Regional growth in Western Europe: an empirical exploration of interactions with agriculture and agricultural policy

Roger S. Bivand and **Rolf J. Brunstad**[*]
Norwegian School of Economics and Business Administration

Discussion Paper 01/02

Abstract

Studies of convergence of regional growth in Western Europe have given varying results, depending on the addition of conditioning variables and estimation methods, as well as underlying models. This exploration adds agricultural variables to the basic convergence model, which suggests that initially poorer regions grow faster than initially more wealthy regions. Subsidising agriculture may be expected to impact growth in different ways than initial gross value added per capita. We find some support for our hypothesis that agricultural support has a negative impact on convergence. The coefficients have the right sign and are significant or nearly significant, but the results of both the basis model and our augmented model may be subject to other mis-specification problems, possibly including non-stationarity, even after including the spatially lagged dependent variable. This is examined using geographically weighted regression.

JEL Classification: R11, O40, C29

Keywords: Convergence, agriculture, spatial econometrics

[*]
Address for correspondence:
Department of Economics,
Norwegian School of Economics and Business Administration
Helleveien 30, N-5045 Bergen, Norway
e-mail: roger.bivand@nmhh.no
rolf.brunstad@nhh.no

This version January 24, 2002

Regional growth in Western Europe: an empirical exploration of interactions with agriculture and agricultural policy

Roger S. Bivand and Rolf J. Brunstad
Department of Economics,
Norwegian School of Economics and Business Administration,
Breiviksveien 40, N-5045 Bergen, Norway.

January 2002

1 Introduction

Much of the analysis of regional growth in Europe in recent years has concentrated on the concept of convergence, whether in the neo-classical model, or in alternative formulations. Empirical work has shifted from more confirmatory non-spatial estimation to the acknowledgement of the importance of spatial factors, including spillovers, and to exploratory spatial data analysis. This has occurred in conjunction with similar work in North America, and has led to a better understanding of the difficulties involved in relying simply on the convergence model without augmentations.

The work discussed in this chapter is an early attempt to throw light on apparent variability in regional convergence in relation to agriculture as a sector subject to powerful political measures. We would like to explore the possibility that some of the observed specification issues in current results are rooted in neglecting agricultural policy interventions, within the limitations imposed by data available at this stage. We would also like to use this as a case setting for evaluating the appropriateness of geographically weighted regression as a technique for assessing coefficient variability, over and above for instance country dummies, but possibly reflecting missing variables or other specification problems.

2 Convergence, agriculture and agricultural policy

Rather than review the convergence literature broadly, we prefer to focus on suggestions pointing up issues to be explored here. We will therefore be taking some positions as given, and will only mention them briefly for clarity. We will be concerned with β-convergence as represented in empirical studies in the following way:

$$\frac{1}{T}\log(\frac{y_{i,T}}{y_{i,0}}) = \alpha + \beta \log(y_{i,0}) + u_i,$$

where α and β are coefficients and u_i is a disturbance term (Paci, 1997, p. 617). Given an estimate of β, the speed of convergence may be represented as: $\theta = -\log(1+T\beta)/T$, with 100θ expressing this speed in percentage points (Baumont et al. 2001, p.8).

The underlying regularity in this representation is that the rate of growth $y_{i,T}/y_{i,0}$ of a regional economy i in the period up to T is related to its initial condition in period 0 for some measure $y_{i,0}$. The measure to be used here is gross value added per capita measured in EUR 1000 at constant 1990 prices. Figure 1 shows the regional distribution of this variable for 1989, the initial period to be used here except where stated. Details of the choice of period and regions will be given below in section 2.2. The contrast between lower values in most of the Iberian peninsula and southern Italy and the rest of the study area is familiar, as are the effects of regional definition artifacts[1] in Benelux and western Germany, with high urban values contrasting with apparently lower surrounding rural values despite commuting.

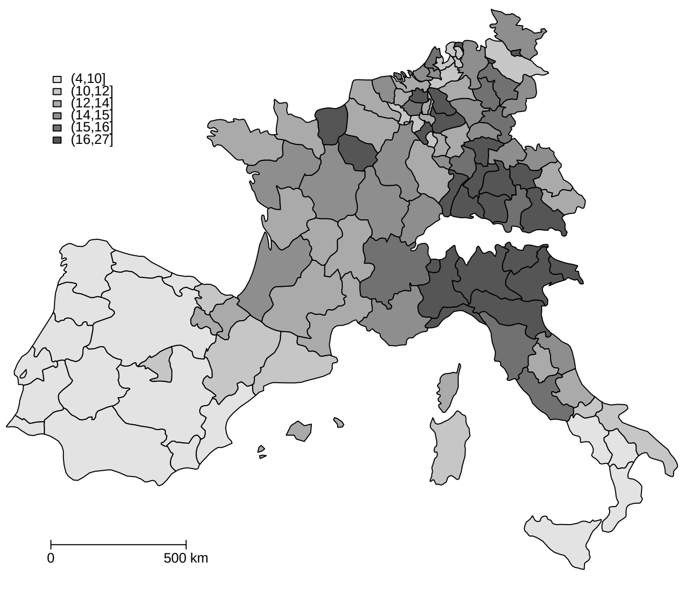

Figure 1: Gross value added per capita, EUR 1000, 1989 ($y_{i,0}$).

[1]For example the underbounding of Hamburg is discussed by Fingleton, where the functional region of the agglomeration is clearly larger than the administrative area (2001, p. 147).

Figure 2 appears to fit well to the initial conditions: regions such as those in Spain and Portugal with low initial condition values have high growth rates, while French regions have low growth rates and medium initial conditions. However, closer inspection suggests that the stories of particular regions, or clubs of regions, are more complex than our simple convergence model indicates. This is strengthened by an examination of Table 1, and by the insignificance of a χ^2 test (9.969, df = 9, p-value = 0.353) on the relationship.

