

Schoiswohl, Florian; Piribauer, Philipp; Gmeinder, Michael; Koch, Matthias; Fischer, Manfred

Conference Paper

Regional income convergence in Europe: a Bayesian model averaging perspective

52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Schoiswohl, Florian; Piribauer, Philipp; Gmeinder, Michael; Koch, Matthias; Fischer, Manfred (2012) : Regional income convergence in Europe: a Bayesian model averaging perspective, 52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/120673>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Regional Income Convergence in Europe: A Bayesian Model Averaging Perspective¹

Michael Gmeinder, Philipp Piribauer, Florian Schoiswohl,

Matthias Koch, Manfred M. Fischer

Vienna University of Economics and Business, Austria

52th European Congress of the Regional Science Association International

August 21-25, 2012, Bratislava, Slovakia

Abstract

Recently Bayesian Model Averaging and spatial filtering techniques have been combined in the study of income convergence, in order to deal with the uncertainties regarding potential growth determinants and the underlying spatial structure that arise in the context of spatial growth regressions. While spatial eigenvector filtering has proven to be computationally advantageous for a limited number of observations, the method also imposes restrictions regarding the interpretation of the parameter estimates. This paper attempts to overcome this issue by using spatial filtering together with Bayesian Model Averaging as a means to identify 10,000 top models (i.e. the models with the highest posterior probability), which can in turn be used for impact estimations within Spatial Autoregressive and Spatial Durbin specifications. These impact estimates allow for a more detailed assessment of convergence between regions, as they account for effects that can be attributed to spatial spillovers. Using a dataset covering 255 European regions (NUTS-2), we illustrate our approach and compare the findings with spatial filtering estimates that neglect the role of spatial spillovers in the convergence process.

¹ The study reported in this paper was integral part of the seminar course Spatial Economics (Instructors: Manfred M. Fischer and Matthias Koch) in the Economics Master's Program of the Vienna University of Economics and Business. We are very grateful for the excellent research environment provided by the Institute for Economic Geography and GIScience. Further we would like to thank Jesús Crespo-Cuaresma and Martin Feldkircher for their helpful comments.

1 Introduction

Spatial regression models have gained popularity in recent years in the empirical study of the determinants of economic growth and the related question of income convergence. Using such models, growth spillovers between regions can be explicitly taken into account. However, when investigating growth determinants and income convergence in a spatial econometric framework, one is confronted with two dimensions of uncertainty. The first dimension, model and parameter uncertainty, arises from what Brock and Durlauf (2001) call the ‘open-endedness’ of growth theories, which refers to the existence of a multitude of theories that aim to explain economic growth which are not mutually exclusive, i.e. finding evidence for the validity of one theory does not render the other growth theories invalid. As a consequence, there are many variables that have the potential to explain growth and are therefore candidates for being included in the model to be estimated. It is, however, inefficient or even infeasible to include all of the potential growth determinants into a single model. Thus, there is the question of which variables should be considered, i.e. we are uncertain about the ‘correct’ model. There is a growing literature that tries to address the issue of model uncertainty by employing Bayesian Model Averaging (BMA) techniques, for example Crespo-Cuaresma et al. (2009) who carry out a BMA exercise in the context of a Spatial Autoregressive (SAR) model.

Even though model and parameter uncertainty is accounted for in the approach taken by Crespo-Cuaresma et al. (2009), the BMA exercise is carried out conditional on a given spatial weight matrix. In fact, most of the empirical studies thus far stick to a single spatial weighting scheme. There are, however, many ways to define such a neighborhood structure, which is why studies sticking to a single spatial weight matrix ignore the second dimension of uncertainty, namely uncertainty with respect to the underlying spatial structure. Noteworthy exceptions are the approaches proposed by LeSage and Fischer (2008) and Crespo-Cuaresma and Feldkircher (2012). Relying on numerical integration techniques, the former consider different spatial structures originating from the same class of weighting schemes (k-nearest neighbors) in the context of a Spatial Durbin Model (SDM). This approach is computationally burdensome when using a large number of variables and spatial weight matrices. A computationally more efficient approach that allows for the inclusion of many different spatial weight matrices is proposed by the latter, who make use of spatial filtering techniques based on the eigenvector approach put forward by Tiefelsdorf and Griffith (2007).

Since spatial filtering treats spatial effects as a mere nuisance that the data is being rid of, a coefficient capturing growth spillovers cannot be elicited. As a consequence, the impacts of

changes in the explanatory variables cannot be properly assessed, i.e. it is not possible to determine direct, indirect and total effects, as suggested by LeSage and Pace (2009). Concerning the question of income convergence this implies that only the part of convergence that is spillover free can be obtained. The calculation of direct, indirect and total effects is therefore necessary, in order to arrive at a more complete picture regarding the magnitude of income convergence, i.e. convergence estimates that include effects stemming from spatial spillovers. Hence, relying on the approach by Crespo-Cuaresma and Feldkircher (2012) and thereby focusing on a number of top models as well as their posterior model probabilities, we propose a procedure which allows the computation of direct, indirect and total effects. We apply this procedure, using the same dataset as Crespo-Cuaresma and Feldkircher (2012) and Crespo-Cuaresma et al. (2009), to investigate income convergence across 255 European regions (NUTS-2) for the period 1995 to 2005.

This paper is organized as follows. While section 2 lays out the econometric setting, sections 3 and 4 introduce the techniques used by Crespo-Cuaresma and Feldkircher (2012) which allows taking uncertain spatial effects into account. Section 5 explains how these findings can be used for obtaining an estimate of the spatial spillover parameter which in turn renders a proper impact assessment possible. The results of our empirical exercise are presented in section 6. Section 7 concludes.

