View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

Unemployment Rates at the Regional and National Levels of the European Union: An Integrated Analysis

Annette S. Zeilstra and J. Paul Elhorst[1]

Department of Economics, University of Groningen, The Netherlands.
Correspondence to: P.O. Box 800, 9700 AV Groningen, The Netherlands.
Tel: +31-50-3638344, Fax: +31-50-3637337, E-mail: a.s.zeilstra@rug.nl.
Tel: +31-50-3633893, Fax: +31-50-3637337, E-mail: j.p.elhorst@rug.nl,
.

Paper presented at the 46th European Regional Science Association (ERSA)
30 August - 3 September 2006, Volos

Abstract

This study investigates the causes of variation in regional unemployment rates in a cross-country perspective. The explanatory variables consist of both regional-level and macro-level variables. An appropriate econometric model of random coefficients for the former and fixed coefficients for the latter variables is developed, further taking into account that observations may be correlated over time and over space and that some of the explanatory variables are not strictly exogenous. On the basis of this model a regional unemployment rate equation is estimated, using data of 143 regions across 11 EU countries derived from Eurostat, 1983-1997, and national data on labour market institutions predominantly derived from the OECD.

Contents

1. Introduction
2. National and regional determinants of the unemployment rate
3. An econometric model of the regional unemployment rate in a cross-country perspective
4. Preliminary results
5. Conclusions and further research
Appendices

[1] We would like to thank Jan Oosterhaven and Erik Meijer for valuable comments.

1. INTRODUCTION

Involuntary unemployment has been one of the major policy concerns in the European Union. Unemployment levels vary widely between countries in the European Union, but the variation in unemployment rates between regions within a country is even larger (OECD, 2005b). Moreover, though unemployment differences between countries in Europe have decreased markedly in the past decade, regional unemployment differences within countries have remained stable. In some European countries regional unemployment differences have even increased (OECD, 2005a). A graphical illustration of the large dispersion of regional unemployment rates around the national average can be found in appendix B.

Regional unemployment rates are determined by both regional and national factors. On the one hand regional unemployment rates are strongly influenced by variables such as the regional composition of the labour force and the regional structure of the labour market. On the other hand, regional labour markets operate in an environment that is heavily influenced by national labour market institutions, such as the level of centralisation of wage bargaining. In the United Kingdom wage bargaining mostly takes place at the company or plant level. By contrast, in a large part of Continental Europe wage bargaining is highly centralised. Sectoral wages are determined at the national level rather than at the company or regional level. In Austria, Belgium, Finland, Germany, Greece, Ireland, The Netherlands, Norway, Spain, and Sweden wage bargaining predominantly takes place at either the sectoral or national level. In Italy and France wages are negotiated at both the industry and company level.[2] Moreover, the sectoral wages as negotiated by the union and the employer federation at the national level apply to a large part of the people working in that sector. The centralisation of wage bargaining reduces the scope for regional wage differences.[3] Other national variables that influence regional unemployment are e.g. employment protection and the tax-benefit system.

Most empirical regional unemployment studies study the effect of regional variables on unemployment for a single country.[4] Since these studies are single country studies, the effect of different national level variables is not measured. Note that though regional unemployment studies that consider multiple countries exist, they usually do not include national labour market

[2] OECD (2004, Table 3.5) and European Commission (2003, Table 27).
[3] Both the European Commission (2003) and the OECD (1994, 2004, 2005a) advocate decentralisation of wage bargaining so that wages can adjust more easily to local labour market conditions.
[4] For an overview see Elhorst (2003).

institutions as explanatory variables[5]. Studies that do investigate the effect of labour market institutions on unemployment for several countries usually do not study regional unemployment rates, but study national unemployment.[6]

By developing a regional econometric model in a cross-country perspective, we are able to model the two components simultaneously and to differentiate between the effects of regional and national characteristics.

The remainder of this paper is structured as follows. In section 2 the regional and national determinants of the (regional) unemployment rate are discussed. In section 3 we delineate the econometric model in which the variable to be explained is the unemployment rate at regional level and both regional-level and national-level variables serve as explanatory variables. Because we are interested in the regional distribution of the unemployment rate within countries as well as among them, we employ a mixed model with random coefficients for the regional-level variables and fixed coefficients for the national-level variables. Additionally, the common problems of space-time data, serial dependence, spatial dependence and heteroskedasticity are accounted for. In section 4 we present and discuss the results of our empirical analysis. Section 5 concludes.

2. NATIONAL AND REGIONAL DETERMINANTS OF THE UNEMPLOYMENT RATE

Regional unemployment rates are determined by both labour demand and supply variables and wage-setting institutions. These variables consist of both regional and national level variables. In section 2.1 we discuss the national variables that influence regional unemployment and in section 2.2 we discuss regional variables that influence regional unemployment.

2.1 NATIONAL VARIABLES

Wage bargaining level

According to Calmfors and Driffill (1988) the relation between the degree of centralisation of wage bargaining and wages is hump-shaped. Both low and high levels of centralisation lead to moderation of wage demands. In the former case wage moderation is caused by market forces. In

[5] See e.g. Baddeley et. al. (1998) and Taylor and Bradley (1997).

[6] Two recent surveys of the literature on the effect of labour market institutions on aggregate performance are Blau and Kahn (1999) and Nickell and Layard (1999).

the latter case moderation is caused by the internalisation of externalities. If the union and employer federation bargain at the national level, it is likely that they account for the effect of higher wages on the general price level and on unemployment. By contrast, if wage bargaining takes place at the industry level, the union and employer federation may not take into account that higher wages and lower employment in the own sector make it more difficult for workers from other sectors to obtain a job as well.[7]

Assuming a downward sloping labour demand curve, the hump-shaped relation between centralisation and wages corresponds to a U-shaped relation between centralisation and unemployment. As a result, we would expect relatively low unemployment rates for low and high levels of centralisation compared to the intermediate case.[8]

Union density

Since higher union density strengthens the bargaining position of the union, we expect higher wage demands, lower employment, and higher unemployment if union density increases. Note that the effect of union density on wage demands may depend on the degree of centralisation. In countries with highly centralised wage bargaining we would expect less pronounced effects of union density on wage demands and unemployment than in countries with firm or plant level wage bargaining. Belot and Van Ours (2004) have studied the interaction effect of several institutions for 18 countries for the period 1960-1994 and only find a significant positive effect of union density on employment for countries with firm or plant level bargaining.

Unemployment benefits

Unemployment benefits influence unemployment in two ways. First, higher unemployment benefits may have a positive influence on wage demands, because the fear for unemployment decreases. Second, higher unemployment benefits make participation in the labour market more attractive and may result in higher labour participation. Both effects may lead to higher unemployment.

[7] A more elaborate overview of negative wage externalities is given by Calmfors (1993).

[8] Note that the effect of intermediate level wage bargaining on unemployment can be offset by the extent to which bargaining is coordinated.

Tax wedge

Although an increase in payroll taxes will have no long run effects on unemployment in a perfectly competitive world, it might raise unemployment if unions have bargaining power and product markets are not perfectly competitive (Scarpetta, 1996). Furthermore, even if capital is perfectly mobile between countries and labour is not, the tax burden of a rise in payroll tax will fall on the employer, if a worker is already receiving the minimum wage (Nickell, 1997). This in turn may lead to lower employment and higher unemployment.

