



Web Appendix

for:

Unemployment clusters across European regions and countries

Henry Overman, Diego Puga

London School of Economics, University of Toronto

Published in:

Economic Policy, 34, Spring 2002

CONTENTS

This appendix contains three sections. The first presents a detailed description of the data we use and various transformations we have undertaken. The next presents a more detailed picture of the evolution of regional unemployment inequality and the last presents two more stochastic kernel mappings

1. DATA DETAILS

1.1. Definition of regions

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 Coordenação 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 Zeeland 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, Baleares 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 IJssel lake (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, BrabantWallon, Hainaut, Liège, 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-WestGlamorgan, Borders-Central-Fife-Lothians- Tayside,Dumfries- Galloway-Strathclyde, Highlands-Islands, Grampian, Northern Ireland.

1.2. Regional labour market data

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 Direcção 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).

1.3. Addition data sources used in the regression analysis

The regression analysis of Section 6 uses the following additional data sources. For the purpose of splitting population by skill, lowskill 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 amore 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.

2. UNEMPLOYMENT INEQUALITY DISTRIBUTIONS IN MORE DETAIL

Figure A1 shows a sequence of kernel estimates of the density of Europe relative unemployment rates for four years: 1986, 1989, 1993, and 1996. Two features are particularly noticeable in Figure 3. First, as we move through the decade, the 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.