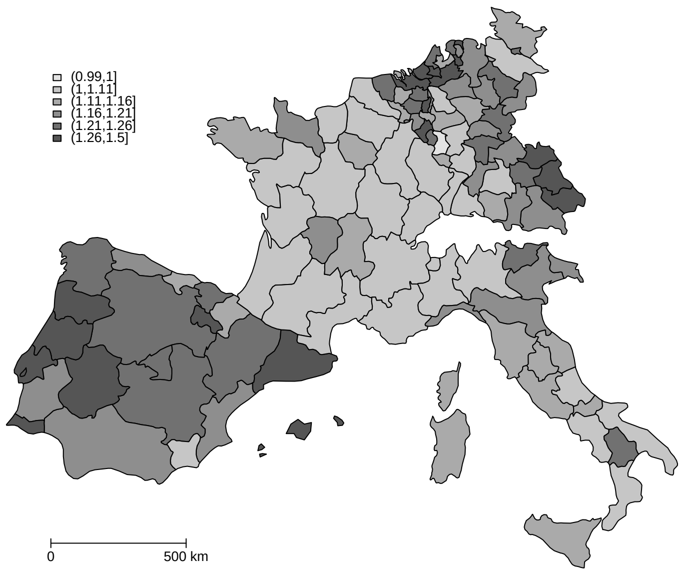

Figure 2: Regional economic change: GVA per capita 1999 as proportion of GVA per capita 1989 ($y_{i,1999}/y_{i,1989}$).

Table 1: Contingency table of initial conditions (1989) by growth rates (1989-1999), both variables cut at quartiles.

	$y_{i,1999}/y_{i,1989}$				
$y_{i,1989}$	0.999-1.11	1.11-1.17	1.17-1.23	1.23-1.45	sum
4.03-11.7	5	6	8	10	29
11.7-13.8	6	10	4	9	29
13.8-15.7	8	8	6	6	28
15.7-26.8	10	5	10	4	29
sum	29	29	28	29	115

To check whether this lack of a clear relationship is restricted to this choice of 1989 as base year, and the period to 1999 to measure growth, similar χ^2 tests for base years from 1989 to 1998, and growth rates from those base years to 1999, in total 55 tests, were conducted. They confirm that this value is not exceptional, and that it is difficult to sustain a general relationship between observed initial conditions and growth rates in European regions during this period. Only 14 of the 55 relationships pass as blunt a test as this. Returning to earlier empirical studies, it is clear that many have already pointed out not just the theoretical difficulties of the simple convergence model, but here more importantly the interesting structural features that it chooses to neglect.

Among others, Paci notes that "the observed process of aggregate convergence can hide important structural change phenomena" (1997, p. 617). In analysing sectoral labour productivity, Paci finds that without using a "Southern" dummy or national dummies, the convergence relationship for agriculture is not significant (1997, p. 627). Others, including Fagerberg et al. (1997), Pons-Novell and Viladecans-Marsal (1999), Paci and Pigliaru (1999), and López-Bazo et al. (1999), also draw attention to the specific structural role of agriculture in empirical analyses of convergence in European regions. There are of course also other structural phenomena of interest, but here we will concentrate on agriculture.

2.1 The impact of agricultural policy

The potential impacts of agricultural policy in the European Union on cohesion have been central in changes in the measures and component parts of the Common Agricultural Policy over the past decade. Cohesion is understood as accelerating regional economic growth in those parts of the EU with region GVA per capita markedly below that of developed regions, and thus most regions receiving cohesion support are agricultural regions, although not all agricultural regions are cohesion regions. The basic features of these measures, and changes taking place, including the Mac Sharry reforms and Agenda 2000, have recently been surveyed by Colman (2001), and links to regional policy are covered by Tondl (2001).

The second EU report on economic and social cohesion stresses the role of convergence, and concludes that specific measures will be needed to eliminate regional disparities (DG REGIO, 2001b). These will address differences in underlying conditions and factor endowments, among which labour force skills are seen as central. In addition, a preliminary study was devoted to the impact of community agricultural policies on cohesion (DG REGIO, 2001a). In this study, and more generally in the economics of tariff system structures, attention is drawn to difficulties in adequately measuring economic assistance.

Agricultural policy could be expected to interact with convergence in two ways. Firstly, and for practical reasons the only relationships to be explored empirically, one could expect the proportion of subsidised agriculture in a regional economy, and the intensity of the support, to influence the region’s growth rate negatively; for present purposes we assume that agriculture may be treated as though it were uniformly subsidised. The reasons for these negative relationships are that subsidies attenuate the movement of labour and capital to other sectors (and/or regions) with higher returns, conserving structures of factor allocation at the cost of those paying for the subsidies. The subsidies may also also be expected to reduce or to distort incentives to farmers to change their mixes of products and/or methods of production.

In this sense the subsidies are counterproductive as they hamper the growth in GVA. However the recent discussion about the so-called multifunctionality of agriculture may indicate that agricultural activities produce benefits over and above the market value of agricultural production. In economic terms agricultural production may have positive external effects on perceived public goods like the amenity value of the cultural landscape. If this is the case, and if agricultural subsidies are used as a means to internalise these externalities, growth is reduced only because we are measuring the wrong thing, traditional GVA instead of an extended GVA including the willingness to pay for such amenities. Whether or not this is the case is of course of vital importance for the policy implications of a negative relation between agricultural support and regional growth. For a recent paper addressing this question see Brunstad et al. (1999).

The second approach not followed up here, would be to consider the impact on speed of convergence of changes in agricultural policy regimes; for such an approach to be considered, regionalised agricultural accounts would be needed. Since they are not available in a systematic form at scales and levels of detail needed for Europe-wide analysis, micro-level studies would probably be required, such as panel studies, to cast light on the detailed relationships between different subsidy regimes and the embedding of agriculture in regional economies.

While governments granting subsidies to producers can account for them from public expenditure, other forms of subsidy cannot be as readily measured. In particular this applies to subsidies based on tariffs, quota systems, import bans, etc., where the transfer is carried out with consumers within the trade barrier system subsidising producers by an amount equivalent to the difference between the local price and the price of the same good delivered to that market from an external source at world market prices. Measurement complications and production distortions here also affect markets in intermediate goods.

