Merseyside
Jihovýchod	Leipzig	Toscana	**Slovakia**	North Eastern Scotland
Jihozápad	Lüneburg	Umbria	Bratislavský kraj	North Yorkshire
Moravskoslezsko	Mecklenburg-Vorpommern	Valle d'Aosta/Vallée d'Aoste	Stredné Slovensko	Northern Ireland (UK)
Praha	Mittelfranken	Veneto	Východné Slovensko	Northumberland and Tyne and
Severovýchod	Münster	**Latvia**	Západné Slovensko	Wear
Severozápad	Niederbayern	Latvija	**Slovenia**	Outer London
Strední Cechy	Oberbayern	**Lithuania**	Vzhodna Slovenija	Shropshire and Staffordshire
Strední Morava	Oberfranken	Lietuva	Zahodna Slovenija	South Western Scotland
Denmark	Oberpfalz	**Luxemburg**	**Sweden**	South Yorkshire
Hovedstaden	Rheinhessen-Pfalz	Luxemburg	Mellersta Norrland	Surrey, East and West Sussex
Midtjylland	Saarland	**Netherlands**	Norra Mellansverige	Tees Valley and Durham
Nordjylland	Sachsen-Anhalt	Drenthe	Östra Mellansverige	West Midlands
Sjælland	Schleswig-Holstein	Flevoland	Övre Norrland	West Wales and The Valleys
Syddanmark	Schwaben	Friesland (NL)	Småland med öarna	West Yorkshire
Estonia	Stuttgart	Gelderland	Stockholm	
Eesti	Thüringen	Groningen	Sydsverige	
Finland	Trier	Limburg (NL)	Västsverige	
Åland	Tübingen	Noord-Brabant	**Spain**	
Etelä-Suomi	Unterfranken	Noord-Holland	Andalucía	
Helsinki-Uusimaa	Weser-Ems	Overijssel	Aragón	
Länsi-Suomi	**Greece**	Utrecht	Cantabria	
Pohjois-ja Itä-Suomi	Anatoliki Makedonia, Thraki	Zeeland	Castilla y León	
France	Attiki	Zuid-Holland	Castilla-la Mancha	
Alsace	Dytiki Ellada	**Norway**	Cataluña	
Aquitaine	Dytiki Makedonia	Agder og Rogaland	Comunidad de Madrid	
Auvergne	Ionia Nisia	Hedmark og Oppland	Comunidad Foral de Navarra	
Basse-Normandie	Ipeiros	Nord-Norge	Comunidad Valenciana	
Bourgogne	Kentriki Makedonia	Oslo og Akershus	Extremadura	
Bretagne	Kriti	Sør-Østlandet	Galicia	
Centre (FR)	Notio Aigaio	Trøndelag	Illes Balears	
Champagne-Ardenne	Peloponnisos	Vestlandet	La Rioja	
Corsica	Sterea Ellada	**Poland**	País Vasco	
Franche-Comté	Thessalia	Dolnoslaskie	Principado de Asturias	
Haute-Normandie	Voreio Aigaio	Kujawsko-Pomorskie	Región de Murcia	
Île de France				

Marginal effects

In our spatial Durbin logit model the interpretation of marginal effects of the k-th explanatory variable (with $k = 1, \ldots, K$) differs from those in linear models. This is due to the fact that (i) the logit model is non-linear in nature and marginal effects differ by the level of the k-th variable, and (ii) the presence of spatial autocorrelation gives rise to an $N \times N$ matrix of partial derivatives, which makes interpretation of marginal effects richer, but also more complicated (see also LeSage and Pace 2009).

The first issue, where the marginal effect of the probability of $p(y_i = 1)$ varies with the level of the explanatory variable z_{ik}, is usually addressed in the logit literature by providing marginal effects in reference to the mean value of the k-th explanatory variable, which is denoted as $\overline{z_k} = \sum_{i=1}^{N} z_{ik}/N$. The marginal effects can thus be interpreted as the change in probability of observing $y = 1$ associated with a change in the average sample observation of the k-th explanatory variable. Note, that this also implies that marginal effects depend on the distribution of the explanatory variable itself.

Partial derivatives of the model in Eq. (2.1), with respect to the k-th coefficient can be written as:

$$\begin{aligned}\boldsymbol{\mu}_k &= \boldsymbol{A}^{-1}\boldsymbol{I}_N\overline{z_k}\gamma_k + \boldsymbol{A}^{-1}\boldsymbol{W}\overline{z_{Wk}}\theta_k,\\ \frac{\partial p(y=1|\overline{z}_k)}{\partial \overline{z}_k'} &= \frac{\exp \boldsymbol{\mu}_k}{1+\exp\left(\boldsymbol{\mu}_k\right)} \odot \left(\boldsymbol{A}^{-1}\boldsymbol{I}_N\beta_k + \boldsymbol{A}^{-1}\boldsymbol{W}\theta_k\right)\\ &= \boldsymbol{\Lambda}_k,\end{aligned} \tag{A.1}$$

where β_k and θ_k denote the k-th element of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$, respectively. $\overline{z_{Wk}}$ denotes the average value of the k-th spatially lagged explanatory variable, and $\odot$ is the Hadamard product. Note that marginal effects of the k-th coefficient, denoted as $\boldsymbol{\Lambda}_k$, are an $N \times N$ matrix due to the presence of the $N \times N$ spatial multiplier $\boldsymbol{A}^{-1}$.

Since interpreting $N \times N$ marginal effects proves cumbersome, we define in accordance with LeSage and Pace (2009) summary impact effects. These can be readily calculated from $\boldsymbol{\Lambda}_k$:

$$direct_k = \frac{1}{N}\boldsymbol{\iota}_N' diag(\boldsymbol{\Lambda}_k) \tag{A.2}$$

$$total_k = \frac{1}{N}\boldsymbol{\iota}_N' \boldsymbol{\Lambda}_k \boldsymbol{\iota}_N \tag{A.3}$$

$$indirect_k = total_k - direct_k, \tag{A.4}$$

where $\boldsymbol{\iota}_N$ denotes an $N \times 1$ vector of ones. The average direct effects summarize the average effect of a marginal change in the k-th explanatory variable on the log-odds in the own region. Average indirect effects, on the other hand, summarize the average impact due to a marginal change in *all other* regions. A third measure is given by the average total effects, which summarizes the own regional change in log-odds due to marginal change of the k-th variable in *all* regions.