Unemployment clusters across European regions and countries

Henry G. Overman[*‡]
London School of Economics

Diego Puga[*§]
University of Toronto

Centre for Economic Policy Research Discussion Paper 2250
October 1999. Revised September 2001

Abstract: European regions experienced a polarisation of their unemployment rates between 1986 and 1996. Regions with high or low initial unemployment saw little change, while regions with intermediate unemployment moved towards extreme values. During this process, regions had similar outcomes to their neighbours. This nearness effect is driven by the fact that neighbouring regions faced similar changes in labour demand. We show that these similarities are only partly explained by regions being part of the same Member State, having similar skill composition, or sectoral specialisation. Put together, our results suggest that EU policy makers need to add a regional and transnational dimension to European unemployment policy. We provide a number of suggestions as to how this might be done.

Key words: European regional unemployment, polarisation.
JEL classification: R12, E24, F15.

[*]We thank Carlos Gradín for help in calculating the polarisation 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.

[‡]Department of Geography and Environment, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom (e-mail: h.g.overman@lse.ac.uk; website: http://cep.lse.ac.uk/~overman). Also affiliated with the Centre for Economic Performance at the London School of Economics and the Centre for Economic Policy Research. Funding from the United Kingdom Economic and Social Research Council is gratefully acknowledged.

[§]Department of Economics, University of Toronto, 150 St. George Street, Toronto, Ontario M5S 3G7, Canada (e-mail: d.puga@utoronto.ca; website: http://dpuga.economics.utoronto.ca). Also affiliated with the Canadian Institute for Advanced Research and the Centre for Economic Policy Research. Funding from 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.

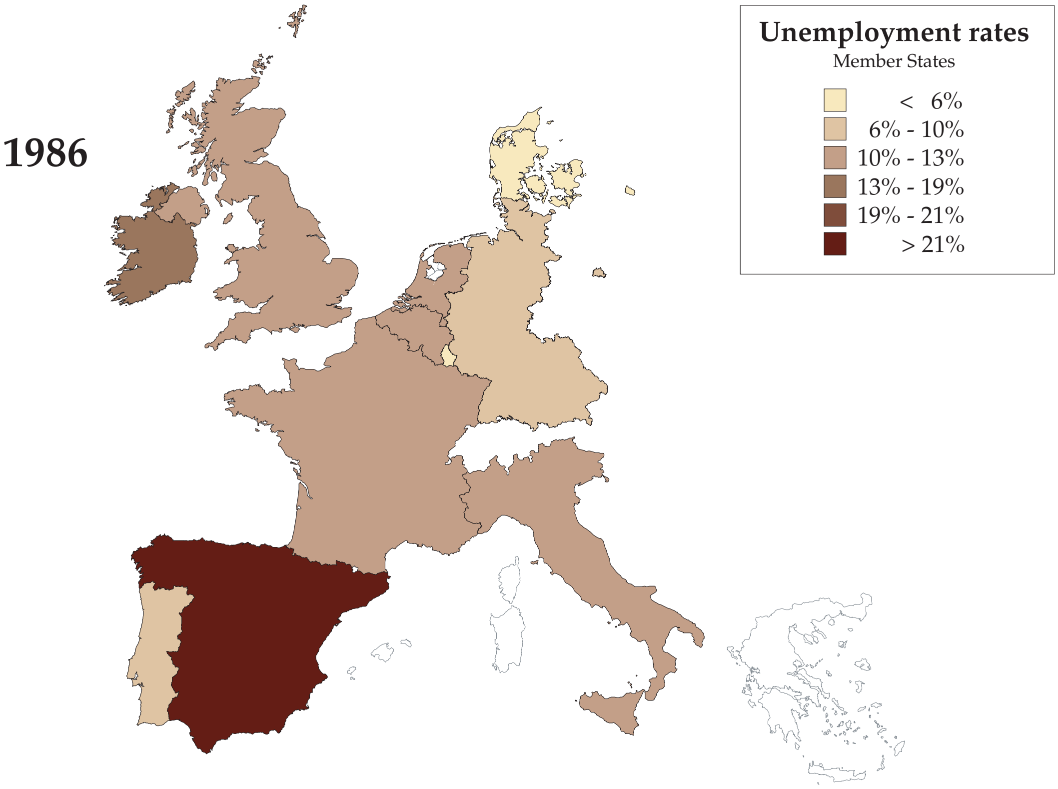

Figure 1. National unemployment rates in Europe

Most European policymakers believe that reducing unemployment is the key challenge currently facing Europe. The Luxembourg process, launched at the 1997 European Job Summit, emphasises the reform of national labour market policies as the key to meeting this challenge. Few people would doubt the importance of such labour market reforms in tackling unemployment. However, looking at unemployment as a purely national phenomena risks ignoring the fact that unemployment has an important regional component. In this paper, we show that there is a regional and transnational dimension to unemployment that is becoming increasingly important. As a result, high unemployment is unlikely to be fully addressed by the kind of national reforms proposed by the Luxembourg process.

The emphasis of the Luxembourg process reflects the fact that when we think about differences in unemployment rates across Europe, we normally think of differences across countries as represented in Figure 1. However, national averages hide large differences in unemployment rates across regions *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 United Kingdom, in 1996, Merseyside had an unemployment rate 3.2 times that of the Surrey-Sussex region; 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.

This paper starts by analysing how these regional differences have evolved over time. Up until the mid 1980s, differences in unemployment rates across European regions were very stable, with

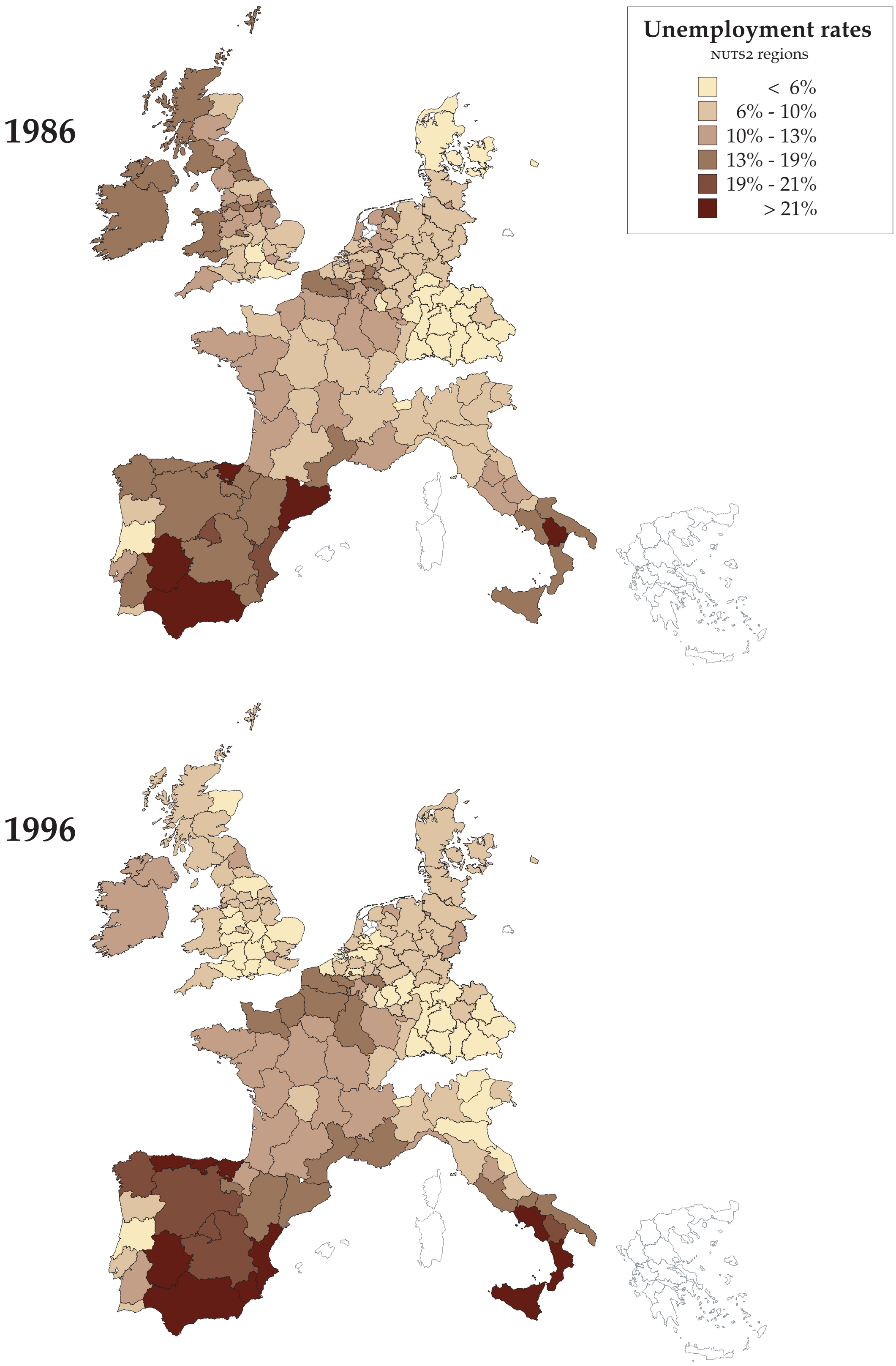

Figure 2. Regional unemployment rates in Europe

regional labour forces adjusting just enough to offset ongoing changes in regional employment (see Chapter 6 in Layard, Nickell, and Jackman, 1991). Figure 2 depicts regional unemployment rates for the contiguous European Community in 1986 and a decade later. The maps suggest that the stability described by Layard *et al.* (1991) up to the mid 1980s no longer holds. The average unemployment rate for regions in these maps was the same, 10.7%, in 1996 as in 1986, and the decade separating them could be thought of as covering a full cycle in unemployment rates.[1] Yet the map for 1996 looks different enough from that for 1986, that one starts to wonder what has happened to the distribution of European regional unemployment rates over this period.

In Section 2 we show that, during the decade from 1986 to 1996, there has been a polarisation of unemployment rates across the regions of the European Union (EU). Regions that in 1986 had a low unemployment rate relative to the EU average still tended to have a relatively low unemployment rate in 1996. Similarly, regions that in 1986 had a relatively high unemployment rate still tended to have a relatively high unemployment rate in 1996. However, regions with intermediate initial unemployment rates had mixed fortunes. Some saw a marked fall in their relative unemployment rate, while others saw it rise, and still others saw it roughly unchanged.

EU policy makers see tackling regional inequalities as a key policy objective. Reflecting this, a third of the EU budget is devoted to tackling regional inequalities, with the focus so far placed mostly on income inequalities — about two thirds of the expenditure is allocated on the basis of low regional income. Regarding unemployment, the emphasis has so far been on national labour market institutions, and how these affect countries average unemployment levels and differences in regional unemployment rates *within* each country.