2 Econometric Setting

Following Crespo-Cuaresma and Feldkircher (2012) our point of departure in the empirical analysis of income convergence across European regions is marked by SAR model, which allows for explicitly modeling potential growth spillovers by introducing a so-called spatial lag of the dependent variable on the right-hand side of the regression equation. The SAR model can be written as:

$$y = \alpha \iota_N + \rho W y + X_k \beta + \varepsilon \quad (1)$$

y is an N -dimensional vector containing the annual income growth rate of each cross-sectional unit, where N denotes the total number of observations (regions). α is the intercept parameter and ι_N is an N -dimensional vector of ones. The second term on the right-hand side of (1) constitutes the spatial lag of the dependent variable, where the $N \times N$ spatial weight matrix W specifies the underlying spatial structure, with its corresponding parameter ρ capturing the intensity of spatial autocorrelation. The elements W_{ij} of the W matrix assume a

non-zero value when observations i and j are neighboring regions and zero otherwise. Since a region cannot be its own neighbor, all entries on the main diagonal are zero, i.e. $W_{ii} = 0$. Positive values for ρ can be interpreted as spillovers from neighboring spatial units, which in our case amounts to income growth spillovers across European regions. If $\rho = 0$, we are left with a non-spatial specification, which is therefore just a special case of the SAR model. β is a k -dimensional vector of parameters corresponding to \mathbf{X}_k , a $N \times k$ matrix whose columns represent the k explanatory variables, which are composed of potential growth determinants in our analysis. A potential growth determinant in \mathbf{X}_k , which is of particular interest, is the level of income per capita at the beginning of the period for which the average annual income growth rate is calculated. If we find the coefficient of initial income to be negative, after controlling for other factors as well as accounting for spatial spillovers, conditional income convergence occurs, i.e. regions with low initial income levels grow on average faster, compared to regions that started out with relatively higher per capita income levels. ε represents the $N \times 1$ vector of iid errors.

The two initially mentioned dimensions of uncertainty affect different parts of the model specification given in (1). While model uncertainty concerns the set of covariates \mathbf{X}_k to be included in the model, uncertainty about the underlying spatial structure has implications for the choice of the spatial weight matrix \mathbf{W} . Non-spatial BMA procedures or BMA procedures with certain spatial effects, i.e. using one \mathbf{W} matrix, as in Crespo-Cuaresma et al. (2009), have been put forward as a means of taking model uncertainty into account. This leaves us, however, with the problem of uncertainty regarding the spatial effects. Therefore, the following two sections introduce the approach by Crespo-Cuaresma and Feldkircher (2012), which, through the use of spatial filtering techniques, allows for uncertain spatial effects in the BMA exercise.

3 Spatial filtering

Two different approaches, which account for uncertainty with respect to the set of covariates *and* the underlying spatial structure \mathbf{W} are worth mentioning: (a) numerical integration techniques as put forward by LeSage and Parent (2007) as well as LeSage and Fischer (2008) and (b) spatial filtering approaches, as suggested by Tiefelsdorf and Griffith (2007) and used within a BMA setting by Crespo-Cuaresma and Feldkircher (2012).

While the literature has shown that numerical integration approaches are computationally expensive, the eigenvector approach of spatial filtering is able to mitigate this problem and

allows for many different spatial structures to be considered in a BMA exercise. This advantage results from the fact that the procedure filters out the spatial part of the data *prior to* estimating. However, spatial filtering has its drawbacks: First, filtering is an approximation technique and the appropriateness of the approximation depends on the properties of the spatial structure \mathbf{W} .² Furthermore, as the spatial information is filtered out, no inference about the spatial parameter ρ can be drawn.

We start by rewriting the SAR model specification given in (1), using the properties of the Neumann series, so that the equation is divided in a spatial and a non-spatial part:

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

The term $\sum_{m=1}^{\infty} \rho^m \mathbf{W}^m (\alpha \iota_N + \mathbf{X}_k \beta + \varepsilon)$ is treated as a nuisance term and is to be filtered out by finding a suitable approximation which replaces that term in the estimation. Following Tiefelsdorf and Griffith (2007), a subset of eigenvectors of the matrix \mathbf{W} can be used as such a proxy. Based on the set of eigenvectors resulting from an eigenvalue decomposition of \mathbf{W} , a minimal subset of E eigenvectors $\{e_1, \dots, e_E\}$ required for the approximation can be identified via a step-by-step procedure. Using the eigenvectors, a simple first-order SAR process can be rewritten as: $y = \alpha \iota_N + \rho \mathbf{W}y + u \sim \alpha \iota_N + \sum_{i=1}^E \gamma_i e_i + \tilde{u}$. Starting with the eigenvector that corresponds to the highest eigenvalue, we continue adding eigenvectors (in decreasing manner according to their corresponding eigenvalues) to this rewritten SAR process as long as the Moran's I statistic of the residuals \tilde{u} lies above the threshold 0.1. The set of E eigenvectors that needs to be included in order to undercut this value represents the aforementioned subset.

Using the extracted subset of eigenvectors of \mathbf{W} as an approximation for the last term in (2), our estimation equation can be rewritten as follows:

$$y = \alpha \iota_N + \sum_{i=1}^E \gamma_i e_i + \mathbf{X}_k \beta + \tilde{\varepsilon} \quad (3)$$

²Pace et.al. (2011) discuss this issue in more detail.

4 Bayesian Model Averaging

Since we use a Bayesian framework, we are able to depart from single model considerations by simply applying Bayes' Theorem. The posterior distribution of the parameters β contains all information about β after seeing the data and is given by:

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

where D denotes the available data and β a certain parameter.³ The sub- and superscripts j and z indicate that we consider more than just a single model, with $M_j^z \in \mathbf{M}$. The model space is given by $\mathbf{M} = \{M_1^1, M_2^1, \dots, M_{2^K}^1, \dots, M_1^2, \dots, M_{2^K}^2, \dots, M_1^Z, \dots, M_{2^K}^Z\}$, with K and Z denoting the set of potential explanatory variables and the potential spatial weight matrices, respectively. This leads to a total number of $2^K \times Z$ potential models.

As shown in (4), the idea of BMA is to evaluate the entire model space \mathbf{M} , and calculate a weighted average of the models using posterior model probabilities $p(M_j^z|D)$. Let $p(M_j^z)$ denote the prior model probability assigned to model M_j^z , then the posterior model probability (PMP) is given by:

$$p(M_j^z|D) = \frac{p(D|M_j^z)p(M_j^z)}{\sum_{j=1}^{2^K} \sum_{z=1}^Z p(D|M_j^z)p(M_j^z)} \quad (5)$$

Since we are not only interested in the relevance of certain models, we can also calculate the relative importance of the inclusion of a certain variable, the posterior inclusion probability (PIP) of variable x_k , which is given by the sum of PMPs including variable x_k .

In order to obtain the integrated likelihood $p(D|M_j^z)$, we have to assign prior distributions to the parameters σ and β . For β we use the well-known g -prior:⁴

$$p(\beta|\alpha, \rho, \sigma, M_j^z) \sim N(0, g\sigma^2(\mathbf{X}_k' \mathbf{X}_k)^{-1}) \quad (6)$$

A g -prior specification provides several advantages. Perhaps the most advantageous property of the g -prior is that the integrated likelihood has a closed form, which reduces the computational burden severely, since we can rely on analytical solutions. Furthermore, it

³See Koop (2003) or Hoeting et al. (1999) for more details about BMA.

