Estimation strategies for spatial dynamic panel using GMM. A new approach to the convergence issue of European regions

Salima Bouayad Agha, Lionel Vedrine

Estimation strategies for Spatial Dynamic Panel using GMM. A new approach to the convergence issue of European regions[1]

Salima Bouayad-Agha[1], Lionel Vedrine[2]

[1] TEPP – GAINS, Université du Maine, Avenue Olivier Messiaen 72085 Le Mans Cedex 9, salima.bouayad-agha@univ-lemans.fr

[2] Cemagref, UMR 1273 Métafort, 24 avenue des Landais, 63172 Aubière Cedex, France, Lionel.vedrine@cemagref.fr

No senior authorship involved.

Abstract:

The recent progress in spatial econometrics offers a number of estimators for models that treat spatial dependence explicitly but techniques for handling spatial dependence appear to be essentially confined to cross-sectional studies. Despite the fact that dynamic panel models have been the object of recent important developments, the econometric analysis of spatial and dynamic panel models is at an early stage.

In this context, the aim of this paper is to review the different methods used in the literature, for such space-time data sets, and to suggest two strategies in order to estimate spatial dynamic panel using GMM. The first is to extend the moment restrictions of Arellano and Bond estimator for spatial autoregressive dynamic panel. The second allows for spatial dependence in the error process in calculating the optimal weighting matrix at the second step of the Arellano and Bond estimator.

These strategies are carried out to estimate the convergence of European regions during the last 25 years.

Keywords: Spatial Econometrics, Dynamic Panel Model, GMM, Regional Convergence

JEL Classification: C21, C23, O52, R11, R15

---

[1] The authors thank Nadine Turpin for useful comments.

1. Introduction

While econometric analysis of dynamic panel models is now fairly standard (Arellano and Bond, 1991; Arellano and Bover, 1995; Blundell and Bond, 1998) and spatial econometric literature is well documented either on cross-sections (Anselin and Kelejian, 1997; Anselin, 2001) or on static panel (Elhorst, 2003), econometric analysis of spatial and dynamic panel models is still at an early stage of development. Modelling space-time data is quite complex and several problems must be tackled: obviously, leaving correlation on each spatial unit over time and spatial dependence on the spatial units at each date aside would lead to misspecification but the main difficulty is that we have *a priori* no reason to believe that one problem is more important than the other is.

Empirically, there are several examples where the presence of a dynamic process with spatial dependence might occur. This is the case with the regional economic convergence/divergence issue. Many convergence studies stems from neoclassical growth model (Solow, 1956; Swan, 1956) and most of them rely on a beta convergence model: regions (countries) do not have the same structural characteristics and thus converge towards different steady states relative per capita incomes. Factor mobility, trade relations and geographical spillovers (like technology spreading) can matter to understand how economic development of a region is likely to be influenced by neighbouring regions. Various recent convergence studies have found evidence for model misspecifications if spatial interdependencies of regional growth are ignored (Arbia *et al.*, 2008). Working on a dynamic panel specification, Badinger *et al.* (2004) applied a GMM estimator to spatially filtered variables; Elhorst (2005) suggests a maximum likelihood estimation of models that are dynamic both in space and time for regional analysis; Piras and Arbia (2007) extend panel-data models with spatial error autocorrelation for a convergence analysis of EU regions. More precisely the main argument of applying the extended GMM in a spatial context is that it corrects for the endogeneity of the spatial lagged dependent variable and other potentially endogenous explanatory variables.

In this context, the aim of this paper is to review the different methods used in the literature, for such space-time data sets, and to suggest an estimation strategy that consider together the dynamic specification and the spatial dimension of the panel with an illustration on regional conditional convergence framework. Estimating a per capita GDP conditional convergence model that

incorporates an explicit consideration of spatial dependence effects, we rely on dynamic panel generalised method of moment estimations that control for endogeneity, variable omission and spatial dependence problems.

The reminder of the paper is organised as follows. In the next section, we address the econometric issues on spatial dynamic panel model and develop an empirical estimation strategy. Section 3 reports the main results of our empirical estimation strategy for the convergence approach on panel data. Section 4 concludes and suggests areas for further research.

2. Econometric issues on spatial dynamic panel data model

2.1. Dynamic panel model

A general dynamic panel model can be described as follows:

$$(1) \qquad y_{it} = \alpha y_{i,t-1} + \left(\eta_i + \nu_{i,t}\right) \qquad |\alpha| < 1 \qquad i = 1,\ldots,N; \qquad t = 2,\ldots,T$$

where $y_{i,t}$ is an observation on some series for individual *i* in period t, $y_{i,t-1}$ is the observation on the same series for the same individual in the previous period and $\eta_i + \nu_{i,t}$ is the usual "error components" decomposition of the error term which allows for unobserved heterogeneity (ε_{it} hereafter). These specific effects are supposed to be stochastic, which here implies that they are necessarily correlated with the lagged dependent variable. With the additional assumption that $\nu_{i,t}$ is serially uncorrelated, the Ordinary Least Squares (OLS) estimator of α in the levels equations is inconsistent. Standard results for omitted variable bias indicate that, at least in large samples, the OLS levels estimator is biased upwards. The Within estimator for panel data that allows to control for time invariant characteristics lead to substantial gains in robustness (compared with cross-section estimator) but it is not without costs if one do not pay attention to the dynamic of adjustment. Standard results for omitted variable bias indicate that, at least in large samples, the Within Group estimator is biased downward (Nickell, 1981). Thus, we might hope that a candidate consistent estimator will lie between the OLS and Within estimates. The most widely-used alternative strategy is to difference the model to eliminate the fixed effects and then applies the GMM using a set of appropriate instruments to address

the correlation between the differenced lagged dependent variable and the induced MA(1) error term (see Arellano and Bond, 1991). The GMM estimator is a two-step estimator. In the first step, an initial positive semi definite weight matrix is used to obtain consistent estimates of the parameters. Given this consistent estimates, a weight matrix can be constructed that is consistent for the efficient weight matrix, which is used for the asymptotically efficient two-step estimates. It is common practice to use the inverse of the moment matrix of the instruments as an initial weight matrix. More precisely the first step in the estimation procedure consists in eliminating the individual effects via a first difference transformation (GMM-DIFF) or a forward orthogonal deviation (Arellano and Bover, 1995).