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 (in Section 4) we develop a non-parametric methodology that shows 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 Member State. 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 polarisation.

This finding — that 'nearness' matters — is consistent with polarisation being driven by changes in demand or supply that are similar across neighbouring regions. Our second step (in Section 5) is to show that polarisation has been demand rather than supply driven. In fact changes in supply have helped mitigate the process of polarisation.

Our third step (in Section 6) 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

[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.

specialised 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 are geographically concentrated this would help explain our findings on a strong geographical component to unemployment. Our regression results 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, nearness matters as much across as within national borders.

To summarise, our results suggest that EU unemployment is a phenomenon that has an important regional *and* transnational component. Over the last decade European regions' unemployment outcomes have closely followed those of neighbouring regions. This is only partly explained by regions being part of the same Member State, having a similar labour force composition, or sectoral specialisation. Remarkably, neighbouring regions across national borders are as important as domestic neighbours in understanding unemployment outcomes. The polarisation of unemployment that has occurred over the last decade has been characterised by the emergence of clusters of high and low unemployment, which show little respect for national borders.

What does all this tell us about the policy measures needed to deal with the polarisation 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 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 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 polarisation 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 8 once we have outlined our empirical evidence and analytical framework.

1. Data

We study *Europe relative* unemployment rates from 1986 to 1996. The time period covered reflects the fact that only a much more limited regional coverage is available before 1986. In addition, serious comparability problems arise when using data for earlier years. The Europe relative unemployment rate is defined as the ratio of the regional unemployment rate to the European-wide average unemployment rate. Working with relative, as opposed to absolute unemployment rates, helps remove co-movements due to the European-wide business cycle and trends in the average unemployment rate. As mentioned in the Introduction, the average European unemployment rate was the same in 1996 as in 1986, 10.7%, and the decade in between can be regarded as covering a full cycle.

The unemployment rate series are computed from the harmonised unemployment rates and labour force data contained in the Regio database produced by Eurostat (Eurostat, 1998). These data are based on the results of the Community Labour Force Survey, carried out in Spring each year.

The analysis focus on the contiguous European Community of 1986. That is, those regions of the EU that satisfy the following three criteria:

1. Have been part of the EU (European Economic Community before 1 November 1993) from 1986 to 1996.

2. Are in a Member State which has a land border with at least one other Member State containing at least one region satisfying (1).

3. Have a land border with at least one other region satisfying (1) and (2).

The definition of regions corresponds to level two of the Nomenclature of Territorial Units for Statistics (NUTS2), a hierarchical classification with three regional levels established by Eurostat to provide comparable regional breakdowns of EU Member States. These NUTS2 regions form the basis for the classification of objective regions with respect to objective 1 of EU policy. There are 150 NUTS2 regions satisfying criteria (1) to (3). The average NUTS2 region in our data set had a land area of 13,800 square kilometres and a population of 2.1 million in 1996 (that is slightly larger than the US State of Connecticut and with two thirds of its population).

The Data Appendix gives full details of the regional coverage and data sources.

2. The polarisation of European regional unemployment rates

Increasing inequalities

What has happened to the distribution of regional unemployment rates over the decade beginning in 1986? In Figure 3, we plot a sequence of kernel estimates of the density of Europe relative unemployment rates for four years: 1986, 1989, 1993, and 1996.[2] By definition, 1 on the horizontal axis indicates the European average unemployment rate, 2 indicates twice the average, and so on. Two features are particularly noticeable in Figure 3. First, as we move through the decade, the

[2]The density plots can be interpreted as the continuous equivalent of a histogram, in which the number of intervals has been let tend to infinity and then to the continuum. All densities are calculated nonparametrically 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.

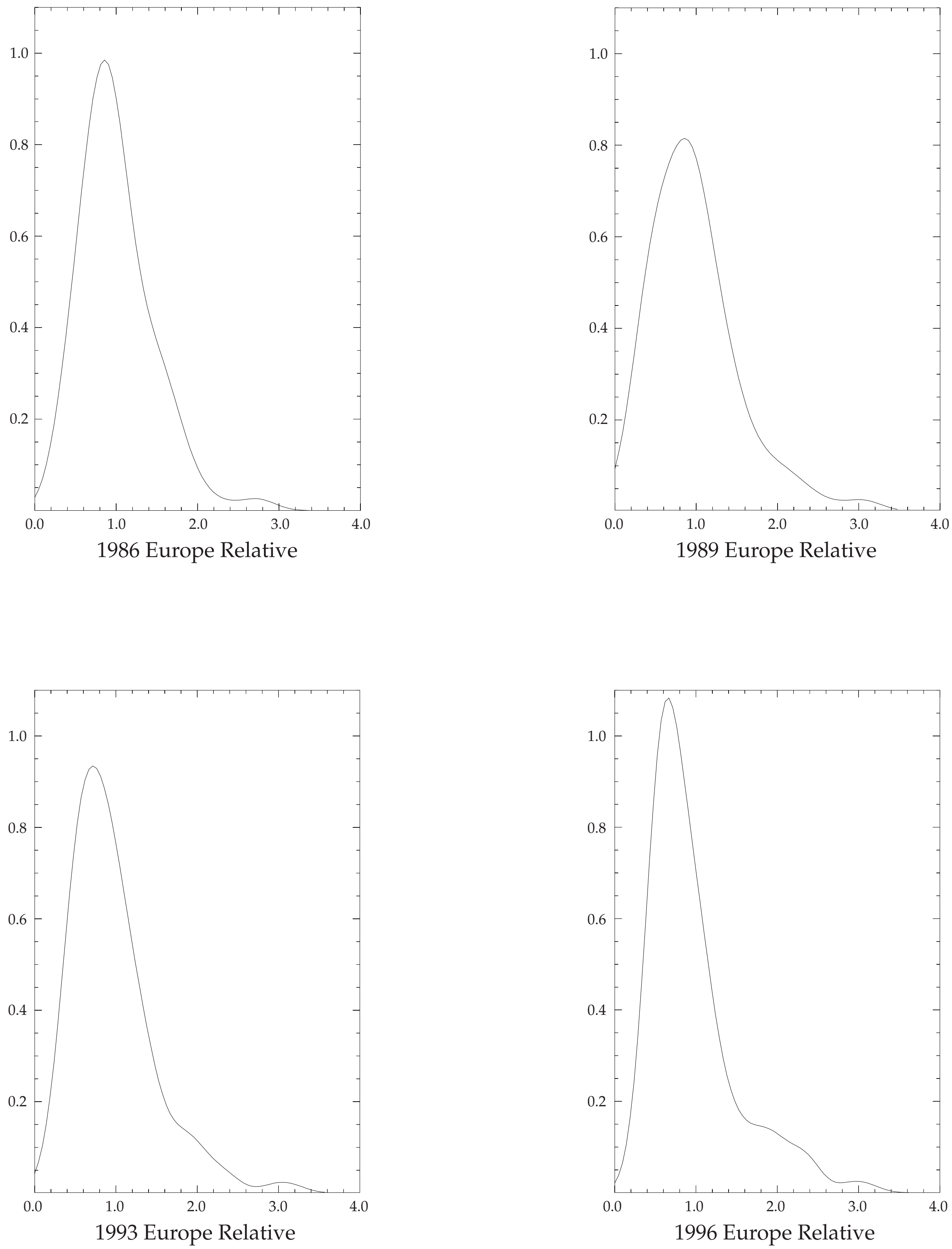

Figure 3. Densities of Europe relative unemployment rates

1986 Europe Relative: n	1986 Europe Relative: class	1996 Europe Relative: [0–0.6)	[0.6–0.75)	[0.75–1)	[1–1.3)	[1.3–∞)
32	[1.3–∞)	0.00	0.00	0.16	0.22	0.62
32	[1–1.3)	0.06	0.22	0.34	0.19	0.19
42	[0.75–1)	0.24	0.29	0.26	0.21	0.00
23	[0.6–0.75)	0.52	0.26	0.09	0.09	0.04
21	[0–0.6)	0.81	0.19	0.00	0.00	0.00

Table 1. 1986 to 1996 Europe relative transition probability matrix

distribution of unemployment rates for a majority of regions becomes more concentrated below the European average: the peak of the distribution, close to the average in 1986, moves slightly leftwards and the mass becomes more narrowly concentrated around that peak. Second, there is a growing group of regions with unemployment rates above twice the European average: these regions produce the 'bulge' in the upper tail of the distribution — to see this clearly, contrast the mass above twice the European average unemployment rate in 1986 and 1996. Looking through the four snapshots we see that these two features have slowly evolved over the decade. Therefore, 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.

Polarisation

The density plots are suggestive of a gradual polarisation of European regional unemployment rates. However, this interpretation cannot be supported by the density plots alone. The collection of densities tell us nothing about the identity of regions in the distribution of regional unemployment rates. Is it true that a group of low unemployment regions and a group of high unemployment regions has slowly emerged, while regions with intermediate unemployment rates have moved closer to the tails of the distribution? Certainly, more regions had low or high unemployment rates in 1996 than in 1986, but what was their relative position in the earlier year? Does this collection of snapshots actually just show churning of the unemployment rate distribution, the random ups and downs of regional fortunes, or are they the result of a more structured process?

The natural way to answer these questions is to track the evolution of each region's relative unemployment rate over time. An easy way to do this is to construct a transition probability matrix. Table 1 reports the transition probability matrix between the 1986 and 1996 distributions of Europe relative unemployment rates. Each row of this matrix takes a given range of 1986 unemployment rates and shows how regions initially in that range have ended up distributed across different ranges of unemployment rates in 1996.[3]

[3]The table gives two additional pieces of information. The first column gives n, the number of regions that begin their transitions in a given state. The second column gives the classes that divide up the state space.