⁴See for instance Zellner (1989).

seems reasonable to specify the prior variance of the estimators in a similar way as a classical OLS estimate does. We therefore only have to choose the scalar g , which measures our uncertainty about the prior mean.

There is a vast literature that deals with the choice of g . In our application we rely on the results of Fernández et al. (2001) and set $g = \max(N, K^2)$, commonly dubbed the BRIC-prior.⁵ For the spatial parameter ρ we follow LeSage and Parent (2007) and use a beta prior distribution.

The only thing that remains to be chosen is a prior on the model space. We depart from the usual choice of a uniformly distributed model prior, and make use of a binomial-beta prior, as proposed by Ley and Steel (2009).

It is computationally infeasible to work through the entire model space. Fortunately, we can ease this computational burden by using Markov Chain Monte Carlo Model Comparison (MC³) to find a subsample of models, which acts as a proper proxy for the entire model space \mathcal{M} . Since we are concerned with uncertainty about the explanatory variables as well as uncertainty about the spatial weighting scheme, we need a sampling procedure which considers both dimensions.⁶ Our sampling algorithm works as follows.

Starting with a given set of eigenvectors, we use a birth-death sampler (Madigan and York 1995) using 7,000,000 iterations and 5,000,000 burn-ins. However, we slightly modify the sampler to ensure a proper treatment of interaction terms, which is in accordance with the strong heredity principle, as proposed by Chipman (1996). Strong heredity means that an interaction between two variables (AB) is only eligible to be active, if both parent variables (A and B) are included in the sampled model. Likewise, squared variables are only eligible to be sampled when the non-squared parent is active. The acceptance probability of the candidate model M_c^z is given by:

$$\tilde{p}_{cj} = \min \left[1, \frac{p(M_c^z)p(D|M_c^z)}{p(M_j^z)p(D|M_j^z)} \right] \quad (7)$$

The second step is concerned with the sampling of the spatial weight matrices, or rather the corresponding set of eigenvectors. Therefore, the sampler also uniformly draws a set of eigenvectors of the other weight matrices. Following Crespo-Cuaresma and Feldkircher (2012), we use a uniform prior on the sets of spatial weight matrices. The

⁵See Fernandez et al. (2001)

⁶For a detailed motivation of MC³-sampling see Madigan and York (1995) or Green (1995).

acceptance probability of the candidate set of eigenvectors representing the spatial structure is given by:

$$\hat{p}_{cz} = \min \left[1, \frac{p(D|M_j^c)}{p(D|M_j^z)} \right] \quad (8)$$

5 Estimation procedure and data

The combination of the methods introduced in the two previous sections, namely spatial eigenvector filtering in a BMA framework (‘spatial filtering BMA’), enables us to deal with uncertainty regarding the covariates on the one hand and uncertainty concerning the underlying spatial effects on the other. However, as all spatial information and therefore the spatial autoregressive parameter ρ , which captures spatial spillovers and in turn induces feedback effects, vanishes due to the filtering procedure, we are no longer able to give a full account of the impacts associated with changes in the explanatory variables. In other words, the parameter estimates and the convergence parameter in particular can no longer be interpreted accordingly, as it is not possible to calculate partial derivatives. In order to regain the information about the spatial parameter ρ , we propose, in a first step, a spatial filtering BMA exercise as a device to identify the 10,000 models with the highest PMPs – the so-called *top models* – each containing a certain set of covariates \mathbf{X}_k and a specific spatial structure \mathbf{W} .

In a second step, we re-estimate the 10,000 top models using SAR and SDM specifications:

$$\begin{aligned} SAR: \quad y &= \alpha \iota_N + \rho \mathbf{W}y + \mathbf{X}_k \beta_k + \varepsilon \quad ; \quad \varepsilon \sim N(0, \sigma^2 \mathbf{I}_N) \\ SDM: \quad y &= \alpha \iota_N + \rho \mathbf{W}y + \mathbf{X}_k \beta_k + \mathbf{W} \mathbf{X}_k \theta_k + \varepsilon \quad ; \quad \varepsilon \sim N(0, \sigma^2 \mathbf{I}_N) \end{aligned} \quad (9)$$

As spatial filtering annihilates all spatial information, we remain agnostic about the particular form of the data generating process (DGP). To mitigate this problem we separately use both model specifications, in order to provide evidence for the robustness of our estimates.

Since there are N^2 partial derivatives of y with respect to each parameter, inference is somewhat more difficult than in a classical linear model specification. For the interpretation of the parameter estimates we use commonly known metrics of average direct, indirect and total effects. For the parameter β_k of variable x_k , these effects are in the SDM case defined as follows:⁷

⁷See LeSage and Pace (2008)

$$\begin{aligned}
\overline{direct}(k) &= N^{-1}\text{tr}(\mathbf{S}_k(\mathbf{W})) \\
\overline{indirect}(k) &= \overline{total}(k) - \overline{direct}(k) \\
\overline{total}(k) &= N^{-1}\mathbf{1}_N' \mathbf{S}_k(\mathbf{W}) \mathbf{1}_N \\
\mathbf{S}_k &= (\mathbf{I}_N - \rho \mathbf{W})^{-1}(\mathbf{I}_N \beta_k + \mathbf{W} \theta_k)
\end{aligned} \tag{10}$$

The effects for the SAR case can be retrieved by setting $\theta_k = 0$. Once we have obtained 10,000 estimates for the impacts from our set of top models, we use the corresponding PMPs obtained from the spatial filtering BMA exercise to calculate weighted averaged impacts. By weighting the 10,000 parameter estimates for ρ with the respective PMPs, we are also able to retrieve an estimate for the spatial parameter.

Since our regression involves several interaction terms, the impact metrics of the interaction terms need to be slightly modified. If the variable x_k of interest is also interacted with a variable x_s , \mathbf{S}_k changes according to (11):

$$\begin{aligned}
\mathbf{S}_k^s &= (\mathbf{I}_N - \rho \mathbf{W})^{-1}(\mathbf{I}_N \beta_k + \mathbf{W} \theta_k + \mathbf{W} \odot (\mathbf{1}_N x'_{ks} \theta_{ks}))(\mathbf{I}_N - \rho \mathbf{W})^{-1} \odot \\
&\quad (\mathbf{1}_N x'_{ks} \beta_{ks}),
\end{aligned} \tag{11}$$

where \odot denotes the Hadamard operator and x_{ks} the resulting variable from the interaction. Again, the impact estimates are averaged using the corresponding PMPs from the first step as weighting factors.