The dynamic panel-data specification has become frequent in growth convergence empirical studies. Since the inclusion of the time-lagged dependent variable in the equation might lead to inconsistent estimates, instrumental variable estimators are needed. A commonly employed estimation procedure to estimate the parameters in a dynamic panel data model with unobserved individual specific heterogeneity is to transform the model into first differences. Sequential moment conditions are then used where lagged levels of the variables are instruments for the endogenous differences and the parameters estimated by GMM (see Arellano and Bond, 1991; adopted by Caselli *et al.*, 1996 in the growth context). The first step in the estimation procedure consists in eliminating the individual effects via a first difference transformation (GMM-DIFF) or a forward orthogonal deviation[2] (as suggested by Arellano and Bover, 1995). Assuming that the error terms ε_{it} are serially uncorrelated, the lagged difference of the endogenous variable is instrumented with the lagged difference of the endogenous variable (Δy_{it-1}) with all lagged levels of the variable y_{it-1} starting with lag two y_{it-2} and going back to earlier lagged levels. Essentially these use an instrument matrix of the form:

$$Z_i = \begin{bmatrix} y_{i1} & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & y_{i1} & y_{i2} & \cdots & 0 & \cdots & 0 \\ . & . & . & \cdots & . & \cdots & . \\ 0 & 0 & 0 & \cdots & y_{i1} & \cdots & y_{iT-2} \end{bmatrix}$$

[2] First differencing and forward orthogonal deviation transformation involve the same procedure (similar instrument matrix) to estimate dynamic panel data specification.

where rows correspond to the first-differenced equations for period t=3,4,…,T for individual i and exploits the moments conditions $E\left(y_{is}\Delta\varepsilon_{it}\right)=0$ for for s=1,…,T-2 and t=2,…,T.

The overall validity of instruments can be checked by a Sargan-Hansen test of over-identifying restrictions.

This GMM estimator extends in a natural way to autoregressive-distributed lag models of the form:

$$(2) \qquad y_{it}=\alpha y_{i,t-1}+\beta x_{i,t}+\left(\eta_i+\nu_{i,t}\right) \quad |\alpha|<1 \quad i=1,\ldots,N; \qquad t=2,\ldots,T$$

where x_{it} can be a vector of current and lagged values of additional explanatory variables. Different moment conditions will be available depending on what is assumed about the correlation between x_{it} and the two components of the error term (Bond, 2002):

- If x_{it} is strictly exogenous, $E\left(x_{is}\Delta\varepsilon_{it}\right)=0$ for s=1,…,T and t=2,…,T.
- If x_{it} is weakly endogenous, $E\left(x_{is}\Delta\varepsilon_{it}\right)=0$ for s=1,…,T-1 and t=2,…,T.
- If x_{it} is strictly endogenous, $E\left(x_{is}\Delta\varepsilon_{it}\right)=0$ for s=1,…,T-2 and t=2,…,T.

Unfortunately in the case of persistent data and for a small number of time series observations, lagged levels are only weak instruments for subsequent first differences and the GMM-DIFF can have very poor finite sample properties in terms of bias, and precision. Blundell and Bond (1998) proposed the system GMM estimator (GMM-SYS) that combines moment conditions for the model in first differences with moment conditions for the model in levels. It has been shown to correspond to the model in levels, with lagged differences of the endogenous variables as instruments. Blundell and Bond (1998) argued that the GMM-SYS estimator performs better than the GMM-DIF estimator because the instruments in the model remain good predictors for the endogenous variables in this model even when the series are very persistent.

The choice between the two estimators (GMM-DIFF, GMM-SYS) is conducted according to a common test of over identifying restrictions: the validity of the additional instruments used by GMM-SYS for the level equation can easily be tested using difference Hansen tests.

2.2. Spatial dynamic panel model

Spatial econometric methods deal with the incorporation of spatial interaction and spatial structure into regression analysis. Spatial error model and spatial lag model are two different approaches to address the issue of spatial dependencies (Anselin, 2001). The first one is a nuisance form of spatial dependence and incorporates a spatial autoregressive process in the error term (as in time series analysis). Ignoring this spatial autocorrelation may lead to inefficient estimates. The second method incorporates spatial dependence in the form of a spatial lag variable, often considered as a spatial autoregression model. This specification is more appropriate for explicitly specifying as completely as possible the impact of nearby observations on the dependent variable so that explicit inferences can be made for spatially lagged variables (Beck *et al*. 2006; Blonigen *et al*., 2007).

The design of these two specifications relies on a spatial weight matrix W describing the spatial arrangement of the cross-section units. As it is standard in spatial econometrics, for ease of interpretation, this weighting matrix W is row-standardized so that each row in W sums to one. Throughout the rest of the paper, we assume that W remains constant over time.

First, we consider a "time-space simultaneous" model (Anselin *et al*., 2007):

$$(3) \qquad y_{it} = \alpha y_{i,t-1} + \beta x_{i,t} + \rho W y_{i,t} + \left(\eta_i + \nu_{i,t}\right) \quad |\alpha| < 1 \quad i = 1,\ldots,N; \qquad t = 2,\ldots,T$$

where $Wy_{i,t}$ is a first order spatial lag. The coefficient ρ stands for the intensity of spatial effects. Thus when ρ =0, the model remains as section 2.1.

According to Anselin (2001) and Abreu *et al.* (2005), including a spatially lagged dependent variable causes simultaneity and endogeneity problems, which in turn means that this variable must be treated as endogenous and thus proper estimation methods must account for this endogeneity. There are only a limited number of available estimators for Dynamic Spatial lag Model on Panel Data. Assuming all explanatory variables are exogenous beside the spatial lag term[3], dynamic spatial lag models are usually estimated using the GMM estimator (see for example, Madriaga and Poncet, 2007). Spatial lag

[3] As underlined by Kukenova and Monteiro (2008) there is no currently available estimator to consider this simultaneity problem in line with the potential endogeneity of other explanatory variables.

term and autoregressive term are instrumented by their lagged values, by lagged values of the dependent variables as well as by spatially weighted explanatory variables.

If the spatial lag is strictly endogenous, the moment restrictions of section 2.1 are not sufficient to provide an unbiased and consistent estimation. However, a natural way to estimate (3) is to assume further moment restrictions of 2.1 by considering $Wy_{i,t}$ as endogenous variable:

$E\left(Wy_{is}\Delta\varepsilon_{it}\right)=0$ for t=2,…,T and s=1,…,T-2.

Moreover, we can use spatially weighted explanatory variables $Wx_{i,t}$ to instrument the spatial lag term. In other terms, we identify the exogenous part of the spatial lag variability by a spatially weighted model.

The validity of this strategy requires the following moment restrictions:

If Wx_{it} is strictly exogenous, $E\left(Wx_{it}\Delta\varepsilon_{it}\right)=0$ for t=2,…,T.

We suggest checking the robustness of this extended GMM estimator by proceeding successively to cross sections, Least Squares Dummy Variables (LSDV), GMM dynamic panel and GMM spatial dynamic panel estimation.

