Fostering the endogenous potential development of European regions:

a panel data analysis of the cohesion policy on regional convergence over the period 1980-2005

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

[1] 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, nadine.turpin@cemagref.fr, lionel.vedrine@cemagref.fr

No senior authorship involved.

Paper prepared for the Regional Studies Association International Conference (Pécs, may 2010).

Granted by the Regional Council of Auvergne and the FP7 (PRIMA project)

Abstract

This paper investigates whether structural funds policy affects the European economies in such a way that poorer regions catch up with the rich ones, estimating a conditional convergence econometric model. In this model, regional convergence depends on the policy treatment and the regional economic structure, proxied by investment per capita and demographic growth rate. The convergence model is specified in a dynamic panel-data form on a dataset of 191 NUTS II EU14 regions observed over more than 25 years (from 1980 to 2005). A Generalized Method of Moment estimation enables obtaining consistent estimates of the beta-parameter along with estimates of the impact of the regional policies and the regional economic structure on regional growth. The analysis highlights that structural funds have a small but significant impact on Objective1 regions convergence rate, whatever the impact model we designed. Moreover, considering for the spatial dimension of the panel leads a still significant, but less important, impact of structural funds on convergence.

Key-words : Dynamic panels, GMM, Regional Convergence, Spatial Dependence, Structural Funds

JEL : C21, C23, O52, R11, R1

Introduction

The main objective of the European cohesion policy is the reduction of the economic and social disparities between the levels of development of the European regions. The policy aims at fostering the endogenous development potential of European regions and so far has focussed on lagging regions. The policy has been renewed in 2007 and a debate has just been initiated to continuously improve the policy, based on a public consultation about the budget review and about the territorial cohesion strategy (European Commission, 2008). At this stage, it is important to evaluate the impact of past Structural Funds (SF) expenditure to assess whether structural policies are effectively leading to a narrowing of disparities of wealth among EU regions. This paper investigates empirically such impact of SF on European regions convergence.

The convergence and growth approach is particularly appropriate to check if the Cohesion Policy affect EU regions in such a way that lagging ones catch up with the rich ones. The large amount of empirical literature published in the last few years on the regional economic convergence dilemma stress that the convergence speed measurements strongly depend upon the model specifications, the data used and the estimation strategy (see for recent surveys: Arbia *et al.* (2008); Esposti and Bussoletti (2008)). The most recent convergence studies are anchored on advances from linear dynamic panel-data models, and use dedicated estimation techniques to obtain a consistent Generalized Method-of-Moments (GMM) estimator (Arellano and Bond, 1991, Arellano and Bover, 1995, Blundell and Bond, 1998). Moreover, working on regional data motivates a stream of literature to account for spatial dependence in regional growth. The spatial econometric literature is well documented either on cross-sections (Anselin and Kelejian, 1997, Anselin, 2001) or on static panel (Elhorst, 2003), but empirical methods for dynamic spatial panels that are required for regional convergence studies, are still at an early stage of development. Working on a dynamic panel specification, Badinger (Badinger *et al.*, 2004) applied a GMM estimator to spatially filtered variables. More recently Elhorst (Elhorst, 2005) suggests a maximum likelihood estimation of models that are dynamic both in space and time for regional analysis and Arbia and Piras (2005) extend panel-data models with spatial error autocorrelation for a convergence analysis of EU regions.

However, measuring a convergence rate is not sufficient to analyse whether SF foster the development of EU lagging regions and we have to design an appropriate econometric method to conduct such an impact analysis. Using space-time data is quite complex and the design of an accurate specification is not evident: obviously, leaving dynamic specification and spatial dependence specification 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. Therefore, conducting an accurate impact analysis of such a policy necessitates gathering and analysing together observations before the policy is implemented and during the policy implementation and thus long space-time series organised in panel data. Such space-time data sets require the design of an estimation strategy to consider together the dynamic specification and the spatial dimension of the panel. In our case, we have to combine this estimation strategy with an accurate specification for the impact analysis of structural funds inside this framework.

In this paper, we estimated 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.

We concentrate on funds targeted towards Objective1 (O1) regions, because they represent two third of the SF expenditure (for the 2000-2006 programme). These regions are lagging behind in terms of GDP per capita (lower than 75 % of the EU average) or population density (below 8 inhabitants per km^2). The analysis is on 191 EU14[1] regions over the period 1980-2005. This period covers observations on EU regions before (80-88) and during (89-93, 94-99 and 2000-2006) the programming periods of SF. Working on observations before the policy applies can be helpful to apply impact evaluation methods is a necessary condition for impact evaluation methods (Ravallion, 2005).

---

[1] Current NUTS 2 classification includes 212 regions (EU15). However, we had to eliminate some regions due to missing data (see annex 1 for a description of the analysed regions) and we choose not to consider Luxembourg.

The analysis highlights that structural funds have a small but significant impact on Objective1 regions convergence rate, whatever the impact model we designed. Moreover, considering for the spatial dimension of the panel leads a still significant, but less important, impact of structural funds on convergence. This last result confirms that the GDP concentration over the region tends to hamper them exploiting their endogenous development potentials and entail their convergence. As such, the cohesion policy that aims at counterbalancing the effects of activities concentration can attain this objective.

The reminder of the paper is organised as follows. The first Section develops some theoretical and empirical considerations on the impact analysis of structural funds on convergence. Section 2 presents the econometric issues on spatial dynamic panel model.. Section 3 describes the data and describes the specificity of O1 regions in Europe. Section 4 presents the results and Section 5 concludes.

1. Theoretical and empirical considerations: impact analysis of structural funds on convergence

1.1. Impact evaluation of Structural Funds

The central question in the impact analysis of structural funds is what would have happened to the regions receiving SF if they had not in fact benefited from the scheme. Since, at any moment in time, a group is either in the scheme under consideration or not, but not both, at the hearth of this kind of policy evaluation is a missing problem. Constructing a counterfactual (a group which is as similar as possible to those benefiting from the scheme) is the major issue that evaluation methods address (for a review see Blundell and Costa Dias, 2002). The design of such a counterfactual is hampered by a selection bias because O1 regions are likely to differ from the other (untreated) regions for some unobserved characteristics. As not all regions in Europe are eligible to O1 scheme, we have to separate this selection effect from the treatment effect we specifically want to analyse (Ravallion, 2005). The average effect of the SF policy (treatment effect on the treated) is measured as the difference between the outcome value under treatment and the counterfactual of no receiving treatment.