For our estimations we use data from Crespo-Cuaresma and Feldkircher (2012). This enables us to directly compare our estimates with those obtained in their similar setting and enrich their conclusions with impacts including spatial spillovers. The dataset includes 56 variables for 255 European regions on NUTS-2 level for all EU-27 member states.⁸ Our dependent variable is defined as the average annual growth rate of real GRP per capita for the period 1995-2005. The explanatory variables can roughly be categorized into six different groups of potential growth determinants, with variables capturing physical factor accumulation and initial GRP per capita, human capital, technological innovation, sectoral structure and employment, infrastructure and socio-geographical variables. Where possible the explanatory variables are measured at the year 1995. Otherwise data for the earliest available year after 1995 is used. Since we are concerned with income convergence in European regions, the explanatory variable of particular interest is the initial GRP per capita (*GRPCAP0*).

⁸A list of all variables including a brief description and data sources as well as an overview of the European regions in the dataset can be found in the data appendix.

The set of possible spatial structures \mathbf{W} consists of a large variety of different specifications. We construct a total of 39 different \mathbf{W} matrices which can be categorized into four different groups: (i) Queen matrices \mathbf{W}_ϕ^Q with $\phi = [1, \dots, 10]$, which reflects queen neighborhood of order ϕ , (ii) k-nearest neighbors \mathbf{W}_ϕ^{K-NN} based on $\phi = [4, 5, 6, 8, 9, 10, 12, 15, 20, 25]$ nearest neighbors, (iii) distance band matrices \mathbf{W}_ϕ^B with $\phi = [50, 100, 150, 200, 400, 500, 600, 700, 800, 1000]$ where a region is considered a neighbor if it lies within a band of ϕ kilometers, and (iv) inverse distance matrices with the typical element $[\mathbf{W}_\phi^e]_{ii} = 0$ and $[\mathbf{W}_\phi^e]_{ij} = d_{ij}^{-\phi}$ for all $i \neq j$, where in this case d_{ij} represents the distance between region i and j and ϕ the decay parameter which takes the values $\phi = [1, \dots, 9]$. A higher decay parameter ϕ causes a more rapid decay of spatial weight.

6 Empirical results

The results of the first stage, i.e. spatial filtering BMA, are overall in line with the results of Crespo-Cuaresma and Feldkircher (2012).⁹ We identify the initial GRP per capita (*GRPCAPO*), the capital city dummy (*Capital*), indicating if a region contains the respective country's capital city, and the share of the low-skilled labor force (*ShSL*) to be robust determinants of the income growth pattern in European regions, where the latter is in contrast to Crespo-Cuaresma and Feldkircher (2012) who find the share of the high-skilled labor force (*ShSH*) to be of predominant importance. However, as the sign of our posterior mean for *ShSL* is negative, the interpretation of the role of education remains similar. Quadratic terms of both human capital variables appear to have low PIPs, hence we do not find support for potential non-linear relationships between human capital and economic growth. The interaction of *Capital* with *ShSL* suggests that regions containing the respective country's capital city are more adversely affected by the share of low-skilled labor than other regions. The interaction term between *GRPCAPO* and *Capital*, however, indicates that regions containing a country's capital city converge, on average, faster to the steady state than other regions.

The interpretation of the results from the first stage is, however, problematic as the parameter estimate for *GRPCAPO* only captures the spillover-free part of convergence and hence may underestimate (or overestimate) the magnitude of the 'real' convergence effects depending on the form of spatial spillovers. Crespo-Cuaresma and Feldkircher (2012) argue, in light of their results, that the estimate of the convergence parameter is reduced in magnitude once uncertainty about the spatial structure is accounted for. When not accounting for spatial

⁹ Results for the first stage are available from the authors upon request.

spillovers, they are absorbed by the convergence coefficient, whereas in the spatial filtering case the spillovers are integrated out of the coefficient. Therefore, they interpret the difference between the coefficients as evidence for the presence of positive spatial growth spillovers. However, since filtering techniques treat spatial effects as a mere nuisance, Crespo-Cuaresma and Feldkircher (2012) are not able to retrieve an estimate for the spatial parameter ρ , which is necessary for explicitly investigating the workings of spatial spillovers, as defined in (10) and (11), respectively. Therefore, the role of spatial spillovers remains elusive. Our re-estimation of the 10,000 top models draws a more detailed picture of the role of growth spillovers in the process of income convergence.

Table 1 summarizes the posterior means of the SAR and SDM specifications using different classes of spatial weight matrices. The first column depicts the *full model* considering all 39 spatial weight matrices. The columns two to five contain the parameter averages for the corresponding classes of weight matrices. We only observe minor changes between the results of the first and the second stage, suggesting that the 10,000 top models are an appropriate subset and therefore proxy for the whole model space.

The role of growth spillovers in the convergence process can be assessed by comparing the posterior means of the full model reported in the first column in Table 1 and the averaged impact estimates in Table 2 (refer also to Figure 1 and Figure 2). For *GRPCAP0*, the direct and total effects in the SDM setting are -0.015 and -0.020, respectively (the difference of -0.004 being the indirect effects), compared to -0.014 for the posterior mean. Direct effects can be interpreted as the average impact on income growth of a typical region resulting from a change in the typical region's initial GRP per capita, including all feedback effects (i.e. growth spillovers) generated by this change. The total effect measures the average impact on a typical region's income growth emanating from a change in *GRPCAP0* in all regions simultaneously. The difference between the *GRPCAP0*'s posterior mean and its direct effect signals only small growth spillovers experienced by a typical region occurring from changes in that region. However, as the total effects suggest, spillovers are not of negligible size, if all regions are considered to converge. This conclusion is therefore in line with the one drawn by Crespo-Cuaresma and Feldkircher (2012). Our results are, however, based on an impact assessment that explicitly deals with spillovers in the income convergence process.

