DISCUSSION PAPER SERIES

No. 8987

THE ROLE OF BORDERS, LANGUAGES, AND CURRENCIES AS OBSTACLES TO LABOR MARKET INTEGRATION

Kevin Bartz and Nicola Fuchs-Schündeln

Centre for Economic Policy Research

www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP8987.asp

ISSN 0265-8003

THE ROLE OF BORDERS, LANGUAGES, AND CURRENCIES AS OBSTACLES TO LABOR MARKET INTEGRATION

Kevin Bartz, Renaissance Technologies
Nicola Fuchs-Schündeln, Goethe University Frankfurt and CEPR

Discussion Paper No. 8987
May 2012

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 8987

May 2012

ABSTRACT

The Role of Borders, Languages, and Currencies as Obstacles to Labor Market Integration*

Based on a modified Spatiotemporal Autoregressive Model (STAR), we analyze whether borders still constitute significant impediments to labor market integration in the European Union, despite the formal law of free movement of labor. Using regional data from the EU-15 countries over 21 years, we find that this is the case. We further investigate whether the abolishment of border checks through the Schengen agreement or the introduction of the Euro improved our measure of labor market integration across borders, and do not find evidence in favor. Last, we investigate the role of languages, and potentially cultures, as obstacles to labor market integration. We find that indeed language borders play a larger role than country borders in explaining the lack of labor market integration across borders.

JEL Classification: C4, J4, J6
Keywords: labor market integration, European integration, spatial econometrics

Kevin Bartz
Renaissance Technologies

Nicola Fuchs-Schündeln
Chair for Macroeconomics and Development
Goethe University Frankfurt
House of Finance
60323 Frankfurt
GERMANY

Email: fuchs@wiwi.uni-frankfurt.de

*The authors are indebted to Rima Izem for very valuable contributions in developing this project. The authors also thank Anders Hopperstedt from the Center of Geographic Analysis at Harvard for calculating the driving times between regions, and Eduardo Morales for providing excellent research assistantship.

Submitted 15 May 2012

The Role of Borders, Languages, and Currencies as Obstacles to Labor Market Integration*

Nicola Fuchs-Schündeln
Goethe University Frankfurt,
and CEPR

Kevin Bartz
Renaissance Technologies

May 15, 2012

Abstract

Based on a modified Spatiotemporal Autoregressive Model (STAR), we analyze whether borders still constitute significant impediments to labor market integration in the European Union, despite the formal law of free movement of labor. Using regional data from the EU-15 countries over 21 years, we find that this is the case. We further investigate whether the abolishment of border checks through the Schengen agreement or the introduction of the Euro improved our measure of labor market integration across borders, and do not find evidence in favor. Last, we investigate the role of languages, and potentially cultures, as obstacles to labor market integration. We find that indeed language borders play a larger role than country borders in explaining the lack of labor market integration across borders.

1 Introduction

The process of economic and political integration in Europe began shortly after the war, and was intended to support the economic development of all participating countries. European integration did not happen overnight, but has been an ongoing piecewise transition process. Two major components of European integration are the free movement of goods and services, as well as labor. In principle, the free movement of labor has been established for many decades in the European Union (see e.g. Vandamme, 2000). Article 48 of the treaty establishing the European Economic Community in 1957 stated that "the free movement of workers shall be ensured within the Community not later than at the date of the expiry of the transitional period. This shall involve the abolition of

*The authors are indebted to Rima Izem for very valuable contributions in developing this project. The authors also thank Anders Hopperstedt from the Center of Geographic Analysis at Harvard for calculating the driving times between regions, and Eduardo Morales for providing excellent research assistantship.

any discrimination based on nationality between workers of the Member States, as regards employment, remuneration and other working conditions." Both migration and commuting should thus in principle be as easy across country borders of EU member countries as within country borders for workers with EU nationality.

Labor market integration through migration and commuting leads to efficiency gains in a model with regional shocks (see e.g. Borjas, 2001). The theory of migration states that workers decide where to reside based on expected future wages in source and destination regions (see e.g. Hunt, 2006). The probability of unemployment as well as the expected wage conditional on being employed enter the calculation of expected future wages. One typically assumes that fixed costs of migration affect the decision whether to migrate or not. When considering a model of commuting, it is more reasonable to assume flexible costs that are increasing in the distance of the commute (see e.g. Fuchs-Schündeln and Izem, 2012). Otherwise, the decision whether to commute or not is influenced by the same factors as the decision whether to migrate or not.

This paper investigates to which degree labor market integration has been achieved and how it has developed over the last two centuries within the 15 member countries of the European Union before the enlargement in 2001, called EU-15 from now on.[1] Commuters should react to differentials in wages and unemployment rates between adjoining regions, and should thereby introduce spatial correlation of unemployment rates and income levels. Similarly, firms should take advantage of nearby regions with lower wages when establishing relationships with suppliers of goods and services. This should lead to regional convergence of wage levels. Due to the free movement of goods and labor, these arguments hold across country borders to the same extent as within country borders. It has already been shown that country borders constitute significant impediments to the flow of goods (see e.g. McCallum, 1995, and Anderson and van Wincoop, 2003). Is the same true for the flow of labor?

To answer this question, we use 21 years of regional data from the EU-15 countries. We build a spatiotemporal autoregressive model (STAR) that analyzes the spatial as well as time-dependence of unemployment rates and GDP per capita in the 207 NUTS2 regions of the EU-15 countries.

[1] Free movement of labor has only been established in 2011 for the new EU member countries.

We modify the model to allow for a different spatial correlation across and within national country borders to answer the question whether borders still matter in the European Union. We also discuss some methodological issues that arise in this modification of the standard STAR model, which could be used to answer similar questions regarding the importance of borders in other contexts.

After presenting the main results, namely that country borders constitute significant impediments to labor market integration in the EU-15, we further analyze why this is the case. Specifically, we investigate three hypotheses. Our first hypothesis is that border controls might be the reason. Border checks not only increase average travel time, but also substantially increase the uncertainty about travel time. This could have an important effect especially on daily commuting. The Schengen agreement, initially signed in 1985, led to the elimination of border controls between many countries in the European Union at different points in time. E.g. Germany implemented the Schengen agreement in 1995, Italy in 1998, and Denmark in 2001. Thus, this agreement allows us to analyze whether labor market integration across borders improved after the abolition of border controls.

The second hypothesis is that different currencies in countries of residence and of work complicate labor market integration. To test this, we analyze whether the adoption of the irrevocable conversion rates of the national currencies towards the Euro in 1999 or the introduction of the Euro coins and notes in 2002 led to higher cross-border labor market integration. To the extent that the implementation of Schengen and the adoption of the Euro were endogenous, we are biased towards finding significant effects of these policies.

Third, we focus on the role of languages as obstacles to labor market integration. This is a natural hypothesis given the low migration rates in the European Union when compared to interstate migration in the US.[2] In the case of the EU, country borders and language borders do not necessarily coincide (e.g. both French and Flamish are spoken in Belgium, while Germany and Austria share the same language). Moreover, using a measure of the closeness of two languages from the lexicostatistical analysis of Dyen et al. (1992), we can generate further variation by analyzing whether labor market integration is harder to achieve e.g. between France and Germany,

[2] One should note, however, that Europeans migrate less in reaction to labor market differentials even within one country than US-citizens (Decressin and Fatas, 1995).

whose languages share only 24% of cognate words, than between France and Spain, whose languages are both romance languages and share 73% cognate words. Last, by allowing regional languages to identify separate language areas, we can also analyze whether any language effect is a true language effect, or rather a cultural effect. If people in a region speak both the national and the regional language, and the regional language borders still matter, it is more likely that we pick up a cultural effect than a true language effect which prevents communication.[3]

Our paper relates to the literature on migration and commuting in response to regional shocks. Blanchard and Katz (1992) demonstrate the importance of labor migration as a reaction to regional shocks within the United States. Applying their analysis to Europe, Decressin and Fatas (1995) find that Europeans are far less willing to migrate in response to economic incentives, and that the reaction to regional shocks in Europe comes mostly through changes in the unemployment rate and participation rate. Overman and Puga (2002) show that the regional distribution of unemployment rates became even more polarized in the European Union between 1986 and 1996.

The rest of the paper is organized as follows. In the next section, we describe our data. Section 3 presents our modification of the Spatiotemporal Autoregressive Model. We show the main results in Section 4, and analyze the effect of two policies, namely the Schengen agreement and the Euro introduction, on labor market integration across borders in Section 5. Section 6 investigates the role of languages by contrasting the importance of country borders to the one of language borders. The last section concludes.

2 Data

We use data from Eurostat on NUTS2 regions in the EU-15 countries.[4] The data cover 21 years from 1986 to 2006. Our main variable of interest is the unemployment rate, defined as the ratio of people aged between 15 and 64 years and declared as unemployed over the sum of employed and unemployed people in this age group. Eurostat calculates unemployment rates on the NUTS2 level based on the European Labour Force Survey, which guarantees harmonization across coun-

[3] Brügger, Lalive, and Zweimüller (2009) exploit the language border in Switzerland to analyze the effect of culture on unemployment.

[4] NUTS stands for Nomenclature of Territorial Units for Statistics. The EU defines regions on the NUTS 1, 2, and 3 levels, with an increasing degree of fineness.

tries through a common questionnaire design (Eurostat, 2008). As control variables, we use the proportion of part-time workers among the employed, the proportion of the population working in agriculture, industry, or services, the educational composition of the work force,[5] the age composition (the share of the active population aged 15 to 24, 25 to 54, and 55 to 64), and the participation rate. Data on GDP per capita, deflated with the CPI, are only available from 1996 on. Ideally, we would analyze the spatial correlation of wages, which predict the probability of commuting, but in the absence of wages we use GDP per capita as a proxy. The data on languages come from the World Language Mapping System of Global Mapping International, and were merged into the data on NUTS2 level using Geographic Information Systems.