Employment protection

On the one hand, stronger employment protection reduces the inflow into unemployment, because firing an employee becomes more difficult. On the other hand, stronger employment protection makes an employer more cautious about hiring a new employee and reduces the flow out of unemployment.[9] In sum one would expect lower short term unemployment and higher long term unemployment if employment protection increases. The overall effect on unemployment can go either way, though it could be quite small, because the short run and the long run effect tend to cancel each other out[10].

Changes in inflation

If the actual price level exceeds the price level as expected during the wage bargaining process, real wages are relatively smaller and as a result employment may be higher and unemployment lower. The opposite holds if the actual price level is lower than the expected price level. We incorporate price surprises (actual price level minus the expected price level) by using the change in inflation as a proxy ($\Delta^2 p$).

Another national level variable is active labour market policy. However, we do not include this variable in our analysis, because the causality between active labour market expenditures and unemployment runs both ways.

---

[9] See e.g. OECD (2004), Table 2.2.

[10] Nickell (1997, Table 6) finds a negative effect of employment protection on short term unemployment, a positive effect of employment protection on long term unemployment and a small negative effect of employment protection on total unemployment.

2.2 REGIONAL VARIABLES

An extensive overview of regional variables that are used in empirical studies to explain regional unemployment differentials is provided by Elhorst (2003a).[11] We do not include all variables that are discussed by Elhorst (2003a) in our analysis for two reasons. First, comparable data on some variables, such as regional commuting and migration, are not available for all regions and countries in our sample. Second, we do not want to include some variables, such as the shares of different industries in employment, in our analysis. The reason why we do not want to include the latter is that the share of, for example, services in total employment may increase for two reasons. First, the share of services may increase because employment in the service sector has increased. Second, the share of services may increase due to a decline in employment in the manufacturing sector. Both phenomena lead to a higher share of services in total employment, but the expected effect on regional unemployment is quite different. Below we discuss the regional variables that we use in our analysis.

Demographic composition of the population

We include two variables to account for regional differences in the demographic composition of the population, the percentage of the working age population aged between 15 and 24 and the percentage of the working age population aged between 55 and 64. We expect a positive effect of the former on unemployment, due to relatively high youth unemployment.[12]. The effect of the percentage of the working age population aged between 55 and 64 on unemployment can go both ways. On the one hand, older workers may use the unemployment benefit system as an early retirement scheme, thus raising unemployment. On the other hand, older workers may use actual early retirement schemes and exit the labour force, which may have a negative effect on unemployment.

Educational attainment

We have included the percentage of individuals having medium or higher education to account for differences in the educational attainment of the labour force. We expect that education has a

[11] In a review of 41 empirical studies, Elhorst (2003) gives a detailed overview of theoretical and empirical explanations of regional unemployment differentials used in the applied literature.

[12] Note that the upward effect on unemployment will be tempered if relatively many young people are enrolled in full-time education.

negative effect on unemployment for three reasons. First, better educated individuals are more likely to possess skills that are demanded in an economy with continued technological progress. Second, better educated individuals are expected to conduct more efficient searches. Third, better educated individuals are less likely to be fired and so exhibit more stable patterns of employment.

Economic structure

We have included two variables that influence the demand for labour in a region in order to capture part of the economic structure of a region, productivity and employment growth. We expect that the economic structure of a region is more favourable and unemployment lower if the productivity is higher. The same holds for the average employment growth over the past two years, higher employment growth is associated with a more favourable economic structure and less unemployment.

Labour supply

Labour participation is included to incorporate the effect of labour supply on unemployment. The effect of labour participation on unemployment that first comes to mind is a positive effect, because according to the accounting identity higher labour supply *ceteris paribus* leads to higher unemployment. However, the effect of labour participation on unemployment can be negative as well, because higher participation may lead to more local jobs. Because labour participation also depends on unemployment this variable is not strictly exogenous, and is estimated with instrumental variables.

Lagged unemployment

Finally, we include the lagged unemployment rate to deal with serial dependence.

3. AN ECONOMETRIC MODEL OF THE REGIONAL UNEMPLOYMENT RATE IN A CROSS-COUNTRY PERSPECTIVE

The data used for the empirical analysis presented later in this paper are regional data across multiple countries. These data may therefore be said to be grouped at two different levels. Regions are so-called level 1 units grouped within countries that are the level 2 units. According to Goldstein (1995, pp. 1-2), the existence of such groupings should not be ignored in the empirical

analysis. Regional unemployment rates are affected by neighbourhood and contextual effects, that is by regional-level and national-level variables respectively. Consequently, working at a single level, estimating a macroeconomic equation based on macro data or a regional economic equation based on regional data, is likely to lead to a distorted representation of reality. A single-level model assumes that the data do not follow a hierarchical structure, that all the relevant variation is at one scale. A two-level model takes the hierarchical structure between regions and countries into account by modelling the variation at both levels, and allows all the regions that belong to a particular country to be more alike than a random sample. Moreover, within each country in the hierarchy, different relationships are allowed around the overall relationship for all regions in all countries.

In a two-level model the parameters with respect to the regional-level variables may vary from one country to another, while the parameters with respect to the national-level variables are the same for all countries

$$U_{crt} = \beta_c^{'} X_{crt} + \alpha' Z_{ct} + \varepsilon_{crt}, \tag{1a}$$

$$\beta_c = \beta + v_c, \tag{1b}$$

$$E(\varepsilon_{crt}) = 0, \ Var(\varepsilon_{crt}) = \sigma_c^2, \tag{1c}$$

$$E(v_c) = 0, \ Var(v_c) = V, \tag{1d}$$

where c (=1,…,N) refers to a country, r refers to a region (=1,…,R_c with R_c the number of regions in country c), and t (=1,…T) refers to a given time period. U_{crt} is unemployment rate in region r of country c at time t, X_{crt} is a vector of explanatory variables measured in region r of country c at time t, and Z_{ct} is a vector of explanatory variables in region r but only observed at the national level of country c at time t, since these variables do not differ between regions within countries.[13] ε_{crt} is a heteroskedastic disturbance term with variance σ_c^2, which is assumed to be different for different countries. β_c represents a vector of random response parameters and α a vector of fixed response parameters in the regression equation. The β_c applying to a particular country is the outcome of a random process with common-mean-

[13] For reasons to be explained later, endogenous explanatory variables will also be classified among Z_{ct}, even when they are observable at regional level.

coefficient vector β and covariance matrix *V*. When the vectors β and β_c (c=1,…,N) are of size *K*, *V* is of size *K*×*K*.

This model belongs to the class of mixed linear models. Frees (2004) gives a general and detailed overview of the mathematical and statistical fundamentals of this class of models, as well as substantive applications across the social sciences. These mixed models are also known as two-level or multilevel models, the difference being that in this type of model as well as in many regional economic applications the error term is assumed to be homoskedastic, $Var(\varepsilon_{crt}) = \sigma^2$ for every country *c*.[14] We reject the assumption of a homoskedastic error term and consider a generalization of the two-level model because we believe a heteroskedastic error term, $\sigma_c^2 \neq \sigma^2$ for different countries *c*, is more realistic (see Frees, 2004, pp. 45-52) .

