Identifying and Interpreting Convergence Clusters Across Europe

Luisa Corrado, Ron Martin and Melvyn Weeks

February 2004

CWPE 0414

Not to be quoted without permission

Identifying and Interpreting Convergence Clusters across Europe*

Luisa Corrado† Ron Martin‡ Melvyn Weeks§

23th January 2004

Abstract

In this paper we examine the spatial and temporal distribution of per capita income across Europe. We base our analysis on a cluster methodology which allows for an endogenous selection of regional clusters using a multivariate test for stationarity where the number and composition of clusters are determined by the application of pairwise tests of regional contrasts. To circumvent the problem of how to interpret the composition of resulting convergence clusters we construct a number of testable hypotheses based upon orderings consistent with the findings of recent studies on regional growth and convergence. We do this using a set of geographical, socio-demographic and political indicators measuring contiguity and institutional similarity, accessibility, specialisation, region specific levels of agglomeration and regional classification according to the European Union Structural Fund objectives. One of the contributions of our study is a method which facilitates the interpretation of the cluster outcomes on the basis of the factors identified above. Unlike previous studies, we present our results using a geographic representation of regions across Europe.

Key-Words: Regional Convergence, Pairwise Regional Comparison, New Economic Geography.

JEL: C51, R11, R15.

*We are grateful to Andrew Harvey, Gernot Dopplehoffer and Vasco Carvalho for their comments on the work. Acknowledgement: The authors acknowledge financial support from the Economic and Social Research Council, award L138251008. Corresponding authors: Luisa Corrado and Melvyn Weeks email:luisa.corrado@uniroma2.it, mw217@econ.cam.ac.uk

†Faculty of Economics, University of Rome Tor Vergata.

‡Department of Geography, University of Cambridge.

§Faculty of Economics and Politics, University of Cambridge.

1 Introduction

Since the mid-1980s long-term growth has re-appeared on economists' research agenda. An important stimulus for this revival has been the renewed interest in the empirics of growth, and especially in the evidence for long-term convergence of per capita output and incomes between nations. This empirical debate has in turn promoted the re-examination and re-orientation of growth theory. Standard neoclassical growth models (Solow (1956) and Swann (1956)) postulate that countries will converge to the same level of per capita income (output) in the long run, independently of initial conditions, as long as there are diminishing returns to capital and labour, and perfect diffusion of technological change (taken as exogenous). However, the prediction of the neoclassical model has appeared increasingly at odds with the lack of evidence for international convergence and the variable nature of the convergence process even within the industrialised group of countries (Abramovitz (1986); Boltho and Holtham (1992)).

Over the past two decades a so-called New Growth theory has emerged (see, for example, Romer (1986), Lucas (1988), Barro and Sala-i-Martin (1991)). This framework extends the canonical neoclassical model by allowing for increasing returns in the production function in order to determine the (endogenous) long-term growth rate. Although there exists a number of variants of this new endogenous growth theory, all permit a wider set of possibilities with regard to convergence behaviour.

Studies by Romer (1986), Lucas (1988) and Grossman and Helpman (1991), inter alia, suggest that under conditions of market competition and increasing returns to accumulable factors such as human capital, convergence cannot spontaneously take place. A number of other theoretical studies in the area of economic geography and regional economic theory have focussed on the tension between central and peripheral areas (see Fujita, Krugman, and Venables (1999), Fujita and Thisse (2001) and Krugman (1991)); on the effect of regional externalities, including congestion effects and increasing returns from agglomeration (Cheshire and Carbonaro (1989)); on the relationship between location and transportation costs (Limao and Venables (2001)); and on the role EU interventions that should allow depressed regions to strengthen their endowment of productive knowledge and of R&D and to catch up with the richer ones (Boldrin and Canova (2001)). Indeed, these new growth variants (including that by Barro and Sala-i-Martin (1995)) imply conditional convergence by allowing for different initial conditions between countries (institutions, economic structures, tastes, etc.); or club convergence between countries with similar structural conditions and/or amongst which there are technological spillovers (see Bertola (1993)). Thus, conditional convergence still implies different degrees of convergence with poorer countries growing faster than richer ones, when other relevant factors are controlled for.[1]

If we believe in the one-sector neoclassical growth models with exogenous technological change then we can use traditional unconditional β convergence methods. If on the other hand we believe that convergence should be found among regions that are relatively similar to each other in terms of economic, political or geographical factors, then clearly unconditional convergence methods are not appropriate, since all kind of countries are pooled together without accounting for common or country-specific effects. In this context a number of issues are central. First, if we consider the existence of multiple steady states (see, for example, Durlauf and Johnson (1992)) there is the question of how these states and their composition are identified. Two polar positions are based upon a priori identification of both the number and composition of regimes, versus a statistical approach which allows for endogenous deter-

[1] For an analysis of the methodological implications of the new growth evidence see Temple (1999).

mination via a clustering algorithm. Second, there is the question of scale, by which we mean that, for a given set of objectives, the appropriate size and heterogeneity of the cross-sectional unit will likely vary. This is particularly relevant in the study of convergence although, with the exception of the work of Boldrin and Canova (2001), it has been ignored in the literature.

Our aim in this paper is to extend this body of literature by providing a representation of the spatial distribution of regional and national per capita income across Europe. We base our analysis on the cluster method introduced by Hobijn and Franses (2000) which allows for an endogenous selection of regional clusters using a multivariate test for stationarity. The number and composition of clusters are determined by the application of pairwise tests of regional contrasts. Inter-regional interactions and co-dependence, which originate from the presence of increasing returns due to regional endowments and/or from the similarity of the institutional and political frameworks, produce multiple convergence paths. However, what is required to support the anecdotal evidence on the presence of these different poles of attraction is a more formal analysis of how these and their composition are determined. To circumvent the problem of how to interpret the composition of resulting convergence clusters we construct a number of testable hypotheses based upon orderings suggested by alternative conceptions of regional growth and convergence. We do this using a set of geographical, socio-demographic and political indicators measuring contiguity and institutional similarity, accessibility, specialisation, region specific levels of agglomeration and regional classification according to the European Union Structural Fund objectives. One of the contributions of our study is a method which facilitates the interpretation of the cluster outcomes on the basis of the factors identified above.

The paper is organised as follows. Section two introduces the existing approaches used to evaluate the extent and composition of convergence clubs. We also examine one of the most critical questions in the literature on convergence and growth: the appropriate level of aggregation of cross-sectional units, both at the geographical and at the industrial level. Section three describes the cluster methodology adopted and introduces a time-varying version of the cluster method. By allowing the cluster composition to be time-varying we are able to examine how the degree of convergence and the cluster composition varies over time. Section four describes the data. Sections five and six comment on the cluster outcomes, and test the outcomes against one or more hypothesised patterns.

2 Identifying and Interpreting Convergence Clubs

The statistical methods usually employed in the literature to detect convergence clubs are typically centered around a cluster analysis of the cross-sectional units. In this context two fundamental areas of enquiry are how these states and their composition are identified. Two extreme positions are based upon (i) a priori identification of the possible orderings of regions and the maximum number of convergence clubs (Canova (1999)); and (ii) a statistical approach centered around a clustering algorithm (see Hobijn and Franses (2000)). The approach of Hobijn and Franses (2000) tests whether countries are converging absolutely to the same long-run level or relatively to their long-run equilibria and allows for an endogenous selection of regional clusters using a multivariate test for stationarity. The number and composition of clusters are determined by the application of pairwise tests of regional differences in the logarithm of per capita Gross Value Added (GVA). The principal drawback of this approach follows immediately from one of its advantages, namely that the clustering model uses no conditioning variables, and simply uses information contained in per capita differences in real gross domestic product recorded at either the country or regional level. In this respect, although the approach is not dependent upon the choice of conditioning variables, and attendant problems of misspecification, there is a real difficulty of interpretation. For example, the basic output of this approach consists of recording the number of clusters found and a classification of regions by clusters. Obviously as the number of regions increases, it becomes increasingly difficult to interpret such outcomes. Hobijn and Franses acknowledge this fact, noting that beyond making references on the effect of country and contiguity (which is observed), there is no additional information to evaluate the composition of the cluster groups.

In this paper we utilise additional information which enables us to characterise regions by, for example, geographical, socio-demographic and policy indicators. The use of these quasi-fixed effects allows us to determine to what extent our estimated cluster pattern is consistent with one or more hypothesised clustering schemes. For example, using data on geographical peripherality, a categorical indicator which classifies regions location on a 10 point scale[2], we are able to confront the observed cluster outcomes with a hypothetical pattern generated by a world in which regions cluster solely on the basis of a core periphery dichotomy. Similarly, we construct hypothetical cluster patterns based on the settlement structure[3] and confront these with the cluster outcomes to see whether increasing urban agglomeration can explain the clustering outcome in certain sectors. We also compare the observed cluster outcomes and the set of policy instruments used by the European Community to promote the development and structural adjustment of European Community regions. For example, underdeveloped regions in Southern Europe that have been targeted under Objective 1 of the reformed European Structural Funds (regions whose GDP per capita is less than 75% of the Community average), are typically peripheral regions with little industry. The priorities for development in these areas have included the promotion of the production sector, modernising infrastructure and encouraging research (Martin and Tyler (2000)). Our analysis will reveal whether the externality-inducing policies implemented by the European Community have changed the profile of the temporal income distribution in these regions.

By considering two time periods we are also able to formally determine whether the effect of these quasi-fixed factors on the degree of convergence and cluster composition is constant. This information represents the most useful output of our study. For example, in the case of the manufacturing sector we might postulate that given increasing globalisation and increasing trade links, the effect of geographical attributes such as country membership and peripherality will fall over time. This represents an extremely useful information from a policy perspective, in the sense of assessing whether the set of region-specific quasi-fixed factors, which help determine a regions' performance (in terms of real per capital GVA), are changing over time.

Common to a number of existing approaches which have been used to evaluate the extent and composition of convergence clubs in models of economic growth, is the implicit recognition that economic theory is relatively impotent in both determining the number and composition of clubs. Durlauf and Johnson (1995) use a regression tree approach to locate different regimes

[2] Using this indicator European regions are classified according to the level of accessibility with respect to core regions. For example, Belgium and regions belonging to the Western and the Southern part of the Netherlands have the highest level of accessibility, whereas regions belonging to the Southern part of Europe the lowest.

[3] Using this indicator European regions are classified according to the number of inhabitants and population density. For example, regions in the North-Eastern and North-Western part of Austria, the region Wallonne in Belgium, the Central and Southern part of Spain and France and the Northern part of Italy have the highest level of population agglomeration.

utilising conditioning information from a set of control variables. The procedure approximates the growth process as the sum of piecewise linear functions where observations are grouped by initial conditions. Using cross-country data over the period 1960-1985, the authors discretize the support over a number of variables, specifically initial output and literacy rates. With each variable then written as the union of a finite number of intervals, an algorithm is employed alongside a standard OLS model to determine the optimal first split over this set of variables. Given that they find the first split is determined by output, the authors interpret this as suggestive of the fact that output dominates literacy in locating multiple convergence clubs. An obvious disadvantage of this approach is that although the use of conditioning information provides automatic and model consistent information for interpreting the resultant clusters, the methodology is reliant on both a correct identification of the mean equation *and* the set of variables controlling the clustering of regions.