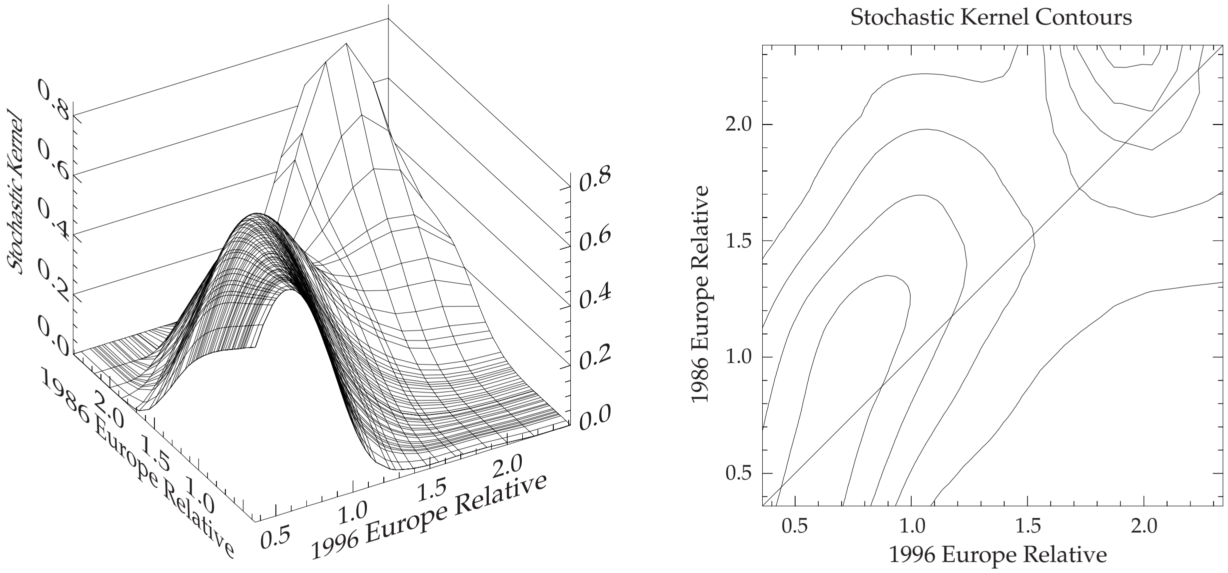

Figure 4. 1986 to 1996 Europe relative Stochastic Kernel

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, 19% had an unemployment rate between 0.6 and 0.75 times the average, and none had a relative unemployment rate higher than that. The next row up tells us that of those regions with an initial unemployment rate between 0.6 and 0.75 times the European average, 26% remained in that range, while 52% saw their unemployment rate fall below 0.6 times the average. Jumping to the top row we also see strong persistence amongst the regions with highest unemployment rates: of the regions with an initial unemployment rate above 1.3 times the European average, 61% remained above 1.3 times the average in 1996, while 23% moved to between the average and 1.3 times the average. However, regions with unemployment rates between 0.75 and 1.3 times the European average (third and fourth rows from the bottom) had experienced much greater mobility — regions with initial unemployment rates between 0.75 times the average and the average ended up almost equally distributed across the four intervals between 0 and 1.3 times the average.

Europe relative unemployment rates are, by nature, a continuous variable. There is a degree of arbitrariness involved in choosing a specific discretisation, and changing from one discretisation to another can easily distort the 'true' picture of transitions. In addition, many interesting details are lost as a result of the discretisation.

Figure 4 resolves these problems by avoiding any discretisation, and plotting the transition kernel from the 1986 distribution of Europe relative unemployment rates to the 1996 distribution of Europe relative unemployment rates. This is, in effect, the continuous equivalent of the transition probability matrix of Table 1. Cutting across the kernel for any given value of 1986 unemployment rates gives us the probability distribution of 1996 unemployment rates for regions that start with that value in 1986. Thus, a 'peak' in the kernel corresponds to a high number in the transition matrix while a 'valley' is the equivalent of a low number.[4]

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 the more familiar contours on a standard geographical map. Lines on the contour plot connect points at the same height on the three-dimensional kernel. An additional 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.

Figure 4 confirms that there has been a polarisation of regional unemployment rates between 1986 and 1996. 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. Similarly, 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 had mixed fortunes: some saw their relative unemployment rate fall, while others saw it rise. Still others saw it roughly unchanged.

Measuring Polarisation

We could calculate and compare a host of summary statistics of the distribution of regional unemployment rates across time. For instance, the Gini coefficient rose from 0.236 in 1986 to 0.281 in 1996. However, the results from the transition kernel in the previous subsection, suggest that the most significant change between 1986 and 1996 has been, not so much an increase in inequality, but rather the polarisation of regions into two groups — one with low unemployment and one with high unemployment.

To quantify this polarisation, we use the generalisation by Esteban, Gradín, and Ray (1999) of the polarisation measure of Esteban and Ray (1994). In the simplest case, for two groups ('high' and 'low' unemployment regions in our case) this polarisation measure is simply

$$P = 2D - G ,$$

where D is the mean deviation and G is the Gini coefficient. This polarisation 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. Between 1986 and 1996 polarisation thus measured increased by 37%, from 0.096 to 0.131.

3. A simple framework for understanding regional unemployment polarisation

To help us think about the polarisation of unemployment rates, let us develop a simple framework that shows how such polarisation might arise (this is a two-region version of that in Chapter 6 of

[4]The three dimensional stochastic kernel plots are drawn so that the density of lines reflects the underlying number of observations on which that part of the kernel is estimated. This procedure makes the pictures easier to read and more informative, but does not change the shape of the kernel.

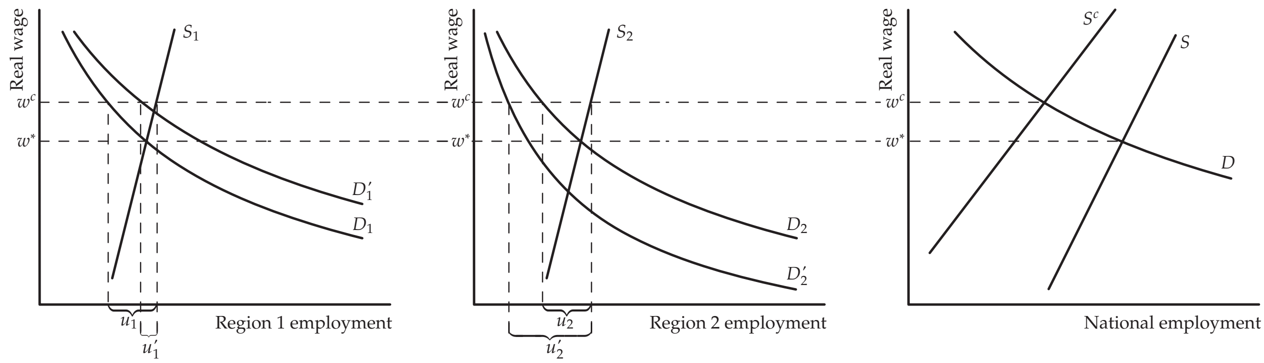

Figure 5. Equilibrium unemployment and its polarisation

Burda and Wyplosz, 2001). In the labour market, equilibrium employment and the wage level are given by the intersection of labour demand and supply. Involuntary unemployment occurs when wages do not adjust to clear the market so that not all labour supplied by individuals is hired by firms. There are a number of reasons why wages might not be fully flexible (see Chapter 6 in Burda and Wyplosz, 2001, for details).

For concreteness, suppose that wages are determined by a national bargaining process between a union representing workers and an association of employers. To keep matters simple, suppose that the union holds all bargaining power. The nation consists of two regions and the institutional setting is such that they are constrained to have a common wage level. We assume that the trade union has joint preferences over both wages and *aggregate* employment, but it does not care about the distribution of employment across regions. The two diagrams on the left of Figure 5 show the labour supply (S_1 and S_2) and demand (D_1 and D_2) in two initially identical regions 1 and 2. Aggregating these gives the national labour supply (S) and demand (D) depicted on the right of Figure 5. When workers bargain collectively over wages through the national union, their collective labour supply curve(S^c) lies above the aggregate supply curve because they feel that together they have more power to achieve a higher wage for any given level of employment. The equilibrium wage is w^c, determined by the intersection between S^c and D. This results in involuntary regional unemployment, $u_1 = u_2$, because individuals would be willing to provide labour at the lower wage w^*, given by the intersection of the two curves S and D. However, they cannot do so because the wage is determined by the national wage bargaining process and unemployed workers cannot underbid their employed colleagues.

One possible reason why we might see polarisation in this context is a change in demand conditions that affects the two regions differently. This is represented in Figure 5 by having demand for labour by firms in region 1 increase from D_1 to D_1' and demand for labour in region 2 decrease from D_2 to D_2'. For simplicity, assume that the changes offset one another when aggregated into national labour demand. As a result, the national wage resulting from collective bargaining remains unchanged at w^c. This results in a low level of unemployment, u_1' in region 1 and a high level of unemployment, u_2' in region 2. With flexible wages, or with collective bargaining at the regional level, the wage level would be lower in region 2 than in region 1, reflecting their different demand conditions. Thus in this context, polarisation of unemployment results from the combination of different demand conditions and an institutional setting that does not take this into consideration. Clearly, a second possible source for this kind of polarisation would be opposite changes in supply conditions in the two regions (a decrease in S_1 and an increase in S_2).

The diagram represents a single country and shows that the institutional setting matters for regional unemployment. A third possibility is that polarisation at the European wide level could result from some countries changing their institutional setting to one that results in lower unemployment and others not doing so. In what follows we provide empirical evidence that discriminates between these three stories.

4. Is unemployment a national or regional phenomenon?

We now show that while there is a national component to unemployment, this is insufficient to explain the distribution of EU wide unemployment rates. In fact, regions have unemployment outcomes that are closer to neighbouring regions than to other regions in the same Member State. This is consistent with polarisation being driven by changes in demand or supply (as shown in the diagram) that are similar across neighbouring regions. The next step will then be to show that polarisation has been demand rather than supply driven and following that, finally, we will discriminate between possible sources of similar demand changes across neighbouring regions.

We want to examine how much of the dispersion in EU regional unemployment rates is the result of countries having very different national unemployment rates, perhaps as the result of different labour market institutions. We also want to allow for the fact that in some countries regional unemployment rates might be tightly concentrated around the State average while in others they might be more dispersed, so that knowing the country a region belongs to might help us understand its unemployment rate to a different extent for regions with high and low unemployment relative to EU average. To this effect, we develop a nonparametric method that allows us to identify groups of regions that have similar unemployment outcomes. The nonparametric approach we develop builds on a collection of tools proposed by Quah (1996, 1997) for studying the dynamics of evolving distributions.

In our case, the idea is to look at how close each region's unemployment rate is to that of some group of regions which we would expect to behave similarly. To do this we establish a mapping from a region's unemployment rate relative to the European average to the same region's unemployment rate relative to the group average. To make this concrete, think about grouping together regions that belong to the same Member State.