Two other problems that frequently occur when using space-time data are serial dependence between the observations on each spatial unit over time and spatial dependence between the observations on the spatial units at each point in time. Both problems have received considerable attention successively in the time-series and spatial cross-section econometrics literature, but not so much in combination (Elhorst, 2001).[15] To deal with serial dependence, the unemployment rate U_{crt} will be regressed on its serial lagged value U_{crt-1}.

To deal with spatial dependence, the econometric model (1) can be estimated using region-by-year cell means, assuming that there is no correlation among the unobserved determinants of unemployment rates across regions. As we pointed out in the introduction to this paper, this is not very likely. Not only people from the same region, but also people from different regions may share the same unobservable characteristics. Variation in unemployment rates among different parts of the country might be explained by unobserved sub-national variables, such as differences in e.g. culture. These unobserved circumstances may have different regional effects, as a result of which regional error terms correlate and their variance is not constant across the country. The best-known model focusing on spatial dependence starts with a first-order spatial autoregressive process generating the error terms, $\varepsilon_{ct} = \delta_c W_c \varepsilon_{ct} + u_{ct}$, where the error terms ε_{ct} and u_{ct} are written in vector form for each cross-section of regions in country

[14] Applications based on the multilevel approach in regional economic research are Jones (1991), Ward and Dale (1992), Gould and Fieldhouse (1997) and McCall (1998).
[15] One exception but in a different setting is discussed in Frees (2004, pp. 283-285).

c at time t, $u_{ct} \sim N(0, \sigma^2 I_{R_c})$ (see Anselin, 1988). W_c (c=1,…,N) is a $R_c \times R_c$ non-negative matrix with zeros on the diagonal describing the spatial arrangement of the regions in country c. δ_c is called the spatial autocorrelation coefficient and is assumed to be fixed but different for different countries. Consequently, the covariance matrix of the error terms in equation (1c) changes to

$$E(\varepsilon_{ct}) = 0, \; Var(\varepsilon_{ct}) = \sigma_c^2 [(I_{R_c} - \delta_c W_c)'(I_{R_c} - \delta_c W_c)]^{-1} = \sigma_c^2 \Omega_c . \qquad (1c')$$

A final problem is that one or more of the explanatory variables might be endogenous. Endogeneity requires instrumental variable methods like two-stage least squares (2SLS) to obtain consistent parameter estimates. Balestra and Negassi (1992), among others, have pointed out that the assumption of a random element in the coefficients of endogenous explanatory variables raises intractable difficulties at the level of identification and estimation. For this reason, coefficients of endogenous variables are assumed to be fixed and part of Z, even though they vary across regions within countries.[16] This also applies to U_{crt-1}. This variable should be treated as an endogenous explanatory variable since it is correlated with the random error terms v_c.

Although Frees (2004) discusses heteroskedasticity, serial dependence, spatial dependence, endogenous explanatory variables, as well as different combinations, this paper is among the first to consider these four problems within one framework. For a better understanding of the estimation method described below, the reader may nonetheless consult Frees' book, especially his description of the mixed linear model and the GLS estimation method (pp. 92-96).

Let us assume that the data are first sorted by country, and then for each country first by time and then by region. Thus the first R_1 observations represent all the regions in country 1 at time $t=1$, the second R_1 observations all the regions in country 1 at time $t=2$, and so on. Let U_{ct}, X_{ct} and Z_{ct} denote the observations and ε_{ct} the error terms of a cross-section of regions stacked within a particular country at a particular point in time, and let U_c, X_c and Z_c denote all the observations and ε_c all the error terms stacked within a particular country. The full model may

[16] If these variables are observed at regional level, the explanatory variables Z in equation (1a) should formally not only be indexed by the subscripts c and t, but also by the subscript r. We omitted this detail to avoid confusion with previous descriptions of similar models in the literature.

then be expressed as

$$\begin{bmatrix} U_1 \\ U_2 \\ . \\ U_N \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ . \\ X_N \end{bmatrix} \beta + \begin{bmatrix} Z_1 \\ Z_2 \\ . \\ Z_N \end{bmatrix} \alpha + \begin{bmatrix} X_1 & 0 & 0 & 0 \\ 0 & X_2 & 0 & 0 \\ . & . & . & . \\ 0 & 0 & 0 & X_N \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ . \\ v_N \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ \varepsilon_N \end{bmatrix} = \mathrm{X}\beta + Z\alpha + \mathrm{diag}(\mathrm{X}_1,...,X_N) \times v + \varepsilon. \quad (2a)$$

The covariance matrix Φ of the composite disturbance term $diag(X_1,...,X_N) \times v + \varepsilon$, if we further take into account that ε_{ct} follows a spatial autoregressive process (1c'), is block-diagonal, with the c^{th} diagonal block given by

$$\Phi_c = X_c V X_c^{'} + \sigma_c^2 (\Omega_c \otimes I_T). \quad (2b)$$

The ML estimator of the response parameters is known to be equivalent to the GLS estimator (Lindstrom and Bates, 1988) and, if we use Hsiao and Tahmiscioglu (1997), to be equal to

$$\begin{bmatrix} \beta \\ \alpha \end{bmatrix}_{GLS} = \left[\sum_{c=1}^{N} [X_c \; Z_c]' \Phi_c^{-1} [X_c \; Z_c] \right]^{-1} \left[\sum_{c=1}^{N} [X_c \; Z_c]' \Phi_c^{-1} L_c \right], \quad (3)$$

where the inverse of Φ_c is

$$\Phi_c^{-1} = \frac{(\Omega_c^{-1} \otimes I_T)}{\sigma_c^2} - \frac{X_c (\Omega_c^{-1} \otimes I_T)}{\sigma_c^2} [V^{-1} + \frac{X_c^{'} (\Omega_c^{-1} \otimes I_T) X_c}{\sigma_c^2}]^{-1} \frac{X_c^{'} (\Omega_c^{-1} \otimes I_T)}{\sigma_c^2}, \quad (4a)$$

$$\Omega_c^{-1} = (I_{R_c} - \delta_c W_c)' (I_{R_c} - \delta_c W_c). \quad (4b)$$

A problem is that this GLS estimator of the response parameters contains unknown parameters that must also be estimated: δ, σ^2 and the elements of *V*. Importantly, an asymptotically efficient feasible GLS estimator of the response parameters does not require that we have the ML estimator of these unknown parameters, only a consistent one (Greene, 2003). Moreover, ML estimation of these parameters, although possible, is laborious (Elhorst, 2003b). To obtain the feasible GLS estimator, we combine the standard procedures of estimating a model with instrumental variables, random coefficients, heteroskedasticity and spatial autocorrelation into one framework.

To deal with endogeneity, we follow Bowden and Turkington (1984, ch. 3) and Amemiya (1985, pp. 240-241). They have pointed out that the GLS analog instrumental variables estimator

of b in a linear model $Y = Xb + \varepsilon,$ with $E(\varepsilon\varepsilon') = \sigma^2\Phi$ and one or more endogenous X variables, equals

$$b = (\hat{X}'\Phi^{-1}\hat{X})^{-1}\hat{X}'\Phi^{-1}Y, \quad \text{with} \quad \hat{X} = P(P'\Phi^{-1}P)^{-1}P\Phi^{-1}X, \tag{5}$$

where P denotes a matrix of instrumental variables (including the exogenous variables from X). In our model, $b = (\beta_{\cdot},..., \beta_N, \alpha)'$ and Φ is block diagonal, with the inverse of the c^{th} diagonal block given by (4).