Canova (1999) also points out the problem of data deficiencies in utilising a similar approach to search for convergence clubs in European regional data[4] Using a predictive density framework, the basic premise is that within a given cross-section there exists clustering of units of unknown form. In this respect both the number of clusters and composition thereof are treated as unknown 'parameters', analogous in part, to the problem of locating endogenous structural breaks in a time-series. Unlike standard tests for determining the number of heterogenous groups in a cross-section, intra-group heterogeneity of parameters is permitted. In fact, instead of assuming that model parameters are the same for all units, Canova (1999) allows the distribution within groups to differ. He finds that there are heterogeneities in European regional per-capita income and a tendency of the steady-state distribution to cluster around four poles of attractions when ordered according to the initial conditions of income per-capita.

The major drawback of this methodology, common also to the approach of Bernard and Durlauf (1995), is that the search for cross-sectional breakpoints is based upon a priori ordering of the data. For example, if we wish to emphasise the importance of location, then we would order regions on the basis of geographical proximity. In addition the maximum number of clusters, or cross-sectional break points, must be chosen a priori; this may generate substantial bias when dealing with disaggregate regional data where there may be a large number of poles of attraction, exceeding the imposed number of groupings.

2.1 A Question of Aggregation

The question as to the appropriate threshold size for geographical units over which it is sensible to test for economic convergence has been largely ignored in the literature. There is some evidence that the degree of convergence varies according to the scale at which regional differences (contrasts) are measured. For example, it is possible to find convergence at one spatial level, but a lack of convergence (or even divergence) at another. This is not just a statistical (aggregation) issue, since it raises the basic question of how economic convergence/divergence growth processes operate at different geographical and sector levels (see Martin and Sunley (1998)). For example, in a study on cross-province convergence for China, Jian, Sachs, and Warner (1996) find that divergence in the post reform period is entirely explained by the variance *between* regions defined as the group of coastal and interior provinces. However, *within* both of these regions there was no finding of divergence. A related issue is that the extent of regional differences at different scales may not be stable over time. In the UK, for

[4] In this study a convergence club represents a group of regions with each member possessing distinct steady states but demonstrating a tendency to cluster relative to other groups.

example, the between-region contribution to the total variance of unemployment rates across regions has steadily declined over the past 15 years, while the intra-region contribution has remained more or less constant (Gregg and Wadsworth (1999)). Whilst the between-region component exceeded the within region component in the 1970s and first half of the 1980s, since the mid-1980s the within-region component has exceeded the between-region component.

In an examination of the extent of inequality and convergence across Europe, the question of geographical scale is obviously central. Boldrin and Canova (2001) criticise the European Commission for utilising inappropriate regional units (the so-called Nomenclature of Statistical Territorial Units or NUTS). The principal measure for their comments is that NUTS1, NUTS2 and NUTS3 regions are neither uniformly large or sufficiently heterogeneous such that a finding of income divergence across regions cannot unequivocally be taken as evidence for the existence of endogenous, cumulative growth processes. In fact, the smaller the geographical scale, the more incomplete and fragmented is the statistical information we can get. These difficulties become more severe if we further disaggregate the information among industries and sector of productions. In conducting our analysis at the NUTS1 level we achieve a compromise between the need for a reliable set of information at a regional level which is sufficiently homogeneous and the need of moving beyond national borders to depict the true process of convergence.

The question of scale is also important in terms of the appropriate level of aggregation of productivity measurement. Bernard and Jones (1996) pose the question of whether trends in *aggregate* productivity are also revealed at the industry level. Relative to a finding of convergence at the aggregate national output level, the authors find that whereas the manufacturing sector has not exhibited signs of convergence, the service sector shows strong evidence of convergence. One possible explanation is that international spillovers, associated mostly with manufacturing, may not be contributing substantially to convergence either through capital accumulation or technological transfer. In fact, in a world with specialisation and spillovers, the non-tradable sectors will behave very much like an aggregate growth model and per capita output will converge over time as the technological diffusion process spreads. As a result in the service sector factor productivity will most probably converge since public services are invariant across countries and the information and communications based technologies used to offer the same services are potentially similar. In contrast within the manufacturing sector comparative advantage leads to specialisation, and since different countries or regions produce different goods, there is no reason to expect convergence in multifactor productivity.

Both the choice of the geographical scale and the sectoral level of aggregation are crucial since the choice of the wrong level of productivity aggregation and/or the choice of the wrong geographical scale may lead to the misleading conclusions in terms of cluster identification and composition. We believe that focussing on convergence of regional (NUTS1) per capita GVA at the sector level goes some way to addressing the aggregation issues highlighted above.

3 A Cluster Methodology

Before applying the cluster algorithm, we first consider the appropriate mapping from the economic to the statistical hypothesis. If we observe that two economies have converged then we might say that the difference between per capita income is stationary over the sample period considered, if starting conditions are unimportant. This variant of the economic concept of convergence has a statistical analogue in the notion of a testing for a stationary difference. In this case the number, N_c, and composition of clusters are determined by the application of tests of regional *contrasts* of the form $y_{it} - y_{jt}\ \forall i, j \in F$, where $i, j = 1, ..., R$ indexes regions, t is the time index, F denotes the set of regions, and y denotes the logarithm of per capita GVA. We define a benchmark alternative hypothesis of no convergence as consistent with the outcome $N_c = R$, namely where each of the geographical units coincides with a single cluster. Given that the likelihood of this alternative is decreasing in the number R, the null hypothesis of convergence is unlikely to be rejected. We are then left with the problem of interpreting the *degree* of convergence.

The notion of *absolute* convergence implies that, independent of the current income levels, regions i and j converge to the same income levels. Hence, two regions are perfectly converging if regions i and j are converging to the same level of output, namely if:

$$\lim_{s\to\infty} E(y_{i,t+s} - y_{j,t+s} \mid I_t) = 0 \qquad \forall i \neq j \in F \tag{1}$$

Asymptotic *relative* convergence implies that the difference between per capita income for i and j converges to a finite constant. Two regions are converging relatively if

$$\lim_{s\to\infty} E(y_{i,t+s} - y_{j,t+s} \mid I_t) = \mu_{i,j} \quad \forall i \neq j \in F, \tag{2}$$

where $\mu_{i,j}$ denotes a specific mean difference for regions i and j. We denote asymptotic perfect and asymptotic relative convergence respectively by τ_0^{ij} and τ_μ^{ij}.

The formation of clusters is described by the following algorithm. The algorithm is initialized by associating the R regions in F with N_c clusters. Pairwise tests of the null hypothesis of zero mean stationarity are conducted for all i, j pairs in F. For $p_0^{ij} > p_{\min}$,[5] where p_0^{ij} denote the empirical p-values, we do not reject the null hypothesis of convergence for regions i and j. We repeat this procedure for all pairs of regions and collect all empirical p-values in the vector $\overset{\wedge}{\mathbf{p}}$. The first cluster, say G_1, is formed by selecting the pair of regions, say l and k such that $p_0^{lk} = \max\limits_{i,j\in F}(\overset{\wedge}{\mathbf{p}})$.

We now repeat this process, where for example, in the second iteration the set of pairwise comparisons over F will include G_1. We iterate until the condition $p_0^{ij} > p_{\min}$ is violated $\forall\ i, j$ pairs. Note that if $p_0^{ij} < p_{\min}\ \forall\ i, j \in F$ we reject asymptotic perfect convergence, and apply the same algorithm to the less restrictive test of asymptotic *relative* convergence.

3.1 An Alternative Time-Varying Framework.

In general the use of stationarity tests requires that the data under analysis is near its long-run equilibria given the assumption that the sample moments of the data accurately approximate limiting moments. Hence, if the economies are in transition towards steady-state or start diverging (as it is implied by the cross-section approach) then the series will not satisfy the property of stationary output differences.

To capture this type of cluster dynamics we apply a time-varying stationarity test. Specifically we utilise a n-years rolling window that shifts progressively until the end of the sample period, T, is reached. Regions i and j are perfectly converging if

$$\lim_{t\to T-n} E(y_{i,t+n} - y_{j,t+n} \mid I_t) = 0 \qquad \forall i \neq j \in F \tag{3}$$

[5] The cluster algorithm requires the choice of a nominal critical p-value, $p_{\min} \in (0, 1)$ which defines the significance level for the tests.

while they are converging relatively if

$$\lim_{t \to T-n} E(y_{i,t+n} - y_{j,t+n} \mid I_t) = \mu_{i,j} \qquad \forall i \neq j \in F. \tag{4}$$

The above argument implies that the cluster composition may not be time invariant in the sense that some countries may exhibit convergence until a certain point of the sample period considered, and divergence thereafter. This approach is particularly useful for policy analysis since it gives a richer set of information on the temporal distribution and composition of the convergence clubs; its main drawback is the short time-horizon which affects the size of the test and the cluster outcome.[6]

4 Data

EUROSTAT has established an administrative map of the European Union called NUTS (Nomenclature of Statistical Territorial Units). The present NUTS nomenclature subdivides the economic territory of the 15 countries of the European Union using three regional and two local levels. The three regional levels are: NUTS3, consisting of 1031 regions; NUTS2, consisting of 206 regions; and NUTS1 consisting of 77 regions. NUTS0 represents the delineation at the national level and comprises France, Italy, Spain, UK, Ireland, Austria, Netherlands, Belgium, Luxemburg, Sweden, Norway, Portugal, Greece, Finland, Denmark and West Germany.[7] A complete list of regions is given in Table 1. The corresponding map of the European regions is reported in Figure 1.

We use regional data on Gross Value Added[8] per worker for the period 1975 to 1999 for agriculture, manufacturing and service. The service sector has been further sub-divided into market and non-market services: market services comprise distribution, retail, banking, and consultancy; non-market services comprise education, health and social work, defence and other government services. This disaggregation encompasses the information of more general aggregate indicators which are based upon measures of total factor productivity, thereby ignoring the possible differential contribution to convergence of different sectors.

The variables employed to interpret our cluster outcomes consist of a number of indicators which may be considered as fixed effects[9]; these are then divided into geographical, socio-demographic and political according to the ordinal classification reported in Table 2. The *geographical* effects comprise country-membership, the geographical location of the region, and its distance with respect to Central European regions. Country-membership defines the institutional setting; geographical location, which classifies regions on a 5 point scale, is a measure of contiguity and institutional similarity, whereas the periphery-core indicator is a measure of accessibility and classifies regions according to their relative distance with respect

[6] Appendix A describes the tests that we have performed to check for the robustness of the results.

[7] For Portugal, Luxemburg and Ireland, data is only available at the NUTS0 level. For Norway we have no data at the NUTS1 level. Time series data for the sample period considered are not available for East Germany, which is therefore excluded from the analysis.

[8] GVA has the comparative advantage with respect to GDP per capita of being the direct outcome of various factors that determine regional competitiveness. Regional data on GVA per-capita at the NUTS1 level for agriculture, manufacturing, market and non-market services, have been kindly supplied by Cambridge Econometrics, and are taken from their European Regional Database. All series have been converted to constant 1985 prices (ECU) using the purchasing power parity exchange rate.

