

Tani, Massimiliano; Manuguerra, Maurizio

**Working Paper**

## The effect of migration and spatial connectivity on regional skill endowments across Europe: 1988-2010

IZA Discussion Papers, No. 7292

**Provided in Cooperation with:**

IZA – Institute of Labor Economics

*Suggested Citation:* Tani, Massimiliano; Manuguerra, Maurizio (2013) : The effect of migration and spatial connectivity on regional skill endowments across Europe: 1988-2010, IZA Discussion Papers, No. 7292, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/71699>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

IZA DP No. 7292

**The Effect of Migration and Spatial Connectivity on  
Regional Skill Endowments across Europe: 1988-2010**

Massimiliano Tani  
Maurizio Manuguerra

March 2013

# **The Effect of Migration and Spatial Connectivity on Regional Skill Endowments across Europe: 1988-2010**

**Massimiliano Tani**

*Macquarie University  
and IZA*

**Maurizio Manuguerra**

*Macquarie University*

Discussion Paper No. 7292  
March 2013

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## **ABSTRACT**

### **The Effect of Migration and Spatial Connectivity on Regional Skill Endowments across Europe: 1988-2010**

This paper investigates the effects of labour migration and openness to trade on regional skill endowments across the European Union at a time of increased economic integration. Using regional data from Eurostat's Regio database and the open web source Openflight for the period 1998-2010 we test whether the spatial concentration of skills has increased or decreased over time. We account for neighboring effects associated with both geographic proximity and links through civilian flights using a random spatial effect model. We find that migration contributes to convergence in regional skill endowments across member states, particularly at the Southern and Eastern periphery of the European Union. We also find that inter-regional connectivity through civilian flights has much stronger effects on the evolution of a region's skill endowment than geographic contiguity.

JEL Classification: F20, J61

Keywords: European Union, migration, skill endowments, convergence,  
spatial connectivity

Corresponding author:

Massimiliano Tani  
Department of Economics  
Macquarie University  
Sydney, NSW 2109  
Australia  
E-mail: [max.tani@mq.edu.au](mailto:max.tani@mq.edu.au)

## 1 Introduction

More than twenty years have passed since 1986, when the then 12 member states of the European Community signed the Single Act, in which they agreed to establish, by January 1<sup>st</sup> 1992, “an area without frontiers in which the free movement of goods, persons, services and capital is ensured ...”<sup>1</sup> [The Single Act aimed to] “promote [the Community's] overall harmonious development, [and strengthen its] economic and social cohesion”<sup>2</sup>. Since then, steady progress towards the formation of the “single market” has led to the reduction of many barriers to international factor movements with the formation of the European Union (Maastricht Treaty, 1992), the adoption of the euro as a single currency in 1999, and the enlargement of the Union to 15 new member states in 2004 and 2007<sup>3</sup>, with others as official or potential candidates<sup>4</sup>.

The international freedom of movement was expected to help Europeans in working age to reallocate across member states given the significant cross-border differences in average income per capita and unemployment rate, as reported in the studies at that time (e.g. Flanagan, 1993). The introduction of the euro, by making it easier to compare prices internationally, was expected to give further impetus to mobility as a channel to smoothen labour market differences across member states.

However, these expectations have hardly materialised. Over the past 25 years the employment share of European citizens working in a different member state from that of birth or nationality has remained stable at around 2% of total employment. By contrast, that of non-European citizens, to whom the Single Act provisions did not apply, has almost doubled to about 6% in 2010 (Martin and Tyler, 2000; Niebuhr and Schlitte, 2004; Cavelaars and Hessel, 2007). Cross-country surveys further suggest that mobility within the EU mostly happens because of issues related to the quality of life rather than better employment prospect (Recchi, 2008). There is also no material sign of the desired convergence in income and unemployment unambiguously pursued by the EU’s regional policies (Boldrin and Canova, 2001; Corrado et al, 2005; Azomahou et al, 2011; Becker et al, 2010).

---

<sup>1</sup> Article 8A, Single European Act, 1986.

<sup>2</sup> Article 130A, Single European Act, 1986.

<sup>3</sup> These are: Austria, Sweden and Finland in 1995; Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, and Slovenia in 2004; Bulgaria and Romania in 2007.

<sup>4</sup> As of 2012 the official candidates are: Iceland, Macedonia, Montenegro, Serbia, and Turkey. Potential candidates are: Albania, Bosnia and Herzegovina, and Kosovo.

The literature has justified the lack of movement with the finding that Europeans respond to a labour demand shock via changes in participation rather than mobility (e.g. Decressin and Fatas, 2005), though there is mixed evidence on this (e.g. Jimeno and Bentolila, 1998; Fredriksson, 1999; Tani, 2003). Little internal migration has also been associated to the limited portability of pension and other benefits (e.g. Recchi, 2008), and the presence of local economies of scale, which lead to income convergence among super-regions across but not within member states (e.g. Giannetti, 2002; Dall’Erba et al, 2008).

One feature of existing analyses is that they tend to be based on measures of ‘price’, such as income and wages. These measures are however often unadjusted for cross-regional differences in purchasing power and productivity. Yet, these adjustments matter in the locational decisions of workers and firms. As an example Mauro and Spilimbergo (1998) present the cases of two Spanish regions: Barcelona, where skilled labour is abundant, and Extremadura, where it is scarce. They find that real wages for skilled labour are higher in Barcelona: this, a priori, ought to attract further skilled workers and deter employers. But they also find that productivity-adjusted real wages for skilled labour are lower in Barcelona than Extremadura, justifying the apparently puzzling locational choice of firms. Productivity and purchasing-power adjusted data, when available, indeed suggest that regional inequalities are mostly due to productivity differences (e.g. Esteban, 2002; Dall’Erba et al, 2008). This type of data is however unavailable for European regions during the transition to a more integrated economy, compromising the attractiveness of analysing regional economic fortunes.

Nevertheless, it is possible to carry out meaningful complementary analyses based on quantity measures of outcomes. We will focus on the ‘skill intensity’ of a region, as measured by the ratio of skilled to unskilled labour. Aside from the advantage of not depending on productivity and purchasing power information, the analysis of regional inequalities in terms of quantity-based outcome measures offers two additional advantages.

First, price measures typically do not account for the degree of a region’s openness to inter-regional trade. As suggested by Mundell (1975) commodity trade can substitute factor movements. Since not all goods and services are traded uniformly across EU regions, price measures may be distorted by the geographic bias associated with spatial distribution of non-tradable industries. The second advantage is that focusing

on prices informs on changes in the skill composition and, with it, the associated changes in the demand and supply needs of a more/less skilled workforce. The stock of human capital is a well-known source of comparative advantage and long-term economic growth for a locale (e.g. Romer, 1991, Dowrick, 2003). It is also a resource that can be directly ‘managed’ by policymakers, as is the case with investments in local schools. A region of highly educated inhabitants may provide different opportunities to a locale than a region abundant in unskilled labour. Wages and incomes however may remain silent on the ‘quality’ effect of a changing skill composition.

We complement the large volume of existing price-based studies of migration across the EU at a time of progressive economic integration with a study focusing on its effect on the regional skill composition. In so doing, we explicitly take into account the extent of the non-tradable sector, which shelters regions from external shocks. We pay particular care to the geographical and connectivity characteristics of each region. However, in addition to account for proximity to other regions (e.g. Dall’Erba et al, 2008; Ramajo et al, 2008), we create a measure that also accounts for a region’s connectivity to others through civilian air flights. It is a known fact that air links act as a channel for the diffusion of information (e.g. Salt, 1992; Sassen, 1991). Our hypothesis is that they also establish links between regional labour markets albeit more so for some sub-groups of workers, such as those with valuable skills.

The empirical analysis is based on data from Eurostat’s Regio database for the period 1998-2010, as they have the advantage of using comparable definitions across member states. We find that immigrants contribute to the overall convergence of regional skill endowments across the EU during the period, though inter-regional differences have arisen within some member states. We also find statistically significant spatial effects, especially thanks to links by air. We interpret this result as stemming from the emergence of a European labour market for skills, where highly educated workers relocate across cities along air corridors rather than geographic proximity (we use the word ‘skilled’ as a synonym for white collar jobs).

The rest of the paper is organised as follows: section 2 presents a brief summary of the reference literature. Section 3 presents the theoretical model, which draws heavily on Blanchard and Katz (1992). Section 4 discusses the empirical approach. Section 5 introduces the data. Section 6 discusses the empirical analysis and concludes.

## 2 Literature

The literature on regional development within the European Union has studied in detail the issue of convergence or divergence in regional GDP per capita (e.g. Quah, 1996; Overman and Puga, 2002). This addressed the aim of Objective 1 of the EU's Regional Development Fund, the largest structural fund, which is the 'economic adaptation of less developed regions'. From a theoretical point of view, convergence follows the results of the well-known Heckscher-Ohlin international trade theorem. This theorem states that trade of commodities and factors are substitute. Hence reduced barriers to trade, as those experienced in the EU since its creation, lead to factor price equalisation. As a result, regions with low income per capita are expected to grow faster than regions with above-average income per capita. This hypothesis, also known as beta-convergence, further purports that regions differing only in their initial levels of income per capita and capital per worker will converge to the same level of per capital income. Besides economic theory, this is a clear desired outcome of the EU's regional policy.

Economic integration however can also exacerbate initial differences in income. Studies in the so-called New Geography Trade literature suggest that lower transport costs accentuate the relevance of vertical linkages between market access and factor costs amongst firm, leading to agglomeration. As economic integration lowers transport costs from a high initial level, competing forces such as increasing returns to scale (Krugman, 1991) or a high degree of vertical integration between upstream and downstream industries (Venables, 1995) may entice firms to agglomerate. When transport is costless, location does not matter. Regional skill endowments may evolve following a U-shape: rising when regions initially 'merge' from an initial autarky level and subsequently fall as economic integration progresses (e.g. Puga, 2002). Generally, when integration is 'intermediate', agglomeration forces dominate. As skilled workers move to skill-intensive regions, migration is predicted to reinforce initial skill differences. A similar conclusion is reached by theoretical models of endogenous growth, where increasing returns to production from activities such as R&D and high tech are predicted to drive firms to concentrate in space. Regions enjoying high R&D investments or having a large proportion of skilled labour would benefit from further agglomeration.

