Spatial Filtering, Model Uncertainty and the Speed of Income Convergence in Europe*

Jesús Crespo Cuaresma†
University of Innsbruck

Martin Feldkircher‡
Oesterreichische Nationalbank

May 29, 2009

Abstract

In this paper we put forward a Bayesian Model Averaging method dealing with model uncertainty in the presence of potential spatial autocorrelation. The method uses spatial filtering in order to account for different types of spatial links. We contribute to existing methods that handle spatial dependence among observations by explicitly taking care of uncertainty stemming from the choice of a particular spatial structure. Our method is applied to estimate the conditional speed of income convergence across 255 NUTS-2 European regions for the period 1995-2005. We show that the choice of a spatial weight matrix - and in particular the choice of a class thereof - can have an important effect on the estimates of the parameters attached to the model covariates. We also show that estimates of the speed of income convergence across European regions depend strongly on the form of the spatial patterns which are assumed to underlie the dataset. When we take into account this dimension of model uncertainty, the posterior distribution of the speed of convergence parameter appears bimodal, with a large probability mass around no convergence (0% speed of convergence) and a rate of convergence of 1%, approximately half of the value which is usually reported in the literature.

Keywords: Model uncertainty, spatial filtering, determinants of economic growth, European regions.

JEL Classifications: C11, C15, C21, R11, O52.

*The opinions in this paper are those of the authors and do not necessarily coincide with those of the Oesterreichische Nationalbank. We would like to thank Ben Gardiner, Stefan Lang, Ron Martin, Michael Pfaffermayr, Tomáš Slačík, Pete Tyler, Stefan Zeugner and participants in the Land Economy Research Seminar at the University of Cambridge for their helpful comments.

†Department of Economics, University of Innsbruck. Universitätstrasse 15, 6020 Innsbruck, Austria. E-mail address: jesus.crespo-cuaresma@uibk.ac.at.

‡Oesterreichische Nationalbank, Otto-Wagner-Platz 3, 1090 Vienna, Austria. E-mail address: martin.feldkircher@oenb.at

1 Introduction

In this paper we propose a framework for analyzing spatially correlated data under model uncertainty concerning both the covariates of the specification and the spatial links existing in the data. Recently, a branch of literature which uses Bayesian tools for the analysis of spatially correlated data under model uncertainty (in terms of variable choice) has developed, mainly around the work of James P. LeSage and coauthors. LeSage and Parent (2007) introduce Bayesian Model Averaging (BMA) methods for spatial econometric models, and the methods are applied in LeSage and Fischer (2007) to study determinants of income in EU regions. These methods are also used by LeSage and Parent (2008) to evaluate the existence of knowledge spillovers from patent activity in the EU.

Most empirical studies in the spatial econometrics literature model spatial spillovers in the framework of spatial autoregressive (SAR) specifications (see Anselin (1988)) conditional on a given spatial weight matrix which parametrizes the geographical links among cross-sections. On the other hand, not much research is existing about the effects of misspecification in the spatial weights matrix on the estimates of the model parameters. Most of the existing studies stick to a single spatial weights matrix and build the econometric model conditioning on the choice of such a spatial structure. Given a spatial structure, some empirical studies perform robustness checks where the estimation is repeated for different spatial link matrices (see Crespo Cuaresma et al. (2009) for a recent example using European regional growth data). A noteworthy exception is the work by LeSage and Fischer (2007), which considers uncertainty in the spatial link matrix. LeSage and Fischer (2007) concentrate on a single class of spatial weight matrices (K nearest neighbor matrices) and consider uncertainty concerning the measurement of distance and the number of neighbors considered. The method put forward by LeSage and Fischer (2007) uses numerical integration techniques to obtain posterior model probabilities, thereby limiting the applicability to very large datasets, which would prove computationally too costly. In this piece of work we develop a simple BMA method to obtain parameter estimates after integrating out the uncertainty over the matrix of spatial weights by means of spatial filtering. Using spatial filtering based on the eigenvector approach (see Getis and Griffith (2002) and Tiefelsdorf and Griffith (2007)), the Markov chain used to obtain BMA estimates can rely on standard (non-spatial) estimation methods to reconstruct the posterior distribution over the model space. This implies that larger sets of covariates and/or spatial weight matrices can be easily incorporated to the analysis.

We apply the BMA method to check the robustness of economic growth determinants among European regions and to obtain estimates of the speed of income convergence in the presence of uncertainty about both the nature of the covariates entering the model and the matrix of spatial weights. Researchers have spent a great deal of effort in trying to assess and quantify the income convergence process across economic units (usually, countries, see Barro (1991) and Barro and Sala-i-Martin (1991)). A vast amount of the existing empirical literature focuses on estimating the income convergence speed using cross-sectional data, with a theoretical setting based on neoclassical economic growth models (see Mathunjwa and Temple (2007), for a thorough analytical account of convergence in the Solow model). Many authors have also approached the issue of income convergence using regional datasets (see Sala-i-Martin (1996)). The use of regional data, however, poses an extra problem to the study of income convergence and the measurement of the speed of convergence. There is widespread evidence (see e.g. Fischer and Stirböck (2006), Niebuhr (2001)) that spatial spillovers have a significant influence on economic growth and therefore observations from regional growth datasets cannot be regarded as independently generated, even after controlling for region-specific determinants. Spatial interactions, such as technological spillovers or factor mobility, both being important forces for the process of convergence, need therefore to be specified explicitly in order to obtain estimates of the speed of income convergence within a group of regional units. In the presence of positive spatial autocorrelation in economic growth data, estimates of the speed of income convergence across geographical units will tend to be biased upwards if the spatial structure of the data is left unmodeled.[1]

In our empirical application, we obtain the posterior distribution of the speed of income convergence across European regions in the presence of model uncertainty concerning the choice of regressors and spatial links. In particular, we average over models containing 16 possible spatial weight matrices corresponding to 4 different classes (Queen contiguity, nearest neighbor, exponential decay and distance band matrices). Our results indicate that the speed of income convergence across regional units is around 1%, approximately half of the value which tends to be obtained with models conditioning on a single spatial weight matrix.

This paper is organized as follows. Section 2 considers the issue of uncertainty about the spatial correlation structure and embeds the problem in a general Bayesian Model Averaging setting where uncertainty about the variables entering the specification is also assumed. Section 3 applies the methodology to a dataset on European regions in order to obtain estimates of the speed of income convergence in Europe. Section 4 concludes.

2 Spatial autocorrelation, spatial filtering and model uncertainty

2.1 The econometric setting

Consider a cross-sectional growth regression from which we aim at extracting the speed of (conditional) income convergence across N geographical units. We explicitly model the potential existence of spatial autocorrelation by using a model of the class of spatial regression models (Anselin (1988)), namely a spatial autoregressive (SAR) model,

$$y = \alpha\iota_N + \rho\mathbf{W}y + \mathbf{X}_k\vec{\chi}_k + \sigma\varepsilon \tag{1}$$

where y is an N-dimensional column vector whose elements correspond to the annualized income growth of each geographical unit, α is the intercept term, ι_N is an N-dimensional column vector of ones, $\mathbf{X}_k = (\mathbf{x}_1 \dots \mathbf{x}_k)$ is a matrix whose columns are stacked data for k explanatory variables and $\vec{\chi}_k = (\chi_1 \dots \chi_k)'$ is the k-dimensional parameter vector corresponding to the variables in $\mathbf{X}_k$. We specify the spatial autocorrelation structure using the matrix $\mathbf{W}$, with its corresponding coefficient ρ reflecting the degree of spatial autocorrela-

[1] The patterns of regional growth and convergence in Europe have also been investigated by Boldrin and Canova (2001).

tion. One of the potential variables in $\mathbf{X}_k$ is the initial income level at the beginning of the period in which the growth rate of income is calculated. Equation (1) constitutes a parametric spatial model where the spatial parameter ρ is often interpreted as a spillover parameter, with positive values indicating the existence of spillovers from neighboring observations. Let us denote the parameter associated to initial income per capita by β. Evidence of conditional convergence is found whenever β is negative, thus implying that, after controlling for other factors, economies with low initial income levels grow on average faster than others having relatively higher initial income. The speed of convergence can be computed using a log-linearization around the steady-state of the Solow model as $\lambda = -(1/\tau)[1 - \exp(-\beta\tau)]$, where τ is the length of the period considered in the growth variable (see for instance Barro and Sala-i-Martin (2003)).

Since growth theory is ambiguous about the set $\mathbf{X}_k$ of explanatory variables to include, we are confronted with a classical situation of model uncertainty concerning the covariates which should enter the model. If the estimate of the coefficient of interest (in our case β) depends on the covariates entering the model, we will eventually overestimate the degree of precision of our estimate if we do not account for this particular source of uncertainty.

In our setting, an extra degree of uncertainty arises if we do not know the actual nature of the spatial interactions which we model through the spatial autoregressive term in (1), that is, if we conduct inference conditional on $\mathbf{W}$. However, besides reflecting the degree of spatial interaction across the data, Anselin (1988) notes that ρ might pick up a range of misspecifications of the general model. Spatial autocorrelation will be observable whenever the phenomenon under study is a spatial process or omitted variables cause spatial variation in the residuals (Tiefelsdorf and Griffith (2007)). Note that both arguments typically apply to economic cross-section data, where economic units interact with each other and omitted variables decrease the level of confidence in econometric analysis. Since inference from the SAR model is conditional on the weight matrix $\mathbf{W}$, which has to be exogenously specified, and in most applications there is little theoretical guidance on which structure to put on the weight matrix, explicitly accounting for this source of model uncertainty is a natural generalization to uncertainty in the nature of $\mathbf{X}_k$ in the framework of BMA.