[9] With the exception of population change which is averaged across the years 1991-1995 for all the NUTS1 regions.

to core regions. The *socio-demographic* effects are indicators of regional-urban agglomeration and group regions according to their settlement structure and population growth. So for example, according to the population indicator, a categorical indicator that classifies regions on a 5 point scale, regions in the North-Western part of Germany, Belgium, Central Spain and Northern Italy have the highest levels of population growth. For agriculture we also use an indicator of regional agricultural specialisation which groups regions on a 5 point scale according to the percentage of land utilisation under agriculture.

Finally regions are classified according to the degree and nature of public assistance, in terms of their designation under the specific EU Cohesion and Structural Fund objective assigned (*political* effect). More specifically we distinguish between non-assisted (Objective 0) regions, Objective 1 regions (underdeveloped regions with structural adjustment problems), and Objective 2 and 5 regions (regions affected by industrial decline and backward rural areas).

In sections 5, 6 and 7 we report our results. In section 5 we begin with an informal analysis of how both the extent and composition of convergence clubs within Europe differ both over time, and across different sectors. We note that, consistent with a number of existing studies (see, for example Bernard and Durlauf (1995)), our initial observations are largely descriptive, and the interpretation is generally linked to geographic proximity. In sections 6 and 7 we generate a number of hypothetical cluster patterns according to a number of alternative models of convergence processes, and in this regard partially circumvent the problems noted above.

5 The Observed Clusters

In Table 4 we report results based on test of asymptotic relative convergence[10] at the country level (NUTS0). The largest clusters in agriculture and manufacturing comprise four countries, whereas non-market services exhibit the highest degree of convergence with a five country cluster. This confirms the findings of Romer (1986), Lucas (1988) and Quah (1995) that convergence is easier to find in the service sector since most countries (and regions) tend to have similar types of basic market and non-market services.

Aggregate national level data may mask the extent of the convergence processes. Subsequently we also analyse the process of convergence at the regional level (NUTS1). Given the large number of EU regions we choose to present the results for asymptotic perfect and relative convergence in mapped form rather in tables. Clusters with the highest number of member regions are indicated with a darker shade on each map. Regions which belong to two-country clusters or do not cluster with any other region have no shading. In the key to the maps, the first number indicates the cluster size and the second letter denotes the different cluster membership.

The full sample results (1975-1999) for the four sectors are displayed in the top panel of Figures 2 to 5. In agriculture (top panel of Figure 2), we find a five-region cluster which comprises regions located in the North-West, and Eastern parts of England and in the South-Western part of Germany. Note also that the four regions located in Southern-Italy, South of Spain and Greece belong to the same cluster (4B). This confirms that agricultural regions with similar climate and technological endowments (Wichmann (1996)) tend to cluster together. A similar result is also present in Durlauf and Johnson (1992).

In the case of manufacturing (top panel of Figure 3) there is one five country cluster and in general we have less convergence than in the other sectors. This is also consistent with

[10] Our analysis focusses on the less restrictive concept of relative convergence.

the findings of Bernard and Jones (1996) who find little evidence of labour productivity[11] convergence in manufacturing. A higher degree of convergence is found for the service sector (top panels of Figures 4 and 5) where there are seven country clusters both for market and non-market services. It could be argued that the extent of convergence would be expected to be more prevalent in manufacturing than in services, because this sector is mainly traded, whereas most services are local population orientated. On the other hand the convergence in services most likely reflects the systemic shift towards a more orientated service economy and society.

5.1 The time-varying results

The time varying results are displayed in the lower panels of Figures 2 to 5. For agriculture (lower panel of Figure 2) there is evidence of a reduction in the degree of convergence as measured by the size of the largest cluster, falling from seven to five from the initial (1975-1993) to the final period (1981-1999). In addition, there is evidence of an increase in the existence of clustering at a smaller scale: the number of four region clusters increases from four in the sub-sample 1975-1993 to eight in the sub-sample 1981-1999. In the manufacturing sector (lower panel of Figure 3) there is a fall in the size of the largest cluster from eight to five regions and an increase in the number of middle-size clusters; in the market-service sector (lower panel Figure 4) there is no change in the size of the largest cluster. However, there is a reduction in the number of five region clusters from the initial (1975-1993) to the final period (1981-1999) and an increase in clustering at the smaller scale. In the non-market service sector (lower panel Figure 5) the degree of convergence is even more pronounced. The size of the largest cluster increases from five to seven from the initial (1975-1993) to the final period (1981-1999). Also there is a substantial increase in the extent of middle and small scale clusters, with four region clusters increasing from three to four and the three region clusters increasing from six to eight. So in this sector the regions not clustering in 1975-1993 are converging in 1981-1999.

Table 5 presents summary information in each sector. The distribution of clusters confirms the lower degree of convergence in the agricultural and manufacturing sector, and the higher degree of convergence of the service sector. The full-sample (1975-1999) results are displayed in the top panel, and indicate that the largest clusters are in the service sector (one seven region cluster in each sector), whereas in the agricultural and manufacturing sector there are mostly middle size (three and two) clusters. The time-varying results are displayed in the middle and lower panel of the table. Examining the final period (1981-1999) we observe that there are no clusters comprising more than five regions in both the agricultural and manufacturing sector, whereas in the market services we have one cluster with six regions, and in the non-market service sector there is one cluster comprising seven regions. Note that the non-market-service sector in the initial sample period (1975-1993) does not have any cluster comprising more than five regions.

A number of studies have detected a slowing down of overall regional convergence across the EU regions from the mid-1980s onwards. Our results suggest that this does not hold when sectoral disaggregations are examined; and that somewhat different processes are at work in manufacturing as compared to services. Although this issue obviously warrants further investigation, beyond casual observations as to the importance of spatial proximity and national (country) effects in influencing the convergence process, the clusters are difficult to interpret.

[11] Our focus is on labour productivity which does not allow for the identification of the contribution of technology and capital. As such a broader measure of multifactor productivity may lead to different results.

In exploring the reasons for this diverse convergence dynamics, and for the change in the cluster membership we construct testable hypotheses that examine the difference between observed cluster patterns as generated by our testing methodology, and hypothetical clusters generated by a number of specific socio-geographical and politico-institutional factors.

6 Comparing Cluster Outcomes with Hypothetical Cluster Patterns

In evaluating the cluster outcomes against one or more hypothetical cluster patterns, we consider the existence of one or more *orderings* based upon economic theory. In this respect we might instructively think of the clusters (and regions therein) as *data* generated by the outcome of a sequence of tests, with a given termination condition. It is therefore possible to confront these outcomes with hypothetical patterns formed on the basis of a number of hypotheses which encapsulate various 'models' of regional economic growth. The geographical, socio-demographic and political indicators used are listed in Table 2. All of them are central components in the new economic geography growth models since they justify the presence of increasing returns and comparative advantage at the sectoral and/or regional level (Fujita, Krugman, and Venables (1999) and Fujita and Thisse (2001)).

The first set of *geographical* factors explains convergence on the basis of country-membership, peripheral-core distribution of the economic activity, geographical location and the intensity of the transportation network. In their earlier work on regional convergence, Barro and Sala-i-Martin (1995) argued that regional convergence is more likely amongst regions within a given nation than it is between regions in different nations. Their argument is that institutional frameworks, regulatory systems, consumer tastes, and technologies are much more similar across regions within a given country than they are between different countries. This line of reasoning would lead us to hypothesise a significant country (national) effect on regional convergence clustering.

At the same time, recent work on the application of endogenous growth theory to regional development, suggests that growth effects arising from knowledge creation and spillovers, on the one hand, and the accumulation of skilled human capital, on the other, tend to exhibit spatial concentration. Strong spatial proximity effects are held to operate, implying a significant degree of spatial autocorrelation effects in the geographical pattern of growth performance. In other words, we should expect convergence clusters to comprise sets of neighbouring or spatially proximate regions. Another important factor for the location of activity is the intensity of transportation network. Since the production in our four sectors differs in the intensity of transportation costs and in their relative distance from the final markets, then regions with a better transport infrastructure might be expected to attract sectors which produce transport intensive commodities. This approach is developed in a trade theory framework in Limao and Venables (2001). On a larger geographical scale, it is often argued that the regional patterns of growth and development in the EU are characterised by a strong and persistent core-periphery structure, in which a core of leading growth regions encompassing the South East region of the UK, parts of the Netherlands, the Paris region, the Brussels region, Southern Germany, and northern Italy, is contrasted with a periphery of slower growing regions. The implication is that regional convergence dynamics should reflect the core-periphery dichotomy.

The second set of *socio-demographic* factors explains convergence on the basis of population growth and agglomeration effects. Along these lines Martin and Ottaviano (2001) show that growth and geographical agglomeration are self-reinforcing processes. In fact, agglomeration increases with growth since it is always more convenient to locate the activity where final market is bigger or the production of knowledge is higher. At the same time growth increases with agglomeration since agglomeration reduces the cost of innovating in the regions where the economic activity concentrates.

The third set of *political* factors explains regional convergence on the basis of political intervention at the EU level. In fact, the role of the European Monetary Union is to "encourage and guide structural adjustment which would help poorer regions to catch up with the wealthier ones" (Delors (1989), p. 22; see also Martin (2001) and Martin (2003)). The instruments used include the European Development Fund, the European Social Fund and the European Agricultural Guidance and Guarantee Fund (Martin and Tyler (2000)). This allows us to test the validity of what Boldrin and Canova (2001) label as the 'weak-divergence' theory. According to this notion of weak divergence, to escape from the poverty trap more depressed regions should be provided with externality-inducing factors. We test the correlation between the observed clusters and an artificially constructed matrix based on whether each region has been recipient of one of the structural funds. Table 3 summarises the implications of each of these indicators for the regional clusters.

7 The Univariate and Multivariate Cluster Correlations

To test for these different hypotheses we calculate cluster correlations across a number of artificially constructed cluster patterns. We write this correlation as

$$\zeta^h = \left(\frac{\sum_{j=1}^{n} \sum_{i \neq 1}^{n} m_{ij}^h \times \widehat{m}_{ij}}{\left(\sum_{j=1}^{n} \sum_{i \neq 1}^{n} m_{ij}^h \right)^{1/2} \left(\sum_{j=1}^{n} \sum_{i \neq 1}^{n} \widehat{m}_{ij} \right)^{1/2}} \right)^{1/2} \tag{5}$$

where $\mathbf{M}^h = \{m_{ij}^h\}$ denotes the artificially constructed cluster matrix based on hypothesis h, and $\widehat{\mathbf{M}} = \{\widehat{m}_{ij}\}$ is the matrix of observed clusters. Elements $\widehat{m}_{ij}$ and m_{ij}^h equal 1 if regions i and j belong to the same cluster.

To evaluate the hypotheses presented above we use data on a set of quasi-fixed effects described in Tables 2 and 3. Given that we would expect each effect to explain a relatively small fraction of the cluster outcome, we also test whether the degree of convergence and the cluster composition in the four sectors is affected by the *joint* interaction of some of the geographical, socio-demographic and policy indicators. To test for these joint effects we construct the multivariate cluster correlation index, ζ_m, which we write as:

$$\zeta_m^h = \left(\frac{\sum_{j=1}^{n} \sum_{i \neq 1}^{n} m_{ij}^{h_1} \times m_{ij}^{h_2} \times m_{ij}^{h_3}}{\left(\sum_{j=1}^{n} \sum_{i \neq 1}^{n} m_{ij}^{h_1} \right)^{1/2} \left(\sum_{j=1}^{n} \sum_{i \neq 1}^{n} m_{ij}^{h_2} \right)^{1/2} \left(\sum_{j=1}^{n} \sum_{i \neq 1}^{n} m_{ij}^{h_3} \right)^{1/2}} \right)^{1/2} \tag{6}$$

where h is a multivariate hypothesis which, in this example, is based on a combination of three univariate hypotheses h_1, h_2, h_3. The correlation coefficient ζ_m^h gives the correlation between our observed outcome and the multivariate hypothesis h.

