DISCUSSION PAPER SERIES

DP12749

REGIONAL ECONOMIC DEVELOPMENT IN EUROPE, 1900-2010: A DESCRIPTION OF THE PATTERNS

Joan R. Rosés and Nikolaus Wolf

ECONOMIC HISTORY, INTERNATIONAL TRADE AND REGIONAL ECONOMICS and MACROECONOMICS AND GROWTH

REGIONAL ECONOMIC DEVELOPMENT IN EUROPE, 1900-2010: A DESCRIPTION OF THE PATTERNS

Joan R. Rosés and Nikolaus Wolf

REGIONAL ECONOMIC DEVELOPMENT IN EUROPE, 1900-2010: A DESCRIPTION OF THE PATTERNS

Abstract

We provide the first long-run dataset of regional employment structures and regional GDP and GDP per capita in 1990 international dollars, stretching over more than 100 years. These data allow us to compare regions over time, among each other, and to other parts of the world. After some brief notes on methodology we describe the basic patterns in the data in terms of some key dimensions: variation in the density of population and economic activity, the spread of industry and services and the declining role of agriculture, and changes in the levels of GDP and GDP per capita. We next discuss patterns of convergence and divergence over time and their explanations in terms of short-run adjustment and long-run fundamentals. Also, we document for the first time a secular decrease in spatial coherence from 1900 to 2010. We find a U-shaped development in geographic concentration and regional income inequality, similar to the finding of a U-shaped pattern of personal income inequality.

JEL Classification: D31, N1, N9, R1

Keywords: Regional Inequality, Europe, Long-Run

Joan R. Rosés - j.r.roses@lse.ac.uk
London School of Economics and Political Science and CEPR

Nikolaus Wolf - nikolaus.wolf@wiwi.hu-berlin.de
Humboldt University Berlin and CEPR

Regional Economic Development in Europe, 1900-2010: a description of the Patterns

Joan R. Rosés[1] and Nikolaus Wolf[2]

February 2018

Abstract

We provide the first long-run dataset of regional employment structures and regional GDP and GDP per capita in 1990 international dollars, stretching over more than 100 years. These data allow us to compare regions over time, among each other, and to other parts of the world. After some brief notes on methodology we describe the basic patterns in the data in terms of some key dimensions: variation in the density of population and economic activity, the spread of industry and services and the declining role of agriculture, and changes in the levels of GDP and GDP per capita. We next discuss patterns of convergence and divergence over time and their explanations in terms of short-run adjustment and long-run fundamentals. Also, we document for the first time a secular decrease in spatial coherence from 1900 to 2010. We find a U-shaped development in geographic concentration and regional income inequality, similar to the finding of a U-shaped pattern of personal income inequality.

[1] London School of Economics and Political Science, UK and CEPR, emaill: J.R.Roses@lse.ac.uk
[2] Humboldt University Berlin, Germany, CEPR and CESifo, email: nikolaus.wolf@wiwi.hu-berlin.de, corresponding author

1. Introduction[3]

Over the last four generations, the European economy went through turbulent changes. In 1900, the UK was still the leading country of the world in more than one respect, with France, Germany and others following and catching up. The First World War marked the end of a long period of both, economic growth and integration. Moreover, if seen from a global perspective the Great War also marked the end of European dominance and the beginning of a decline of the continent in weight and influence. During the interwar years, the European growth record was rather poor since erroneous policies and coordination failures prevented Europe from fully realizing its economic potential (Rosés and Wolf 2010). After the Second World War, Europe's economy started another long period of rapid economic expansion (the 'Golden Era'), which slowed down in the 1970s but nevertheless continued until today. Again, this expansion was accompanied by a process of integration across states, notably with the formation of the European Economic Community and, later, the Eurozone. More recently, the project of European integration has been fundamentally questioned, partly in consequence of the Global Financial crisis and the European Debt crisis that followed in its wake. Moreover, it seems that forces of economic and political disintegration are gaining momentum not only in Europe but also in other major developed economies. These different historical tendencies have been described and analysed by a substantial literature elsewhere (see, for example, Crafts and Toniolo 1996; Eichengreen 2007; Berend 2016; James 2017).