The most commonly adopted approach at the national scale is to estimate agricultural assistance using producer subsidy equivalents, an approach used in a number of international organisations, and described in detail by Cahill and Legg (1989). The data requirements are however substantial, not just for price and quantity series for the chosen commodities, but also insight into the intermediate agricultural goods involved in regionally varying production processes. Some of the work required to estimate regional PSE series has been carried out in the DG REGIO study (2001a), yet more in Heckelei and Britz (2000). Since the present study is only intended to flag the importance of the agricultural sector, it was found more appropriate to use less adequate but more accessible data, rather than further reduce the number of regions under consideration or increase uncertainties associated with estimating or interpolating variables.

2.2 Sources of data in agriculture

European agricultural accounts are available in two versions, EAA 89/92 and EAA 97 REV.1.1, and regional versions at the level of NUTS level 2 are available from Eurostat[2]. The accounting data used here is for agriculture gross value added at market prices, subsidies, taxes linked to production (including VAT balance), and

[2] The data used here are taken from Eurostat theme: theme1, domain: regio, collect: agri-r, table: a2acct. We are grateful to Lucy McKeever of RCADE, University of Durham, for help in accessing this and other Eurostat and GISCO datasets.

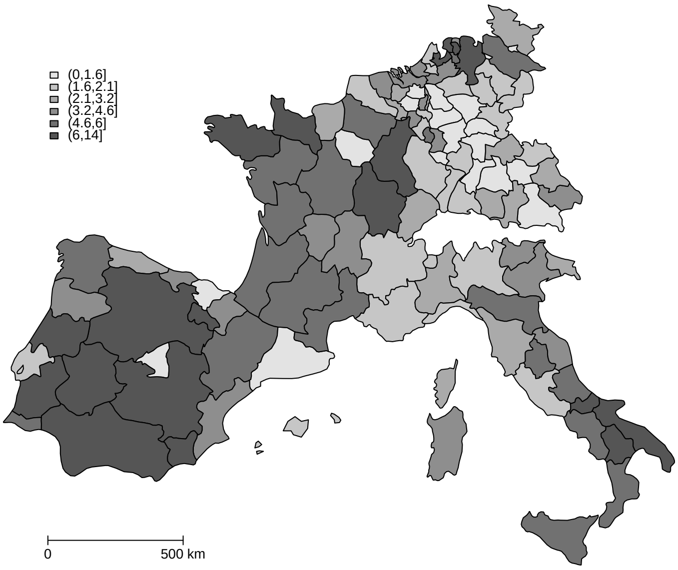

Figure 3: GVA in agriculture at factor costs (average 1988-90) as percentage of total GVA 1989.

gross value added at factor costs, for the period 1988-1998, in million ECU. From the middle of the period, there are very many missing values, and some countries either only report at NUTS level 1, or level 0. For this reason, The study area does not include Ireland, Denmark, Greece, UK, Berlin and former East Germany, and also drops overseas dependencies and Atlantic islands. It was further found necessary to aggregate three regions in Belgium (BE1 Brussels, BE24 Vlaams Brabant, and BE31 Brabant Wallon) in order to maintain the spatial series. In order to smooth the agricultural accounts data somewhat, and to accommodate further missing data problems, the variables to be used below are averages of values reported during 1988-1990, in most cases but not all, the averages of three values. [3]

Figure 3 shows average agricultural GVA measured at factor costs 1988-90 (in EAA 89/92 nomenclature, GVA at factor costs is GVA at market prices plus subsidies minus taxes linked to production) as a percentage of total GVA for 1989. The underlying total GVA and population datasets have been taken from Cambridge Econometrics' European Regional Databank, and are measured in 1990m ECU and 1000 persons [4]. The two datasets have been merged after dropping NUTS level 2 regions that could not be used because of missing agricultural accounts data. There is a potential problem of double-counting involved in using regional GVA series, because they can and most often do include subsidies as a component of gross value added. [5]

[3]It is unfortunate that the data are not more complete, because the choice of initial conditions means that some older members will appear to have higher support than newer members, and also because support to the older members declined during the chosen period while support to newer members probably rose.

[4]We would like to thank Cambridge Econometrics and Sasha Thomas for their help.

[5]In addition, changes introduced in revisions of the System of National Accounts, knocked on to EAA

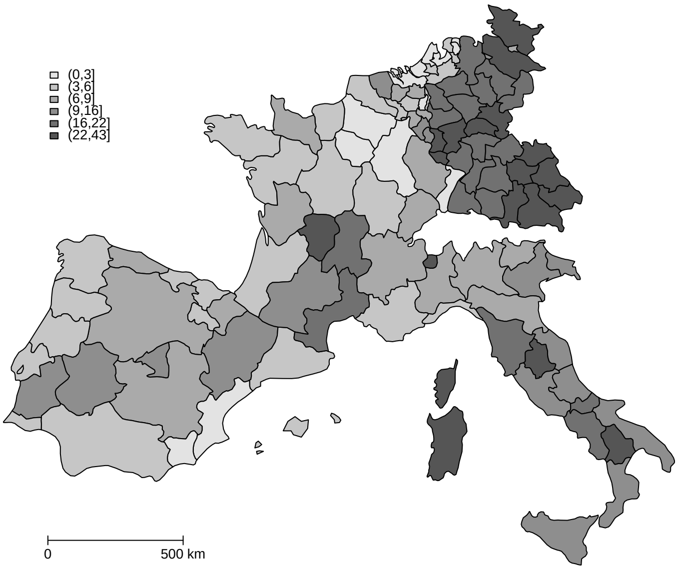

Figure 4: Relative subsidy levels 1988-90: agricultural subsidies (average 1988-90) as proportion of GVA in agriculture at market prices (average 1988-90).

Figure 4 displays the other variable to be used to represent the influence of agricultural policy, agricultural subsidies as a proportion of agricultural GVA at market prices. This is necessarily a very inadequate substitute for a properly constructed measure of producer subsidy equivalents, because it does not encompass the effects of tariff structures, which are not simply proportional to subsidy payments, but vary regionally with produced quantities of commodities. It has been chosen to restrict the agricultural variables to the 1988-90 period, partly because of missing data in the mid-1990's in agricultural accounts, but also because the base year date for the convergence model is 1989 - also chosen to match the years with least missing agricultural data.