2.2 Spatial filtering

The spatial filtering literature seeks to remove residual spatial autocorrelation patterns prior to estimation and is in principle not interested in directly estimating ρ in (1). Getis and Griffith (2002) propose two (nonparametric) approaches of filtering the data before applying regression analysis. The method utilizes a local spatial statistic (the G_i statistic, see Anselin (1988)) to decompose the data into a purely spatial and a non-spatial part. Limitations to this approach are that (a) it is restricted to non-negative data and (b) each variable entering the regression has to be filtered separately. The approach put forward by Getis and Griffith (2002) and Tiefelsdorf and Griffith (2007), on the other hand, is based on an eigenvector decomposition of a transformed $\mathbf{W}$ matrix, where the transformation depends on the underlying spatial model.

Assume that the data follows a SAR model as in equation (1) and thus can be written as

$$
\begin{aligned}
y &= (I-\rho\mathbf{W})^{-1}(\alpha\iota_N + \mathbf{X}_k\vec{\chi}_k + \sigma\varepsilon) = \\
&= \alpha\iota_N + \mathbf{X}_k\vec{\chi}_k + \sigma\varepsilon + \sum_{m=1}^{\infty}\rho^m\mathbf{W}^m(\alpha\iota_N + \mathbf{X}_k\vec{\chi}_k + \sigma\varepsilon).
\end{aligned} \tag{2}
$$

Spatial filtering methods aim at finding a good approximation for the last term in (2) which allows to remove the residual spatial autocorrelation induced by either a pure spatial autoregressive process or omitted variables that tie the residuals spatially together. The spatial link matrix is first transformed to satisfy symmetry and then multiplied by the demeaning projector $M_1 = I - \iota_N(\iota_N'\iota_N)^{-1}\iota_N'$ in order to extract eigenvectors with underlying SAR structure. Each extracted eigenvector $\vec{e_i}$ of $[M_1\frac{1}{2}(\mathbf{W}+\mathbf{W}')M_1]$ reflects a distinctive spatial pattern and is associated with a specific spatial autocorrelation level. Thus instead of equation (1) we may estimate

$$
y = \alpha\iota_N + \sum_{i=1}^{E}\gamma_i\vec{e_i} + \mathbf{X}_k\vec{\chi}_k + \sigma\varepsilon, \tag{3}
$$

where each eigenvector $\vec{e_i}$ spans one of the spatial dimensions. By introducing the eigenvectors into the regression, we explicitely take care of (remaining) spatial patterns in the residuals. Furthermore spatial commonalities among the covariates in $\mathbf{X}_k$ are conditioned out. This reduces the degree of multicollinearity and further separates spatial effects from the "intrinsic" impact the employed regressors exert on the dependent variable.

The fact that the transformation of the spatial weight matrix does not involve the design matrix $\mathbf{X}_k$ is an important advantage in the framework of model uncertainty, since the calculation of the eigenvectors has to be carried out only once.[2] In our application, we identify the set of eigenvectors needed (E) with the algorithm proposed by Tiefelsdorf and Griffith (2007). This algorithm identifies the minimal subset of eigenvectors until the residual spatial correlation as measured by Moran's I statistic (see Anselin (1988)) drops below a certain threshold value.

2.3 Bayesian Model Averaging with uncertain spatial effects

From a Bayesian perspective, the problem of obtaining estimates of the parameter associated to a covariate under uncertainty in both the nature of $\mathbf{W}$ and $\mathbf{X}_k$ can be handled in a straightforward manner using spatial filtering techniques. Let us assume that we are interested in the parameter corresponding to the initial income level, β. Denote the set of potential models by $\mathcal{M} = \{M_1^1, M_2^1, \ldots, M_{2^K}^1, \ldots M_1^2, \ldots, M_{2^K}^2, \ldots, M_1^Z, \ldots, M_{2^K}^Z\}$, where K stands for the number of potential explanatory variables and Z the number of potential

[2] Notice that this would not be the case for models involving spatially lagged errors (see Tiefelsdorf and Griffith (2007)). In this case, the projection matrix used is a function of $\mathbf{X}_k$. Although our method is not affected by the use of this projector, the implementation for large datasets can be computationally very costly.

spatial weighting matrices $\mathbf{W}_z$, $z = 1, \ldots, Z$ each with associated set of eigenvectors E_z. The cardinality of $\mathcal{M}$ is therefore $2^K \times Z$. A particular model, say M_k^z, is characterized by its parameter vector $\theta_k^z = (\alpha,\ \chi_k, \gamma_z)$ corresponding to the intercept term included in all models, the coefficients on the regressors entering the model and the coefficients on the set of eigenvectors E_z related to $\mathbf{W}_z$. In the BMA framework[3], the posterior distribution of β takes now the form of

$$p(\beta|y) = \sum_{j=1}^{2^K} \sum_{z=1}^{Z} p(\beta|M_j^z, y) p(M_j^z|y) \tag{4}$$

with y denoting the data and β the coefficient of interest. Inference on β is based on single inferences under models $j = 1, \ldots, 2^K \times Z$ weighted by their respective posterior model probabilities, $p(M_j^z|y)$, which in turn depend on the corresponding matrix of spatial weights. We can construct (4) making use of the fact that

$$p(M_j^z|y) = \frac{p(y|M_j^z)\overline{p}(M_j^z)}{\sum_{j=1}^{2^K} \sum_{z=1}^{Z} p(y|M_j^z)\overline{p}(M_j^z)}. \tag{5}$$

where $\overline{p}(M_j^z)$ denotes the prior distribution assigned to model M_j^z and $p(y|M_j^z)$ is the integrated likelihood. For the sake of illustration, consider the particular case of two competing models. In this case, the posterior odds are simply given by the product of the Bayes Factor with the prior odds. In order to obtain (5) and thus (4), we need to specify priors for the regression coefficients, for the variance σ and over the model space $\mathcal{M}$. As is common practice in the applied literature, we use Zellner's g-prior structure on the regression slopes, which merely requires the choice of one hyper parameter g, thus specifying $(\vec{\chi}, \alpha)|\sigma^2 \sim N\left(0, \sigma^2[gX'X]^{-1}\right)$. Following Ley and Steel (2009), we move away from assuming an uninformative prior over the model space, as many other BMA studies tend to do. Instead, we assume that the prior on the model space ($\overline{p}(M)$) is a binomial-beta prior, which we elicit by anchoring the prior on an expected model size. The technical appendix presents a discussion on the specific prior choices for g and on the model space.

In many applications, such as the one we present here, the cardinality of the model space renders the evaluation of (4) intractable. Several methods have been proposed to overcome this problem and Markov Chain Monte Carlo Model Composition (MC^3) algorithms have become a useful tool to evaluate subsets of the model space which account for a large posterior model probability mass (see Fernández et al. (2001b) for an application to economic growth determinants). Throughout the paper we rely on a random walk MC^3 search algorithm to evaluate the model space. We slightly modify the usual MC^3 method in order to account for uncertainty over a set of spatial weight matrices. Our algorithm proceeds in the following modified way:

1. Starting with a model as defined by a group of regressors and the set of eigenvectors E_z associated to a spatial weight matrix $\mathbf{W}_z$, in each iteration step a candidate regressor is drawn from the set of potential covariates. We add the candidate regressor to the current model M_j^z if that model did not already include it. On the other hand, the

[3] For a textbook introduction to BMA, see for instance Koop (2003), chapter 11.

candidate regressor is dropped from the model if it is already contained in M_j^z. Note that both models are conditional on the *same* set of eigenvectors E_z. The candidate model is thus always drawn from a neighborhood of the current one, defined as the subset of the model space formed by models which differ only by a single regressor. The candidate model M_c^z is then subject to the following acceptance probability:

$$\tilde{p}_{cj} = \min\left[1, \frac{\overline{p}(M_c^z)p(y|M_c^z)}{\overline{p}(M_j^z)p(y|M_j^z)}\right]. \tag{6}$$

Notice that the potential punishment for model size is embedded in both the model prior and the Bayes Factor.

2. In the second step a candidate weighting matrix $\mathbf{W}_c$ (and hence its associated set of eigenvectors E_c) is drawn uniformly from the set of remaining matrices $\mathbf{W}_{(-z)} := \{\mathbf{W}_i\}_{i=1}^Z, i \neq z$. Since we are interested in handling uncertainty across different specifications of $\mathbf{W}$ the eigenvectors belonging to E_c are always forced to enter the regression jointly. The accepted model from step 1), denote it by M_j^z, is then compared with the model containing the same regressors but a *different* set of eigenvectors E_c. The acceptance probability is given by:

$$\hat{p}_{cz} = \min\left[1, \frac{p(y|M_j^c)}{p(y|M_j^z)}\right]. \tag{7}$$

Since both models consist of the same number of regressors subject to sampling the prior odds on model size cancel. The reward for parsimonity with respect to the spatial weight matrix is solely governed by the Bayes Factor. It is straightforward to introduce a further informative prior on the space of weight matrices instead of the uniform prior employed in our analysis.

We repeat steps 1) and 2) a large number of times and compute the corresponding BMA statistics based on the set of models visited, instead of the full model space. We are especially interested in the posterior distribution of the parameters of the covariates in (1), their posterior variance and the posterior inclusion probability (PIP) of the covariates. The latter one is defined as the sum of posterior model probabilities of the specifications including a particular covariate, $\text{PIP}_l = \sum_{j=1}^{2^K} \sum_{z=1}^{Z} p(M_j^z$ such that $\chi_l \neq 0|y)$. The performance of this method is assessed in the following subsection by means of a simulation study.