7.1 The Tests Outcome

Tables 6 and 7 report the results of the univariate and multivariate cluster correlation analysis described in the previous section. In order to be able to discriminate among the different outcomes we test for the significance in the difference of the correlation coefficients between the two sub-periods 1975-1993 and 1981-1999 using the following statistic:

$$z = \frac{Z_1 - Z_2 - \mu_{Z_1-Z_2}}{\sigma_{Z_1-Z_2}} \qquad \text{where} \qquad Z_i = \frac{1}{2}\ln\frac{1+\zeta_i}{1-\zeta_i}\ (i=1,2) \tag{7}$$

where the subscripts 1 and 2 refer, respectively, to the sub-periods 1975-1993 and 1981-1995.[12] The values in bold of Tables 6 and 7 indicate that the correlation coefficients are significant.

The results in Table 6 indicate that, as we would predict, the importance of the geographical factors in determining the cluster outcomes is falling over time. In the case of agriculture, the correlation between the observed outcome and the hypothesised cluster pattern generated by country-membership falls from 0.36 to 0.29; manufacturing from 0.34 to 0.28; and non-market services from 0.43 to 0.35. The correlation between the observed outcome and the hypothesised cluster pattern generated by a peripheral-core distribution of regions in manufacturing falls from 0.38 to 0.27. In addition the correlation with the hypothesised cluster pattern generated by the geographical location falls from 0.35 to 0.28 in manufacturing, and from 0.39 to 0.36 in the non-market services. Also the role played by the intensity of the transportation network is decreasing across the two sub-samples both in agriculture (from 0.29 to 0.22) and in the non-market services (from 0.33 to 0.29). In the agricultural sector we find that in the second sub-period only 34% of the cluster outcome can be explained by the intensity of land used (percentage of land used by total area). This is in contrast with the prediction that regions with a similar level of agricultural intensification should also display a similarity in the productivity levels and in the productive knowledge used (see Durlauf and Johnson (1992)).

Finally, utilising information on the policy objective status assigned by the EU to each NUTS1 region, we are also interested in assessing whether the provision of the EU Structural and Cohesion Funds have played any role in reducing divergence between richer and poorer regions[13] across the sample period considered (recognising of course that the strength of those funds has varied over time). The correlation results indicate that, in the agricultural sector, the importance of the political factor in determining the cluster outcome is falling over time and in the final sub-period only 35% of the regions which are classified under the same objective of the EU structural fund intervention are clustering together.

Overall our results indicate that external economies derived from the geographical factors listed in Tables 2 and 3 and which constitute an important source of increasing returns and comparative advantage, are progressively becoming less important in explaining the cluster outcomes, particularly in the manufacturing sector where the geographical location and the peripheral-core distribution of the industrial activity is becoming less concentrated and more dispersed. This observation is also consistent with the notion that increasing integration appears to lower the costs of trade between regions, thereby encouraging the dispersion of economic activity (see Krugman and Venables (1995) and Baldwin and Forslid (1999)).

[12] z is normally distributed with mean $\mu_{Z_1-Z_2}$ and variance $\sigma_{Z_1-Z_2} = \sqrt{\sigma^2_{Z_1} + \sigma^2_{Z_2}} = \sqrt{\frac{2}{n(n-1)-3}}$

[13] To date, as Boldrin and Canova (2001) emphasise, the main criticism is that regional policies have served mostly "a redistributive purpose motivated by the nature of the political equilibria upon which the European Union is built" (p. 4).

Turning to socio-demographic factors, in agriculture population growth by area is less relevant in explaining the change over time of the cluster outcomes; its correlation with the observed cluster outcome falls from 0.37 to 0.27. The settlement structure is becoming less relevant both in agriculture, where the coefficient falls from 0.36 to 0.29, and in manufacturing, where the correlation falls from 0.31 to 0.26. This result indicates that the two sectors tend to be less concentrated in areas where the settlement structure is more agglomerated, hence becoming progressively less local-demand induced.

We now analyse the multivariate correlation coefficients between the observed cluster outcomes and the hypothesised cluster patterns generated by a composite indicator based on combining a number of the quasi-fixed factors. In this case we examine whether these fixed effects have jointly combined to reduce or increase divergencies. In the case of agriculture we examine the extent to which country-membership (C), location (L) and the agricultural intensification (AG) indices can explain, jointly, the observed set of cluster outcomes. This multivariate correlation index allows us to establish whether regions with similar intensity of land utilisation, climate conditions and institutional settings are aligned in terms of productivity. For manufacturing, we examine whether regional cluster outcomes are informed by the canonical 'new economic geography view' of joint interaction among country-membership, location and the periphery-core distribution (PC) of industrial activity. For market and non-market services we examine the extent to which the joint interaction of socio-demographic and geographical factors (country, location and the indicator of intensity of population growth (P)) play a role in explaining the cluster outcomes as a possible source of agglomeration externalities. Finally, for all the four sectors we analyse the joint interaction between country-membership, location and the political-intervention at the EU level (EU).

Table 7 shows that for market-services the correlation between the observed outcome and the composite indicator based on country membership, location and the provision of EU Structural Funding increases from 0.24 to 0.27 from the initial to the final period, whereas for non-market services the correlation between the same composite indicator and the observed cluster outcomes decreases from 0.33 to 0.24. In the market-service sector (which includes retail, distribution and banking) the high correlation between the observed clusters and the EU indicator may be explained by considering that in recent years the sector has been targeted by the allocation of Cohesion Funds[14], thereby implying that the productivity levels of the regions where the service sector is predominant have gradually converged.

A potentially noteworthy result is obtained in the manufacturing sector. The multivariate correlation results confirm that the concentration of production is becoming less localised and more fragmented. The joint effect of country location and peripheral-core patterns falls substantially from 0.29 to 0.16, suggesting that the provision of poorer areas with a number of fundamental endowments to promote structural adjustment and development, has progressively eroded the traditional notion of geographical periphery. For the non-market service sector, the correlation of the observed cluster outcome with the hypothesised joint indicator comprising country-membership, location and population, increases from 0.23 to 0.29. Since this sector comprises education and health, it is clear that country-membership and location, combined with a specific population structure, explains most of the change of the relative convergence cluster outcomes.

[14] Particularly to Greece, Portugal, Ireland and Spain.

8 Conclusions

There is a wide debate whether the rejection or non rejection of stationarity is informative about the process of convergence. As St. Aubyn (1999) notes, if countries with different initial points are converging, which is in line with a significant negative coefficient for the initial value in cross-section regressions, then pair-wise stationarity tests may well reject the null. A similar outcome could also be reached if the economies are *truly* diverging, such that a test of stationarity does not provide us with a unique answer. Secondly, it is also true that there is as much interest in *how* these patterns evolve over time and space, as in the general question of whether or not convergence has taken place. We have addressed this issue by examining the pattern of European convergence using pairwise stationarity tests on regional contrasts over the full sample period and across sub-samples, using a time-varying framework to capture the full set of converging and diverging patterns among the regional economies over time.

Unlike previous studies on the identification of convergence clusters we have utilised a methodology which places no constraints upon both the number and composition of clusters. In addition, and in contrast to Canova (1999), we do not impose any ordering on the data. The resulting cluster outcomes have been interpreted by comparisons with hypothesised cluster patters informed by economic theory. Using a set of geographical, socio-demographic and politico-institutional indicators available at the NUTS1 regional level, we have calculated the correlation between the hypothetical cluster patterns implied by each of these quasi-fixed effects and the observed outcomes. The results are quite different across the four sectors considered. For agriculture, the convergence clubs are affected by geographical factors. This may depend either on similar starting conditions, which define local convergence clubs (Durlauf and Johnson (1992)), or on similar climate which facilitates technological spillovers across regions (Wichmann (1996)). For manufacturing, the observed outcomes show, in general, a minor degree of convergence; in fact, in this sector the production of different goods implies that technology cannot diffuse easily; this produces patterns of regional specialisation across Europe which, however, do appear to weaken over time (Bernard and Jones (1996)). In the service sector, the observed clusters display more convergence, since typically this is a knowledge and information-based sector (the weightless economy) that transcends physical distance and geographical proximity (Quah (1996)). When comparing the observed outcomes with the different hypothesised pattern informed by economic theory, we find that productivity is more concentrated in the areas where the population growth is higher. Linked to this findings there is probably an 'agglomeration' effect which produces certain types of location externalities.

Finally, we have tested whether the provision of EU Structural and Cohesion Funds have had any role in affecting the composition of the convergence clubs. For agriculture, the correlation of the EU funding indicator in the two sub-samples remains stable, indicating that regions under the same objective had a similar cluster pattern across the two sub-periods. The only sector where the correlation between the observed outcomes and the EU funding indicator increases, is the market-service sector which represents, in the EU perspective one of the key externalities-inducing sectors to promote regional development.

A. Robustness Results

The cluster algorithm described in Section 4 requires the choice of a critical p-value ($p_{\min}$), and a bandwidth parameter, l. As noted, as we reduce $p_{\min}$ the less likely is the rejection of the null hypothesis of convergence. The choice of the bandwidths as demonstrated by Hobijn and Franses (2000), and Hobijn, Franses, and Ooms (1998), turns out to be critical with small samples since it affects the size of the test. We examined the robustness of the cluster algorithm with respect to various choices of l to assess both the degree of convergence (the sensitivity of the number of clusters with respect to the bandwidth); and the composition of the convergence clubs. To analyse the sensitivity of cluster composition we follow Hobijn and Franses (2000) and use a cluster correlation index which measures the degree of overlap between the two outcomes. To this end we construct a matrix $\mathbf{M}^* = \{m_{i,j}\}, i, j = 1, ..., R$, where m_{ij} is 1 if regions i and j belong to the same cluster and zero otherwise. Let $\mathbf{M}^a$ denote a particular value of $\mathbf{M}$ generated by a sequence of pairwise tests of stationarity with bandwidth parameter $l = a$; $\mathbf{M}^b$ is similarly defined. In varying the bandwidth parameter, $m_{ij}^a \times m_{ij}^{b\neq a}$ equals 1 if countries i and j are in the same convergence club for values of the bandwidth parameters a and b. The statistic $\zeta \in (0,1)$, given below, represents the correlation parameter between the two outcomes.

$$\zeta = \left(\frac{\sum_{j=1}^{n} \sum_{i\neq 1}^{n} m_{ij}^a \times m_{ij}^{b\neq a}}{\left(\sum_{j=1}^{n} \sum_{i\neq 1}^{n} m_{ij}^a\right)^{1/2} \left(\sum_{j=1}^{n} \sum_{i\neq 1}^{n} m_{ij}^{b\neq a}\right)^{1/2}} \right)^{1/2} \qquad \text{for} \qquad a, b = 1, ..., 6 \qquad \text{(A.1)}$$

Table A.1 reports the sensitivity results for all the sectors in the time-varying cluster algorithm considering a bandwidth parameter value ranging from 1 to 6. These outcomes have been generated by allowing for an eighteen-year rolling window spanning the years 1975-1999. In order to get information on the change in composition of the convergence clubs over our sample period, we consider the initial (1975-1993) and final window (1981-1999). This proves to be particularly useful in terms of policy analysis since we can assess whether the cluster size and its composition has changed over time. The first window captures the years following the creation of the European Regional Development Fund set up in 1975. The second window gives a picture of the fall-out of the intervention in 1986 aimed at reforming both the European Regional Development Fund, the European Social Fund and the European Agricultural Guidance and Guarantee Fund. We note that an application of the Hobijn and Franses (2000) procedure in small sample must be conducted with care, since the size of the test depends significantly on the choice of l. The first row in each matrix for the two sub-samples reports the number of clusters in the two sub-periods considered. For agriculture and non market-services there are more convergence clubs, hence less convergence, in the sub-sample (1981-1999). The reverse is true for manufacturing and market-services where there are fewer convergence clubs, and hence more convergence in the final sub-period. A bandwidth $l = 2$ is confirmed to be robust with respect to various choices of l. These results are in line with those reported by Hobijn and Franses (2000) for the Penn World Table convergence clusters, described in Summers and Heston (1991) where the cluster correlation between the various outcomes never exceeds 0.66. Hence, we apply a bandwidth $l = 2$ and we set $p_{\min}$ equal to 0.01.