Traditionally, approaches aiming at evaluating the impact of schemes in the presence of non-random scheme placement combine two types of data to elaborate counterfactual outcomes, observations on the non-treated regions and observations on the treated regions collected on periods preceding the scheme reception. These two types of observations are analysed with various regression approaches:

- The before-after approach could be a solution to this unobservable counterfactual problem in a non-experimental context (Todd, 2006). A major drawback of this estimator is that it is sensitive to time effects since O1 regions can have different economic situations between the two periods (cyclical bias).
- The cross-section uses information only on non O1 regions. It is affected by a selection bias for non random placement.
- Finally, the difference in difference approach is the most often used one. It requires pre and post-programs data on participants and non participants and measures the impact of the SF policy by the difference in the before-after change in outcomes between participants and non participants (Todd, 2006). It is based upon the hypothesis that the outcome difference over time for eligible regions in absence of subsidy is the same as the outcome difference over time for non eligible regions that will never receive any subsidy[2]. We also assume that the control group regions are not affected by the scheme that applies only in the treated regions. In other words, we assume that there is no spillover effect of the structural funds targeted to O1 regions towards the non treated ones. So we retain this last approach.

1.2. The growth approach

Once the impact analysis empirical approach delimited with the associated selection bias controlled, we can turn to the convergence issues that the cohesion policy aims at. 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.

[2] This assumption can be considered as plausible if both regions groups have similar trends prior to program implementation

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 denoted as *conditional*: economies converge towards the same growth rate. Accordingly, the following general model is in line with the empirical growth literature:

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

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 an "equation à la Barro": for a dynamic panel-data specification of the convergence model, current output is regressed on lagged output and control variables (Durlauf et al., 2006, Islam, 1995):

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

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.

1.3. Impact analysis in a convergence model

So far, none of the variables considers the structural policies. We need to extend the model towards impact evaluation of the policy, including additional dummies variables T and P to assess the policy impact on conditional convergence. The regional projects have been designed in a bottom-up approach and thus we assume in our analysis that not all the means have been activated in all regions at all time. Such an assumption is translated in the estimation strategy by the choice of no specification for the way the policy fosters regional development. In other words, we shall not investigate whether the EU SF policy increases O1 regions' income by mean of the investment rate. We shall concentrate on assessing whether belonging to the O1 group enables valorising projects dynamics that can result in higher income levels (OECD, 2006).

The difference-in-difference treatment effect approach can be obtained rewriting the equation (1) as:

$$\ln\left(\frac{Y_{i,t}}{pop_{i,t}}\right) = (1+\beta_1)\ln\left(\frac{Y_{i,t-1}}{pop_{i,t-1}}\right) + \beta_2 \ln\left(\frac{I_{i,t}}{pop_{i,t}}\right) + \beta_3 \ln\left(\frac{pop_{i,t}}{pop_{i,t-1}}\right) + \beta_4 T_t + \beta_5 P_i + \beta_6 P_i T_t + \alpha_i + \mu_t + \varepsilon_{i,t} \qquad (2)$$

T and P are dummy variable, the first one indicating if we are or not in an O1 program (T=1 when the O1 start up (after 1989) and 0 otherwise) and the second one indicating if the region under consideration is eligible at this O1 program. These two variables allow controlling on cyclical and selection bias: the average effect of O1 regions are measured by the coefficient β_6 while β_4 and β_5 capture respectively cyclical and selection bias.

Because of our aim is to predict O1 impact on regional convergence, we also last introduce an interaction term between the participation of O1 and the lagged GDP p.c. in equation (2):

$$\ln\left(\frac{Y_{i,t}}{pop_{i,t}}\right) = (1+\beta_1)\ln\left(\frac{Y_{i,t-1}}{pop_{i,t-1}}\right) + \beta_2 \ln\left(\frac{I_{i,t}}{pop_{i,t}}\right) + \beta_3 \ln\left(\frac{pop_{i,t}}{pop_{i,t-1}}\right) + \beta_4 T_t + \beta_5 P_i + \beta_6 P_i T_t + (1+\beta_7) D_{i,t-1} \ln\left(\frac{Y_{i,t-1}}{pop_{i,t-1}}\right) + \alpha_i + \mu_t + \varepsilon_{i,t} \qquad (3)$$

With $D_{i,t-1}$=1 when the region i are eligible to O1 at time t-1 and 0 otherwise. The interaction term tests if the policy affects O1 regions convergence process.

To sum up, the first model (equation 1) explains convergence process without structural policy, the second one (equation 2) provides an average effect of the policy on O1 regions (β_6) and the third analyses the marginal impact of SF on O1 regions convergence ($\beta_1 + \beta_7$).

2. Econometric issues

2.1. Estimation in dynamic panel

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 must be used. The most commonly used estimator for dynamic panels in the literature is the GMM estimator (Arellano and Bond, 1991) adopted by Caselli *et al.* (1996) in the growth context. Lagged levels of the series are used to instrument lagged first differences (GMM-DIFF estimator). The first step in the estimation procedure consists in eliminating the individual effects via a first difference transformation or forward orthogonal deviation[3] (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 in question y_{it-1} starting with lag two y_{it-2} and going back to earlier lagged levels.

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 estimator behaves poorly. The system GMM estimator developed by Arelano and Bover (1995) and Blundell and Bond (1998) allows the removal of fixed effects via the first-differenced equation and to maintain the distance variable in the level equation, besides the technical gains of additional level moment conditions and increased efficiency. Additional assumptions about the initial conditions, not too restrictive in empirical growth frameworks, yields moment conditions still informative even for persistent time series. The GMM-SYS estimator combines the standard set of equations in first-differences with suitably lagged first-differences as instruments, with an additional set of equations in levels with appropriate lagged first-differences as instruments. In this stacked system, the standard set of equations in first-differences is instrumented with suitably lagged first differences while the instruments for the level equation are lagged differences of the variables.

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

The choice between the two estimators (GMM-DIFF, GMM-SYS) can finally be done 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 dependencies[4]

Since the end of the 1990s various convergence studies have found evidence for serious model misspecifications when spatial interdependencies of regions are not considered in the analysis (Abreu et al., 2005) As discussed in Arbia *et al.* (2008) the application of spatial econometric methods is an essential tool for proper statistical inference on regional data. Spatial error model and spatial lag model are two different approaches to address this issue of spatial dependencies. The first one includes a spatial autoregressive process in the error term. The second method captures spatial dependence with a regressor that is similar to a lagged dependent variable, which is often considered as a spatial autoregression model. In line with recent literature (Beck et al., 2006, Blonigen et al., 2007), we reckon that this last model is more appropriate to quantify how the growth rate of a region is affected by the growth rate in the surrounding regions.

We adopted a space-time simultaneity specification (Anselin *et al.*, 2007) and introduced a spatial lag term $\rho W \ln\left(\frac{Y_{i,t}}{pop_{i,t}}\right)$ in equations (1) to (3). The modified equations (not detailed here) are quoted (4) to (6) in the estimation section. The design of the models relies on a weight matrix W (k-nearest neighbours' matrix). The way this matrix is designed is of importance since it defines how space is accounted for (see appendix for a discussion of the construction of our weight matrix). The spatial lag coefficient captures the impact of the endogenous variable from neighbourhood locations.