Empirical price-based studies in the convergence-divergence debate provide mixed results. Beta convergence occurs up to the mid-1990s, after which convergence



disappears. When the regression includes national dummy variables, no convergence emerges, suggesting income inequalities across member states reduce because of the converging activity of a few super-regions, though regional incomes within countries increase (Niebuhr and Schlitte, 2004). The emergence of this result on both ‘old’ member states and those which have accessed the EU since 2004 acts as a strong reminder that economic policies privileging income redistribution (like those currently in place) rather than economic growth do not lead an egalitarian inter-regional convergence in incomes per capita. Hence their pursuit is not justified on economic principles (Boldrin and Canova, 2001).

Quantity-focused macroeconomic-based analyses of labour markets are scarcer than price analyses, and tend to focus on either regional unemployment rates or employment growth. In either case (e.g. Overman and Puga, 2002; Martin and Tyler, 2000; Becker et al, 2010) the evidence of convergence is mixed: EU regions appear to converge until the 1990s, after which divergence prevails. The process of integration started with the Single Act appears to drive regions apart, against the stated objectives of EU’s regional policies.

Aside from the convergence-divergence debate, a different literature stream has instead focused on the dynamic workings of the labour market, providing insights on how local labour markets respond when hit by a shock. The main theoretical model in this area is the work of Blanchard and Katz (1992), who, using a partial equilibrium model (they focus only on one region), investigate the evolution of regional labour markets following shocks to labour demand. Here differences between regions remain constant over time as these are determined by the preferences of firms and workers. Shocks to labour demand or supply can be absorbed by unemployment, changes in participation and migration. Blanchard and Katz find that a shock to labour demand is initially absorbed by unemployment, and over time by migration: workers losing their job do not remain in a state, but move to another, re-equilibrating relative wages and unemployment rates across US states.

An influential study carried out the EU regions (Decressin and Fatas, 1995), shows that during the 1980s they absorbed labour demand shocks through changes in participation. European workers losing a job withdraw from the labour force rather than migrating, as occurs in the US. This finding was confirmed in more recent work

extending the original data to 2003 (Cavelaars and Hessel, 2007)<sup>5</sup>. In a previous article focusing on Spain, it was also found that workers with low levels of educational attainment typically withdraw from the labour force following a negative labour demand shock, while workers with high educational attainment respond by migrating (Mauro and Spilimbergo, 1998).

In this study we address the issue of convergence in regional skill endowments by building on the Blanchard and Katz model, as it provides an ideal framework to analyse the evolution of a regional labour market using a quantity, rather than price, approach. In particular, we extend the Blanchard and Katz model:

- ‘horizontally’, by introducing tradable and non-tradable firms in the labour demand, and by distinguishing between native and foreign workers in the labour supply. These elements will enrich how regional skill endowments react to a shock, such as an increased economic integration, and enable one to understand its effect on sub-components of labour demand and supply;
- ‘vertically’, by expanding the analysis across regions differently endowed with skills, rather than focusing on one only, to determine whether economic integration reduces or expand initial differences in skill endowments in the long-term.

### 3 Theoretical approach

The model formalises a two-sector region  $i$ , which produces two bundles of goods under a constant returns to scale technology. Of the two sectors, one produces tradable goods, while the other sector produces only non-tradable items. The symbols used in the model are identical to those used by the original model of Blanchard and Katz, so that it is possible to follow the different predictions of the extensions presented.

#### *Labour Demand*

The relative labour demand for skilled labour in each region  $i$  at time  $t$  is:

$$w_{it} = \frac{w_{ist} / w_{eust}}{w_{iut} / w_{eunt}} = -d(n_{it} - u_{it}) + z_{it} \quad (1)$$

---

<sup>5</sup> This result however is controversial as does not emerge for a number of member states and EU data during the 1990s. See Jimeno and Bentolila, 1998; Fredriksson, 1999; Mauro and Spilimbergo, 1998; Tani, 2003; Nahus and Parikh, 2001.

where  $w_{it}$  is the logarithm of region  $i$ 's ratio of the skilled ( $s$ ) to the unskilled ( $u$ ) wage relative to the corresponding average ratio across the EU (subscript  $eu$ ) average at time  $t$ . The variable  $n_{it}$  is the logarithm of the regional relative skilled labour force ( $\frac{n_{ist}}{n_{iut}}$ ) relative to the corresponding figure across the EU ( $\frac{n_{eust}}{n_{euut}}$ ). We refer to this

measure as ‘relative skill intensity’ throughout the rest of the paper.

The variable  $n_{it}$  includes natives (N), intra- (I) and extra-EU (E) and hence can be written as:

$$n_{it} = \ln \frac{N_{ist} + I_{ist} + E_{ist}}{N_{eust} + I_{eust} + E_{eust}} - \ln \frac{N_{iut} + I_{iut} + E_{iut}}{N_{euut} + I_{euut} + E_{euut}} \quad (2)$$

A region's relative unemployment rate  $u_{it}$  is defined as:

$$u_{it} \equiv \left[ \left( \frac{UN_{ist}}{EM_{ist}} \right) / \left( \frac{UN_{iut}}{EM_{iut}} \right) \right] / \left[ \left( \frac{UN_{eust}}{EM_{eust}} \right) / \left( \frac{UN_{euut}}{EM_{euut}} \right) \right] \quad (3)$$

where  $UN_{it}$  and  $EM_{it}$  represent a region's total number of unemployed and employed (i.e. native plus intra- and extra-EU), respectively. This definition of unemployment implies that the difference ( $n_{it} - u_{it}$ ) in (1) is approximately equal to the logarithm of relative employment<sup>6</sup>. It is assumed that  $d > 0$ .

The variable  $z_{it}$  denotes the overall number of firms in the region relative to the corresponding number across the EU, and is defined as:

$$z_{it} = \ln \frac{\bar{z}_{it}^T + \bar{z}_{it}^{NT}}{\bar{z}_{eut}^T + \bar{z}_{eut}^{NT}} \quad (4)$$

where superscripts  $T$  and  $NT$  indicate ‘tradable’ (meaning goods and services that can be transported across regions) and ‘non-tradable’, respectively. In turn, the tradable and non-tradable sectors hire labour according to:

$$\Delta \bar{z}_{it}^T = \ln \frac{\bar{z}_{it+1}^T}{\bar{z}_{eut+1}^T} - \ln \frac{\bar{z}_{it}^T}{\bar{z}_{eut}^T} = -a_T w_{it} + X_i^{dT} + \varepsilon_{it+1}^{dT} \quad (4a)$$

$$\Delta \bar{z}_{it}^{NT} = \ln \frac{\bar{z}_{it+1}^{NT}}{\bar{z}_{eut+1}^{NT}} - \ln \frac{\bar{z}_{it}^{NT}}{\bar{z}_{eut}^{NT}} = -a_{NT} w_{it} + X_i^{dNT} + \varepsilon_{it+1}^{dNT} \quad (4b)$$

---

<sup>6</sup> If UN, EM and LF denote the numbers of unemployed, employed and those in the labour force, then  $u_s / u_u = (UN/EM)_s / (UN/EM)_u \approx \ln(1 + UN/EM)_s / \ln(1 + UN/EM)_u = (\ln(LF) - \ln(EM))_s / (\ln(LF) - \ln(EM))_u$ . Hence  $(n_s/n_u - u_s/u_u) \approx (\ln(LF) - \ln(LF) + \ln(EM))_s / (\ln(LF) - \ln(LF) + \ln(EM))_u = \ln(EM)_s / \ln(EM)_u$ .

where the parameter  $a$  is the sensitivity of firms to the region's relative wage. By assumption  $a_T > a_{NT} > 0$ , to highlight that tradable firms are more sensitive to relative labour costs: they have a higher elasticity of relative skilled labour demand with respect to the regional relative wage than non-tradable firms.

The variables  $X_i^{dT}$  and  $X_i^{dNT}$  represent the relative attractiveness of a region to tradable and non-tradable firms (for simplicity constant over time), and  $\varepsilon_{it+1}^{dT}$  and  $\varepsilon_{it+1}^{dNT}$  are white noise stochastic processes representing unexpected changes in technology, the bundle of goods produced, and their relative prices, including transportation costs.

We assume that tradable and non-tradable firms do not distinguish between native, intra- and extra-EU labour<sup>7</sup>. As long as regional relative skilled wages are below their long-run equilibrium level, firms using skilled labour intensively will move in, and vice-versa.

### *Labour Supply*

The relative demand for native, intra- and extra-EU in a region labour evolves according to the dynamics below:

$$\Delta \tilde{N}_{it} = \ln \frac{N_{ist+1}}{N_{eust+1}} - \ln \frac{N_{ist}}{N_{eust}} - \left( \ln \frac{N_{iut+1}}{N_{euit+1}} - \ln \frac{N_{iut}}{N_{euit}} \right) = b_1 w_{it} - g_1 u_{it} + X_i^s + \varepsilon_{1t+1}^s \quad (5a)$$

$$\Delta \tilde{I}_{it} = \ln \frac{I_{ist+1}}{I_{eust+1}} - \ln \frac{I_{ist}}{I_{eust}} - \left( \ln \frac{I_{iut+1}}{I_{euit+1}} - \ln \frac{I_{iut}}{I_{euit}} \right) = b_2 w_{it} - g_2 u_{it} + X_i^s + \varepsilon_{2t+1}^s \quad (5b)$$

$$\Delta \tilde{E}_{it} = \ln \frac{E_{ist+1}}{E_{eust+1}} - \ln \frac{E_{ist}}{E_{eust}} - \left( \ln \frac{E_{iut+1}}{E_{euit+1}} - \ln \frac{E_{iut}}{E_{euit}} \right) = b_3 w_{it} - g_3 u_{it} + X_i^s + \varepsilon_{3t+1}^s \quad (5c)$$

where  $b_1$ ,  $b_2$ , and  $b_3$  are the labour supply elasticities of skilled natives, intra- and extra-EU (by assumption:  $b_1 < b_2 < b_3$ ), while  $g_1$ ,  $g_2$ , and  $g_3$  (by assumption:  $g_1 < g_2 < g_3$ ) are their responsiveness to the regional relative skilled unemployment rate  $u_{it}$ . It is also assumed that unskilled foreign workers supply labour more elastically than unskilled natives.