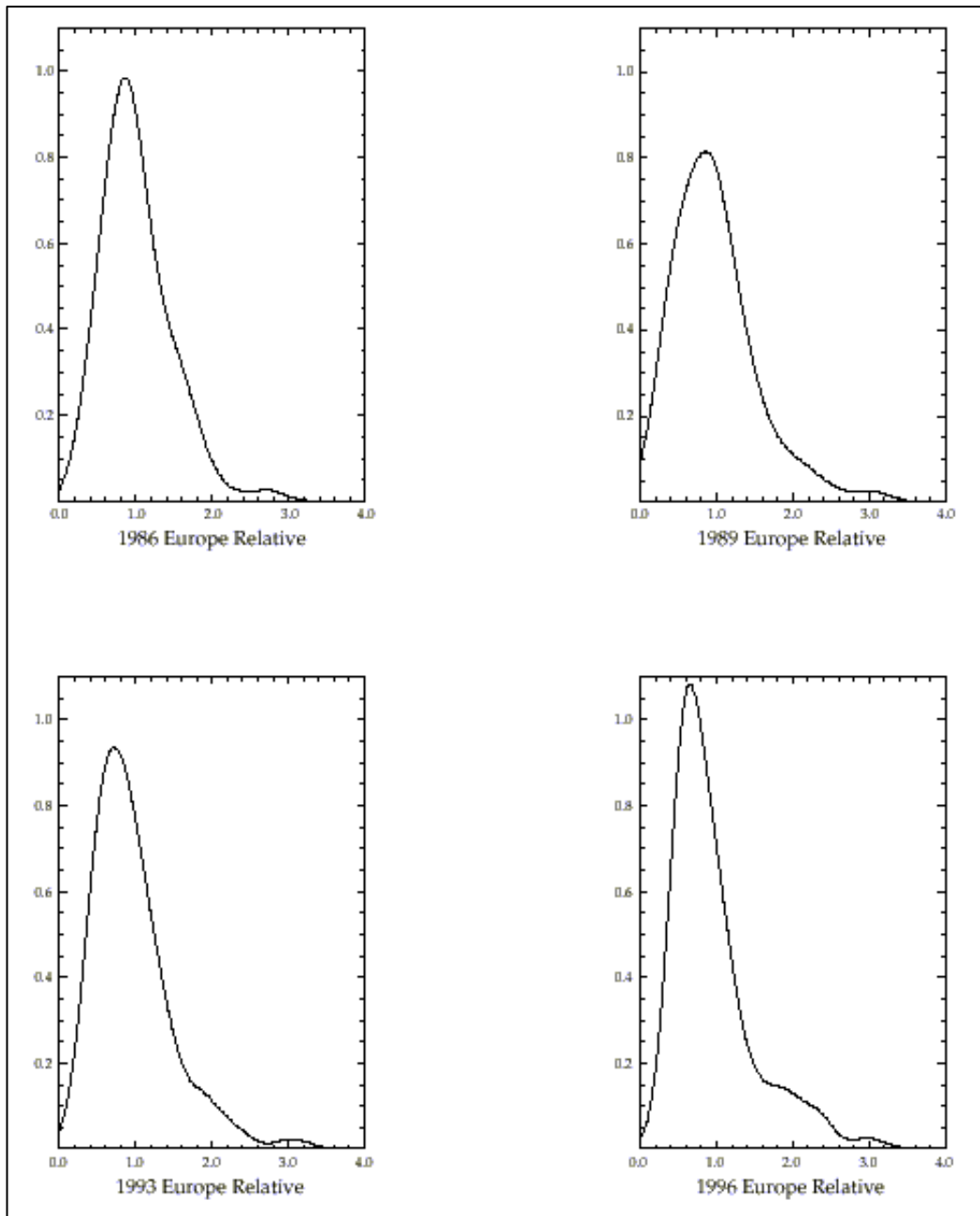


Figure A1: Densities of 'Europe relative' unemployment rates, 1986, 1989, 1994 and 1996.

3. STOCHASTIC KERNEL MAPPINGS FOR SECTOR AND SKILL GROUPINGS

The main text presents the mappings for national groupings and neighbour groups. Here we present the stochastic kernel mappings of Europe-relative unemployment rates to two other likely groupings, namely the initial sectoral specialisation of regions.

3.1. Sectoral Groupings

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. Exactly this sort of reasoning explains the inclusion of unemployment as a component of the EU's eligibility criterion for objective 2 status. Could the similarity across neighbours be a result of regions with heavy industrial or primary employment being contiguous?