	$W_{\phi}^Q, W_{\phi}^{K-NN}, W_{\phi}^B, W_{\phi}^e$			W_{ϕ}^Q			W_{ϕ}^{K-NN}			W_{ϕ}^B			W_{ϕ}^e		
	PIP	PM SAR	PM SDM	PIP	PM SAR	PM SDM	PIP	PM SAR	PM SDM	PIP	PM SAR	PM SDM	PIP	PM SAR	PM SDM
Capital	1.000	0.077	0.076	1.000	0.072	0.072	1.000	0.137	0.139	1.000	0.140	0.147	1.000	0.042	0.040
GRPCAP0	0.986	-0.014	-0.015	0.985	-0.014	-0.015	1.000	-0.016	-0.017	0.999	-0.016	-0.018	0.998	-0.015	-0.013
ShSL	0.675	-0.009	-0.013	0.715	-0.009	-0.014	0.074	-0.001	-0.001	0.037	0.000	0.000	0.783	-0.003	-0.010
AccessAir	0.454	0.004	0.005	0.458	0.004	0.005	0.290	0.002	0.002	0.895	0.006	0.008	0.291	0.001	0.002
TELF	0.352	-0.001	-0.001	0.378	-0.001	-0.001	0.002	0.000	0.000	0.012	0.000	0.000	0.041	0.000	0.000
AccessRoad	0.219	-0.002	-0.002	0.236	-0.002	-0.002	0.003	0.000	0.000	0.005	0.000	0.000	0.005	0.000	0.000
ShSH	0.201	0.008	0.007	0.211	0.009	0.007	0.068	0.003	0.002	0.019	0.001	0.001	0.300	0.011	0.009
AirportDens	0.110	0.829	0.789	0.111	0.842	0.802	0.027	0.166	0.165	0.510	4.226	3.873	0.023	0.128	0.114
POPDENS0	0.098	-0.001	-0.001	0.098	-0.001	-0.001	0.014	0.000	0.000	0.525	-0.007	-0.010	0.003	0.000	0.000
EMPDENS0	0.086	0.001	0.002	0.086	0.001	0.002	0.014	0.000	0.000	0.486	0.009	0.013	0.003	0.000	0.000
Distde71	0.081	0.000	0.000	0.004	0.000	0.000	0.002	0.000	0.000	0.005	0.000	0.000	0.002	0.000	0.000
RegPent27	0.062	0.000	0.000	0.067	0.000	0.000	0.003	0.000	0.000	0.005	0.000	0.000	0.002	0.000	0.000
RoadDens	0.047	0.001	0.001	0.051	0.001	0.001	0.006	0.000	0.000	0.013	0.000	0.000	0.004	0.000	0.000
INTF	0.024	0.001	0.001	0.018	0.001	0.000	0.113	0.003	0.002	0.008	0.000	0.000	0.143	0.005	0.005
URTO	0.021	-0.001	-0.001	0.023	-0.001	-0.001	0.002	0.000	0.000	0.003	0.000	0.000	0.025	-0.002	-0.001
ConnectAir	0.019	0.000	0.000	0.019	0.000	0.000	0.029	0.000	0.000	0.008	0.000	0.000	0.011	0.000	0.000
PatentHT	0.017	0.001	0.001	0.018	0.001	0.001	0.003	0.000	0.000	0.007	0.000	0.000	0.003	0.000	0.000
ShLLL	0.015	0.000	0.000	0.016	0.000	0.000	0.002	0.000	0.000	0.003	0.000	0.000	0.002	0.000	0.000
PatentICT	0.015	0.001	0.000	0.016	0.001	0.000	0.003	0.000	0.000	0.007	0.000	0.000	0.003	0.000	0.000
OUTDENS0	0.015	0.000	0.000	0.016	0.000	0.000	0.003	0.000	0.000	0.044	0.000	0.000	0.003	0.000	0.000
URL0	0.014	0.000	0.000	0.015	0.000	0.000	0.002	0.000	0.000	0.003	0.000	0.000	0.014	0.000	0.000
shGFCF	0.014	0.000	0.000	0.007	0.000	0.000	0.014	0.000	0.000	0.122	0.000	0.002	0.562	0.003	0.008
ERETO	0.013	0.000	0.000	0.014	0.000	0.000	0.003	0.000	0.000	0.003	0.000	0.000	0.004	0.000	0.000
ARH0	0.008	0.000	0.000	0.009	0.000	0.000	0.002	0.000	0.000	0.003	0.000	0.000	0.006	0.000	0.000
ARL0	0.008	0.000	0.000	0.006	0.000	0.000	0.055	0.000	0.000	0.015	0.000	0.000	0.003	0.000	0.000
PatentT	0.007	0.000	0.000	0.007	0.000	0.000	0.002	0.000	0.000	0.003	0.000	0.000	0.003	0.000	0.000
PatentBIO	0.006	0.001	0.001	0.006	0.001	0.001	0.005	0.001	0.001	0.006	0.001	0.001	0.004	0.000	0.000
RegCoast	0.006	0.000	0.000	0.006	0.000	0.000	0.005	0.000	0.000	0.008	0.000	0.000	0.003	0.000	0.000
ShCE0	0.006	0.000	0.000	0.006	0.000	0.000	0.002	0.000	0.000	0.003	0.000	0.000	0.004	0.000	0.000
Airports	0.005	0.000	0.000	0.005	0.000	0.000	0.009	0.000	0.000	0.004	0.000	0.000	0.002	0.000	0.000
PatentShICT	0.005	0.000	0.000	0.005	0.000	0.000	0.002	0.000	0.000	0.005	0.000	0.000	0.002	0.000	0.000
EREL0	0.005	0.000	0.000	0.004	0.000	0.000	0.025	0.000	0.000	0.006	0.000	0.000	0.003	0.000	0.000
EREH0	0.004	0.000	0.000	0.005	0.000	0.000	0.003	0.000	0.000	0.003	0.000	0.000	0.005	0.000	0.000
RegObj1	0.004	0.000	0.000	0.005	0.000	0.000	0.002	0.000	0.000	0.018	0.000	0.000	0.003	0.000	0.000
DistCap	0.004	0.000	0.000	0.087	0.000	0.000	0.008	0.000	0.000	0.003	0.000	0.000	0.009	0.000	0.000
gPOP	0.004	0.001	0.001	0.004	0.001	0.001	0.002	0.000	0.000	0.003	0.001	0.000	0.004	0.001	0.000
URH0	0.004	0.000	0.000	0.003	0.000	0.000	0.009	0.000	0.000	0.004	0.000	0.000	0.091	0.003	0.002
HRSTcore	0.004	0.000	0.000	0.004	0.000	0.000	0.002	0.000	0.000	0.002	0.000	0.000	0.007	0.000	0.000
RailDens	0.004	0.000	0.000	0.003	0.000	0.000	0.006	0.000	0.000	0.005	0.000	0.000	0.003	0.000	0.000
PatentShHT	0.003	0.000	0.000	0.003	0.000	0.000	0.002	0.000	0.000	0.005	0.000	0.000	0.002	0.000	0.000
ART0	0.003	0.000	0.000	0.003	0.000	0.000	0.004	0.000	0.000	0.003	0.000	0.000	0.003	0.000	0.000
Temp	0.003	0.000	0.000	0.003	0.000	0.000	0.002	0.000	0.000	0.004	0.000	0.000	0.002	0.000	0.000
PatentShBIO	0.003	0.000	0.000	0.003	0.000	0.000	0.003	0.000	0.000	0.009	0.000	0.000	0.008	0.000	0.000
ShAB0	0.002	0.000	0.000	0.003	0.000	0.000	0.002	0.000	0.000	0.006	0.000	0.000	0.004	0.000	0.000
Seaports	0.002	0.000	0.000	0.002	0.000	0.000	0.002	0.000	0.000	0.003	0.000	0.000	0.002	0.000	0.000
ConnectSea	0.002	0.000	0.000	0.002	0.000	0.000	0.002	0.000	0.000	0.003	0.000	0.000	0.002	0.000	0.000
TELH	0.002	0.000	0.000	0.002	0.000	0.000	0.002	0.000	0.000	0.003	0.000	0.000	0.002	0.000	0.000
Hazard	0.002	0.000	0.000	0.002	0.000	0.000	0.002	0.000	0.000	0.003	0.000	0.000	0.003	0.000	0.000
Settl	0.002	0.000	0.000	0.002	0.000	0.000	0.002	0.000	0.000	0.002	0.000	0.000	0.005	0.000	0.000
RegBoarder	0.002	0.000	0.000	0.002	0.000	0.000	0.002	0.000	0.000	0.002	0.000	0.000	0.003	0.000	0.000
ShSH ²	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000
ShSL ²	0.017	0.000	-0.001	0.018	0.000	-0.001	0.000	0.000	0.000	0.001	0.000	0.000	0.002	0.000	0.000
Cap.xGRPCAP0	0.438	-0.005	-0.005	0.406	-0.005	-0.005	0.975	-0.012	-0.012	0.939	-0.012	-0.013	0.109	-0.001	-0.001
Capital x ShSH	0.004	0.000	0.000	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.038	-0.003	-0.003
CapitalxShSL	0.628	-0.028	-0.028	0.667	-0.029	-0.030	0.031	-0.001	-0.001	0.022	-0.001	-0.001	0.757	-0.034	-0.036
ρ	1.000	0.217	0.187	1.000	0.195	0.156	1.000	0.233	0.200	1.000	0.267	0.147	1.000	0.337	0.454

