Regional unemployment clusters

Nearness matters within and across Europe's borders

SUMMARY

High unemployment and regional inequalities are major concerns for European policy-makers, but so far connections between policies dealing with unemployment and regional inequalities have been few and weak. We think that this should change. This paper documents a regional and transnational dimension to unemployment – i.e., geographical unemployment clusters that do not respect national boundaries. Since the mid 1980s, regions with high or low initial unemployment rates saw little change, while regions with intermediate unemployment moved towards extreme values. During this polarization, nearby regions tended to share similar outcomes due, we argue, to spatially related changes in labour demand. These spatially correlated demand shifts were due in part to initial clustering of low-skilled regions and badly performing industries, but a significant neighbour effect remains even after controlling for these, and the effect is as strong within as it is between nations. We believe this reflects agglomeration effects of economic integration. The new economic geography literature shows how integration fosters employment clusters that need not respect national borders. If regional labour forces do not adjust, regional unemployment polarization with neighbour effects can result. To account for these 'neighbour effects' a cross-regional and transnational dimension should be added to national anti-unemployment policies. Nations should consider policies that encourage regional wage setting, and short distance mobility, and the EU should consider including transnational considerations in its regional policy, since neighbour effects on unemployment mean that an anti-unemployment policy paid for by one region will benefit neighbouring regions. Since local politicians gain no votes or tax revenues from these 'spillovers', they are likely to underestimate the true benefit of the policy and thus tend to undertake too little of it.

— Henry G. Overman and Diego Puga

Economic Policy April 2002 Printed in Great Britain

Unemployment clusters across Europe's regions and countries

Henry G. Overman and Diego Puga

London School of Economics; University of Toronto

1. INTRODUCTION

Reducing unemployment and redressing regional inequalities are two key challenges currently facing Europe. The Luxembourg process, launched at the 1997 European Job Summit, emphasizes reform of national labour market policies as the key to lowering unemployment. The European Union's structural spending – which accounts for one-third of the EU budget – aims to reduce regional inequalities, mainly income inequalities. So far, connections between policies dealing with unemployment and regional inequalities have been few and weak. This should change. In this paper, we document an important regional and transnational dimension to unemployment, and argue that this 'neighbour effect' means that national-level reforms envisaged in the Luxembourg process are unlikely to be fully successful.

The Luxembourg process's emphasis on national reforms reflects the fact that one typically thinks about differences in unemployment rates as differences across countries. However, national averages hide large differences in unemployment rates across regions

We thank Carlos Gradín for help in calculating the polarization measure, and Danny Quah for making his econometric shell tsrf publicly available. We are also grateful to Gilles Duranton, Maia Güell, Hubert Jayet, Angelo Melino, Andrew Oswald, Nadia Soboleva, Dan Trefler, the Editors and Panel Members of this journal, two referees, and participants at various seminars for helpful discussions. This paper was written in part while both authors were visiting the Norges Handelshøyskole. Funding from the United Kingdom Economic and Social Research Council, the Social Sciences and Humanities Research Council of Canada, and from the Connaught Fund and the Joint Research Initiative of the University of Toronto is gratefully acknowledged.

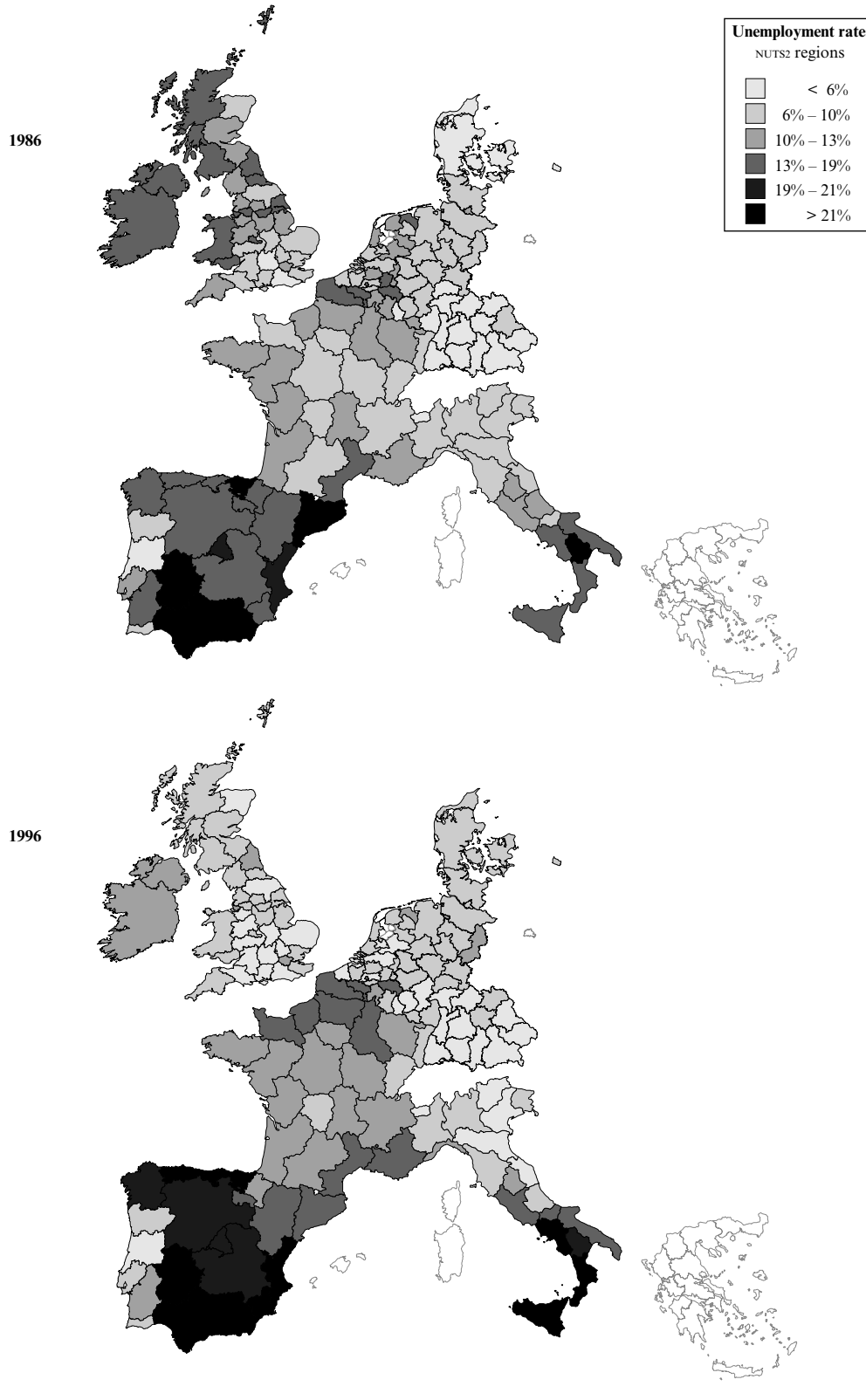

Figure 1. EU regional unemployment rates, 1986 and 1996

within countries. The case of Italy is best known, with Campania having a 1996 unemployment rate 4.4 times as high as Valle d'Aosta, but large regional differences exist in all European countries. In the UK, Merseyside had an unemployment rate 3.2 times that of the Surrey-Sussex region in 1996, and in Belgium, the unemployment rate of Hainut was 2.2 times that of Vlaams Brabant; in France, Languedoc-Roussillon had a rate twice that of Alsace; and so on.

Moreover, differences in regional unemployment have evolved importantly since the mid 1980s. Figure 1 depicts regional unemployment rates for the contiguous European Community in 1986 and a decade later. Roughly speaking, regions with unemployment rates that were either very high or very low relative to the EU average tended to stay that way, but many regions with intermediate initial unemployment rates saw their rates either rise markedly or fall markedly from 1986 to 1996. Thus while aggregate unemployment traced out a full cycle between 1986 and 1996,[1] the spatial pattern of regional unemployment rates became more polarized (Section 2 is devoted to confirming this visual impression more thoroughly).

1.1. Two border regions in Belgium

Before turning to the argument and a thorough look at the data, we illustrate regional and transnational aspects of unemployment clustering using the example of two border regions in Belgium.

In 1986 the Belgian region of Limburg had an unemployment rate 1.2 times the Belgian average and 1.3 times the EU average. By 1996 its unemployment rate had fallen below both the Belgian and EU averages. Just across the border from Limburg (Belgium), two Dutch regions had similar experiences. The unemployment rates of Limburg (Netherlands) and Noord-Brabant fell relative to both the Dutch and EU averages.

Back in Belgium, 90 km south-west of Limburg and on the border with France, the region of Hainaut started with a similar unemployment rate in 1986. However, instead of falling as it did in Limburg, this rate rose both in absolute terms and relative to both the Belgian and EU averages. Just across the border from Hainaut, the French region of Nord-Pas de Calais also saw its unemployment rate increase in both absolute and relative terms.

The different fortunes of these two Belgian regions were not driven by changes in their labour forces. Both regions had growing labour forces, but Limburg's actually grew more than twice as fast. The reason for Limburg's fall in unemployment is that its employment grew even faster than its labour force, and over four times faster than Hainaut's. A similar process occurred in the two Dutch neighbours of the Belgian

[1] The average European unemployment rate in 1986 (for regions belonging to what was then the European Economic Community) was 10.7%, starting to come down from a peak of 10.8% one year before that. It kept coming steadily down to 8.1% in 1990, and then steadily up to a new peak of 11% in 1994, after which it fell back to its 1986 rate of 10.7% in 1996.

Limburg. These regions that did relatively well had large and growing labour forces. But they also had a rate of employment growth that more than matched their labour force growth, and that brought their unemployment rates down. By contrast Nord-Pas de Calais, the French neighbour of Hainaut that did relatively badly, lost employment while its labour force was rising.