Spatial error model in dynamic panel context is more complicated to estimate than spatial lag model. Nevertheless, Elhorst (2005) propose maximum likelihood estimation for dynamic panel including spatial error autocorrelation. In a static panel model, a consistent estimator can be obtained from a set of moment conditions as demonstrated in Kapoor *et al.* (2007). Mutl (2006) extends this approach to a dynamic panel case.

We consider the following specification:

(4) $$y_{it}=\alpha y_{i,t-1}+\beta x_{i,t}+\varepsilon_{i,t} \quad |\alpha|<1 \quad i=1,\ldots,N; \qquad t=2,\ldots,T$$

$$\varepsilon_{it}=\rho W\varepsilon_{i,t}+\left(\eta_i+\nu_{i,t}\right) \quad |\rho|<1$$

where the error term consists of the spatially lag simultaneous error and the previous error component $(\eta_i + \nu_{i,t})$.

We assume that is $\nu_{\mathrm{i,t}}$ independently distributed, with a constant variance σ_ν^2 .

We estimate a GMM estimator using the same moment as the Arellano Bond estimator but with optimal weight matrix that allows for spatial error autocorrelation. According to Mutl (2006), this weighting matrix is constructed using first step estimation of σ_ν^2, σ_1^2 and ρ based on the Kapoor *et al* (2007) spatial GM estimator[4].

3. Empirical application: β-convergence of European regions

3.1. Modelling β-convergence on spatial dynamic panel specification

Many convergence studies stems from traditional neoclassical growth model (Solow, 1956; Swan, 1956). Following the "model based" specification of Mankiw *et al.*, (1992) Barro and Sala-i-Martin (1992), we rely on a β-convergence model where the GDP per capita (hereafter GDP p.c.) growth depends not only on the initial GDP level, but also on other conditioning variables. Regions do not have the same structural characteristics and thus converge towards different steady–state income levels. The further a region finds itself from its own steady state, the faster its growth rate will be. In this case, convergence is *conditional*: economies converge towards the same growth rate. Accordingly, the following general model is in line with the empirical growth literature:

$$\ln\left(\frac{y_{i,t}}{y_{i,0}}\right) = \beta_0 + \beta_1 \ln y_{i,0} + \sum_k \beta_k X_{k,i,t} + \sum_g \pi_g Z_{g,i,t} + \varepsilon_{i,t} \quad (5)$$

[4] Mutl (2006) demonstrates the consistency of the spatial GM in a dynamic panel data model.

where $y_{i,t}(i=1,\ldots,n;t=1\ldots,T)$ is the GDP p.c. of region i at time t=; β_0 is an intercept term, β_1 is the convergence coefficient, X is a set of *k* explanatory variables related to the growth model of Solow (Solow, 1956) and Z is a set of variables that may affect the convergence process but are not directly related to the model of Solow (1956) (Durlauf *et al.*, 2006).

We assume that the variables sets in Z_i and X_i are independent (see Mankiw *et al.*, 1992 for a discussion on this point). It is now possible to extend Mankiw *et al.* (1992) model towards a dynamic panel-data specification of the convergence model, current output is regressed on lagged output and control variables (Durlauf et al., 2006):

$$(6) \qquad \ln\left(\frac{Y_{i,t}}{pop_{i,t}}\right)=(1+\beta_1)\ln\left(\frac{Y_{i,t-1}}{pop_{i,t-1}}\right)+\beta_2\ln\left(\frac{I_{i,t}}{pop_{i,t}}\right)+\beta_3\ln\left(\frac{pop_{i,t}}{pop_{i,t-1}}\right)+\alpha_i+\mu_t+\varepsilon_{i,t}$$

where $\frac{Y_{i,t}}{pop_{i,t}}$, $\frac{I_{i,t}}{pop_{i,t}}$ are respectively the gross domestic product and the investment per capita and $\ln\left(\frac{pop_{i,t}}{pop_{i,t-1}}\right)$ is the demographic growth rate. We introduce individual and time specific intercepts (respectively α_i and μ_t) in order to control for unobserved heterogeneity. Thereby, β_1 measures the GDP convergence conditionally of investment per capita and population growth rate.

However, for both econometrics and economic theory, it is too restrictive to assume no cross-sectional dependence in testing convergence hypothesis. The restriction implies that the economies are closed, which is obviously inappropriate to understand regional convergence process. Regions are units naturally open to economic flows like trade, technology diffusion and factor mobility, which may affect regional convergence process.

Following the model specification of (3) to address with spatial dependence, we estimate successively:

$$(7) \qquad \ln\left(\frac{Y_{i,t}}{pop_{i,t}}\right)=(1+\beta_1)\ln\left(\frac{Y_{i,t-1}}{pop_{i,t-1}}\right)+\rho W\ln\left(\frac{Y_{i,t}}{pop_{i,t}}\right)+\beta_2\ln\left(\frac{I_{i,t}}{pop_{i,t}}\right)+\beta_3\ln\left(\frac{pop_{i,t}}{pop_{i,t-1}}\right)+\alpha_i+\mu_t+\varepsilon_{i,t}$$

and

$$(8)\qquad \ln\left(\frac{Y_{i,t}}{pop_{i,t}}\right) = (1+\beta_1)\ln\left(\frac{Y_{i,t-1}}{pop_{i,t-1}}\right) + \beta_2 \ln\left(\frac{I_{i,t}}{pop_{i,t}}\right) + \beta_3 \ln\left(\frac{pop_{i,t}}{pop_{i,t-1}}\right) + \alpha_i + \mu_t + \varepsilon_{i,t}$$

$$\varepsilon_{it} = \rho W \varepsilon_{i,t} + (\eta_i + \nu_{i,t})$$

3.2. Data description

We use a panel dataset of 191 regions in 14 member states of EU-15 (see appendix A, which describes the set of regions, included and excluded in the sample) over a period to 25 years (1980-2005).

Data variables come from the Cambridge Econometrics database[5]. The gross domestic product (GDP) and investment (provided by Cambridge Econometrics in 1995 constant euro) have been transformed into logarithm of per capita term ($\ln\left(\frac{Y_{i,t}}{pop_{i,t}}\right)$, $\ln\left(\frac{I_{i,t}}{pop_{i,t}}\right)$) in order to consider the scale effect. The demographic growth rate is measured from the total population data dynamics ($\ln\left(\frac{pop_{i,t}}{pop_{i,t-1}}\right)$).