To deal with random coefficients, we follow Swamy (1970). Swamy's estimation procedure consists of three steps: (i) Estimate the model assuming that all the response parameters are fixed and that the coefficients of the X variables are different for different countries. This can be carried out by regressing U on $diag(X_1,\ldots,X_N)$ and Z. We use the mnemonic F to refer to these estimates; (ii) Estimate V by (see also Hsiao and Tahmiscioglu, 1997)

$$V = \frac{1}{N-1}\sum_{c=1}^{N}(\hat{\beta}_c^F - \frac{1}{N}\sum_{c=1}^{N}\hat{\beta}_c^F)(\hat{\beta}_c^F - \frac{1}{N}\sum_{c=1}^{N}\hat{\beta}_c^F). \tag{6}$$

(iii) Estimate the common-mean coefficient vector β and the other response coefficient parameters, given V, using (4).

To deal with heteroskedasticity and spatial error autocorrelation, step (i) of Swamy's procedure should also include the estimation of σ_c^2 and δ_c (c=1,…,N), step (ii) remains unchanged, while step (iii) should be carried out not only given V, but also given $\Omega_c = \Omega(\sigma_c^2, \delta_c)$. The extension of step (i) can be carried out along an iterative two-stage procedure taken from Anselin (1988): (Stage 1) Given the initial estimate of the response parameters α^F and β^F, estimate σ_c^2 by $\sigma_c^2 = \sum_{t=1}^{T} e_{ct}' e_{ct} / R_c T$ and estimate δ_c (c=1,…,N), given α, β_c and σ_c^2, by maximizing the (concentrated) log-likelihood function $-\frac{R_c T}{2} ln(2\pi\sigma_c^2) + T\sum_{j=1}^{R_c} ln(1 - \delta_c\omega_{cj}) - \frac{1}{2\sigma_c^2}\sum_{t=1}^{T} e_{ct}' e_{ct}$, where ω_{cj} denote the characteristic roots of the spatial weight matrix W_c of country c; (Stage 2) Re-estimate the response parameters α and β_c, given the estimates of σ_c^2 and δ_c. Note that

$X_c^{'}\Omega_c^{-1}X_c/\sigma_c^2 = X_c^{*'}X_c^{*}$ when applying the following transformations to the data $X_{ct}^{*} = (1/\sigma_c^2 \times \Omega)^{1/2} X_{ct} = 1/\sigma_c \times (I_{R_c} - \delta_c W_c)X_{ct}$

$= 1/\sigma_c \times X_{ct} - \delta_c/\sigma_c \times W_c X_{ct}$. Similar transformations apply to Z_{ct} and U_{ct}. This two-stage procedure within Swamy's estimation procedure must be repeated until convergence occurs. Given these transformations, step (iii) reduces to

$$\begin{bmatrix}\beta \\ \alpha\end{bmatrix}_{GLS} = \left[\sum_{c=1}^{N}\sum_{t=1}^{T} S_{ct}^{*'}S_{ct}^{*} - \sum_{c=1}^{N}\sum_{t=1}^{T} S_{ct}^{*'}X_{ct}^{*}[V^{-1} + \sum_{c'=1}^{N}\sum_{t'=1}^{T} X_{c't'}^{*'}X_{c't'}^{*}]^{-1}X_{ct}^{*'}S_{ct}^{*}\right]^{-1} \times \quad (7)$$

$$\left[\sum_{c=1}^{N}\sum_{t=1}^{T} S_{ct}^{*'}L_{ct}^{*} - \sum_{c=1}^{N}\sum_{t=1}^{T} S_{ct}^{*'}X_{ct}^{*}[V^{-1} + \sum_{c'=1}^{N}\sum_{t'=1}^{T} X_{c't'}^{*'}X_{c't'}^{*}]^{-1}X_{ct}^{*'}L_{ct}^{*}\right]$$

where $S_{ct}^{*} = [X_{ct}^{*} Z_{ct}^{*}]$ and the matrix which is inverted at the right-hand side of (7) is the covariance matrix of the feasible GLS estimator of the common-mean-coefficient vector β and α.

4. PRELIMINARY RESULTS

Data

Our primary data source is Eurostat's regional database. We have used the regional division of Eurostat on the NUTS2 level, with the exception of the UK since Eurostat only provides NUTS 1 data for the UK[17]. The data set covers the period 1983-1997. The total number of observations is 1549 divided over 143 regions across 11 countries. It should be noted that the data set is not complete, mainly because some countries became member states of the EU after 1983 and in some regions Eurostat started data registration after 1983. Data registration in the former East Germany started in 1991. Due to large differences in the data, East and West Germany are treated as two different level 2 units. Moreover, because the OECD does not collect data on Greece and Luxembourg for some of the institutional variables we use, Greece and Luxembourg are excluded from our analysis.

More detailed information about the data can be found in the appendices. Appendix A contains a more detailed description of the data. Appendix B shows the variation of regional unemployment rates around the national average. Appendix C contains the correlation coefficients of the variables.

[17] During the 1980s and the first half of the 1990s.

Preliminary estimation results

Preliminary estimation results are shown in table 1, Appendix D. The coefficient estimates in the first column reflect short-term effects. Long-term effects can be obtained from the short-term estimated coefficients by multiplying the latter by $1/(1-\hat{\tau})$, where $\hat{\tau}$ is the coefficient estimate of lagged unemployment rate. The last column reports the standard deviation of the variables with random coefficients based on the V estimate.

The coefficients of the regional variables have the expected sign, though the effect of the percentage of working age individuals aged between 15 and 24 is insignificant. The percentage of working age individuals aged between 55 and 64 has a negative effect on unemployment, which might be caused by early retirement from individuals in this age group. As expected, educational attainment, productivity and employment growth have a negative effect on unemployment. Labour participation also has a negative effect on unemployment. Note that there is considerable heterogeneity of the coefficients across countries. As can be seen from table 1, the standard deviation of the variables with random coefficients exceeds the coefficient estimate of these variables.

The national variables unemployment benefits and tax wedge have the more or less expected positive effect on unemployment, whereas the change in inflation has the expected negative effect on unemployment. We did not include the centralisation index directly as an explanatory variable, because of the little variation in this index given our relatively short time period. Instead, we use the centralisation index to study the effect of union density for different levels of centralisation. We expected to find a positive effect of union density on unemployment for low bargaining levels and insignificant effects for higher wage bargaining levels. By contrast, we find insignificant effects of union density for low bargaining levels, whereas we find a significant negative effect on unemployment for high bargaining levels. An explanation for this counterintuitive result could be that high levels of union density in a situation with centralised wage bargaining lead to lower wage demands, because the wage negotiations are more broadly based. Finally, employment protection has a negative effect on unemployment.

5. CONCLUSIONS AND FURTHER RESEARCH

The model developed in this paper is a mixed model of random coefficients for regional-level

variables and fixed coefficients for national-level variables. A striking result is the considerable heterogeneity of the coefficients across countries, this indicates that regional unemployment rates in the EU are not determined by a common structure.