2.4 A simulation study

In order to test the ability of our sampler to both identify model covariates and unveil spatial structures present in the data we conduct a small simulation study. Our focus is the posterior distribution over the spatial link matrices and we choose a rather simple setting for the data generating process. We draw 10 potential explanatory variables $(\mathbf{x}_1, \ldots, \mathbf{x}_{10})$ using N =255 draws from a standard normal distribution for each covariate, so as to match the sample size of our empirical application. The spatial autocorrelation level is fixed at $\rho = 0.6$, a typical level of spatial dependence present in economic data sets. Data on the dependent variable are generated according to

$$y = \rho \mathbf{W}_z y + 1.5\mathbf{x}_1 + 2\mathbf{x}_4 - 0.5\mathbf{x}_{10} + 0.5\varepsilon, \quad (8)$$

where ε is a standard normal variable. We restrict our space of potential weighting matrices to three different classes (for a textbook discussion on weighting schemes see Anselin (1988)): Queen contiguity matrices, K-nearest neighbor matrices and distance band matrices.

Queen contiguity matrices assign equal positive weights to observations sharing a common border (including cases where the common border is just a vertex). We will consider a first-order contiguity definition for neighbors in this class, and denote the spatial weighting matrix as $\mathbf{W}_1^Q$.[4] The K-nearest neighbor coding scheme evaluates airline-distances between all observations and assigns a positive weight to the K nearest neighbors. From this class of weighting matrices, we consider a weighting scheme based on four neighbors ($\mathbf{W}_4^{K-NN}$) for the simulation. Finally, distance band matrices regard geographical units that lie within a distance band of d kilometers as neighbors. Our space of spatial weight matrices in the simulation includes a distance band matrix based on a band of 400 kilometers ($\mathbf{W}_{400}^b$). All these alternative space weighting matrices belong to the class of binary weight matrices and solely differ with respect to the definition of the set of neighbors.[5]

We impose the spatial weights corresponding to each matrix computed on the dataset of 255 NUTS-2 regions, and thus replicate spatial patterns in our simulated data which reproduce the geographical structure of the European regional dataset analyzed in section 3. For the simulation we consider five cases, each corresponding to a $\mathbf{W}_z$ matrix in (8):

- case $z = 1$: $\mathbf{W}_z$ is a first order Queen contiguity matrix ($\mathbf{W}_1^Q$),
- case $z = 2$: $\mathbf{W}_z$ is a four nearest neighbor weight matrix ($\mathbf{W}_4^{K-NN}$),
- case $z = 3$: $\mathbf{W}_z$ is a 400 km distance band weight matrix ($\mathbf{W}_{400}^b$),
- case $z = 4$: $\mathbf{W}_z$ is given by $0.3\mathbf{W}_1^Q + 0.6\mathbf{W}_4^{K-NN} + 0.1\mathbf{W}_{400}^b$,
- case $z = 5$: $\mathbf{W}_z$ is given by $0.5\mathbf{W}_1^Q + 0\mathbf{W}_4^{K-NN} + 0.5\mathbf{W}_{400}^b$.

The set of potential covariates in the simulation has cardinality 10, and the set of potential spatial weighting schemes has a cardinality of 3 ($\mathbf{W}_1^Q$, $\mathbf{W}_4^{K-NN}$ and $\mathbf{W}_{400}^b$), thus leading to a model space composed by 3072 models. We repeat the exercise for 50 simulated datasets for each setting z, using an MC^3 search method over the model space with 5000 replications each time. The averaged results by case are presented in Table 1. Since the inclusion probabilities of the variables included in the model were all very close to one, and the estimated parameters also very close to the true values, we do not report them and concentrate exclusively on the inclusion probabilities (percentage of models visited by the MC^3 algorithm by $\mathbf{W}$ matrix) for each one of the spatial weighting matrices.[6]

[4] Note that this weighting scheme might create "spatial islands" (i.e. observations without any neighbors).

[5] All matrices used in the analysis are row-standardized.

[6] We use the BRIC prior (Fernández et al., 2001a) for g ($g = 1/N$) and the beta-binomial prior over the model space with prior expected model size $K/2$. The results on the inclusion probabilities of the explanatory variables and the corresponding posterior distributions over parameters are available from the authors upon request. For all five settings, the sampler identified the true variables and the associated coefficients with high precision.

	$\mathbf{W}_1^Q$	$\mathbf{W}_4^{K-NN}$	$\mathbf{W}_{400}^b$
Case $j = 1$			
Percentage visited	99.66	0.34	0.00
Adj. R^2	0.47	0.37	0.23
# eigenvectors	25.50	23.46	9.02
Case $j = 2$			
Percentage visited	0.00	100.00	0.00
Adj. R^2	0.29	0.42	0.18
# eigenvectors	16.98	33.82	7.18
Case $j = 3$			
Percentage visited	0.00	0.00	100.00
Adj. R^2	0.09	0.12	0.19
# eigenvectors	2.56	6.56	10.44
Case $j = 4$			
Percentage visited	3.87	96.06	0.07
Adj. R^2	0.31	0.36	0.19
# eigenvectors	16.94	27.24	8.44
Case $j = 5$			
Percentage visited	32.96	16.89	50.15
Adj. R^2	0.25	0.22	0.21
# eigenvectors	11.84	13.50	9.66

The results in each case refer to averages over 50 simulated datasets. "Percentage visited" is the percentage of times a model with a given spatial weight matrix was visited in the MC^3 algorithm, and is thus interpreted as the corresponding posterior inclusion probability. "Adj. R^2" is the average adjusted R^2 of regressions based exclusively on the eigenvectors corresponding to a particular spatial weighting matrix, and "# eigenvectors" is the average number of eigenvectors extracted using the method by Tiefelsdorf and Griffith (2007).

Table 1: Simulation results

The results indicate that the method can identify the underlying spatial structures with extremely high precision for the cases where the true spatial weighting matrix is a member of a single class. Not surprisingly, the results for cases 4 and 5, where the spatial weighting matrix is a weighted average of matrices from different classes, are less spectacular, but still very satisfactory.

3 Income convergence and spatial interactions across European Regions

In this section we assess the robustness of growth determinants and estimate the speed of income convergence of European regions in presence of both model uncertainty in terms of model covariates and the form of spatial interactions. Our dataset contains information on 50 potential covariates for 255 NUTS-2 European regions. The dependent variable refers to the average annual growth rate of real income per capita over the period 1995-2005, deflated using national price data. Information about coverage and definitions of the variables and abbreviations is presented in the Data Appendix. We consider linear models such as (1) in the spatial filtering representation given by (3).

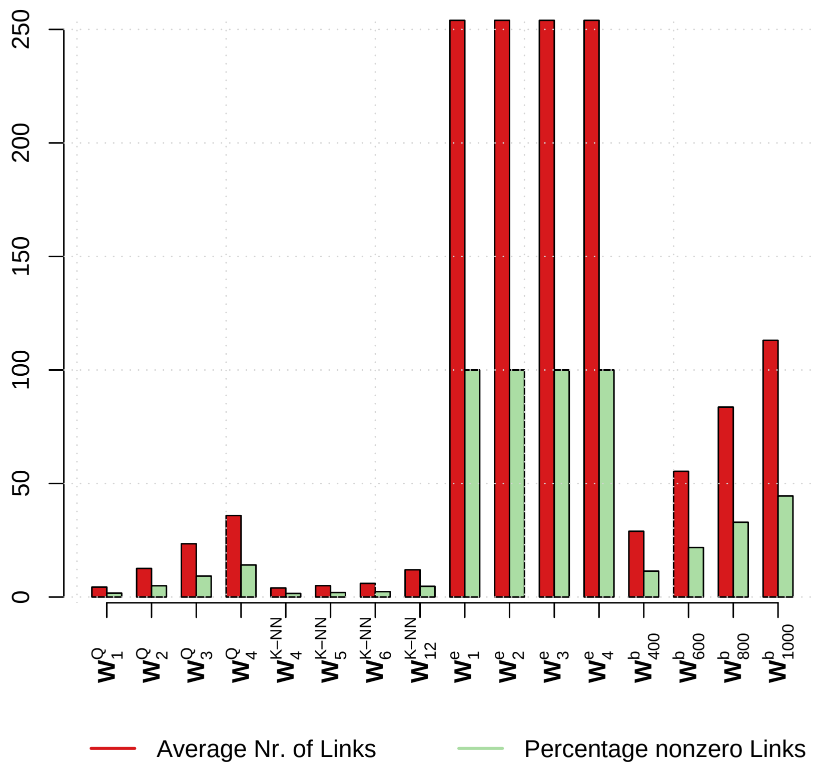

Figure 1: Summary statistics for different weight matrices: total links and non-zero links

We allow for four types of spatial weighting matrices for each one of the following classes: Queen matrices, K-nearest neighborhood, distance band and exponential decay. For the class of Queen matrices we consider $\mathbf{W}_z^Q$, $z = 1, 2, 3, 4$, ranging from a first order neighborhood matrix up to a fourth order neighborhood. The class of K-nearest neighborhood matrices is represented by four variants, $\mathbf{W}_z^{K-NN}$, $z = 4, 5, 6, 12$, each one based on z neighbors. The space of spatial weight matrices in our empirical study includes further four distance band matrices $\mathbf{W}_z^b$, $z = 400, 600, 800, 1000$, where each one of them identifies neighbors based on bands of z kilometers. Finally, the set of exponential decay matrices has a representative element given by $[\mathbf{W}_\phi^e]_{ij} = [d_{ij}]^{-\phi}$, where d_{ij} is the (airline) distance between observations i and j and the parameter ϕ governs the decay of the weighting scheme. We consider four possible exponential decay matrices, given by $\phi = 1, 2, 3, 4$. A unit ϕ parameter implies that observations are weighted according to inverse distances, while higher values of ϕ lead to a sharper decay of weights as distance increases. Figure 1 summarizes the number of links and the percentage of strictly positive links for each one of the matrices in the set of potential spatial weight matrices. As Figure 1 exemplifies, there are strong differences in the spatial structure underlying each one of the matrices in the sense of the amount of neighboring units assumed to affect economic performance in a given region. The correlation of spatially lagged income ($\mathbf{W}^z y$) for the 16 matrices ranges from 0.50 to 0.96.