For the estimation, we consider five aggregated time-periods (1980-84, 1985-89, 1990-94, 1995-99 and 2000-2005) to avoid short run variations in GDP growth rates due to business-cycle effects. The accurate number of years required to avoid short-run variations is still under discussion in the literature (see Temple, 1999, for an analysis). Temple (1999) recommends 5 or 10 years long periods, but we preferred to follow the approach from Badinger *et al.* (2004) and chose quinquennial time periods to collect information on at least 5 periods. Thereby, we have a panel on 955 observations of 191 regions during 5 periods. Of course, the dynamic panel specification restricts this panel to 4 periods because of the autoregressive term ($\ln\left(\frac{Y_{i,t-1}}{pop_{i,t-1}}\right)$).

The spatial lag term $W \ln\left(\frac{Y_{i,t}}{pop_{i,t}}\right)$ is constructed from a weighted average of neighbouring regions' income level. Thus, we have chosen a geographical definition of neighbourhood based on the Euclidean distance between regions in order to construct the spatial weight matrix (W). More

[5] The Cambridge Econometrics database is available at http://www.camecon.com.

precisely we have chosen a k-nearest neighbours weight specification, $w_{ij}(k)$ represents the element of W matrix in row i and column j:

$w^*_{ij}(k) = 0$ if $i = j$

$w^*_{ij}(k) = 1$ if $d_{ij} \leq d_i(k)$

$w^*_{ij}(k) = 0$ if $d_{ij} \geq d_i(k)$

d_{ij} is the distance between the regions i and j centroids, and $d_i(k)$ is a cut-off distance based on the distance of k-nearest neighbour for region i. The interactions are assumed to be negligible above this distance. Although we have constructed W with k=10, the results are similar with k=5, 15 and 20.

So, the matrix is row-standardised $w_{ij}(k) = \frac{w^*_{ij}(k)}{\sum_j w^*_{ij}(k)}$ to provide easier interpretation (each weight may be interpreted as the region's share in the total spatial effect of the sample) and make parameter estimates more comparable (see Kelejian and Robinson, 1995 for a good discussion on the spatial weighting matrix).

k-nearest neighbours' weight matrix has the most fitted to representing spatial interaction of our sample: this specification lead to each region has the same number of neighbouring regions (k) including islands on our sample and reduce the heterogeneity problem of regional superficies (Anselin, 2002).

Table 1 *Descriptive statistics*

Variable	1980-84	1985-89	1990-94	1995-99	2000-04
$ln(\frac{Y_{i,t}}{pop_{i,t}})$					
Sample mean	9.44	9.54	9.63	9.71	9.81
Standard deviation	0.40	0.39	0.39	0.38	0.37
$ln(\frac{I_{i,t}}{pop_{i,t}})$					
Sample mean	7.82	7.95	7.99	8.08	8.21
Standard deviation	0.47	0.44	0.41	0.37	0.34
$ln(\frac{pop_{i,t}}{pop_{i,t-1}})$					
Sample mean		2.84	3.37	3.28	3.29
Standard deviation		1.39	1.19	1.24	1.30
$Wln(\frac{Y_{i,t}}{pop_{i,t}})$					
Sample mean	9.46	9.55	9.64	9.72	9.83
Standard deviation	0.34	0.34	0.34	0.34	0.32

Table 1 depicts the dynamics of GDP p.c., investment per capita, demographic growth rate and spatially lagged GDP for European regions. One can see the regular evolution of each variable. The sample mean of regional GDP pc increases from 9.44 to 9.81, what represents a growth of 3.92% during the whole period. In the same time, the decreasing of GDP p.c. standard deviation points out the reduction of GDP p.c. disparities across time. This evolution is confirmed by the analysis of the distribution of that variable (Figure 1).

Figure 2 graphs the GDP p.c. geographic pattern relative to the EU-14 average GDP level for the 5 periods. The regions are split into 6 classes, from below 50 % of the European average to more than 150 % of this average. For the first period, regions with income below 50% of the EU average can be found mainly in the southern periphery and most of them are in Greece or Portugal. A few number (7) of these regions had GDP p.c. below 50% of the EU average over the whole period. More precisely these are in Spain (1), Greece (3) and Portugal (3). Except these particular regions, the per capita GDP spatial pattern between 1980-1984 and 2000-2005 is more dynamic in the periphery, indicating a small catching process. Most regions in Spain, Greece, Ireland or Portugal experienced growth rates above the average EU-14 growth rate but one can note that for regions in Spain, Greece and Portugal the average in 2000-2005 is still below the EU-14 1980-1984 average, while the most spectacular result is for Ireland, even if only two regions are concerned.