In Appendix A, we describe a few necessary imputations due to missing values. Moreover, some NUTS change their boundaries during our sample period, and we describe how we deal with them in the appendix. Whenever two NUTS regions were merged into one or one NUTS region was split into two during the sample period, we use the larger NUTS region as the unit of analysis and reconstruct data on this region from the two underlying NUTS by building population weighted averages.[6]

We do our analysis on two different data sets. The longer data set covers the years 1986 to 2006 and contains 148 NUTS2 regions. It misses GDP per capita, as well as all information on Austria, Finland, and Sweden, which joined the European Union only in 1995, and East Germany. Moreover, it splits the UK into eleven NUTS2 regions (those that were in place up to the reform in 1996). The shorter data set covers the years 1996 to 2006 and contains 207 NUTS2 regions. It includes GDP per capita, information on Austria, Finland, Sweden, and East Germany, and splits the UK into 37 NUTS2 regions. We use this data set whenever we want to analyze GDP per capita, and whenever it is appropriate to answer the question of interest on the shorter time period. Figure 1 shows the unemployment rates in the NUTS2 regions in 1986 and 2006. Figure 2 shows the GDP data for 2006. It is worth noting that Greece does not share borders with any other

[5] Specifically, we use the proportion of the employed who have as highest educational degree at most lower secondary education, the proportion who have as highest educational degree at most upper secondary, non-tertiary education, and the proportion of individuals with tertiary education.

[6] This issue is especially important for the UK, where the NUTS2 definitions changed completely in 1996. For all other countries, there are only a few cases of reforms.

EU-15 country, and some countries (Finland, Sweden, UK, Ireland) have only a very small number of NUTS2 regions that share borders with other EU-15 countries. The across-border correlation is mostly identified through the Central European[7] and Romance-language speaking countries in the sample.

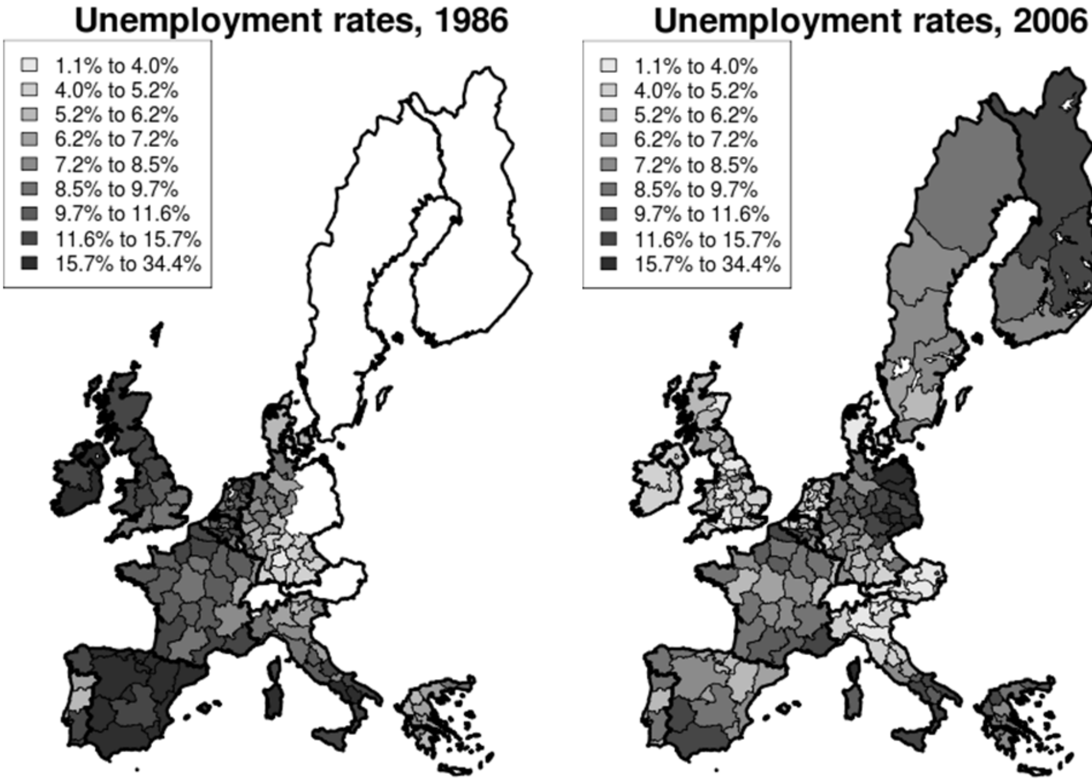

Figure 1: Unemployment Rates in NUTS2 Regions in 1986 and 2006

3 The Spatiotemporal Autoregressive Model

The spatiotemporal autoregressive (STAR) model (see e.g. Cressie 1993, and Kelley et al. 2000) is widely used in the spatial analysis of aerial data. STAR specifies the relationship of the variable of interest to covariates collected in the region of interest, while accounting for the correlation structure in time and space. Time correlation describes a unit's dependency on its own past – i.e.,

[7] With Central European countries, we mean Austria, Belgium, Germany, Luxembourg, and the Netherlands.

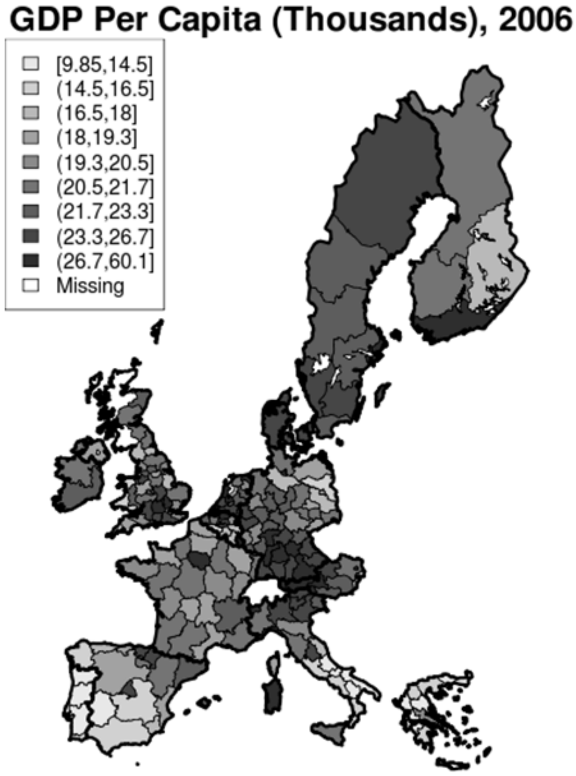

Figure 2: GDP per capita (in thousand Euros) in NUTS2 Regions in 2006

a time-autoregressive process – while spatial correlation describes its dependency on neighboring units.

Subsection 3.1 presents a specification for STAR's spatiotemporal neighborhood matrix in the case of discrete-time data. Our discrete times are the calendar years over which unemployment is measured. In Subsection 3.2, we develop a parameterized STAR model that allows us to measure the effect of borders on spatial correlation. We discuss possible extensions of our model in Subsection 3.3.

3.1 Discrete-time STAR

Suppose there are n_s spatial units observed at each of n_t discrete times. The STAR model can be specified as

$$
\begin{aligned}
Y(s,t) &= X(s,t)\beta + U(s,t),\ s \in \{1,\ldots,n_s\},\ t \in \{1,\ldots,n_t\} && (1)\\
U &= WU + \varepsilon && (2)\\
\varepsilon(s,t) &\overset{\text{iid}}{\sim} \mathcal{N}(0,\sigma^2), && (3)
\end{aligned}
$$

Equation (1) gives the main linear model, where s denotes a spatial unit and t time. In our application, $Y(s,t)$ is the unemployment rate or GDP per capita of unit s at time t. $X(s,t)$ is the unit's vector of covariates, and $U(s,t)$ is its residual.

Expressions (2) and (3) specify the residual correlation structure. Each residual $U(s,t)$ is composed of two parts. The first is WU, a weighted average of neighboring residuals in space and time. The second is $\varepsilon(s,t)$, an independent innovation that allows $U(s,t)$ to deviate from the average of its neighbors.

The key to modeling correlation is the spatiotemporal neighborhood matrix W. Each row of W contains all the neighbor weights associated with a specific unit s at a specific time t. In the standard STAR model, nontrivial neighbors include adjacent spatial units at time t (spatial neighbors), the unit s itself at time $t-1$ (time neighbor), or adjacent spatial units at time $t-1$ (space-time neighbors). Combining these parts, W is assumed to take the form

$$W = \rho_S S + \rho_T T + \rho_{ST} ST \qquad (4)$$

S is an adjacency matrix containing all the spatial neighbors at every time. Matrix T contains the time neighbors. S and T are responsible for the space and time correlation in U, respectively, while the matrix product ST provides a space-time interaction term.