Similarities in unemployment within Member States

Consider the extreme case, where all regions within each State have (almost) identical unemployment rates. In that case, any differences in regional unemployment rates correspond to regions being in States with different national unemployment rates. In this extreme benchmark case, regardless of a region's *Europe relative* unemployment rate, its unemployment relative to the average for other regions in the same Member State (*State relative*) will be close to one. The stochastic

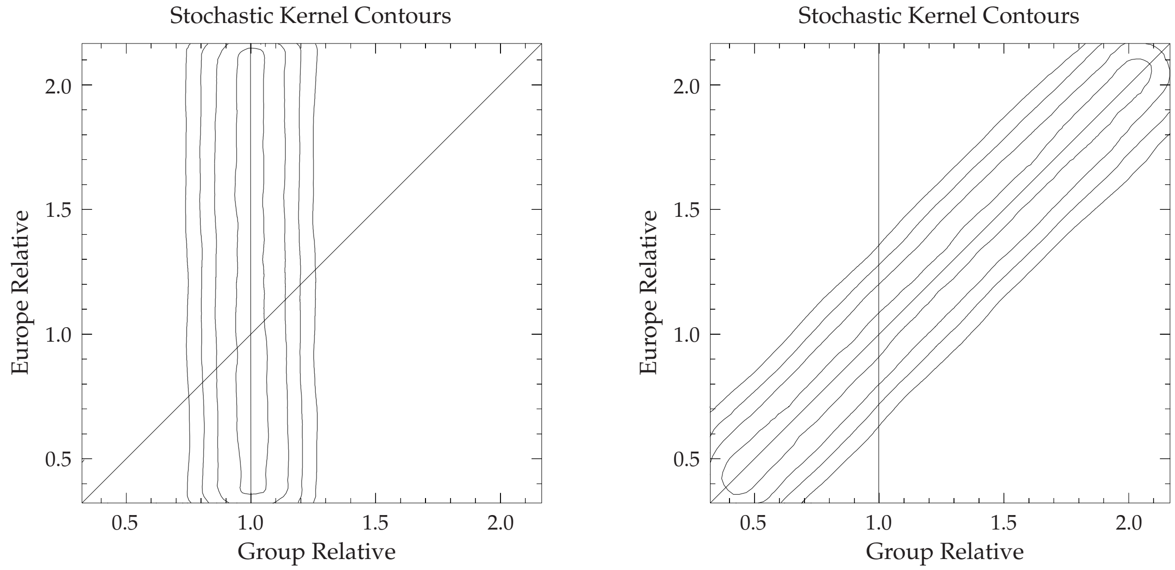

Figure 6. Benchmark stochastic kernels

kernel mapping Europe relative to State relative unemployment rates will then have (almost) all mass centred around one. The contour plot on the left of Figure 6 illustrates this benchmark. To read the contour plot, you need to picture a 'mountain range' running 'north-south' with the crest of the range around one — the figure plots the contours of this mountain range.

To understand the concept behind the three-dimensional kernels in this section, imagine taking a cross-section perpendicular to the Europe relative unemployment axis at some value: this gives the distribution of State relative (more generally, group relative) outcomes conditional on that value of Europe relative unemployment. The stochastic kernel plots this conditional distribution for all possible values of Europe relative unemployment. A related way to think about the kernels is to imagine taking ranges of Europe relative and State relative unemployment rates and integrating under the kernel. Just like a cell in a transition probability matrix, this value gives us the probability of a region in the chosen range of Europe relative outcomes also being in the chosen range of State relative outcomes.

The contour plot on the right of Figure 6 illustrates the opposite extreme. For this benchmark, there is a similar regional distribution within each State and (almost) identical State averages. In this case, high Europe relative unemployment rates correspond to high State relative unemployment rates (and vice-versa). The stochastic kernel mapping Europe relative to State relative unemployment rates then has (almost) all mass concentrated on the diagonal.

In reality we see neither of these two extremes. Figure 7 shows the actual Europe relative to State

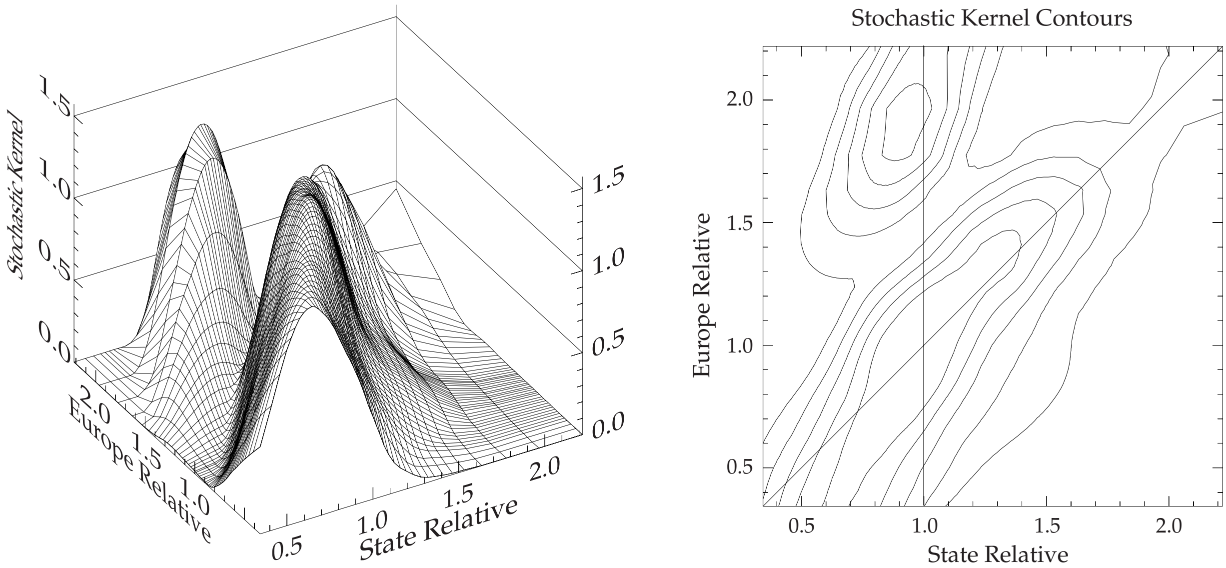

Figure 7. Europe relative to State relative stochastic kernel

relative stochastic kernel. The kernel is calculated using data for all eleven years[5]. 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 State average. That is, a regions 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 State outcomes. The spike around the vertical line in this range corresponds to approximately one half of Spanish regions with unemployment rates close to the Spanish average, plus Ireland[6] prior to 1994. However, there are also regions in this range whose outcome differs as much from their State average as from the European average, leading to a wide spread of mass to the right of the vertical line at one and close to the diagonal. This was 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.

[5]To estimate the kernel, we first derive 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/`). The three dimensional stochastic kernel plots are drawn so that the density of lines reflects the underlying number of observations on which that part of the kernel is estimated. This procedure makes the pictures easier to read and more informative without changing the shape of the kernel.

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

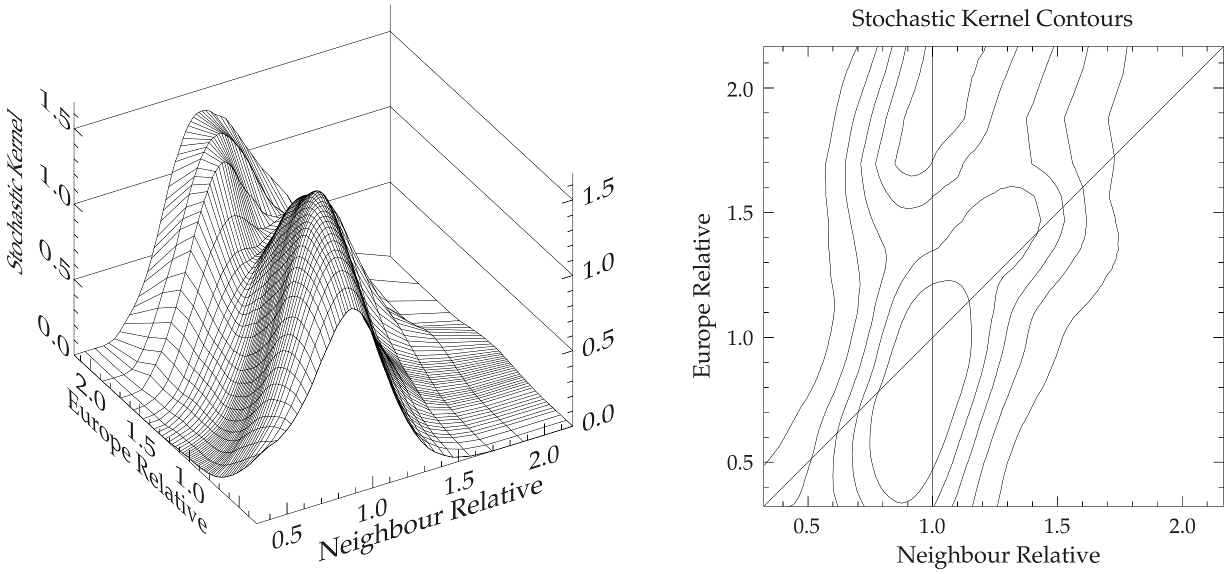

Figure 8. Europe relative to neighbour relative stochastic kernel

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 Member State. We now show that unemployment outcomes are closer to those of neighbouring regions.

To show this, we construct a kernel mapping Europe relative to *neighbour relative* unemployment rates, defined as each region's unemployment rate divided by the labour force weighted average of the unemployment rates of contiguous regions (including foreign neighbours, but not including the region itself). Note that this definition means that the neighbourhood groupings are overlapping. To take a specific example, the Spanish region Catalunya is a neighbour to the French regions of Midi-Pyrennes and Languedoc-Roussillon and to the Spanish regions Aragón and Comunidad Valenciana. The neighbours of Comunidad Valenciana include Aragón and Catalunya but not Midi-Pyrennes and Languedoc-Roussillon. Thus, unlike the other groupings, the resulting groups are not a partition of the set of regions. Any attempt to partition the set of regions in to neighbouring groups would require arbitrary definitions of neighbours.

Comparison of Figure 8 with Figure 7 shows that regional unemployment outcomes are much closer to the outcomes of neighbouring regions than to the outcomes of regions in the same Member State. This difference is particularly clear when one contrasts figures 7 and 8 in the '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 7 and 8 (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 State relative kernel. The depth of the valley is particularly relevant, because, as we have seen in Section 2, it is in this intermediate unemployment range that regions with similar starting positions had the very different evolutions that drove polarisation.

Also, note that a regions' domestic neighbours are part of the groups used to construct both

	n		State Relative				
Europe Relative	240	[1.45–∞)	0.00	0.06	0.43	0.13	0.38
	201	[1.15–1.45)	0.00	0.13	0.16	0.52	0.18
	577	[0.75–1.15)	0.00	0.05	0.68	0.18	0.09
	330	[0.55–0.75)	0.06	0.30	0.50	0.12	0.02
	302	[0–0.55)	0.30	0.36	0.32	0.01	0.00
			[0–0.55)	[0.55–0.75)	[0.75–1.15)	[1.15–1.45)	[1.45–∞)

	n		Neighbour Relative				
Europe Relative	240	[1.45–∞)	0.00	0.04	0.45	0.30	0.21
	201	[1.15–1.45)	0.00	0.06	0.40	0.37	0.16
	577	[0.75–1.15)	0.01	0.07	0.63	0.21	0.07
	330	[0.55–0.75)	0.04	0.16	0.72	0.07	0.01
	302	[0–0.55)	0.22	0.17	0.55	0.06	0.00
			[0–0.55)	[0.55–0.75)	[0.75–1.15)	[1.15–1.45)	[1.45–∞)

Table 2. Europe relative to State and neighbour relative transition probability matrices

kernels. In Figure 8, however, other regions in the same State are included. In Figure 7 they are not, but foreign neighbours are. This suggest 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.