The drop in Limburg's unemployment rate versus Hainaut's rise cannot be put down to differences in the skill composition of their labour force. Both these Belgian regions had a similar percentage of their population with less than upper secondary education. And the French region of Nord-Pas de Calais, despite having a smaller fraction of people with less than upper secondary education than either of the Belgian regions, had a worse unemployment outcome.

Further, the evolution of these regions was not due to their different initial sectoral composition. Admittedly in 1986 Nord-Pas de Calais was a predominantly industrial region. But Hainaut also saw its unemployment rate rise and in 1986 was concentrated in services. In contrast, the Belgian success story, Limburg, was concentrated in industry and of its two neighbours, one was mainly industrial (Noord-Brabant), the other service based (Limburg). No simple story of sectoral changes explains the relative performance of these regions. Possible differences between the Flemish and French-speaking regions of Belgium cannot explain these changes either. Contiguous to both the Flemish speaking Belgian Limburg and to the Dutch Limburg is the French-speaking Belgian region of Liège, which also experienced a reduction in its unemployment rate.

Given the small flows of workers across these borders, both in terms of commuting and permanent moves, one can hardly argue that there are functional labour markets extending across these regions. However, firms do seem to find it attractive to exploit other advantages of location close to these borders, such as the ability to use suppliers from different countries. The areas on the borders between Belgium and France and Belgium and the Netherlands have provided traditional locations for industry. In recent years, however, these two borders have experienced very different evolutions. The most publicized case came in 1997 as Renault announced the closure of its Vilvoorde plant in Belgium. This raised protests at the loss of 3,100 jobs, at a time when Renault was planning to expand operations in other parts of Europe. At about the same time in Limburg (Netherlands), Volvo introduced a three-shift working schedule in its Nedcar plant, to double production over the following three years, drawing on suppliers from both sides of the Belgian–Dutch border. And on the Belgian side of this border, General Motors was also expanding production at its Antwerp plant.

Starting from similar intermediate unemployment rates, the Belgian regions Limburg and Hainaut have moved towards opposite extremes of the European distribution, but in each case have gone along with their foreign neighbours. Below we show that this story is representative of a broad pattern that has developed across Europe.

But what does this mean for policy?

1.2. Why should we care?

We will argue that policy needs to be broadened to reflect the fact that regional unemployment has a strong geographical component that goes *across* national borders. We build our argument in three steps. First we develop a new, highly flexible tool that allows us to show that while there is a national component to unemployment, this is insufficient to explain the distribution of EU-wide unemployment rates. Regions have unemployment outcomes that are closer to neighbouring regions than to other regions in the same nation. Our methodology allows us to look at different parts of the distribution separately and indicates that this geographical component dominates most clearly for regions with intermediate levels of unemployment. These are the regions that have driven polarization. This finding – that 'nearness' matters – is consistent with polarization being driven by changes in demand or supply that are similar across neighbouring regions. Our second step is to show that polarization has been demand rather than supply driven (indeed, supply changes mitigated polarization).

Our third step is to discriminate between possible sources of similar demand changes across neighbouring regions. Regions may be similar in terms of the sectoral composition of their employment or in the age, sex and skill structure of their populations. Regions initially specialized in agriculture or manufacturing may have seen their unemployment rates rise as the EU production structure moves away from those sectors. Similarly, regions with a high proportion of low-skilled workers may have seen their unemployment rates rise as production shifts from low-skilled to high-skilled employment. If these types of regions were geographically concentrated this would help explain our findings on a strong geographical component to unemployment. Using standard statistical analysis, we show that the geographical concentration of these characteristics does matter, but that nearness continues to matter even after we control for the impact of these characteristics. Most surprisingly, the neighbour effect seems to work as powerfully across national borders as it does within them.

1.3. Policy implications

What does all this tell us about the policy measures needed to deal with the polarization in EU regional unemployment rates? The fact that regional characteristics matter for unemployment outcomes is consistent with a framework where real wages do not fully adjust to reflect regional economic conditions. In such a framework, the geographical dimension to regional unemployment simply reflects the fact that neighbouring regions face similar economic conditions. Thus, the policy conclusions to be drawn from our analysis depend crucially on why economic conditions tend to be similar across neighbouring regions.

Our results suggest that some of the similarity in changes in demand across neighbouring regions is due to clusters of low-skilled regions and badly performing industries. However, a significant neighbour effect remains after controlling for these

factors. We believe that this may reflect a process of relocation across the EU with ever-closer economic integration acting as the catalyst. Recent theoretical developments – the so-called new economic geography – suggest that such a process can be associated with the emergence of spatial concentrations of employment, and that with falling barriers to trade these may extend across national borders. If regional labour forces do not fully adjust to such employment changes, then geographical location may be important in explaining the increased polarization of unemployment rates.

Two types of policy conclusions emerge from our analysis. The first relates to the spatial scale at which these unemployment clusters occur. The second relates to the fact that unemployment clusters extend across national borders. We return to these policy conclusions in Section 5 once we have outlined our empirical evidence and analytical framework.

2. THE POLARIZATION OF EUROPEAN REGIONAL UNEMPLOYMENT RATES

Characterizing changes in spatial patterns involves some non-standard tools, so we shall have to spend some energy explaining these tools before using them. We start with the notion of relative unemployment and the definition of regions.

2.1. Relative unemployment and the definition of regions

Being interested in the spatial distribution of unemployment, our key variable will be regional unemployment rates relative to the average unemployment rate in all European regions. Although it may seem odd to look at relative unemployment, an analogy with income distributions shows that it is really the natural thing to study. When studying income distribution, one can consider incomes per individual in absolute terms. Alternatively, one can study individual incomes normalized by the population-wide income average. Although there are merits to using the absolute income distribution, it is more natural to use relative incomes when considering changes in income distributions over time (this neutralizes overall changes in income levels). In the same way, relative unemployment rates allow us to abstract from overall changes in unemployment rates.

We must also address the question of spatial categories. Plainly, we would like to have as fine a geographic partition of Europe as possible, but we are constrained by data to looking at what is known as the NUTS2 level. This is the second least disaggregated level of Eurostat's hierarchical classification, Nomenclature of Territorial Units for Statistics, known by its French acronym NUTS. Level 2 regions are moderately large sub-national regions (the regions in Figure 1 are NUTS2 regions). For example, the Ile-de-France in France, Piemonte in Italy and, Saarland in Germany are individual NUTS2 regions. The average NUTS2 region in our data set had a land area of 13,800 sq. km and a population of 2.1 million in 1996.

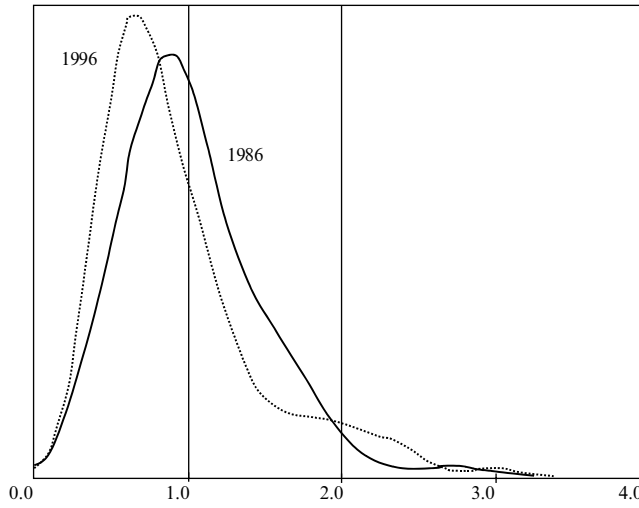

Figure 2. Distributions of relative European regional unemployment, 1986 versus 1996

Technical note: The plots are densities and can be interpreted as the continuous equivalent of a histogram, (i.e. where the number of intervals has been let tend to infinity and then to the continuum). All densities are calculated non-parametrically using a Gaussian kernel with bandwidth set as per Section 3.4.2 of Silverman (1986). The range is restricted to the positive interval using the reflection method proposed there.

Source: Authors' calculations on Eurostat NUTS2 data for 1986 and 1996.

The final data issue – one that is unique to our spatial focus – concerns the issue of neighbours. Our main focus is on neighbour effects, be they domestic or international, so we limit our data to regions that have European neighbours. This leads us to exclude all of Greece (which has no land borders with other EU nations as well as many data problems) and a series of islands, such as the Balearic Islands and Guadeloupe. After this selection, we have 150 NUTS2 regions to work with.

Our data runs from 1986 to 1996; 1996 was the latest year available to us and regional unemployment data prior to 1986 is limited and what there is suffers from serious comparability problems across nations.

2.2. Increasing spatial inequality

Figure 1 already hinted at an increase in inequality of Europe's regional unemployment rates. Here we want to verify this with techniques that allow more precise conclusions. Figure 2 plots the distribution of regional unemployment rates relative to the average of all regions – what we call the 'Europe relative' unemployment rates. The solid line shows the distribution in 1986, the dashed line shows it in 1996. To read this type of diagram, note that 1.0 on the horizontal axis indicates the European average unemployment rate, 2.0 indicates twice the average, and so on. The height of the curve over any point gives the probability that any particular region will have that relative rate. Since the height of the curve at any particular point gives the probability,

the area under the curve between, say, 0.0 and 1.0, gives the total likelihood that a region will have a relative unemployment rate that is between 0.0 and 1.0.