The parameter  $X_i^s$  denotes a region's relative attractiveness to workers (e.g. clean environment, climate, taxation, access to public goods and other institutional factors), which, for simplicity, is assumed to be identical for all people, and time-invariant. For

---

<sup>7</sup> It is possible to further differentiate tradable and non-tradable firms by augmenting equation (4b) with a term:  $+\gamma \Delta n_{it}$  (with  $1 > \gamma > 0$ ) to highlight the skill-bias of the non-tradable sector: the higher the relative share of skilled workers in the region, the more non-tradable firms will find it attractive to establish there. This assumption would reflect the observation that non-tradable services generally cater for people with higher incomes, viewed as a proxy for skill level.

a given  $X_i^s$ , skilled natives, intra- and extra-EU people emigrate if the regional skilled relative wage is below its long-run equilibrium level. The parameters  $\varepsilon_{1t+1}^s$ ,  $\varepsilon_{2t+1}^s$ , and  $\varepsilon_{2t+1}^s$  are white noise stochastic components reflecting unexpected changes in the relative skilled labour supply for each national group, such as a new migration law introducing a minimum educational requirement or specific occupational experience. Finally, the regional skilled relative wage is assumed to follow unemployment through:

$$w_{it} = -\frac{u_{it}}{c} \quad (6)$$

where  $c > 0$ .

We first transform the aggregate regional relative employment  $n_{it}$  and the relative regional number of firms  $z_{it}$  in terms of their components, exploiting the fact that:

$$\ln \frac{a+b}{A+B} = \frac{1}{2} \left[ \ln \left( \frac{a}{A} \right) \right] + \frac{1}{2} \left[ \ln \left( \frac{b}{B} \right) \right] - \frac{1}{2} \left[ \frac{\left( \frac{a}{a+b} \right)}{\left( \frac{A}{A+B} \right)} \right] - \frac{1}{2} \left[ \frac{\left( \frac{b}{a+b} \right)}{\left( \frac{B}{A+B} \right)} \right]$$

to get:

$$\begin{aligned} \Delta n_{it} = & \frac{1}{2} \Delta \tilde{N}_{it} + \frac{1}{4} \Delta \tilde{I}_{it} + \frac{1}{4} \Delta \tilde{E}_{it} - \frac{1}{2} \Delta \tilde{\theta}_{ist}^N - \frac{1}{2} \Delta (1 - \tilde{\theta}_{ist}^N) - \frac{1}{4} \Delta \tilde{\theta}_{ist}^I - \frac{1}{4} \Delta (1 - \tilde{\theta}_{ist}^I) \\ & + \frac{1}{2} \Delta \tilde{\theta}_{iut}^N + \frac{1}{2} \Delta (1 - \tilde{\theta}_{iut}^N) + \frac{1}{4} \Delta \tilde{\theta}_{iut}^I + \frac{1}{4} \Delta (1 - \tilde{\theta}_{iut}^I) \end{aligned} \quad (7)$$

$$\Delta z_{it} = \frac{1}{2} \Delta \tilde{z}_{it}^T + \frac{1}{2} \Delta \tilde{z}_{it}^{NT} - \frac{1}{2} \Delta \tilde{\theta}_{it}^T - \frac{1}{2} \Delta (1 - \tilde{\theta}_{it}^T) \quad (8)$$

where:

$$\begin{aligned} \tilde{\theta}_{ist}^N &= \ln \frac{\frac{N_{ist}}{N_{ist} + I_{ist} + E_{ist}}}{\frac{N_{eust}}{N_{eust} + I_{eust} + E_{eust}}} \\ (1 - \tilde{\theta}_{ist}^N) &= \ln \frac{\left(1 - \frac{N_{ist}}{N_{ist} + I_{ist} + E_{ist}}\right)}{\left(1 - \frac{N_{eust}}{N_{eust} + I_{eust} + E_{eust}}\right)} \\ \tilde{\theta}_{ist}^I &= \ln \frac{\frac{I_{ist}}{I_{ist} + E_{ist}}}{\frac{I_{eust}}{I_{eust} + E_{eust}}} \\ (1 - \tilde{\theta}_{ist}^I) &= \ln \frac{\left(1 - \frac{I_{ist}}{I_{ist} + E_{ist}}\right)}{\left(1 - \frac{I_{eust}}{I_{eust} + E_{eust}}\right)} \\ \tilde{\theta}_{it}^T &= \ln \frac{\left(\frac{z_{it}^T}{z_{it}^T + z_{it}^{NT}}\right)}{\left(\frac{z_{eut}^T}{z_{eut}^T + z_{eut}^{NT}}\right)} \\ (1 - \tilde{\theta}_{it}^T) &= \ln \frac{\left(1 - \frac{z_{it}^T}{z_{it}^T + z_{it}^{NT}}\right)}{\left(1 - \frac{z_{eut}^T}{z_{eut}^T + z_{eut}^{NT}}\right)} \end{aligned}$$

### Equilibrium Paths

The equilibrium paths of regional skilled relative wage, unemployment rate, and labour force growth are derived from the intersection of relative labour demand and supply curves. Natives, intra- and extra-EU workers face identical equilibrium paths with respect to the regional skill relative wage and unemployment rate. These are respectively:

$$w_{it+1}^* = \frac{1 + dc - dR - Q}{1 + dc} w_{it} + \frac{(X_i^{Td} + X_i^{NTd}) - dX_i^s}{1 + dc} + \frac{1}{1 + dc} (dO_i - J_i^T + J_i^d - dJ_i^s) \quad (9)$$

$$u_{it+1}^* = -dw_{it+1}^* \quad (10)$$

where:

$$R = 1/2(b_1 + cg_1) + 1/4(b_2 + cg_2) + 1/4(b_3 + cg_3)$$

$$Q = 1/2a_T + 1/2a_{NT}$$

$$O_t = \frac{1}{2}(\Delta\tilde{\theta}_{ist}^N - \Delta\tilde{\theta}_{iut}^N) + \frac{1}{2}(\Delta(1 - \tilde{\theta}_{ist}^N) - \Delta(1 - \tilde{\theta}_{iut}^N)) + \frac{1}{4}(\Delta\tilde{\theta}_{ist}^I - \Delta\tilde{\theta}_{iut}^I) + \frac{1}{4}(\Delta(1 - \tilde{\theta}_{ist}^I) - \Delta(1 - \tilde{\theta}_{iut}^I))$$

$$J_t^T = \frac{1}{2}(\Delta\tilde{\theta}_t^T - \Delta\tilde{\theta}_t^{NT})$$

$$J_t^d = \frac{1}{2}\varepsilon_{it+1}^{dT} + \frac{1}{2}\varepsilon_{it+1}^{dNT}$$

$$J_t^s = \frac{1}{2}\varepsilon_{1t+1}^s + \frac{1}{4}\varepsilon_{2t+1}^s + \frac{1}{4}\varepsilon_{3t+1}^s$$

Expression (9) suggests that the equilibrium path of skilled relative wages depends on the regional relative attractiveness for firms (positively) and workers (negatively). It also depends on changes in the relative shares of skilled natives and foreign citizens (positively) and the relative openness of the region to trade (negatively).

The equilibrium path for the regional relative employment growth for skilled labour is given by:

$$\Delta n_{it+1}^* = \frac{1+dc-dR-Q}{1+dc}\Delta n_{it} + \frac{Q}{1+dc}X_i^s + \frac{R}{1+dc}(X_i^{Td} + X_i^{NTd}) - \frac{RJ_t^T}{1+dc} + dO_{t+1} + \frac{d(1+dc-dR-Q)-dR}{1+dc}O_t + \frac{R}{1+dc}(J_t^d - dJ_t^s) + J_{t+1}^s \quad (11)$$

The equilibrium paths for native, intra- and extra-EU relative employment growth can be found by replacing the equilibrium path for wages described by equation (9) into equations (5a)-(5c), respectively, as shown in the Appendix.

Equation (11) formalises the hypothesis that the skill intensity of a region is less persistent over time if the region's labour force includes migrants (R) and its industries include tradable firms (Q).

### *Europe's Economic Integration*

We assume that economic integration occurring since the Single Act takes the form of a positive shock to the labour demand for tradable firms:  $\varepsilon_{it+1}^{dT} > 0$ . This choice tries to capture that the enlargement of the potential market is more likely to affect tradable vis-à-vis non-tradable firms, as the former are more susceptible to labour costs. We study the effect of this shock using comparative statics on two regions that are differently endowed with skills keeping constant their level of attractiveness. To do so we rewrite equations (9) and (11) as:

$$w_{it+1}^* = \sum_{i=0}^{\infty} \lambda^i \left\{ \frac{(X_i^{Td} + X_i^{NTd}) - dX_i^s}{1+dc} + \frac{1}{1+dc} (dO_t - J_t^T + J_t^d - dJ_t^s) \right\}_{t-i}$$

$$\Delta n_{it+1}^* = \sum_{i=0}^{\infty} \lambda^i \left\{ \frac{Q}{1+dc} X_i^s + \frac{R}{1+dc} (X_i^{Td} + X_i^{NTd}) - \frac{RJ_t^T}{1+dc} + dO_{t+1} + \left( d\lambda - \frac{dR}{1+dc} \right) O_t + \frac{R}{1+dc} (J_t^d - dJ_t^s) + J_{t+1}^s \right\}_{t-i}$$

where  $\frac{1+dc-dR-Q}{1+dc} = \lambda$ .

The positive shock to tradable firms in regions with abundant unskilled workers

affects wages through  $\frac{\partial w_{it+1}^*}{\partial \varepsilon_{it}^{Td}} = \sum_{i=0}^{\infty} \lambda^i \left\{ \frac{1}{2(1+dc)} \right\}_{t-i}$  and relative skilled employment

*growth* through  $\frac{\partial \Delta n_{it+1}^*}{\partial \varepsilon_{it}^{Td}} = \sum_{i=0}^{\infty} \lambda^i \left\{ \frac{R}{1+dc} \right\}_{t-i}$ . Both effects tend to zero as time goes on, as

$\lambda < 1$ , resulting only in temporary effects.