In order to check the visual ranking of the kernels, we discretise the state space of relative unemployment rates and calculate the transition matrices that are the discrete versions of the continuous stochastic kernels. These discretisations, presented in Table 2, allow us estimate the relative mass in different areas of the kernels without having to integrate explicitly. To interpret these matrices it is useful to compare them with the same benchmarks we used to interpret the corresponding stochastic kernel: large numbers on the column for the interval containing one, versus large numbers on the diagonal. We see that the Europe relative to neighbour relative matrix has all diagonal elements smaller than those of the Europe relative to State relative matrix. At the same time, all other elements in the central column are larger in the Europe relative to neighbour relative 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. This finding — that 'nearness' matters more — is consistent with polarisation being driven by changes in labour demand or supply that are similar across neighbouring regions. We now check whether we should be focusing on demand or supply changes.

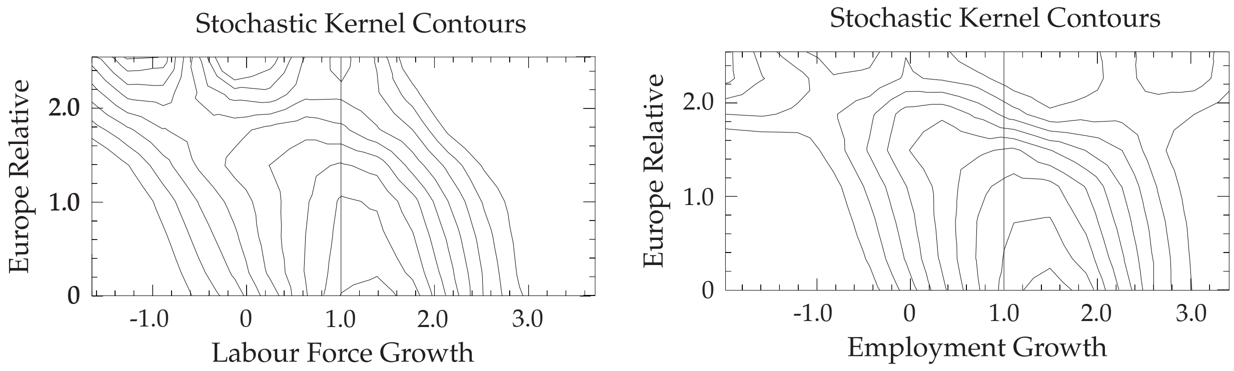

Figure 9. Labour force and employment growth

5. Is polarisation demand or supply driven?

We can use our data to test whether it is labour force or employment changes that are driving polarisation. The plot on the left hand side of Figure 9 graphs the stochastic kernel mapping the distribution of 1996 Europe relative unemployment rates to the distribution of labour force changes between 1986–1996 (relative to the average growth in the European labour force over the decade). The vertical line at one marks regions with labour force growth equal to the European average, 6.3%. The concentration of mass at the bottom right of the figure shows that most regions that ended up with relatively low unemployment had relatively high labour force growth. Similarly, the concentration of nearly all mass at the top of the figure to the left of one shows that regions that ended up with relatively high unemployment generally had below average labour force growth. Thus, labour force changes have actually worked *against* polarisation. They have prevented high unemployment regions from having even higher unemployment and low unemployment regions from having even lower unemployment.

The plot on the right hand side of Figure 9 graphs the stochastic kernel mapping the distribution of 1996 Europe relative unemployment rates to the distribution of employment changes between 1986–1996 (again, relative to the European average). The vertical line at one marks regions with labour force growth equal to the European average. The concentration of mass at the bottom right of the figure shows that most regions that ended up with relatively low unemployment had relatively high employment growth. Similarly, the concentration of mass at the top of the figure to the left of one 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* polarisation. 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 polarisation process like that represented in Figure 5.

Therefore, the polarisation of the distribution of 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 and, to a much lesser extent, labour force, and thus seeing their unemployment rate rise. We now turn to discriminating between possible sources of demand changes and establish why these changes might be similar across neighbouring regions.

6. The factors driving polarisation

In this section, we consider a number of factors that may be changing the demand for labour across regions and assess which of these factors is driving the 'nearness' result that we identified in Section 4.

A first possibility arises 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 skill workers will have experienced similarly high unemployment outcomes, while regions with a small proportion of low skill 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 polarisation are the result of changes in the sectoral composition of EU industry. The period 1986 to 1996 saw the continuation of an ongoing shift of European employment from agriculture, mining, and industry into services. In the absence of counteracting labour force changes, this may have resulted in high unemployment rates for regions with initial specialisation in declining sectors. In this case the similarity across neighbours could be a result of regions with declining sectors being contiguous.

A third possibility is that the location of regions within the EU matters because of the relocation of firms that is occurring as Europe becomes increasingly integrated. Recent location theories (Fujita, Krugman, and Venables, 1999) predict significant changes in the spatial distribution of employment in response to such a process of economic integration. In the European context, this is likely to take the form of an increasing agglomeration of employment Puga (1999). This could take the form of agglomerations of specific activities or, if there are significant cross-sector linkages, agglomerations of overall employment. 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 (Midelfart-Knarvik, Overman, Redding, and Venables, 2000) find evidence of both types of changes. If this is what is driving polarisation, then we would expect to see changes in regions' unemployment closely related to those of their neighbours even after we control for State membership and for changes related to their skill and sectoral composition or other similar factors.

To discriminate between these possibilities, we examine the cross-section of changes in regional unemployment rates as a function of State and regional characteristics. Table 3, column 1, shows ordinary least squares results for our first empirical specification. The dependent variable is the (logarithm of the) change in the regional unemployment rate between 1986 and 1996. We consider a number of different explanatory variables. We include the regions initial unemployment rate to allow for some conditional convergence to the mean in EU unemployment rates. We include two

Dependent variable Δ unemployment 1986–1996	1 OLS	2 IV	3 OLS	4 IV	5 OLS	6 IV	7 OLS 3-year MA	8 IV 3-year MA
initial unemployment	-0.128^{**} (0.057)	-0.100^{**} (0.061)	-0.215^{**} (0.063)	-0.202^{**} (0.066)	-0.231^{**} (0.063)	-0.219^{**} (0.068)	-0.188^{**} (0.043)	-0.160^{**} (0.046)
% low skill	0.247^{**} (0.106)	0.241^{**} (0.101)	0.220^{**} (0.093)	0.241^{**} (0.098)	0.196^{**} (0.070)	0.216^{**} (0.075)	0.155^{**} (0.062)	0.182^{**} (0.072)
% medium skill	−0.080 (0.158)	0.025 (0.184)	0.093 (0.126)	0.137 (0.139)	0.082 (0.121)	0.118 (0.129)	0.087 (0.100)	0.137 (0.109)
initial % agriculture	−0.026 (0.019)	−0.037 (0.021)						
initial % manufacturing	-0.278^{**} (0.090)	-0.230^{**} (0.093)						
initial % NACE17 sectors	No	No	Yes	Yes	Yes	Yes	Yes	Yes
country dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Δ neighbours' unemployment	0.541^{**} (0.111)	0.864^{**} (0.264)	0.550^{**} (0.084)	0.833^{**} (0.257)				
Δ domestic neighbours					0.630^{**} (0.089)	0.830^{**} (0.181)	0.759^{**} (0.094)	1.000^{**} (0.199)
Δ foreign neighbours					0.266^{**} (0.134)	0.527^{*} (0.285)	0.399^{**} (0.139)	0.777^{**} (0.271)
Adjusted R^2	0.73	0.71	0.84	0.83	0.85	0.84	0.82	0.81
No. observations	147	147	147	147	147	147	147	147

Heteroscedastic robust standard errors in parenthesis
** denotes coefficient significantly different from zero with 5% confidence level, * with 10% level

Table 3. Regression results

variables that capture the skill composition of the region — the percentage of adult population with low skills (highest educational qualification: less than upper secondary education), and the percentage with medium skills (highest educational qualification: completed upper secondary education). To capture the initial structure of employment in the region we use the percentage of employment in each sector. We start with a very aggregate classification with just three sectors and include two of them in the regression — the percentage of regional employment in primary sectors (agriculture, mining, forestry, and fishing), and the percentage of regional employment in manufacturing. Below, we move on to a more detailed sectoral classification at the NACE17 level. Country dummies are included, but not reported, in this and all other specifications. These dummies should capture the effects of national characteristics such as cross-country differences in labour market institutions. We exclude Member States classified as a single NUTS2 region (Denmark, Ireland, and Luxembourg) from the regressions. To capture the location 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). All explanatory variables are expressed in logarithms. Further details on data definitions and sources are given in the Data Appendix. Heteroscedastic robust standard errors are reported in parenthesis.

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 polarisation of unemployment rates that we documented in Section 2 must be explained by something other than purely the initial unemployment rate. The coefficient on the percentage of adult population with low skills is positive, large, and significant, as would be expected. After conditioning on the other variables, a high proportion of 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 the percentage of initial employment in agriculture and other primary sectors is not significantly different from zero. However, the percentage of initial employment in industry has a negative effect on unemployment rate changes. To interpret this sign, it is worth noting that for most of the Northern and Central European regions traditionally specialised in heavy industry, the worst part of the adjustment to their decline was over by the mid 1980s. Since then many of these regions have seen their unemployment rate fall. Adjustment has taken place later in heavy industrial regions in Southern Europe. Since the latter have a higher proportion of population with low skills, this can explain why, after controlling for skills, the effect of manufacturing specialisation on unemployment changes is negative.

The most remarkable aspect of these results, however, is that the evolution of the unemployment rate in neighbours has a very strong and significant effect, even after controlling for common characteristics. Thus, common state membership, common skills and sectoral composition are not driving the nearness effect. Before interpreting this result further, we discuss a number of econometric issues and then conduct a number of robustness checks.

We capture the neighbour effect through changes in neighbours' unemployment, rather than through covariance assumptions on the error structure, because we want to capture the impact of predictable increases in neighbouring unemployment. Such expected increases are, by definition, orthogonal to the error, and thus best captured through the inclusion of a 'spatially lagged' dependent variable (see Anselin, 1988, for further discussion). However, spatially lagged dependent variables are correlated with the error (a region's unemployment affects its neighbour's unemployment, which in turn affects the region's unemployment, and so on). To solve this endogeneity problem, we instrument for the spatially lagged dependent variable.