Two features of these curves show that the distribution of regional unemployment rates has become more uneven. First, the diagram tells us that the likelihood that any particular region's rate is below average has increased from 1986 to 1996. We see this by noting that the area under the 1996 curve that is to the left of the average, 1.0, is greater than the corresponding area under the 1986 curve. Secondly, there is also an increase in the probability of a region having more than twice the European average. Again we see this by noting that the area under the 1996 curve that is to the right of 2.0 exceeds that of the 1986 curve. Thus over time more regions have experienced unemployment rates below the European average, or above twice that average, and less regions have unemployment rates between the average and twice the average. This provides a clear indication that the spatial distribution of unemployment has been more unequal. In our Web Appendix (which can be found on http://www.economic-policy.org) we provide the distributions for 1989 and 1993 as well; these show that the increased spatial inequality has slowly evolved over the decade.

2.3. Polarization of regional unemployment

Figure 2 shows that there were more regions with very low or very high unemployment rates in 1996 than in 1986, and fewer regions with intermediate rates. But what caused this? Does it result from a structured evolution of regional unemployment, or do these snapshots merely show churning of the unemployment rate distribution – the random ups and downs of regional fortunes? To show that there is indeed a structured process in which regions with either high or low unemployment rates have seen little change while regions with intermediate rates have moved towards the extremes of the distribution, we must track the evolution of each region's relative unemployment rate over time. With 150 regions, doing this would be unwieldy, so we rely on a technique – i.e. a transition matrix, or transition table – that lets us summarize the behaviour of all regions in a handful of numbers.

A transition matrix categorizes the various unemployment rate outcomes into a manageable number of ranges, and lists these ranges both across the top and down the side of the matrix. Any cell of the matrix corresponds to a pair of ranges – the 'row range' and the 'column range' – and number in each cell gives the probability that a region that started out in the row range ends up in the column range. Of course, every region ends in one column range or another, so the probabilities along each row sum up to 100%.

Table 1 reports the transition probability matrix linking the 1986 and 1996 distributions of European relative unemployment rates. Reading along the bottom row of the matrix, we observe strong persistence for regions starting with an unemployment rate below 0.6 times the European average. By 1996, 81% remained below 0.6 times the European average, and none had a relative unemployment rate higher than 75%

Table 1. Polarization evidence from the 1986 to 1996 transition matrix

	1996 ranges of relative unemployment rates (%)				
	0.0–0.6	0.6–0.75	0.75–1.0	1.0–1.3	1.3–highest
1986 ranges:					
1.3–highest	0	0	16	22	62
1.0–1.3	6	22	34	19	19
0.75–1.0	24	29	26	21	0
0.6–0.75	52	26	9	9	4
0.0–0.6	81	19	0	0	0

Notes: In 1986, there were 32 regions in the 1.3–highest range, with the corresponding number for the other ranges being (in descending order): 32, 42, 23 and 21. The ranges were chosen to give a similar number of regions per range while keeping the ranges as round numbers.

Source: Authors' calculations using Eurostat data.

of the average. Jumping to the top row we also see strong persistence among the regions with highest unemployment rates: of the regions with an initial unemployment rate above 1.3 times the European average, 62% remained above 1.3 times the average, while 22% moved to between the average and 1.3 times the average. However, regions with unemployment rates between 0.6 and 1.3 times the European average (second, third and fourth rows from the bottom) experienced much greater mobility – regions with initial unemployment rates between 0.75 and 1.0 times the average ended up almost equally distributed across the four intervals between 0 and 1.3 times the average. The second to last tells us that of regions in the 0.6 to 0.75 range, 26% remained in that range, while 52% saw their unemployment rate fall below 0.6 times the average.

Taken together, the Table 1 results confirm that the Europe's regions have become polarized in terms of their unemployment experiences.

2.3.1. Continuous 'transition matrices'. Relative unemployment rates are, by nature, a continuous variable and this means that any categorization of unemployment rates into specific ranges is arbitrary. Experience from the study of income distributions shows that this arbitrariness can matter in the sense that setting out different ranges may lead us to draw different conclusion about what actually happened. In addition, many interesting details are lost when we limit ourselves to a handful of ranges.

Fortunately, there is a tool that addresses these shortcomings. We can calculate what amounts to a transition matrix with an infinite number of infinitely small ranges. The results from this tool – known by its technical name 'stochastic kernel' – are displayed as a three-dimensional diagram, or a two-dimensional contour map. Figure 3 shows the results for the transition from the 1986 distribution of European relative unemployment rates to the 1996 distribution. Cutting across the kernel for any given 1986 unemployment rate is like reading across one row of Table 1 in that

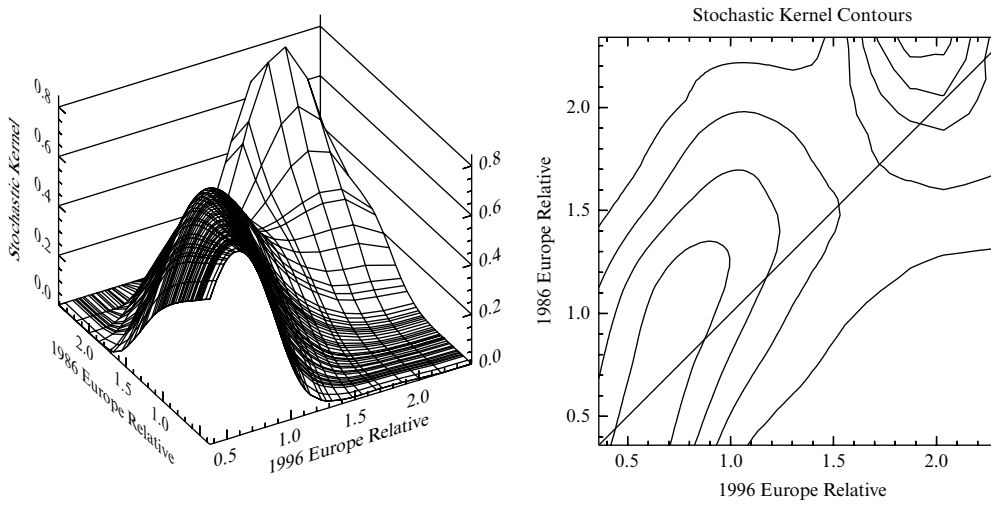

Figure 3. Polarization, evidence from a stochastic kernel

it gives us the probability distribution of 1996 unemployment rates for regions that start with that value in 1986. For this reason, a 'peak' in the kernel corresponds to a high number in the transition matrix while a 'valley' is the equivalent of a low number.

The plot on the right-hand side of the figure is a contour plot of the three-dimensional kernel on the left. The contour plot works in exactly the same way as on a standard geographical map. Lines on the contour plot connect points at the same height on the three-dimensional kernel. A straight line is drawn in the contour plot to mark the diagonal, where all mass would be concentrated if there was complete persistence in the distribution.

The twin-peak nature of Figure 3 confirms that there has been a polarization of regional unemployment rates. Regions that had a low unemployment rate relative to the European average in 1986 tended to maintain or reduce their unemployment rate over the next decade. Regions that had a high unemployment rate relative to the European average in 1986 still tended to have a relatively high unemployment rate in 1996. However, regions with intermediate unemployment rates were unlikely to remain there; most saw their Europe-relative rate either fall or rise.

2.3.2. Measuring polarization. In addition to these visual methods, we can calculate more formal measures of inequality and polarization. For instance, one very common measure of inequality – the so-called Gini coefficient – rose by 19%, from 0.236 in 1986 to 0.281 in 1996. However, Figure 3 suggest that the most significant change between 1986 and 1996 has been not so much an increase in inequality, but rather the polarization of regions into two groups – one with low unemployment and one with high unemployment. To quantify this, we calculate a polarization measure

that was introduced by Esteban *et al.* (1999).[2] Between 1986 and 1996 polarization thus measured increased by 37%, from 0.096 to 0.131.

3. A SIMPLE FRAMEWORK FOR UNDERSTANDING REGIONAL UNEMPLOYMENT POLARIZATION

The geographic polarization of European unemployment rates demonstrated above is a new fact. To better understand its causes and suggest appropriate policy responses, we need a simple framework to organize our thinking about how such a phenomenon could arise. For this purpose, a slightly modified version of the textbook model of European unemployment by Burda and Wyplosz (2001) serves well.

In the Burda–Wyplosz framework, unemployment is caused by a failure of wages to adjust to the point where labour supply matches labour demand. There are a number of well-known reasons for this, but for the sake of concreteness we suppose that wages are determined by national bargaining between a trade union and an employers' association, and – to keep matters simple – we suppose that the union holds all bargaining power.

The situation is depicted in Figure 4. This stylized nation consists of two regions and an institutional setting where wages are set at the national level (the national labour market is shown in the right-most panel; the regional labour markets are shown in the two left-most panels). The bargaining stance of the national trade union, which we assume cares about wages and *national* employment (but not its regional distribution), is summarized by the 'collective' labour supply curve S^c. Thus S^c shows the real wage that the Union would demand for any given amount of labour supplied. The true labour supply curve S shows the total amount of labour that workers would, acting individually, supply at any given real wage (S^c lies above S since workers bargaining collectively can achieve a higher wage for any employment level). The bargaining stance of the employers' association is captured by the labour demand schedule D. This shows the maximum real wage that employers can pay for any given amount of labour hired. The bargaining outcome, which sets the national wage rate, is shown in the right-hand panel as the intersection of S^c and D. With the real wage set at w^c, employers choose to hire L workers. Individual workers, however, would supply L^o at w^c. The difference, equal to 'u', is equilibrium unemployment.

Turning to regional unemployment, we take the two regions to be initially identical (for simplicity) so they have identical labour demand (D_1 and D_2) and supply (S_1 and S_2) curves (there are no regional collective labour-supply curves; bargaining is

[2] More technically, the polarization measure is high when the density takes the shape of two groups of regions with small differences in unemployment rates within each group and large differences across groups. It increases as regions within each group become more homogenous in terms of their unemployment rates and/or as the two groups move further apart from each other. In the simplest case of two groups this polarization measure is simply $P = 2D - G$, where D is the mean deviation and G is the Gini coefficient.