For discrete times, all three matrices are square with dimension $n_s n_t \times n_s n_t$, each row and column corresponding to a specific unit and time. By convention, rows are ordered first by time and then by spatial unit, which gives the matrices the special sparse forms shown in (5), which illustrates the structure of the three matrices S, T, and ST for the case of $n_t = 3$. A mathematical definition

Figure 3: NUTS2 Regions Used for the Impulse Response Analysis

is given above each matrix. S is a block-diagonal matrix. Assuming the spatial neighbors do not change over time, each block holds the same single-time spatial adjacency matrix S_1 (dimension $n_s \times n_s$). T is a lower shift matrix. Each row contains a single nonzero entry corresponding to the same unit at the previous time. ST is a lower shift block matrix with S_1 in the blocks.

$$
\begin{array}{cccc}
 & S = I_{n_t} \otimes S_1 & T = L_{n_t} \otimes I_{n_s} & ST = L_{n_t} \otimes S_1 \\
\begin{array}{c}\text{Time 1}\\ \text{Time 2}\\ \text{Time 3}\end{array} &
\left(\begin{array}{c|c|c} S_1 & & \\ \hline & S_1 & \\ \hline & & S_1 \end{array}\right) &
\left(\begin{array}{c|c|c} & & \\ \hline I_{n_s} & & \\ \hline & I_{n_s} & \end{array}\right) &
\left(\begin{array}{c|c|c} & & \\ \hline S_1 & & \\ \hline & S_1 & \end{array}\right)
\end{array}
\tag{5}
$$

Our primary focus is on the specification of S_1. As an example, consider the five regions highlighted in Figure 3, which are drawn from Belgium and the Netherlands. S_1 in the standard

STAR model is given by

$$S_1 = \begin{array}{c} \\ \text{NL41} \\ \text{NL42} \\ \text{BE21} \\ \text{BE24} \\ \text{BE10} \end{array} \begin{array}{c} \begin{array}{ccccc} \text{NL41} & \text{NL42} & \text{BE21} & \text{BE24} & \text{BE10} \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array} \tag{6}$$

The first row has nonzero entries for the neighbors of NL41, namely NL42 and BE21, which are the only two units adjacent to NL41 among the five units under consideration. In the second row, we see NL42's single neighbor, NL41.

Note that S_1 given in (6) is a binary matrix. It is common to row-standardize S_1 (or, equivalently, S) so that the sum of each row is one, which fixes each unit's total spatial dependency to a common constant. We investigate the implications of standardization in Appendix B and decide to work without standardization.

3.2 Modeling Border Effects

In the standard STAR model, all spatial neighbors are given equal weight in the autocorrelation of U. Our principal contribution is to further parameterize the spatial neighborhood matrix S to capture the impact of neighbors within and across country borders.[8] We write

$$W = \rho_W S_W + \rho_A S_A + \rho_T T + \rho_{WT} S_W T + \rho_{AT} S_A T \tag{7}$$

S_W is the within-borders neighborhood matrix, comprised of spatial neighbors that lie within a common country border. Each nonzero entry in S_W hence corresponds to a pair of spatial neighbors within the same country. In contrast, S_A is the across-borders neighborhood matrix, comprised of spatial neighbors that lie on different sides of a country border – i.e., adjacent pairs from different countries.

[8] Sun et al. (2005) decompose a spatial effect into building and neighborhood effects in an analysis of multi-unit residential real estate data. This is somewhat related to our approach.

To understand this, we again consider the example from Figure 3. Our parameterization yields (column headers omitted for brevity)

$$S_A = \begin{array}{c} \text{NL41} \\ \text{NL42} \\ \text{BE21} \\ \text{BE24} \\ \text{BE10} \end{array} \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}; \; S_W = \begin{array}{c} \text{NL41} \\ \text{NL42} \\ \text{BE21} \\ \text{BE24} \\ \text{BE10} \end{array} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Only two entries of S_A are nonzero – those corresponding to the country border between BE21 and NL41. All other spatial adjacencies are within-border pairs and are relegated to S_W.

Our parameterization aims at describing the impact of country borders on spatial correlation. Specifically, we would like to test the hypothesis that borders do not matter, such that spatial neighbors within a country border exhibit the same correlation than those crossing the country border, $\rho_W = \rho_A$. If $\rho_W = \rho_A$, then the border structure makes no difference. At the opposite extreme, if hypothesis testing establishes that $\rho_A = 0$, then there is no significant spatial correlation at all across borders.

We provide a brief sketch of how we estimate the model via maximum likelihood. At the top level, we use Nelder-Mead to optimize over the ρ terms. At each iteration, W is given by the linear combination (4) using the current ρ. Then, following Cressie (1993), a normal distribution for Y is derived from equations (1), (2) and (3), through which we evaluate the likelihood. An R package implementing our model and the basic STAR model is available upon request. We make widespread use of several optimizations for band- and block-diagonal matrices (Bates and Maechler, 2010).

3.3 Extensions

We introduce two natural extensions of the STAR model used in this paper. First, our setup allows for the use of any type of border, not just country borders. In Section 6, we examine the option of using language borders instead. We also consider language sets of differing granularity (national or regional languages, and cultural effects).

Second, the spatial neighborhood matrix S need not be binary. It can be continuous, with the weight on a pair of units based on a measure of distance between them. The lower the distance, the greater the weight. Under this approach, even non-adjacent pairs are considered "neighbors," albeit with lower weights. Note that this makes S a dense matrix, unlike the sparse matrix generated by using the spatial adjacency matrix as S.

We experiment with two spatial neighborhood matrices, one binary and one continuous. The first is the spatial adjacency matrix between NUTS2 units, as described in Section 3.2. The second is based on driving distances between major cities in each pair of units, calculated using Geographic Information Systems. To transform the distances into weights, we try three functions: the inverse of the driving time, the inverse squared, and the inverse cubed. In terms of the likelihood of the baseline model described below in Section 4, the spatial adjacency matrix outperforms all three alternatives.

In our specifications, neighborhood matrices based on driving distances tend to place too much weight on faraway regions. This is especially notable in East Germany, where substantial neighborhood weight is assigned to a dense block of nearby West German regions which, while close in distance, exhibit markedly different unemployment rates. As a result, the residuals ε have substantially greater magnitude all across Germany under the driving distance matrices. In comparison, under the spatial adjacency matrix, only units at the former East-West German border are impacted. While the problem is somewhat mitigated by using a higher-order inverse (square or cube) in the transformation of distances to weights, the simple spatial adjacency matrix still performs best in terms of the likelihood. As a result, we work exclusively with the spatial adjacency matrix as S for the remainder of our analysis.[9]

[9] We also try interacting the adjacency matrix with some function of distance. That way, the distance between adjacent regions is taken into account. Some of these specifications slightly outperform the simple adjacency matrix in terms of the likelihood. However, results are very robust to the weighting matrix, and for simplicity reasons we decide to present results with the simple adjacency matrix.

4 Do Borders Matter?

4.1 The Border Effect in the EU

4.1.1 Unemployment Rates

We first investigate whether borders matter in the European Union by analyzing whether the spatial dependency of unemployment rates differs within and across borders. We use the logit specification of the unemployment rate as the dependent variable.[10] Table 1 shows the results. In the first specification, we do not include any controls but time fixed effects. We then successively add more control variables that can potentially explain the unemployment rate. Specification (ii) adds the participation rate, as well as the proportion of the work force that is working part-time, since both variables could affect reported regional unemployment rates. In specification (iii), we additionally control for the proportion of the work force in agriculture and industry (the omitted variable being services). Specification (iv) adds the population density (measured as the work force per square kilometer) to capture differences in unemployment rates between rural and urban areas, and the proportion of the labor force aged 15 to 24, and aged 25 to 54, to capture systematic differences in unemployment rates for different age groups. Finally, in specification (v) we include country fixed effects. These could be important if there were systematic differences in the way unemployment rates are defined across countries. As explained in Section 2, this should however not be the case since Eurostat reports harmonized unemployment rates. A downside of including country fixed effects is that they capture systematic differences in labor market policies across countries that lead to differences in unemployment rates, and we would not like to control for these differences. If for example in country A the unionization rate is higher and leads to higher unemployment than in neighboring country B, then the model of commuting predicts that many workers would commute from a region in country A to a neighboring one in country B, and this would lead to a positive spatial correlation in unemployment rates across borders. Therefore, specification (iv) is our preferred specification and will be used for further analyses in the following

[10] The dependent variable in the model is allowed to vary between minus and plus infinity, while the unemployment rate varies between 0 and 1. Therefore, we use as the dependent variable the logit transformation of the unemployment rate, $\ln\left(\frac{UR}{1-UR}\right)$.

sections.[11] The last specification in column (vi) replaces the country and year fixed effects with country-year fixed effects. The model thus analyzes the spatial correlation of regional deviations of the unemployment rate from the country-year average. Theory would predict that it is the absolute level of the unemployment rate, not the deviation from the annual national mean, that induces commuting between regions. Moreover, adding country-year fixed effects means that we are estimating a model with many parameters. On the other hand, country-year fixed effects capture country-specific shocks. Therefore, one cannot argue that a potentially lower spatial correlation across borders is caused by country-specific shocks in this specification.

The first striking result of Table 1 is that the intertemporal autocorrelation of regional unemployment rates ρ_T is quite high. This high intertemporal autocorrelation also holds up after controlling for country-year fixed effects in specification (vi). This shows that some regions have systematically higher or lower unemployment rates than the country average. With regard to the two main parameters of interest, namely ρ_A, capturing the spatial correlation across borders, and ρ_W, capturing the spatial correlation within borders, in all specifications is the spatial correlation across borders rather low and borderline significant.[12] By contrast, the spatial correlation within borders is 4 to 6 times larger in specifications (i) to (v) and highly significant. Thus, while unemployment rates are highly spatially correlated within national borders, national borders still seem to constitute significant impediments to labor market integration. The spatial correlations across and within borders are significantly different in all specifications but the last one. The within-border correlation parameter ρ_W drops significantly when country-year fixed effects are added to the model in specification (vi). This seems to be the case because, if a country exhibits a country-wide shock, and thus country-year fixed effects are important, part of this shock might be captured in the contemporaneous spatial correlation if we do not allow for country-year fixed effects.

The log-likelihood increases by 27 points in specification (iv) when allowing for different spatial correlation across and within borders, as compared to the standard STAR model that does not

[11] Nevertheless, as a robustness check we included controls for employment protection, the gross unemployment replacement rate, union density, product market regulation, and active labor market policies in specifications (iv) and (v). Results are robust and are available from the authors upon request.