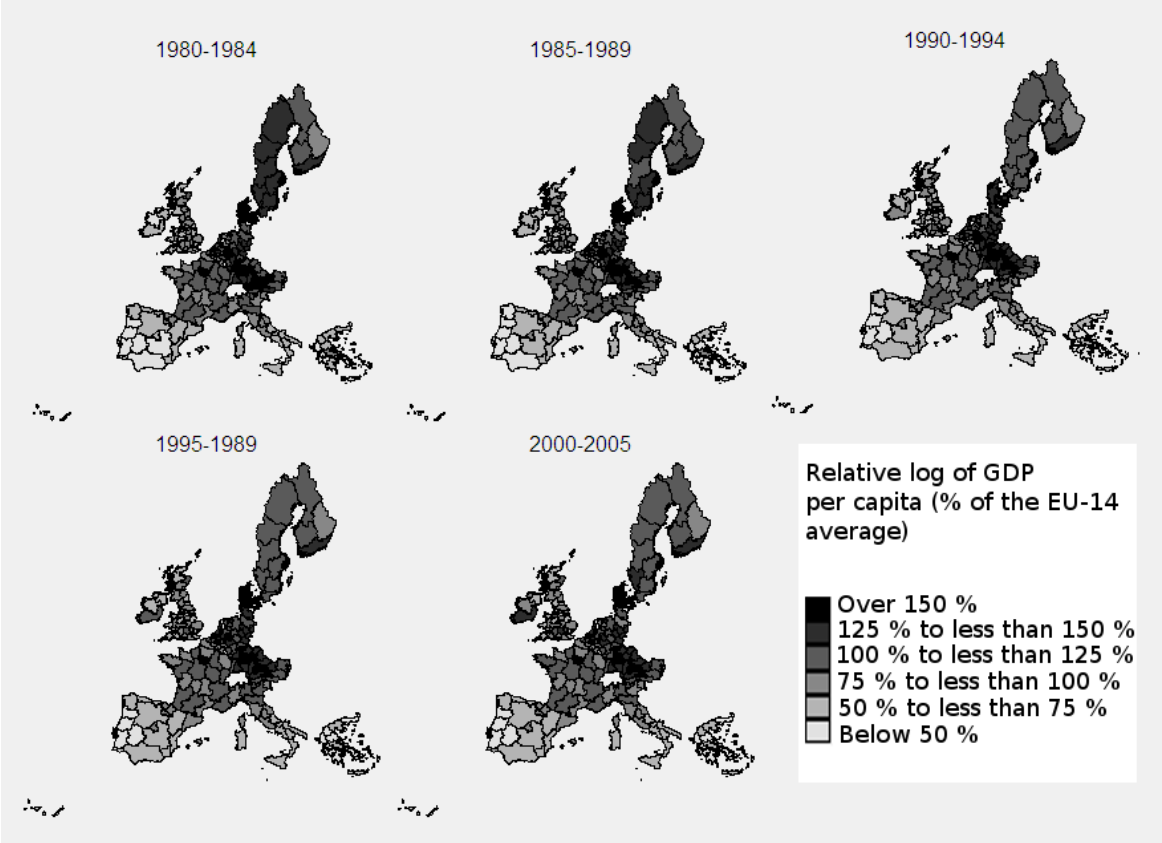

Figure 2 *Geographic pattern of the GDP per capita relative to the EU-14 average GDP level for the 5 periods (adapted from Bouayad-Agha et al, 2009, Cambridge Econometrics database)*

The next section will enable us to assess the strength of this process and to test the various factors which affect it.

3.3. Econometric analysis

Table 2 and 3 report respectively estimates of conditional convergence model for the period 1980-2005. This time lag has been split into 5 periods (1980-84, 1985-89, 1990-94, 1995-99 and 2000-05).

It has been shown that ignorance of the spatial correlation leads to potentially misleading estimates and tests, so we will first analyse the spatial properties of the residuals, and then present the estimation results and validity tests.

3.3.1. Validity tests

The presence of spatial correlation has been tested using the Moran's statistics on the GMM residuals (Z_{moran}, Table 2 and 3). This test is the most commonly used to detect spatial correlation and its application has been extended to residuals regression in Anselin and Kelejian (1997).

Strong significant spatial correlation can be detected on GMM residuals of model (1) that ignore spatial dependence (Table 2). This result confirms the presence of spatial effects in the European regional convergence. As expected, the introduction of a spatial lag term (model (2)) reduces significantly the spatial correlation of residuals for the GMM estimator. Moran's test fail to reject the null hypothesis of no spatial correlation for 1990-1994, 2000-05 and for 1985-89, 1995-99 respectively with a 10% and 5% significance level (Table 3). Similar results are obtained with a spatial error specification (Table 3). These results confirm the presence of spatial dependencies which leads to inconsistent results and inefficient validity tests in model (1). Hence, we are not looking on the remaining tests in Table 2 and we therefore proceed with the spatial lag model, concentrating on Table 3 tests results.

The consistency of the GMM estimator depends on whether lagged values of the autoregressive and spatial autoregressive terms are valid instruments for the regression. In the estimation process, we are using the orthogonality conditions between the error term in first difference and lagged values of the dependent variables. In order to test these conditions, we report tests for first and second-order serial correlation (Arellano and Bond, 1991) and we consider three specification tests (AR(1), AR(2) and Hansen tests).

AR(1) and AR(2) tests provides further support to the model and its estimation since this statistics fails to reject this hypothesis of no second order correlation while reject no first-order serial correlation (Table 3). The overall validity of the instruments can be tested by the Hansen test of over-identifying restrictions. This test confirms the overall validity of the instrumental variables (Table 3) at the 10% level of significance in the spatial lag model. Nevertheless, the result of Hansen test on spatial error specification points out the weakness of the instruments in this context (0.023). Hence, the results must be interpreted carefully.

The Hansen test can be also informative on the validity of additional instruments by comparing the difference of Hansen statistic between two sets of instruments. The Hansen-diff reported in Table 3 check the validity of additional instruments used by system-GMM. We report first-differenced GMM results because Hansen-diff (21.01) rejects the validity of additional instruments used by system-GMM for the level equation.

We assume that the explanatory variables are exogenous using the Hansen-diff test between exogenous and predetermined set of instruments (as suggested by Bond, 2001). Overall validity tests do not indicate problems with instrument validity and orthogonality conditions used by first-differenced GMM estimators. We do not use system-GMM because additional instruments of the level equation are not valid.

3.3.2. Results

Within the framework of a dynamic panel specification that considers the spatial dimension of GDP per capita, we find empirical evidence on conditional convergence of European regions. This process is strongly affected by spatial dimension.

As explained above, the GMM estimator controls for both endogeneity and other econometric problems; it is expected to address the inconsistency of POLS and LSDV estimators which provide however upper and lower bounds for the autoregressive parameter (Bond *et al.*, 2001)[6]. For Model (6) and (7), LSDV values are respectively lies between 0.48 and 0.39, and around 0.89 when the POLS estimator is used. As expected, for these two models, the estimated autoregressive parameters fall between these bounds: close to 0.8 for the model (6) and 0.5 for the model (7).

The estimated autoregressive parameter (Table 2) are close to those obtained in other studies (Caselli *et al.*, 1996). this result points out that there is a significant European regional convergence, conditionally to investment per capita (which have a significant positive impact on regional

[6] With fixed T, POLS gives an estimate of the coefficient of the lagged income that is biased upward in the presence of individual specific effects (Hsiao, 1986) and LSDV gives an biased downwards estimate for the same coefficients (Nickell, 1981).

development). The non significant impact of demographic growth suggests that the evolution of labour force does not affect regional development.

Table 2 *Estimation results of Dynamic model (6)*

	POLS	LSDV	GMM-DIFF
$ln(\frac{Y_{i,t-1}}{pop_{i,t-1}})$	0.898***	0.479***	0.829***
	(0.01)	(0.05)	(0.19)
$ln(\frac{I_{i,t}}{pop_{i,t}})$	0.089***	0.218***	0.193***
	(0.01)	(0.02)	(0.03)
$ln(\frac{pop_{i,t}}{pop_{i,t-1}})$	0.006***	-0.000	-0.002
	(0.00)	(0.00)	(0.00)
constant	0.335***	3.291***	
	(0.07)	(0.55)	
Convergence speed	2.15%	14.7%	4.0%
Observations	764	764	573
r^2	0.983	0.926	
AR(1)			-5.88
AR(1) (p.v.)			(0.00)
AR(2)			0.54
AR(2) (p.v.)			(0.59)
Hansen J			20.77
Hansen J (p.v.)			(0.00)
Hansen-Diff J			6.44
Hansen-Diff J (p.v.)			(0.04)
Z_{moran}			
Z_{85-89}	10.517	4.579	11.831
Z_{85-89}(p.v.)	(0.00)	(0.00)	(0.00)
Z_{90-94}	9.691	12.504	9.087
Z_{90-94}(p.v.)	(0.00)	(0.00)	(0.00)
Z_{95-99}	14.578	6.601	14.632
Z_{95-99}(p.v.)	(0.00)	(0.00)	(0.00)
Z_{00-04}	6.617	4.147	7.631
Z_{00-04}(p.v.)	(0.00)	(0.00)	(0.00)