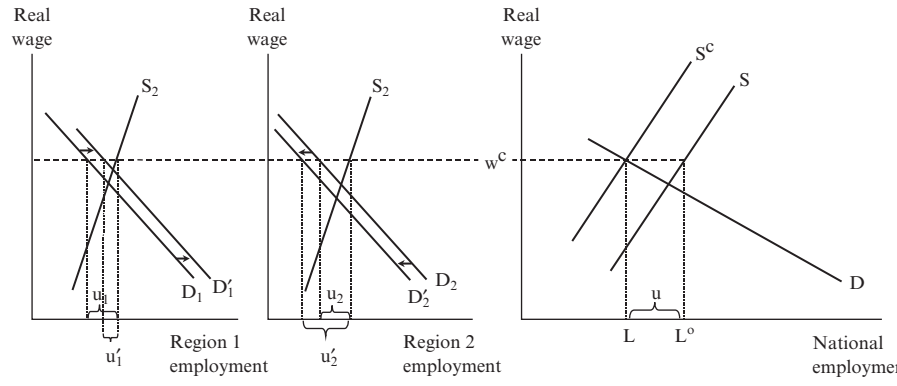

Figure 4. A simple model of regional unemployment

nationwide). With the wage fixed at w^c, the regions' unemployment level are initially equal, namely u_1 and u_2.

3.1. Three sources of unemployment polarization

Polarization in this framework could arise from three distinct but not mutually exclusive sources: (1) a within-nation polarization of labour demand; (2) a within-nation polarization of labour supply; or (3) a between-nation change in labour market institutions.

Demand-driven polarization is represented in Figure 4 by a transfer in labour demand from region 2 to region 1. This shows up in region 1 as a shift from D_1 to D_1' and in region 2 as a shift from D_2 to D_2'. For simplicity, the demand shifts are assumed to be just offsetting, so w^c remains the national wage. Plainly, this labour-demand polarization polarizes regional unemployment, lowering region 1 unemployment to u_1', and raising region 2's to u_2'. Note that with flexible wages, or with collective bargaining at the regional level, this polarization could have been avoided by changing regional wages to reflect regional demand conditions. Thus, polarization of unemployment results from the combination of different demand conditions and an institutional setting that does not take this into consideration. Clearly, a second source of unemployment polarization would be analogous polarization of regional labour supplies, for example an inward shift of S_1 and an outward shift of S_2.

A third possibility is that polarization is nation-driven. At the European-wide level regional unemployment polarization could result from different national institutional reforms. Our diagram, which represents a single country, shows that the national institutions matter for regional unemployment. Plainly, if we presented similar diagrams for several countries and supposed that the gap between S^c and S narrowed in some nations, the relative unemployment rates of European regions taken together would become polarised as unemployment fell in some nations but remained unchanged in others.

In what follows we provide empirical evidence that discriminates between these three stories.

4. WHAT CAUSED REGIONAL UNEMPLOYMENT TO POLARIZE?

To answer this question, we develop a highly flexible tool for identifying which groups of regions have similar unemployment outcomes.[3] This tool allows us to display massive amounts of information in a single diagram, but before it can be useful, we must explain how to read it.

4.1. How to read a stochastic kernel mapping

The tool – called stochastic kernel mapping – is a method that allows us to look at how close each region's unemployment rate is to that of some group of regions that we would expect to behave similarly. In essence, it provides a way of judging the appropriateness of various ways of grouping regions.

To make this concrete, suppose (for the sake of illustration) that the best way to think about regional unemployment was national groupings. In this hypothetical case, unemployment is an almost purely national phenomenon in the sense that nations have different rates, but all regions within a nation have (almost) identical unemployment rates. For this extreme benchmark case, a region's unemployment rate relative to its national average – what we call the region's nation-relative rate – would be close to one regardless of what its 'Europe-relative' rate is. A convenient way of displaying the relationship between a region's Europe-relative and its nation-relative rates is to use a transition matrix akin to Table 1 above. That is, we calculate the probability that a region with a particular Europe-relative rate will have any given nation-relative rate. To reveal as much information as possible, and to avoid the problems of arbitrary ranges, we use a diagram like Figure 3 rather than a table like Table 1.

The left-panel of Figure 5 exhibits what the diagram would look like in the extreme benchmark example we are considering. The point 1.0 on the horizontal axis indicates that a region's rate is exactly equal to the average rate of its nation; 2.0 indicates its rate is twice the national average, and so on. The numbers on the vertical axis indicate a region's rate relative to the Europe-wide average in an analogous manner. As in Figure 3, we can show the probability that a region with a particular Europe-relative rate has any given nation-relative rate with the help of a three-dimensional diagram (illustrated here as a contour map as in the right-panel of Figure 3). In the benchmark case, the stochastic kernel mapping will then have (almost) all mass centred on the vertical line at 1.0. (To read the contour plot, you need to picture a

[3] The non-parametric approach we develop builds on a collection of tools proposed by Quah (1996, 1997) for studying the dynamics of evolving distributions.

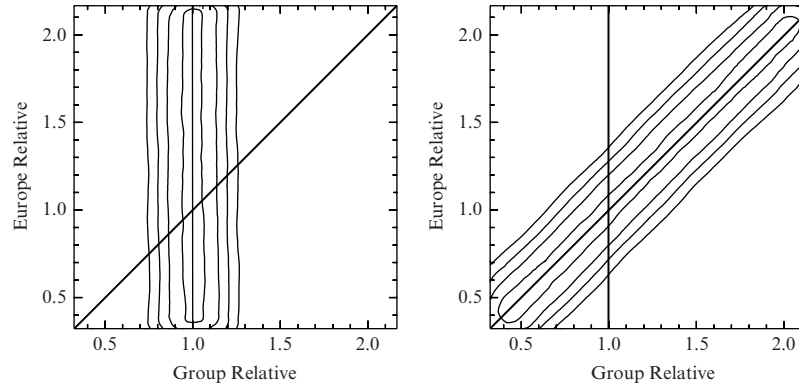

Figure 5. How to read a stochastic kernel: two benchmark examples

'mountain range' running 'north–south' with the crest of the range around 1.0 – the figure plots the contours of this mountain range.)

To further fix ideas, consider the opposite hypothetical extreme where national boundaries have nothing to do with regional unemployment, i.e. the distribution of regional unemployment within each nation is quite similar to the Europe-wide distribution. In this extreme, knowing that a region has a high Europe-relative rate tells us that it is also likely to have a high nation-relative rate. The contour plot in the right panel of Figure 5 illustrates this. Notice that the kernel looks like a mountain range running up the diagonal. Contrasting our two extreme benchmarks we note that when a region's Europe-relative rates can be usefully grouped according to the scheme proposed on the horizontal axis, the kernel will look like the left panel. When the grouping scheme on the horizontal axis is useless, the kernel will look like the right panel.

We turn next to using this tool to illustrate whether unemployment is a regional or a national phenomenon.

4.2. Is unemployment a national or regional phenomenon?

It is often asserted that European unemployment results primarily from national labour market institutions, so the first story to test is that polarization stems from changes at the national level. To this end, we calculate the stochastic kernel mapping for Europe-relative rates to nation-relative rates, using data for all regions and all eleven years. If regional unemployment in Europe is primarily a national matter, then the plots of data are going to look something like the left panel of Figure 5. If it has nothing to do with national institutions, it will look something like the right panel. Plainly, there would be easier ways to proceed if we were only interested in these two extremes. The great merit of the stochastic kernel approach is that it is flexible enough to show a variety of relationships. To cite one of an infinity of examples, it might be that regions with extremely high unemployment tend to correspond to the national benchmark while most regions corresponded to the other benchmark.

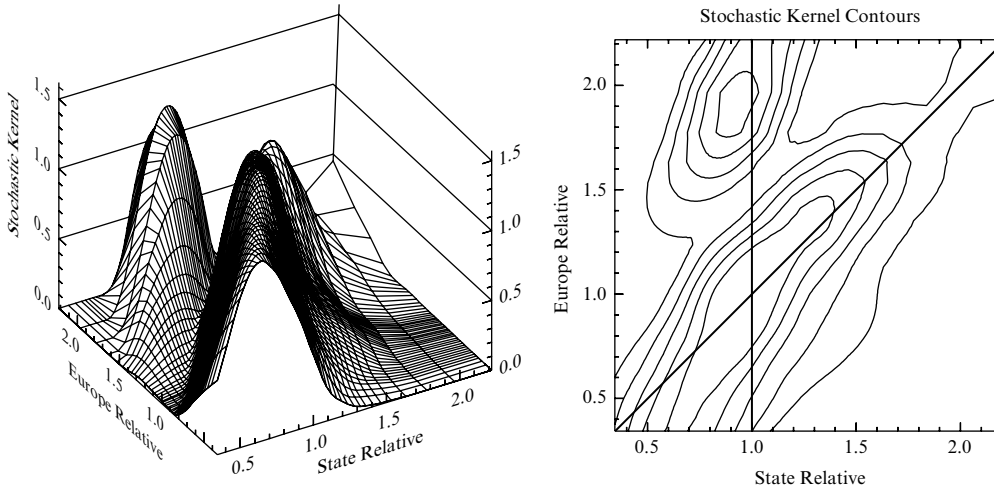

Figure 6. Is unemployment mainly national? Evidence from a stochastic kernel mapping

Technical notes: The kernel is estimated by first deriving the joint distribution of Europe-relative and Group-relative unemployment rates. We then numerically integrate under this joint distribution with respect to Group-relative rates, to get the marginal distribution of Europe-relative rates. Finally, we estimate the marginal distribution of Group-relative rates conditional on Europe-relative rates by dividing the joint distribution by the marginal distribution. Calculations were performed with Danny Quah's tsrf econometric shell (available from http://econ.lse.ac.uk/~dquah/).