3 Estimating convergence: specification issues

Attention has been drawn in a series of studies to specification problems found in estimating the standard convergence model using OLS. The roots of these problems are partly related to substantive spatial relationships, such as spillovers (Vayá et al., 1998), but may also involve missing variables, structural differences across the chosen study area, and functional form. Fingleton and McCombie (1998) and Fingleton (1999a, 2001) draw attention to the clear need to pay attention to specification, over and above the introduction of spatial econometric techniques also made by Vayá et al.

97, moving some subsidies into the definition of basic prices (OECD 2000); consequently the regional total GVA series may include varying amounts of subsidies depending on whether they are based on market prices or basic prices.

(1998), Baumont et al. (2001), and in the North American context by Rey and Montouri (1999) and Rey (2001). Details of the estimation methods and tests are not repeated here for brevity, and may be found in the cited articles.

Table 2: Modelling convergence 1989-1999 (t-values or z-values in parentheses)

	OLS	ML lag	OLS	ML lag
Intercept	0.0346	0.0147	0.0400	0.0203
	(6.56)	(2.86)	(5.52)	(2.91)
log GVA pc 1989	-0.00725	-0.00288	-0.00987	-0.00498
	(-3.55)	(-1.61)	(-3.99)	(-2.24)
% speed of convergence	0.753	0.292	1.039	0.511
log % agriculture 1988-90			-0.00200	-0.00133
			(-2.19)	(-1.68)
log subsidy/GVA 1988-90			-0.00138	-0.000699
			(-1.90)	(-1.11)
σ	0.00704	0.00597	0.00687	0.00591
log likelihood	407.8	421.1	411.7	422.8
AIC	-809.6	-836.1	-813.3	-835.7
Chow F	3.772		1.284	
p-value	0.026		0.281	
RESET F	3.921		1.573	
p-value	0.0226		0.187	
Breusch-Pagan	0.0626		2.195	
p-value	0.969		0.699	
Moran's I	0.332		0.308	
Z	5.596		5.377	
p-value	< 0.00001		< 0.00001	
Robust LM (err)	0.444		0.092	
p-value	0.505		0.762	
Robust LM (lag)	3.860		2.414	
p-value	0.049		0.120	
spatially lagged dependent variable		0.546		0.512
		(5.99)		(5.41)
LM (err)		2.00		0.485
p-value		0.157		0.486

Results from the estimation of the standard convergence model, with the annual log GVA per capita growth rate (1989-1999) as the dependent variable and log GVA per capita in 1989 as the independent variable, are presented in Table 2. The table also shows the results of estimating the same model including the spatially lagged dependent variable[6] using maximum likelihood, and of estimating an augmented model including the two agriculture variables defined above in section 2.2 for both estimation methods. The percentage speed of convergence for the OLS standard model, 0.75%, is comparable with that reported in Baumont et al. (2001, p. 26) of 0.84% for 1980-1995 and 138 regions.

Concentrating first on the OLS estimates of the standard model, it is worth noting that while the Breusch-Pagan test does not reject homoskedasticity[7], the RESET test (Johnson and DiNardo, 1997, p. 121) using the second, third, and fourth powers of the fitted values as extra variables indicates the presence of some specification error (the same result is found for all T=1990-1999 for base year 1989, and often for other base years). Figure 5 gives us a view of the relationship: it seems that the regions

[6]The spatially lagged dependent variable can be understood as the average value of neighbours of the region examinied, for some definition of neighbour and row-standardised neighbourhood weights.

[7]Tests of residual terms in all estimated models for non-normality were not significant.

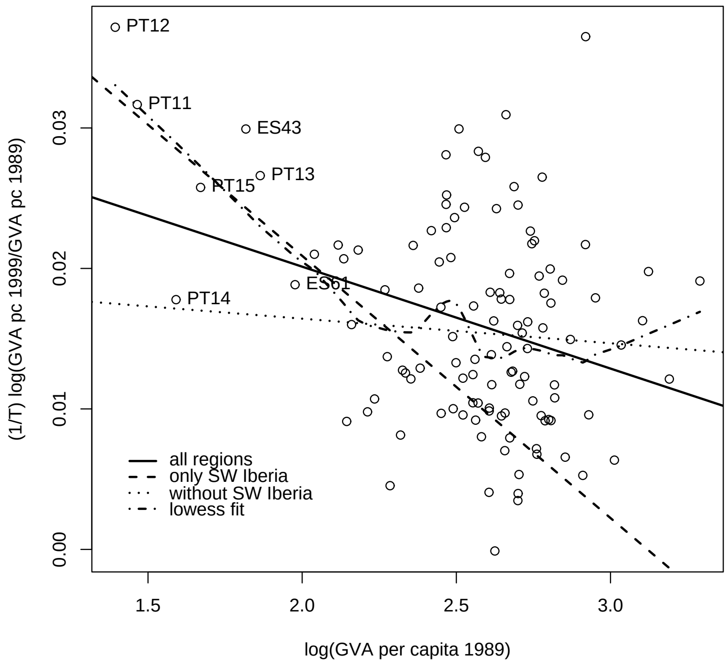

Figure 5: Plot of standard OLS convergence model (1989-1999) with regression fitted lines for all regions, only SW Iberia, all without SW Iberia, and a Lowess fit.

belonging to South West Iberia (the Portuguese regions and the Spanish regions of Extramadura and Andalucia) differ structurally from the remainder. They have much lower initial condition values and much higher growth rates, and the slopes of regression lines for all regions, just SW Iberia, and all without SW Iberia are quite different. The plotted Lowess fit (a robust local regression fit passing a moving window across the data set, see Cleveland, 1979) confirms that there are structural differences in the data set, which could be modelled by including the square of the independent variable, but may better be considered in a Chow test context (Johnson and DiNardo, 1997, p. 113-116). Subsetting the data set into SW Iberia and not SW Iberia, we can conduct a Chow test for structural difference in both slope and intercept coefficients, differences found to be significant as shown in Table 2. Again, this result applies to all T=1990-1999 for base year 1989, and for some other base years (1990, 1991, 1993, 1994 and 1995, and some T.