In future research we may also include interactions between other institutional variables (e.g. interaction between the tax wedge and the benefit replacement rate). Studies by Belot and Van Ours (2001 and 2004) indicate that interactions between institutions may play an important role in explaining differences in unemployment.

REFERENCES

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

Amemiya, T. (1985) *Advanced Econometrics*. Basic Blackwell, Oxford.

Baddeley, M., R. Martin, and P. Tyler (1998) European regional unemployment disparities: convergence or persistence?, *European Urban and Regional Studies*, 5: 195-215.

Balestra, P. and S. Negassi (1992) A Random Coefficient Simultaneous Equation System with an Application to Direct Foreign Investment by French Firms. *Empirical Economics* 17: 205-220.

Belot, M,. and J.C. van Ours (2004) Does the recent success of some OECD countries in lowering their unemployment rates lie in the clever design of their labor market reforms?, *Oxford Economic Papers*, 56: 621-642.

Belot, M. and J.C. van Ours (2001), Unemployment and labor market institutions: An empirical analysis, *Journal of the Japanese and International Economies*, 15: 403-418.

Blau, F. and L. Kahn (1999) Institutions and laws in the labor market, In: Ashenfelter, O. and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3A, Elsevier Science Publishers BV, Amsterdam, 1399-1461.

Bowen, W.G. and T.A. Finegan (1969) *The Economics of Labor Force Participation*. University Press, Princeton.

Calmfors, L. (1993) Centralisation of wage bargaining and macroeconomic performance: A survey, *OECD Economic Studies*, 21:161-191.

Calmfors, L. and J. Driffill (1988) Bargaining structure, corporatism and macroeconomic performance, *Economic Policy*, 6:13-61.

Elhorst, J.P. (2003a) The mystery of regional unemployment differentials: theoretical and

empirical explanations, *Journal of Economic Surveys*, 17(5): 709-748.

Elhorst, J.P. (2003b) Specification and Estimation of Spatial Panel Data Models. *International Regional Science Review* 26: 244-268.

Elhorst, J.P. (2001) Dynamic Models in Space and Time. *Geographical Analysis* 33: 119-140.

European Commission (2003) *Employment in Europe 2003*. Office for official publications of the European Communities, Luxembourg.

Frees, E.W. (2004) *Longitudinal and Panel Data*. Cambridge University Press, Cambridge.

Goldstein, H. (1995) Multilevel Statistical Models, second edition. Arnold/Oxford University Press, New York.

Gould, M.I. and E. Fieldhouse (1997) Using the 1991 Census SAR in a Multilevel Analysis of Male Unemployment. *Environment and Planning A* 29: 611-628.

Greene, W.H. (2003) *Econometric Analysis*, fifth edition. Prentice Hall, New Jersey.

Hsiao, C. and A.K. Tahmiscioglu (1997) A Panel Analysis of Liquidity Constraints and Firm Investments. *Journal of the American Statistical Association* 92: 455-465.

Jones K. (1991) Multi-Level Models for Geographical Research. *Concepts and Techniques in Modern Geography* 54, University of East Anglia, Norwich.

Lindstrom, M.J. and D.M. Bates (1988) Newton-Raphson and EM Algorithms for Linear Mixed-Effects Model for Repeated-Measures Data. *Journal of the American Statistical Association* 83: 1014-1022.

McCall, L. (1998) Spatial Routes to Gender Wage (In)Equality: Regional Restructuring and Wage Differentials by Gender and Education. *Economic Geography* 74: 379-404.

Nickell, S. (1997), Unemployment and labor market rigidities: Europe versus North America, *Journal of Economic Perspectives*, 11(3): 55-74.

Nickell, S. and R. Layard (1999) Labour market institutions and economic performance, In: Ashenfelter, O. and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3C, Elsevier Science Publishers BV, Amsterdam, 3029-3084.

OECD (1994) *The OECD Jobs Study: Facts, analysis, strategies.* OECD, Paris.

OECD (2004) *Employment Outlook*. OECD, Paris.

OECD (2005a) *Employment Outlook*. OECD, Paris.

OECD (2005b) *OECD Regions at a Glance. OECD*, Paris.

Scarpetta, S. (1996), Assessing the role of labour market policies and institutional settings on

unemployment: A cross-country perspective, *OECD Economic Studies, 26: 43-98.*

Schürmann, C. and A. Talaat (2000) *Towards a European Peripherality Index: Final Report*, Institut für Raumplanung, Universität Dortmund, Berichte aus dem Institut für Raumplanung 53, Dortmund.

Swamy, P.A.V.B. (1970) Efficient Inference in a Random Coefficient Regression Model. *Econometrica* 38: 311-323.

Taylor, J. and S. Bradley (1997) Unemployment in Europe: A comparative analysis of regional disparities in Germany, Italy and the UK, *Kyklos*, 50: 221-245.

Ward, C. and A. Dale (1992) Geographical Variation in Female Labour Force Participation: An Application of Multilevel Modelling. *Regional Studies* 26.3: 243-255.

APPENDIX A: Data

Both regional and national data are used. The primary regional level data were selected from the Eurostat file called 'Regions'. Eurostat uses a classification in NUTS 1, NUTS 2, and NUTS 3 level regions. Where NUTS is an abbreviation of "Nomenclature of Territorial Units for Statistics". The higher the NUTS level the smaller the regions.

Most regional data are available for NUTS 2 level regions, with the exception of the UK for which NUTS1 level data are used, since Eurostat only provides NUTS 1 data for the UK up to 1995.

The NUTS classification by Eurostat changes over time. The NUTS 2 classification of 1999 is used for all countries except for the UK. For the UK the NUTS 1 classification of 1995 is used. Due to extensive changes in the territorial breakdown of the UK the old NUTS 1 classification can not be linked to the 1999 classification and no historical series are available for the UK in the NUTS 1 classification of 1999 (see, Eurostat, 2002).

The regional division we have used consists of a maximum of 143 regions. The data set covers the period 1983-1997, but is incomplete because in some countries or regions Eurostat started data registration after 1983, in part because some countries became member states of the EU after that time (in Spain registration started in 1986; in Portugal registration started in 1986 and in 1988 for the Algarve; in France registration started in 1988 for Provence and Corse; and in the Netherlands registration started in 1988 for Overijssel, Gelderland and Flevoland, while for all other regions data for 1984 and 1986 are lacking). Finally, one region has been left aside (Ceuta y Melilla) and some regions are joined with other regions (Bremen with Lüneburg and Hamburg with Schleswig-Holstein). As a result, the number of observations is different for each year. The total number of observations is 1549, taking into account the effect of the lagged unemployment rate.

Countries included in the analysis (observation period and number of regions within parentheses) are Belgium (1983-1997, 9), Denmark (1983-1997, 1), West Germany (1983-1997, 29), East Germany (1990-1997, 5), Spain (1986-1997, 17), France (1983-1997, 22), Ireland (1983-1997, 1), Italy (1983-1997, 20), the Netherlands (1983,1985,1987-1997, 12), Austria (1995-1997, 9), Portugal (1986-1997, 7) and the UK (1983-1993, 11).