*Notes: POLS: Pooled Ordinary Least Squares; LSDV: Least Squares Dummy Variable; GMM-DIFF: Generalized Method of Moments on forward orthogonal deviation equation. *, ** and *** indicate significance at the 1%, 5% and 10% level. Robust standard errors are displayed in parentheses. GMM-DIFF estimations are obtained by a forward orthogonal deviation transformation with instrument set composed by* $\left(\ln\left(\frac{Y_{i,1}}{pop_{i,1}}\right), \ln\left(\frac{Y_{i,2}}{pop_{i,2}}\right),\ldots,\ln\left(\frac{Y_{i,T-2}}{pop_{i,T-2}}\right)\right)$ *and the overall set of other explanatory variables (considered as exogenous) .*

The results from the dynamic spatial panel data models are displayed in Table 3. The introduction of spatially lagged variable leads to significantly different results. The autoregressive parameter (0.453) in model (7) drops sharply regarding to model (6) indicating that we measure a faster regional conditional convergence when we consider the impact of neighbouring income on regional development. Spatial lagged income coefficient (0.417) suggests a strong significant impact of spillover effect between European regions on their development dynamics.

We can explain this drop by the fact that a part of y_{it} variability previously explained by the variability of the lagged term y_{it-1} in (6) is due to spatial interdependence Wy_{it} variability in (7). The sum of $(1+\beta_1)$ and ρ in (7) is close to $(1+\beta_1)$ in (6).

ρ reflects how spillover affect income growth rate for all regions with the same initial income. The development of the European regions is strongly affected by their spatial interdependence, thus suggesting that the convergence process is not only a "temporal" process but also a spatial process. In the framework of the Solow model, we can assume that speed of convergence to the steady state is affected by regional spillover. The rate at which the regions converge towards their steady state depends on the assumption of decreasing marginal product. Regional spillovers may induce this decline, so, the dynamics toward the steady state doesn't depend only of the gap between $Y_{i,0}$ and its steady state value but also on the strength of spatial spillover. Thus, spatial concentration of GDP p.c. must impede the convergence process. However, Egger and Pfaffermayr (2006) explain that spatial spillover implies that the convergence speed depends on the region's location. Thus, they suggest amending the computation of the convergence speed in order to allow for this issue. we can decompose (As suggested by Egger and Pfaffermayr, 2006) the convergence speed into its "classical" part, a remoteness effect and the impact of initial gap.

The spatial error model provides results similar to the other studies. The convergence speed (4.8%) is in line with the values computed by Badinger et al. (2004).

The introduction of a spatial lag does not affect the impact of investment per capita spending on regional development. The $\beta_2(\ln\frac{I_{i,t}}{pop_{i,t}})$ coefficient is still significant around 0.16. Its value increases to 0.236 when we allow for spatial error dependence.

The demographic growth rate is not significant for any specification.

Table 3 *Estimation results of the Spatial Dynamic model (7) and (8)*

	POLS Spatial autoregressive lag	FE2SLS Spatial autoregressive lag	GMM-DIFF Spatial autoregressive lag	GMM-DIFF Spatial autoregressive error
$ln(\frac{Y_{i,t-1}}{pop_{i,t-1}})$	0.887***	0.398***	0.453***	0.786***
	(0.02)	(0.08)	(0.09)	(0.04)
$Wln(\frac{Y_{i,t}}{pop_{i,t}})$	0.018	0.433***	0.417***	
	(0.02)	(0.09)	(0.11)	
$ln(\frac{I_{i,t}}{pop_{i,t}})$	0.087***	0.178***	0.159***	0.236***
	(0.01)	(0.02)	(0.03)	(0.03)
$ln(\frac{pop_{i,t}}{pop_{i,t-1}})$	0.006***	-0.001	-0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.001)
constant	0.288***			
	(0.08)			
$Wu_{i,t}$				0.8439
Convergence speed	2.4%	19.0%	15.7%	4.8%
Observations	764	764	573	573
r^2	0.983	0.928		
AR(1)			-2.152	
AR(1) (p.v.)			(0.031)	
AR(2)			1.005	
AR(2) (p.v.)			(0.315)	
Hansen J		1.352	13.69	17.77
Hansen J (p.v.)		(0.245)	(0.134)	(0.023)
Hansen-Diff J			32.93	
Hansen-Diff J (p.v.)			(0.000)	
Z_{moran}				
Z_{85-89}	10.057	4.201	1.651	-1.309
Z_{85-89}(p.v.)	(0.00)	(0.00)	(0.049)	(0.095)
Z_{90-94}	9.128	11.881	0.567	-2.068
Z_{90-94}(p.v.)	(0.00)	(0.00)	(0.285)	(0.019)
Z_{95-99}	14.234	6.195	1.373	-1.094
Z_{95-99}(p.v.)	(0.00)	(0.00)	(0.085)	(0.137)
Z_{00-04}	6.315	3.696	-0.021	-2.693
Z_{00-04}(p.v.)	(0.00)	(0.00)	(0.492)	(0.004)

*Notes: *, ** and *** indicate significance at the 1%, 5% and 10% level. Robust standard errors are displayed in parentheses. SGMM-DIFF estimations are obtained by a forward orthogonal deviation transformation with instrument set composed by* $\left(\ln\left(\frac{Y_{i,1}}{pop_{i,1}}\right), \ln\left(\frac{Y_{i,2}}{pop_{i,2}}\right), \ldots, \ln\left(\frac{Y_{i,T-2}}{pop_{i,T-2}}\right); W\ln\left(\frac{Y_{i,1}}{pop_{i,1}}\right), W\ln\left(\frac{Y_{i,2}}{pop_{i,2}}\right), \ldots, W\ln\left(\frac{Y_{i,T-2}}{pop_{i,T-2}}\right)\right)$ *and the overall set of other explanatory variables (considered as exogenous) .*

4. Conclusion

The aim of this paper is to review strategies to estimate Spatial Dynamic panel data model using GMM.

Then, we estimate a conditional convergence model of European regions which allows both space and dynamic dimension to illustrate these strategies. Using a panel dataset of 191 regions over 1980-2005, we estimate a Spatial Dynamic Panel using spatial lag strategy estimation in addition to standard dynamic model on panel data techniques (Arellano and Bond, 1991; Blundell and Bond, 1998).

We find empirical evidence on conditional convergence of European regions. European convergence is conditional to the investment rate. In analysing the spatial dimension, we find that convergence is significantly affected by the spatial disparities.