However, most authors have treated the European economy as a group of national economies, stressing the role of national governments and international organizations. Such an approach has several advantages. First, it naturally ties in with the political history of Europe, based on the emergence of territorial national states during the early modern period and their international relations. Second, most quantitative evidence has been collected and described at the level of nation states based on the work of national statistical offices, which developed during the 19th century. Yet this approach comes at some costs. It neglects the often considerable variation within states (sometimes larger than between states) and it tends to attribute differences in development to differences in national institutions or policies without being able to test this. As we will show in this paper, there are several clusters of regions, which are sometimes highly developed, that transcend national boundaries, such as the enlarged Rhine–Meuse–Scheldt delta. Furthermore, differences of income per capita (and labour productivity) within countries are larger, and sometimes more resilient, than differences across countries. In particular, the process of income

convergence across European nations was not always accompanied by a similar process of convergence of regions within countries. We can show that the distribution of activity across regions shifted over time, first until 1980 converging to a more equal distribution, and from 1980 diverging back to a less equal distribution.

In this paper we want to reconsider the economic development of Europe since 1900 from the perspective of European regions as pioneered by Pollard (1981), and to provide a quantitative basis for more work along these lines. We do this using modern regional units (following the European NUTS classification as far as we can), which we trace back over time with comparable indicators of economic development. We provide a set of new estimates of regional employment structures and regional GDP and GDP per capita in 1990 international dollars, stretching over more than 100 years. These data allow us to compare regions over time, among each other, and to other parts of the world. After some brief notes on our methodology we describe the basic patterns in the data in terms of some key dimensions: variation in the density of population and economic activity, the spread of industry and services and the declining role of agriculture, and changes in the levels of GDP and GDP per capita. We next discuss patterns of convergence and divergence over time and show how the geography of activity has changed with a long-run decrease in spatial coherence from 1900 to 2010 and a U-shaped development in geographic concentration and regional income inequality. The latter seems to be related to the finding of a U-shaped pattern of personal income inequality as documented by Piketty and Saez (2003), Piketty (2014) and others.

2. Data and Methodology

Our data set contains 173 regions covering 16 European nation states at the level of NUTS-2 (as of 2014) and spans 11 benchmark years between 1900 and 2010.[4] We could not at this stage include the long-run development of states in Central and Eastern Europe, because the reconstruction of historical data that would stretch back until 1900 is still under way. Eight of our 173 sample regions are aggregated from two or three NUTS-2 regions in order to trace the regions over time in constant borders. Moreover, some of our regions belonged to different political entities over time, such as Alsace or Lorraine, which provides us with some interesting case studies on the potential role of national institutions for economic development. One of our regions – Flevoland in the Netherlands – consists mostly of land that has been reclaimed from ocean beds only in the 1950s and 1960s and therefore enters the data only in 1970. Lastly, for two states, Luxembourg and the Republic of Ireland, we have no further regional breakdown.

---

[4] We cover Austria, Belgium, Denmark, Finland, France, Germany, Italy, Ireland, the Netherlands, Norway Portugal, Spain, Sweden, Switzerland, and the United Kingdom and also consider newly estimated data for Luxembourg in our empirical work. For details on the data see the relevant country chapters in Joan R. Rosés and Nikolaus Wolf (eds.), The Economic Development of Europe's Regions: A Quantitative History Since 1900 (forthcoming, Routledge Explorations in Economic History).

To reconstruct regional GDPs, we have resorted to two different types of methods and sources. After 1960, we have mostly employed official data on the regional distribution of income. Specifically, from 1960 to 1990, national statistic offices provided that kind of information and since then regional data is being provided by Eurostat, the statistical office of the European Union. For the majority of countries before 1960, we have employed the Geary and Stark's methodology (Geary and Stark 2002). Notable exceptions here are Austria and the Netherlands, where a more direct approach could be used. Geary and Stark's methodology has two main advantages: (1) it requires readily available data (employment by sector and region, wages by sector and region, and historical national accounts) and (2) has an easy interpretation within the national accounting framework. The basic principle is that a country's GDP is equal to the sum of all regional GDPs. More specifically, the total GDP of any country Y_i is the sum of n regional GDPs (Y_j):

$$Y_i = \sum_j^n Y_j.$$

Furthermore, regional GDP (Y_j) can be decomposed into the contributions from all sectors in the economy:

$$Y_j = \sum_k^K y_{jk} L_{jk},$$

y_{jk} being the output, or the average added value, per worker in each region j, in sector k, and L_{jk} the number of workers in each region j and sector k. As we have no direct data for y_{jk}, its value is approximated by assuming that regional differences in labour productivity in each industry are reflected in the regional industry wage level relative to the national industry wage level $\frac{w_{jk}}{w_k}$.