Regional variables

*Unemployment rate (%):*Ratio of people being unemployed (harmonised unemployment) and the active labour population. It relates to persons who are aged at least 15 at a certain point in time. A person is considered unemployed if he/she is without work, currently available for work and seeking work, that is, if he/she has taken specific steps in a specified recent period to seek paid employment or self-employment. Available for the period: 1983-1997. Regional Division: NUTS2. Source: Regions database 1999, tables UN3PERS and LF2ACT, Eurostat. Source data UK and region Brabant (for the period 1983-1992) in Belgium: Regions database 1995, tables CHOM3ABSOLU and EFDT2ACTIV, Eurostat.

Pop15-24 (%): Number of people aged between 15-24 as a percentage of the total working age population aged between 15-64. Available for the period: 1983-1997. Regional Division: NUTS2. Source: Regions database 1999, table LF2POP, Eurostat. Source data UK and region Brabant (for the period 1983-1992) in Belgium: Regions database 1995, table EFDT2POP, Eurostat.

Pop55-64: Number of people aged between 55-64 as a percentage of the total working age population aged between 15-64. Available for the period: 1983-1997. Regional Division: NUTS2. Source: Regions database 1999, table LF2POP, Eurostat. Source data UK and region Brabant (for the period 1983-1992) in Belgium: Regions database 1995, table EFDT2POP, Eurostat.

Education (%): Percentage of people aged 25-59 having medium or higher education. Medium education is defined as upper secondary education, International Standard Classification of Education (ISCED) 3 and higher education is defined as tertiary education, ISCED 5-7. Available for the period: 1992-1997. For the period 1983-1991 the numbers of 1992 were used. Regional Division: NUTS2. Source: Labour Force Survey, Eurostat.

Productivity: GDP divided by total employment. Availability: 1983-1996. Source: Eurostat.
The GDP figures in current local prices are converted to 1990 ECU's with the help of Purchasing

Power Parities, GDP deflators and exchange rates developed by Eurostat.

Employment growth (%): Average growth of employment over the past two years as a percentage of previous employment. Available for the period: 1983-1997. Regional Division: NUTS2. Source: Regions database 1999, table LF2EMP, Eurostat. Source data UK and region Brabant (for the period 1983-1992) in Belgium: Regions database 1995, table EFDT2EMPL Eurostat.

Labour participation: Ratio of the economically active population aged between 15-64 years and the total population aged between 15-64 years. The economically active population is the sum of employed and unemployed people. Available for the period: 1983-1997. Regional Division: NUTS2. Source: Regions database 1999, tables LF2ACT and LF2POP, Eurostat. Source data UK and region Brabant (for the period 1983-1992) in Belgium: Regions database 1995, tables EFDT2ACTIV and EFDT2POP, Eurostat.

Wage: The hourly labour costs (average hourly earnings of manual and non-manual workers in manufacturing) are converted to 1990 ECU's with the help of Purchasing Power Parities and Consumer Price Indices developed by Eurostat. Available for years 1984, 1988, 1992, and 1996. The intermediate years are computed by using national figures, or in case these figures were not available by linear interpolation. Regional Division: NUTS1. Source: Labour Costs Survey, Eurostat.
The hourly labour costs are transformed into net wages by using the total tax wedge. The total tax wedge is given by employees' and employers' social security contributions and personal income tax less transfer payments as percentage of gross labour costs as faced by married couples. Available for the period 1983-1997. Source: OECD.

Spatial weight matrix: The spatial weight matrices used in the estimations are symmetric inverse travel time matrices for passenger traffic. Travel time over land depends on road type, urban and mountainous speed constraints and national car speed limits. Overseas travel time depends on embarkation waiting time and the travel time by ferry. Source: Institut für Raumplanung, see Schürmann and Talaat (2000). The travel time between region A and region B in one direction

sometimes differs from the travel time in the opposite direction. The spatial weight matrix is made symmetric by taking average travel times. The effect of taking average travel time is limited. The difference in travel time between both directions is less than 10 percent in 96 percent of the cases. Alternatively, the difference between the travel time in one direction and the average travel time is less than 10 percent in 99 percent of the cases.

National variables

Centralisation index: Index {0-5}of centralisation of wage bargaining ranging between. The index increases in the level of wage bargaining. The index is 1 in case of firm/plant level bargaining and 5 in case of national wage bargaining. Availability: Five year averages are available for the periods 1980-1984, 1985-1989, 1990-1994 en 1995-2000. Source: OECD. (2004), Table 3.5.

Tax wedge: Employees' and employers' social security contributions and personal income tax less transfer payments as percentage of gross labour costs as faced by married couples. Available for the period 1983, 1985, 1987, 1989, 1991 and 1993-1997. The intermediate years are computed by linear interpolation. Data for France are only available starting 1993, for the period 1983-1992 the numbers of 1993 are used. Source: OECD.

*Unemp. Benefits (%):*Ratio between the unemployment benefit and the median wage. This ratio presents the average of the unemployment benefit replacement rates over five years for two earnings levels, three family situations. Available for the period 1983-1997 (only odd years are available, even years are calculated by linear interpolation). Source: OECD.

Union density|cen≤ 2 (%): The number of union members as a percentage of the number of employees, in countries with a centralisation index of 2 or less. Available for the period 1983-1997. Source: OECD Labour Market Statistic Database.

Union density|2<cen<4 (%): The number of union members as a percentage of the number of employees, in countries with a centralisation index of between 2 and 4. Available for the period

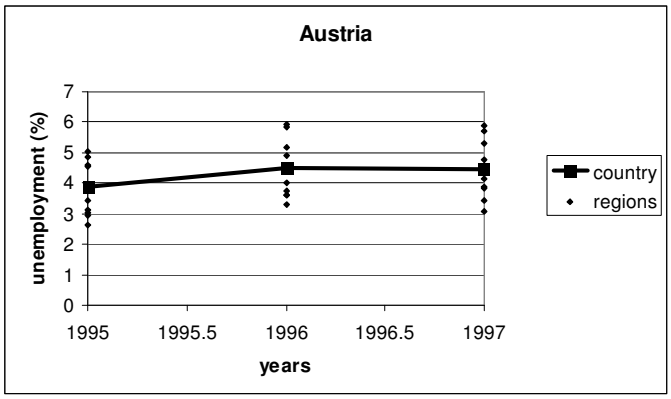
1983-1997. Source: OECD Labour Market Statistic Database.

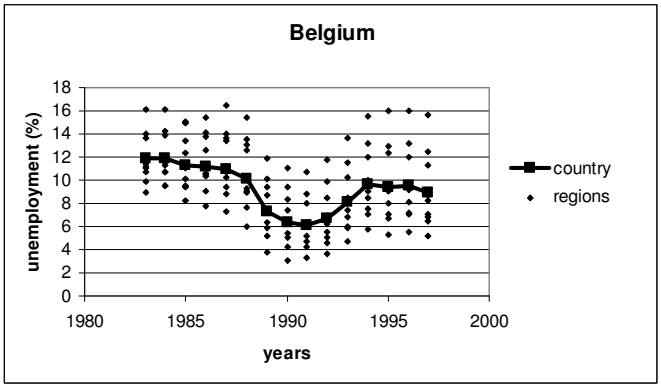
Union density|cen≥4 (%): The number of union members as a percentage of the number of employees, in countries with a centralisation index of at least 4. Available for the period 1983-1997. Source: OECD Labour Market Statistic Database.