[4] Spatial heterogeneity can also create low values cluster and lead to inconsistent estimates. This is very difficult to distinguish between spatial dependence and spatial heterogeneity in cross section context. But in a panel data, spatial heterogeneity can be treated by standard panel methods (fixed effect or first differencing transformation).

As underlined by 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[5], dynamic spatial lag models are usually estimated using the GMM estimator (see for example, Madriaga and Poncet, (2007)). The spatial lag term is instrumented by its lagged values, by lagged values of the dependent variables as well as by spatially weighted explanatory variables. One can check 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 estimations.

3. EU structural funds and O1 regions

The European investments from the cohesion policy aim at improving the competitive position of regional policies by encouraging regions to provide public goods, like networks of transport and energy, environmental quality, investments in education and research-development. In other words, the convergence policy fosters regional development by various means, like competitiveness enhancement, infrastructure improvement, active labour market facilities, innovation enhancement or sustainable development. The public goods provided result from public and private expenditure. The policy seeks to add value beyond simple investments, with a multi-levels governance model to involve local and regional actors in the design and delivery of the policy. This governance model enables regions to activate the most appropriate drivers to foster their development and design projects in a bottom-up approach.

From 1988 onwards, the cohesion policy relies on the same principles. The policy concentrates funds on a limited number of "objectives", with a focus on the least developed regions; funding is based on multi-annual programming with ongoing analysis and evaluation; the design and implementation of the programs involve

[5] As underlined by Kukenova and Monteiro KUKENOVA, M. & MONTEIRO, J. A. (2008) Spatial dynamic panel model and system GMM: a Monte-Carlo investigation. *Munich personnal RePEc archive paper n°11569*. there is no currently available estimator to consider this simultaneity problem in line with the potential endogeneity of other explanatory variables

regional, national and EU actors, and additionality ensures that EU expenditure is not substituted to national investment. The focus on the least developed regions concentrates funding on the Objective 1 regions, that represent about 25 % of the European population, and benefit from 64 % of the allocated funds (for the last program).

3.1. Data description

The analysed dataset has been designed according to econometric issues described in Section2. Because we address impact analysis, this dataset comprises O1 and non-O1 regions, including data collected before the implementation of the policy. 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). A small number of regions are excluded due to missing data which several are eligible on Objective 1 programs (New German Lander, French overseas etc...). Finally, our dataset represents 90% of overall EU-15 regions and 80 % of Objective1 regions. The 191 regions are observed over a period to 25 years (1980-2005).

Data variables related to equations (1) to (6) come from the Cambridge Econometrics database[6]. 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 discussed 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

[6] the Cambridge Econometrics database is available at http://www.camecon.com

(2004) and chose quinquennial time periods to collect information at least two periods before the beginning of the policy. 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 last variable, Objective1 eligibility, comes from the Commission cohesion reports[7].

3.2. Income dynamics in European regions

Table 1: Descriptive statistics

Variable	1980-84	1985-89	1990-94	1995-99	2000-05
$ln(\frac{Y_{i,t}}{pop_{i,t}})$					
Total sample mean (std)	9.44(0.40)	9.54(0.39)	9.63(0.39)	9.71(0.38)	9.81(0.37)
Objective 1 regions mean (std)	9.01(0.32)	9.09(0.31)	9.18(0.28)	9.27(0.29)	9.41(0.29)
Non treated regions mean (std)	9.60(0.29)	9.71(0.27)	9.80(0.27)	9.87(0.26)	9.96(0.26)
$ln(\frac{I_{i,t}}{pop_{i,t}})$					
Total sample mean (std)	7.82(0.47)	7.95(0.44)	7.99(0.41)	8.08(0.37)	8.21(0.34)
Objective 1 regions mean (std)	7.44(0.36)	7.54(0.33)	7.62(0.30)	7.73(0.27)	7.93(0.28)
Non treated regions mean (std)	7.96(0.43)	8.11(0.36)	8.14(0.36)	8.21(0.30)	8.32(0.30)
$ln(\frac{pop_{i,t}}{pop_{i,t-1}})$					
Total sample mean (std)		2.84(1.39)	3.37(1.19)	3.28(1.24)	3.29 (1.30)
Objective 1 regions mean (std)		3.01(1.45)	2.73(1.26)	3.12(1.37)	3.42(1.34)
Non treated regions mean (std)		2.78(1.36)	3.55(1.11)	3.33(1.20)	3.26(1.29)
$Wln(\frac{Y_{i,t}}{pop_{i,t}})$					
Total sample mean (std)	9.46(0.34)	9.55(0.34)	9.64(0.34)	9.72(0.34)	9.83(0.32)
Objective 1 regions mean (std)	9.10(0.33)	9.18(0.33)	9.26(0.31)	9.35(0.32)	9.48(0.30)
Non treated regions mean (std)	9.60(0.24)	9.70(0.22)	9.78(0.23)	9.86(0.21)	9.96(0.21)

Table 1 depicts the dynamics of GDP p.c., investment per capita, demographic growth rate and spatially lagged GDP for O1 regions and other regions in Europe. For every time period, O1 regions exhibit a GDP per capita far lower than the European average. The difference between O1 and non-O1 region increases from 1980-84 period to 85-89 on, and then slowly decreases till today (for a graphical representation of this dynamics, refer to figures 1 and 3).

[7] cohesion reports are available here: http://ec.europa.eu/regional_policy/policy/object/index_gb.htm

Investment par capita is lower in O1 regions despite the difference between O1 and non-O1 regions' investment is reduced in the last two periods. It is worth noting that non-O1 regions tend to concentrate income more than O1 regions, for the all periods, but still the difference decreases along time.