We apply the spatial filter proposed by Tiefelsdorf and Griffith (2007) to each of our **W** matrices and extract the relevant subsets of eigenvectors based on a cut-off of 0.1 in Moran's I statistic. Table 4 shows the number of selected eigenvectors for each spatial link matrix and the adjusted R^2 resulting from the regression of the dependent variable solely on a constant and the full set of eigenvectors. The results reveal that a large part of variation in the data can be explained exclusively by spatial patterns as proxied by the eigenvectors. This complicates the estimation of the "pure" speed of income convergence, free from the spatial effects created by economic growth poles.

	$\mathbf{W}_1^Q$	$\mathbf{W}_2^Q$	$\mathbf{W}_3^Q$	$\mathbf{W}_4^Q$	$\mathbf{W}_4^{K-NN}$	$\mathbf{W}_5^{K-NN}$	$\mathbf{W}_6^{K-NN}$	$\mathbf{W}_{12}^{K-NN}$
Number of eigenvectors	15	15	10	8	18	14	18	19
Adj. R^2	0.5888	0.5342	0.5250	0.5126	0.5515	0.5398	0.5248	0.5131
	$\mathbf{W}_1^e$	$\mathbf{W}_2^e$	$\mathbf{W}_3^e$	$\mathbf{W}_4^e$	$\mathbf{W}_{400}^b$	$\mathbf{W}_{600}^b$	$\mathbf{W}_{800}^b$	$\mathbf{W}_{1000}^b$
Number of eigenvectors	19	17	13	17	13	12	12	13
Adj. R^2	0.5630	0.5663	0.5887	0.5778	0.5211	0.4877	0.4759	0.4017

Table 2: Number of eigenvectors and adjusted R^2: Spatially filtered income per capita growth

In a first stage, we obtain BMA estimates for the effect of the covariates on economic growth conditioning individually on each one of the different classes of spatial weighting matrices[7]. Since sensitivity analyses reported in the spatial econometric literature are often restricted to one particular class of spatial weight matrices, the differences in inference resulting across classes of weighting matrices is of particular interest. We therefore assess the dependence of the relative importance of different covariates with respect to spatial weighting matrices. For that purpose, we obtained posterior inclusion probabilities for each variable in our dataset in six different BMA settings. We first obtain BMA statistics based on a linear model without spatial interactions (equation (1) with the constraint $\rho = 0$ imposed or, alternatively, equation (3) with $\gamma_i = 0$ for $i = 1, \ldots, E$). Secondly we obtain BMA statistics based on spatial weight uncertainty but constraining the spatial links to belong to each one of the individual classes of spatial weight matrices (Queen, exponential decay, K-neighborhood and distance band). Finally we calculate BMA statistics where the space of spatial weight matrices is composed by all 4 classes and hence 16 **W** matrices. Table 3 presents the results in terms of posterior inclusion probability (PIP), mean (PM) and standard deviation (PSD) of the posterior distribution of the parameters. Figure 2 plots the posterior inclusion probability of the variables which achieve the highest values in this statistic for the BMA exercises conditioning on different classes of spatial weight matrices. Table 4 presents the posterior inclusion probabilities for each one of the spatial weight matrices in the analysis.

The results in Figure 2 and Table 3 present interesting differences across estimates depending on the class of spatial weight matrices which is conditioned upon. The choice of a particular class of spatial weight matrices as a parametrization of the links across regions may have an important effect on the resulting posterior inclusion probabilities, as can be seen in Figure

[7] The benchmark BRIC prior implies setting $g = 1/K^2$. Following Sala-i-Martin et al. (2004) we expect a typical growth model to be composed of seven regressors a priori and therefore set the prior expected model size equal to seven.

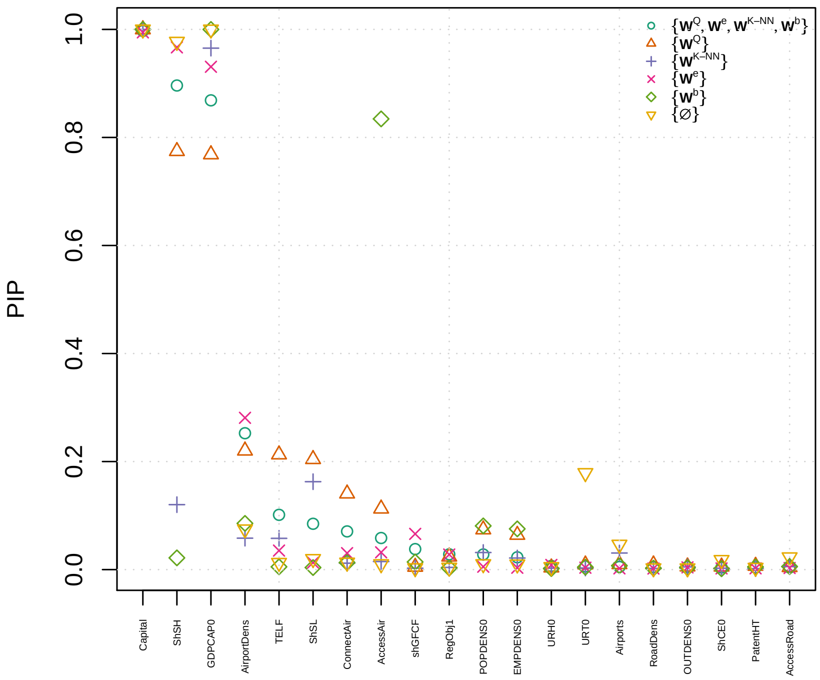

Figure 2: Posterior inclusion probabilities of covariates based on different classes of spatial weight matrices: all classes ($\{\mathbf{W}^Q, \mathbf{W}^e, \mathbf{W}^{K-NN}, \mathbf{W}^b\}$), Queen spatial matrices ($\{\mathbf{W}^Q\}$), K-nearest neighbors spatial matrices ($\{\mathbf{W}^{K-NN}\}$), exponential decay distance matrices ($\{\mathbf{W}^e\}$), distance band matrices ($\{\mathbf{W}^b\}$) and no spatial structure ($\{\emptyset\}$).

2 for the case of the human capital variable ShSH (share of working age population with high education). The results of BMA using the class of K-nearest neighbor matrices and BMA using distance band spatial weight matrices imply that the importance of ShSH as an explanatory factor of differences in income growth is small to negligible. On the other hand, the results based on BMA using one of the remaining spatial weight matrices (Queen and exponential decay), as well as the results based on models without spatial autoregression, depict ShSH as one of the most important variables for explaining income growth in European regions. The results of our preferred specification, where uncertainty is generalized to take place both within and across classes of spatial weight matrices imply that the human capital variable is indeed a robust determinant of economic growth in European regions. Similarly, the results for AccessAir (potential air accessibility) differ extremely if the spatial link is parametrized using a spatial weight matrix with exponential decay as compared to any of the other classes.

The posterior probabilities of models averaged across spatial weighting matrices are presented in Table 4. There are three individual weighting matrices which receive practically all of the evidence in terms of posterior probability: $\mathbf{W}^Q_4$, $\mathbf{W}^e_3$ and, to a minor extent, $\mathbf{W}^Q_3$. These

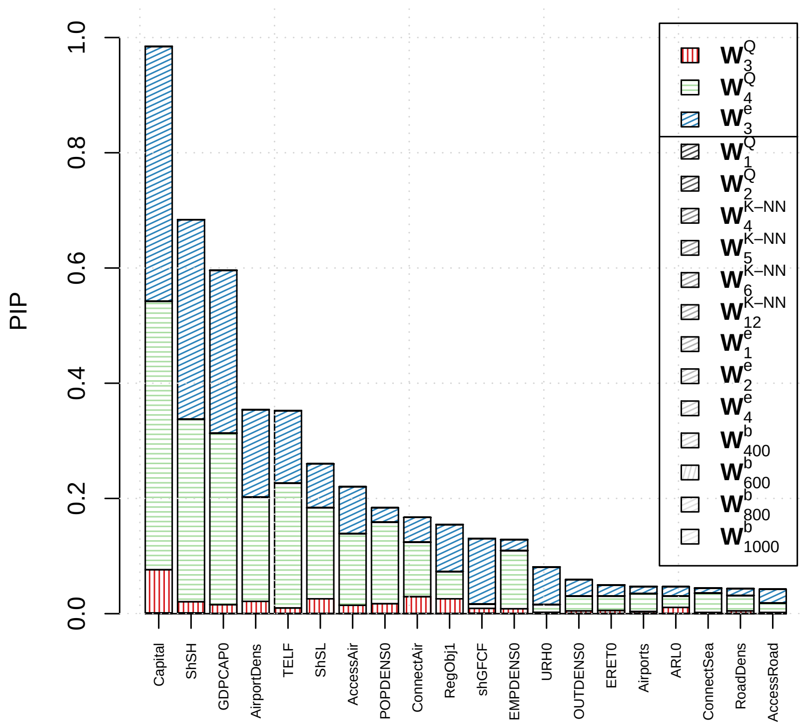

Figure 3: Joint distribution of regressors and weighting matrices based on the model involving all 16 matrices. The distribution is shown for the 20 most important variables according to the associated posterior inclusion probabilities (PIP).

matrices present a relatively different number of links (see Figure 1). The effects of spatial weight uncertainty on the relative importance of explanatory variables as determinants of regional growth can be grasped by examining the joint posterior inclusion distribution of covariates and spatial weight matrices, which is depicted in Figure 3.[8] Figure 3 shows that the importance of regressors in terms of posterior inclusion probability tends to remain similar across spatial structures. An interesting exception is the physical capital investment variable (shGFCF), whose relevance as a robust determinant of growth is exclusively concentrated in models including the exponential decay weighting matrix with $\phi = 3$, $\mathbf{W}_3^e$. The result sheds a particularly interesting light on the modeling choice of spatial links for cross-sectional regional growth regressions, since most empirical applications blindly condition on one of the elements of the $\mathbf{W}$ space, and the choice in the case of economic growth applications tends to be a spatial weighting matrix with exponentially decaying weights with distance.