Neighbour's 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 Gross Domestic Product (GDP)s[7]. Instrumental variables (IV) results using this set of instruments are presented in Table 3, column 2 (in this, and all subsequent specifications we cannot reject the validity of our instrument set at the 5% confidence level using the test proposed by Davidson and MacKinnon, 1993). This shows that instrumenting does not change our initial results. The proportions of low educated and initial industrial employment remain significant. The effect of neighbours' unemployment remains strong and significant.

Our initial specification used a very crude characterisation of employment structure in the region. The level of aggregation reflects data availability — at the NUTS2 regional level, sectoral employment data is only available for the three NACE3 sectors. However, if one is willing to disaggregate data in some instances from the NUTS1 to the NUTS2 level, it is possible to construct approximations of the structure of each region based on value added data at a finer level of sectoral disaggregation corresponding to the NACE17 classification. Further details on the assumptions necessary are available in the Data Appendix. Column 3 shows results when we use these NACE17 sectoral variables. Allowing for greater sectoral disaggregation does not change any of our results. Column 4 shows, again, that instrumenting doesn't change these results.

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 (the percentage of population that reached working age during the period (those aged between 15 and 25 in 1996) and female participation rates. We included the first to capture the impact of high and rising European youth unemployment rates. We included the second because in the mid-1980s female participation rates differed widely across regions. Some regions, in Spain, had participation rates as low as 18%, while others, in the UK, had rates above 50%. Over the decade, female participation rates

[7] Thus, for region i, market potential is defined as $mp_i = \sum_{j \neq i} \text{GDP}_j / d_{i,j}$, where $d_{i,j}$ is the great circle distance between region i and region j, and GDP_j is the GDP of region j, and the sum is over all regions in the European Union excluding region i itself.

significantly converged across regions. This has resulted in large labour force increases in some regions, potentially affecting unemployment rates (see Wasmer, 1998, for an exposition of this argument). Both coefficients had the expected sign, but were insignificant, and so we do not report detailed results here. In addition we have tried including the average change in unemployment for regions with a similar initial sectoral specialisation, a similar skill composition of adult population, and so on. The results are still remarkably robust.

We have seen that the neighbour effect is strong and significant even after conditioning out the fact that neighbouring regions share similar characteristics. We can also test whether that neighbour effect extends across national borders. To do this, we split the neighbours variable for border regions into two components, that due to domestic neighbours and that 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 a third of the sample. If we drop out 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 columns 5 and 6.

Column 5 provides OLS results for the basic specification. Both neighbour effects are strong and significant. Again, the two neighbour effects are endogenous, so we instrument for them. The results are reported in column 6. Both neighbour effects remain strong and significant. Although the domestic neighbours have a higher coefficient, it is noticeable that, once we instrument for the effects, we are unable to reject the hypothesis that the coefficients are identical — the test has a p-value of 0.18. 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 7 and 8 shows that this only strengthens the results, presumably because it removes business-cycle-related noise. Again, all results are identical when we instrument (we also time-average instruments where appropriate).

Our regression results provide us with a better understanding of the neighbour effect. First, we see that neighbours are important for understanding the evolution of unemployment rates during this process of polarisation. Second, the neighbour effect is *not* driven by neighbouring regions sharing similar characteristics. Third, the neighbour effect transcends national borders.[8] To further illustrate these findings, we consider the concrete example of two border regions in Belgium.

[8]The neighbour effect is also not driven by functional labour markets extending across neighbouring regions 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). Second, 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 commutes 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.

7. An 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 European Union 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 kilometres 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 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. However, in recent years these two borders have experienced very different evolutions. The most publicised 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. In this paper we have shown that this story is not unique, but representative of a broader pattern that has developed across Europe.

8. Regional disparities and regional policies

This paper has documented the following facts about the regional unemployment in the EU:

- There has been a polarisation of unemployment rates.
- This polarisation 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 nearness effect, since it is only partly explained by State membership or common regional characteristics.
- This geographical component is as strong within as across national borders.

We now discuss the policy implications of these findings. We begin by comparing the evolution of unemployment with the evolution of income per capita which has been the main focus of EU regional policies.

A comparison with income inequalities

Our results will be particularly surprising to anyone that has followed recent work on the evolution of regional income disparities across European regions. Since the 1980s there have been few changes in the distribution of regional income per person (on this respect see, amongst others, Boldrin and Canova, 2001, Canova and Marcet, 1995, López-Bazo, Vayá, Mora, and Suriñach, 1999, Marcet, 1994, Neven and Gouyette, 1995, Quah, 1999, and Rodríguez-Pose, 1999).

This immobility in the income distribution is evident from the transition probability matrix for relative GDP per person presented in Table 4. To ease the comparison with unemployment rates, we have divided the state space for relative GDP per person into the same classes in which we divided the state space for relative unemployment rates in Table 1 (the slight difference in years covered is due to data availability).

The diagonal of the transition matrix in Table 4 gives, once again, the proportion of regions that were in the same range of the distribution in the mid-1990s as in the mid-1980s. For regions

	n		1995 GDP per capita				
1987 GDP per capita	18	[1.3–∞)	0.00	0.00	0.00	0.17	0.83
	53	[1–1.3)	0.00	0.00	0.13	0.72	0.15
	50	[0.75–1)	0.00	0.18	0.68	0.14	0.00
	19	[0.6–0.75)	0.21	0.47	0.32	0.00	0.00
	30	[0–0.6)	0.83	0.17	0.00	0.00	0.00
			[0–0.6)	[0.6–0.75)	[0.75–1)	[1–1.3)	[1.3–∞)

Table 4. 1987 to 1995 Europe-relative GDP per capita transition probability matrix

initially below 60% of the European average, the proportion of regions that remained in the same range is as high for income (83%) as for unemployment (81%). For regions initially above 130% of the European average, persistence is also very high in both distributions, although it is more marked for income than for unemployment (83% against 62%). But for regions initially around the middle of the distribution the difference is striking: most regions with intermediate income levels remained in the same range whereas most regions with intermediate initial unemployment rates moved to a different range.

GDP per capita and unemployment rates are not two separate dimensions of regional inequalities. They are related to each other by the following identity:

$$\frac{\text{GDP}}{\text{population}} \equiv \frac{\text{GDP}}{\text{employment}} \times \frac{\text{employment}}{\text{labour force}} \times \frac{\text{labour force}}{\text{working-age population}} \times \frac{\text{working-age population}}{\text{population}}$$
$$= \text{productivity} \times (1 - \text{unemployment rate}) \times \text{participation rate} \times \text{working-age pop. share}\ .$$

Esteban (1999) decomposes inequalities in GDP per capita, as measured by a Theil index, into the components in this identity. He finds that, while differences in productivity are still the largest component, differences in unemployment rates are increasingly important.

In contrast European policy is still primarily focused on income differences between the core and the periphery. This emphasis on income differences is reflected in the definition of the EU's regional policy objectives and the allocation of funds across the different objectives. Since 2000, the EU identifies three objectives of 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 adaption and modernisation of policies and systems of education, training and employment. Specific regions are identified as eligible for assistance under the first two objectives. The key requirement for eligibility under objective 1 is that the region has a Gross Domestic Product per capita less than 75% of the EU average. These objective 1 regions then receive the lions share — approximately 70% — of total EU regional expenditure. This criterion has the benefit of being clearly stated and relatively easily verifiable. There is an understandable reluctance to allocating a larger share of this money on the basis of unemployment inequalities. At the same time there seem to be no adequate policies in place to tackle a polarisation of regional unemployment with the characteristics that we have demonstrated. To conclude the paper, we want to offer a few policy recommendations in this respect.

Reinforcing skills

Our regression results in Section 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 emphasise 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 coordinate these expenditures across neighbouring regions with similar skill compositions. However, this policy cannot on its own tackle the polarisation that we document in this paper. Firstly, because education and training dimensions to regional policy have been present throughout the period that we study; and secondly, because polarisation is not purely driven by shifts in demand away from low skilled workers. We now turn to policies that should help tackle these additional causes of polarisation.

Adding a crossregional 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 because we have seen that high unemployment problems extend across regional and national borders. The clusters of high and low unemployment that we have found suggest that if a region was to implement a policy that managed to reduce unemployment, part of the benefits would be felt in neighbouring regions. 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.

In such circumstances the cost of a policy is completely born by some jurisdiction, but the benefits partly accrue to neighbouring jurisdictions, from which local politicians gain no votes or tax revenues. This is particularly true for regions in different countries. The standard analysis of externalities tells us that this will lead to an under provision of policies. One clear area where such an externality is evident is in the provision of training discussed above. Two neighbouring regions may suffer from a common lack of skills. If one of these regions was to provide additional relevant 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. Since the cost of their training would have to born by taxpayers in the region that they left, while these workers would pay taxes and vote in a different region, each region would tend to provide insufficient training. All of this suggest that it is important to add a crossregional *and* transnational dimension to EU unemployment policies.

Regarding the practical implementation of such coordination, 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 coordinate some elements of policy across neighbouring regions.

Coordination of policies across neighbouring regions and a greater focus on training will go some way towards mitigating polarisation 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.

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 polarisation. In fact, this reflects the actual institutional setting in several European countries, where various constraints severely limit interregional wage differentials. In Germany, reunification led to a sharp reduction in regional wage differences between the Eastern and Western Lander. According to Akerlof, Rose, Yellen, and Hessenius (1991), this was largely due to strong union bargaining for wage equalisation to prevent large scale migration, and the perception on the part of workers that higher unemployment in the East did not justify lower wages. In Italy and Spain, wage setting at the national level also limits the responsiveness of wages to regional economic conditions (on this respect, see Jimeno and Bentolila, 1998, who stress the impact on the regional structure of wages in Spain of wage floors set at the national sectoral level). As a result, amongst economists studying European labour markets 'the common view is that local supply and demand conditions play only a limited role in the determination of regional wages' (Bentolila and Dolado, 1991, pp. 215-216). This is worsened by the fact that interregional equalisation is often imposed on nominal as opposed to real wages, and that the cost of living tends to be higher in regions that have more favourable labour market conditions — so instead of being higher, real wages there are pulled below real wages is high unemployment regions. Thus, promoting regional wage setting could do much to alleviate the polarisation of European unemployment.

Some countries have tried to compensate for national wage setting by subsidising 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 subsidies 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 and rigidities arising from the institutional framework prevent this, then the first-best solution is to try to reduce or eliminate those rigidities.

Second, the long life that this type of subsidies tends to have shows that they may result in a strong dependency and be difficult to eliminate.

Third, 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 localised 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 equalisation adds to the rigidities preventing this. We now turn to looking at this mobility argument in more detail.