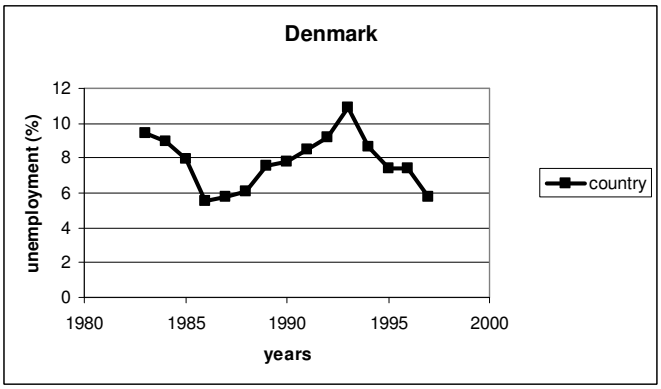
Δ^2 *p (%):* Absolute annual change in inflation, where inflation is defined as the relative change in consumer prices. Available for the period 1983-1997. Source: OECD, Main economic indicators.

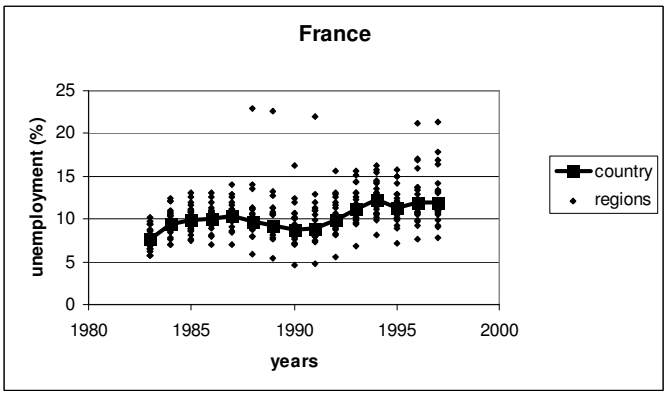
Employment protection: Index {0-2} of the strictness of employment protection. The index increases with the strictness of employment protection. Availability: annual data for the period 1985-1995 and data for the years 1998 en 2003. For the years 1983-1984 data of 1985 are used. For the years 1996-1997 and 1999-2001 data are constructed by linear interpolation. Source data 1985-1995: "Labour Market Institutions Database (LMIDB)" by Nickell. Source data 1998 and 2003: OECD data from the "OECD Labour Market Statistics Database" (variable: EPL1). The OECD data are made comparable with the LMIDB data with the use of 1990 data. In 1990 employment protection data was available in both the Nickell and OECD database.

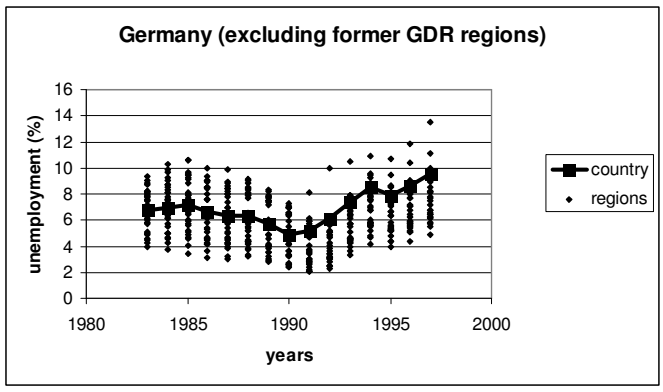
APPENDIX B: Unemployment (%)

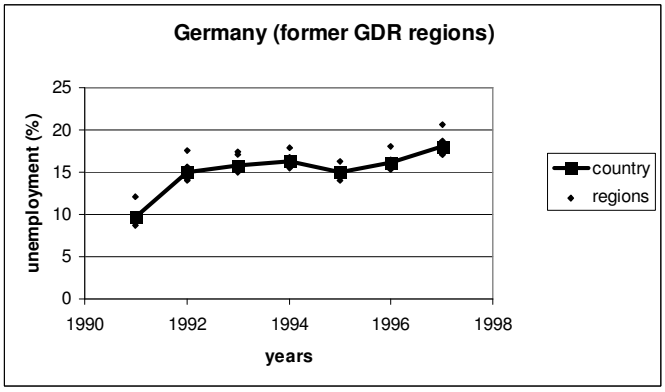


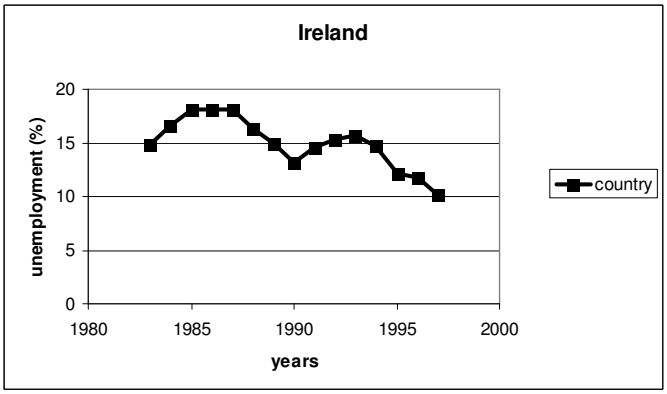


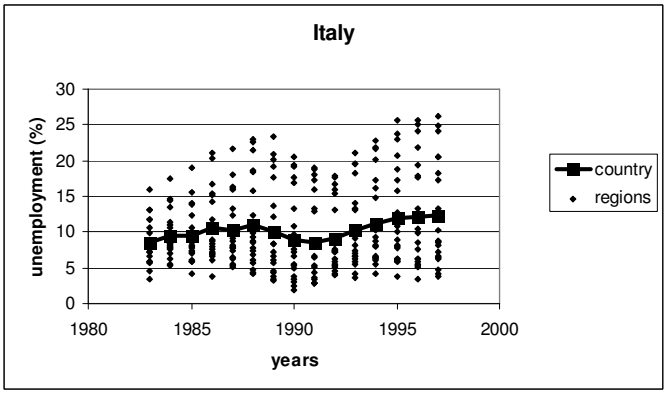


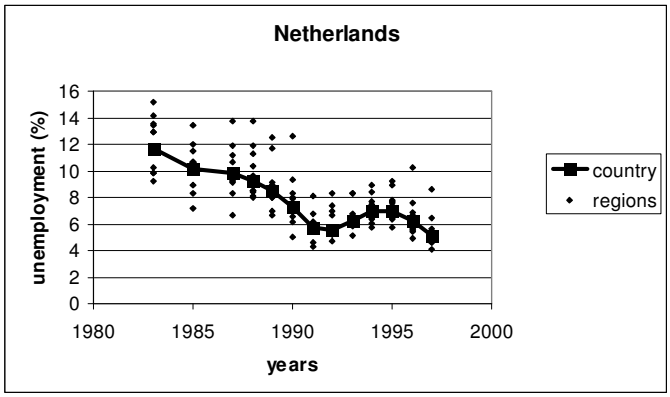


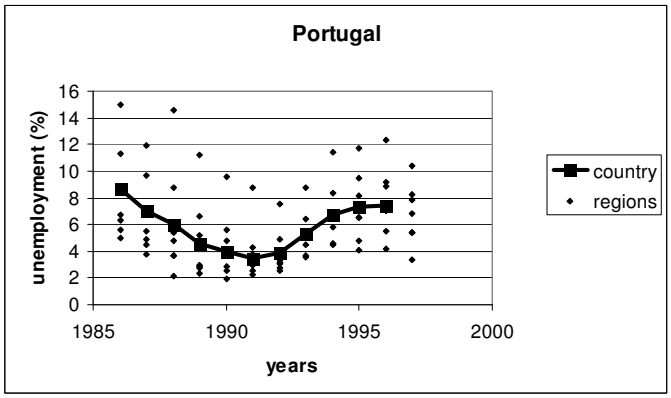

Germany (excluding former GDR regions)
unemployment (%)
16
14
12
10
8
6
4
2
0
1980
1985
1990
1995
2000
years
country
regions