In consequence, we can estimate regional GDP as:

$$Y_j = \sum_k^K [y_k \beta_k (\frac{w_{jk}}{w_k})] L_{jk}$$

where, as suggested by Geary and Stark (2002), y_k is value added per worker in sector k at the national level, w_{jk} is the wage paid in region j in sector k, w_k is the country average wage in each sector k, and β_k is a scalar that preserves the relative region differences but scales the absolute values so that the regional total for each sector adds up to the country totals. So, in the absence of regional output figures, Geary and Stark (2002) suggest a framework for an indirect estimation based on variation in employment and wages, which allows for an approximation of GDP by region at country factor cost. Hence, the basic data involved in this estimation procedure are national estimates of GDP, value added per worker by sector, and nominal wages and employment, by sector and region.

However, we could (and did in some cases) replace indirect estimates with direct ones whenever the data was available. It should be noted that this methodology allows us to compute not only regional GDPs but also regional figures for the different industries. The validity of this methodology against government-based estimations has been often tested with the result that differences between two alternative approaches are

typically small and within the range of errors commonly accepted in official national accounting estimates. In one case – the Netherlands - the method proved less reliable for the period before 1950, but it was possible to use existing regional GDP estimates from van Zanden (1987) for the years 1820 – 1910 and thereafter estimates from the Dutch Central Bureau of Statistics. For Austria regional GDP could be estimated with a more direct approach based on existing regional production data.

Our employment data derive until 1990 from a variety of national sources, mainly population and employment censuses but for the last two benchmarks (2000 and 2010) has been taken directly from Eurostat databases. Obviously, employment data have been made homogeneous to originate in the same sectors across countries (broadly speaking, we have reduced employment to the three basic sectors – primary, secondary and tertiary).

To make our data homogeneous and comparable across different countries, we constructed for each country and year a regional breakdown of national GDP aggregates. Next, we used national-level GDP estimates from the Maddison Project (Bolt and van Zanden 2014), which provides data on GDP of European nation states expressed at purchasing power parity in 1990 international dollars. This database incorporated recent updates to national GDP estimates, such as Germany (Burhop and Wolff 2005), Sweden (Schön and Krantz 2012) or Italy (Baffigi 2013). In the case of Germany, we used the corrected data from Broadberry and Klein (2012) for estimates of the national GDP of the GDR and FRG for the years 1950-1980. Most national GDP figures for 1990-2010 in the Maddison Project data are in turn taken from the Total Economy Database of the Conference Board. It should be noted that country-level estimates of GDP at international prices from the Maddison Project differ slightly from alternative estimates such as those from Eurostat. In consequence, our GDP and GDP per capita calculations at the regional level for the years after 1990 are slightly different from those furnished by Eurostat, even though we have employed regional Eurostat data as base of our regional distribution of GDP within countries. Our main results are robust to the latest Maddison Project Database (2018) with multiple benchmarks.[5]

Our methodology also implies that – similar to estimates provided by Eurostat and the OECD - we have no regional price deflators but, instead, we use national deflators. Hence, we assume that all regions have the same prices within countries. This introduces some bias in our results. The first bias is that our calculations (like all official calculations) overestimate regional differences in living standards since *ceteris paribus* the richer regions tend to have higher prices than the poorer ones given that the non-tradable goods

---

[5] We use the Maddison Data with 1990 International Dollars to ensure comparability with the bulk of research on long-run development. However, we have recalculated all our main results using the more recent data from the Maddison Project Database (MDP) from 2018 with the new real GDP per capita measures termed "CGDPpc" based on multiple benchmark comparisons to improve historical income comparisons across countries. All our results are robust to this, except some minor changes in terms of long-run rankings and a less pronounced change in the locational Gini-coefficient for GDP. Notably the change in distributions (figures 5), in the share of capital-regions over time (fig 7) and the development of sigma-convergence (figure 8) remain the same if we use the MDP2018 data instead.

(e.g. housing) tend to be more expensive. The second bias is that regional price differentials have probably changed over time: with the integration of goods markets, prices of tradable goods have become more homogenous within countries while differences in housing prices may have increased over time. But our methodology has an important advantage: a substantial part of the price differential across regions of non-tradable goods is due to the monopoly power of real-estate owners (Moretti, 2013), who can extract rents from producers (workers). Therefore, our price-unadjusted regional per capita GDPs is an imperfect measure of "welfare" differential across regions in the same country, but it more likely reflects the "true" differences in labour productivity across regions.