The effect of  $\varepsilon_{it+1}^{dT} > 0$  on the relative skilled employment *level* however is permanent, and it is given by:

$$\begin{aligned} \frac{\partial n_{it+1}^*}{\partial \varepsilon_{it}^{Td}} &= \frac{1 - \sum_{i=0}^{\infty} \lambda^i}{1 - \lambda} \left\{ \frac{Q}{1+dc} X_i^s + \frac{R}{1+dc} (X_i^{Td} + X_i^{NTd}) - \frac{RJ_t^T}{1+dc} + dO_{t+1} + \left( d\lambda - \frac{dR}{1+dc} \right) O_t + \frac{R}{1+dc} (J_t^d - dJ_t^s) + J_{t+1}^s \right\}_{t-i} = \\ &= \frac{1}{1 - \lambda} \frac{R}{2} = \frac{R(1+dc)}{2(dR+Q)} > 0 \end{aligned}$$

The relative skill endowment of the region, originally abundant in unskilled labour, has changed to one where there are more skilled workers than before. The effect of integration is larger for regions abundant in unskilled labour, as they are the ones receiving the main influx of tradable firms' investments. With limited scope for dramatic changes in skill composition elsewhere, pre-integration regional skill endowments will converge.

A narrative summary is as follows: as a result of the new 'single market', tradable firms in each region will face a higher demand for their products, and those using intensively the cheaper inputs and able to do so (typically large firms) will relocate to the regions where unskilled labour is abundant. The increase in tradable firms in regions rich in unskilled labour will bring with it also an increased demand for medium and high skilled labour. Research on technological change has repeatedly reported that new technology, such as the one used in new factories, is accompanied



by a bias towards hiring skilled employees. Investment in new plants will be carried out with more modern processes, as well as ancillary services related to the new factories' supply chain. This is likely to raise the demand for skilled labour, attracting foreign immigrants, both skilled and unskilled.

In the case of skill-abundant regions there is no such effect, as the 'single market' will generate an increase in the demand for skilled labour from trading firms using skills intensively (e.g. consulting services) as well as the relocation of some non-tradable firms. These developments are unlikely to be accompanied by large-scale investment in infrastructure and ancillary services as those occurring in regions abundant in unskilled labour. Furthermore, high relative wages for unskilled labour will act as a magnet for unskilled foreign workers, whose relocation will likely lower the skill-rich region's skill intensity.

The transition to the new equilibrium for skilled and unskilled labour rich regions may be accompanied by a temporary convergence in their relative regional wages. Over time however, initial differences in regional wages will return to their original trends if, as empirical evidence supports, there is persistence in regional wages and unemployment (e.g. Decressin and Fatas, 2005). A researcher measuring the long-term effects of the 'single market' would not be able to report any long-run trends towards wage convergence. In contrast, the skill intensity in the two regions would record the changes in the local skill composition, and show a higher similarity over time: the single market would have led to converging regional skill endowments, with lasting effects on each region's local demand for public goods as well as tax revenue base<sup>8</sup>.

#### 4 Empirical approach

To estimate the effect of the European integration on regional skill intensities, we estimate the evolution of regional relative employment (equation (11)) using the following functional form:

$$\Delta n_{it+1}^* = \alpha_1 \Delta n_{it} + \alpha_2 X_i^s + \alpha_3 X_i^d + \alpha_4 (\Delta \% T_t + \Delta \% NT_t) + \alpha_5 (\Delta \% N_{t+1} + \Delta \% I_{t+1}) + \alpha_6 (\Delta \% N_t + \Delta \% I_t) + error_{it} \quad (15)$$

---

<sup>8</sup> Further predictions of the model, not discussed in this paper, include that foreigners will fill the jobs created more than proportionally relative to their share in the destination region, due to their high elasticity to supply labour relative to natives. In particular, there will be more extra-EU than intra-EU and natives in recently opened jobs, as a higher incidence of over-education due to extra-EU's higher labour supply elasticity.

The sign and statistical significance of the parameter  $\alpha_1$  inform on whether or not regional skill endowments have converged. In particular convergence has occurred if  $\alpha_1 < 1$ .

In estimating equation (15) we are conscious of possible neighboring effects: namely that a region's skill intensity depends also on the geographic location of the region and its degree of connectivity with other regions, both within the member state to which it administratively belongs as well as across the EU (e.g. Ramajo et al, 2008; Dall'Erba et al, 2008; Martin, 2001; Canova, 2004; Corrado et al, 2005). If this is the case, then the error term in equation (15) contains spatial elements that may be related to nearby regions as well as the other covariates (e.g. a region's location would influence its openness to trade and presence of tradable firms). Regional skill endowments may not evolve independently across space after all. As a result, we cater for spatial autocorrelation using a random spatial effects model. However, rather than limiting our attention to the geographic neighbors of a region, we construct the region-specific component that reflects its location as a combination of a geographic element and the connectivity established by civilian air links. We recognise that migrations do not necessarily respond to regions' geographic contiguity. In the case of skilled labour the pull and job opportunities arising from relatively large cities with good infrastructure and communication and transport connections are also likely to play a pivotal role in the decision of where to locate (e.g. Gaspar and Glaeser, 1996; Sassen, 1991). We hence add to the physical distance, traditionally captured by gravity models, the neighboring effect arising from connectivity by air links. We do so by applying well-established statistical techniques.

In the past two decades, conditional autoregressive (CAR) models have been widely used to analyze areal data (Besag, 1974). In this study we apply them to a high dimensional case, where standard estimation techniques converge slowly. In particular, we fit equation (15) with a Bayesian Adaptive Independence Sampler (BAIS) (Keith, 2008). The algorithm uses multiple parallel chains and periodically updates the proposal distribution based on population of current elements across all chains, which accelerates convergence and mixing properties. The CAR model we apply is described elsewhere (see for example Carlin, 2003; Gelfand, 2003), but can be summarised as follows: following Besag (1974), the CAR specification for a region  $i$  with  $N_i$  the set of neighbors at time  $t$  is:

$$y_{t,i} | y_{t-1,i} \sim N\left(\mu_{t,i} + \sum_{j \in N_i} b_{ij}(y_{t,j} - \mu_{t,j}), \tau_i^2\right), \quad (16)$$

where  $b_{ij}$  are some weights. Following Brook's lemma in Brook (1964) the formulation in (16) can be conveniently rewritten in matrix form as:

$$y_{t,i} | y_{t-1,i} \sim N\left(\mu_{t,i}, [M_{\tau_i}^{-1}(I - B)]^{-1}\right)$$

where  $M_{\tau_i} = \text{diag}(\tau_i^2)$  and  $(B)_{ij} = b_{ij}$ , with  $b_{ii} = 0$ . The requirement of  $M_{\tau_i}^{-1}(I - B)$  symmetric yields the conditions  $b_{ij}\tau_j^2 = b_{ji}\tau_i^2$ , which in a popular CAR model (Besag, 1991) are satisfied involving only adjacency relationships and imposing  $\tau_i^2 = \sigma^2 / w_{i+}$  and  $b_{ij} = w_{ij} / w_{i+}$ .  $W$  is the adjacency matrix and  $w_{i+}$  are the  $i$ -row sums of  $W$ . Defining the scaled adjacency matrix  $W_s = \text{diag}(w_{i+})$ , the  $CAR(1, \sigma^2)$  model formulation can be written as:

$$y_{t,i} | y_{t-1,i} \sim N\left(\mu_{t,i}, \left[\frac{1}{\sigma^2}(W_s - W)\right]^{-1}\right). \quad (17)$$

To avoid the possibility of a singular  $(W_s - W)$  in the  $CAR(1, \sigma^2)$  formulation, making (17) improper, we introduce a parameter ( $\rho$ ) which lies between  $-1$  and  $1$ , and controls the spatial dependence (Carlin and Banerjee, 2003):

$$y_{t,i} | y_{t-1,i} \sim N\left(\mu_{t,i}, \left[\frac{1}{\sigma^2}(W_s - \rho W)\right]^{-1}\right)$$

The adjacency matrix then has a component accounting for physical contiguity ( $W_a$ ), and another component accounting for civilian flight connections ( $W_f$ ). The resulting adjacency matrix is then  $W = W_a + \eta W_f$ , where  $\eta$  measures the relevance of flight connections relative to pure geographic proximity.

Finally, the joint spatio-temporal process  $Y$  can be factored into:

$$[Y | \beta] = [y_0] \prod_{t=1}^T [y_t | y_{t-1}, \beta]$$

where  $y_t = (y(t,1), \dots, y(t,n))'$ . We estimate equation (15) by maximum likelihood.

## 5 Data

Most data used in the empirical analysis are extracted from Eurostat's Regio database ([http://epp.eurostat.ec.europa.eu/portal/page/portal/region\\_cities/regional\\_statistics/data/database](http://epp.eurostat.ec.europa.eu/portal/page/portal/region_cities/regional_statistics/data/database)), which has the advantage of collecting information using the same definitions across all member states, while information on air flight connectivity is obtained by the

open source Openflight (<http://openflights.org>). We use regional data at NUTS 2 level for the period 1988-2010. From the initial 4525 data points covering each region across the current 27 member states of the EU, we extract a relatively complete panel of 163 regions across 15 member states.

Table 1 presents a statistical summary of the variables used in the estimations.

**TABLE 1 SUMMARY STATISTICS: 1988-2010**

<b>Variable</b>	<b>Observations</b>	<b>Mean</b>	<b>STD</b>
<i>Dependent variable</i>			
Relative skill intensity	3740	.0012	.0774
<i>Controls</i>			
Share tradables	3707	.00217	.1023
Value added	2042	92.616	47.149
Unemployment rate	2501	100.605	52.516
Participation rate	2321	99.151	9.760
Labour force	4548	2007.48	2093.61
Life expectancy	3103	100.426	3.2096
GDP per capita (PPP)	2628	96.067	36.637

The relative skill intensity formalised in Equation (2) is calculated as the ratio of the skilled/unskilled ratio of a region divided by the corresponding ratio for the EU. The definition of ‘skilled’ and ‘unskilled’ reflect occupational aggregates as suggested by Keesing (1966). Namely, the top three categories of the ISCO occupational scale at 1-digit level (ISCO 1-3) are considered as skilled, and the bottom three categories (ISCO 7-9) are defined as unskilled. Alternative indices (e.g. adding the top three categories and half of those in the three intermediate skill groups ISCO 4-6, and dividing it by the bottom three categories plus the remaining 50% of ISCO 4-6) do not change the spatial distribution of the relative skill intensity obtained using only the top and bottom three ISCO categories.