Note that other drivers of regional development can affect convergence process of European regions, as European policies (Esposti and Bussoletti, 2008; Bouayad *et al.*, 2009).

The Spatial Dynamic Panel data model is at early stage. In this paper, we concentrate on studying strategies to estimate spatial dynamic panel data. But this approach needs further development.

First, the usual tests on spatial specification (LM-test) must be extended to time dynamic dimension.

Furthermore, there are several possibilities to crossing Dynamic and spatial literature. In the spatial error model, using GMM system estimator as suggested by Blundell and Bond (1998) in time dynamic can improve its efficiency.

References

ABREU, M., H.L.F. DE GROOT & FLORAX, R. J. G. M. (2005) Space and Growth: a Survey of Empirical Evidence and Methods. *Region et Développement,* 21**,** 13-44.

ANSELIN, L. (2001) Spatial Econometrics. IN BALTAGI, B. H. (Ed.) *Theoretical Econometrics* Blackwell Publishing.

ANSELIN, L. (2002) Under the hood issues in the specification and interpolation of spatial regression models. *Agricultural Economics,* 27**,** 247-267.

ANSELIN, L. & KELEJIAN, H. (1997) Testing for spatial error autocorrelation in the presence of endogenous regressors. *International Regional Science Review,* 20**,** 153-182.

ANSELIN, L., L. LE GALLO & JAYET, H. (2007) Spatial panel econometrics. IN MATYAS, L. & SEVESTRE, P. (Eds.) *The Econometrics of Panel Data Fundamentals and Recents Developments in Theory and Practice.* Springer Berlin Heidelberg

ARBIA, G., LE GALLO, J. & PIRAS, G. (2008) Does evidence on regional economic convergence depend on the estimation strategy? Outcomes from analysis of a set of NUTS2 regions. *Spatial Economic Analysis,* 3**,** 209-224.

ARELLANO, M. & BOND, S. (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies,* 58**,** 277-297.

ARELLANO, M. & BOVER, O. (1995) Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics,* 68**,** 29-51.

BADINGER, H., MÜLLER, W. G. & TONDL, G. (2004) Regional convergence in the European Union, 1985-99: a spatial dynamic panel analysis. *Regional Studies,* 38**,** 541-253.

BARRO, R. J. & SALA-I-MARTIN, X. (1992) Convergence. *Journal of Political Economy,* 100**,** 407-443.

BECK, N., GLEDITSCH, K. S. & BEARDSLEY, K. (2006) Space Is More than Geography: Using Spatial Econometrics in the Study of Political Economy. *International Studies Quarterly,* 50**,** 27-44.

BLONIGEN, B. A., DAVIES, R. B., WADDELL, G. R. & NAUGHTON, H. T. (2007) FDI in space: Spatial autoregressive relationships in foreign direct investment. *European Economic Review,* 51**,** 1303-1325.

BLUNDELL, R. & BOND, S. (1998) Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics,* 87**,** 115-143.

BOND, S. (2002) Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal*, 1, 141-62.

BOND, S., HOEFFLER, A. & TEMPLE, J. (2001) GMM estimation of empirical growth models. *CEPR Discussion Paper*, 3048.

BOUAYAD-AGHA, S., TURPIN, N. & VEDRINE, L. (2009) Fostering the endogenous potential development of European regions: a panel data analysis of the cohesion policy on regional convergence over the period 1980-2005. *International Regional Science Review*, under review.

CASELLI, F., ESQUIVEL, G. & LEFORT, F. (1996) Reopening the convergence debate: A new look at cross-country growth empirics. *Journal of Economic Growth,* 1**,** 363-389.

DURLAUF, S., JOHNSON, P. & TEMPLE, J. (2006) *Growth econometrics,* Amsterdam, North Holland, P. Aghion and S. Durlauf (Eds), Handbook of Economic Growth.

ECKEY, H.F., DREGER C., & TURCK M. (2006) European Regional Convergence in an Human Capital Augmented Solow Model. *University of Kassel working paper,* 88, 1-19.

EGGER, P. & PFAFFERMAYR, M. (2006) Spatial convergence. *Papers in regional science,* 85(2), 200-215.

ELHORST, J. P. (2003) Specification and Estimation of Spatial Panel Data Models. *International Regional Science Review,* 26**,** 244-268.

ELHORST, J. P. (2005) Unconditional Maximum Likelihood Estimation of Linear and Log-Linear Dynamic Models for Spatial Panels. *Geographical Analysis,* 37**,** 85-106.

ESPOSTI, R. & BUSSOLETTI, S. (2008) Impact of Objective1 funds on regional growth convergence in the European Union: a panel-data approach. *Regional Studies,* 42**,** 159-173.

HSIAO, C. (1986) *Analysis of Panel Data*. Cambridge University Press.

KAPOOR, M., KELEJIAN, H. & PRUCHA, I. (2007) Panel data models with spatially correlated error components. *Journal of Econometrics,* 140(1), 97-130.

KELEJIAN, H. & ROBINSON, D.P. (1995) Spatial correlations: a suggested alternative to the autoregressive model. IN ANSELIN, L. & FLORAX, R. (Eds) *New Directions in Spatial Econometrics*, Springer, Berlin.

KUKENOVA, M. & MONTEIRO, J. A. (2008) Spatial dynamic panel model and system GMM: a Monte-Carlo investigation. *Munich personnal RePEc archive paper n°11569*.

MADRIAGA, N. & PONCET, S. (2007) FDI in Chinese Cities: Spillovers and Impact on Growth. *The World Economy,* 30**,** 837-862.

MANKIW, G. N., ROMER, D. & WEIL, D. N. (1992) A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics,* 107**,** 407-437.

MUTL, J. (2006) Dynamic panel data models with spatially correlated disturbances. *Phd dissertation of the University of Maryland.*

NICKELL, S. (1981) Biases in dynamic models with fixed effects. *Econometrica*, 49, 1417-1426.

PIRAS, G. & ARBIA, G. (2007) Convergence in per-capita GDP across EU NUTS2 regions using panel data models extended to spatial autocorrelation effects. *Statistica*.

REY, S.J. & JANIKAS, M.V. (2005) Regional convergence, inequality, and space. *Journal of Economic Geography* ,5(2), 155-176.

SOLOW, R. M. (1956) A contribution to the theory of economic. *The Quarterly Journal of Econometrics*, 70**,** 65-94.

SWAN, T. W. (1956) ECONOMIC GROWTH and CAPITAL ACCUMULATION. *The Economic Record,* 32**,** 334-361.

TEMPLE, J. (1999) The New Growth Evidence. *Journal of Economic Literature,* 37, 112-156.