The fact that the SW Iberia constitutes a block of outliers is also seen in Figure 6, a Moran scatterplot of the annual log growth rates 1989-1999 (see also Rey and Montouri, 1999, p.150). This plot of the values of the variable in question on the x-axis against its spatially lagged values on the y-axis lets us see how the inclusion of the spatially lagged variable might give more insight into the convergence process, where growth rates would depend on the average of growth rates across neighbours of regions, possibly representing spillover or perhaps underbounding of functional

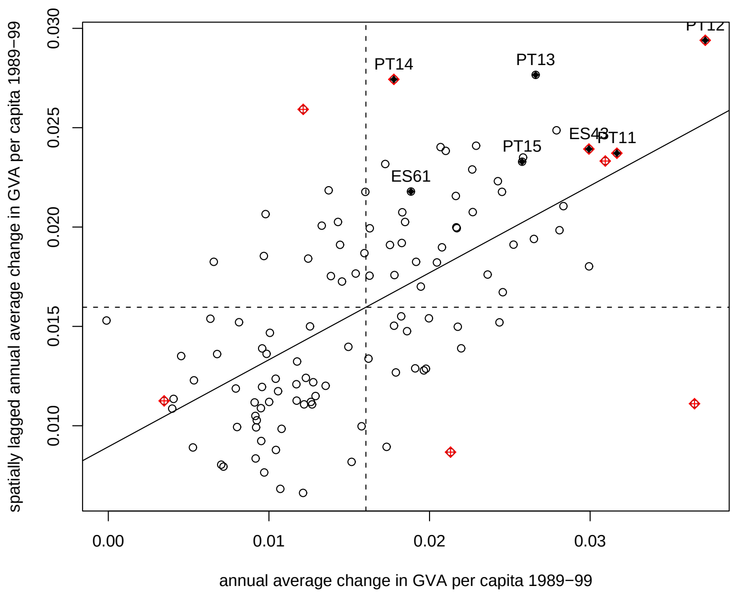

Figure 6: Moran scatterplot of convergence model dependent variable $1/T\log(y_{i,T}/y_{i,0})$

regions. Here neighbours have been defined using a sphere of influence graph [8] rather than boundary contiguity (Vayá et al. 1998) or distance (Baumont et al. 2001). Given a set of points representing the regions[9], a row standardised weights matrix was constructed and used in all analyses presented here.

The tests for spatial specification problems in the standard model, Moran's I (Cliff and Ord, 1981), and LM tests not reported here, are all highly significant. In addition, the robust LM test for a missing spatially lagged dependent variable in the presence of spatial error dependence is significant, but not the other way round (Anselin et al. 1996, Anselin and Bera 1998). It is not impossible that this lack of clarity in the results of the spatial specification tests is related to Fingleton's finding, that Moran's I may also detect spatial non-stationarity (1999b).

Following the augmentation of the standard model with the two agricultural variables, we can see that the non-spatial specification problems encountered in the standard model and reflected in the results of the Chow test and the RESET test are alleviated (column 3 in Table 2). Neither variable is strongly significant, but both have the expected signs, with both higher proportions of agriculture in regional GVA, and higher ratios of subsidies to agricultural regional GVA being associated with lower growth rates. The percentage speed of convergence is somewhat greater, just over 1%.

Stepping back to examine the ML lag estimates of the standard model, we can see that

[8]Thanks to Nicholas Lewin-Koh for this suggestion - his implementation is released as `soi.graph()` in the R contributed package `spweights`, available from http://cran.r-project.org.

[9]GISCO NUTS2 label points projected to Lambert Azimuthal Equal Area using EU standard parameters.

the inclusion of the spatially lagged growth rate reduces the percentage speed of convergence markedly, with the underlying coefficient value becoming much less significantly different from zero. The coefficient of the spatially lagged dependent variable is itself highly significant, and no residual spatial autocorrelation was found to remain in the model. Further details of estimation and testing methods may be found in Anselin (1988).

Introduction of the agricultural variables in the augmented ML lag model sees a further, but much less marked improvement in σ and log likelihood; however the value of the Akaike Information Criterion (AIC) rises a little compared to the standard ML lag model, suggesting possible interaction between the spatially lagged growth rate and the added variables. Adding the spatially lagged independent variables did not help, with AIC rising further to -831.2 (AIC for the standard ML lag model with spatially lagged initial conditions was -836.9).

4 Estimating convergence: exploring non-stationarity

As pointed out, Fingleton (1999b) has indicated that Moran's I may detect spatial non-stationarity in addition to residual autocorrelation. Because Moran's I continues to be significant in the augmented model, perhaps implying non-stationarity, and in the poor improvement of the augmented ML lag model over the standard ML lag model, it seems appropriate to explore the standard and augmented models using geographically weighted regression. Having originally been developed as a tool to exploratory spatial data analysis, its status is under revision at present, and it has been suggested that it may be used to indicate the presence of spatial non-stationarity in more formal ways. Exploratory methods have been used in the analysis of convergence models by Rey and Montouri (1999), Le Gallo and Ertur (2000), Rey (2001) and Arbia (2001).

4.1 Geographically weighted regression

Geographically weighted regression (GWR) is a technique for examining possible variability in coefficient estimates across study areas composed of regions represented by points, and was introduced by Brunsdon et al. (1996, a full description is found in Fotheringham et al. 2000). Instead of simply estimating global coefficient values over the whole data set, local parameter values are estimated for each region/point in the data set:

$$y_i = \beta_{i0} + \sum_{k=1}^{p} \beta_{ik} x_{ik} + \varepsilon_i$$

where $y_i, i = 1, n$ are observations of the dependent variable, $x_{ik}, i = 1, n, k = 1, p$ are observations of the p independent variables, ε_i are disturbance terms, and $\beta(i)$ are unknown parameter vectors which are functions of location i.