[12] The p-value on ρ_A amounts to 0.05 in specifications (iii) and (vi), 0.10 in specification (i), and 0.14 in the other specifications.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
ρ_A	0.018	0.019	0.025	0.019	0.021	0.033
	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)	(0.018)
ρ_W	0.117	0.113	0.116	0.115	0.100	0.046
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
ρ_T	0.919	0.909	0.908	0.903	0.896	0.862
	(0.008)	(0.009)	(0.009)	(0.009)	(0.010)	(0.011)
ρ_{AT}	-0.018	-0.017	-0.022	-0.016	-0.012	-0.008
	(0.013)	(0.014)	(0.013)	(0.013)	(0.016)	(0.020)
ρ_{WT}	-0.101	-0.095	-0.097	-0.095	-0.07	-0.012
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.008)
Participation		-1.507	-1.348	-1.787	-1.236	-1.476
		(0.247)	(0.249)	(0.247)	(0.251)	(0.242)
Part-time		0.335	0.183	0.068	0.235	0.608
		(0.213)	(0.216)	(0.211)	(0.242)	(0.251)
Agriculture			-0.687	-0.397	-0.110	-0.094
			(0.137)	(0.147)	(0.158)	(0.150)
Industry			0.087	-0.158	-0.052	-0.052
			(0.167)	(0.166)	(0.162)	(0.150)
Proportion aged 15 to 24				3.415	3.019	3.201
				(0.379)	(0.374)	(0.379)
Proportion aged 25 to 54				0.293	0.334	0.450
				(0.311)	(0.306)	(0.324)
Population Density				0.130	0.169	0.176
				(0.029)	(0.029)	(0.026)
country fixed effects	no	no	no	no	yes	no
year fixed effects	yes	yes	yes	yes	yes	no
country-year fixed effects	no	no	no	no	no	yes
log-likelihood	755.7	774.2	788.5	848.7	1030.1	1355.4

Note: Standard errors are in parentheses.

Table 1: STAR Model of Unemployment Rates, 1986-2006

distinguish by borders when considering spatial correlation (results available from the authors upon request). Thus, allowing for different spatial correlation across and within borders significantly increases the performance of the model.

The space-time autocorrelation parameters are all negative, and again larger in absolute terms within borders than across borders. The effect of a positive shock in one region would blow up over time if the space-time autocorrelation parameters were not negative. To gain an economic understanding, recall that the spatial parameters (ρ_A and ρ_W) represent the effect of spatial neighbors in the current time period, while the space-time interaction terms (ρ_{AT} and ρ_{WT}) represent the effect of spatial neighbors in the previous time period. Their opposite signs mean that what drives the neighbor effect is not the neighboring residuals themselves, but the change in neighboring residuals from the previous time period. In other words, the model says that areas with high residuals are likely to be near areas with growing residuals, and vice versa.

Among the control variables, regions with a higher participation rate display significantly lower unemployment rates, indicating that booming regions have both higher participation and lower unemployment rates. A higher population density is associated with a higher unemployment rate. Last, regions with a high share of the young work force exhibit significantly higher unemployment rates, reflecting the problem of youth unemployment across Europe. The other estimated coefficients on the control variables are not significant.

We investigate the standardized residuals ε resulting from the regressions, which should show no correlation over time and space, for their spatial dependence. To test the assumption of spatial independence, we conduct a Moran's I test of spatial autocorrelation of the residuals for each year. The p-values all indicate that there is no significant spatial correlation in the residuals.[13]

We also repeat the regressions on the shorter time sample 1996-2006 that includes more NUTS. Here, we can also include the proportion of the working age population with low and medium education as controls (the omitted group being the proportion of the work force with tertiary education). Results are reported in Table 2. In this regression, the across border correlation parameter ρ_A is always positive and significant, but also always significantly smaller than ρ_W. Otherwise, the results are rather similar.

4.1.2 GDP

The first two columns of Table 3 show the results of a STAR regression with log GDP per capita as the dependent variable, using GDP as a proxy for wages. In these regressions, we have to use the shorter sample from 1996 to 2006 that comprises all 15 countries. We run regressions analogous to specifications (iv) and (vi) in Table 2. As in the case of the unemployment rate, the time auto-

[13] An exception is the first time period, i.e. 1986, which is lacking any temporal neighbor in the model.

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
ρ_A	0.056	0.054	0.047	0.051	0.036	0.052
	(0.014)	(0.014)	(0.015)	(0.015)	(0.018)	(0.021)
ρ_W	0.127	0.123	0.123	0.126	0.116	0.095
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
ρ_T	0.893	0.875	0.872	0.866	0.849	0.821
	(0.012)	(0.012)	(0.013)	(0.013)	(0.015)	(0.015)
ρ_{AT}	-0.050	-0.046	-0.040	-0.042	-0.028	-0.021
	(0.015)	(0.015)	(0.016)	(0.016)	(0.019)	(0.023)
ρ_{WT}	-0.105	-0.097	-0.096	-0.098	-0.082	-0.056
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.008)
full set of controls	yes	yes	yes	yes	yes	-yes
country fixed effects	no	no	no	no	yes	no
year fixed effects	yes	yes	yes	yes	yes	no
country-year fixed effects	no	no	no	no	no	yes
log-likelihood	258.8	281.4	298.1	317.1	452.9	625.7

Note: Standard errors are in parentheses.

Table 2: STAR Model of Unemployment Rates, 1996-2006

correlation is very high, indicating a high persistence in the GDP series. The parameter capturing the contemporaneous spatial correlation across borders, ρ_A, is positive and significant. However, ρ_W, which captures the contemporaneous spatial correlation within borders, is 3.5 times (in specification (iv)) and 1.5 times (in specification (vi)) larger than ρ_A, and the difference is significant throughout.[14] This difference is also reflected in the negative space-time interaction coefficients, where in absolute terms the lagged within-border correlation (ρ_{WT}) is always larger than the lagged across-border correlation (ρ_{AT}). The log-likelihood is 10 points larger in specification (iv) than in a standard STAR model that does not allow for different spatial correlations across and within borders (results available upon request), indicating a significant improvement ($p \approx 6 \times 10^{-5}$).[15]

However, the results regarding GDP are sensitive to the inclusion of Luxembourg. Luxembourg's GDP per capita is more than twice the size of the average GDP in all other NUTS. While it is only

[14] Equality of ρ_A and ρ_W can be rejected with a very low p-value in specification (iv), and a p-value of 0.007 in specification (vi).

[15] As for the controls, a higher participation rate is associated with significantly higher GDP. Regions with a higher proportion of low or medium educated workers have lower GDPs than regions with a higher share of workers with tertiary education. Moreover, agricultural regions have a significantly lower GDP, as well as regions with a high share of the young work force. Population density is positively correlated with GDP.

one single NUTS, it is very well connected through five neighbors, all across country borders, and thus it is central enough to the across-border network, which features a lower number of pairs than the within-border network, to influence the results.[16] The last two columns in Table 3 show results for the same regressions as the first two columns, but omitting Luxembourg in the data. Likelihood ratio tests comparing models with only one spatial correlation parameter to the models shown in Table 3 with different spatial correlation parameters within and across borders do not indicate a significant improvement of the latter models over the former ones once Luxembourg is excluded.[17] Therefore, we conclude that there are no significant differences in the spatial correlation of GDP per capita across and within borders in the EU-15. This could be the case either because wage differentials do not matter for commuting, or because GDP per capita is only a non-perfect proxy for wages.[18] We conduct the remaining analyses in this paper solely using the unemployment rate as the dependent variable.

	with Luxembourg		omitting Luxembourg	
	(iv)	(vi)	(iv)	(vi)
ρ_A	0.032	0.073	0.082	0.076
	(0.017)	(0.020)	(0.017)	(0.021)
ρ_W	0.113	0.106	0.113	0.105
	(0.006)	(0.008)	(0.006)	(0.008)
ρ_T	0.961	0.948	0.952	0.948
	(0.009)	(0.010)	(0.010)	(0.010)
ρ_{AT}	-0.029	-0.060	-0.070	-0.063
	(0.019)	(0.023)	(0.018)	(0.023)
ρ_{WT}	-0.107	-0.099	-0.106	-0.098
	(0.007)	(0.008)	(0.006)	(0.008)
full set of controls	yes	yes	yes	yes
country fixed effects	no	no	no	no
year fixed effects	yes	no	yes	no
country-year fixed effects	no	yes	no	yes
log-likelihood	2,723.2	3,024.7	2,776.0	3,004.4

Note: Standard errors are in parentheses.

Table 3: STAR Model of GDP per capita, 1996-2006

[16] See also Footnote 27 for further results confirming this intuition for the importance of Luxembourg.
[17] The p-values of the likelihood ratio tests are 0.2 for specification (iv) and 0.39 for specification (vi).
[18] Labor shares might differ across countries and over time.

4.2 Impulse Response Functions

In order to visualize the importance of the border effect, we show impulse response functions of a shock of size 1 in the residual ε in one region. The impulse responses obviously depend on the neighborhood structure. We present two cases. In the first one, we pick out the five regions in Belgium and the Netherlands shown in Figure 3 and used already for illustrative purposes in Section 3, and show the impulse response function if the entire network consisted just of these five regions. In the second case, we use all the NUTS2 regions in Belgium and the Netherlands and show the impulse response function after a shock of 1 to the same region. Figure 3 shows the five regions, as well as the other regions in Belgium and the Netherlands. We shock the region BE21. The region BE24 is one node away from it, while the region BE10 is two nodes away. The region NL41 is one node away from BE21, but on the other side of the border, while the region NL42 is two nodes away.