The choice of a spatial weight matrix (or a group of them) is particularly important when it comes to obtaining an estimate for the speed of income convergence, embodied in the esti-

[8]Figure 3 is constructed with estimates based on the 3,000 models with highest inclusion probability, while the results reported in Table 4 are based on MC^3 frequencies. This implies that some small quantitative differences exist between the two.

mate of the parameter attached to the initial level of income per capita in the regional growth regression. In Figure 4 we present the posterior density of the speed of income convergence parameter computed over the whole model space (unconditional distribution) and over the set of models which include initial income as a regressor (conditional distribution), averaged over different conditioning sets of spatial weighting matrices. In particular, in order to exemplify the differences depending on the conditioning set in terms of $\mathbf{W}$ matrices, Figure 4 shows the results for the BMA exercise in three different settings for the spatial links: (a) models without spatial weighting matrix, (b) models with spatial weighting matrices of the class of distance bands matrices and (c) models with spatial weighting based on all classes put forward above.

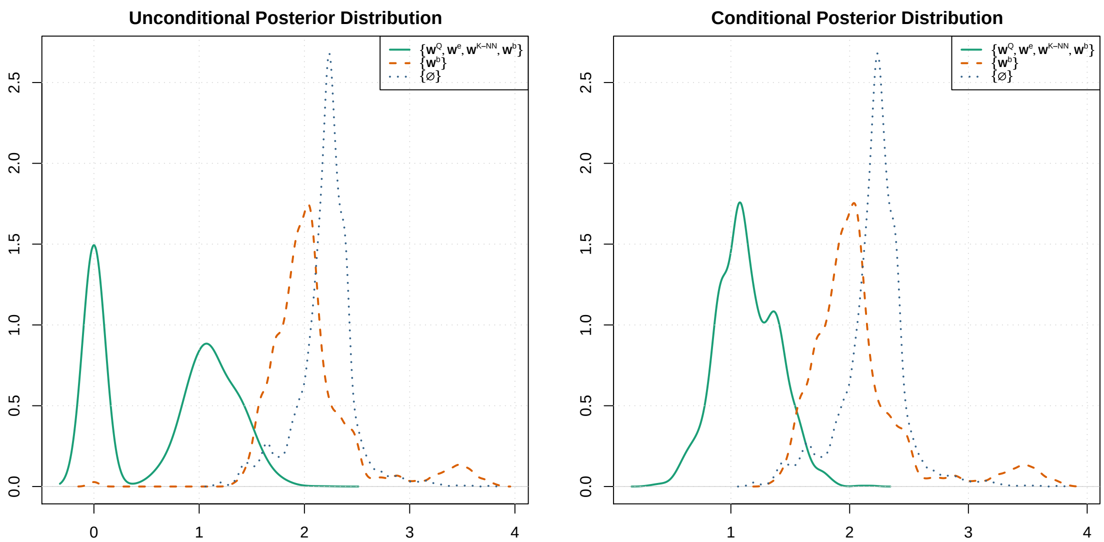

Figure 4: Unconditional and conditional posterior distribution over the speed of income convergence parameter based on 3,000 models with highest inclusion probabilities a) without spatial effects ($\{\emptyset\}$), b) with distance band matrices ($\{\mathbf{W}^b\}$) and c) with the full set of spatial weighting matrices ($\{\mathbf{W}^Q, \mathbf{W}^e, \mathbf{W}^{K-NN}, \mathbf{W}^b\}$)