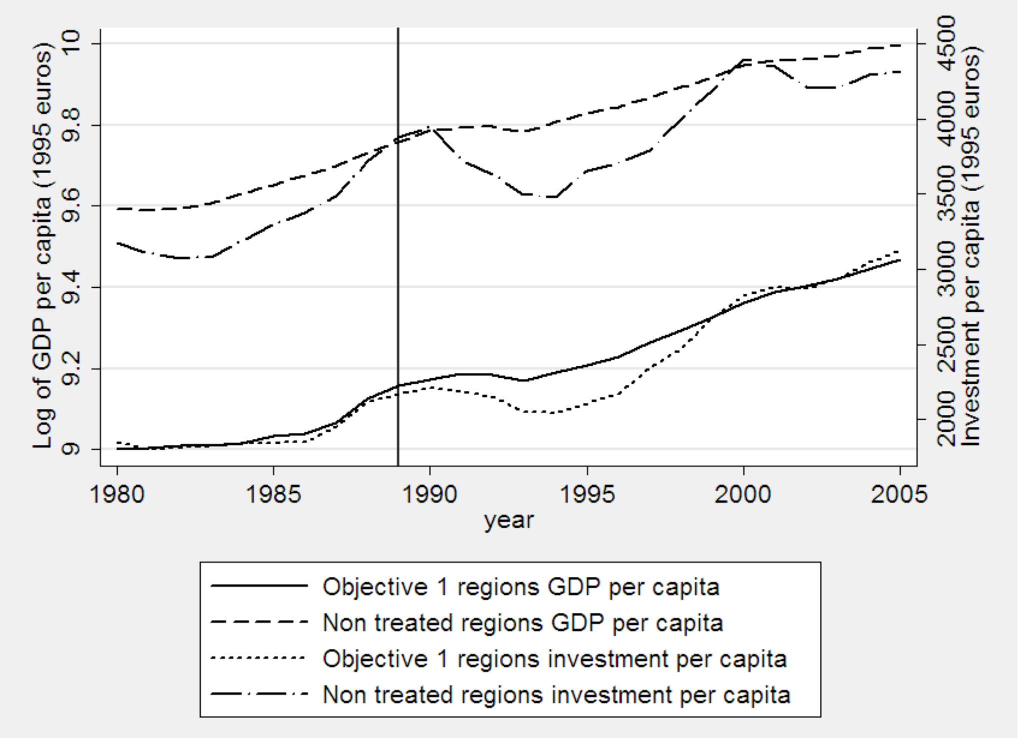

Figure 1: Dynamics of GDP and investment per capita average by treatment status (authors' calculation, Cambridge Econometrics database)

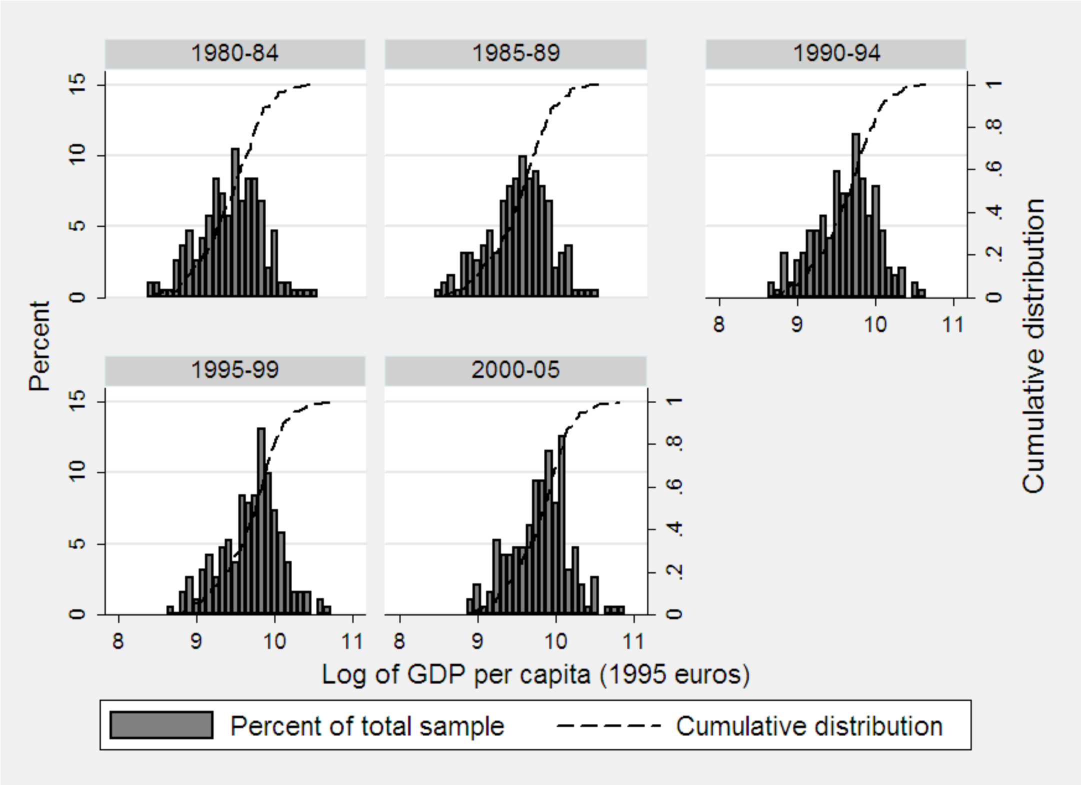

Figure 2: Evolution of income distribution for the set of 191 studied regions (authors' calculation, Cambridge Econometrics database)

Figure 2 depicts the cross-sectional distributions of EU-14 regions for each period, revealing a multimodal structure of the distributions, less obvious at the end of the period. The informations depicted on this Figure suggest a very slow process of catching up of the poorest regions with the richer ones and a process of shifting away of a small group of very rich regions

Figure 3: Annual versus Before/After intervention income and growth by treatment status (authors' calculation, Cambridge Econometrics database)

Figure 3 highlights the difference in GDP p.c. and growth rate between Objective 1 regions and Non treated regions before and after the implementation of the cohesion policy. Even if the GDP per capita gap remains only slowly decreases Before and After intervention, one can see that the growth rates are slightly different and much more important for treated regions and this especially at the end of the period considered, suggesting that the catching up process is at work within a conditional convergence framework : regions seems to converge towards country specific steady-state GDP levels. This process is liable to spread out first in the neighbour regions and then disseminate over the whole European space. Observed spatial correlations highlight an obvious spatial dimension of regional convergence (see Figure 4).

Figure 4 graphs the GDP-per-capita 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.