We present three different scenarios for the impulse responses. First, we work with the actual border structure. In the second scenario, we assume that each of the NUTS belongs to a separate country, such that all borders are country borders. In the third one, we assume that all NUTS belong to the same country, and such the country border between the Netherlands and Belgium is assumed away.

Figure 4 shows the impulse response function of a shock of 1 to region BE21 under the three different scenarios if the five NUTS under consideration were the only ones in the network. The first thing that is visible is that the shock of 1 to the residual ε leads to a contemporaneous effect larger than 1 in the residual U of the shocked region BE21. This is caused by the contemporaneous spatial effects. Secondly, the effect in the original network structure is decreasing the further the NUTS are away within Belgium and within the Netherlands, and for each given number of nodes it is stronger in Belgium than in the Netherlands. The decay of the effect with distance is so strong that the region in Belgium that is two nodes away is affected less than the Dutch region that is one node away. The third thing that is clearly visible is that the effect on the other spatial units would be much smaller if they were all located in different countries, i.e. in scenario 2. If they were all within the same country (scenario 3), then the contemporaneous effect would amount to around

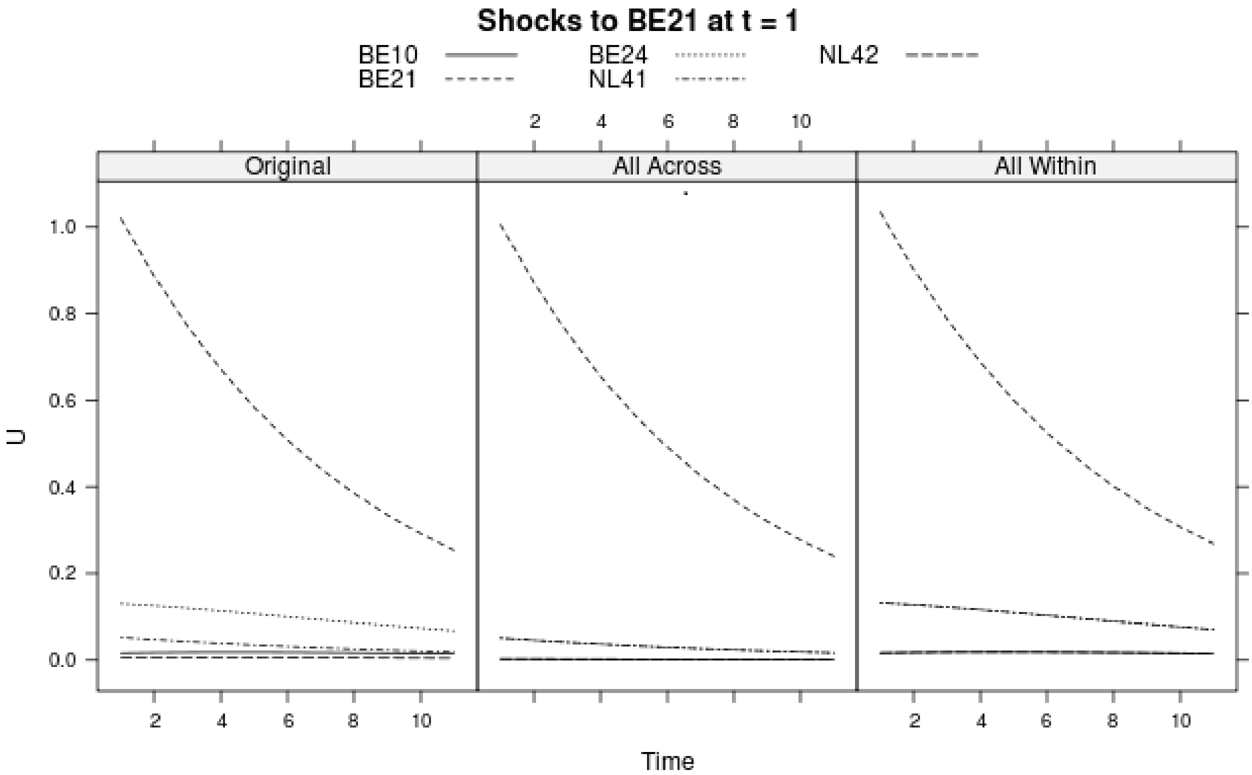

Figure 4: Impulse Response Function for 5 NUTS

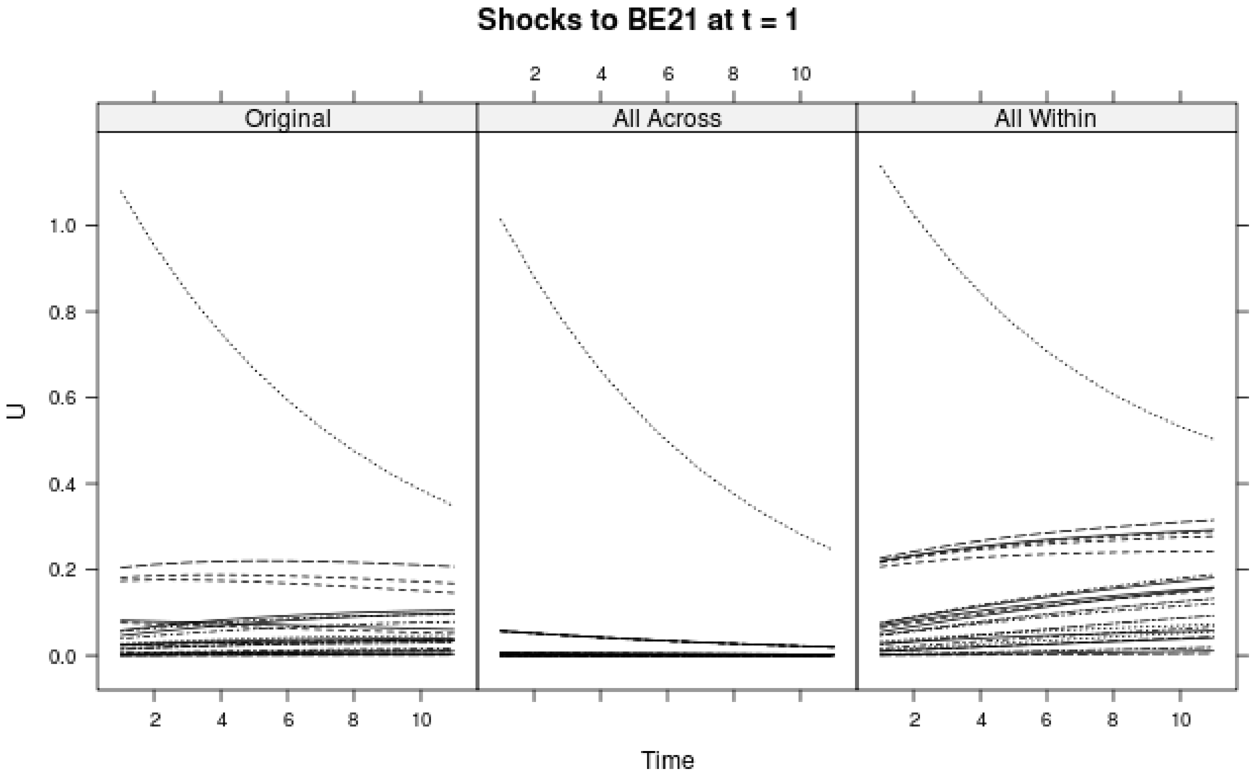

Figure 5: Impulse Response Function for All Belgian and Dutch NUTS

0.15 on the two NUTS that are one node removed, but would already be close to zero for the two NUTS that are two nodes removed.

Considering only these 5 NUTS, the differences in the magnitudes of the effects do not seem to be very large between the three scenarios. Yet, the larger and denser the network, the larger the effects become. Figure 5 shows the impulse response of the same shock to region BE21, considering a network structure that includes all NUTS in Belgium and the Netherlands. The differences in the magnitudes of the effects across the three scenarios now become much larger, and also in absolute terms the shock has a much larger effect on all NUTS (except in specification 2). It now also becomes visible that the border between Belgium and the Netherlands does matter: the effects are on average larger in scenario 3, where this border is assumed away, than in scenario 1, where it exists.

4.3 The Border Effect over Time

After having found that labor market integration is still weaker across borders than within borders in the EU-15 despite the free movement of labor, we now analyze whether labor market integration across borders at least improved over our sample period 1986 to 2006. Specifically, we test whether relative to the within border correlation the across border correlation has increased over time. We do that by allowing for a linear time trend αt in ρ_A. We also allow for a linear time trend βt in ρ_W, in case that there was a change in the overall spatial correlation. The regression includes the full set of controls but no country fixed effects, as in specification (iv) of the baseline results.

	(i) logit(UR) 1986-2006	(ii) logit(UR) 1996-2006
ρ_A	0.024	0.055
	(0.014)	(0.016)
α	-0.0006	-0.0016
	(0.0009)	(0.0025)
ρ_W	0.126	0.143
	(0.005)	(0.005)
β	-0.0016	-0.0052
	(0.003)	(0.0007)
ρ_T	0.906	0.866
	(0.009)	(0.013)
ρ_{AT}	-0.014	-0.040
	(0.014)	(0.016)
ρ_{WT}	-0.091	-0.085
	(0.005)	(0.006)
full set of controls	yes	yes
country fixed effects	no	no
year fixed effects	yes	yes
log-likelihood	859.6	859.6

Note: Standard errors are in parentheses.