Table 1: SAR and SDM coefficients for all classes of W of 10,000 top models

	PIP	Direct	SAR Indirect	Total	Direct	SDM Indirect	Total
Capital	1.00	0.017	0.006	0.023	0.017	0.006	0.023
GRPCAP0	0.99	-0.015	-0.004	-0.019	-0.015	-0.005	-0.020
ShSL	0.68	-0.012	-0.002	-0.014	-0.015	0.009	-0.006
AccessAir	0.45	0.004	0.001	0.005	0.013	0.001	0.014
TELF	0.35	-0.001	0.000	-0.001	0.000	0.000	0.000
AccessRoad	0.22	-0.002	0.000	-0.002	-0.005	-0.001	-0.006
ShSH	0.20	0.008	0.002	0.010	0.007	-0.006	0.001
AirportDens	0.11	0.833	0.482	1.315	0.792	0.456	1.248
POPDENS0	0.10	-0.001	-0.001	-0.002	-0.004	-0.002	-0.006
EMPDENS0	0.09	0.001	0.001	0.002	0.006	0.004	0.009
Distde71	0.08	0.000	0.000	0.000	0.000	0.000	0.000
RegPent27	0.06	0.000	0.000	0.000	0.000	0.000	0.000
RoadDens	0.05	0.001	0.000	0.001	-0.002	0.000	-0.003
ConnectAir	0.02	0.000	0.000	0.000	0.000	0.000	0.000
INTF	0.02	0.001	0.000	0.001	0.001	0.001	0.002
PatentHT	0.02	0.001	0.000	0.001	-0.002	0.000	-0.003
PatentICT	0.02	0.001	0.000	0.001	-0.002	0.000	-0.003
ShLLL	0.02	0.000	0.000	0.001	0.001	0.000	0.001
URT0	0.02	-0.001	0.000	-0.001	-0.002	0.000	-0.002
Airports	0.01	0.000	0.000	0.000	0.000	0.000	0.000
ARH0	0.01	0.000	0.000	0.000	0.001	0.000	0.001
ARL0	0.01	0.000	0.000	0.000	0.000	0.000	0.000
EREH0	0.01	0.000	0.000	0.000	0.000	0.000	0.000
ERET0	0.01	0.000	0.000	0.000	-0.001	0.000	-0.002
gPOP	0.01	0.001	0.000	0.001	0.005	0.001	0.006
HRSTcore	0.01	0.000	0.000	0.000	0.000	0.000	0.000
OUTDENS0	0.01	0.000	0.000	0.000	0.000	0.000	0.000
PatentBIO	0.01	0.001	0.000	0.001	-0.001	0.000	-0.001
PatentShICT	0.01	0.000	0.000	0.000	0.000	0.000	0.000
PatentT	0.01	0.000	0.000	0.000	0.000	0.000	-0.001
RegCoast	0.01	0.000	0.000	0.000	0.000	0.000	0.000
ShCE0	0.01	0.000	0.000	0.000	-0.001	0.000	-0.001
shGFCF	0.01	0.000	0.000	0.000	0.000	0.000	0.000
URH0	0.01	0.000	0.000	0.000	0.000	0.000	-0.001
URL0	0.01	0.000	0.000	0.000	0.000	0.000	0.000
ART0	0.00	0.000	0.000	0.000	0.000	0.000	0.000
ConnectSea	0.00	0.000	0.000	0.000	0.000	0.000	0.000
DistCap	0.00	0.000	0.000	0.000	0.000	0.000	0.000
EREL0	0.00	0.000	0.000	0.000	0.000	0.000	0.000
Hazard	0.00	0.000	0.000	0.000	0.000	0.000	0.000
PatentShBIO	0.00	0.000	0.000	0.000	0.001	0.000	0.001
PatentShHT	0.00	0.000	0.000	0.000	0.000	0.000	0.000
RailDens	0.00	0.000	0.000	0.000	-0.001	0.000	-0.001
RegBoarder	0.00	0.000	0.000	0.000	0.000	0.000	0.000
RegObj1	0.00	0.000	0.000	0.000	0.000	0.000	0.000
Seaports	0.00	0.000	0.000	0.000	0.000	0.000	0.000
Settl	0.00	0.000	0.000	0.000	0.000	0.000	0.000
ShAB0	0.00	0.000	0.000	0.000	0.000	0.000	0.000
TELH	0.00	0.000	0.000	0.000	0.000	0.000	0.000
Temp	0.00	0.000	0.000	0.000	0.000	0.000	0.000

Table 2: Impact estimates of a SAR and SDM model using 10,000 top models and all spatial weight matrices.