Figure A2: Stochastic kernel mapping for sectoral composition

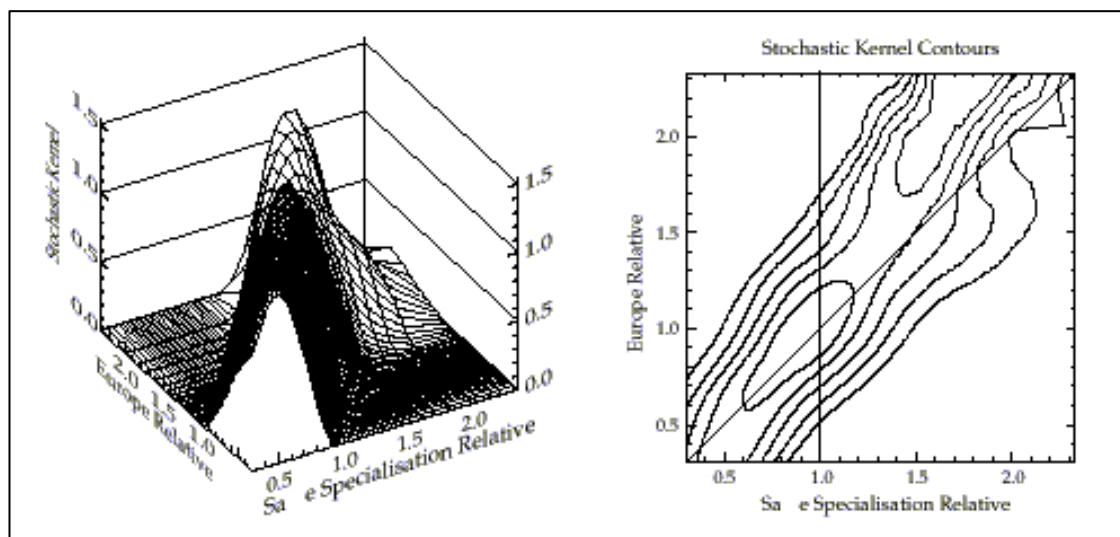


Figure A2 suggests that the answer is no. This figure shows the stochastic kernel mapping Europe relative unemployment rates to same specialisation relative unemployment rates. To do this, we group regions by the sector (agriculture and other primary sectors, manufacturing, or services) in which the initial share of regional employment was highest, relative to the average European share. The concentration of mass on the diagonal of figure 6 suggests that regions with similar initial specialisation have seen very different outcomes. This is probably due to the fact that the largest drop in agricultural and manufacturing employment had already taken place before the beginning of the period we consider. In the 15 years between 1971 and 1986 the share of manufacturing in European employment fell from 41 % to 33 %, while the share of services rose from 45 % to 59 %. In the next ten years to 1996, the share of manufacturing only fell by another three percentage points to 30 %, while that of

services rose to 65 %. Spatial concentrations of declining sectors are not the key component driving the neighbour effect.

3.2. Grouping by similar skill composition

There has been some discussion as to whether changes in the patterns of relative labour demand and supply in Europe have resulted in a rise in unemployment rates for the low skilled relative to the high skilled (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. Apparently this possible link between skills and unemployment underlies the EU's inclusion of objective 3 as a regional policy dimension. Have regions with a large proportion of low skill workers experienced similarly high unemployment outcomes, while regions with a small proportion of low skill workers experienced similarly low unemployment outcomes?

Figure A3: Stochastic kernel mapping for skill composition

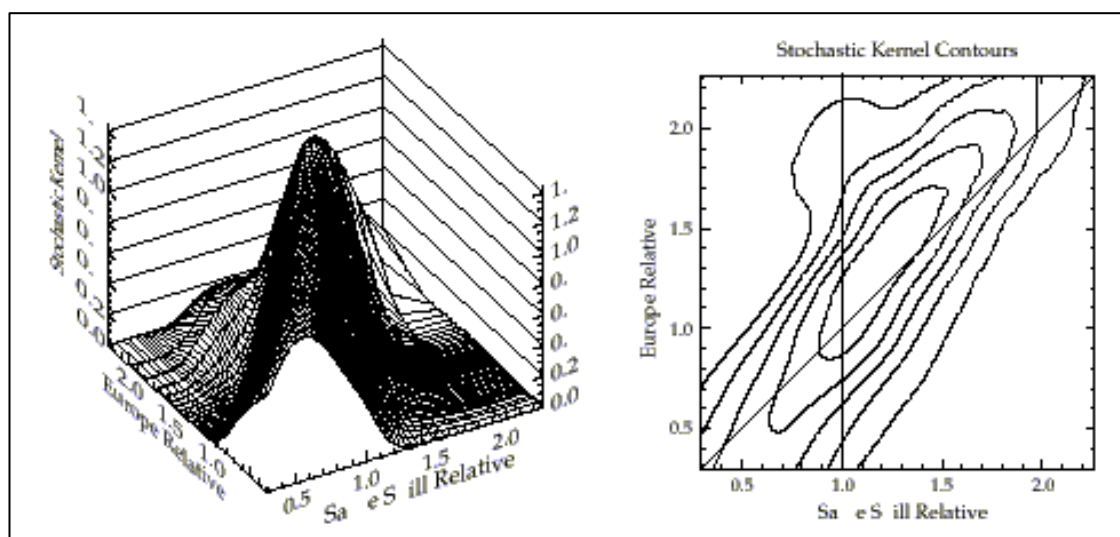


Figure A3 plots the stochastic kernel mapping Europe relative to same skill relative unemployment rates. We construct the kernel using nine groups of regions that have a similar percentage of adult population with less than upper secondary education (divided into equally spaced intervals between 0% and 90%). The concentration of mass on the diagonal reflects the fact that the distribution of unemployment rates across each of our nine groups of regions with similar skill composition is not dissimilar from the distribution of unemployment rates across all European regions. Spatial concentrations of similarly skilled regions are clearly not the key component driving the neighbours effect either. Discretisation 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 1 , allow us estimate the relative mass in different areas of the kernels without having to integrate explicitly. The matrices give two additional pieces of information. The second column gives the ranges that we have used to discretise the distribution. The first column gives n , the number of regions that have Europe relative unemployment in any given range.