Table 4: STAR Model with Linear Time Trends

As the first column of Table 4 shows, there has been no significant time trend in the across-border correlation. α is negative, insignificant, and very close to zero, and thus the across-border correlation has not improved over time in the EU-15 countries between 1986 and 2006. A surprising result comes when we observe the trend in the within-border correlation. β is negative, and much larger than α in absolute terms. This indicates that within-border correlation has actually decreased in the EU over time. This is an effect that plays mainly out in the last 10 years of our sample: in column (ii), where we run the same regression on the shorter sample 1996-2006, β is three times larger in absolute terms than in column (i), and significant. The result that within-border correlation decreased over time is in line with the results of Overman and Puga (2002) that polarization in the EU increased over time.[19]

5 The Role of Policies

While we find that there is no general linear time trend that indicates an increasing cross-border correlation of labor markets in the European Union, we still investigate whether specific policies have improved the cross-border integration of labor markets, namely the abolishment of border controls through the Schengen agreement, as well as the introduction of the Euro.

5.1 The Schengen Agreement

In the Schengen Agreement, originally signed in 1985, five countries of the European Union, namely Belgium, France, Germany, Luxembourg, and the Netherlands, decided on the abolition of border controls. The first countries to implement the abolition of border controls in 1995 were the five original signers plus Portugal and Spain. Italy and Austria followed in 1998, Greece in 2000, and Denmark, Finland and Sweden in 2001 (European Commission, 2010).[20] Border checks could constitute significant impediments to labor market integration. Even if the average time spent at the control station was short, the uncertainty about it could significantly deter daily commuting. To analyze the effect of the abolishment of border controls, we allow for a different across-border correlation if both countries adopted Schengen than if none or only one country adopted Schengen. Due to the different times at which the countries implemented the Schengen agreement, there is a fair amount of spatial and temporal variation in the data.

[19] They analyze the time period 1986 to 1996.

[20] If the Schengen agreement was implemented sometime between January and June, we count the year as the first Schengen-year; if the implementation took place July to December, we count the following year as the first Schengen-year.

To test for an effect of the Schengen agreement, we interact the across-border correlation ρ_A with a Schengen-dummy that is equal to one if both countries sharing the border adopted the Schengen agreement. We call the corresponding coefficient $\alpha_{Schengen}$. Additionally, we interact the within-border correlation with a Schengen dummy if the own country adopted Schengen, calling the corresponding coefficient $\beta_{Schengen}$. We do the latter based on the observation that there seems to be a time trend in within-border correlations.

The main question of interest is whether the Schengen agreement led to an improvement in cross-border correlations, i.e. whether $\alpha_{Schengen} > 0$. In order to control for overall changes in correlation over time, e.g. due to a general increase in the willingness to commute in the Schengen countries, we also investigate whether $\alpha_{Schengen} > \beta_{Schengen}$, i.e. whether the improvement in cross-border correlation in the Schengen countries is larger than the improvement in within-border correlation in the same countries.

	logit(UR) 1986-2006
ρ_A	0.014
	(0.014)
$\alpha_{Schengen}$	0.007
	(0.011)
ρ_W	0.106
	(0.005)
$\beta_{Schengen}$	0.013
	(0.004)
ρ_T	0.904
	(0.009)
ρ_{AT}	-0.014
	(0.014)
ρ_{WT}	-0.093
	(0.005)
full set of controls	yes
country fixed effects	no
year fixed effects	yes
log-likelihood	854.8

Note: Standard errors are in parentheses.

Table 5: STAR Model of the Schengen agreement

The coefficient on $\alpha_{Schengen}$ is positive and around half the size of ρ_A, but not significant. Thus, we cannot reject the hypothesis that the implementation of the Schengen agreement did not increase across-border correlations. Yet, the point estimate goes in the expected direction. Interestingly, the spatial correlation increased significantly within countries that allowed for the Schengen implementation, and in fact $\alpha_{Schengen} < \beta_{Schengen}$. This is in contrast with the general negative trend in the within correlation documented in Section 4.3. The negative trend documented there must arise primarily in non-Schengen countries or pre-Schengen periods. The likelihood ratio test significantly favors the current model compared to a model that does not allow for a Schengen effect. However, given that the increase in within-border correlation past Schengen is larger than the increase in cross-border correlation, and only the former is significant, this seems to be driven by the change in within-border correlation. We conclude that the abolishment of border controls did not lead to a significant increase in the across-border correlations of unemployment rates.[21]

5.2 The Euro Introduction

Eleven member states of the EU-15 fixed their exchange rate irrevocably on January 1st, 1999. Greece joined as the twelfth country one year later. Fixed exchange rates should significantly facilitate both cross-border commuting and supplier relationships across borders. The common Euro coins and notes were then introduced in these twelve countries in 2002, which should further enhance labor market integration, since it abolished the need to exchange currencies. We investigate whether the irrevocable fixing of the exchange rates or the introduction of the Euro coins and notes facilitated labor market integration across borders by allowing for a different cross-border correlation parameter ρ_A if both countries fixed their exchange rate towards the Euro or introduced the Euro coins and notes, respectively. We do this analysis on the shorter sample of all 15 EU countries. While some EU-15 countries, namely the UK, Denmark, and Sweden, did not adopt the Euro until today, the adopting countries all did it at the same time (with the exception of Greece, which however does not contribute to identifying the cross-border correlation). Therefore, there is less variation in the data in this test than in the Schengen test.

[21] As a robustness check, we also estimate a model that allows for different space-time interaction effects after the Schengen agreement. Such a model qualitatively confirms the results of the current model.

As in the previous section, we interact the across-border correlation ρ_A with a Euro-dummy that is equal to one if both countries either irrevocably fixed their exchange rates towards the Euro or introduced the Euro coins and notes, and call the corresponding coefficients $\alpha_{EuroExch}$ and $\alpha_{EuroNotes}$, and interact the within-border correlation with a Euro dummy if the own country either irrevocably fixed the exchange rate towards the Euro or introduced the coins and notes, calling the corresponding coefficients $\beta_{EuroExch}$ and $\beta_{EuroNotes}$. We conduct the tests on both the irrevocable fixing of the exchange rate and the introduction of the coins and notes separately, and analyze as in Section 5.1 whether $\alpha > 0$ and whether $\alpha > \beta$.

	(i) logit(UR) 1996-2006	(ii) logit(UR) 1996-2006
ρ_A	0.051	0.054
	(0.015)	(0.017)
$\alpha_{EuroExch}$	-0.003	
	(0.015)	
$\alpha_{EuroNotes}$		-0.009
		(0.016)
ρ_W	0.131	0.134
	(0.005)	(0.005)
$\beta_{EuroExch}$	-0.020	
	(0.005)	
$\beta_{EuroNotes}$		-0.018
		(0.005)
ρ_T	0.868	0.869
	(0.013)	(0.013)
ρ_{AT}	-0.041	-0.038
	(0.016)	(0.017)
ρ_{WT}	-0.095	-0.095
	(0.006)	(0.006)
full set of controls	yes	yes
country fixed effects	no	no
year fixed effects	yes	yes
log-likelihood	325.4	323.7

Note: Standard errors are in parentheses.

Table 6: STAR Model of Euro introduction

As Table 6 shows, the introduction of the Euro did not lead to a significant change in the across-border correlation, neither the fixing of the exchange rate nor the introduction of the actual coins and notes. $\alpha_{EuroExch}$ and $\alpha_{EuroNotes}$ are both very small and insignificant. The within-country correlation however always significantly decreases after the Euro introduction, in line with the result of decreasing within-border correlation in Section 4.3. Thus, relatively speaking, the cross-border correlation deteriorated less after the Euro introduction than the within-border correlation.[22] As for the Schengen analysis, likelihood ratio tests prefer the current models over one where $\alpha = \beta = 0$, but this seems to be driven mostly by changes in the within-border correlations. Overall, we conclude that the Euro introduction did not lead to a significant improvement in labor market integration across borders, but relatively to within border correlation the across border correlation improved.

6 The Role of Languages

So far, we find that country borders still matter in the European Union when it comes to labor market integration, and that this situation has not changed over time or improved through specific policy measures. This raises the natural question whether the use of different languages impedes labor market integration in Europe. While in the trade of goods English is usually used as the common language, in the daily workplace national languages still play a very dominant role. Are language borders instead of country borders the true impediments to labor market integration?

6.1 Language Borders and Cultural Borders

In order to investigate this question, we exploit the fact that language and country borders do not necessarily coincide in the European Union. For this analysis, we have to assign a language to each NUTS2 region. We use data on languages from the World Language Mapping System of Global Mapping International. It is collected in cooperation with SIL International, which publishes the same information in the *Ethnologue*. These data include a large number of languages (almost 7000 worldwide), and allow for more than one language in a region. The Ethnologue distinguishes between dialects and languages based on spoken intelligibility.[23] The information that

[22] This result is not robust to allowing for changes in the space-time interaction terms for the test based on the fixation of the exchange rates.

[23] Obviously, there are marginal cases in these data.

the Ethnologue collects comes from over 15,000 references from books, journal articles, dissertations, other academic papers about languages and cultures, and from unpublished sources (personal communications, census data, etc.).

In the 15 EU countries considered in our study, the Ethnologue lists 94 distinct languages. Clearly, many of these are spoken only in multi-language areas. How to deal with these multi-language areas is the primary difficulty in using these data. We opt for three different approaches. For all three approaches, using GIS we first assign a language to a NUTS3 region (i.e. the smaller regions than NUTS2) whenever a language is covering more than half of the NUTS3 region in the World Language Mapping System software.[24] Based on population numbers from the NUTS3 region, we then calculate which percentage of the population in the NUTS2 region speaks the respective language. These percentages do not have to add up to 100%. For example, in the Spanish NUTS2 region Castilla-La Mancha, 100% of the population speak Spanish, and 58% speak additionally Extremaduran. In the NUTS2 Comunidad Valenciana, everyone speaks Spanish and additionally Catalan-Valencian-Balear.