	Spatial weight matrices: $\{\mathbf{W}^Q, \mathbf{W}^e, \mathbf{W}^{K-NN}, \mathbf{W}^b\}$			Spatial weight matrices: $\{\mathbf{W}^Q\}$			Spatial weight matrices: $\{\mathbf{W}^e\}$			Spatial weight matrices: $\{\mathbf{W}^{K-NN}\}$			Spatial weight matrices: $\{\mathbf{W}^b\}$			Spatial weight matrices: $\{\emptyset\}$		
Var. name	PIP	PM	PSD	PIP	PM	PSD	PIP	PM	PSD	PIP	PM	PSD	PIP	PM	PSD	PIP	PM	PSD
AccessAir	0.0585	0.0004	0.0019	**0.1127**	0.0009	0.0029	0.0320	0.0002	0.0012	0.0149	0.0001	0.0010	**0.8344**	0.0092	0.0048	0.0102	0.0001	0.0009
AccessRoad	0.0039	0.0000	0.0002	0.0044	0.0000	0.0002	0.0030	0.0000	0.0001	0.0055	0.0000	0.0004	0.0056	0.0000	0.0005	0.0229	-0.0001	0.0007
AirportDens	**0.2525**	1.2955	2.3880	**0.2201**	1.2667	2.5757	**0.2810**	1.3720	2.3277	0.0582	0.3071	1.3499	0.0855	0.5245	1.8186	0.0746	0.3805	1.4553
Airports	0.0049	0.0000	0.0000	0.0098	0.0000	0.0001	0.0023	0.0000	0.0000	0.0307	0.0000	0.0001	0.0072	0.0000	0.0000	0.0463	0.0000	0.0002
ARH0	0.0026	0.0000	0.0012	0.0029	0.0000	0.0013	0.0023	0.0000	0.0012	0.0014	0.0000	0.0008	0.0026	0.0000	0.0014	0.0313	0.0014	0.0084
ARL0	0.0034	0.0000	0.0012	0.0052	-0.0001	0.0013	0.0029	0.0000	0.0006	0.0072	-0.0001	0.0016	0.0016	0.0000	0.0004	0.0097	-0.0002	0.0035
ART0	0.0026	0.0000	0.0010	0.0025	0.0000	0.0013	0.0023	0.0000	0.0008	0.0015	0.0000	0.0007	0.0021	0.0000	0.0019	0.0078	0.0001	0.0026
Capital	**0.9972**	0.0115	0.0024	**0.9999**	0.0128	0.0022	**0.9945**	0.0107	0.0022	**1.0000**	0.0153	0.0020	**1.0000**	0.0171	0.0023	**1.0000**	0.0178	0.0019
ConnectAir	0.0707	-0.0003	0.0010	**0.1406**	-0.0005	0.0014	0.0304	-0.0001	0.0005	0.0119	0.0000	0.0003	0.0129	0.0000	0.0004	0.0132	0.0000	0.0004
ConnectSea	0.0038	0.0000	0.0001	0.0046	0.0000	0.0001	0.0019	0.0000	0.0000	0.0017	0.0000	0.0000	0.0020	0.0000	0.0000	0.0018	0.0000	0.0000
DistCap	0.0030	0.0000	0.0000	0.0044	0.0000	0.0000	0.0019	0.0000	0.0000	0.0061	0.0000	0.0000	0.0021	0.0000	0.0000	0.0024	0.0000	0.0000
Distde71	0.0025	0.0000	0.0000	0.0032	0.0000	0.0000	0.0019	0.0000	0.0000	0.0031	0.0000	0.0000	0.0031	0.0000	0.0000	0.0036	0.0000	0.0000
EMPDENS0	0.0228	0.0003	0.0020	0.0641	0.0008	0.0034	0.0034	0.0000	0.0005	0.0214	0.0002	0.0018	0.0754	0.0011	0.0039	0.0097	0.0001	0.0012
EREH0	0.0022	0.0000	0.0006	0.0027	0.0000	0.0008	0.0029	0.0000	0.0008	0.0021	0.0000	0.0007	0.0015	0.0000	0.0006	0.0247	0.0006	0.0042
EREL0	0.0021	0.0000	0.0006	0.0028	0.0000	0.0006	0.0025	0.0000	0.0005	0.0031	0.0000	0.0008	0.0024	0.0000	0.0004	0.0069	0.0000	0.0030
ERET0	0.0035	0.0000	0.0015	0.0039	0.0000	0.0009	0.0024	0.0000	0.0009	0.0019	0.0000	0.0010	0.0034	0.0000	0.0021	0.0412	0.0009	0.0049
GDPCAP0	**0.8687**	-0.0094	0.0044	**0.7685**	-0.0096	0.0058	**0.9308**	-0.0093	0.0033	**0.9653**	-0.0107	0.0030	**0.9999**	-0.0179	0.0029	**1.0000**	-0.0202	0.0016
gPOP	0.0018	-0.0001	0.0065	0.0030	-0.0002	0.0088	0.0013	0.0000	0.0040	0.0016	0.0000	0.0050	0.0021	0.0000	0.0062	0.0045	0.0007	0.0133
Hazard	0.0017	0.0000	0.0000	0.0029	0.0000	0.0000	0.0014	0.0000	0.0000	0.0075	0.0000	0.0000	0.0030	0.0000	0.0000	0.0015	0.0000	0.0000
HRSTcore	0.0032	0.0000	0.0008	0.0034	0.0000	0.0008	0.0025	0.0000	0.0007	0.0022	0.0000	0.0006	0.0024	0.0000	0.0006	0.0020	0.0000	0.0005
INTF	0.0037	0.0000	0.0011	0.0057	0.0001	0.0013	0.0037	0.0000	0.0011	0.0110	0.0002	0.0027	0.0200	0.0006	0.0045	0.0046	0.0001	0.0014
OUTDENS0	0.0048	0.0000	0.0000	0.0055	0.0000	0.0000	0.0041	0.0000	0.0000	0.0050	0.0000	0.0000	0.0034	0.0000	0.0000	0.0025	0.0000	0.0000
PatentBIO	0.0021	0.0001	0.0061	0.0029	0.0003	0.0085	0.0017	0.0000	0.0042	0.0016	0.0001	0.0052	0.0031	0.0005	0.0115	0.0018	0.0000	0.0052
PatentHT	0.0041	0.0001	0.0030	0.0067	0.0003	0.0042	0.0023	0.0000	0.0014	0.0018	0.0000	0.0016	0.0049	0.0002	0.0034	0.0036	0.0001	0.0028
PatentICT	0.0032	0.0001	0.0017	0.0061	0.0002	0.0026	0.0025	0.0000	0.0010	0.0033	0.0001	0.0016	0.0053	0.0002	0.0025	0.0043	0.0001	0.0021
PatentShBIO	0.0034	0.0000	0.0011	0.0030	0.0000	0.0009	0.0037	0.0001	0.0012	0.0015	0.0000	0.0006	0.0026	0.0000	0.0008	0.0022	0.0000	0.0007
PatentShHT	0.0019	0.0000	0.0003	0.0023	0.0000	0.0004	0.0019	0.0000	0.0003	0.0020	0.0000	0.0003	0.0020	0.0000	0.0004	0.0027	0.0000	0.0005
PatentShICT	0.0027	0.0000	0.0003	0.0027	0.0000	0.0004	0.0018	0.0000	0.0002	0.0025	0.0000	0.0003	0.0025	0.0000	0.0004	0.0032	0.0000	0.0005
PatentT	0.0028	0.0000	0.0005	0.0032	0.0000	0.0007	0.0014	0.0000	0.0003	0.0068	0.0001	0.0012	0.0049	0.0001	0.0010	0.0022	0.0000	0.0004
POPDENS0	0.0280	-0.0003	0.0018	0.0742	-0.0008	0.0032	0.0053	0.0000	0.0003	0.0316	-0.0003	0.0017	0.0809	-0.0010	0.0035	0.0102	-0.0001	0.0011
RailDens	0.0023	0.0000	0.0008	0.0024	0.0000	0.0009	0.0028	0.0000	0.0008	0.0017	0.0000	0.0006	0.0024	0.0000	0.0008	0.0017	0.0000	0.0005
RegBoarder	0.0017	0.0000	0.0000	0.0025	0.0000	0.0001	0.0016	0.0000	0.0000	0.0018	0.0000	0.0001	0.0020	0.0000	0.0001	0.0023	0.0000	0.0001
RegCoast	0.0021	0.0000	0.0001	0.0040	0.0000	0.0002	0.0016	0.0000	0.0001	0.0021	0.0000	0.0001	0.0022	0.0000	0.0001	0.0021	0.0000	0.0001
RegObj1	0.0283	0.0001	0.0009	0.0239	0.0001	0.0008	0.0279	0.0001	0.0009	0.0040	0.0000	0.0002	0.0029	0.0000	0.0002	0.0038	0.0000	0.0002
RegPent27	0.0026	0.0000	0.0001	0.0051	0.0000	0.0002	0.0022	0.0000	0.0001	0.0015	0.0000	0.0001	0.0176	0.0001	0.0006	0.0032	0.0000	0.0002
RoadDens	0.0049	0.0000	0.0007	0.0093	0.0001	0.0011	0.0020	0.0000	0.0002	0.0022	0.0000	0.0003	0.0027	0.0000	0.0003	0.0026	0.0000	0.0003
Seaports	0.0023	0.0000	0.0001	0.0027	0.0000	0.0001	0.0029	0.0000	0.0001	0.0048	0.0000	0.0002	0.0021	0.0000	0.0001	0.0031	0.0000	0.0002
Settl	0.0017	0.0000	0.0001	0.0018	0.0000	0.0001	0.0018	0.0000	0.0001	0.0014	0.0000	0.0000	0.0020	0.0000	0.0001	0.0021	0.0000	0.0001
ShAB0	0.0022	0.0000	0.0012	0.0023	0.0000	0.0011	0.0028	0.0000	0.0014	0.0019	0.0000	0.0012	0.0024	0.0000	0.0012	0.0018	0.0000	0.0010
ShCE0	0.0041	0.0000	0.0011	0.0053	0.0001	0.0015	0.0022	0.0000	0.0006	0.0025	0.0000	0.0008	0.0016	0.0000	0.0005	0.0180	0.0005	0.0038
shGFCF	0.0380	0.0009	0.0047	0.0049	0.0001	0.0012	0.0661	0.0015	0.0062	0.0026	0.0000	0.0007	0.0141	0.0003	0.0024	0.0028	0.0000	0.0006
ShLLL	0.0019	0.0000	0.0006	0.0024	0.0000	0.0008	0.0013	0.0000	0.0004	0.0020	0.0000	0.0009	0.0019	0.0000	0.0007	0.0164	0.0005	0.0046
ShSH	**0.8962**	0.0424	0.0173	**0.7745**	0.0365	0.0216	**0.9667**	0.0458	0.0131	**0.1202**	0.0043	0.0123	0.0217	0.0006	0.0047	**0.9773**	0.0476	0.0117
ShSL	0.0848	-0.0020	0.0068	**0.2043**	-0.0050	0.0101	0.0130	-0.0002	0.0020	**0.1627**	-0.0036	0.0087	0.0038	0.0000	0.0006	0.0197	-0.0003	0.0023
TELF	**0.1014**	-0.0003	0.0009	**0.2130**	-0.0007	0.0014	0.0352	-0.0001	0.0004	0.0579	-0.0002	0.0008	0.0052	0.0000	0.0001	0.0127	0.0000	0.0002
TELH	0.0017	0.0000	0.0000	0.0019	0.0000	0.0000	0.0017	0.0000	0.0000	0.0017	0.0000	0.0000	0.0023	0.0000	0.0000	0.0019	0.0000	0.0000
Temp	0.0028	0.0000	0.0001	0.0033	0.0000	0.0001	0.0022	0.0000	0.0001	0.0019	0.0000	0.0001	0.0020	0.0000	0.0001	0.0037	0.0000	0.0001
URH0	0.0068	0.0002	0.0036	0.0036	0.0000	0.0014	0.0086	0.0003	0.0044	0.0027	0.0001	0.0014	0.0019	0.0000	0.0010	0.0049	0.0000	0.0023
URL0	0.0033	0.0000	0.0007	0.0048	-0.0001	0.0010	0.0028	0.0000	0.0006	0.0019	0.0000	0.0006	0.0091	-0.0002	0.0022	0.0284	-0.0006	0.0041
URT0	0.0060	-0.0001	0.0020	0.0092	-0.0002	0.0024	0.0035	-0.0001	0.0016	0.0029	0.0000	0.0011	0.0044	-0.0001	0.0019	**0.1788**	-0.0067	0.0151

PIP stands for “posterior inclusion probability", PM stands for “posterior mean" and PSD stands for “posterior standard deviation". All calculations based on MC^3 sampling using 2 million posterior draws.

PIPs over 10% in bold. The form of the different weight matrices is presented in the text.

Table 3: BMA results based on different classes of spatial weight matrices

	$\mathbf{W}_1^Q$	$\mathbf{W}_2^Q$	$\mathbf{W}_3^Q$	$\mathbf{W}_4^Q$	$\mathbf{W}_4^{K-NN}$	$\mathbf{W}_5^{K-NN}$	$\mathbf{W}_6^{K-NN}$	$\mathbf{W}_{12}^{K-NN}$
PIP	0.0234	0.0000	1.7728	36.2936	0.0000	0.0016	0.0000	0.0000
	$\mathbf{W}_1^e$	$\mathbf{W}_2^e$	$\mathbf{W}_3^e$	$\mathbf{W}_4^e$	$\mathbf{W}_{400}^b$	$\mathbf{W}_{600}^b$	$\mathbf{W}_{800}^b$	$\mathbf{W}_{1000}^b$
PIP	0.0000	0.0000	61.8796	0.0000	0.0000	0.0290	0.0000	0.0000

Table 4: Posterior inclusion probability over space of weighting matrices (in %)

The convergence speed estimates for the model space which does not include spatial effects appear systematically higher than in the cases where spatial spillovers are explicitly modeled, as is expected if unmodeled positive spatial autocorrelation is present in the data. As we allow for more flexibility when modeling spatial autocorrelation patterns, the estimate of the "pure" speed of convergence (free of the effect of spatial autocorrelation) decreases. The differences are quantitatively very large. If we concentrate on the median of the conditional posterior distributions in Figure 4, the estimate of the speed of convergence falls from levels above 2% (a value which has become something of a stylized fact when it comes to cross-sectional growth regressions) in the posterior distribution based on models without spatial effects to 1% when the full set of 16 spatial weight matrices is conditioned upon. If a smaller set of spatial weighting matrices is conditioned upon, the resulting posterior distribution over the speed of convergence parameter lies somewhat between the two extremes, with a modal value of 2%.