The local estimates, one regression for each region, are made using weighted regression, with the weights assigned to observations being a function of the distance between the region for which coefficient estimates are required and all the other regions. The distance function is directly comparable to those used in kernel density methods, and here use has been made of a bisquare function:

$$w(i)_{jj} = (1-(d_{ij}^2/d^2))^2, d_{ij} \leq d$$

where $w(i)_{jj} = 0$ when $d_{ij} > d$; $w(i)_{jj}$ is the weight assigned to observation j in estimating the parameters for observation i. The parameter d is termed the bandwidth, and for the bisquare function is the maximum distance for non-zero weights; it may be found by cross-validation and the value used here is 563 km, found by cross-validation for the augmented model. The parameters are estimated by weighted least squares:

$$\hat{\beta}(i) = [\mathbf{X}^T\mathbf{W}(i)\mathbf{X}]^{-1}\mathbf{X}^T\mathbf{W}(i)\mathbf{y}$$

where $\mathbf{W}(i)$ is the diagonal weights matrix for observation i; $\mathbf{X}^T\mathbf{W}(i)\mathbf{X}$ is assumed to be invertible.

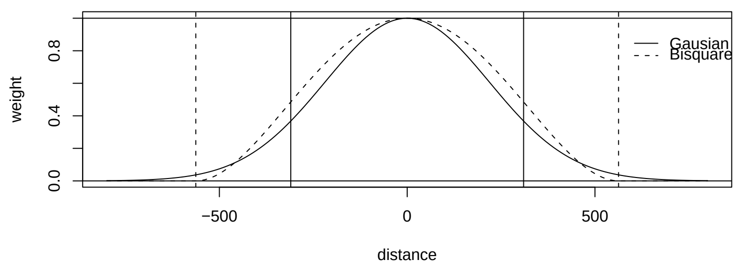

Figure 7: Values of weighting functions by distance for bandwidths chosen by cross-validation ofr the augmented model; bandwidth values shown by vertical lines. The very similar areas under the curves for the bisquare and Gaussian functions show that cross-validation has established adequate bandwidths in both cases.

Figure 7 shows the spatial range of this weighting scheme, and for comparison the Gaussian scheme and its equivalent cross-validated bandwidth (310 km). In current GWR techniques, bandwidth is first established globally, meaning that regions in denser parts of the study area have a larger number of weighted neighbours contributing to their estimates than in less dense parts. The bandwidth used here is about twice the greatest nearest neighbour distance in the data set (278 km), and between the lower quartile (420 km) and median (773 km) of all inter-region distances.

Figure 8 shows the range of coefficient values estimated for the percentage speed of convergence by GWR for the standard and augmented models, plotted with vertical

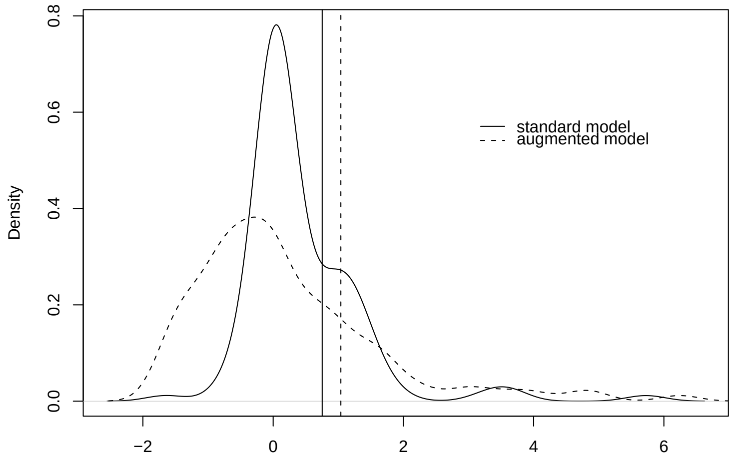

Figure 8: Density plots (smoothed histograms) of percentage θ (speed of convergence) coefficient estimates from standard and augmented GWR models; OLS model point estimates are shown by vertical lines (see Table 2). Bandwidths in both cases 563 km.

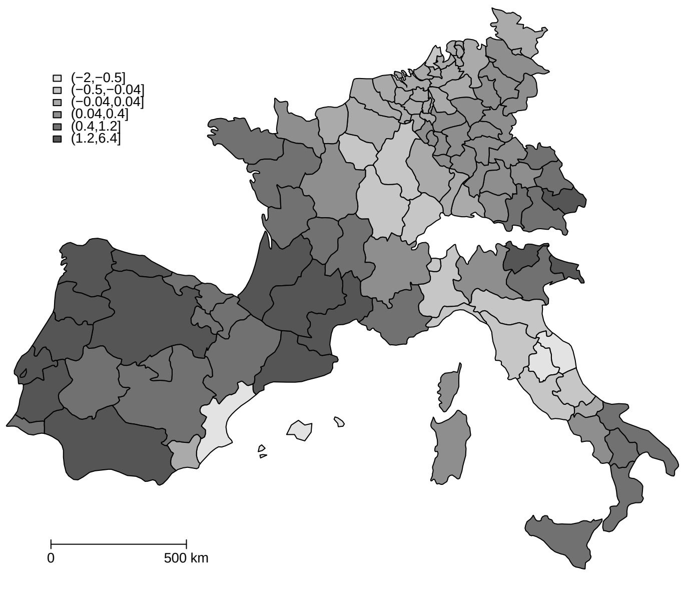

Figure 9: Percentage θ (speed of convergence) coefficient estimates from standard GWR model. Bandwidth 563 km, bisquare weighting function.

lines marking the global parameter estimates from the OLS models described above. As can be seen, in neither case do the GWR estimates group tightly or symmetrically around the global estimates. This is in itself not necessarily an issue, since some variation between multiples of standard errors is expected, but the extent and shape of the dispersion is worrying. For this weighting scheme (bisquare, bandwidth = 563km), it seems clear that in the standard model there is evidence of structural instability, with a peak just above zero, and a shelf centred at 1, with the global estimate at the root of the shelf. In the augmented model, the distribution is centred on zero, with many negative estimates, and a long right tail - the global estimate does not give a good description of the distribution at all. Of course here growth is also being driven (or rather restrained) by the values of the agricultural variables, but even so it is difficult to avoid the impression that non-stationarity may be present.