In our first approach, we only consider official national languages. In this case, we deal with eleven languages. Some country borders do not coincide with language borders (Luxembourg speaks French, the southern part of Belgium speaks French, the northern part of Belgium speaks Dutch, and Austria speaks German). Belgium is the only country that is divided by a language border.

In our second approach we consider "regional languages". Now, whenever there are two languages spoken by all inhabitants of a NUTS2 (as e.g. in the region Comunidad Valenciana mentioned above), the regional language is picked for the determination of language borders. This captures already more of a cultural effect than a language effect, since if indeed all inhabitants speak the national language as well, language should not constitute a barrier in the labor market. This specification adds six languages to the eleven above, namely Basque, Catalan-Valencian-Balear, Galician, Irish Gaelic, Scots, and Welsh. The new language borders thus all arise in Spain and the UK.

Last, we capture "cultural effects", since cultural hurdles might also matter for economic actions

[24] NUTS3 regions are the smallest regions for which we have population data, allowing us to build population-weighted averages when assigning languages to NUTS2 regions.

(see e.g. Falck et al., 2010, or Guiso et al., 2006). In that set-up, we assign as a language to a NUTS2 any language that is spoken by at least 75% of the population as defined above, and that is not the national or regional language, or any language whose name coincides with the name of the NUTS2.[25] For all NUTS2 for which there is no such language, we assign the regional languages as described above. In this approach, we get 53 languages in the 15 EU countries. While the Ethnologue defines all of these as languages, as opposed to dialects, it is very likely that the speakers might be understood by someone who only knows the national language.[26] Therefore, we likely capture cultural hurdles, rather than true language barriers.

	(i) country borders	(ii) national languages	(iii) regional languages	(iv) cultural effects
ρ_A	0.051	0.017	0.060	0.109
	(0.015)	(0.017)	(0.014)	(0.008)
ρ_W	0.126	0.127	0.126	0.121
	(0.005)	(0.005)	(0.005)	(0.006)
ρ_T	0.866	0.860	0.862	0.870
	(0.013)	(0.013)	(0.013)	(0.013)
ρ_{AT}	-0.042	-0.024	-0.061	-0.086
	(0.016)	(0.018)	(0.015)	(0.009)
ρ_{WT}	-0.098	-0.097	-0.096	-0.094
	(0.006)	(0.006)	(0.006)	(0.007)
full set of controls	yes	yes	yes	yes
country fixed effects	no	no	no	no
year fixed effects	yes	yes	yes	yes
log-likelihood	317.1	327.9	318.0	304.8

Note: Standard errors are in parentheses.

Table 7: STAR Model Allowing for Language Borders

Table 7 shows the results of the estimations with language borders. It uses the 1996 to 2006 data on unemployment. For comparison, specification (i) uses country borders and thus repeats the results from column (iv) in Table 2. When national language borders replace country borders in column (ii), the likelihood rises by 10 points. Also, the correlation across borders becomes smaller and insignificant, while it is significant in the specification with country borders (i). The within

[25] E.g., "Extremaduran" is spoken in the NUTS2 region "Extremadura".
[26] Examples for these languages are Venetian, Veluws, Picard, Alemannisch.

correlation remains largely unchanged. When regional languages are used to determine border effects (column(iii)), the likelihood rises only very slightly above the one in the model with country borders, and remains almost 10 points below the one with national language borders. The model with cultural borders (column (iv)) performs even worse: its likelihood is 12 points below the one with country borders. From this analysis, we conclude that different national languages seem to impede labor market integration in the European Union more than any hurdles associated with crossing country borders. This could also explain why we see no improvement in the labor market integration across countries over time or through policy experiments: the languages stay constant.[27]

We also estimate a model that allows for country and national language borders at the same time. The likelihood ratio test gives as a result that it significantly outperforms the model with only country borders, but does not significantly outperform the model with only national language borders.[28] Thus, adding language borders to country borders matters, while the opposite is not true. We can then also calculate correlation coefficients for all four possible combinations of within and across country and language borders.[29] As Table 8 shows, there is no significant spatial correlation between two units that are separated by a country and a language border. If two adjacent units lie within one country, but do not share a common language (two regions at the language border in Belgium), the point estimate for their spatial correlation doubles, but is still insignificant. On the other hand, if two adjacent units share the same language, but are separated by a country border (as e.g. two regions at the border between Germany and Austria), they exhibit a significantly positive spatial correlation. Thus, country borders matter less than language borders for labor market integration. The highest spatial correlation can be found within country and language borders.

[27] We also run the estimations presented in Table 7 using GDP per capita as the dependent variable. If Luxembourg is included, the model with country borders outperforms the one with national language borders, confirming our intuition above that Luxembourg matters in the model with country borders because of its many across-border neighbors and its outlier value of GDP per capita. When Luxembourg is excluded, models with country and national language borders perform equally well, and both indicate no significant difference in correlation across and within borders.

[28] Results are available from the authors upon request.

[29] Omitting for simplicity of exposition the time terms, we can write the spatial neighborhood matrix in this exercise as $W = \rho_{WC}S_{WC} + \rho_{AC}S_{AC} + \rho_{WL}S_{WL} + \rho_{AL}S_{AL}$, where C stands for country and L stands for language, and the notation is otherwise unchanged. The estimate in the upper left corner of Table 8 corresponds to $\rho_{AL} + \rho_{AC}$, in the upper right corner to $\rho_{AL} + \rho_{WC}$, in the lower left corner to $\rho_{WL} + \rho_{AC}$, and in the lower right corner to $\rho_{WL} + \rho_{WC}$.

		Country	
		Across	Within
	Across	0.015	0.029
Language		(0.017)	(0.025)
	Within	0.113	0.127
		(0.022)	(0.005)

Table 8: Correlations Across and Within Country and Language borders

6.2 The Closeness of Languages

Last, we investigate whether language borders between two languages that have a lot of similarities are more permeable than borders between two languages that are very distinct. Our measure of closeness of two languages comes from the lexicostatistical analysis of Dyen, Kruskal, and Black (1992). The idea behind the closeness measure is to derive the percentage of words that are cognate between two languages. The translations of a word in two languages are "cognate" if within both languages they have an unbroken history of descent from a common ancestral form. Otherwise, the translations are defined as "not cognate".[30] Two hundred commonly used words have been assessed in this way.[31] If there are more than two translations for a word (as there often are), the highest degree of cognation judged between any of the translations is used (i.e. as long as two possible translations in the different languages are judged as cognate, the word is counted as cognate). From these two hundred words, the percentage of cognate words over all words that were assigned either the status of cognate or not cognate is used as a measure of the closeness of the two languages; this measure is called lexicostatistical percentage. E.g. French and Spanish, two romance languages, share 73% cognate words, while French and German, a Germanic language, share only 24%.

Table 9 gives a complete matrix of the measure of closeness of languages used in Europe, where the closeness is transformed to a number such that 0 indicates zero distance and 327 the maximum distance in our data set.[32] Finnish is missing in the analysis by Dyen et al. (1992). One can clearly

[30] Translations that are believed to be related by borrowing or by accidental similarity are thus not treated as cognate. For example, since the English "flower" and the French word "fleur" are known to be related by borrowing, they have been defined as not cognate. There exists a third category, "doubtfully cognate", if it was not possible to establish whether the translations are cognate or not.

[31] Examples for these words include "animal", "to hit", "all", "belly".

[32] The exact transformation is described in Dyen et al. (1992).

	Italian	French	Spanish	Port.	German	Dutch	Swedish	Danish	English	Greek
Italian	0									
French	17	0								
Spanish	19	24	0							
Port.	21	28	10	0						
German	170	188	175	186	0					
Dutch	170	180	164	173	14	0				
Swedish	169	182	168	167	29	30	0			
Danish	168	184	170	175	27	33	11	0		
English	181	187	181	184	48	43	46	45	0	
Greek	273	327	293	315	219	204	237	251	276	0

Table 9: Lexicostatistical Closeness of European Languages

see in the table the closeness among the Romance languages on the one hand and the Germanic languages on the other hand.

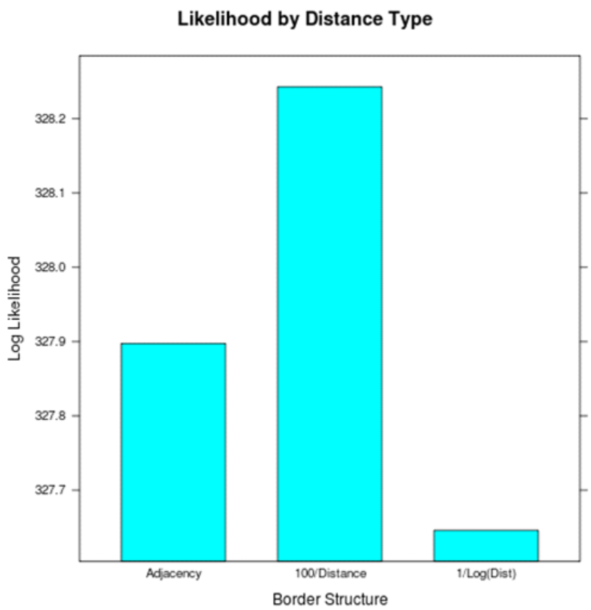

Figure 6: Likelihoods for different language distance specifications

We perform two analysis. In one, we replace the simple adjacency matrix of languages with a weight matrix that takes 100 over the lexicostatistical distance as an input, in the second one we use a weight matrix that takes the inverse of the logarithm of the lexicostatistical distance as an input. In these two adjacency matrices, language borders between two distinct languages are less penetrable than language borders between two similar languages. As Figure 6 shows, the likelihood is higher for the estimation that uses as a weight matrix 100 divided by the lexicostatistical distance than for the weight matrix that uses the inverse of the logarithm of the lexicostatistical distance, or for the simple adjacency weight matrix. However, the differences are very small, amounting to only 0.6 between the two specifications using lexicostatistical distance. Thus, there is only very weak indication that language distances matter. Overall, it seems to be the case that any two different languages impose similar hurdles to labor market integration.