We obtain an interesting insight to the importance of assessing model uncertainty in the framework of spatially autocorrelated data by comparing conditional and unconditional posterior distributions of the speed of convergence across European regions in Figure 4. While the two distributions do not differ strongly for the BMA estimates based on standard OLS and those based on the class of distance band matrices, as we allow for more flexibility when modeling spatial links a bimodal unconditional distribution with a zero mode emerges. This is the case since some of the spatial structures allow us to model the growth differences between rich and poor regions based on purely geographical patterns, without the need of including conditional convergence. This results in a group of models which do not include the initial level of income per capita as a regressor but are able to explain growth differentials relatively well. Table 3 and Table 4 show that such an effect is related to the inclusion of Queen contiguity matrices in the conditioning set.

This implies that the estimates obtained in previous research concerning the regional speed of income convergence using models which condition on a single spatial link matrix tended to overestimate the extent of the income convergence process. This result is particularly important for economic policy exercises implying the *ceteris paribus* condition in models with spatial autocorrelation. The interpretation of the income convergence speed in neoclassical economic growth models is based on the assumption that all factors affecting economic growth remain constant with the exception of income itself. In the framework of models with spatial spillovers in the form of spatial autoregressive specifications, an extra assumption for the interpretation of the speed of convergence is that the income levels of other regions remain constant, so that ensuring the lack of spatial correlation in the residuals of the growth regression is an important prerequisite for the analysis. Our results give strong evidence that the quantitative assessment of income convergence across regions requires a systematic treatment of uncertainty in the nature of spatial growth spillovers.

4 Conclusions

We put forward a Bayesian Model Averaging method for dealing with model uncertainty in the presence of potential spatial autocorrelation of unknown form. We propose using spatial filtering methods to, on the one hand, exploit large sets of possible classes of spatial weighting matrices and, on the other hand, achieve computational gains as compared to the direct estimation of spatial autoregressive models. Using simulations, we show that the method is able to identify correctly covariates and spatial patterns present in the data.

We use our method to evaluate the robustness of growth determinants across European regions for the period 1995-2005 and, in particular, to estimate the speed of income convergence. We show that the choice of a type of specification in terms of a particular class of spatial weighting matrices can have an important effect on the estimates of the parameters attached to the model covariates. We also show that estimates of the speed of income convergence across European regions depend strongly on the form of the spatial patterns which are assumed to underly the dataset. When we take into account this dimension of model uncertainty, the posterior distribution of the speed of convergence parameter appears bimodal, with a large probability mass around no convergence (0% speed of convergence) and a rate of convergence of 1%, approximately half of the value which is usually reported in the literature. Our results indicate that previous research concerning the regional speed of income convergence in models which condition on single spatial link matrices tended to overestimate the catching-up process in income levels.

References

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

Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *The Quarterly Journal of Economics*, 106, No. 2:407–443.

Barro, R. J. and Sala-i-Martin, X. (1991). Convergence across States and Regions. *Brookings Papers on Economic Activity*, 1:107–182.

Barro, R. J. and Sala-i-Martin, X. (2003). *Economic Growth*. MIT Press.

Bernardo, J. and Smith, A. (1994). *Bayesian Theory*. John Wiley and Sons, New York.

Boldrin, M. and Canova, F. (2001). Inequality and Convergence in Europe's Regions: Reconsidering European Regional Policies. *Economic Policy*, 16:205–253.

Crespo Cuaresma, J., Doppelhofer, G., and Feldkircher, M. (2009). The determinants of economic growth in european regions. *Working Papers 2008-26, Faculty of Economics and Statistics, University of Innsbruck, and CESifo Working Paper Series.*

Fernández, C., Ley, E., and Steel, M. F. (2001a). Benchmark Priors for Bayesian Model Averaging. *Journal of Econometrics*, 100:381–427.

Fernández, C., Ley, E., and Steel, M. F. (2001b). Model Uncertainty in Cross-Country Growth Regressions. *Journal of Applied Econometrics*, 16:563–576.

Fischer, M. and Stirböck, C. (2006). Pan-European Regional Income Growth and Club-Convergence. Insightsfrom a Spatial Econometric Perspective. *Annals of Regional Science*, 40(3):1–29.

Foster, D. P. and George, E. I. (1994). The Risk Inflation Criterion for Multiple Regression. *The Annals of Statistics*, 22:1947–1975.

Getis, A. and Griffith, D. A. (2002). Comparative Spatial Filtering in Regression Analysis. *Geographical Analysis*, 34:2:130–140.

Koop, G. (2003). Introduction to Bayesian Econometrics. *John Wiley*.

LeSage, J. P. and Fischer, M. (2007). Spatial Growth Regressions, Model Specification, Estimation, and Interpretation. *SSRN Working Paper*, 980965.

LeSage, J. P. and Parent, O. (2007). Bayesian Model Averaging for Spatial Econometric Models. *Geographical Analysis*, 39:3:241–267.

LeSage, J. P. and Parent, O. (2008). Using the Variance Structure of the Conditional Spatial Specifcation to Model Knowledge Spillovers. *Journal of Applied Econometrics*, 23:235–256.

Ley, E. and Steel, M. F. (2009). On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regressions. *Journal of Applied Econometrics*, (forthcoming).

Liang, F., Paulo, R., Molina, G., Clyde, M. A., and Berger, J. O. (2008). Mixtures of g Priors for Bayesian Variable Selection. *Journal of the American Statistical Association*, 103:410–423.

Mathunjwa, J. S. and Temple, J. (2007). Convergence behaviour in exogenous growth models,. *Bristol Economics Discussion Papers*, 06/590.

Niebuhr, A. (2001). Convergence and the Effects of Spatial Interaction. *Jahrbuch für Regionalwissenschaft*, 21:113–133.

Raftery, A. E. (1995a). Bayesian Model Selection in Social Research. *Sociological Methodology*, 25:111–163.

Raftery, E. A. (1995b). *Sociological Methodology*. Blackwell Publishing.

Sala-i-Martin, X. (1996). Regional Cohesion: Evidence and Theories of Regional Growth and Convergence. *European Economic Review*, 40:1325–52.

Sala-i-Martin, X., Doppelhofer, G., and Miller, R. I. (2004). Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. *American Economic Review*, 94:813–835.

Schwarz, G. (1978). Estimating the Dimensions of a Model. *Annals of Statistics*, 6(2):461–464.

Tiefelsdorf, M. and Griffith, D. A. (2007). Semiparametric filtering of spatial autocorrelation: the eigenvector approach. *Environment and Planning A*, 37:1193–1221.

Technical Appendix

Priors over the parameter space conditional on a model specification

BMA belongs to the class of shrinkage estimators, where shrinkage over models is governed by the parameter g, which elicits our prior over slope parameters. The choice of g is thus crucial for posterior inference. Fernández et al. (2001a) propose an automated way to choose g based on an exhaustive simulation study. The benchmark prior advocated by Fernández et al. (2001a) amounts to setting $g = 1/\max(N, K^2)$. This prior structure bridges between the unit information prior (UIP, $g = 1/N$) proposed by Kass and Raftery (1995a) and Raftery (1995b) and the risk information criterion (RIC, $g = 1/K^2$) by Foster and George (1994). The use of UIP implies that the Bayes factor can be interpreted asymptotically (and therefore approximated) as the difference of the Schwarz information criterion (Schwarz (1978)) values for the two corresponding models. Other approaches include mixtures of g-priors and variants of the Zellner-Siow prior (Liang et al. (2008)). Throughout the paper we use the benchmark prior in Fernández et al. (2001a), which implies that for the setting in our empirical application the RIC is preferred when choosing g.

Priors over the model space

A prior over models in $\mathcal{M}$ has to be chosen in order to obtain BMA estimates of the parameters. Two typical prior specifications have been usually imposed in the literature: a) an uninformative flat prior over all models, which implies that the posterior odds ratio resembles solely the Bayes factor and comparison of models is governed by their relative marginal likelihoods, and b) a prior that discriminates among models according to the number of regressors they include, so that a larger prior probability mass falls over models of a given size (see Sala-i-Martin et al. (2004)). This second alternative is instrumentalized by assuming that each covariate enters the regression with probability ϑ, which implies that the prior mass for model j which includes k_j variables (in addition to the eigenvectors used for spatial filtering) amounts to $\overline{p}(M_j^z) = \vartheta^{k_j}(1-\vartheta)^{K-k_j}$. The uninformative prior in a) is nested in b) by imposing $\vartheta = 1/2$, which results into equal model probabilities of 2^{-K} for all models for each spatial matrix, thus $2^{-K\times Z}$ is the prior inclusion probability of each model in our case.