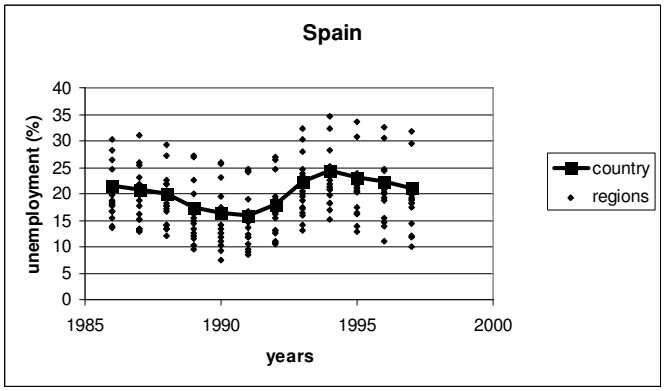

Germany (former GDR regions)
unemployment (%)
25
20
15
10
5
0
1990
1992
1994
1996
1998
years
country
regions

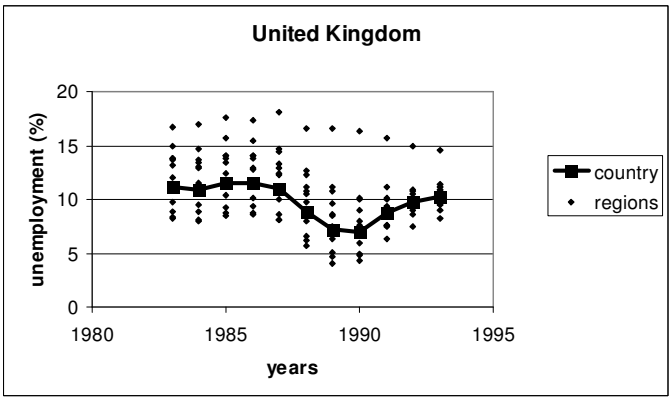

Ireland
unemployment (%)
20
15
10
5
0
1980
1985
1990
1995
2000
years
country


Italy
unemployment (%)
30
25
20
15
10
5
0
1980
1985
1990
1995
2000
years
country
regions


Netherlands
unemployment (%)
0
2
4
6
8
10
12
14
16
1980
1985
1990
1995
2000
years
country
regions


Portugal
unemployment (%)
0
2
4
6
8
10
12
14
16
1985
1990
1995
2000
years
country
regions


Spain
unemployment (%)
0
5
10
15
20
25
30
35
40
1985
1990
1995
2000
years
country
regions


United Kingdom
unemployment (%)
0
5
10
15
20
1980
1985
1990
1995
years
country
regions

APPENDIX C: Correlation coefficients between the variables

	Unemployment rate	Pop15-24	Pop55-64	Education	Productivity	Employment growth	Labour participation	Tax wedge	Unemp. Benefits	Union density|cen≤ 2	Union density|2<cen≤4	Union density|cen≥2	Δ^2 p (%)	Employment protection
Unemployment rate	1	0.38	-0.01	-0.44	0.00	-0.14	-0.51	0.06	-0.04	0.09	-0.24	-0.10	-0.12	0.05
Pop15-24		1	-0.36	-0.55	-0.30	0.11	-0.43	-0.19	-0.04	0.15	-0.22	0.20	-0.01	0.07
Pop55-64			1	-0.12	-0.01	-0.23	-0.10	0.07	-0.32	0.02	-0.02	-0.02	-0.07	0.32
Education				1	0.07	0.11	0.57	0.00	0.34	-0.41	0.51	-0.31	0.15	-0.31
Productivity					1	-0.02	-0.26	0.49	-0.08	0.15	-0.04	-0.35	0.06	0.10
Employment growth						1	0.06	-0.07	0.17	-0.12	0.09	-0.03	0.20	-0.08
Labour participation							1	-0.43	0.19	-0.15	0.13	0.02	0.06	-0.52
Tax wedge								1	-0.17	0.27	-0.14	-0.37	-0.07	0.48
Unemp. Benefits									1	-0.71	0.41	0.07	0.12	-0.26
Union density|cen≤ 2										1	-0.64	-0.14	-0.12	-0.08
Union density|2<cen≤4											1	-0.19	0.16	-0.07
Union density|cen≥2												1	-0.12	0.08
Δ^2 p (%)													1	-0.04
Employment protection														1

APPENDIX D: PRELIMINARY ESTIMATION RESULTS

Dependent variable: regional unemployment rate

	Short term		Long term		
Explanatory variables	Coef.	T Value	Coef.	T Value	S.D.
Pop15-24	0.04	0.96	0.08	0.91	0.63
Pop55-64	-0.30	-8.23	-0.51	-5.37	0.98
Education	-0.03	-2.57	-0.05	-2.10	0.17
Productivity	-0.05	-5.09	-0.08	-3.59	0.46
Employment growth	-0.15	-7.78	-0.26	-4.85	0.35
Lagged unemp. rate	0.42	6.46	-	-	
Labour participation	-0.31	-9.62	-0.53	-5.86	
Tax wedge	0.09	2.48	0.15	2.28	
Unemp. Benefits	0.03	3.39	0.06	2.63	
Union density|cen≤ 2	0.00	0.20	0.00	0.20	
Union density|2<cen<4	0.01	0.93	0.01	0.96	
Union density|cen≥4	-0.05	-3.52	-0.08	-3.06	
Δ^2 p (%)	-0.16	-3.55	-0.27	-3.32	
Employment protection	-3.06	-7.05	-5.26	-5.01	

Notes: Countries (observation period and number of regions within parentheses) are Belgium (1983-1997, 9),Denmark (1983-1997, 1), West Germany (1983-1997, 29), East Germany (1990-1997, 5), Spain (1986-1997, 17), France (1983-1997, 22), Ireland (1983-1997, 1), Italy (1983-1997, 20), the Netherlands (1983,1985,1987-1997, 12), Austria (1995-1997, 9), Portugal (1986-1997, 7) and the UK (1983-1993, 11). Total number of observations is 1549, taking into account the effect of the lagged unemployment rate. 'Labour participation' and 'lagged unemployment rate' are treated as endogenous explanatory variables. To estimate the coefficients of 'labour participation' and 'lagged unemployment rate', exogenous variables in X and Z are measured at time t-1 and t-2 as instrumental variables. In addition, 'wage' measured at time t-1 is used as an instrument for 'labour participation'. Note that we have to estimate the lagged unemployment rate with instrumental variables, because it is correlated with the random error terms. Coefficient estimates are corrected for heteroskedasticity and spatial autocorrelation. The average spatial autocorrelation coefficient is 0.75.