Table A.1: Sensitivity Results: Relative Convergence

Agriculture

			1975-1993			
clubs	19	21	21	22	25	27
$l =$	1	2	3	4	5	6
1	–	0.7010	0.6671	0.5261	0.5685	0.6035
2	–	–	0.5718	0.5843	0.5400	0.5747
3	–	–	–	0.5290	0.5167	0.4740
4	–	–	–	–	0.6267	0.4395
5	–	–	–	–	–	0.5982
			1981-1999			
clubs	24	24	24	25	27	28
$l =$	1	2	3	4	5	6
1	–	0.6177	0.5218	0.6218	0.5866	0.5788
2	–	–	0.6496	0.5876	0.5487	0.5531
3	–	–	–	0.6325	0.5022	0.5490
4	–	–	–	–	0.5754	0.6305
5	–	–	–	–	–	0.5195

Manufacturing

			1975-1993			
clubs	23	22	26	26	29	27
$l =$	1	2	3	4	5	6
1	–	0.3850	0.5240	0.3762	0.3741	0.4359
2	–	–	0.5014	0.4671	0.4741	0.5173
3	–	–	–	0.3893	0.5626	0.3539
4	–	–	–	–	0.6934	0.4920
5	–	–	–	–	–	0.5491
			1981-1999			
clubs	25	23	25	26	26	32
$l =$	1	2	3	4	5	6
1	–	0.4990	0.5363	0.4350	0.4520	0.4033
2	–	–	0.6152	0.5388	0.5139	0.4237
3	–	–	–	0.5982	0.4965	0.3993
4	–	–	–	–	0.5644	0.4741
5	–	–	–	–	–	0.4759

Market Services

			1975-1993			
clubs	21	22	22	25	27	28
$l =$	1	2	3	4	5	6
1	–	0.6394	0.6135	0.3732	0.4277	0.4671
2	–	–	0.6049	0.4123	0.5691	0.5035
3	–	–	–	0.4182	0.5000	0.5108
4	–	–	–	–	0.6044	0.3862
5	–	–	–	–	–	0.5702
			1981-1999			
clubs	24	24	25	27	28	30
$l =$	1	2	3	4	5	6
1	–	0.7469	0.6621	0.5822	0.5020	0.5839
2	–	–	0.7275	0.6475	0.5097	0.5509
3	–	–	–	0.6126	0.5034	0.5449
4	–	–	–	–	0.5390	0.5976
5	–	–	–	–	–	0.7438

Non-Market Services

			1975-1993			
clubs	23	24	25	27	27	29
$l =$	1	2	3	4	5	6
1	–	0.5838	0.4611	0.5435	0.4527	0.5404
2	–	–	0.6855	0.6672	0.6019	0.5403
3	–	–	–	0.5758	0.5945	0.5860
4	–	–	–	–	0.7311	0.5925
5	–	–	–	–	–	0.5825
			1981-1999			
clubs	22	22	22	24	26	26
$l =$	1	2	3	4	5	6
1	–	0.4205	0.4230	0.5413	0.4582	0.5400
2	–	–	0.4913	0.6276	0.4654	0.4789
3	–	–	–	0.4990	0.4541	0.4817
4	–	–	–	–	0.5634	0.4699
5	–	–	–	–	–	0.5345

References

Abramovitz, M. (1986): "Catching Up, Forging Ahead, and Falling Behind," *Journal of Economic History*, 46:2, 385–406.

Baldwin, R. E., and R. Forslid (1999): "The Core-Periphery Model and Endogenous Growth: Stabilizing and Destabilizing Integration," *Economica*, 67, 307–24.

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

Barro, R. J., and X. Sala-i-Martin (1995): "Technological Diffusion, Convergence and Growth," NBER working paper 5151.

Bernard, A. B., and S. N. Durlauf (1995): "Convergence in International Output," *Journal of Applied Econometrics*, 10, 97–108.

Bernard, A. B., and C. I. Jones (1996): "Comparing Apples to Oranges: Productivity Convergence and Measurement Across Industries and Countries," *The American Economic Review*, 86(5), 1216–1238.

Bertola, G. (1993): "Models of Economic Integration and Localized Growth," in *Adjustment and Growth in the European Monetary Union*, ed. by F. Torres, and F. Giavazzi, pp. 156–179, Cambridge. Cambridge University Press.

Boldrin, M., and F. Canova (2001): "Inequality and Convergence: Reconsidering European Regional Policies," *Economic Policy: A European Forum*, 0(32), 205–45.

Boltho, A., and G. Holtham (1992): "The Assessment: New Approaches to Economic Growth," *Oxford Review of Economic Policy*, 8(4), 1–14.

Canova, F. (1999): "Testing for Heterogeneity in the Cross Sectional Dimension of a Panel: A Predictive Density Approach," Working Paper, Universitat Pompeu Fabra, May.

Cheshire, P., and G. Carbonaro (1989): *Convergence-Divergenece in Regional Growth Rates: An Empty Black Box?* Unwin Hyman, London.

Delors, J. (1989): "Report on Economic and Monetary Union in the European Community," Discussion paper, Report to the European Council by a Committee chaired by J. Delors, Brussels, Commission of the European Community.

Durlauf, N., and P. Johnson (1992): "Local versus Global Convergence Across National Economies," Discussion Paper 3996, NBER.

Durlauf, S., and P. Johnson (1995): "Multiple Regimes and Cross-Country Growth Behavior," *Journal of Applied Econometrics*, 10, 365–84.

Fujita, M., P. Krugman, and A. J. Venables (1999): *The Spatial Economy: Cities, Regions and International Trade*. MIT Press, Cambridge, MA.

Fujita, M., and J. Thisse (2001): "Agglomeration and Market Interaction," .

Gregg, P., and J. Wadsworth (1999): *The State of Working Britain*. Manchester University Press.

Grossman, G. H., and E. Helpman (1991): "Trade, Knowledge Spillovers, and Growth," *European Economic Review*, 35(2-3), 517–26.

Hobijn, B., P. Franses, and M. Ooms (1998): "Generalisation of the KPSS-Test for Stationarity," Discussion Paper Econometric Institute Report 9802, Erasmus University Rotterdam.

Hobijn, B., and P. H. Franses (2000): "Asymptotically Perfect and Relative Convergence of Productivity," *Journal of Applied Econometrics*, 15, 59–81.

Jian, T., J. D. Sachs, and A. M. Warner (1996): "Trends in Regional Inequality in China," *National Bureau of Economic Research Working Paper 5412*, p. 24.

KRUGMAN, P. (1991): "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99(3), 483–499.

KRUGMAN, P., AND A. VENABLES (1995): "Globalisation and the Inequality of Regions," *Quarterly Journal of Economics*, 60, 857–80.

LIMAO, N., AND A. VENABLES (2001): "Infrastructure, Geographical Disadvantage, Transport Costs, and Trade," *World Bank Economic Review*, 15(3), 451–79.

LUCAS, R. (1988): "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22, 3–42.

MARTIN, P., AND G. I. P. OTTAVIANO (2001): "Growth and Agglomeration," *International Economic Review*, 42(4), 947–68.

MARTIN, R. (2001): "EMU versus the Regions: Regional Convergence and Divergence in Euroland," *Journal of Economic Geography*, (1), 51–80.

——— (2003): "EMU and Enlargement: Twin Threats to European Regioanl Cohesion?," *in Arnull, A. and Wincott, D. (Eds) Accountability and Legitimacy in the European Union*, pp. 345–363.

MARTIN, R., AND P. SUNLEY (1998): "Slow Convergence? The New Endogeneous Growth Theory and Regional Development," *Economic Geography*, 74, 201–227.

MARTIN, R., AND P. TYLER (2000): "Regional Employment Evolutions in the European Union: A Preliminary Analysis," *Regional Studies*, 34(7), 601–16.

QUAH, D. (1995): "Ideas Determining Convergence Clubs," Working paper, Economics Department, London School of Economics (August).

——— (1996): "Aggregate and Regional Disaggregate Fluctuations," *Empirical Economics*.

ROMER, P. (1986): "Increasing Returns and Long-Run Growth," *Journal of political economy*, 94, 1002–1037.

SOLOW, R. M. (1956): "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, 70(1), 65–94.

ST. AUBYN, M. (1999): "Convergence Across Industrialised Countries (1890-1989): New Results Using Time Series Methods," *Empirical Economics*, 24(1), 23–44.

SWANN, T. W. (1956): "Economic Growth and Capital Accumulation," *Economic Record*, 32, 334–361.

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

WICHMANN, T. (1996): "Technology Adoption in Agriculture and Convergence Across Economies," .