7 Conclusion

In this paper, we analyze whether borders still constitute significant impediments to labor market integration in the EU-15 countries, and find that this is the case. We further investigate whether labor market integration across borders might have changed over time. We find that the cross-border correlation did not increase in absolute terms from 1986 to 2006, but did increase relative to the within-border correlation, which actually decreased over time. Specific policy measures like the Schengen agreement and the Euro introduction did not have a positive effect on our measure of labor market integration.

The lack of improvement of cross-border correlation over time seems to suggest that some permanent forces prevent labor market integration across borders in the European Union. A natural candidate for such a permanent force are languages. Therefore, we analyze whether the true impediments to labor market integration are languages rather than country borders, and find that this is the case. A downside of this analysis is that country and language borders only differ in a few cases, and thus the difference between both is not perfectly identified. Regional languages or cultural effects play less of a role than national languages as impediments to labor market integration. Moreover, language borders are of very similar importance whether two languages are close in a lexicostatistical sense or not.

References

[1] Anderson, James E. and Eric van Wincoop (2003): Gravity with Gravitas: A Solution to the Border Puzzle, American Economic Review, 93(1), 170-192.

[2] Bates, Douglas and Maechler, Martin (2010). Matrix: Sparse and Dense Matrix Classes and Methods. R package version 0.999375-44, http://CRAN.R-project.org/package=Matrix.

[3] Blanchard, Olivier and Lawrence Katz (1992): Regional Evolutions, Brookings Papers on Economic Activity, 1992(1), 1-75.

[4] Borjas, George J. (2001): Does Immigration Grease the Wheels of the Labor Market? Brookings Papers on Economic Activity, 1, 69-133.

[5] Brügger, Beatrix, Rafael Lalive and Josef Zweimüller (2009): Does Culture Affect Unemployment? Evidence from the Röstigraben, IZA Discussion Paper No. 4283.

[6] Cressie, Noel A.C. (1993): Statistics for Spatial Data (revised edition), New York: J. Wiley.

[7] Decressin, Jörg W. and Antonio Fatas (1995): Regional Labor Market Dynamics in Europe, European Economic Review, 39(9), 1627–1655.

[8] Dyen, Isidore, Joseph B. Kruskal, and Paul Black (1992): An Indoeuropean Classification: A Lexicostatistical Experiment, Transactions of the American Philosophical Society, 82(5), 1-132.

[9] European Commission (2010): The European Union Celebrates the 25th Anniversary of the Schengen Area, MEMO/10/249.

[10] Eurostat (2008): European Regional and Urban Statistics. Reference Guide, Eurostat Methodologies and Working Papers.

[11] Falck, Oliver, Stephan Heblich, Alfred Lameli, and Jens Südekum (2010): Dialects, Cultural Identity, and Economic Exchange, IZA Discussion Paper No. 4743.

[12] Fuchs-Schündeln, Nicola, and Rima Izem (2012): Explaining the Low Labor Productivity in East Germany - A Spatial Analysis, Journal of Comparative Economics, 40(1), 1-21.

[13] Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2006): Does Culture Affect Economic Outcomes?, Journal of Economic Perspectives, 20(2), 23-48.

[14] Hunt, Jennifer (2006): Staunching Emigration from East Germany: Age and the Determinants of Migration, Journal of the European Economic Association, 4(5), 1014-1037.

[15] Kelley, Pace, Ronald Barry, Otis Gilley, and C.F. Sirmans (2000): A Method for Spatial-Temporal Forecasting with an Application to Real Estate Prices, International Journal of Forecasting, 16, 229-246.

[16] McCallum, John (1995): National Borders Matter: Canada-U.S. Regional Trade Patterns, American Economic Review, 85(3), 615-623.

[17] Overman, Henry G. and Diego Puga (2002): Unemployment Clusters across Europe's Regions and Countries, Economic Policy, April 2002, 117-147.

[18] Sun, Hua, Yong Tu, and Shi-Ming Yu (2005): A Spatio-Temporal Autoregressive Model for Multi-Unit Residential Market Analysis, Journal of Real Estate Finance and Economics, 31(2), 155-187.

[19] Vandamme, François (2000): Labour Mobility Within the European Union: Findings, Stakes and Prospects, International Labour Review, 139(4), 437-455.

Appendix

A Data Transformations

Data on the educational composition of the work force is missing for Germany, Luxembourg, and the United Kingdom for the year 1998. We impute these data as the average value of the corresponding NUTS2 in the years 1997 and 1999. For very few NUTS-year observations, data on some subgroups that are used to built sums are missing (e.g. data on the number of unemployed with high education might be missing). In these cases, we set the data to zero. This introduces some measurement error. However, the number of such cases is small, only affecting around 0.5% of the overall data.

Some NUTS2 regions are merged to larger ones or split up during our sample period. In these cases, we merge the data from the smaller units into the one from the larger unit, using population-weighted averages (where the population corresponds to the working age population). The UK is most affected by the procedure, since it consists of eleven NUTS2 until 1995, and 37 since then. Affected by this procedure are moreover Belgium (Vlaams-Brabant), Germany (Berlin and Brandenburg), Ireland (Border Midland and Western Ireland, Southern and Eastern Ireland), Italy (Trentino-Alto Adige), the Netherlands (Flevoland), Portugal (Sud), Finland (Etela-Suomi and Lansi-Suomi), and Sweden (Smaland, Vastsverige).

CPI data come from the World Development Indicators for all countries except Germany. For Germany, we use separate CPIs for East and West German regions up to the year 1999 from the German Statistical Office.

B Standardization

As mentioned in Section 3, in the standard STAR model, the rows of the spatial adjacency matrix are typically standardized, making the resulting S row-stochastic. The issue is more complicated in our model (7), which decomposes S into an across-border term S_A and a within-border term S_W. We must decide not only whether to row-standardize S_A and S_W individually, but also whether to row-standardize their sum. To encapsulate these choices we write

$$S = H(\rho_A G(S_A) + \rho_W G(S_W))$$

When G is a function dividing a matrix by its row sums, we call the parameterization *pre-standardized*. When H is such a function, we call it *post-standardized*. When both G and H are null functions (i.e., $G(M) = M \; \forall \; M$), we call the parameterization *non-standardized*.

The different options lead to different implications with regard to the overall dependency of a unit on its spatial neighbors. By standardizing the rows of S, post-standardization fixes the total magnitude of each unit's neighbor weights. This means that a unit with many neighbors will be as tied to their average as a unit with few neighbors. Without post-standardization, a unit's overall level of dependency increases with its neighbor count. The number of neighbors is an imperfect proxy for the potential spillovers between a region and its surroundings.[33]

The issue of whether to post-standardize thus hinges on whether the overall dependency of a unit is increasing in the number of neighbors. Figure 7 shows the root median squared difference of the logit unemployment rate of a NUTS and the average one of its neighbors by the number of neighbors, together with the 95% confidence intervals. The root median squared difference displays a decrease between 1 and 6 neighbors, and then an increase between 6 and 8 neighbors. Hence, the graph does not give a uniform guidance regarding post-standardization. However, with the exception of the NUTS with seven or eight neighbors, assuming that the overall dependency is increasing in the number of neighbors seems reasonable. We therefore opt against post-standardization.[34]

Pre-standardization standardizes the rows of S_A and S_W individually. This fixes the overall level of across- and within-border dependency to the same proportion (ρ_A and ρ_W) for all units. The across- and within-border weight totals are the same regardless of the relative counts of neighbors. In other words, ρ_A is always divided between the set of across-border neighbors and ρ_W between the within-border neighbors. In our application, we quickly observe many flaws with this approach. When there is only a lone across-border neighbor, pre-standardization tends to overweight it. Often a lone across-border neighbor can receive greater weight than each of many within-border neighbors. Consequently, pre-standardization usually leads to extreme residuals around the boundaries. We

[33] The actual spillovers should depend on the percentage of shared kilometers of border, as well as population density close to the border, and connectedness through routes. A region with only one neighbor might be completely surrounded by its single neighbor, or might border the sea, leaving most of its border without any neighbors.

[34] Another problem with post-standardization is numerical stability. When ρ_A and ρ_W are of opposite sign, the row sums of S can be near zero, which leads to blowups in the determinant of the covariance matrix in the likelihood.

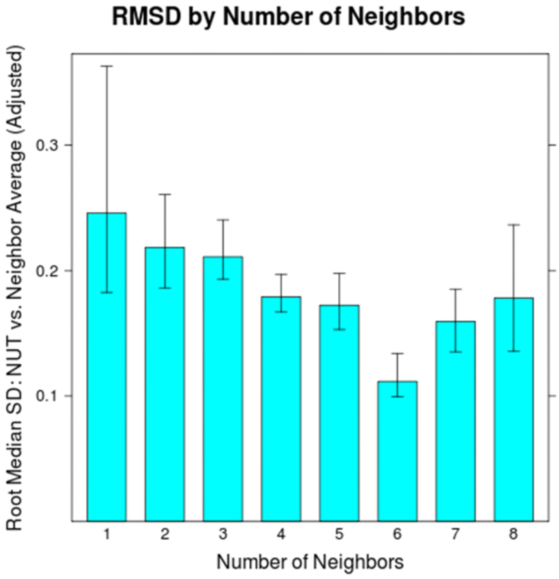

Figure 7: Root Median Squared Difference of the Unemployment Rate of a NUTS region and the average one of its neighbors.

choose not to employ pre-standardization. Having adopted neither form of standardization, we use non-standardization for all later analyses.