The relative skill intensity varies significantly across regions and time, as the data include very different geographic areas in economic structure, history and level of development. The data also cover a heterogeneous period, which includes the introduction of the euro, the disintegration and subsequent absorption into the EU of countries of the former Soviet block, and the dramatic economic effects of the Global Financial Crisis. Notwithstanding the volatile historical period, high skill intensity generally features in the high-income urban areas of Northern Europe (South East of England, Ile de France, Brussels, Hamburg, Nord-Rheinland-Pfaltz, Westfalia,

Flanders). By contrast, low skill intensity is common amongst rural or tourist regions along the Mediterranean coast, especially in Greece, Portugal, and Spain.

The variability and evolution of relative skill intensity during the period is depicted in Figure 1 below.

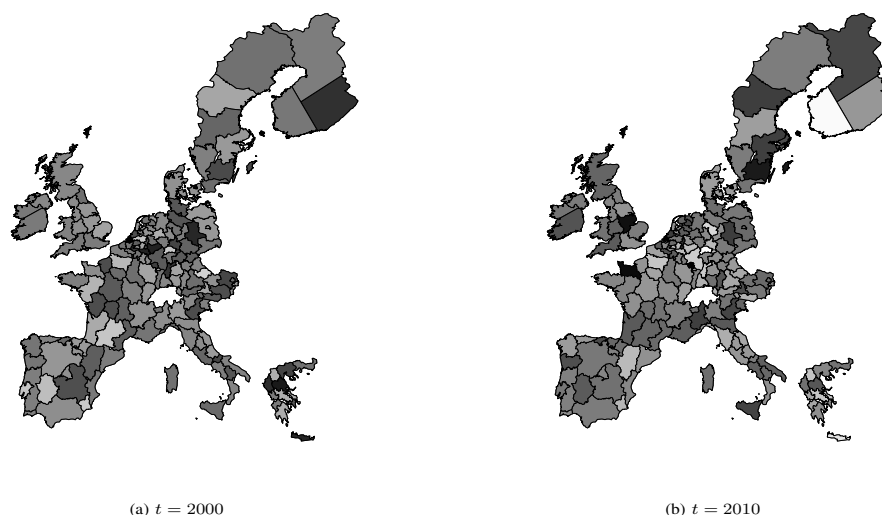


Fig. 1. Growth of the ratio between skilled and unskilled workers in the EU in 2000 and 2010

As can be seen, the relative skill intensity has not been geographically persistent: the regions with the largest changes in 2000 have experienced more modest changes ten years later.

Table 1 also reports the control variables used. Aside from the size, proxied by the labour force, the other variables are measured relative to the EU average, which is either centred at zero for growth variables (share of tradable industries) or scaled to 100 (unemployment and participation rates, GDP per capita at PPP, value added, and life expectancy). Values away from zero or 100 reflect outliers and prevalence. As a result, there seems to be a slight prevalence of regions with higher unemployment (100.605) and lower participation (99.151) than the EU average<sup>9</sup>. There seem to be a prevalence of regions where life expectancy is longer than the average (100.426) as well as lower GDP per capita at constant PPP (96.067). The average EU region has

---

<sup>9</sup> Regional fortunes, as proxied by the unemployment rate, vary significantly both across space (low in Northern Europe; high in the South), though the average for the period is a relatively high 8.9%. Participation also varies significantly both across space and time: it tends to be higher in Northern member states, especially the UK and Denmark, while it is much lower in places like Italy, Spain and Greece where women traditionally are involved in housework and raising children, and where the retirement age tends to be low.

about 2 million inhabitants of working age (18-64). We use an index for value added to account for different regional industrial structures.

The information on civilian air transportation between regions is used with geographic proximity to generate the adjacency matrix  $W$ , which is used to account for spatial effects. The strength of the link between pairs of regions based on civilian flights is shown in Fig.2 below. The regions are ordered by country and darker dots correspond to a greater value of the connection. It is interesting to see how well some regions are connected with other European regions.

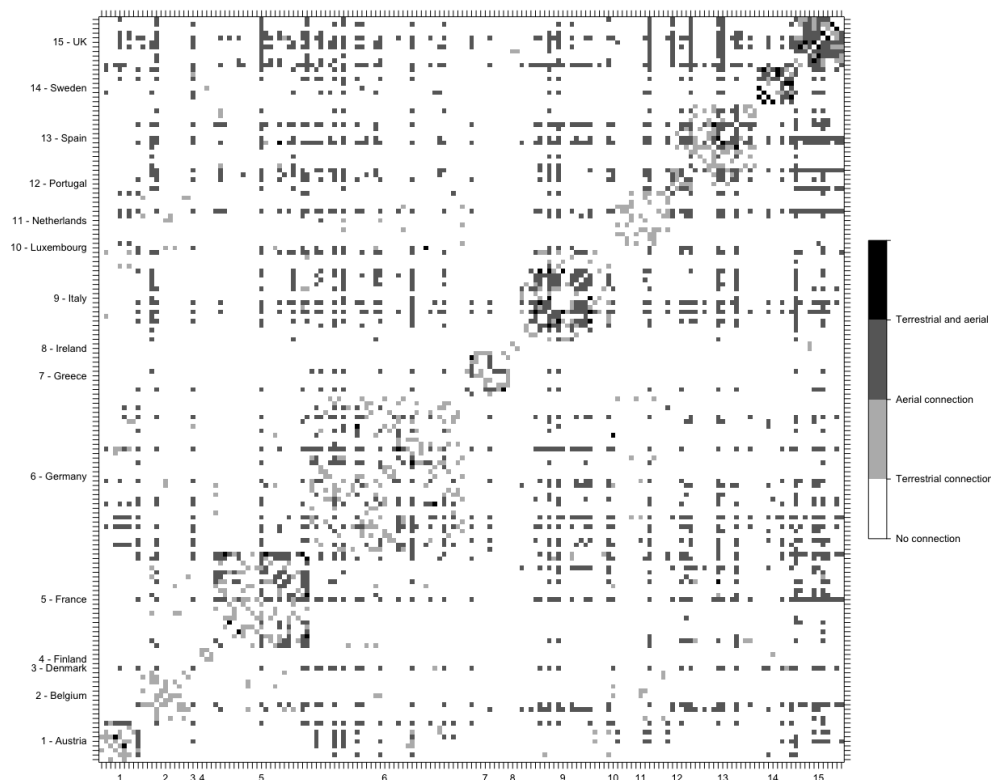


Fig. 2. Visual representation of the proximity matrix  $W$ .

Fig. 2 highlights that air links mainly connect regions within a country (along the diagonal), while only a handful of regions are also well connected to locales in other member states.

Table 2 summarises an index of the strength of the air link of each of the regions in the sample. The most connected regions are the East Anglia ( $W$ : 169.5), Cataluna (166.5), Lazio (158.5), the South East of England and Ile de France (157.5 each), Madrid (144.5), Noord Holland and Lombardia (141.5 each). The least connected regions include several regions at the geographic periphery of the EU.