Table 1: NUTS1 code

Code	Country		Code	Country	
AT	*Austria*		IE	*Ireland*	
AT1		Ostosterreich			
AT2		Sudosterreich	IT	*Italy*	
AT3		Westosterreich	IT1		Nord Ovest
BE	*Belgium*		IT2		Lombardia
BE1		Region Bruxelles-Capital-Brussels	IT3		Nord Est
		Hoofdstedelijke Gewest	IT4		Emilia-Romagna
BE2		Vlaams Gewest	IT5		Centro
BE3		Region Wallonne	IT6		Lazio
DE	*Germany*		IT7		Abruzzo-Molise
DE1		Baden-Wurttemberg	IT8		Campania
DE2		Bayern	IT9		Sud
DE3		Berlin	ITA		Sicilia
DE5		Bremen	ITB		Sardegna
DE6		Hamburg	LU	*Luxembourg*	
DE7		Hessen			
DE9		Niedersachsen	NL	*Netherlands*	
DEA		Nordrhein-Westfalen	NL1		Noord-Nederland
DEB		Rheinland-Pfalz	NL2		Oost-Nederland
DEC		Saarland	NL3		West-Nederland
DEG		Thuringen	NL4		Zuid-Nederland
DK	*Denmark*		PT	*Portugal*	
			PT1		Continente
ES	*Spain*		SE	*Sweden*	
ES3		Comunidad de Madrid			
ES4		Centro(E)	UK	*United Kingdom*	
ES5		Este	UKC		North East
ES6		Sur	UKD		North West
ES7		Canarias	UKE		Yorkshire and
F1	*Finland*				Humber
			UKF		East Midland
FR	*France*		UKG		West Midlands
FR1		Ile de France	UKH		East of England
FR2		Bassin-Parisien	UK1		London
FR3		Nord Pas de Calais	UKJ		South East
FR4		Est	UKK		South West
FR5		Ouest	UKL		Wales
FR6		Sud-Ouest	UKM		Scotland
FR7		Centre-Est			
FR8		Mediterranee			
GR	*Greece*				
GR1		Voreia Ellada			
GR2		Kentriki Ellada			
GR3		Attiki			
GR4		Nisia Aigaiou, Kriti			

Table 2: Data Source and Description

	Factors	Mechanism	Year	Coverage	Measurement level	Source	Comments
Geographical	Country	Institutional Setting			nominal		Manual classification
	Peripheral Core	Accessibility		EU15	nominal	German Federal Office for Building and Planning (BBR)	1 Peripheral core, 2 Central and central metropolitan regions, 3 Tourist regions, 4 Brussels and Bremen, 5 German New Lander, 6 Central and Eastern UK, 7 Nordic countries, 8 Peripheral Southern Europe, 9 Mediterranean plus Ireland, 10 Northern Italy.
	Geographic location of region	Contiguity and Institutional Similarity		EU15	nominal	BBR	Manual classification according to geographic location: 1 The North, 2 Atlantic, 3 Mediterranean, 4 Eastern EU-border, 5 Centre.
	Dissection: Length of transportation	Accessibility	1996(94/9)	EU15	ordinal	University of Trier	The classes are: 1 Very low, 2 Low, 3 medium, 4 High, 5 Very High
	Agricultural intensification	Specialisation	1989-96	EU15	ordinal	Greek and Dutch NFPs	Composite indicator of percent growth of agricultural accounts, percent of agricultural holdings > 50 and percent of land use by total area (see report of the working group). The classes are: 1 High pressure, 2 Important pressure, 3 Eventual presence of pressure, 4 Neutral pressure, 5 Negative pressure.
Socio-demographic	Population growth by total area	Agglomeration	1991-95	EU15	ordinal		No data for Madeira, Acores, Canarias, Ceuta y Melilla; the classes are 1 very low, 2 low, 3 medium, 4 high, 5 very high.
	Settlement structure	Agglomeration		EU15	nominal	BBR	I. Agglomerated regions with a centre > 300,000 and a population density >(I.1) or < (I.2) 300 inhabitants/km^2; II. Urbanised regions with a centre > 150,000 inhabitants with a population density > (II.1) or < (II.2) 150 inhabitants/km^2. III Rural regions with a population density < 100 inhabitants/km^2 and a centre > (III.1) or < (III.2) 125,000 inhabitants.
Political	Type of EU Structural and Cohesion Funds	Externalities-Inducing Policies		EU15	ordinal	Eurostat	The classes are 0 = No special status, 1 = Objective 1 status only, 2 = Objective 2 status only, 5 = Objective 5b status only, 6 = Objective 6 status only, 7 = Partially Objective 5b, 8 = Partially Objective 2, 9 = Partially Objective 2 and 5b, 10 = Partially Objective 2, 5b and 6, 11 = Partially Objective 1 and 5b, 12 = Partially Objective 1 and 2, 13 = Partially Objective 1, 2 and 5.

Table 3: The Quasi-Fixed Factors

	Factors	Description
Geographical		
	Country membership	Regions cluster solely on the basis of their nation-state membership. The associated mechanisms include a shared institutional framework, the same set of political institutions and a well defined geographic boundary.
	Periphery-core	Regions are classified according to their relative distance with respect to a core of European regions.
	Geographic location	Regional clusters are determined by a broader geographical classification of regions: Northern European, Atlantic, Mediterranean, Central or Eastern European. Here, it is assumed that contiguity and institutional similarity may affect regional convergence
	Transportation network by total area	Regions are classified according to the intensity of transportation network.
	Agricultural intensification	Regions are classified according to a composite indicator of percent growth of agricultural accounts, percent of agricultural holdings greater than 50% and percentage of land use by total area.
Socio-demographic		
	Population growth by area	Regions are classified based upon the average of population growth between 1991 and 1995. Changes in population growth and population density capture the role of urban agglomeration in shaping real GVA per capita convergence.
	Settlement structure	Regions are classified according to the number of inhabitants and population density. This may reflect, for example, different levels of urbanisation and agglomeration dynamics.
Political		
	EU Structural Funds objectives	Regions are classified according to the different EU Structural Funds objectives. The EU Cohesion and Structural Fund objectives are:
		Objective 1. To promote the development and structural adjustment of underdeveloped regions
		Objective 2. To redevelop regions or areas within regions (local labour markets or urban communities) which are seriously affected by industrial decline
		Objective 3. To combat long term unemployment, to provide career prospects for young people (aged under 35) and to reintegrate persons at risk of being excluded from the labour market.
		Objective 4. To facilitate the adaption of workers to industrial change and developments in the production system.
		Objective 5a. To speed up the adaption of production, processing and marketing structures in agriculture and forestry and to help modernise and restructure the fisheries and aquaculture sector
		Objective 5b. To promote the development of rural areas
		Objective 6. To promote the development of northern regions in the new member states in Scandinavia (since 1995 Finland and Sweden)

Table 4: Asymptotic Relative Convergence: NUTS0

Agriculture	Manufacturing	Market Services	Non-Market Services
1. FR IE LU BE	1. LU AT UK NO	1. ES IE PT NO	1. DK DE ES AT UK
2. IT PT UK	2. DK FR PT	2. GR SE BE	2. IT LU NL
3. DK NL	3. ES IE BE	3. DE UK	3. IE FI BE
4. GR SE	4. GR NL	4. IT AT	4. SE NO
5. FI NO	5. DE FI	5. FR FI	
Single Country Clusters			
AT ES DE	SE IT	DK LU NL	FR GR PT

Table 5: Cluster Summary Information

Cluster size	1	2	3	4	5	6	7	8	Total Clusters
				Number of Clusters					
				1975-1999					
Agriculture	1	9	7	4	1	1	0	0	23
Manufacturing	2	7	11	4	1	0	0	0	25
Market Service	6	11	6	1	2	0	1	0	27
Non-market Service	1	8	6	2	1	2	1	0	21
				1975-1993					
Agriculture	2	5	4	6	0	2	1	0	20
Manufacturing	1	7	6	4	2	0	0	1	21
Market Services	1	10	4	2	4	1	0	0	22
Non-market Services	2	10	6	3	3	0	0	0	24
				1981-1999					
Agriculture	2	8	8	5	1	0	0	0	24
Manufacturing	3	7	4	7	2	0	0	0	23
Market Services	1	14	5	3	1	1	0	0	25
Non-market Services	1	6	9	4	1	0	1	0	22

Table 6: Univariate Analysis

		Agriculture	Manufacturing	Market Services	Non-Market Services
Geographical					
	Country Membership				
	(1975-1993)	**0.363****	**0.346****	0.340	**0.431****
	(1981-1999)	**0.295****	**0.297****	0.316	**0.352****
	z	*(3.59)*	*(2.57)*	*(1.26)*	*(4.40)*
	Periphery-Core				
	(1975-1993)	**0.349****	**0.380****	0.308	0.354
	(1981-1999)	**0.382****	**0.272****	0.318	0.359
	z	*(-1.79)*	*(5.70)*	*(-0.52)*	*(-0.27)*
	Geographic Location				
	(1975-1993)	0.384	**0.351****	0.318	**0.394***
	(1981-1999)	0.392	**0.284****	0.321	**0.364***
	z	*(-0.44)*	*(3.51)*	*(-0.15)*	*(1.65)*
	Transportation Network				
	(1975-1993)	**0.296****	0.279	**0.292****	**0.338****
	(1981-1999)	**0.229****	0.266	**0.329****	**0.293****
	z	*(3.39)*	*(0.66)*	*(-1.93)*	*(2.35)*
	Agricultural Intensification†				
	(1975-1993)	**0.402****	−−	−−	−−
	(1981-1999)	**0.348****	−−	−−	−−
	z	*(2.96)*			
Socio-Demographic					
	Population Growth by Area				
	(1975-1993)	**0.371****	**0.312****	0.340	0.324
	(1981-1999)	**0.270****	**0.350****	0.316	0.321
	z	*(5.31)*	*(-2.01)*	*(1.26)*	*(0.15)*
	Settlement Structure				
	(1975-1993)	**0.363****	**0.317****	0.302	0.304
	(1981-1999)	**0.297****	**0.267****	0.321	0.321
	z	*(3.49)*	*(2.57)*	*(-0.99)*	*(-0.88)*
Political					
	EU Structural Fund Objectives				
	(1975-1993)	**0.393***	0.310	0.342	0.335
	(1981-1999)	**0.358***	0.320	0.334	0.312
	z	*(1.92)*	*(-0.52)*	*(0.42)*	*(1.21)*

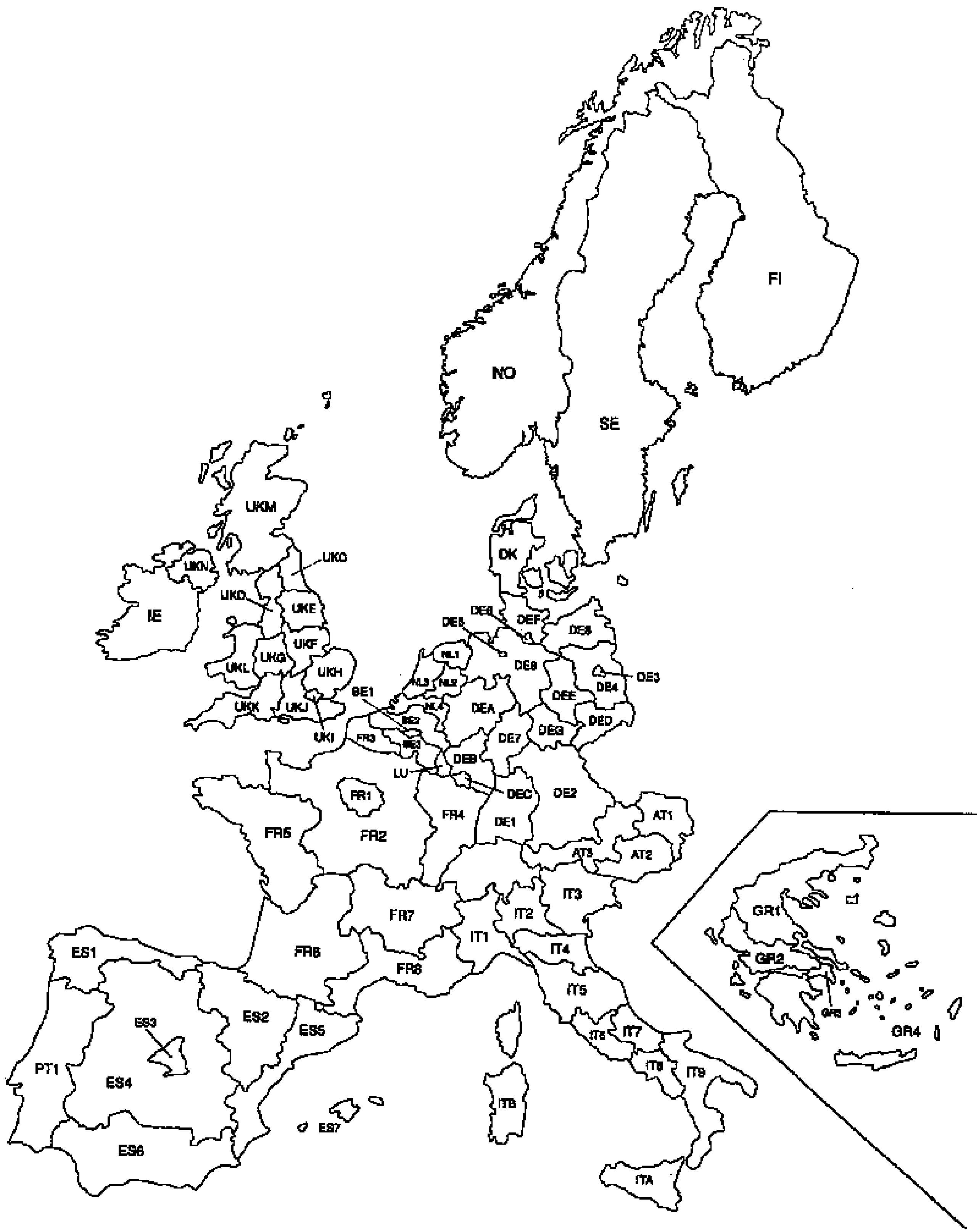
Significance Level ** = 5% and * = 10%. The values in bold are significant.
† Data are available only for the agricultural sector.