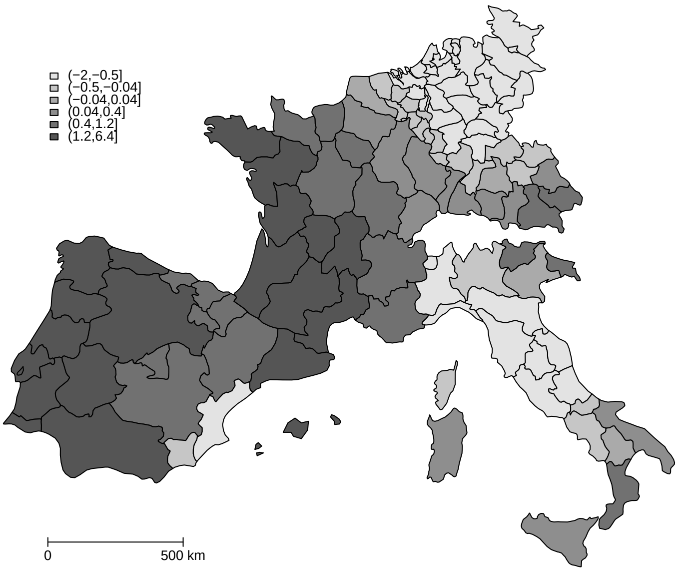

Figure 10: Percentage θ (speed of convergence) coefficient estimates from augmented GWR model. Bandwidth 563 km, bisquare weighting function.

Mapping the standard GWR percentage speed of change estimates (Figure 9), we can see that our tentative conclusion of structural differences is sustained. The banded West-East pattern is even more pronounced in the mapped values of estimates from the augmented GWR model when agricultural variable have been taken into account (Figure 10). Introducing the agricultural variables has a major impact in Germany, for which values of the agricultural subsidies variable are relatively high (see Figure 4). In both cases, had the maps not displayed spatial patterning, we could have risked concluding that the variability seen in Figure 8 was of little importance. What remains is to try to establish whether the GWR estimates add substantially to the global results from Table 2.

4.2 Testing stationarity using geographically weighted regression

So far, very few tests for GWR models have been published. Two, based on the use of analysis of variance, use generalised degrees of freedom (no longer integer because of the varying weights sums per estimate) to compare the improved sum of squares accounted for by the GWR estimates as compared with the global OLS estimates. Both Brunsdon et al. (1999) and Leung et al. (2000a) establish that a hat matrix may be constructed from rows from the GWR model, each row using its appropriate weights. The two F-tests they derive differ in that Brunsdon et al. (1999) compare the difference in the residual sums of squares of the OLS and GWR models with the residual sum of squares of the GWR model, while Leung et al. (2000a) compare the same difference with the residual sum of squares of the OLS model, choosing a different denominator for the F-ratio, as well as slightly different degrees of freedom Results of these tests are shown in Table 3, from which we can see that the GWR estimates for both standard and augmented models significantly reduce the residual sum of squares over and above the OLS estimates.

Table 3: Comparison of standard and augmented OLS and GWR convergence models.

	standard		augmented	
	OLS	GWR	OLS	GWR
σ	0.00704	0.00562	0.00687	0.00501
AIC	-809.6	-853.9	-813.3	-884.6
Brunsdon et al. ANOVA		2.85		2.16
p-value		0.0000		0.0003
Leung et al. F(2) ANOVA		2.20		1.64
p-value		0.0002		0.0102
Moran's I	0.332	0.122	0.308	0.062
Leung et al. p-value	< 0.00001	< 0.00001	< 0.00001	0.00148
Z under randomisation p-value	< 0.00001	0.0365	< 0.00001	0.2624

Table 3 also contains estimates of σ and the Akaike Information Criterion, based on work in progress reported in Brunsdon et al. (2000). Here again the properties of the GWR hat matrix are used to derive the residual sum of squares of the GWR model, and the appropriate degrees of freedom for the AIC measure. The reported values suggest that the level of non-stationarity present in the data under analysis is such that GWR does provide a more adequate representation than global parameter estimates, even when structurally-varying agricultural variables are included. Finally, following Brunsdon et al. (1998), the spatial autocorrelation of the GWR residuals is reported. The results of the test procedure described by Leung et al. (2000b), a three-moment χ^2 approximation, is shown, together with the probability values of Moran's I under randomisation ignoring the fact that we are testing regression residuals. Both tests, and the estimated values of Moran's I statistic, suggest that the GWR models display less global spatial autocorrelation in their residuals than the OLS models, and that the GWR model augmented with agricultural variables performs best.

5 Conclusions and directions for further research

Following up the impacts of agricultural policy on the estimation of empirical models of regional convergence does seem worthwhile, because on balance more of the variance of the average annual regional growth rates is accounted for in the augmented model than the standard model estimated by OLS, ML lag, and GWR. Since it is possible that the non-stationarity is related to further missing variables, including the inadequacy of the way in which agricultural subsidies are represented, it will be important to try to replace the agriculture variables with better estimates of producer subsidy equivalents for the base year. An issue that deserves further attention is how to handle changes in agricultural policy regime occurring between years 0 and T. On balance however, we would argue that this exploratory analysis does give support to the role of agricultural subsidies in accounting for variations in regional growth. Coefficient estimates for the agricultural variables in the OLS and ML lag global models had the expected negative signs, and were either significant or almost significant at conventional levels.

On the technical side, some further tests for specification problems in the spatial context are becoming available, but have not been applied to convergence models yet (Baltagi and Li, 2001, de Graff et al., 2001). The tests on GWR estimates also need to be more firmly established, although some Monte Carlo studies on the convergence models reported above did not give cause for concern that the conclusions drawn were inappropriate. The GWR results also need to be tested for spatial autocorrelation, and if necessary to accommodate a spatially lagged dependent variable in some form, although GWR does already involve a spatial weighting of the observations themselves.

It is clear that addressing concerns about non-stationarity is one of the major tasks which has to be tackled in order to promote spatial econometrics; it may also be the case that panel estimation techniques or multi-level modelling will provide further tools that may be employed to test for specification adequacy. It does seem to be clear, though, that the full spatial econometrics toolbox, including both exploratory and confirmatory techniques, needs to be applied to the unravelling of the stories involved in European regional growth.