**Table 2 The Strength of Air Connections across the European Union**

REGION'S NAME	CONNECT	COUNTRY	REGION'S NAME	CONNECT	COUNTRY	REGION'S NAME	CONNECT	COUNTRY	REGION'S NAME	CONNECT	COUNTRY
BURGENLAND	2	Austria	ARNSBERG	19.5	Germany	ABRUZZO	20.5	Italy	ANDALUCIA	122.5	Spain
KARNTEN	25	Austria	BERLIN WEST BERLIN EAST	78.5	Germany	BASILICATA	3	Italy	ARAGON	25.5	Spain
NIEDEROESTERREICH	4	Austria	BRANDENBURG	81.5	Germany	CALABRIA	18.5	Italy	ASTURIAS	3	Spain
OBEROESTERREICH	20	Austria	BRAUNSCHWEIG	6	Germany	CAMPANIA	54	Italy	CANTABRIA	23	Spain
SALZBURG	53.5	Austria	BREMEN	40.5	Germany	EMILIA ROMAGNA	76	Italy	CASTILLALA MANCHA	7	Spain
STEIERMARK	27.5	Austria	DARMSTADT	127.5	Germany	FRIULI	22	Italy	CASTILLA LEON	26	Spain
TIROL	37	Austria	DETMOLD	11	Germany	LAZIO	158.5	Italy	CATALUNA	166.5	Spain
VORARLBERG	2	Austria	DUESSELDORF	135	Germany	LIGURIA	34	Italy	COMUNIDAD VALENCIANA	106.5	Spain
WIEN	93.5	Austria	FREIBURG	3	Germany	LOMBARDIA	141.5	Italy	EXTREMADURA	9	Spain
ANTWERPEN	10	Belgium	GIESSEN	4	Germany	MARCHE	22.5	Italy	GALICIA	30.5	Spain
BRUXELLES BRUSSELS	1	Belgium	HAMBURG	74.5	Germany	MOLISE	4	Italy	MADRID	144.5	Spain
HAINAUT	82	Belgium	HANNOVER	57.5	Germany	PIEMONTE	61	Italy	MURCIA	15.5	Spain
HALLEVLVOORDE	107	Belgium	KARLSRUHE	24.5	Germany	PUGLIA	55.5	Italy	NAVARRA	9	Spain
LEUVEN	3	Belgium	KASSEL	7	Germany	SARDEGNA	50	Italy	PAISVASCO	47.5	Spain
LIEGE	9	Belgium	KOBLENZ TRIER RHEINHESSEN PFALZ	60	Germany	SICILIA	67.5	Italy	RIOJA	6.5	Spain
LIMBURGB	5	Belgium	KOELN	82.5	Germany	TOSCANA	70	Italy	MELLERSTA NORRLAND	12	Sweden
LUXEMBOURG	5	Belgium	LUENEBURG	8	Germany	TRENTINO	6.5	Italy	NORRAMELLAN SVERIGE	18	Sweden
NAMUR	5	Belgium	MAGDEBURG DESSAU HALLE	6	Germany	UMBRIA	5.5	Italy	OSTRAMELLAN SVERIGE	61.5	Sweden
OOST VLAANDEREN	5	Belgium	MECKLENBURG	11.5	Germany	VALLE D'AOSTA	2	Italy	OVRENORRLAND	12	Sweden
WEST VLAANDEREN	4	Belgium	MITTELFRANKEN	53.5	Germany	VENETO	88.5	Italy	SMALAND MEDYARNA	13	Sweden
DENMARK	94.5	Denmark	MUENSTER	18.5	Germany	LUXEMBOURG GRAND DUCHE	52.5	Luxembourg	STOCKHOLM	76	Sweden
ALAND	0	Finland	NIEDERBAYERN	3	Germany	DRENTHE	11.5	Netherlands	SYDSVERIGE	15.5	Sweden
ITA SUOMI	2	Finland	OBERBAYERN	137	Germany	FLEVOLAND	3	Netherlands	VASTSVERIGE	74	Sweden
POHJOIS SUOMI	3	Finland	OBERFRANKEN	7.5	Germany	FRIESLAND	5	Netherlands	EAST ANGLIA	169.5	UK
REST OF FINLAND	2	Finland	OBERPFALZ	4	Germany	GELDERLAND	7	Netherlands	EAST MIDLANDS	50	UK
ALSACE	30	France	SAARLAND	13	Germany	GRONINGEN	3	Netherlands	NORTH	40.5	UK
AQUITAINE	48.5	France	SACHSEN	34	Germany	LIMBURGNL	8.5	Netherlands	NORTHERN IRELAND	44.5	UK
AUVERGNE	21	France	SCHLESWIG HOLSTEIN	14	Germany	NOORD BRABANT	46	Netherlands	NORTH WEST UK	121	UK
BASSE NORMANDIE	11.5	France	SCHWABEN	28.5	Germany	NOORD HOLLAND	141.5	Netherlands	SCOTLAND	92	UK
BOURGOGNE	11	France	STUTTGAERT	87.5	Germany	OVERIJSSSEL	6	Netherlands	SOUTH EAST UK	157.5	UK
BRETAGNE	22	France	THUERINGEN	6	Germany	UTRECHT	3	Netherlands	SOUTH WEST UK	63	UK
CENTRE	13	France	TUEBINGEN	24	Germany	ZEEELAND	5	Netherlands	WALES	23	UK
CHAMPAGNE ARDENNE	8	France	UNTERFRANKEN	7	Germany	ZUID HOLLAND	40	Netherlands	WEST MIDLANDS	60	UK
FRANCHE COMTE	5	France	WESER EMS	8	Germany	ALGARVE	52	Portugal	YORKSHIRE AND HUMBERSIDE	43	UK
HAUTE NORMANDIE	4	France	ANATOLIKI MAKEDONIA THRAKI	3.5	Greece	CENTRO ALENTEJO	6	Portugal			
ILE DE FRANCE	157.5	France	ATTIKI	59.5	Greece	LISBOA	83.5	Portugal			
LANGUEDOC ROUSSILLON	32.5	France	DYTIK IELLADA	4	Greece	NORTE	68	Portugal			

## 6 Results

The median values of the parameters' estimates of equation (15) are reported in Table 3. All the covariates that enter the model are statistically significantly different from zero, and have the expected signs. The relative skill intensity of a region depends positively on the net inflow of skilled foreigners ( $\alpha_5$ ), a large labour force, and participation. By contrast, it grows negatively with the region's openness to trade ( $\alpha_4$ ), high value added, high unemployment, and, perhaps surprisingly, life expectancy.

There are significant differences in the evolution of the skill intensity across member states. Over the period examined, skilled labour has grown above the EU average in Germany (the reference) and in three main geographic areas. These are: (1) Belgium and Luxembourg, possibly due to the arrival of new member states and an enlarged EU administrative centre; (2) the Nordic countries (Denmark, Finland, Sweden), which enjoy a common labour market and include two new EU member countries (Finland and Sweden); and (3) many regions in the South of the EU (France, Greece, Italy and Spain), which were the beneficiaries of large EU funds reallocations as well as foreign direct investments from both within and external the EU.

In contrast, initially skilled-intensive member states such as Austria, the UK, and the Netherlands experienced a slower growth in skill intensity.

The parameter of interest,  $\alpha_1$ , is negative and statistically significantly different from zero. During the period in exam, differences in skill intensity amongst European regions have reduced, implying convergence in regional skill endowments. Since the period examined coincides with a time of intensified economic integration with reforms promoting greater capital and labour mobility within the EU, the results support the presence of a negative relationship between high barriers to factor movements and skill intensity, as predicted by the Heckscher-Ohlin international trade theorem. We find no evidence of intensified skills' agglomeration in few locales. This is the first main result of our analysis.

The second novel result is about the sign and statistical significance of the parameter  $\rho$ , which accounts for spatial random effects. The estimate shows that the role of geographic adjacency is statistically significantly different from zero, and is positive, as expected (+0.457). Neighbouring a region experiencing an increase in skill intensity also increases a region's own skill intensity.



Table 3: Medians and 95% credible intervals of the estimated parameters.

Covariate	2.5%	50%	97.5%
$\alpha_1$	-0.2336	-0.2272	-0.2174
$\alpha_5$	0.0071	0.0081	0.0093
$\alpha_6$	-0.0021	-0.0013	-0.0003
$\alpha_4$	-0.0855	-0.0817	-0.0774
Unemployment rate	-0.0016	-0.0008	-0.0001
Life expectancy	-0.0022	-0.0013	-0.0006
Value added	-0.0276	-0.0268	-0.0262
Labour force	0.0022	0.0025	0.0028
Participation rate	0.0158	0.0171	0.0180
Austria	-0.0087	-0.0071	-0.0056
Belgium	0.0117	0.0134	0.0151
Denmark	0.0164	0.0180	0.0202
Finland	0.0043	0.0069	0.0087
France	0.0190	0.0200	0.0209
Greece	0.0067	0.0083	0.0106
Ireland	-0.0092	-0.0079	-0.0062
Italy	0.0111	0.0122	0.0131
Luxembourg	0.0880	0.0891	0.0906
Netherlands	-0.0054	-0.0040	-0.0029
Portugal	-0.0057	-0.0035	-0.0029
Spain	0.0101	0.0111	0.0119
Sweden	0.0111	0.0118	0.0132
UK	-0.0029	-0.0019	0.0003
$\sigma$	0.4199	0.4197	0.4195
$\rho$	0.4552	0.4570	0.4585
$\eta$	2.6707	2.6728	2.6748

The most surprising result however has been the magnitude and statistical significance of the parameter  $\eta$ , which measures the strength of neighbouring effects stemming from civilian air links vis-à-vis geographic contiguity. As  $\eta = 2.673$ , we find that the regions with a stronger increase in skill intensity are more likely to have similarities with other advanced regions which they are connected to by air transport

than with contiguous regions. This is probably the effect of the presence of a non-homogeneous distribution of skilled workers in Europe coinciding with a few centres in the EU that attract most of the skilled labour available. If we look at the results when the flight connections are not considered and only the adjacency relationship enters the matrix  $W$  ( $\eta = 0$ , results not shown), the spatial random effects are significant but very weak ( $\rho = 0.011$ ). This can be interpreted as the consequence of the geographical isolation of regions attracting skilled workers. The EU does not have clusters of regions with similar levels of economic development and industrial structure located in a restricted area, but, thanks to its history of nation-states, is characterized by a few regions with high skill intensity (e.g. the regions where London, Paris, Milan, Frankfurt, Madrid are located) embedded in much less attractive economic contexts for skilled workers. As a result, when we augment the adjacent matrix with information on inter-regional flight connections, the spatial random effects become much stronger ( $\rho = 0.457$ ). This result suggests that an increase in the skill intensity in a region is most likely associated with a similar increase in regions that are comparable for economic structure, and similarly well served by air transport links. Geographical proximity is not a sufficient condition to solely determine the evolution of skill endowments across space. Economic fortunes in the EU depend, amongst others but crucially, on the existence of good air transport infrastructure and networks in a locale.

## References

- Azomahou, T.T., Phu N.V., and Thi K.C.P. (2011). Testing convergence of European regions: A semiparametric approach. *Economic Modelling* 28(3): 1202-1210.
- Becker, S.O., Egger, P. and M. Von Ehrlich (2010). Going NUTS: The effect of EU Structural Funds on regional performance. *Journal of Public Economics* 94(9): 578-590.
- Besag, J. (1974): Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)* 36, 192–236
- Besag, J., York, J., and A. Molli (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 43(1), 1–59
- Blanchard, O.J. and L.F. Katz (1992). Regional Evolutions, *Brookings Papers on Economic Activity* 0(1): 1-75.
- Boldrin, M., and F. Canova (2001). Inequality and convergence in Europe's regions: reconsidering European regional policies. *Economic policy* 16(32): 205-253.
- Brook, D. (1964): On the distinction between the conditional probability and the joint probability approaches in the specification of nearest-neighbour systems. *Biometrika* 51, 481–483
- Canova, F. (2004) Testing for Convergence Clubs in Income Per Capita: A Predictive Density Approach. *International Economic Review* 45(1): 49-77.
- Carlin, B.P. and S. Banerjee (2003): Hierarchical multivariate CAR models for spatio-temporally correlated survival data. In: Bayesian statistics, 7 (Tenerife, 2002), pp. 45–63. Oxford Univ. Press, New York
- Cavelaars, P., and J. Hessel (2007). *Regional Labour Mobility in the European Union: Adjustment Mechanism or Disturbance?* De Nederlandsche Bank.
- Chauveau, D., and P. Vandekerkhove (2002): Improving convergence of the Hastings-Metropolis algorithm with an adaptive proposal. *Scandinavian Journal of Statistics* 29(1), 13–29
- Corrado, L., Martin, R., and M. Weeks (2005). Identifying and Interpreting Regional Convergence Clusters across Europe. *The Economic Journal* 115(502): C133-C160.
- Dall'Erba, S., Percoco, M., and G. Piras (2008). The European regional growth process revisited." *Spatial Economic Analysis* 3(1): 7-25.
- Decressin, J. and A. Fatas (1995) Regional Labour Market Dynamics in Europe, *European Economic Review* 39: 627-1655.
- Dowrick, S. (2003). *Ideas and education: level or growth effects?* No. w9709. National Bureau of Economic Research.