Promoting short distance mobility

Referring back to Figure 5 it should be clear that polarisation of unemployment due to shifts in demand can be offset by corresponding shifts in supply. We have shown that these offsetting changes in supply have taken place only to a very small extent. Further changes in supply would require more mobility in and out of the labour force, or increased migration. In Europe temporary labour demand changes appear to trigger larger changes in labour participation than migration in contrast to the United States where the opposite occurs (Decressin and Fatàs, 1995, Blanchard and Katz, 1992). Clearly, from a policy perspective, reducing unemployment by having discouraged job seekers stop looking for a job altogether provides only statistical comfort. This suggests that adjustment needs to come through increased migration.

The problem is that international and interregional mobility in Europe has been very low in recent decades(see Chapter 4 in Braunerhjelm, Faini, Norman, Ruane, and Seabright, 2000). This tendency has been compounded by the fact that nation states seem to be remarkably reluctant to accept the changes in population distribution that such migration would entail. This seems to be particularly true with regard to migration between Member States, but is also true of migration flows between broad regions within Member States (North-South migration in Italy and the UK, or

East-West migration in Germany). However, our results suggest that even facilitating migration at a finer geographical level might go a long way towards decreasing polarisation.

Indeed, our finding that unemployment outcomes are so much more homogenous across neighbours, foreign and domestic, than across regions in the same Member State highlights the spatial dimensions of the emerging clusters of high and low unemployment in Europe. The average Member State has 13.6 regions, while the average neighbourhood has 5.6 regions. Hence these are clusters of typically less than one half of the size of the average Member State of the European Union, but often extend across national borders and include regions from more than one Member State. 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.

One of the key motivations behind European integration is to achieve a more efficient organisation of EU wide production. Inevitably this entails changing labour demand patterns across regions which have resulted in a polarisation of unemployment outcomes. Let us not fall into the temptation of giving up the important gains from integration for lack of courage or imagination to implement the right policies.

References

Akerlof, George A., Andrew K. Rose, Janet L. Yellen, and Helga Hessenius. 1991. East germany in from the cold: the economic aftermath of currency union. *Brooking Papers on Economic Activity* (1):1–105.

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.

Bentolila, Samuel and Juan José Dolado. 1991. Mismatch and internal migration in Spain, 1962–86. In Fiorella Padoa Schioppa (ed.) *Mismatch and Labour Mobility*. Cambridge: Cambridge University Press, 182–234.

Blanchard, Olivier Jean and Lawrence F. Katz. 1992. Regional evolutions. *Brooking Papers on Economic Activity* (1):1–61.

Boldrin, Michele and Fabio Canova. 2001. Europe's regions: Income disparities and regional policies. *Economic Policy* (32):205–253.

Braunerhjelm, Pontus, Riccardo Faini, Victor Norman, Frances Ruane, and Paul Seabright. 2000. *Integration and the Regions of Europe: How the Right Policies Can Prevent Polarization*. London: Centre for Economic Policy Research.

Burda, Michael and Charles Wyplosz. 2001. *Macroeconomics. A European Text*. Third edition. Oxford: Oxford University Press.

Canova, Fabio and Albert Marcet. 1995. The poor stay poor: Non-convergence across countries and regions. Discussion Paper 1265, Centre for Economic Policy Research. URL `http://www.cepr.org/`.

Centraal Bureau Voor de Statistiek. 1987. *Regionaal Statistich Zakboek 1986*. The Hague: Centraal Bureau Voor de Statistiek.

Cheshire, Paul and Gianni Carbonaro. 1996. Urban economic growth in Europe: Testing theory and policy prescriptionsl. *Urban Studies* 33(7):1111–1128.

Davidson, Russell and James G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. New York: Oxford University Press.

de Falleur, M. and V. Vandeville. 1996. Cross-border flows of workers in Europe: Facts and determinants. In Harvey Armstrong and R. W. Vickerman (eds.) *Convergence and Divergence Among European Regions*. London: Pion, 177–189.

Decressin, Jörg and Antonio Fatàs. 1995. Regional labor market dynamics in Europe. *European Economic Review* 39(9):1627–1655.

Esteban, Joan María. 1999. L'euro y la desigualtat territorial: implicacions per a Catalunya. In Joan María Esteban and Jordi Gual (eds.) *Catalunya dins l'Euro*. Barcelona: Antoni Bosh Editor, 165–210.

Esteban, Joan María, Carlos Gradín, and Debraj 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.

Esteban, Joan María and Debraj Ray. 1994. On the measurement of polarization. *Econometrica* 62(4):819–851.

Eurostat. 1989. *Regions: Statistical Yearbook 1988*. Luxembourg: Office for Official Publications of the European Communities.

Eurostat. 1990. *Regions: Statistical Yearbook 1989*. Luxembourg: Office for Official Publications of the European Communities.

Eurostat. 1995. *Regions: Nomenclature of Territorial Units for Statistics (NUTS)*. Luxembourg: Office for Official Publications of the European Communities.

Eurostat. 1997. *Education Across the European Union: Statistics and Indicators 1996*. Luxembourg: Office for Official Publications of the European Communities.

Eurostat. 1998. *Regio Database*. Luxembourg: Eurostat. May 1998 CD-ROM version.

Fujita, Masahisa and Paul R. Krugman. 1995. When is the economy monocentric? Von Thünen and Chamberlin unified. *Regional Science and Urban Economics* 25(4):508–528.

Fujita, Masahisa, Paul R. Krugman, and Anthony J. Venables. 1999. *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge: MIT Press.

Hanson, Gordon H. 1998. Market potential, increasing returns, and geographic concentration. Working Paper 6429, National Bureau of Economic Research. URL http://www.nber.org/.

Jimeno, Juan Francisco and Samuel Bentolila. 1998. Regional unemployment persistence (spain, 1976–1994). *Labour Economics* 5(1):25–51.

Layard, Richard, Stephen Nickell, and Richard Jackman. 1991. *Unemployment: Macroeconomic Performance and the Labour Market*. Oxford: Oxford University Press.

López-Bazo, Enrique, Esther Vayá, Antonio J. Mora, and Jordi Suriñach. 1999. Regional economic dynamics and convergence in the European Union. *Annals of Regional Science* 33(3):343–370.

Manacorda, Marco and Barbara Petrongolo. 1998. Skill mismatch and unemployment in OECD countries. Discussion Paper 307, Centre for Economic Performance, London School of Economics.

Marcet, Albert. 1994. Los pobres siguen siendo pobres: Convergencia entre regiones y países, un análisis bayesiano de datos de panel. In *Crecimiento y convergencia regional en España y Europa*, volume 2. Bellaterra: Instituto de Análisis Económico (CSIC).

Midelfart-Knarvik, Karen Helene, Henry G. Overman, Stephen J. Redding, and Anthony J. Venables. 2000. The location of European industry. Economic Papers 142, European Commission Directorate-General for Economic and Financial Affairs.

Neven, Damien and Claudine Gouyette. 1995. Regional convergence in the European Community. *Journal of Common Market Studies* 33(1):47–65.

Nickell, Stephen and Brian Bell. 1995. The collapse in demand for the unskilled and unemployment across the OECD. *Oxford Review of Economic Policy* 11(1):40–62.

Overman, Henry G. and Diego Puga. 1999. Unemployment clusters across European regions and countries. Discussion Paper 2250, Centre for Economic Policy Research.

Puga, Diego. 1999. The rise and fall of regional inequalities. *European Economic Review* 43(2):303–334.

Quah, Danny T. 1996. Regional convergence clusters across Europe. *European Economic Review* 40(3–5):951–958.

Quah, Danny T. 1997. Empirics for growth and distribution: Stratification, polarisation, and convergence clubs. Discussion Paper 324, Centre for Economic Performance, London School of Economics. URL http://econ.lse.ac.uk/~dquah/.

Quah, Danny T. 1999. Regional cohesion from local isolated actions — historical outcomes. In *Study of the Socio-economic Impact of the Projects Financed by the Cohesion Fund — A Modelling Aproach*, volume 2. Luxembourg: Office for Official Publications of the European Communities.

Rodríguez-Pose, Andrés. 1999. Convergence or divergence? Types of regional responses to socio-economic change. *Tijdschrift voor Economische en Sociale Geografie* 90(4):363–378.

Silverman, B. W. 1986. *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall.

Wasmer, Etienne. 1998. Can labour supply explain the rise of unemployment and inequality in OECD countries? Discussion Paper 410, Centre for Economic Performance, London School of Economics. URL http://www.ecare.ulb.ac.be/ecare/etienne/home.htm.

Data Appendix

Our definition of regions corresponds to level two of the Nomenclature of Territorial Units for Statistics (NUTS), 1995 version (Eurostat, 1995). The NUTS was established by Eurostat to provide comparable regional breakdowns of the Member States of the European Union. It is a hierarchical classification with three regional levels: each Member State is partitioned into an integral number of NUTS1 regions, each of which is in turn partitioned into an integral number of NUTS2 regions, each of which is in turn partitioned into an integral number of NUTS3 regions. (There are two additional sub-regional or local levels, NUTS4 and NUTS5, of which only the latter, consisting of Communes or their equivalent, is defined for all Member States). In 1996 the EU had 77 NUTS1 regions, 206 NUTS2 regions, and 1,031 NUTS3 regions. Eurostat (1995) also calls NUTS2 regions 'Basic Regions', and describes these as the appropriate level for analysing regional-national problems; it is also the level at which both national and Community regional policies are generally implemented.

NUTS2 regions correspond to national administrative units in Austria (Bundesländer), Belgium (Provinces), Finland (Suuralueet), Germany (Regierungsbezirke), Greece (Development Regions), Italy (Regioni), Netherlands (Provincies), Portugal (Comissaoes de Coordenaçao Regional), and Sweden (Riksområden). NUTS2 regions also correspond to national administrative units, but with exceptions, in France (Régions, plus the four Departements d'Outre Mer), and Spain (Comunidades Autónomas, plus Ceuta y Melilla). Three Member States are classified as a single NUTS2 region: Denmark, Ireland, and Luxembourg. In the UK, Groups of Counties have been introduced as an intermediate (NUTS2) level between NUTS1 (Standard Regions) and NUTS3 (a combination of Counties and Local Authority Regions) units.

The data set includes (with a single exception, documented below) all the NUTS2 regions of the EU that satisfy the following three criteria:

1. Have been part of the EU (European Economic Community before 1 November 1993) from 1986 to 1996.
2. Are in a Member State which has a land border with at least one other Member State containing at least one region satisfying (1).
3. Have a land border with at least one other NUTS2 region satisfying (1) and (2).