References

Anselin L (1988) Spatial econometrics: methods and models. Kluwer, Dordrecht

Anselin L, Bera AK (1998) Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah A, Giles DEA (eds) Handbook of applied economic statistics. Marcel Dekker, New York, pp. 237-289

Anselin L, Bera AK, Florax R, Yoon MJ (1996) Simple diagnostic tests for spatial dependence. Regional Science and Urban Economics 26: 77-104

Arbia G (2001) The role of spatial effects in the empirical analysis of regional concentration. Journal of Geographical Systems 3: 271-281

Baltagi BH, Li D (2001) LM tests for functional form and spatial error correlation. International Regional Science Review 24: 194-225

Baumont C, Ertur C, Le Gallo J (2001) A spatial econometric analysis of geographic spillovers and growth for European regions, 1980-1995. (Unpublished working paper 2001-04, LATEC UMR, University of Burgundy)

Brunsdon C, Fotheringham AS, Charlton ME (1996) Geographically weighted regression: a method for exploring spatial nonstationarity. Geographical Analysis 28: 281-298

Brunsdon C, Fotheringham AS, and Charlton, ME (1998) Spatial nonstationarity and autoregressive models. Environment and Planning A 30: 957-973

Brunsdon C, Fotheringham AS, and Charlton, ME (1999) Some notes on parametric signficance tests for geographically weighted regression. Journal of Regional Science 39: 497-524

Brunsdon C, Fotheringham AS, Charlton ME (2000) Geographically weighted regression as a statistical model. (Unpublished working paper, Department of Geography, University of Newcastle)

Brunstad RJ, Gaasland I, Vårdal E (1999) Agricultural production and the optimal level of landscape preservation. Land Economics 75: 538-546

Cahill C, Legg W (1989) Estimation of agricultural assistance using producer and consumer subsidy equivalents: theory and practice. OECD Economic Studies 13: 13-43

Cleveland WS (1979) Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association 74, 829-836.

Cliff AD, Ord JK (1981) Spatial processes. Pion, London

Colman D (2001) The Common Agricultural Policy. In: Artis M, Nixson F (eds) The economics of the European Union. Oxford University Press, Oxford, pp. 97-124

DG REGIO (2001a) Study on the impact of community agricultural policies on economic and social cohesion. (http://www.inforegio.cec.eu.int/wbdoc/docgener/studies/pac_en.htm, accessed 2 Jan 2002)

DG REGIO (2001b) Unity, solidarity, diversity for Europe, its people and its territory: Second report on economic and social cohesion. (http://www.-inforegio.cec.eu.int/wbdoc/docoffic/official/reports/contentpdf_en.htm, accessed 2 Jan 2002)

Fagerberg J, Verspagen B, Caniels M (1997) Technology, growth and unemployment across European regions. Regional Studies 31: 457-466

Fingleton B (1999a) Estimates of time to convergence: an analysis of the regions of the European Union. International Regional Science Review 22: 5-34

Fingleton B (1999b) Spurious spatial regression: some Monte Carlo results with a spatial unit root and spatial cointegration. Journal of Regional Science 39: 1-19

Fingleton B (2000) Spatial econometrics, economic geography, dynamics and equilibrium: a 'third way'? Environment and Planning A 32: 1481-1498

Fingleton B (2001) Equilibrium and economic growth: spatial econometric models and simulations. Journal of Regional Science 41: 117-147

Fingleton B, McCombie JSL (1998) Increasing returns and economic growth: some evidence for manufacturing from the European Union regions. Oxford Economic Papers 50: 89-105

Fotheringham AS, Brunsdon C, Charlton ME (2000) Quantitative geography: perspectives on spatial data analysis. Sage, London

de Graff T, Florax RJGM, Nijkamp P, Reggiani A (2001) A general misspecification test for spatial regression models: dependence, heterogeneity, and nonlinearity. Journal of Regional Science 41: 255-276

Heckelei T, Britz W (2000) Concept and explorative application of an EU-wide, regional agricultural sector model (CAPRI project). (http://www.agp.uni-bonn.de/agpo/rsrch/capri/eaaecapri.pdf, accessed 2 Jan 2002)

Johnston J, DiNardo J (1997) Econometric methods. McGraw-Hill, New York

Le Gallo J, Ertur C (2000) Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980-1995. (Unpublished working paper 2000-09, LATEC UMR, University of Burgundy)

Leung Y, Mei CG, Zhang WX (2000a) Statistical tests for spatial non-stationarity based on the geographically weighted regression model. Environment and Planning A 32: 9-32

Leung Y, Mei CG, Zhang WX (2000b) Testing for spatial autocorrelation among the residuals of the geographically weighted regression. Environment and Planning A 32: 871-890

López-Bazo E, Vayá E, Mora AJ, Suriñach J (1999) Regional economic dynamics and convergence in the European Union. Annals of Regional Science 33: 343-370

OECD (2000) Eurostat’s EAA97: Main changes and the impact on the EAA-data. OECD Statistics Directorate STD/NA/AGR(2000)10 (http://www1.oecd.org/std/AGRMeet2000/docs/naa00_10.pdf, accessed 2 Jan 2002)

Paci R (1997) More similar and less equal: economic growth in the European regions. Weltwirtschaftliches Archiv 133: 609-634

Paci R, Pigliaru F (1999) Is dualism still a source of convergence in Europe? Applied Economics 31: 1423-1436

Pons-Novell J, Viladecans-Marsal E (1999) Kaldor’s Laws and spatial dependence: evidence for the European regions. Regional Studies 33: 443-451

Rey SJ (2001) Spatial empirics for economic growth and convergence. Geographical Analysis 33: 195-214

Rey SJ, Montouri BD (1999) US regional income convergence: a spatial econometric perspective. Regional Studies 33: 143-156

Tondl G (2001) Regional policy. In: Artis M, Nixson F (eds) The economics of the European Union. Oxford University Press, Oxford, pp. 180-212

Vayá E, López-Bazo E, Artís M (1998) Growth, convergence and (why not?) regional externalities. (Unpublished working paper E98/31, "Anàlisi Quantitativa Regional" Research Group, University of Barcelona)