- Esteban, J. (2002). *Economic polarization in the Mediterranean Basin*. CREI, Centre de Recerca en Economia Internacional.
- Eurostat (1992) *Labour Force Survey: Methodology and Definitions*, Luxembourg: Eurostat.
- Ezcurra, R., and M. Rapún (2006). Regional Disparities and National Development Revisited The Case of Western Europe. *European Urban and Regional Studies* 13(4): 355-369.
- Flanagan, R.J. (1993). European Wage Equalisation since the Treaty of Rome. In U.Lloyd, B.Eichengreen & W.T. Dickens (Eds.), *Labor in an Integrated Europe* (pp.167-187). Washington, D.C.: The Brookings Institution.
- Fredriksson, P. (1999) ‘The Dynamics of Regional Labor Markets and Active Labor Market Policy: Swedish Evidence’, *Oxford Economic Papers* 51(4): 623-648.
- de Freitas, M., Pereira, F., and F. Torres (2003). Convergence among EU regions, 1990–2001. *Intereconomics* 38(5): 270-275.
- Gaspar, J., and E. Glaeser (1998). “Communications Technology and the Future of Cities”. *Journal of Urban Economics*, 43(1), 136-156.
- Giannetti, M. (2002). The effects of integration on regional disparities: Convergence, divergence or both? *European Economic Review* 46(3): 539-567.
- Gelfand, A., and P. Vounatsou (2003): Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics* 4(1), 11–15–11–15
- Gelfand, A.E., and S.K. Sahu (1994): On Markov chain Monte Carlo acceleration. *Journal of Computational and Graphical Statistics* 3(3), 261–276
- Gelman, A., Carlin, J.B., Stern, H.S., and D.B. Rubin (2003): *Bayesian Data Analysis*, second edn. Chapman and Hall, London
- Gilks, W., Roberts, G., and E. George (1994): Adaptive direction sampling. *The Statistician* 43(1), 179–189
- Glaesner, H.J. (1986): The Single European Act. *Yearbook of European Law* 6(1), 283–312
- Haario, H., Saksman, E., and J. Tamminen (1999): Adaptive proposal distribution for random walk Metropolis algorithm. *Computational Statistics* 14, 375–395
- Heckerman, D., Smith, A.F.M., and M. West (eds.) (2003). *Bayesian Statistics* 7, pp. 277–292. Clarendon Press, Oxford
- Jimeno, J.F. and Bentolila, S. (1998). Regional Unemployment Persistence (Spain, 1976-1994), *Labour Economics* 5(1): 25-51.
- Keesing, D.B. (1966). Labor Skills and Comparative Advantage, *American Economic*

*Review* 56(2): 249-258.

Keith, J., Kroese, D.P., and D. Bryant (2004): A generalized Markov sampler. *Methodology and Computing in Applied Probability* 6(1), 29–53

Keith, J.M., Kroese, D.P., and G.Y. Sofronov (2008): Adaptive independence samplers. *Statistics and Computing* 18(4), 409–420

Keith, J.M., Sofronov, G.Y., and D.P. Kroese (2008): The generalized gibbs sampler and the neighborhood sampler. In: A. Keller, S. Heinrich, H. Niederreiter (eds.) *Monte Carlo and Quasi-Monte Carlo Methods 2006*, pp. 537–547. Springer-Verlag Berlin Heidelberg, Germany, Heidelberg

Krugman, P.R. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy*, 99(3), 483-499.

Maastricht Treaty (1992)  
[http://europa.eu/legislation\\_summaries/institutional\\_affairs/treaties/treaties\\_maastricht\\_en.htm](http://europa.eu/legislation_summaries/institutional_affairs/treaties/treaties_maastricht_en.htm), accessed March 22

Martin, R., and P. Tyler (2000). Regional employment evolutions in the European Union: a preliminary analysis. *Regional Studies* 34(7): 601-616.

Martin, R. (2001). EMU versus the regions? Regional convergence and divergence in Euroland. *Journal of Economic Geography* 1(1): 51-80.

Mauro, P., and A. Spilimbergo (1998). How Do the Skilled and the Unskilled Respond to Regional Shocks?: The Case of Spain. *IMF Staff Papers*, 46(1), 1-17.

Mengersen, K., and C.P. Robert, C.P. Iid sampling using self-avoiding population monte carlo: the pinball sampler. In: J.M. Bernardo, M.J. Bayarri, J.O. Berger, A.P. Dawid

Mundell, R.A. (1957). 'International Trade and Factor Mobility'. Reprinted in J.Bhagwati (Ed.), *International Trade* (pp. 321-335). Cambridge, MA and London: The MIT Press.

Nahuis, R. and Parikh, A. (2001) Factor Mobility and Regional Disparities: East, West, Home's Best?, Paper presented at the Economics Program Seminars, Research School of Social Sciences, Canberra: Australian National University, 27<sup>th</sup> April.

Niebuhr, A., and F. Schlitte (2004). Convergence, trade and factor mobility in the European Union—implications for enlargement and regional policy. *Intereconomics* 39(3): 167-176.

Obstfeld, M. and G. Peri (1998) Regional non-adjustment and Fiscal Policy, *Economic Policy: A European Forum* 0(26): 205-247.

Overman, H., and D. Puga (2002). Unemployment clusters across Europe's regions and countries. *Economic policy* 17(34): 115-148.

- Puga, D. (2002). European regional policies in light of recent location theories. *Journal of Economic Geography* 2(4): 373-406.
- Quah, D. (1996). Regional convergence clusters across Europe. *European Economic Review* 40(3): 951-958.
- Ramajo, J., Marquez, M., Hewings, G. and M. Salinas (2008). Spatial heterogeneity and interregional spillovers in the European Union: Do cohesion policies encourage convergence across regions?. *European Economic Review* 52(3): 551-567.
- Recchi, E. (2008). Cross-state mobility in the EU. *European Societies* 10(2): 197-224.
- Romer, P. (1991). *Endogenous technological change*. No. w3210. National Bureau of Economic Research.
- Salt, J. (1992). Migration Processes among the Highly Skilled in Europe. *International Migration Review*, 26(2), 484-505.
- Sassen, S. (1991). *The Global City*. Princeton: Princeton University Press.
- Tani, M. (2003). Have Europeans become more mobile? A note on regional evolutions in the EU: 1988–1997. *Economics Letters* 80(1): 23-30.
- Ter Braak, C.J.F. (2006). A Markov chain Monte Carlo version of the genetic algorithm differential evolution: easy Bayesian computing for real parameter spaces. *Statistics and Computing* 16
- Tierney, L., and A. Mira (1999): Some adaptive Monte Carlo methods for Bayesian inference. *Statistics in Medicine* 18, 2507–2515
- Venables, A.J. (1995). Economic Integration and the Location of Firms. *American Economic Review*, 85(2), 296-300.

## Appendix:

### Theoretical model – solving the model

The evolution of a region's relative skill intensity (7) and labour demand (8) enable us to rewrite equation (1), after bringing it forward one period and subtracting from it its value at time  $t$  and replacing the regional relative unemployment rate with equation (6), as:

$$w_{it+1} - w_{it} = -d\Delta n_{it} - dc(w_{it+1} - w_{it}) + \Delta z_{it}$$

We will also rewrite equation (2) brought forward one period as:

$$\begin{aligned} \Delta n_{it+1} = & \left[ \frac{1}{2}(b_1 + cg_1) + \frac{1}{4}(b_2 + cg_2) + \frac{1}{4}(b_3 + cg_3) \right] w_{it+1} + X_i^s + \\ & d \left\{ \frac{1}{2}(\Delta \tilde{\theta}_{ist+1}^N - \Delta \tilde{\theta}_{iut+1}^N) + \frac{1}{2}(\Delta(1 - \tilde{\theta}_{ist+1}^N) - \Delta(1 - \tilde{\theta}_{iut+1}^N)) \right\} + \\ & d \left\{ \frac{1}{4}(\Delta \tilde{\theta}_{ist+1}^I - \Delta \tilde{\theta}_{iut+1}^I) + \frac{1}{4}(\Delta(1 - \tilde{\theta}_{ist+1}^I) - \Delta(1 - \tilde{\theta}_{iut+1}^I)) \right\} + \\ & \frac{1}{2}\varepsilon_{1t+2}^s + \frac{1}{4}\varepsilon_{2t+2}^s + \frac{1}{4}\varepsilon_{3t+2}^s \end{aligned}$$

The equilibrium is found for values of  $w_{it+1}$ , which are then replaced in  $\Delta n_{it+1}$  to find the equilibrium regional relative employment growth.