Table 7: Multivariate Analysis

	(1979-1993)	(1981-1999)	z	(1979-1993)	(1981-1999)	z
	$C \cap L \cap AG$			$C \cap L \cap EU$		
Agriculture	0.294	0.284	*(0.51)*	0.258	0.253	*(0.25)*
	$C \cap L \cap PC$			$C \cap L \cap EU$		
Manufacturing	**0.296****	**0.162****	*(6.68)*	0.275	0.244	*(1.56)*
	$C \cap L \cap P$			$C \cap L \cap EU$		
Market Services	0.264	0.236	*(1.40)*	**0.242****	**0.279****	*(-1.87)*
Non-Market Services	**0.237****	**0.296****	*(-2.99)*	**0.331****	**0.243****	*(4.52)*

C = Country Membership, PC = Periphery-Core, L = Geographical Location, AG = Agricultural Intensification, P = Population Growth by Area, EU = European Union Structural Funds Objectives. Significance Level ** = 5% and * = 10%. The values in bold are significant.

Figure 1

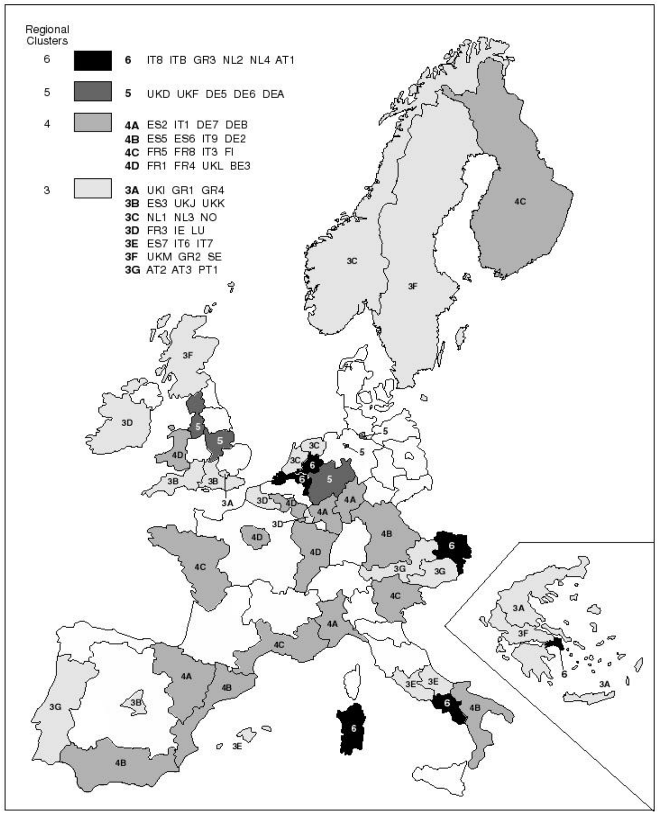


FI
NO
SE
UKM
UKC
UKN
IE
UKD
UKE
UKF
UKG
UKL
UKH
UKK
UKJ
UKI
DK
DEF
DE8
DE9
DE6
DE3
DE4
DEE
DED
DEG
DEA
DE7
DEB
DEC
DE2
DE1
NL1
NL2
NL3
NL4
BE1
BE2
BE3
FR3
LU
FR1
FR2
FR4
FR5
FR7
FR6
FR8
AT1
AT2
AT3
IT1
IT2
IT3
IT4
IT5
IT6
IT7
IT8
IT9
ITA
ITB
ES1
ES2
ES3
ES4
ES5
ES6
ES7
PT1
GR1
GR2
GR3
GR4

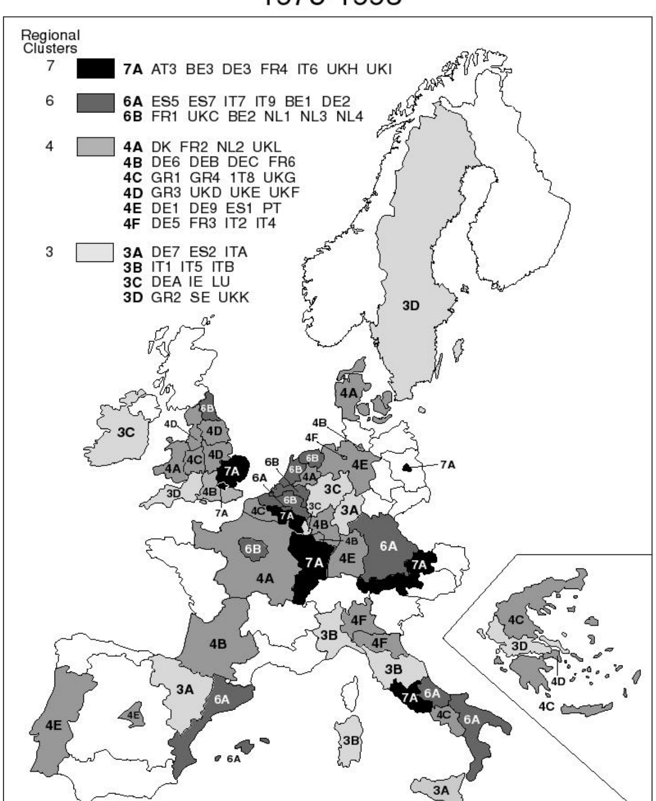
Figure 2

ASYMPTOTIC RELATIVE CONVERGENCE

AGRICULTURE 1975-1999



1975-1993



1981-1999

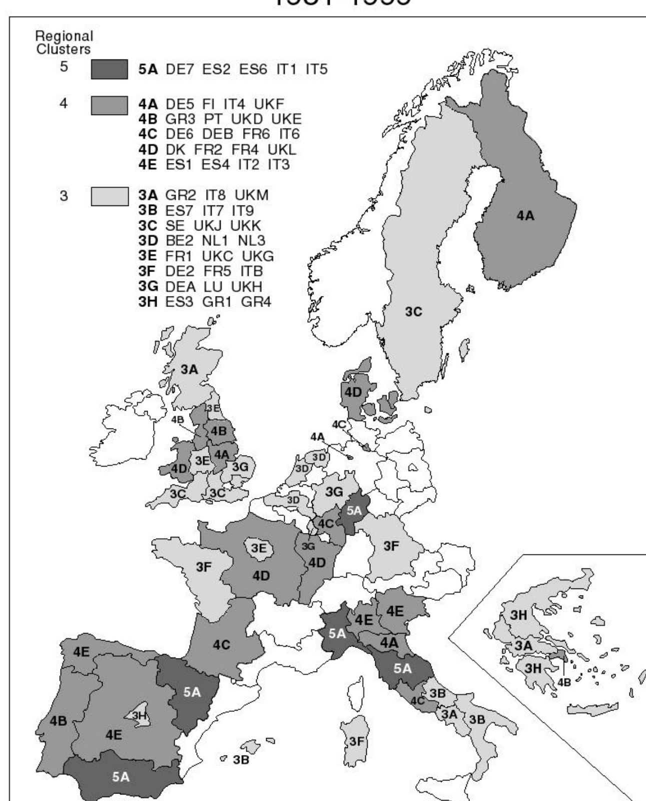
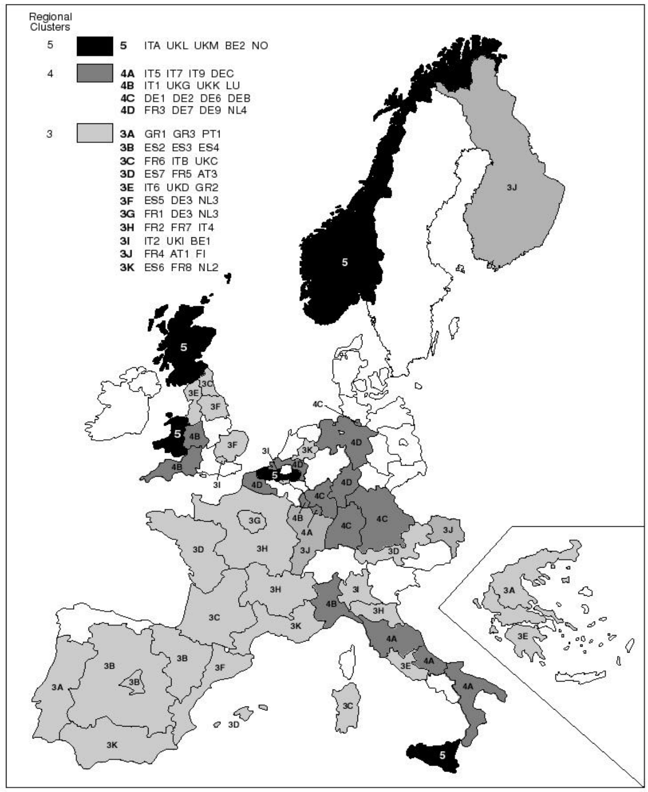