We include as land borders water borders less than five kilometres wide. This leads us to consider as geographical neighbours regions separated by a river (such as Zeelland and Zuid-Holland in Netherlands). It also leads to the inclusion of Sicilia (Italy), which, although an island, is only separated from Calabria (Italy) by the 3,300 metres-wide Strait of Messina — soon to be joined by a single span suspension bridge (see `http://www.strettodimessina.it/`).

From the 206 NUTS2 regions that formed the EU in 1996, 30 are excluded from the analysis because they were not part of the European Economic Community in 1986: the nine NUTS2 regions of Austria, the six NUTS2 regions of Finland, and the eight NUTS2 regions of Sweden, all of which became part of the EU with the accession of these three Member States in 1995; and the seven NUTS2 regions of Germany that were part of the former Democratic Republic of Germany (Brandenburg, Mecklenburg-Vorpommern, Sachsen, Dessau, Halle, Magdeburg, and Thüringen), which only became part of the EU with German reunification in 1990.

Greece has no land border with any other Member State, so its 13 NUTS2 regions are also excluded.

Finally, another 12 NUTS2 regions are excluded because they have no land border with any other NUTS2 region satisfying criteria (1) and (2): Baleares, Ceuta y Melilla, and Canarias (Spain), Corse, Guadeloupe, Martinique, Guyane, and Réunion (France), Sardegna (Italy), Açores, and Madeira (Portugal), are all entirely surrounded by water and/or by territories which are not part of the

EU; Berlin (Germany) is entirely surrounded by NUTS2 regions which were part of the former Democratic Republic of Germany.

Flevoland (Netherlands) is the only region that satisfies criteria (1)-(3) above but has been excluded due to lack of data: there is no labour force or unemployment data for Flevoland for 1986, even from national sources (see Centraal Bureau Voor de Statistiek, 1987). Flevoland was created as a separate administrative unit (Provincie) in 1986 from the union of the Noordoost, Oostelijk Flevoland, and Zuidelijk Flevoland polders, reclaimed from the IJsselllake (a lake that used to be part of Zuiderzee, a former inlet of the North Sea), and in 1996 accounted for 1.8% of the population and 5.8% of the land area of Netherlands.

The 150 NUTS2 regions used are:

Belgium (11)	Brussels, Antwerpen, Limburg (Belgium), Oost-Vlaanderen, Vlaams Brabant, West-Vlaanderen, Brabant Wallon, Hainaut, Liége, Luxembourg (Belgium), Namur.
Denmark (1)	
France (21)	Ile-de-France, Champagne-Ardenne, Picardie, Haute-Normandie, Centre, Basse-Normandie, Bourgogne, Nord-Pas-de-Calais, Lorraine, Alsace, Franche-Comté, Pays de la Loire, Bretagne, Poitou-Charentes, Aquitaine, Midi-Pyrénées, Limousin, Rhône-Alpes, Auvergne, Languedoc-Roussillon, Provence-Alpes-Côte d'Azur.
Germany (30)	Stuttgart, Karlsruhe, Freiburg, Tübingen, Oberbayern, Niederbayern, Oberpfalz, Oberfranken, Mittelfranken, Unterfranken, Schwaben, Bremen, Hamburg, Darmstadt, Giessen, Kassel, Braunschweig, Hannover, Lüneburg, Weser-Ems, Düsseldorf, Köln, Münster, Detmold, Arnsberg, Koblenz, Trier, Rheinhessen-Pfalz, Saarland, Schleswig-Holstein.
Ireland (1)	
Italy (19)	Piemonte, Valle d'Aosta, Liguria, Lombardia, Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Emilia-Romagna, Toscana, Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia.
Luxembourg (1)	
Netherlands (11)	Groningen, Friesland, Drenthe, Overijssel, Gelderland, Utrecht, Noord-Holland, Zuid-Holland, Zeeland, Noord-Brabant, Limburg (Netherlands).
Portugal (5)	Norte, Centro (Portugal), Lisboa e Vale do Tejo, Alentejo, Algarve.
Spain (15)	Galicia, Asturias, Cantabria, País Vasco, Navarra, Rioja, Aragón, Madrid, Castilla-León, Castilla-La Mancha, Extremadura, Cataluña, Comunidad Valenciana, Andalucía, Región de Murcia.

United Kingdom (35)	Cleveland-Durham, Cumbria, Northumberland-Tyne and Wear, Humberside, North Yorkshire, South Yorkshire, West Yorkshire, Derbyshire-Nottinghamshire, Leicestershire-Northamptonshire, Lincolnshire, East Anglia, Bedfordshire-Hertfordshire, Berkshire-Buckinghamshire-Oxfordshire, Surrey-East-West Sussex, Essex, Greater London, Hampshire-Isle of Wight, Kent, Avon-Gloucestershire-Wiltshire, Cornwall-Devon, Dorset-Somerset, Hereford-Worcestershire-Warwickshire, Shropshire-Staffordshire, West Midlands (County), Cheshire, Greater Manchester, Lancashire, Merseyside, Clwyd-Dyfed-Gwynedd-Powys, Gwent-Mid-South-West Glamorgan, Borders-Central-Fife-Lothians-Tayside, Dumfries-Galloway-Strathclyde, Highlands-Islands, Grampian, Northern Ireland.

Regional unemployment rates and labour force from 1986 to 1996 are taken from the harmonised unemployment rates (Table `regio/unemp/un3rt`) and labour force (Table `regio/unemp/un3wpop`) in the May 1998 version of the Regio database published by Eurostat (Eurostat, 1998).

These data are based on the results of the Community Labour Force Survey (LFS). The Community LFS is carried out in Spring each year and for each Member State provides the number of the unemployed (in accordance with the definition of the International Labour Office), and the labour force (labelled 'working population') for April. The national unemployment data are subsequently regionalised to NUTS2 level on the basis of the number of persons registered at unemployment offices in April of the reference year (with the exceptions of Greece, Spain, Italy, Portugal, Finland, and Sweden, where the regional unemployment structures are taken from the Community LFS). The national labour force data are regionalised to NUTS2 level according to the results of the Community LFS. The regional unemployment rates are then obtained by dividing the number of the unemployed by the labour force.

The Regio database has no data on unemployment rates or labour force for two years, 1986 and 1987, for 13 of the targeted regions: all the NUTS2 regions of Netherlands, and Algarve (Portugal). For all of them (except the Dutch region of Flevoland, as documented above) comparable data has been obtained as follows. For the NUTS2 regions of the Netherlands in 1986 and 1987, the total number of the unemployed in the Netherlands in Table `/regio/unemp/un3pers` of the Regio database has been regionally disaggregated to NUTS2 level, on the basis of the number of the unemployed in each region from Table II.4 of Eurostat (1989), which are also derived from the Community LFS. Similarly, the total labour force of the Netherlands in Table `/regio/unemp/un3wpop` of the Regio database has been regionally disaggregated to NUTS2 level, on the basis of regional labour force figures from Table II.2 of Eurostat (1990) (for 1986), and of regional labour force figures computed by dividing the number of the unemployed by the corresponding unemployment rates in Table II.4 of Eurostat (1989) (for 1987). Regional unemployment rates have then been calculated by dividing the number of the unemployed by the labour force. For Algarve (Portugal) in 1986 and 1987, employment and unemployment figures have been privately obtained from national sources (Portugal's Instituto Nacional de Estatística for employment, and Direcçao de Serviços de Estudos de Mercado de Emprego for unemployment), and corrected for the factor by which each of these sources underestimates the corresponding Community LFS data for all the other NUTS2 regions that, together with Algarve, constitute the NUTS1 region Continente (Norte, Centro, Lisboa e Vale do Tejo, and Alentejo). Labour force has been calculated as the sum of the employed and the unemployed, and the unemployment rate by dividing the number of the unemployed by the labour force.

Regional unemployment rates and labour force are used to construct three series of relative unemployment rates: unemployment rates relative to the European average (*Europe relative* for brevity), i.e. *regional unemployment rate / European wide average unemployment rate*; unemployment rates relative to the average for other regions in the same Member State (*State relative*); i.e. *regional unemployment rate / Member state average unemployment rate*; unemployment rates relative to the average for contiguous regions (*neighbour relative*); i.e. *regional unemployment rate / labour force weighted average of contiguous regions*. In all cases averages used to construct the relative series refer only to regions included in the analysis. The information on State membership and contiguity is taken off the paper maps in Eurostat (1995).

The regression analysis of Section 6 uses the following additional data sources. For the purpose of splitting population by skill, low skill is taken to be an educational attainment of less than upper secondary education (below level 3 of the International Standard Classification of Education (ISCED) classification). Medium skill is an educational attainment of upper secondary education (level 3 of the ISCED classification). High skill is an educational attainment of higher education (levels 5, 6, and 7 of the ISCED classification). These data are from Table E14 in Eurostat (1997). These data are not ideal in that they refer to the adult population and not to the labour force, and they are only available for the 150 regions we are interested in for a single year, 1995. However, they are the best available at this level of regional disaggregation.

To calculate the percentage of young population, the young are taken to be those that reached working age during the sample period (those aged between 15 and 25 in 1996). These data are obtained from Table `/regio/lfs-r/lf2emp`) in Eurostat (1998). Initial female participation rates are those for 1986 from Table `/regio/lfs-r/lf2actrt`) in Eurostat (1998), completed with Eurostat (1989). For the calculation of the measure of initial market potential, used as one of the instruments in the instrumental variable estimations of Section 6, 1986 regional GDP levels are from Table `/regioecon-r/egdp/e2gdp`) in Eurostat (1998). The distance between each pair of NUTS 2 regions is the great circle distance between their geographical centres, the coordinates of which have been obtained from `http://shiva.pub.getty.edu/tgn_browser/`.

Data on employment by broad sectoral specialisation at level 3 of the Nomenclature generale des Activites economiques dans la Communaute Europeenne (NACE) classification (agricultural, forestry and fishery products; manufactured products; and market services) comes from Table `/regio/lfs-r/lf2emp` in Eurostat (1998). These data are available for the 150 regions we are interested in only for 1988, but this is close enough to the beginning of the time frame considered to describe early specialisation.

Sectoral value added by NACE17 is available at the NUTS2 level for Belgium, Denmark, Spain, France, Ireland, Italy, the Netherlands and Portugal. For the UK, sectoral value added by NACE17 is available at the NUTS1 level. For the regression, we assume that the industrial structure of each UK

NUTS2 is identical to the industrial structure of the UK NUTS1 containing that NUTS2. We would have liked to use information on sectoral value added at NACE3 to provide a more accurate estimate, but this is not available for the 1986-1996 time frame for which we want data. For Germany, we have information on employment at the NUTS1 level at NACE17 provided by the 11 individual West German Lander. We use this to break down NACE3 sectoral value added which Regio provides for German NUTS1. Then as for the UK, we assume that the industrial structure of each German NUTS2 is identical to the industrial structure of the German NUTS1 containing that NUTS2.