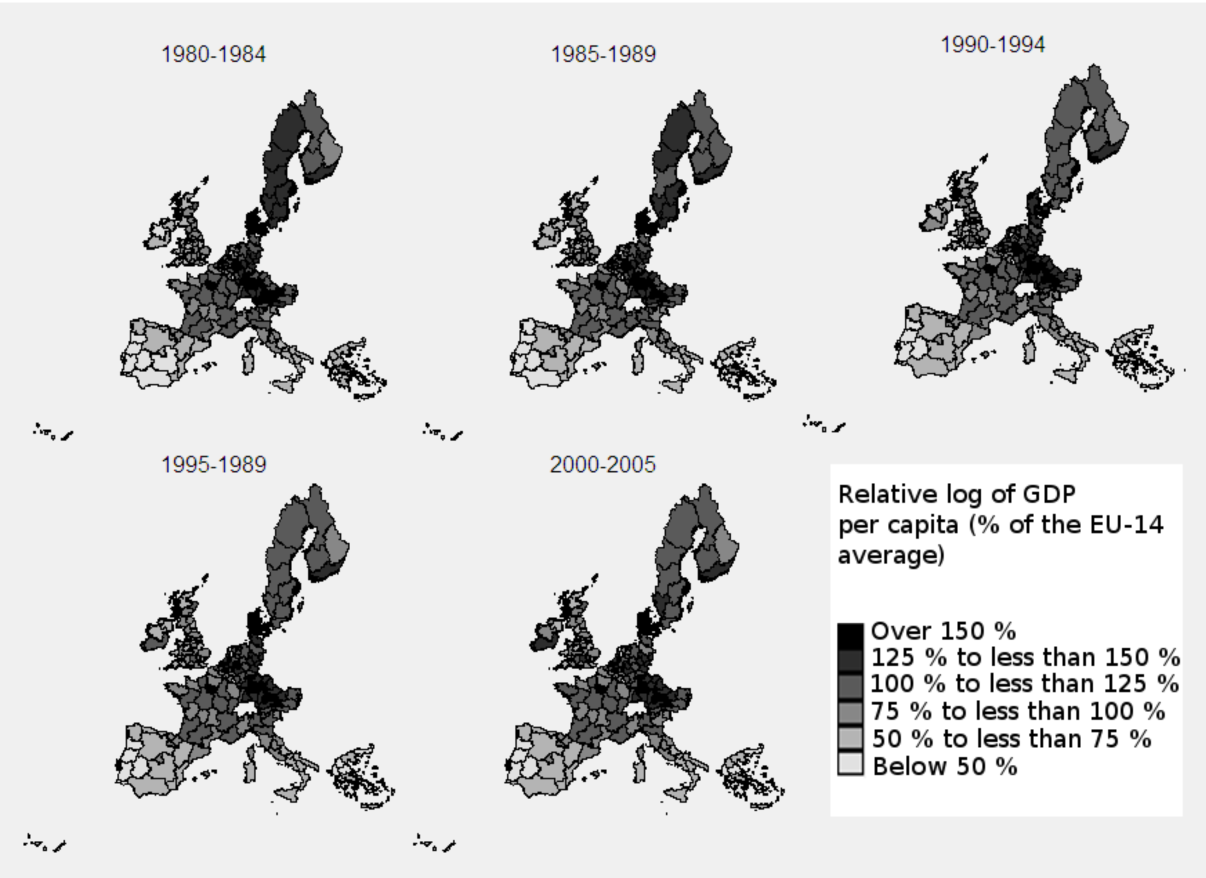

Figure 4: Geographic pattern of the GDP per capita relative to the EU-14 average GDP level for the 5 periods (authors' calculation, Cambridge Econometrics database)

4. Estimation results: impact analysis of structural funds on regional convergence

Table 2 and 3 report respectively estimates of conditional convergence model over 1980-2005 period. This time lag has been split into 5 periods (1980-84, 1985-89, 1990-94, 1995-99 and 2000-05) that includes three different policy programs (1989-93, 1994-99 and 2000-06).

Because the presence of spatial dependencies into residuals leads to misleading estimates and tests, we will first analyse the spatial properties of the residuals, and then present the estimation results and validity tests.

Table 2: Results: estimation of model (1) to (3 with: traditional to dynamic panel data estimators

	POLS			LSDV			GMM-DIFF		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
$ln(\frac{Y_{i,t-1}}{pop_{i,t-1}})$	0.893***	0.876***	0.877***	0.700***	0.645***	0.645***	0.790***	0.818***	0.924***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.06)	(0.04)	(0.06)
$ln(\frac{I_{i,t}}{pop_{i,t}})$	0.095***	0.096***	0.094***	0.240***	0.244***	0.244***	0.188***	0.174***	0.207***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.04)	(0.02)	(0.03)
$ln(\frac{pop_{i,t}}{pop_{i,t-1}})$	0.005***	0.006***	0.006***	-0.000	-0.002	-0.002	-0.001	-0.001	0.003
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
T_t		-0.009	-0.009		0.019*	0.019*		-0.004	-0.028*
		(0.01)	(0.01)		(0.01)	(0.01)		(0.01)	(0.01)
P_i		-0.041***	-0.041***						
		(0.01)	(0.01)						
$P_i.T_t$		0.025*	0.017		0.005	0.005		0.010	0.086**
		(0.01)	(0.01)		(0.01)	(0.01)		(0.01)	(0.03)
$D_{i,t-1}.ln(\frac{Y_{i,t-1}}{pop_{i,t-1}})$			0.001			0.001			-0.016*
			(0.00)			(0.00)			(0.01)
constant	0.330***	0.501***	0.502***	1.039***	1.511***	1.514***			
	0.06)	(0.09)	(0.09)	(0.15)	(0.22)	(0.23)			
N	598	598	598	598	598	598	405	405	405
r2	0.983	0.983	0.983	0.912	0.915	0.915			
ar1							-4.390	-4.549	-4.165
ar1(p.v.)							(0.000)	(0.000)	(0.000)
ar2							0.941	1.029	1.542
ar2(p.v.)							(0.347)	(0.304)	(0.123)
hansen							20.635	18.656	6.485
hansen (p.v.)							(0.000)	(0.000)	(0.011)
hansen-diff							25.11	11.85	9.57
hansen-diff(p.v.)							(0.00)	(0.00)	(0.02)
nb of instruments							5	7	7
Z_{moran}									
Z_{e85-89}							9.612	9.478	13.161
Z_{e85-89}(p.v.)							(0.00)	(0.00)	(0.00)
Z_{e90-94}							8.792	9.273	12.130
Z_{e90-94}(p.v.)							(0.00)	(0.00)	(0.00)
Z_{e95-99}							15.221	15.726	19.800
Z_{e95-99}(p.v.)							(0.00)	(0.00)	(0.00)
Z_{e00-04}							6.982	7.516	16.237
Z_{e00-04}(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* ($\ln\left(\frac{Y_{i,1}}{pop_{i,1}}\right)$, $\ln\left(\frac{Y_{i,2}}{pop_{i,2}}\right)$,..., $\ln\left(\frac{Y_{i,T-2}}{pop_{i,T-2}}\right)$) *and the overall set of other explanatory variables (considered as exogenous) .*

4.1. Spatial correlation of residuals

Spatial error model and spatial lag model are two different approaches to address the issue of spatial dependencies. Spatial dependencies of residuals occur when the spatial lag term is omitted in the regression, or if the spatial autocorrelated component is ignored. In the first case, the estimates are biased (omitted variable bias) and all statistical inferences are invalid, whereas ignoring the error spatial autocorrelation, in the second case, leads to inefficient estimates (unbiased parameter estimates but overall statistical tests are invalid). 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 models (1) to (3) that ignore spatial lag variables (Table 2). This result confirms the presence of spatial effects in the European regional convergence. As expected, the introduction of a spatial lag term in the models (4) to (6) reduces considerably the spatial correlation of residuals. Moran's test fail to reject the null hypothesis of no spatial correlation for 1990-1994, 2000-05 and for 1985-89 respectively with a 10% and 5% significance level (Table 3). We therefore proceed with the spatial lag model.