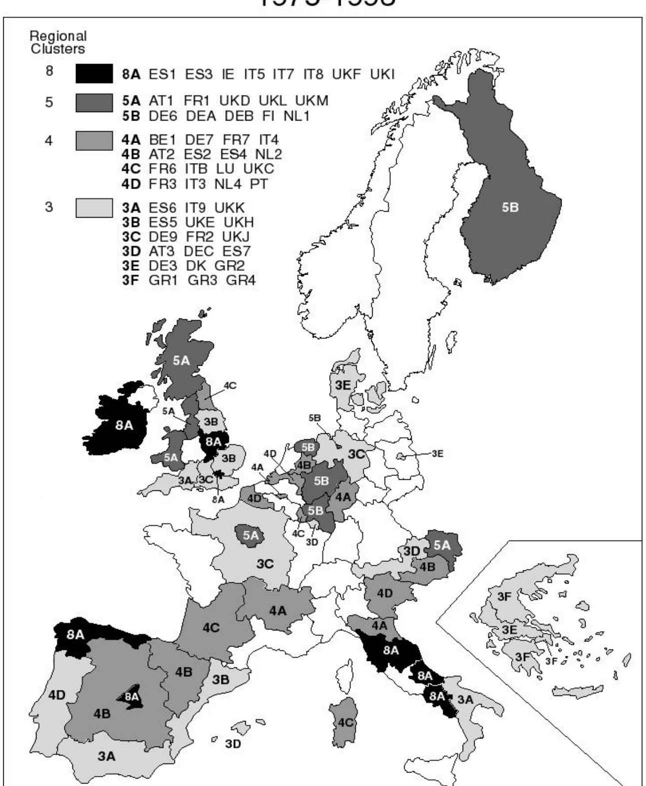
Figure 3

ASYMPTOTIC RELATIVE CONVERGENCE

MANUFACTURING 1975-1999



1975-1993



1981-1999

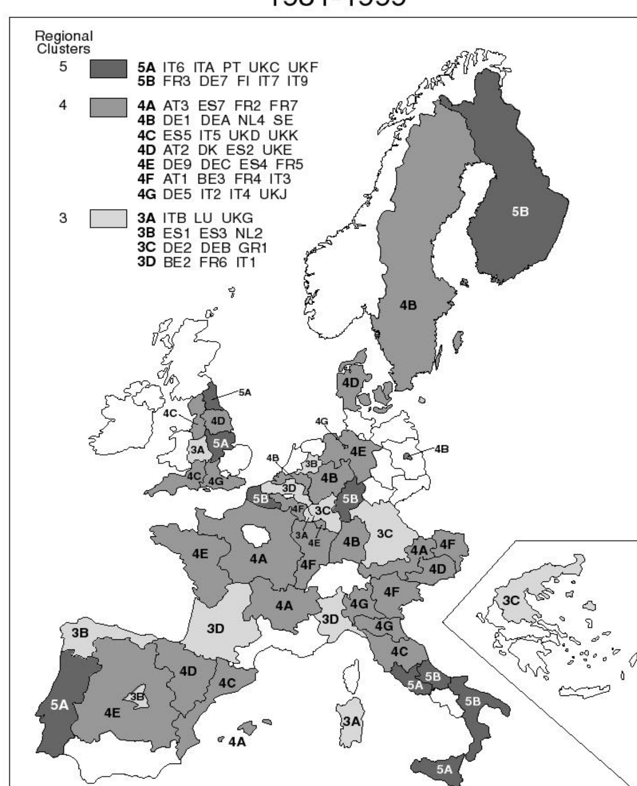
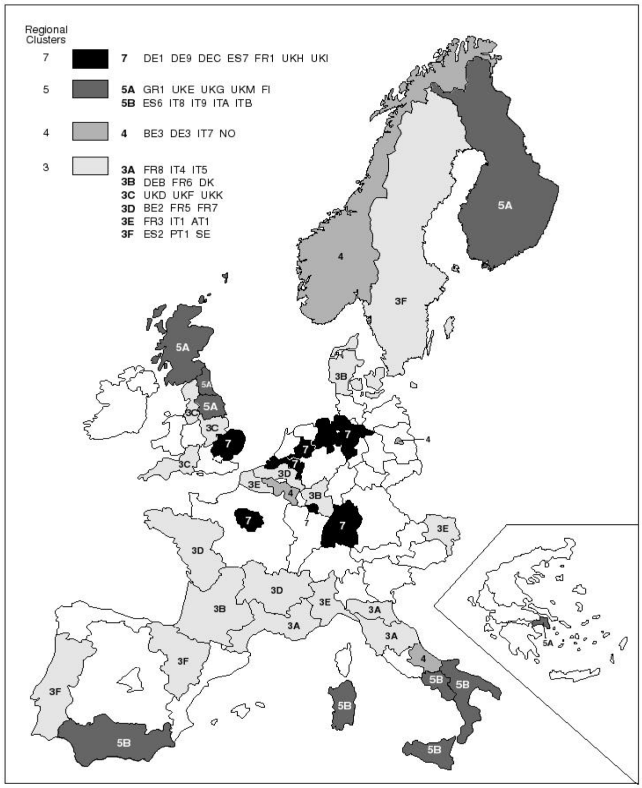


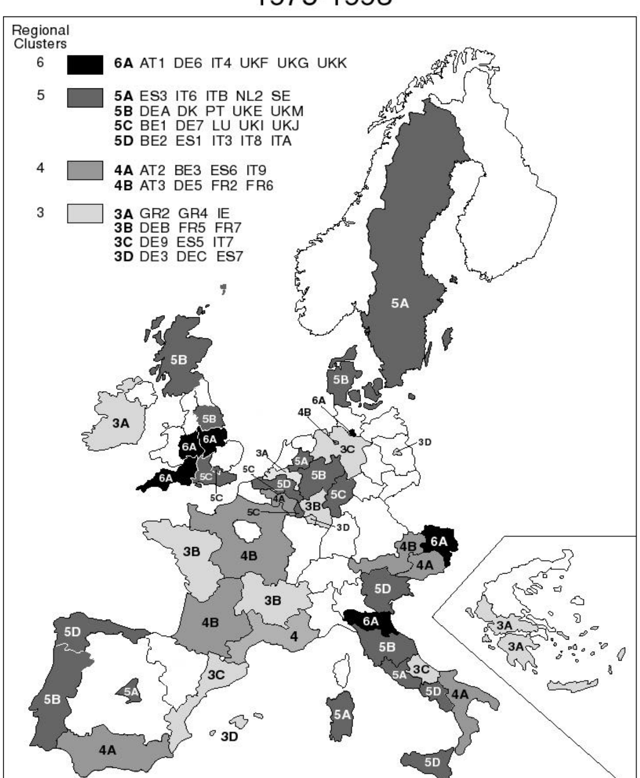
Figure 4

ASYMPTOTIC RELATIVE CONVERGENCE

MARKET SERVICES 1975-1999



1975-1993



1981-1999

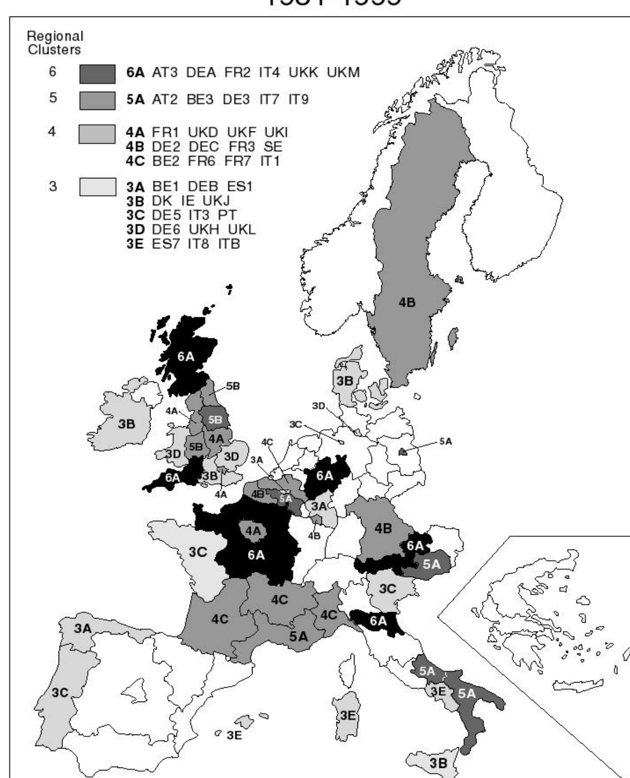
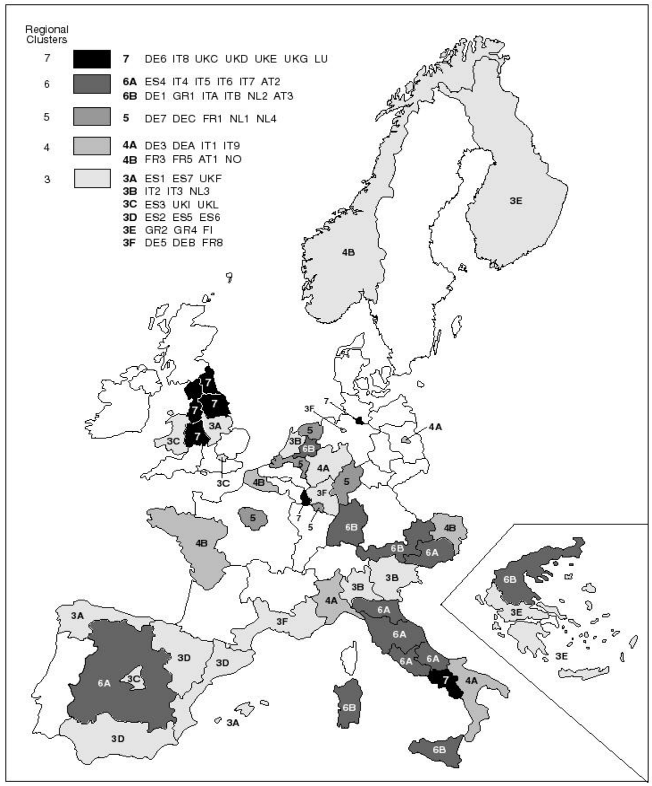


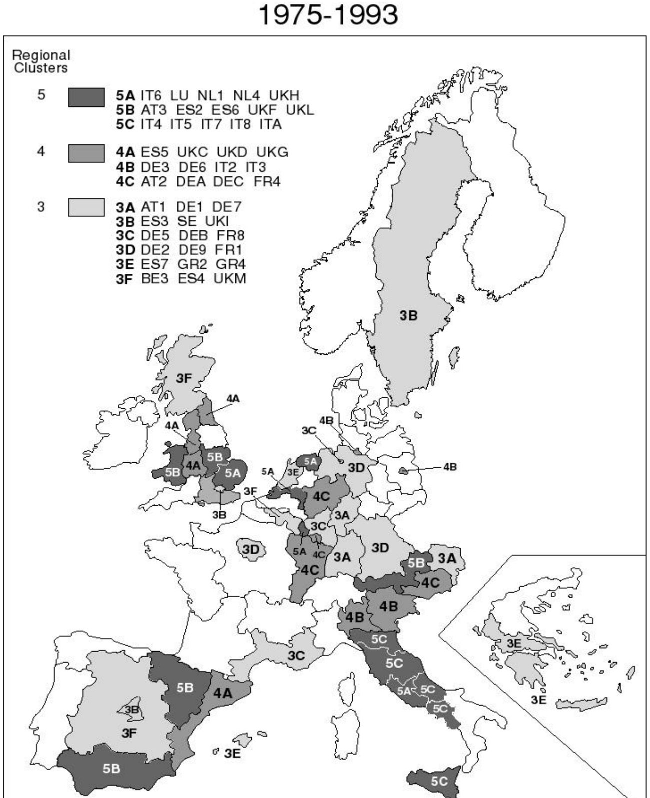
Figure 5

ASYMPTOTIC RELATIVE CONVERGENCE

NON-MARKET SERVICES 1975-1999



1975-1993



1981-1999