Spatial Statistics 6 (2013) 78–90

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

A Gaussian Markov random field approach to convergence analysis

L. Ippoliti [a,*], L. Romagnoli [b], G. Arbia [c]

[a] Department of Economic Studies, University G. d'Annunzio, Chieti-Pescara, Viale Pindaro 42, 65127 Pescara, Italy

[b] EGSI department, University of Molise, Campobasso, Italy

[c] Department of Statistical Science, Catholic University of the Sacred Heart, Rome, Italy

ARTICLE INFO

Article history:
Received 14 January 2013
Accepted 23 July 2013
Available online 6 August 2013

Keywords:
Gaussian Markov random fields
Convergence analysis
Simultaneous autoregressions
β-convergence model

ABSTRACT

Spatial models have been widely applied in the context of growth regressions with spatial spillovers usually modelled by simultaneous autoregressions (SAR). Although largely used, such a class of models present some logical difficulties connected with the error behaviour, the lack of identifiability of the model parameters and their substantive interpretation. To overcome these logical pitfalls, in this paper we introduce a new specification of regional growth regressions by applying multivariate Gaussian Markov random fields (GMRFs). We discuss the theoretical properties of the proposed model and show some empirical results on the economic growth pattern of 254 NUTS-2 European regions in the period 1992–2006. We show that the proposed GMRF model is able to capture the complexity of the phenomenon including the possibility of estimating site-specific convergence parameters which may highlight clustering of regions and spatial heterogeneities in the speed of convergence.

1. Introduction

In recent years there has been an increasing awareness on problems related to the economic growth and on the conditions under which the per-capita income levels of European countries tend to converge over time towards a common level. Many empirical works focus on estimating the income

* Corresponding author. Tel.: +39 0854537531; fax: +39 0854537542.
E-mail addresses: ippoliti@unich.it (L. Ippoliti), luca.romagnoli@unimol.it (L. Romagnoli), giuseppe.arbia@rm.unicatt.it (G. Arbia).

2211-6753/$ – see front matter
http://dx.doi.org/10.1016/j.spasta.2013.07.005

convergence speed using cross-sectional data, with a theoretical setting based on the Solow–Swan neoclassical growth model. The applied econometric literature has devoted growing attention on classical convergence and for a review see, for example, Temple (1999), Islam (2003), Abreu et al. (2005a), Rey and Janikas (2005) and Mathunjwa and Temple (2007). The book by Barro and Sala-i-Martin (1995) also provides an excellent starting point for researchers in this area.

Many authors have also approached the issue of income convergence using regional datasets (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. A key property of many economic data at the regional level is that observations at nearby sites tend to be similar to one another. In fact, there is strong evidence (see for example, 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. Hence, spatial interactions, such as technological spillovers or factor mobility, both being important forces for the process of convergence, need 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 unmodelled.

Most empirical studies in the spatial econometrics literature model spatial spillovers in the framework of simultaneous autoregressive (SAR) specifications (Anselin, 1988; Arbia, 2006; LeSage and Pace, 2009) conditional on a given spatial contiguity matrix which specifies the spatial interactions among the regions. Let $\xi(\mathbf{s}_i, t_0)$ and $\xi(\mathbf{s}_i, t_j)$ be the per-capita GDP variable observed at a specific region $\mathbf{s}_i$ and time points t_0 and t_j, respectively. Then, a widely used SAR specification of the Solow–Swan model, also known as the β-convergence model (Barro and Sala-i-Martin, 1995), is represented by a simple formulation of the spatial Durbin model (Anselin, 1988; LeSage and Pace, 2009)

$$y_i = \varphi + \beta x_i + \phi \sum_{j \in S} x_j + \rho \sum_{j \in S} y_j + u_i \tag{1}$$

where $y_i = \ln\left[\xi(\mathbf{s}_i, t_j)/\xi(\mathbf{s}_i, t_0)\right]$ and $x_i = \ln\left[\xi(\mathbf{s}_i, t_0)\right]$, S is the set of neighbours of site $\mathbf{s}_i$, ϕ and ρ are spatial dependence parameters and $u(\mathbf{s}_i)$ is a zero mean independent Gaussian error term with variance σ_u^2. Positive values of ϕ and ρ denote the existence of spillovers from neighbouring observations; in the case in which they are equal to zero, model (1) nests the classical β-convergence model with no spatial effects. The parameter φ is the intercept term while β is the regression coefficient associated to the initial income per-capita. Variants of model (1) are possible (see Anselin, 1988, Arbia, 2006, LeSage and Pace, 2009) where we can also let different sets of neighbours be associated with the endogenous and exogenous variables. Additional regressors can also be included and the 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.

Model (1), and in general the class of SAR models, forms an important cornerstone of the available models for irregular lattice data. They have an intuitive appeal and enjoy several advantages. For example, the mean of the process can be represented by a linear combination of the values observed within a suitably chosen finite set S; second, it can describe a wide variety of observed spatial correlation structures. Third, several different methods are now available for estimating the parameters of this model in a computationally effective fashion, and also, for validating this model by judging its adequacy for a set of data, especially by model criticism and diagnostic testing (Anselin, 1988; LeSage and Pace, 2009).

However, in two or higher dimensions, SAR models have some logical difficulties. A first pitfall is represented by the fact that the errors in the model are correlated with all the observations and cannot be regarded as innovations. Hence, the uncorrelated errors have no physical meaning and have no link with the process y. The class of SAR processes is not theoretically underpinned by the Wiener–Kolmogorov prediction theory (Whittle, 1963) and, accordingly, as with simultaneous