These results confirm the presence of spatial dependencies which leads to inconsistent results and inefficient validity tests in model (1) to (6). Hence, we are not looking on the remaining tests in Table 2 and we will concentrate on Table 3 tests results.

4.2. Validity tests

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). 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 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.

4.3. Results

Within the framework of a dynamic panel specification that considers the spatial dimension of GDP per capita, we find empirical evidence that cohesion policy fosters the endogenous development of Objective1 regions in Europe and contributes to narrow GDP disparities over EU regions. The results are presented in line with the specifications already exposed in the previous section: we examine first the dynamic model estimates (model (1) to (3,) Table 2) that we got by pooled least squares (POLS), least squares dummy variables (LSDV) and GMM estimations. Then, we examine the estimates without residual spatial dependencies (model (4) to (6), Table 3).

Table 3: Results of estimation of models (4) to (6): dynamic spatial GMM

	S-GMM-DIFF		
	(4)	(5)	(6)
$ln(\frac{Y_{i,t-1}}{pop_{i,t-1}})$	0.468***	0.429***	0.440***
	(0.11)	(0.11)	(0.12)
$ln(\frac{I_{i,t}}{pop_{i,t}})$	0.175***	0.177***	0.197***
	(0.03)	(0.03)	(0.04)
$ln(\frac{pop_{i,t}}{pop_{i,t-1}})$	0.001	0.001	0.002
	(0.00)	(0.00)	(0.00)
$Wln(\frac{Y_{i,t}}{pop_{i,t}})$	0.355**	0.385**	0.405**
	(0.12)	(0.12)	(0.13)
T_t		0.003	-0.005
		(0.01)	(0.01)
P_i			
$P_i.T_t$		0.001	0.032
		(0.01)	(0.02)
$D_{i,t-1}.ln(\frac{Y_{i,t-1}}{pop_{i,t-1}})$			-0.007*
			(0.00)
N	405	405	405
ar1	-1.924	-1.734	-2.122
ar1(p.v.)	(0.054)	(0.083)	(0.034)
ar2	0.878	0.806	0.790
ar2(p.v.)	(0.380)	(0.420)	(0.430)
hansen	13.109	13.386	10.722
hansen(p.v.)	(0.158)	(0.203)	(0.295)
hansen-diff	30.41	16.10	31.70
hansen-diff(p.v.)	(0.00)	(0.00)	(0.00)
nb of instruments	13	16	16
Z_{moran}			
Z_{e85-89}	1.496	1.209	1.680
Z_{e85-89}(p.v.)	(0.07)	(0.113)	(0.046)
Z_{e90-94}	-0.313	-0.646	0.561
Z_{e90-94}(p.v.)	(0.377)	(0.259)	(0.287)
Z_{e95-99}	0.988	0.373	1.691
Z_{e95-99}(p.v.)	(0.162)	(0.355)	(0.045)
Z_{e00-04}	-0.707	-0.876	-0.477
Z_{e00-04}(p.v.)	(0.240)	(0.191)	(0.317)

*Notes: *, ** 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 (* $\ln\left(\frac{Y_{i,1}}{pop_{i,1}}\right)$, $\ln\left(\frac{Y_{i,2}}{pop_{i,2}}\right)$,..., $\ln\left(\frac{Y_{i,T-2}}{pop_{i,T-2}}\right)$; $W\ln\left(\frac{Y_{i,1}}{pop_{i,1}}\right)$, $W\ln\left(\frac{Y_{i,2}}{pop_{i,2}}\right)$, ..., $W\ln\left(\frac{Y_{i,T-2}}{pop_{i,T-2}}\right)$ *) and the overall set of other explanatory variables (considered as exogenous) .*

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)[8]. For Model (1) and (2), LSDV values are around 0.7 (0.70 and 0.64) and grow to values of around 0.89 (0.89 and 0.88) when the POLS estimator is used. As expected, for these 2 models, the estimated autoregressive parameters, lying close to 0.8 (0.79 and 0.81) fall between these bounds. Unfortunately, this is not the case in model (3) where the GMM autoregressive parameter (0.92) exceeds POLS coefficient (0.88) probably due to a misspecification bias (omitted variables). These results are in line our previous residual spatial correlation analysis (section 4.1): it is important to include a spatially lagged variable in our specification. Although Table 2 provides biased and inefficient results, it makes a benchmark for dynamic spatial panel data model results.

The estimated autoregressive parameter are close to those obtained in other studies (Caselli et al., 1996). Its value points out that there is a significant European regional convergence, conditionally to investment per capita (which have a significant positive impact on regional development). The non significant impact of demographic growth suggests that the evolution of labour force does not affect regional development. Last, we can note that in model (3) the treatment effect has a very small (-0.016) but significant effect on Objective 1 regions convergence.

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 (0.47) in model (4) drops sharply regarding to model (1) indicating that we measure a faster regional convergence when we consider the impact of neighbouring income on regional development. Spatial lagged income coefficient (0.35) suggests a strong significant impact of spill over effect between European regions on their development dynamics. 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. The introduction of a spatial

[8] 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).

lag does not affect the impact of investment per capita spending on regional development.($\frac{I_{i,t}}{pop_{i,t}}$ coefficients keep significant and values around 0.18). The demographic growth rate has still no still significant impact in this specification.

Model (6) results provide evidence of policy impact on Objective 1 convergence. The $D_{i,t-1}.\ln\left(\frac{Y_{i,t-1}}{pop_{i,t-1}}\right)$ coefficient assesses weak but significant impact of O1 programs on European convergence of treated regions.

It is now possible to use the estimation results to depict the impact of the structural funds on the income for the Objective1 regions. Results are depicted in Figure 5 for models (3) and (6) for the periods 85-89 (0, before the policy applies), 90-94 (1), 95-99 (2) and 2000-2005 (3). Figure 5 depicts the log difference for Objective1 regions between the predicted value of GDP p.c. when the policy applies and its counterfactual. If the policy impact pattern is similar between the two estimates, model (3) provides an upward biased measure of its impact because this model doses not consider a spatial autoregressive term. Moreover, the figure 5 highlights that policy has affected Objective 1 regions convergence for the first two programming periods (periods 1 and 2, 1990-94 and 1995-99). Although European Union experienced a slow income growth during the first programming period, policy has supported Objective 1 regions development. Last, the difference between O1 predicted income and its counterfactual is constant for the third period, between 1999 and 2005 (around 0.25).