Source: Authors' calculations from Eurostat data.

Figure 6, which displays our results, indicates that the only-nations-matter story is not correct. Except for a probability 'peak' in the northwest corner of the diagram, Figure 6 looks very much like the second benchmark case discussed above. In other words, it seems that with some exceptions, the distribution of unemployment rates within a typical nation is not very different to the distribution of rates across Europe and this suggests that variations in national institutions cannot be the main explanation for variations in Europe's regional unemployment rates. More precisely, for unemployment rates below 1.5 times the European average, the kernel is concentrated close to the diagonal, showing that each region's position with respect to the European average is not dissimilar from its position with respect to its national average. That is, a region's EU relative unemployment rate tends to be quite independent of unemployment in its nation state.

In contrast, for the range above 1.5 times the European average, some high Europe-relative unemployment outcomes do correspond to high national outcomes. The spike around the vertical line in this range corresponds to approximately half the Spanish regions with unemployment rates close to the Spanish average, plus Ireland prior to 1994.[4] However, not all of the probability mass for this range of high Europe-

[4] Ireland is classified as a single NUTS2 region, so by construction its regional unemployment rate is always the state average.

relative rates is concentrated around 1.0 on the horizontal axis. There is also a concentration of mass to the right of the vertical line and near the diagonal indicating that some regions with high Europe-relative rates differ as much from their national average as they do from the European average. This is made up of a small group of regions in 1986, formed by Basilicata and Campania in southern Italy, Northern Ireland, and five regions in the north of England and the south of Scotland. Over the next decade the British regions dropped from this group as their unemployment rates came closer to those of their southern neighbours. At the same time, this group expanded to include regions on both sides of the French–Belgian border, all of southern Italy, and the regions on France's Mediterranean coast.

4.3. Similarities in unemployment across geographical neighbours

We have seen that only regions with the very highest unemployment have outcomes similar to other regions in the same nation. We now show that unemployment outcomes are close to those of neighbouring regions.

To this end, we construct a kernel mapping Europe-relative rate to neighbour-relative rates, where neighbour-relative rates are defined as each region's unemployment rate divided by the labour-force-weighted average of the unemployment rates of its neighbours, i.e. regions that are contiguous to it. Importantly, these neighbours include foreign neighbours but exclude the region itself. Note that this definition is driven by pure spatial considerations, so the neighbour groupings are overlapping. As a consequence, some of the neighbours of two neighbouring regions may not themselves be neighbours.

To take a specific example, Catalunya and Comunidad Valenciana in Spain are neighbours and share some neighbours, but not all of Catalunya's neighbours are neighbours of Comunidad Valenciana. Catalunya's neighbours include the French regions of Midi-Pyrennes and Languedoc-Roussillon and the Spanish regions Aragón and Comunidad Valenciana. Comunidad Valenciana's neighbours also include Aragón but not Midi-Pyrennes and Languedoc-Roussillon. Thus, the resulting groups are not mutually exclusive categorizations, with each region falling into only one category. Although this may sound odd at first, it is standard in the consideration of geographical data. After all, any attempt to partition the set of regions into neighbouring groups would require an arbitrary definition.

Figure 7 presents the mapping of Europe-relative rate to neighbour-relative rates. The first thing to notice is that the probability is massed around the vertical line at 1.0. This suggests that 'neighbour effects' on unemployment rates are very strong, since regardless of how high or low a region's rate is relative to the European average, its rate tends to be close to that of its neighbours. Moreover, comparison of Figure 7 with Figure 6 shows that regional unemployment outcomes are much closer to the outcomes of neighbouring regions than to the outcomes of regions in the same nation. This difference is particularly clear when one contrasts Figures 6 and 7 in the

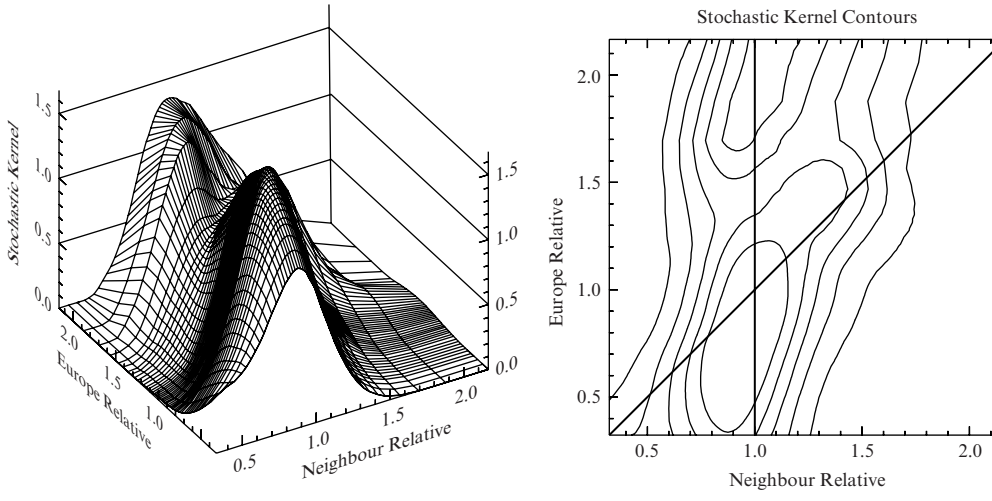

Figure 7. Is unemployment subject to a neighbour effect? Evidence from a stochastic kernel mapping

'twist' of the bottom peak and the 'depth' of the valley between the two peaks in the three-dimensional plot. Alternatively, one can count up the number of lines from the 'bottom' of the contour plot in Figures 6 and 7 (they are plotted at the same heights). Both the lower peak and the valley between the peaks in the neighbour-relative kernel incorporate far more mass than the corresponding areas in the nation-relative kernel. The depth of the valley is particularly relevant, because, as Section 2 showed, polarization was driven by regions with intermediate unemployment ranges experiencing very different evolutions over time.

Also, note that a region's domestic neighbours are part of the groups used to construct both kernels. In Figure 6, however, other regions in the same nation are included. In Figure 7 they are not, but foreign neighbours are. This suggests that foreign neighbours may be more closely related to a region than regions in the same state that are not contiguous – an issue to which we will return below.

To check the visual ranking of the kernels, we calculate transition matrices that correspond to Figures 6 and 7 in the same way that Table 1 corresponds to Figure 3. These transition matrices, presented in Table 2, allow us to more directly gauge the relative mass in different areas of the kernels. To interpret these matrices it is useful to compare them with the same benchmarks we used to interpret the corresponding stochastic kernel, i.e. large numbers on the column for the interval containing 1.0, versus large numbers on the diagonal. We see that the top matrix has all diagonal elements larger than those of the bottom matrix. At the same time, all other elements in the central column are larger in the bottom matrix. This confirms our earlier conclusion that the unemployment outcomes of individual regions are much closer to the outcomes of their neighbours than to the average outcomes of other regions within the same Member State.

Table 2. Is unemployment national or regional? Evidence from transition probability matrices

Transition probabilities table: Europe-relative versus state-relative

	Nation-relative unemployment rates (%)				
	0–0.55	0.55–0.75	0.75–1.15	1.15–1.45	1.45–highest
Europe-relative:					
1.45–highest	0	6	43	13	38
1.15–1.45	0	13	16	52	18
0.75–1.15	0	5	68	18	9
0.55–0.75	6	30	50	12	2
0–0.55	30	36	32	1	0

Transition probabilities table: Europe-relative versus neighbour-relative

	Neighbour-relative unemployment rates (%)				
	0–0.55	0.55–0.75	0.75–1.15	1.15–1.45	1.45–highest
Europe-relative:					
1.45–highest	0	4	45	30	21
1.15–1.45	0	6	40	37	16
0.75–1.15	1	7	63	21	7
0.55–0.75	4	16	72	7	1
0–0.55	22	17	55	6	0

4.4. Is polarization demand or supply driven?

Our finding that 'nearness' matters more than nationality implies that national changes are not the key to understanding the polarization of regional unemployment. The two other stories suggested by our Section 3 theory were that polarization was primarily demand-driven or primarily supply-driven. We next use kernel mappings to investigate which of these stories better fits the facts.

The plot in Figure 8's left panel graphs the stochastic kernel mapping the distribution of 1996 Europe-relative unemployment rates to the distribution of labour force changes between 1986 and 1996 (relative to the average growth in the European labour force over the decade). The vertical line at 1.0 marks regions with labour force growth equal to the European average, namely 6.3%. The concentration of mass at the south-east of the diagram shows that most regions that ended up with relatively low unemployment had relatively high labour force growth. Similarly, regions that ended up with relatively high unemployment generally had below average labour force growth, as can be seen from the concentration of mass at the north-west of the diagram.

What this means is that polarization does not seem to have been supply driven. Indeed, labour force changes have actually worked *against* polarization in the sense that if labour force growth had been more evenly distributed, high unemployment regions would have had, all else equal, even higher unemployment, and low unemployment regions would have had even lower unemployment.