### *Solutions for native, intra- and extra-EU employment*

The equilibrium paths for native, intra- and extra-EU relative employment growth can be found by replacing the equilibrium path for wages described by equation (9) into equations (5a)-(5c), respectively. This yields:

$$\Delta\tilde{N}_{it+1} = (b_1 + dg_1) \frac{1+dc-dR-Q}{1+dc} \Delta\tilde{N}_{it} + \frac{Q}{1+dc} X_i^s + \frac{R}{1+dc} X_i^d - \frac{RJ_t^T}{1+dc} + dO_{t+1} + \frac{d(1+dc-dR-Q)-dR}{1+dc} O_t + \frac{R}{1+dc} (J_t^d - dJ_t^s) + J_{t+1}^s \quad (11a)$$

$$\Delta\tilde{I}_{it+1} = (b_2 + dg_2) \frac{1+dc-dR-Q}{1+dc} \Delta\tilde{I}_{it} + \frac{Q}{1+dc} X_i^s + \frac{R}{1+dc} X_i^d - \frac{RJ_t^T}{1+dc} + dO_{t+1} + \frac{d(1+dc-dR-Q)-dR}{1+dc} O_t + \frac{R}{1+dc} (J_t^d - dJ_t^s) + J_{t+1}^s \quad (11b)$$

$$\Delta\tilde{E}_{it+1} = (b_3 + dg_3) \frac{1+dc-dR-Q}{1+dc} \Delta\tilde{E}_{it} + \frac{Q}{1+dc} X_i^s + \frac{R}{1+dc} X_i^d - \frac{RJ_t^T}{1+dc} + dO_{t+1} + \frac{d(1+dc-dR-Q)-dR}{1+dc} O_t + \frac{R}{1+dc} (J_t^d - dJ_t^s) + J_{t+1}^s \quad (11c)$$

### Steady States

The steady states are obtained by replacing the variables  $w$ ,  $u$ , and  $\Delta n$  with their long-term values, with average shares (constant over time) and a zero error term:

$$\bar{w}_i = \frac{X_i^{Td} + X_i^{NTd} - dX_i^s}{dR + Q} \quad (12)$$

$$\bar{u}_i = -d\bar{w}_i \quad (13)$$

$$\bar{\Delta n}_i = \frac{R(X_i^{Td} + X_i^{NTd}) + QX_i^s}{dR + Q} \quad (14)$$

As in Blanchard-Katz, the attractiveness of regions to workers and firms are two underlying sources of regional relative wages and employment growth. An increase in the attractiveness to skilled workers ( $X^s$ ) in general reduces the regional relative skilled wage but raises the growth rate of skilled employment and the relative unemployment rate. An increase in attractiveness to firms ( $X^{Td} + X^{NTd}$ ) typically increases the relative skilled wage and skilled employment, and decreases the relative unemployment rate, though the exact effects on wages depend on whether  $X^{Td} + X^{NTd} > dX^s$  (stronger demand effects) or vice-versa.

However, in the model above there are two extra sources of employment growth: foreign labour and the tradable sector. Their effect on skilled relative wages and employment growth is ambiguous, as it depends on the net effect on labour supply and demand due to a region's attractiveness, and the magnitude of the elasticity of labour demand to wages ( $d$ ). If  $X^{Td} + X^{NTd} > dX^s$  (stronger demand effects), then an increase in intra- and extra-EU skilled labour will increase the region's steady state



skilled employment growth, though at lower steady state wages. As intra- and extra-EU are characterised by different labour supply elasticities, relative wages will be higher if intra- rather than extra-EU workers move in. The presence of tradable firms (Q) lowers the steady state of regional skilled wages and employment growth, but raises the steady state of relative unemployment, and vice-versa.

### **Estimation strategy: Bayesian Adaptive Independence Sampler**

Markov chain Monte Carlo (MCMC) is a popular computational technique for generating samples from non-standard probability distributions. One of the main issues in the design of MCMC methods is to attain both efficient convergence and mixing properties. One way to achieve this is to adapt the proposal distribution in light of previously sampled points.

Various adaptive MCMC algorithms have been considered in the literature. A popular approach to adapting the proposal distribution is to adjust the proposal based on the pre-runs (see, for example, Gel, 1994; Haa, 1999). An alternative approach is based on multiple chains (see Gil, 1994).

In this section we describe the Bayesian Adaptive Independence Sampler (Keith, 2008)], which is based on running multiple parallel chains with a common proposal distribution. The proposal is periodically updated using the set of current elements across all chains. Adapting the proposal does not change the limiting distribution of each chain.

Let  $f$  be the target pdf and let  $g(x|\theta)$  be a proposal distribution, defined up to a parameter  $\theta$ , which is to be updated. Let  $\theta_0, \theta_1, \dots$  be the parameters for the sequence of proposals. Suppose we have  $N$  parallel chains  $\{X_{1,j}, j = 1, 2, \dots\}, \dots, \{X_{N,j}, j = 1, 2, \dots\}$ , which are referred as the *sampling chains*. At each step of the algorithm,  $\theta_j$  is updated, using the set of current elements  $(X_{1,j}, \dots, X_{N,j})$ . This means that after updating each of the  $N$  chains, the proposal is updated. The algorithm cycles through updates for  $N+1$  chains, since  $\theta_0, \theta_1, \dots$  may also be regarded as values of an underlying Markov chain, which is referred as the *parameter chain*.

Let  $X$  denote the target space, that is, the space on which the target distribution  $f$  is defined. Let  $\Theta$  denote the space of parameters for the proposal distribution. We may regard the  $N+1$  parallel chains as a single chain defined on a space  $Z := \Theta \times X^N$ . So the algorithm target  $\pi$  on  $Z$  is given by

$$\pi(\theta, x_1, \dots, x_N) = h(\theta | x_1, \dots, x_N) \prod_{i=1}^N f(x_i).$$

The algorithm can be described using either a framework of the Generalized Gibbs Sampler (Kei, 2008) or a Metropolis-Within-Gibbs scheme. This scheme involves iterating two steps, known as the Gibbs step and the Metropolis-Hastings step. The Gibbs step is used to propose either a new element  $y \in X$  or a new parameter  $\theta \in \Theta$ . The Metropolis-Hastings step is used to accept or reject it in accordance with an acceptance probability. Both steps satisfy detailed balance conditions.

We use a multivariate normal distribution as the proposal:

$$g(x|\theta) = N(x|\mu, \Sigma) \propto |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right\}.$$

Then  $X = \mathfrak{R}^d$  for some positive integer  $d$ . Using a non-informative prior, the posterior distribution can be obtained (see Gel, 2003 for further details):

$$h(\theta | x_1, \dots, x_N) = h(\mu, \Sigma | x_1, \dots, x_N) = N(\mu | \bar{x}, \Sigma / N) \cdot \text{Inv-W}_{N-1}(\Sigma | S),$$

$$S = \sum_{n=1}^N (x_n - \bar{x})(x_n - \bar{x})^T, \quad \bar{x} = \frac{1}{N} \sum_{n=1}^N x_n.$$

This means that in order to obtain parameters  $\mu$  and  $\Sigma$ , first  $\Sigma$  is drawn from an Inverse-Wishart distribution  $\text{Inv-W}_{N-1}(\Sigma | S)$ , then  $\mu$  is drawn from a Normal distribution  $N(\mu | \bar{x}, \Sigma / N)$ .

The BAIS can be described as the following two steps performed iteratively:

**The Gibbs step:** Cycle over the  $N$  sampling chains and the parameter chain:

- If  $i \in \{0, N-1\}$ , generate  $Y \sim N(y | \mu, \Sigma)$
- If  $i=N$ , generate  $Y \sim N(\mu | \bar{x}, \Sigma / N) \cdot \text{Inv-W}_{N-1}(\Sigma | S)$

**The Metropolis-Hastings step**

- If  $i \in \{0, N-1\}$ ,  $x_i = \begin{cases} Y & \text{if } U \leq \alpha_i(x_i, Y) \\ x_i & \text{otherwise} \end{cases}$

$$\alpha_i(x_i, Y) = \min(\rho_i(x_i, Y), 1), \quad U \sim U(0, 1)$$

- If  $i=N$ ,  $(\mu, \Sigma) = Y$

$$\rho_i(x_i, Y) = \left( \frac{f(y)h(\theta | y)g(x_i | \theta)}{f(x_i)h(\theta | x_i)g(y | \theta)} \right)^{-1} =$$

$$\left( \frac{f(y)N(\mu | \bar{x}_y, \Sigma / N) \cdot \text{Inv-W}_{N-1}(\Sigma | S_y) \times N(x_i | \mu, \Sigma)}{f(x_i)N(\mu | \bar{x}, \Sigma / N) \cdot \text{Inv-W}_{N-1}(\Sigma | S) \times N(y | \mu, \Sigma)} \right)^{-1}$$

where

$$\begin{aligned}\bar{x}_Y &= \frac{1}{N}(x_1 + \dots + x_{i-1} + Y + x_{i+1} + \dots + x_N) \\ S_Y &= (x_1 - \bar{x}_Y)(x_1 - \bar{x}_Y)^T + \dots + (x_{i-1} - \bar{x}_Y)(x_{i-1} - \bar{x}_Y)^T \\ &+ (Y - \bar{x}_Y)(Y - \bar{x}_Y)^T + (x_{i+1} - \bar{x}_Y)(x_{i+1} - \bar{x}_Y)^T + \dots \\ &+ (x_N - \bar{x}_Y)(x_N - \bar{x}_Y)^T\end{aligned}$$

With reference to the statistical properties of the model, the application of the BAIS method has proven to outperform a standard Metropolis sampler in both speed of convergence and autocorrelation properties (see Figures 1 and 2). Figure 1 shows the log-likelihood values for 1000 iterations of the BAIS and the Random Walk Metropolis-Hastings (RWMH) algorithm, confirming fast convergence of the adaptive sampler to an optimal value.

**Figure 1 Burn-in phases for BAIS and RWMH: log-likelihood curves**

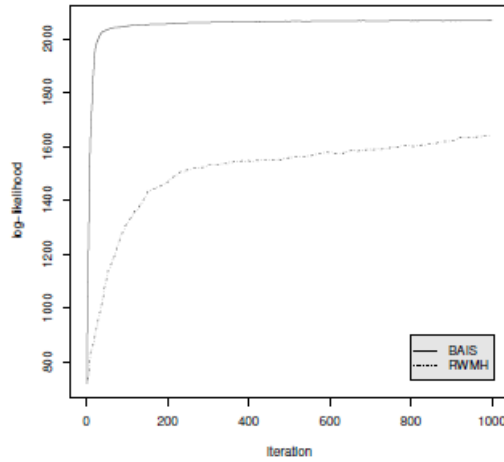


Figure 2 plots the curves of autocorrelations for lags  $0, 1, \dots, 400$ , with the number of chains  $N = 50$ , supporting the claim that the BAIS is more efficient than the RWMH, in terms of autocorrelation of the Markov chain.

**Figure 2** Autocorrelations for  $\alpha_0$  for BAIS and RWMH