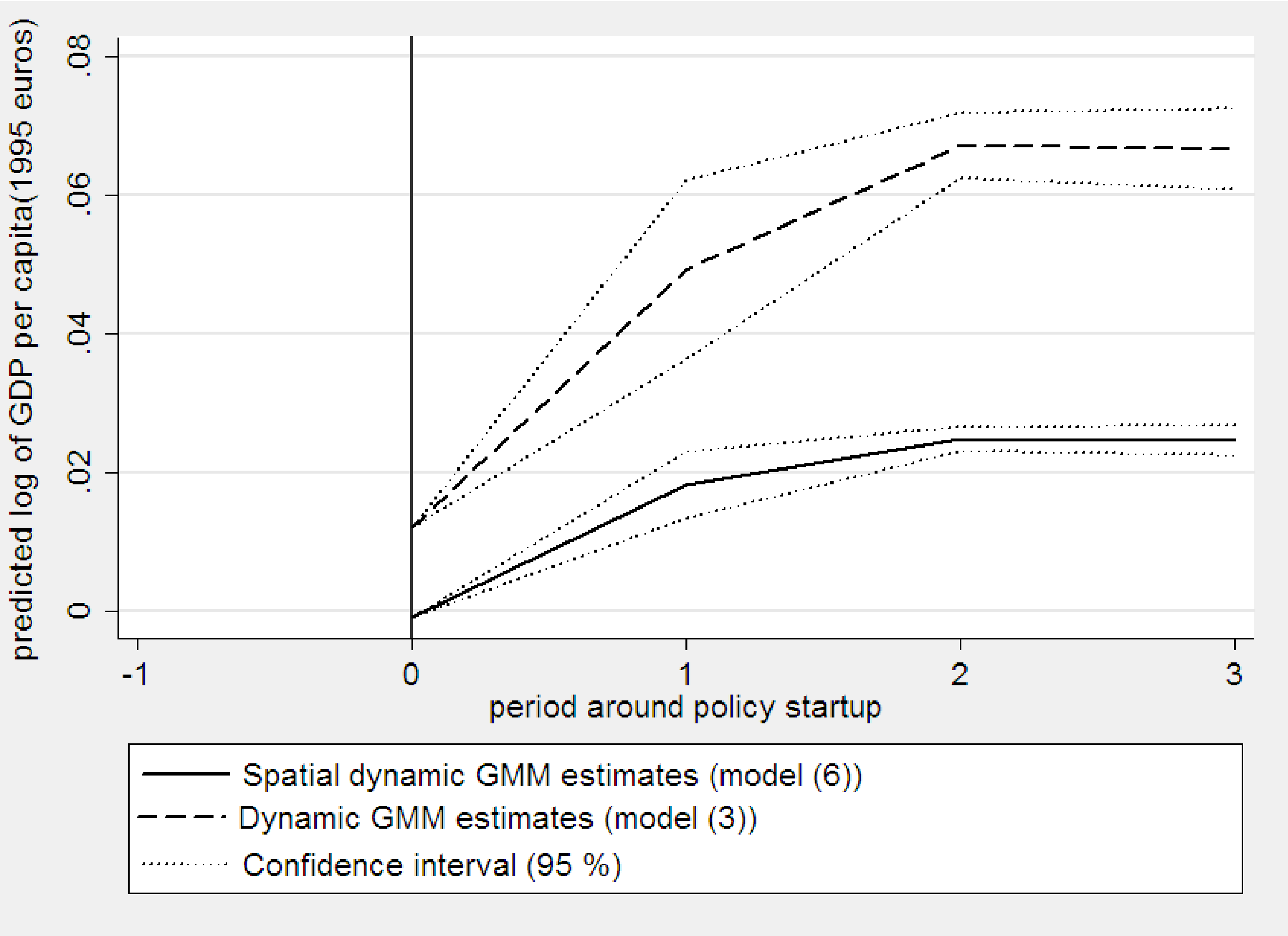

Figure 5: Average of policy impact on convergence of Objective 1 regions for the periods 85-89 (0, before the policy applies), 90-94 (1), 95-99 (2) and 2000-2005 (3 (Authors' calculation, Cambridge Econometrics data)

5. Conclusion

The aim of this paper is to investigate empirically the impact of cohesion policy on European regional convergence. Using a dynamic panel dataset of 191 regions over the period 1980-2005, including information before the policy applies, we extent the current literature by considering together spatial dependencies and impact analysis inside a dynamic panel specification. The broadness of our datasets enables such consideration.

We focussed on Objective1 regions, that receive two third of the structural funds. Because development projects have been designed with a bottom-up approach and we assumed that not all drivers are activated in the same way in all Objective1 regions. As such, we retained a specification that does not explicit the way the policy can foster regional development. Of course, the growth impact does not appear immediately and we considered a 5 years time lag in our analysis.

Within the framework of a dynamic panel specification that considers the spatial dimension of GDP per capita, we find empirical evidence that cohesion policy fosters the endogenous development of Objective1 regions in Europe and contributes to narrow GDP disparities over EU regions. Moreover, our results confirm that the cohesion policy, that aims at counterbalancing the effects of GDP concentration over the richest regions, attains this objective as considering for the spatial dimension of the panel still leads to a significant effect of the structural funds on regional convergence.

Such insights confirm that the bottom-up design of projects, along with the involvement of regional, national and EU actors in the design, implementation and evaluation of the programs, has the capacity of fostering the endogenous development potential of the lagging regions in Europe. As such, it may not be that important to concentrate on the specific drivers in each region in an overall convergence analysis.

Analysing the spatial dimension of the panel data we find that regional spillover do have an impact on regional development. The cohesion policy counterbalances a negative effect on regional development when the richest regions concentrate income and activities. It is however our opinion that improving regional spillovers can contribute to foster the endogenous development of regional clusters, as has been demonstrated with Interreg programs. The last point stresses that the cohesion policy is implemented along with other EU policies (like agricultural policies, industrial regulations) that can favour or hamper the effects of this policy. Extending our analysis towards national redistributive effects, national pensioning strategies and regional clustering can help designing more efficient policies toward regional development.

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 Développment,* 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.

ARBIA, G. & PIRAS, G. (2005) Convergence in per-capita GDP across European regions using panel data models extended to spatial autocorrelation effects *ISAE Working Paper No. 51*.

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.

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.

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-dat approach. *Regional Studies,* 42**,** 159-173.

EUROPEAN COMMISSION (2008) Growing regions, growing Europe - Fifth progress report on economic and social cohesion - communication from the Commission to the European Parliament and the Council (COM(2008) 371). European commission, Directorate general for regional policy.

ISLAM, N. (1995) Growth Empirics: A Panel Data Approach. *The Quarterly Journal of Economics,* 110**,** 1127-1170.