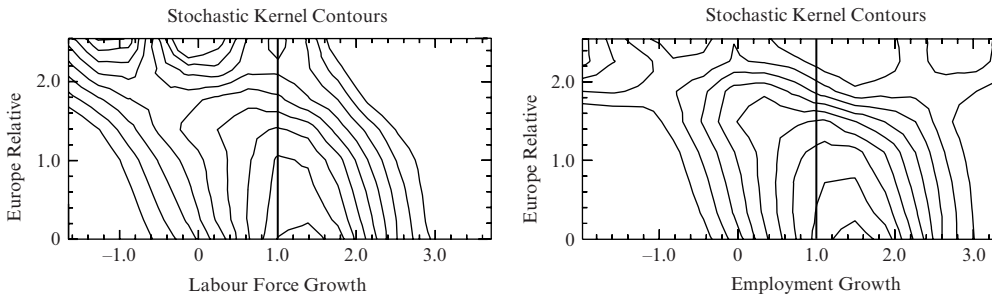

Figure 8. Is unemployment polarization supply driven or demand driven?

To test the demand-driven explanation, we check whether the level of a typical region's unemployment at the end of the period tends to be closely associated with the rate of job creation it experienced from 1986 to 1996. This test of the demand-driven explanation is graphed in the right panel of Figure 8. Specifically, the right panel shows the contours for the relationship between Europe-relative rates in 1996 and the distribution of 1986–96 employment changes (again, relative to the European average). The vertical line at 1.0 marks regions with labour force growth equal to the European average. From the concentration of mass at the south-east of the diagram, we see that most regions that ended up with relatively low unemployment had relatively high employment growth. Similarly, the concentration of mass at the north-west of the diagram shows that regions that ended up with relatively high unemployment generally had below average employment growth. Thus, contrary to labour force changes, employment changes have worked *for* polarization. It is employment changes that have driven high unemployment regions to their high rates and low unemployment regions to their low rates, leading to a polarization process like that represented in Figure 3.

In summary, the sequence of mappings in Figures 6, 7 and 8 showed that the polarization of the European unemployment rates has been the result of some groups of neighbouring regions gaining employment and, to a much lesser extent, labour force, and thus seeing their unemployment rate fall; and other groups of neighbouring regions losing employment (at least in relative terms) and, to a much lesser extent, labour force, and thus seeing their unemployment rate rise.

4.5. What drives the polarization of labour demand?

To understand the polarization of unemployment more deeply, we turn to discriminating between possible sources of demand changes and establishing why these changes might be similar across neighbouring regions.

A first possibility concerns the skill-composition of regional labour forces. Polarization of job creation might arise from the fall in demand for low-skilled workers relative to high-skilled workers in a context in which the supply of skills is unevenly

distributed across regions (see, for instance, Nickell and Bell, 1995; Manacorda and Petrongolo, 1998). If this is the case, regional unemployment outcomes may reflect the underlying skill composition of regional labour forces – regions with a large proportion of low-skilled workers will have experienced similarly high unemployment outcomes, while regions with a small proportion of low-skilled workers will have experienced similarly low unemployment outcomes. If these high- and low-skilled regions are close to one another, then this could explain why neighbouring regions have had similar unemployment outcomes.

A second possibility is that the demand changes driving polarization are the result of changes in the sectoral composition of EU industry. Agriculture, mining and industry employment tends to be regionally concentrated in Europe and employment has shifted continuously from these sectors to services. Without counteracting labour migration, the result could be high unemployment in regions initially specialized in declining sectors. In this case the similarity across neighbours could be a result of regions with declining sectors being contiguous.[5]

A third possibility relies on the predictions of the so-called new economic geography (see, for example, Fujita *et al.*, 1999). These theories predict that economic integration will produce significant changes in the spatial employment pattern, and in the European context, this is likely to take the form of an increasing agglomeration of employment (see Puga 1999). This could take the form of agglomerations of specific activities – i.e. sector-specific industrial clusters – or, if there are significant cross-sector linkages, agglomerations of overall employment, i.e. boom regions and rust belts. While the detailed data required to directly study location and relocation patterns at the regional level is simply not available, studies with country-level data find evidence of both types of changes (see, for example, Midelfart-Knarvik *et al.*, 2000). If this is what is driving polarization, we would expect to see changes in a region's unemployment closely related to those of their neighbours even after we control for nationality and for changes related to their skill and sectoral composition or other similar factors.

4.5.1. Formal statistical analysis. To discriminate between these three possibilities, we use standard statistical analysis to see how the change in the regional unemployment rate between 1986 and 1996 depends upon (1) regional skill composition, (2) the initial sectoral structure of employment, and (3) neighbour effects, i.e. on the change in the unemployment rates of neighbouring regions. More specifically, we measure the skill-composition with two variables – the percentage of adult population with low skills (less than upper secondary education), and the percentage with medium skills (upper secondary education). We capture sectoral influences with the

[5] Stochastic kernel mappings in the Web Appendix that look at the first and second possibilities suggest that neither skill-groups nor sector-groups are important in understanding the pattern of Europe's regional unemployment.

percentage of employment in primary sectors, namely agriculture, mining, forestry and fishing, and the percentage of regional employment in manufacturing. Note that since the sum of employment percentages in primary, secondary and tertiary sectors must equal 100%, we can estimate independent contributions for only two of the three sectors; similar reasoning leads us to exclude data on high-skill workers. To capture the neighbour effect, we include the change in neighbours' unemployment rate. This is defined as the labour-force-weighted average of changes in the unemployment rates of contiguous regions (including foreign neighbours, but not including the region itself).

In addition to these main variables, we include the regions' initial unemployment rates to allow for some conditional convergence to the mean in EU unemployment rates, and we control for national characteristics, such as cross-country differences in labour market institutions, with dummy variables.

4.5.2. Results. The first column of Table 3 shows results when we use the most direct statistical technique (OLS). The most remarkable aspect of these results is that the evolution of the unemployment rate in neighbours has a very strong and significant effect, even after controlling for other regional characteristics. This shows that common nationality, common skills and sectoral composition are not driving the nearness effect. The coefficient on the percentage of adult population with low skills

Table 3. Determinants of changes in regional unemployment, 1986–96

	OLS	IV	OLS	IV
Change in neighbours unemployment	0.541**	0.864*	0.550**	0.833**
	(0.111)	(0.264)	(0.084)	(0.257)
% of low skill workers in regional labour force	0.247**	0.241**	0.220**	0.241**
	(0.106)	(0.101)	(0.093)	(0.098)
% of medium skill workers in regional labour force	−0.080	0.025	0.093	0.137
	(0.158)	(0.184)	(0.126)	(0.139)
Initial % of regional employment in primary sectors	−0.026	−0.037		
	(0.019)	(0.021)		
Initial % of regional employment in manufacturing sectors	−0.278**	−0.230**		
	(0.090)	(0.093)		
Initial % NACE17 sectors	No	No	Yes	Yes
Country dummies	Yes	Yes	Yes	Yes
Initial unemployment	−0.128**	−0.100**	−0.215**	−0.202**
	(0.057)	(0.061)	(0.063)	(0.066)
Adjusted R^2	0.73	0.71	0.84	0.83

Notes: ** and * denote coefficient significantly different from zero with 5% and 10% confidence level, respectively. Heteroscedastic robust standard errors are reported in parentheses. The dependent variable is the difference in logs of unemployment rates. We exclude nations that are classified as a single NUTS2 region (Denmark, Ireland, and Luxembourg) from the regressions, so we have 147 observations in each regression. All explanatory variables are expressed in logarithms. Changes in neighbour unemployment rates are the labour-force weighted averages of changes in the unemployment rates of contiguous regions including foreign neighbours, but not including the region itself. Further details on data definitions and sources are given in the Data Appendix. See Anselin (1988) for further discussion of endogeneity problems with 'spatially lagged' dependent variables. In all specifications we cannot reject the validity of our instrument set at the 5% confidence level using the test proposed by Davidson and MacKinnon (1993).

is positive, large, and significant, as would be expected; controlling for the other factors, a high proportion of the regional population with low skills is associated with an increase, or less of a decrease, in regional unemployment. The coefficient on medium skills, however, is not significantly different from zero. This suggests that it is the lower end of the skill distribution that most markedly affects regional labour market outcomes. The coefficient on primary sectors employment is not significantly different from zero; however, employment in industry seems to matter, as the negative and significant coefficient shows. To interpret this sign, note that many northern and central regions which traditionally specialized in heavy industry had already gone through the worst part of the adjustment by the mid 1980s and have since seen declines in unemployment. Adjustment has taken place later in heavy industrial regions in southern Europe.

Finally, the coefficient on initial unemployment rate suggests that after allowing for workforce characteristics, employment structure, and the evolution of neighbours, we see conditional mean convergence. This suggests that the unconditional polarization of unemployment rates that we documented in Section 2 must be explained by something other than purely the initial unemployment rate.

4.5.3. Statistical refinements. The direct statistical analysis in column 1 may potentially suffer from a confusion of causality. Supposing that a region's unemployment affects its neighbour's unemployment, it is difficult to separate the impact of region A's neighbours on region A's unemployment from the impact of region A's unemployment on that of its neighbours. To solve this direction-of-causality problem, we apply the well-known statistical technique of instrumental variables. Neighbours' initial sectoral employment shares, and the skill composition of their workforces are all possible instruments for the spatially lagged unemployment rates. We would also like to instrument for the movement of firms and workers across regions. Location theories suggest that such movements will be related to some measure of 'market potential' (see Fujita and Krugman, 1995, for theoretical foundations, and Hanson, 1998, for a recent empirical implementation). Thus, we construct an additional instrument based on a simple market potential variable, defined as the inverse distance weighted sum of European regional GDPs.

Results using this more refined statistical technique (IV), which are presented in column 2 of Table 3, show that the confusion-of-causality problem was not severe enough to significantly alter the main column 1 findings. In particular, the effect of neighbours' unemployment remains strong and significant and the proportions of low educated and initial industrial employment remain significant.