Ley and Steel (2009) show that fixing $\vartheta = 1/2$ puts most mass on models with $K/2$ regressors, since they are dominant in number. Their recommendation is thus to treat ϑ as random and placing a (hyper)prior on it. The proposal of Ley and Steel (2009) is to impose that the model size follows a Binomial-Beta(a, b) distribution (Bernardo and Smith (1994)) with $a = 1$, so that

$$P(k = k_j) = \frac{\Gamma(1+b)}{\Gamma(1) + \Gamma(b) + \Gamma(1+b+K)} \binom{K}{k_j} \Gamma(1+k_j)\Gamma(b+K-k_j) \quad k_j = 0, \ldots, K. \tag{9}$$

The prior can be elicited by anchoring the prior expected model size, m.[9] Ley and Steel (2009) quantify the influence that a poorly specified prior exerts on posterior results when ϑ

[9] Note that b is then implicitly defined through $b = (K-m)/m$.

is fixed, which leads to the relative merits of BMA being less pronounced and its predictive power deteriorating. In contrast, the results in Ley and Steel (2009) indicate that the choice of m has no influential impact on posterior inference and the prior over models is purely non-informative.

Data Appendix

Country	Region	
Austria	Burgenland	Salzburg
	Kärnten	Steiermark
	Niederösterreich	Tirol
	Oberösterreich	Vorarlberg
	Wien	
Belgium	Prov. Antwerpen	Prov. Luxembourg (B)
	Prov. Brabant Wallon	Prov. Namur
	Prov. Hainaut	Prov. Oost-Vlaanderen
	Prov. Liège	Prov. Vlaams Brabant
	Prov. Limburg (B)	Prov. West-Vlaanderen
	Région de Bruxelles-Capitale	
Bulgaria	Severen tsentralen	Yugoiztochen
	Severoiztochen	Yugozapaden
	Severozapaden	Yuzhen tsentralen
Cyprus	Cyprus	Severovýchod
Czech Republic	Jihovýchod	Severozápad
	Jihozápad	Strední Cechy
	Moravskoslezsko	Strední Morava
	Praha	
Denmark	Denmark	
Estonia	Estonia	
Finland	Åland	Länsi-Suomi
	Etelä-Suomi	Pohjois-Suomi
	Itä-Suomi	
France	Alsace	Île de France
	Aquitaine	Languedoc-Roussillon
	Auvergne	Limousin
	Basse-Normandie	Lorraine
	Bourgogne	Midi-Pyrénées
	Bretagne	Nord - Pas-de-Calais
	Centre	Pays de la Loire
	Champagne-Ardenne	Picardie
	Corse	Poitou-Charentes
	Franche-Comté	Provence-Alpes-Côte d'Azur
	Haute-Normandie	Rhône-Alpes
Germany	Arnsberg	Lüneburg
	Berlin	Mecklenburg-Vorpommern
	Brandenburg - Nordost	Mittelfranken
	Brandenburg - Südwest	Münster
	Braunschweig	Niederbayern
	Bremen	Oberbayern
	Chemnitz	Oberfranken
	Darmstadt	Oberpfalz
	Detmold	Rheinhessen-Pfalz
	Dresden	Saarland
	Düsseldorf	Schleswig-Holstein
	Freiburg	Schwaben
	Giessen	Stuttgart
	Hamburg	Thüringen
	Hannover	Trier
	Karlsruhe	Tübingen

	Kassel	Unterfranken
	Koblenz	Weser-Ems
	Köln	Leipzig
Greece	Anatoliki Makedonia, Thraki	Kriti
	Attiki	Notio Aigaio
	Dytiki Ellada	Peloponnisos
	Dytiki Makedonia	Sterea Ellada
	Ionia Nisia	Thessalia
	Ipeiros	Voreio Aigaio
	Kentriki Makedonia	
Hungary	Dél-Alföld	Közép-Dunántúl
	Dél-Dunántúl	Közép-Magyarország
	Észak-Alföld	Nyugat-Dunántúl
	Észak-Magyarország	
Ireland	Border, Midlands and Western	
	Southern and Eastern	
Italy	Abruzzo	Molise
	Basilicata	Piemonte
	Calabria	Bolzano-Bozen
	Campania	Trento
	Emilia-Romagna	Puglia
	Friuli-Venezia Giulia	Sardegna
	Lazio	Sicilia
	Liguria	Toscana
	Lombardia	Umbria
	Marche	Valle d'Aosta
	Veneto	
Latvia	Latvia	
Lithuania	Lithuania	
Luxembourg	Luxembourg (Grand-Duché)	
Malta	Malta	
Netherlands	Drenthe	Noord-Brabant
	Flevoland	Noord-Holland
	Friesland	Overijssel
	Gelderland	Utrecht
	Groningen	Zeeland
	Limburg (NL)	Zuid-Holland
Poland	Dolnoslaskie	Podkarpackie
	Kujawsko-Pomorskie	Podlaskie
	Lódzkie	Pomorskie
	Lubelskie	Slaskie
	Lubuskie	Swietokrzyskie
	Malopolskie	Warminsko-Mazurskie
	Mazowieckie	Wielkopolskie
	Opolskie	Zachodniopomorskie
Portugal	Alentejo	Lisboa
	Algarve	Norte
	Centro (PT)	
Romania	Bucuresti - Ilfov	Sud - Muntenia
	Centru	Sud-Est
	Nord-Est	Sud-Vest Oltenia
	Nord-Vest	Vest
Slovak Republic	Bratislavský kraj	Východné Slovensko
	Stredné Slovensko	Západné Slovensko

Slovenia	Slovenia	
Spain	Andalucia	Extremadura
	Aragón	Galicia
	Cantabria	Illes Balears
	Castilla y León	La Rioja
	Castilla-la Mancha	Pais Vasco
	Cataluña	Principado de Asturias
	Comunidad de Madrid	Región de Murcia
	Comunidad Foral de Navarra	Comunidad Valenciana
Sweden	Mellersta Norrland	Småland med öarna
	Norra Mellansverige	Stockholm
	Östra Mellansverige	Sydsverige
	Övre Norrland	Västsverige
United Kingdom	Bedfordshire, Hertfordshire	Kent
	Berkshire, Bucks and Oxfordshire	Lancashire
	Cheshire	Leicestershire, Rutland and Northants
	Cornwall and Isles of Scilly	Lincolnshire
	Cumbria	Merseyside
	Derbyshire and Nottinghamshire	North Yorkshire
	Devon	Northern Ireland
	Dorset and Somerset	Northumberland, Tyne and Wear
	East Anglia	Outer London
	East Riding and North Lincolnshire	Shropshire and Staffordshire
	East Wales	South Western Scotland
	Eastern Scotland	South Yorkshire
	Essex	Surrey, East and West Sussex
	Gloucestershire, Wiltshire and North Somerset	Tees Valley and Durham
	Greater Manchester	West Midlands
	Hampshire and Isle of Wight	West Wales and The Valleys
	Herefordshire, Worcestershire and Warks	West Yorkshire
	Inner London	

Table A.1: European regions in the sample

Variable name	Description	Source
Dependent variable		
gGDPCAP	Growth rate of real GDP per capita	Eurostat
Factor accumulation/convergence		
GDPCAP0	Initial real GDP per capita (in logs)	Eurostat
gPOP	Growth rate of population	Eurostat
shGFCF	Share of GFCF in GVA	Cambridge Econometrics
Infrastructure		
INTF	Proportion of firms with own website regression	ESPON
TELH	A typology of levels of household telecommunications uptake	ESPON
TELF	A typology of estimated levels of business telecommunications access and uptake	ESPON
Seaports	Regions with seaports	ESPON
AirportDens	Airport density	ESPON
RoadDens	Road density	ESPON
RailDens	Rail density	ESPON
ConnectAir	Connectivity to commercial airports by car	ESPON
ConnectSea	Connectivity to commercial seaports by car	ESPON
AccessAir	Potential accessibility air	ESPON
AccessRoad	Potential accessibility road	ESPON
Socio-geographical variables		
Settl	Settlement structure	ESPON
OUTDENS0	Initial output density	
EMPDENS0	Initial employment density	
POPDENS0	Initial population density	
RegCoast	Coast	ESPON
RegBorder	Border	ESPON
RegPent27	Pentagon EU 27 plus 2	ESPON
RegObj1	Objective 1 regions	ESPON
Capital	Capital city	
Airports	Number of airports	ESPON
Temp	Extreme temperatures	ESPON
Hazard	Sum of all weighted hazard values	ESPON
Distde71	Distance to Frankfurt	
DistCap	Distance to capital city	
Technological innovation		
PatentT	Number of patents total	Eurostat
PatentHT	Number of patents in high technology	Eurostat
PatentICT	Number of patents in ICT	Eurostat
PatentBIO	Number of patents in biotechnology	Eurostat
PatentShHT	Share of patents in high technology	Eurostat
PatentShICT	Share of patents in ICT	Eurostat
PatentShBIO	Share of patents in biotechnology	Eurostat
HRSTcore	Human resources in science and technology (core)	Eurostat LFS
Human capital		
ShSH	Share of high educated in working age population	Eurostat LFS

ShSL	Share of low educated in working age population	Eurostat LFS
ShLLL	Life long learning	Eurostat LFS

Sectoral structure/employment

ShAB0	Initial share of NACE A and B (Agriculture)	Eurostat
ShCE0	Initial share of NACE C to E (Mining, Manufacturing and Energy)	Eurostat
EREH0	Employment rate - high	Eurostat LFS
EREL0	Employment rate - low	Eurostat LFS
ERET0	Employment rate - total	Eurostat LFS
URH0	Unemployment rate - high	Eurostat LFS
URL0	Unemployment rate - low	Eurostat LFS
URT0	Unemployment rate - total	Eurostat LFS
ARH0	Activity rate high	Eurostat LFS
ARL0	Activity rate low	Eurostat LFS
ART0	Activity rate total	Eurostat LFS

Table A.2: Variables, description and sources