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.

OECD (2006) *The new rural paradigm: policies and governance,* Paris, OECD Publishing.

RAVALLION, M. (2005) Evaluating Anti-Poverty Programs. IN EVENSON, R. E., SCHULTZ, T. P. & STRAUSS, J. (Eds.) *Handbook of Development Economics.* Amsterdam, North Holland.

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.

TODD, P. E. (2006) Evaluating social program placement and selection of the treated. IN EVENSON, E. E. & SCHULTZ, T. P. (Eds.) *Handbook of Agricultural Economics.* Amsterdam, North Holland.

Appendix A

Name	Objective 1	Code	Name	Objective 1	Code	Name	Objective 1
Cantabria	1989-2006	IE02	Southern and Eastern	1989-2004	UKD4	Lancashire	
Pais Vasco		ITC1	Piemonte		UKD5	Merseyside	1994-2006
Comunidad Foral de Navarra		ITC2	Valle d'Aosta		UKE1	East Riding and North Lincolnshire	
La Rioja		ITC3	Liguria		UKE2	North Yorkshire	
Aragón		ITC4	Lombardia		UKE3	South Yorkshire	2000-06
Comunidad de Madrid		ITD1	Provincia Autonoma Bolzano-Bozen		UKE4	West Yorkshire	
Castilla y León	1989-2006	ITD2	Provincia Autonoma Trento		UKF1	Derbyshire and Nottinghamshire	
Castilla-la Mancha	1989-2006	ITD3	Veneto		UKF2	Leicestershire, Rutland and Northants	
Extremadura	1989-2006	ITD4	Friuli-Venezia Giulia		UKF3	Lincolnshire	
Cataluña		ITD5	Emilia-Romagna		UKG1	Herefordshire, Worcestershire and Warks	
Comunidad Valenciana	1989-2004	ITE1	Toscana		UKG2	Shropshire and Staffordshire	
Illes Balears		ITE2	Umbria		UKG3	West Midlands	
Andalucia	1989-2006	ITE3	Marche		UKH1	East Anglia	
Región de Murcia	1989-2006	ITE4	Lazio		UKH2	Bedfordshire, Hertfordshire	
Canarias (ES)	1989-2006	ITF1	Abruzzo	1989-96	UKH3	Essex	
Itä-Suomi	1995-2006	ITF2	Molise	1989-2006	UKI1	Inner London	
Etelä-Suomi		ITF3	Campania	1989-2006	UKI2	Outer London	
Länsi-Suomi		ITF4	Puglia	1989-2006	UKJ1	Berkshire, Bucks and Oxfordshire	
Pohjois-Suomi	1995-2006	ITF5	Basilicata	1989-2006	UKJ2	Surrey, East and West Sussex	
Île de France		ITF6	Calabria	1989-2006	UKJ3	Hampshire and Isle of Wight	
Champagne-Ardenne		ITG1	Sicilia	1989-2006	UKJ4	Kent	
Picardie		ITG2	Sardegna	1989-2006	UKK1	Gloucestershire, Wiltshire and North Somerset	
Haute-Normandie		NL11	Groningen		UKK2	Dorset and Somerset	
Centre		NL12	Friesland		UKK3	Cornwall and Isles of Scilly	2000-06
Basse-Normandie		NL13	Drenthe		UKK4	Devon	
Bourgogne		NL21	Overijssel		UKL1	West Wales and The Valleys	2000-06
Nord - Pas-de-Calais	1994-2004	NL22	Gelderland		UKL2	East Wales	
Lorraine		NL23	Flevoland	2000-06	UKM1	North Eastern Scotland	
Alsace		NL31	Utrecht		UKM2	Eastern Scotland	
Franche-Comté		NL32	Noord-Holland		UKM3	South Western Scotland	
Pays de la Loire		NL33	Zuid-Holland		UKM4	Highlands and Islands	1994-2006
Bretagne		NL34	Zeeland		UKN0	Northern Ireland	1989-2006
Poitou-Charentes		NL41	Noord-Brabant			Excluded regions (21)	
Aquitaine		NL42	Limburg (NL)		DE30	Berlin	
Midi-Pyrénées		PT11	Norte	1989-2006	DE41	Brandenburg - Nordost	1991-2006
Limousin		PT15	Algarve	1989-2006	DE42	Brandenburg - Südwest	1991-2006
Rhône-Alpes		PT16	Centro (PT)	1989-2006	DE80	Mecklenburg-Vorpommern	1991-2006
Auvergne		PT17	Lisboa	1989-2004	DED1	Chemnitz	
Languedoc-Roussillon		PT18	Alentejo	1989-2006	DED2	Dresden	
Provence-Alpes-Côte d'Azur		PT20	Região Autónoma dos Açores (	1989-2006	DED3	Leipzig	
Corse	1989-2004	SE01	Stockholm		DEE1	Dessau	
Anatoliki Makedonia, T	1989-2006	SE02	Östra Mellansverige		DEE2	Halle	
Kentriki Makedonia	1989-2006	SE04	Sydsverige		DEE3	Magdeburg	
Dytiki Makedonia	1989-2006	SE06	Norra Mellansverige		DEG0	Thüringen	1991-2006
Thessalia	1989-2006	SE07	Mellersta Norrland	1995-2006	ES63	Ciudad Autónoma de Ceuta (ES)	1994-2006
Ipeiros	1989-2006	SE08	Övre Norrland	1995-2006	ES64	Ciudad Autónoma de Melilla (ES)	1994-2006
Ionia Nisia	1989-2006	SE09	Småland med öarna		FI20	Åland	
Dytiki Ellada	1989-2006	SE0A	Västsverige		FR91	Guadeloupe (FR)	1989-2006
Sterea Ellada	1989-2006	UKC1	Tees Valley and Durham		FR92	Martinique (FR)	1989-2006
Peloponnisos	1989-2006	UKC2	Northumberland, Tyne and Wear		FR93	Guyane (FR)	1989-2006
Attiki	1989-2006	UKD1	Cumbria		FR94	Reunion (FR)	1989-2006
Kriti	1989-2006	UKD2	Cheshire		GR41	Voreio Aigaio	1989-2006
Border, Midlands and W	1989-2006	UKD3	Greater Manchester		GR42	Notio Aigaio	1989-2006
					PT30	Região Autónoma da Madeira (PT)	1989-2006

Appendix B. Spatial weight matrix specification

The spatial weight matrix is used to evaluate the covariance of characteristics across region locations. While a variety of weighting matrix may be constructed, in order to allow spatial interaction, the empirical literature chooses weights based on Euclidean distance or Contiguity between regions (Abreu et al., 2005).

Thus, we have chosen a geographical definition of neighbourhood based on Euclidean distance between regions in order to the W matrix. More 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 cutt-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.

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).