APPENDIX A

Code	Name	Code	Name	Code	Name	Code	Name
	Included regions (190)						
AT11	Burgenland	ES13	Cantabria	IE02	Southern and Eastern	UKD5	Merseyside
AT12	Niederösterreich	ES21	Pais Vasco	ITC1	Piemonte	UKE1	East Riding and North Lincolnshire
AT13	Wien	ES22	Comunidad Foral de Navarra	ITC2	Valle d'Aosta	UKE2	North Yorkshire
AT21	Kärnten	ES23	La Rioja	ITC3	Liguria	UKE3	South Yorkshire
AT22	Steiermark	ES24	Aragón	ITC4	Lombardia	UKE4	West Yorkshire
AT31	Oberösterreich	ES30	Comunidad de Madrid	ITD1	Provincia Autonoma Bolzano-Bozen	UKF1	Derbyshire and Nottinghamshire
AT32	Salzburg	ES41	Castilla y León	ITD2	Provincia Autonoma Trento	UKF2	Leicestershire, Rutland and Northants
AT33	Tirol	ES42	Castilla-la Mancha	ITD3	Veneto	UKF3	Lincolnshire
AT34	Vorarlberg	ES43	Extremadura	ITD4	Friuli-Venezia Giulia	UKG1	Herefordshire, Worcestershire and Warks
BE10	Région de Bruxelles-Capitale	ES51	Cataluña	ITD5	Emilia-Romagna	UKG2	Shropshire and Staffordshire
BE21	Prov. Antwerpen	ES52	Comunidad Valenciana	ITE1	Toscana	UKG3	West Midlands
BE22	Prov. Limburg (B)	ES53	Illes Balears	ITE2	Umbria	UKH1	East Anglia
BE23	Prov. Oost-Vlaanderen	ES61	Andalucia	ITE3	Marche	UKH2	Bedfordshire, Hertfordshire
BE24	Prov. Vlaams Brabant	ES62	Región de Murcia	ITE4	Lazio	UKH3	Essex
BE25	Prov. West-Vlaanderen	ES70	Canarias (ES)	ITF1	Abruzzo	UKI1	Inner London
BE31	Prov. Brabant Wallon	FI13	Itä-Suomi	ITF2	Molise	UKI2	Outer London
BE32	Prov. Hainaut	FI18	Etelä-Suomi	ITF3	Campania	UKJ1	Berkshire, Bucks and Oxfordshire
BE33	Prov. Liège	FI19	Länsi-Suomi	ITF4	Puglia	UKJ2	Surrey, East and West Sussex
BE34	Prov. Luxembourg (B)	FI1A	Pohjois-Suomi	ITF5	Basilicata	UKJ3	Hampshire and Isle of Wight
BE35	Prov. Namur	FR10	Île de France	ITF6	Calabria	UKJ4	Kent
DE11	Stuttgart	FR21	Champagne-Ardenne	ITG1	Sicilia	UKK1	Gloucestershire, Wiltshire and North Somerset
DE12	Karlsruhe	FR22	Picardie	ITG2	Sardegna	UKK2	Dorset and Somerset
DE13	Freiburg	FR23	Haute-Normandie	NL11	Groningen	UKK3	Cornwall and Isles of Scilly
DE14	Tübingen	FR24	Centre	NL12	Friesland	UKK4	Devon
DE21	Oberbayern	FR25	Basse-Normandie	NL13	Drenthe	UKL1	West Wales and The Valleys
DE22	Niederbayern	FR26	Bourgogne	NL21	Overijssel	UKL2	East Wales
DE23	Oberpfalz	FR30	Nord - Pas-de-Calais	NL22	Gelderland	UKM1	North Eastern Scotland
DE24	Oberfranken	FR41	Lorraine	NL23	Flevoland	UKM2	Eastern Scotland
DE25	Mittelfranken	FR42	Alsace	NL31	Utrecht	UKM3	South Western Scotland
DE26	Unterfranken	FR43	Franche-Comté	NL32	Noord-Holland	UKM4	Highlands and Islands
DE27	Schwaben	FR51	Pays de la Loire	NL33	Zuid-Holland	UKN0	Northern Ireland
DE50	Bremen	FR52	Bretagne	NL34	Zeeland		Excluded regions (22)
DE60	Hamburg	FR53	Poitou-Charentes	NL41	Noord-Brabant	DE30	Berlin
DE71	Darmstadt	FR61	Aquitaine	NL42	Limburg (NL)	DE41	Brandenburg - Nordost
DE72	Gießen	FR62	Midi-Pyrénées	PT11	Norte	DE42	Brandenburg - Südwest
DE73	Kassel	FR63	Limousin	PT15	Algarve	DE80	Mecklenburg-Vorpommern
DE91	Braunschweig	FR71	Rhône-Alpes	PT16	Centro (PT)	DED1	Chemnitz
DE92	Hannover	FR72	Auvergne	PT17	Lisboa	DED2	Dresden
DE93	Lüneburg	FR81	Languedoc-Roussillon	PT18	Alentejo	DED3	Leipzig
DE94	Weser-Ems	FR82	Provence-Alpes-Côte d'Azur	SE01	Stockholm	DEE1	Dessau
DEA1	Düsseldorf	FR83	Corse	SE02	Östra Mellansverige	DEE2	Halle
DEA2	Köln	GR11	Anatoliki Makedonia, Thraki	SE04	Sydsverige	DEE3	Magdeburg
DEA3	Münster	GR12	Kentriki Makedonia	SE06	Norra Mellansverige	DEG0	Thüringen
DEA4	Detmold	GR13	Dytiki Makedonia	SE07	Mellersta Norrland	ES63	Ciudad Autónoma de Ceuta (ES)
DEA5	Arnsberg	GR14	Thessalia	SE08	Övre Norrland	ES64	Ciudad Autónoma de Melilla (ES)
DEB1	Koblenz	GR21	Ipeiros	SE09	Småland med öarna	FI20	Åland
DEB2	Trier	GR22	Ionia Nisia	SE0A	Västsverige	FR91	Guadeloupe (FR)
DEB3	Rheinhessen-Pfalz	GR23	Dytiki Ellada	UKC1	Tees Valley and Durham	FR92	Martinique (FR)
DEC0	Saarland	GR24	Sterea Ellada	UKC2	Northumberland, Tyne and Wear	FR93	Guyane (FR)
DEF0	Schleswig-Holstein	GR25	Peloponnisos	UKD1	Cumbria	FR94	Reunion (FR)
DK00	denmark	GR30	Attiki	UKD2	Cheshire	GR41	Voreio Aigaio
ES11	Galicia	GR43	Kriti	UKD3	Greater Manchester	GR42	Notio Aigaio
ES12	Principado de Asturias	IE01	Border, Midlands and Western	UKD4	Lancashire	PT20	Região Autónoma dos Açores (PT)