4.5.4. More refined sectoral composition. The column 1 and 2 results were based on an admittedly crude characterization of regional employment structure into primary, secondary or tertiary sectors. This was forced on us by the lack of more refined data availability. We can, however, make some assumptions that allow us to

construct approximations of the data we need (see the Web Appendix for details). As the column 3 and 4 results show, allowing for greater sectoral disaggregation does not change any of our main results (column 3 uses the direct statistical technique and column 4 uses the column 2 refinement).

We have also tried a number of alternative specifications, not reported in the table. For instance, in an earlier version of this paper (Overman and Puga, 1999) we also reported results from including two additional regional characteristics – the age structure of the region and female participation rates. These meant to control for the impact of high and rising European youth unemployment rates and regional variations in female participation rates. In addition we have tried including the average change in unemployment for regions with a similar initial sectoral specialization, a similar skill composition of adult population, and so on. None of these refinements changes our main results.

4.6. Are national or international neighbour effects stronger?

We have seen that the neighbour effect is strong and significant even after controlling for the fact that neighbouring regions often share similar characteristics. We turn next to studying the extent to which neighbour effects extend across national borders. To do this, we split the neighbour variable for border regions into two components, one due to domestic neighbours and one due to foreign neighbours. For the domestic and foreign neighbours variables, the labour force weights are those used when constructing our original neighbourhood variable. This ensures that the sum of the two variables is the original neighbourhood variable, and that the coefficients are directly comparable. There are 51 border regions, representing around one-third of the sample. If we drop the UK's 35 regions, which include only one border region, then border regions make up nearly half the sample (the results do not change for this restricted sample).

The results from these regressions are reported in column 1 of Table 4. These show that both domestic and foreign neighbour effects are strong and significant. To adjust for the two-way causality, we again use the IV technique and the results, shown in column 6 reveal that both neighbour effects remain strong and significant. Although the domestic neighbours have a higher coefficient, it is noticeable that, once we correct for two-way causality in column 2, we are unable to reject the hypothesis that the coefficients are identical. There is also the possibility that our results are affected by differences of perhaps a couple of years between some national business cycles and the European aggregate cycle. While the low frequency of our data does not allow for any sophisticated smoothing, we can nevertheless address this possibility by repeating our regressions with a three-year moving average. Columns 3 and 4 show that this only strengthens the results, presumably because it removes business-cycle-related noise. Again, all main results are identical for the OLS and IV techniques.

Table 4. Comparing the domestic and international neighbour-effects

			3-yr moving average	
	OLS	IV	OLS	IV
Domestic-neighbour effect	0.630**	0.830**	0.759**	1.000**
	(0.089)	(0.181)	(0.094)	(0.199)
Foreign-neighbour effect	0.266**	0.527*	0.399**	0.777**
	(0.134)	(0.285)	(0.139)	(0.271)
% of low skill workers in regional labour force	0.196**	0.216**	0.155**	0.182**
	(0.070)	(0.075)	(0.062)	(0.072)
% of medium skill workers in regional labour force	0.082	0.118	0.087	0.137
	(0.121)	(0.129)	(0.100)	(0.109)
Initial % NACE17 sectors dummies	Yes	Yes	Yes	Yes
Country dummies	Yes	Yes	Yes	Yes
Initial unemployment	−0.231**	−0.219**	−0.188**	−0.160**
	(0.063)	(0.068)	(0.043)	(0.046)
Adjusted R^2	0.85	0.84	0.82	0.81

Notes: ** and * denote coefficient significantly different from zero with 5% and 10% confidence level, respectively. Heteroscedastic robust standard errors are reported in parentheses. We time-average instruments where appropriate.

4.6.1. Are 'functional' labour markets the explanation? The neighbour effect is not driven by labour markets extending across neighbouring regions (so-called functional labour markets) with different characteristics. First, from existing work, we know that functional labour markets tend not to extend across NUTS2 regions (see Cheshire and Carbonaro, 1996, for further discussion). Secondly, neighbourhood effects are equally strong across national borders, and cross-border commuting flows are tiny – in 1990 they represented only 0.2% of the total European labour force (de Falleur and Vandeville, 1996). Of these, roughly 50% are commuters to Switzerland (not an EU member). Only approximately 100,000 cross-border commuting flows occur across border regions in our sample. Even on the German–French border, where commuting flows are strongest, they represent less than 0.8% of the combined border region labour force.

4.6.2. Summary. We have documented the following facts about the regional unemployment in the EU:

- There has been a polarization of unemployment rates.
- This polarization has been driven by changes in relative labour demand.
- Such changes have been similar across geographical neighbours.
- There is a truly geographical component to this neighbour effect, since it is only partly explained by national or regional characteristics.
- This geographical component is as strong within as across national borders.

5. IMPLICATIONS FOR POLICY

Since 2000, the EU identifies three objectives of its regional policy. Objective 1 is to promote the development and structural adjustment of regions whose development is

lagging behind; objective 2 is to support the economic and social conversion of areas facing structural difficulties; objective 3 is to support the adoption and modernization of policies and systems of education, training and employment. The main focus of EU regional policy, however, is on regional income differences, with approximately 70% of total EU regional expenditure spent on 'Objective 1' regions (Objective 1 regions should have per capita GDPs below 75% of the EU average). There are no adequate policies in place to tackle a polarization of regional unemployment. To conclude the paper, we offer a few recommendations for policies that could address regional unemployment.

5.1. Reinforcing skills

Our regression results in Section 4.6 show that regions with low-skilled workforces have had significantly worse unemployment outcomes than other regions, and that it is the very low end of the skill distribution that matters the most. These regions also tend to be geographically concentrated. The existing regional policy framework already has a skills component (Objective 3), but as we pointed out above this does not target regions specifically. Our results emphasize the importance of this objective and suggest that it should have a regional dimension and that the effort should be concentrated on regions with the lowest skilled workers – possibly with attempts to co-ordinate these expenditures across neighbouring regions with similar skill compositions. However, this policy cannot on its own tackle the polarization that we document in this paper; first, because education and training dimensions to regional policy have been present throughout the period that we study; and secondly, because polarization is not purely driven by shifts in demand away from low skilled workers.

5.2. Adding a cross-regional and transnational dimension to tackling unemployment

If EU unemployment policies continue to be focused at the national level, as they are in the Luxembourg process, it will be difficult to tackle the regional and cross-border component that our empirical work has uncovered. However, simply adding a regional component, as has recently been proposed, is not enough. Due to the neighbour effects we documented above, the cost of a policy is completely born by some jurisdiction, but the benefits partly accrue to neighbouring jurisdictions. Since local politicians gain no votes or tax revenues from these spillovers, they are likely to underestimate the true benefit of the policy and thus tend to undertake too little of it.

For example, consider two neighbouring regions that suffer from a common lack of skills. If one of these regions were to subsidize training, some of the workers with newly acquired skills might then move to the other region, particularly since mobility tends to increase with the level of training. Moreover, to the extent that the training lowers unemployment within the region, the neighbour effect suggests that the other

region will also benefit. While there is nothing wrong with this *per se*, it does suggest that regional governments will do too little training. The point is that governments will tend to underestimate the full benefit of the training programme since they will care less about the beneficial effects that accrue to the neighbouring region; after all, neighbouring region tax payers are not the ones bearing the burden of the programme. Note that since foreign neighbour effects are as large as domestic neighbour effects, this sort of 'spillover' will arise even in nations where regional policies are financed centrally. It should also be clear that such spillovers and the attendant under-provision of policy apply to all sorts of regional unemployment policies, not just training.

All of this suggests that it is important to add a cross-regional *and* transnational dimension to EU unemployment policies. This is particularly true if unemployment clusters are at least in part the result of the agglomeration of economic activity because then linkages tend to tie together labour supply and demand conditions across nearby areas.

Regarding the practical implementation of such co-ordination, it is worth noting that our nearness results comes from considering overlapping groups of neighbours rather than a partition of the set of regions. This suggests that the best way to incorporate this cross-regional dimension is not to define some new aggregate of existing regions, but rather to co-ordinate some elements of policy across neighbouring regions.

Co-ordination of policies across neighbouring regions and a greater focus on training will go some way towards mitigating polarization of unemployment, but more fundamental changes may need to be considered if the underlying problem is to be truly addressed. It is to two such policies that we now turn.

5.3. Encouraging regional wage setting

The diagrams we introduced in Section 3 highlight the fact that when regions in a country experience different local labour market conditions, such as rising labour demand in some regions and falling demand in others, wage setting at the national level can foster unemployment polarization. Clearly, promoting regional wage setting could do much to alleviate the polarization of European unemployment. While this has so far proven politically difficult, some EU nations have tried to compensate for national wage setting by subsidizing firms in certain areas. This is the case of Italy, where firms in the south have been allowed to pay lower social security contributions than their northern counterparts – although these subsidies are being phased out, following an agreement reached in 1995 between the Italian Government and the European Commission. There are five reasons to believe that this type of wage subsidy is not the best way to address the near equality of labour costs across locations with very different labour market conditions.

First, and most obvious, it is clearly a second-best approach. If the problem is that market forces would normally result in regional wage differentials but rigidities arising from the institutional framework prevent this, then the first-best solution is to try

to reduce or eliminate those rigidities. Secondly, the long life that this type of subsidy tends to have shows that it may result in a strong dependency and be difficult to eliminate. Thirdly, giving subsidies in the form of exemptions of social security contributions creates a labour cost gap whose size may or may not be adequate, and which is difficult to change in response to changes in underlying conditions. But perhaps the most important argument against exemptions or similar mechanisms is that they force labour cost differences at a very aggregate geographical level (say north/south in the case of Italy). However, we have seen that clusters of high and low unemployment arise at a rather fine geographical level. Thus, it is important that labour cost differences too can arise at a much finer geographical level, and that they can change swiftly. This is particularly crucial if, as our results suggest, employment changes are being driven by a clustering of activities. The so-called 'new economic geography' has highlighted the important role that wage differences play in the emergence of such clusters. A dense network of similar firms with important local buyer/supplier relationships and/or a shared labour pool results in lower costs and also fosters innovation. This attracts more similar firms and puts pressure on wages to differ relative to other sectors in the area and relative to nearby areas. This in turn helps attract workers and closes a virtuous circle. Wage rigidities can easily dampen this mechanism.