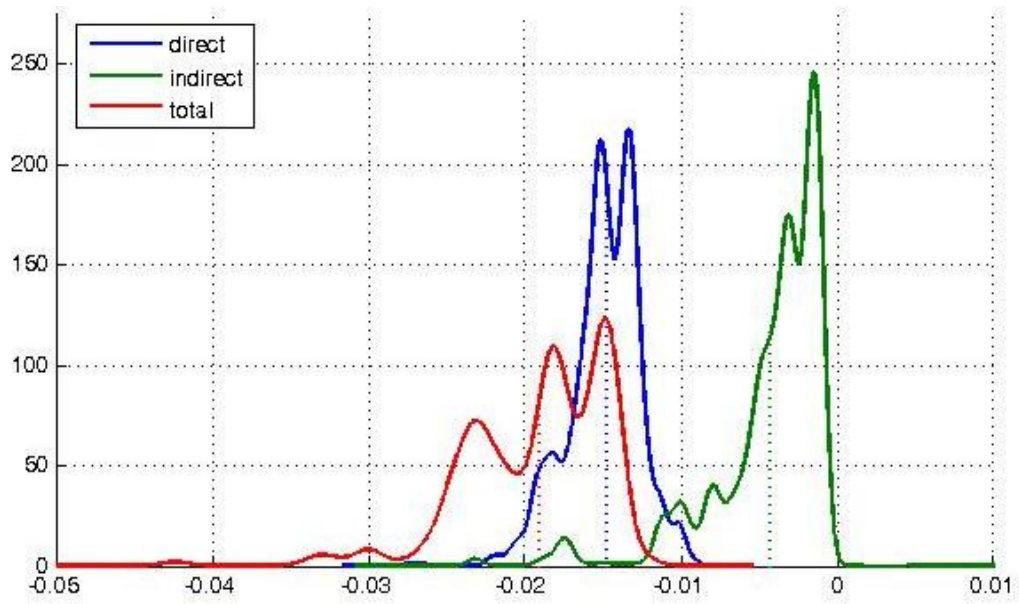


Figure 1: Conditional posterior distribution of impacts of *GRPCAP0*; SAR

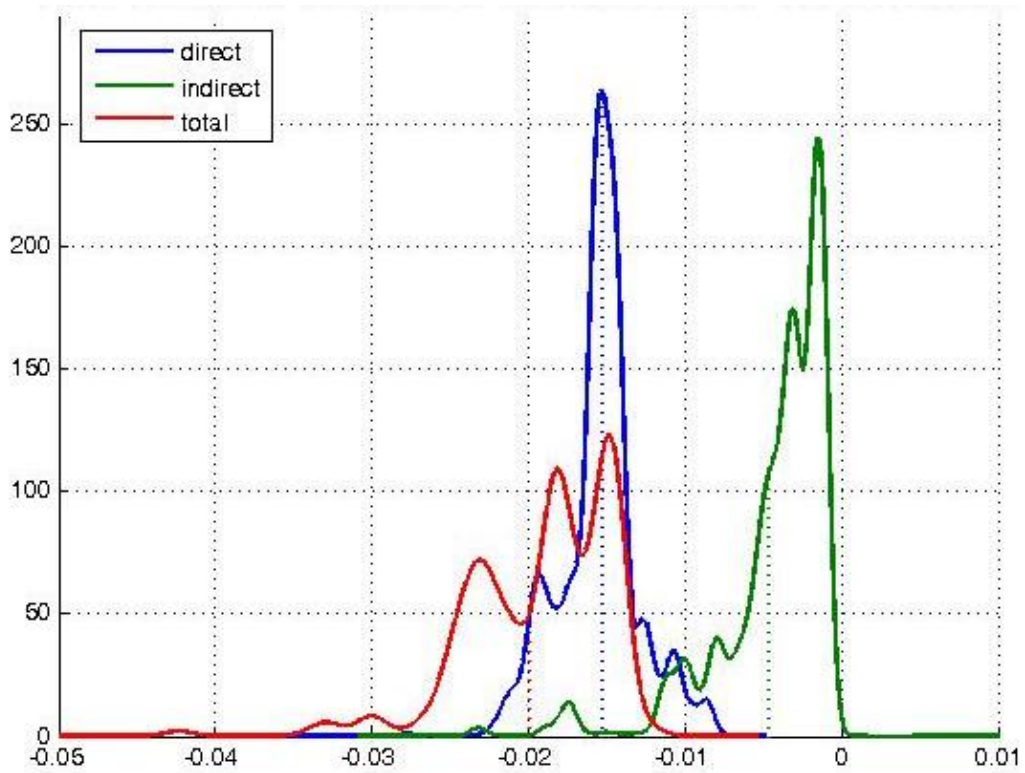


Figure 2: Conditional posterior distribution of impacts of *GRPCAP0*; SDM

7 Conclusions

For the purpose of studying income convergence across European regions, we propose an estimation procedure that allows a) for the consideration of many potential growth determinants as well as a multitude of spatial weight matrices and b) for a proper assessment of spatial effects, which results in a more detailed picture of the workings of spillovers in convergence processes.

In a first step, we rely on a combination of spatial filtering and BMA techniques as put forward by Crespo-Cuaresma and Feldkircher (2012). While model and parameter as well as spatial uncertainty are taken care of this way, a proper impact assessment is no longer possible, since spatial filtering annihilates all spatial information. The novelty of our approach lies in the second step, in which we re-estimate a number of top models from the spatial filtering BMA stage using SAR and SDM specifications and construct averages over the individual models' impact estimates using the respective PMPs as weighting factors.

Overall, our results confirm the findings of Crespo-Cuaresma and Feldkircher (2012). The inverse relationship between economic growth and initial GRP per capital implies that a convergence process across all European regions is observable. We further find that the presence of a country's capital city is an important driver of regional growth processes and that these capital city regions converge on average faster than other regions. Moreover, while a high share of low-skilled labor deters regional income growth, this is even more so for capital city regions.

The assessment of spatial effects, which our proposed approach allows for, underlines the important role of growth spillovers in the convergence process of European regions. Even though our findings are similar to the ones of Crespo-Cuaresma and Feldkircher (2012), our conclusions are based on a more detailed and explicit treatment of spillovers.

8 References

- Brock, W.A. and Durlauf, S.N. (2001), *Growth Empirics and Reality*, The World Bank Economic Review 15 (2), 229-272.
- Crespo-Cuaresma, J., Doppelhofer G., Feldkircher M. (2009), *The determinants of economic growth in European regions*, CESifo Working Paper 2519.
- Crespo-Cuaresma, J., Feldkircher, M. (2012), *Spatial Filtering, Model Uncertainty and the Speed of Income Convergence in Europe*, Journal of Applied Econometrics, forthcoming.
- Fernández, C., Ley, E., Steel, M. F. (2001), *Benchmark Priors for Bayesian Model Averaging*, Journal of Econometrics 100, 381-427.
- Fernández, C., Ley, E., Steel, M. F. (2001), *Model uncertainty in cross-country growth regressions*, Journal of Applied Econometrics 16, 563-576.
- Green, P. (1995), *Reversible jump Markov chain Monte Carlo computation and Bayesian model determination*, Biometrika 82 (4), 711-732.
- Hoeting, J., Madigan, D., Raftery, A., Volinsky, C. (1999), *Bayesian Model Averaging: A Tutorial*, Statistical Science 14 (4), 382-417.
- Koop, G. (2003), *Bayesian Econometrics*, John Wiley & Sons.
- LeSage, J.P., Fischer, M.M. (2008), *Spatial Growth Regressions: Model Specification, Estimation and Interpretation*, Spatial Economic Analysis 3 (3), 275-304.
- LeSage, J., Pace, K. (2009), *Introduction to Spatial Econometrics*, Statistics: Textbooks and Monographs.
- LeSage, J. P., Parent, O. (2007), *Bayesian Model Averaging for Spatial Econometric Models*, Geographical Analysis 39 (3), 241-267.
- Madigan, D., York, J. (1995), *Bayesian Graphical Models for Discrete Data*, International Statistical Review 63, 215-232.
- Pace, R., LeSage, J., Zhu, S. (2011), *Interpretation and Computation of Estimates from Regression Models using Spatial Filtering*, written for Spatial Econometrics Association 2011.
- Raftery, A. (1995), *Bayesian model selection in social research*, Sociological Methodology 25, 111-163.

Tiefelsdorf, M. and Griffith, D.A. (2007), *Semiparametric filtering of spatial autocorrelation: the eigenvector approach*, Environment and Planning A 39, 1193-1221.

Zellner, A. (1989), *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, North-Holland: Amsterdam.

9 Data Appendix

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

	Hamburg Hannover Karlsruhe Kassel Koblenz Köln	Stuttgart Thüringen Trier Tübingen Unterfranken Weser-Ems
Greece	AnatolikiMakedonia, Thraki Attiki DytikiEllada DytikiMakedonia IoniaNisia Ipeiros KentrikiMakedonia	Kriti NotioAigaio Peloponnisos StereiaEllada Thessalia VoreioAigaio
Hungary	Dél-Alföld Dél-Dunántúl Észak-Alföld Észak-Magyarország	Közép-Dunántúl Közép-Magyarország Nyugat-Dunántúl
Ireland	Border, Midlands and Western Southern and Eastern	
Italy	Abruzzo Basilicata Bolzano-Bozen Calabria Campania Emilia-Romagna Friuli-VeneziaGiulia Lazio Liguria Lombardia Marche	Molise Piemonte Puglia Sardegna Sicilia Toscana Trento Umbria Valle d'Aosta Veneto
Latvia	Latvia	
Lithuania	Lithuania	
Luxembourg	Luxembourg (Grand-Duché)	
Malta	Malta	
Netherlands	Drenthe Flevoland Friesland Gelderland Groningen Limburg (NL)	Noord-Brabant Noord-Holland Overijssel Utrecht Zeeland Zuid-Holland
Poland	Dolnoslaskie Kujawsko-Pomorskie Łódzkie Lubelskie Lubuskie Malopolskie Mazowieckie Opolskie	Podkarpackie Podlaskie Pomorskie Ślaskie Świętokrzyskie Warmińsko-Mazurskie Wielkopolskie Zachodniopomorskie
Portugal	Alentejo Algarve Centro (PT)	Lisboa Norte

Romania	Bucuresti – Ilfov Centru Nord-Est Nord-Vest	Sud-Muntenia Sud-Est Sud-VestOltenia Vest
Slovak Republic	Bratislavskýkraj StrednéSlovensko	VýchodnéSlovensko ZápadnéSlovensko
Slovenia	Slovenia	
Spain	Andalucia Aragón Cantabria Castilla y León Castilla-la Mancha Cataluña Comunidad de Madrid Comunidad Foral de Navarra	Comunidad Valenciana Extremadura Galicia Illes Balears La Rioja Pais Vasco Principado de Asturias Región de Murcia
Sweden	MellerstaNorrländ NorraMellansverige ÖstraMellansverige ÖvreNorrländ	Smålandmedöarna Stockholm Sydsverige Västsverige
United Kingdom	Bedfordshire, Hertfordshire Berkshire, Bucks and Oxfordshire Cheshire Cornwall and Isles of Scilly Cumbria Derbyshire and Nottinghamshire Devon Dorset and Somerset East Anglia East Riding and North Lincolnshire East Wales Eastern Scotland Essex Gloucestershire, Wiltshire and North Somerset Greater Manchester Hampshire and Isle of Wight Herefordshire, Worcestershire and Warks Inner London	Kent Lancashire Leicestershire, Rutland and Northants Lincolnshire Merseyside North Yorkshire Northern Ireland Northumberland, Tyne and Wear Outer London Shropshire and Staffordshire South Western Scotland South Yorkshire Surrey, East and West Sussex Tees Valley and Durham West Midlands West Wales and The Valleys West Yorkshire

Table 3: European Regions in the sample

Variable name	Description	Source
Dependent variable		
gGRPCAP	Growth rate of real GRP per capita	Eurostat
Factor accumulation/convergence		
GRPCAP0	Initial real GRP 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	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	ESPON
EMPDENS0	Initial employment density	
POPDENS0	Initial population density	
RegCoast	Coast	
RegBorder	Border	
RegPent27	Pentagon EU 27 plus 2	
RegObj1	Objective 1 regions	
Capital	Capital city	
Airports	Number of airports	
Temp	Extreme temperatures	
Hazard	Sum of all weighted hazard values	
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 4: Variables, description and sources