A fifth aspect worsens this. Exemptions from social security contributions act on the side of employers, not employees. The localized nature of clusters in terms of both employment by activity, and unemployment and employment rates suggests that sharing and exploiting the benefits of agglomeration may rely on reasonable mobility across nearby regions. Artificial spatial wage equalization adds to the rigidities preventing this. We now turn to looking at this mobility argument in more detail.

5.4. Promoting short distance mobility

Referring back to Figure 4 it should also be clear that polarization of unemployment due to shifts in demand can be offset by corresponding shifts in supply. International and inter-regional mobility in Europe, however, has been very low in recent decades (see Braunerhjelm *et al.*, 2000). Moreover, nations seem to be remarkably reluctant to accept the changes in population distribution that such migration would entail. Our results, however, suggest that even facilitating migration at a finer geographical level might go a long way towards decreasing polarization. We found that unemployment is more homogenous across neighbouring regions than it is across regions within the same nation. The average EU member has 13.6 regions, while the average neighbourhood in our data has 5.6 regions. Hence these are clusters of typically less than one half of the size of the average EU nation, but often extend across national borders and include regions from more than one EU member. Given that unemployment clusters are not very large and are scattered across Europe, it may be politically viable as well as more efficient to implement policies that accept some clustering of firms coupled with larger mobility of workers within the immediate neighbourhood.

Discussion

Hylke Vandenbussche

University of Leuven

The Overman–Puga paper provides empirical evidence of polarization of regional unemployment rates in Europe in the period 1986–96, using both non-parametric and parametric techniques. On the basis of their findings the authors suggest that policy-makers should shift their attention away from country-based policies to tackle unemployment and emphasize more on regional unemployment policies that can transcend one individual member state. The value added of the paper clearly lies in the profound empirical research that has gone into it.

One of the main messages of this paper is that average unemployment rates at the country level may hide regional disparities. Therefore, the question is raised whether *countries* are the best grouping criteria when studying and remedying for regional unemployment. In order to answer that question the authors look at alternative grouping criteria. Their findings show that *neighbours* is the best grouping criterion to explain a region's fortune in terms of unemployment rates. In short, when neighbouring regions have low unemployment rates, the chances are high that your region ends up with a relatively low unemployment rate too and visa versa. If neighbours with high unemployment surround your region, this is likely to spill over to your region too. Other grouping criteria such as *nationality* of the neighbour region appear to be relatively unimportant. Also the *sectoral* composition and the proportion of *skills* in a region do not seem to explain a region's fortunes as well as the *neighbours* criterion.

In terms of the methodology used, I feel that too much emphasis lies on the non-parametric approach, which is less well known and understood. Questions like whether deviations from the 1-line of the kernel graphs are statistically significant come to mind. Are there robustness checks for kernel results? Can a different scaling change the two and three-dimensional figures?

The regression analysis is quite short and condensed in comparison to the elaborate kernel type of evidence. To make the regression analysis richer, it seems to me that one could experiment with additional explanatory variables in order to better understand the neighbour effect. For example do neighbour regions share common language, geographical or climate characteristics? Is decision-making in neighbouring regions centralized or decentralized?

While the evidence on polarization of unemployment rates in the Overman–Puga paper is very convincing and well documented, the economic causes of this polarization get much less attention. To dismiss a paper due to a lack of underlying theory to explain the facts, is too easy. In analogy, when the first empirical papers on intra-industry trade came out, opponents pointed to the lack of theoretical underpinnings for this type of trade and dismissed it as a 'statistical regularity'. However, theory soon caught up with the facts.

The same may be true here; it is possible that the Overman–Puga paper has pointed out an important empirical regularity that will be more fully explained by theory later on. Or, alternatively it may just be that when including in the regression a dummy controlling for 'meters above sea level', one gets a very good fit of what explains changes in regional unemployment rates. From Figure 1, it really looks like in particular some of the sunny beach areas in Europe are having the highest unemployment rates.

Panel discussion

Andy Rose suggested that while the graphs and transition matrices are interesting, some elements of confidence intervals or indicators for the degree of precision should be added to the paper. He further pointed out that the paper should make the distinction between temporary and permanent shocks in the context of the optimal currency area literature, which is a natural policy context for the question of regional diversion.

Joachim Winter pointed out that the size of the regions is very uneven across countries and asked how this affected the empirical results. Patrick Honohan asked whether the bivariate plots could mask the true underlying multivariate relationship. For example, in the stochastic kernel figure that looks at the Europe-relative unemployment versus Europe-relative skill levels variable (see the Web Appendix), it is quite diagonal and it looks as if there is no significant effect, yet in the regression results there was a strong significant skills variable. He asked whether it is true that a regression, by partialling out other factors, allows us to identify a skill effect that was not reflected in the bivariate plots. Furthermore, Patrick Honohan suggested that the paper downplays the country effects. For instance, comparing Spain and Portugal leads one to believe that there are indeed significant country effects, and not only effects of the neighbouring region. There seems to be too sharp a boundary in Portugal for this effect to be negligible. Overall, it seems that both country effects and neighbour effects are present and only a multivariate analysis can uncover the relationship, rather than a bivariate plot.

Michele Boldrin was concerned with why the regional analysis is relevant for policy issues. Nearby regions seem to matter more than the country average, but most of the neighbouring regions are also from different countries. He suggested that the study should be limited to those regions that are neighbours, but do not belong to different countries, in order to isolate the neighbour effect. Furthermore, he pointed out that most of the observations come from Italy and Spain, where the regions are the most polarized with respect to the country average. In both countries the southern regions are highly subsidized from central government and this is an important issue. Without these subsidies the regional differences would even be substantially higher. Lars Feld added that the subsidies of the central government may to a certain

extent prevent more useful local government policies that may be better informed about the local industry structure and local policy issues. On the econometric approach Paul Seabright asked the authors to outline more clearly what the advantages of a non-parametric test in the context of studying unemployment really are.

In the reply Henry Overman acknowledged that it is not possible to give confidence intervals in the type of spatial econometrics as used in the paper, although he pointed out the trade-off with a higher degree of information in the point estimates is worthwhile in the context of this paper. Regarding the issue of bivariate versus multivariate techniques, he pointed out that the authors did try to control for variables such as age structure of the region or female labour participation, but that it was very difficult to find a good regional database for this. Regarding the border regions he pointed out that about one out of three regions in the EU have a foreign neighbour. If the UK is excluded, this figure goes up to nearly one out of two. However, the results are robust to excluding the UK as well as to limiting the sample to those that do not have any foreign neighbours.

REFERENCES

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwer Academic Publishers.

Blanchard, O.J. and L.F. Katz (1992). 'Regional evolutions', *Brooking Papers on Economic Activity*.

Braunerhjelm, P., R. Faini, V. Norman, F. Ruane and P. Seabright (2000). *Integration and the Regions of Europe: How the Right Policies Can Prevent Polarization*, London: Centre for Economic Policy Research.

Burda, M. and C. Wyplosz (2001). *Macroeconomics. A European Text*, Oxford: Oxford University Press, 3rd edn.

Cheshire, P. and G. Carbonaro (1996). 'Urban economic growth in Europe: Testing theory and policy prescriptions', *Urban Studies*.

Davidson, R. and J.G. MacKinnon (1993). *Estimation and Inference in Econometrics*, New York: Oxford University Press.

Decressin, J. and A. Fatàs (1995). 'Regional labor market dynamics in Europe', *European Economic Review*.

de Falleur, M. and V. Vandeville (1996). 'Cross-border flows of workers in Europe: Facts and determinants', in H. Armstrong and R.W. Vickerman (eds.), *Convergence and Divergence Among European Regions*, London: Pion.

Esteban, J.M., C. Gradín and D. Ray (1999). 'Extensions of a measure of polarization, with an application to the income distribution of five OECD countries', Processed, Institut d'Anàlisi Econòmica-CSIC, Barcelona.

Fujita, M. and P.R. Krugman (1995). 'When is the economy monocentric? Von Thünen and Chamberlin unified', *Regional Science and Urban Economics*.

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

Hanson, G.H. (1998). 'Market potential, increasing returns, and geographic concentration', Working Paper 6429, National Bureau of Economic Research.

Manacorda, M. and B. Petrongolo (1999). 'Skill mismatch and unemployment in OECD countries', *Economica*

Midelfart-Knarvik, K., H.G. Overman, S.J. Redding and A.J. Venables (2000). 'The location of European industry', Economic Papers 142, European Commission Directorate-General for Economic and Financial Affairs.

Nickell, S. and B. Bell (1995). 'The collapse in demand for the unskilled and unemployment across the OECD', *Oxford Review of Economic Policy*.

Overman, H.G. and D. Puga (1999). 'Unemployment clusters across European regions and countries', Discussion Paper 2250, Centre for Economic Policy Research.
Puga, D. (1999). 'The rise and fall of regional inequalities', *European Economic Review*.
Quah, D.T. (1996). 'Regional convergence clusters across Europe', *European Economic Review*.
—— (1997). 'Empirics for growth and distribution: Stratification, polarisation, and convergence clubs', *Journal of Economic Growth*.
Rodríguez-Pose, A. (1999). 'Convergence or divergence? Types of regional responses to socio-economic change', *Tijdschrift voor Economische en Sociale Geografie*.
Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*, New York: Chapman and Hall.