A final problem with our estimates of regional shares in national GDP is the fact that these are based on census years which vary across nation states. We make the data comparable across countries by interpolating regional shares to several common benchmark years (namely 1900, 1910, 1925, 1938, 1950 and decades thereafter). Next we use these shares together with national-level GDP data from the Maddison Project for these respective benchmark years to construct regional data. We have always avoided interpolation across war periods. Regional shares in national aggregates tend to change very slowly and we find it unlikely that regions within a state follow different business cycle dynamics.

3. Basic Facts on Regional Economic Development: Density of Population, Employment and GDP

Let us start with a look at population density, (see maps 1 and 2). As expected, the density of population measured as persons per km^2 shows considerable variation across regions and over time. The average density increased from 150 (1900) to 282 (2010), the median from 84 (1900) to 149 (2010), indicating that a few very densely populated regions have a large effect. These outliers with extreme population density are basically the same back then and now, namely London and surroundings (UK1), Berlin (DE30) and Hamburg (DE60), followed at some distance by Bremen (DE50), Düsseldorf (DEA1), Brussels and Brabant (BE10, with BE24 and BE31), the Île de France with Paris (FR10) as well as North- and South-Holland (NL32 and NL33). A few regions with very high density in 1900 however experienced a dramatic decline over time, including Hainaut in Belgium (BE32) as well as Chemnitz (DED4) and Leipzig (DED12) in Germany and we will come back to their destiny further below. There was more stability at the bottom of the distribution, with the least densely populated regions both then and now being located in the northern parts of Sweden, Norway and Finland, followed by Alentejo (PT18) in Portugal and regions in central Spain, namely Castile-Leon (ES41), Castile-La Mancha (ES42) and Extremadura (ES43).

Map 1: Population Density, 1900

„Map1_popdensityEU_1900.png"

Source: see text.

Map 2: Population Density, 2010

„Map2_popdensityEU_2010.png"

Source: see text.

A simple intuition from these maps is that very low population densities are related to climatic extremes. More generally, natural geography, notably mean temperature, extreme values of average temperature and average precipitation, but also the suitability of soil for agriculture and distance to major sea ports are indeed very strongly correlated with variation in population density. Moreover, the correlation between the density of population and these geographical variables in 2010 is only very slightly weaker than it was in 1900. Apparently, the impact of natural geography on the location of population across Europe 1900-2010 is strong and persistent. We consider some of these factors in more detail below.

In a next step, we look at the density of economic activity as measured by total employment per km^2 (see maps 3 and 4). The average employment density roughly doubled from 67 persons per km^2 (1900) to 132 (2010). Not surprisingly, employment density is closely correlated with the density of population. However, population and employment density among regions could differ for several reasons. First, the distribution of dependency rates might be uneven due to different demographic trends (where regions with relatively more children and older people tend to have lower employment). Next, female participation rates could differ and regions with lower female participation rates also tend to have a lower overall share of their working-age population employed. And third, unemployment rates might vary. These three factors are not independent of each other as thriving regions tend to have lower unemployment and dependency rates (since attract working-age migrants), and higher female participation rates (given the abundance of labour opportunities), while the contrary holds for poorer regions. Furthermore, the relative importance of these three factors changed over time. Maps 3 and 4 show how employment density developed between 1900 and 2010.

Map 3: Employment Density, 1900

„Map3_empdensityEU_1900.png"

Source: see text.

Map 4: Employment Density, 2010

„Map4_empdensityEU_2010.png"

Finally, the density of GDP, that is GDP per km^2, is again closely related to the pattern of population and employment but the relationship is changing (maps 5 and 6). In this case, differences between employment density and GDP density reflect differences in productivity across regions: more productive regions generate more GDP per km^2 with the same employment per km^2 than less productive regions.

Map 5: GDP Density, 1900

„Map5_GDPdensityEU_1900.png"

Source: see text.

Map 6: GDP Density, 2010

„Map6_GDPdensityEU_2010.png"

Source: see text.

Figure 1 plots the cross-sectional correlation between population density and GDP density over time. While the correlation between the density of population and GDP is generally high, it clearly declines over time as reflected in the linear trend (dotted line). Put differently, variation in GDP per area is increasingly due to variation in GDP per capita and less to variation in population density. Another effect visible in figure 1 is the economic turmoil after World War 2. In 1950 several densely populated regions have quite low GDP per capita, partly due to destruction such as bombing, and partly due to migration of working population. Already by 1960 this effect has largely disappeared, which confirms the resilience of economic activity in certain regions. The drop in correlation in 1990 is due to the collapse of the GDR economy in the wake of German unification, which was followed by a strong recovery thereafter.

Figure 1: Correlation between Population density and GDP density, 1900-2010
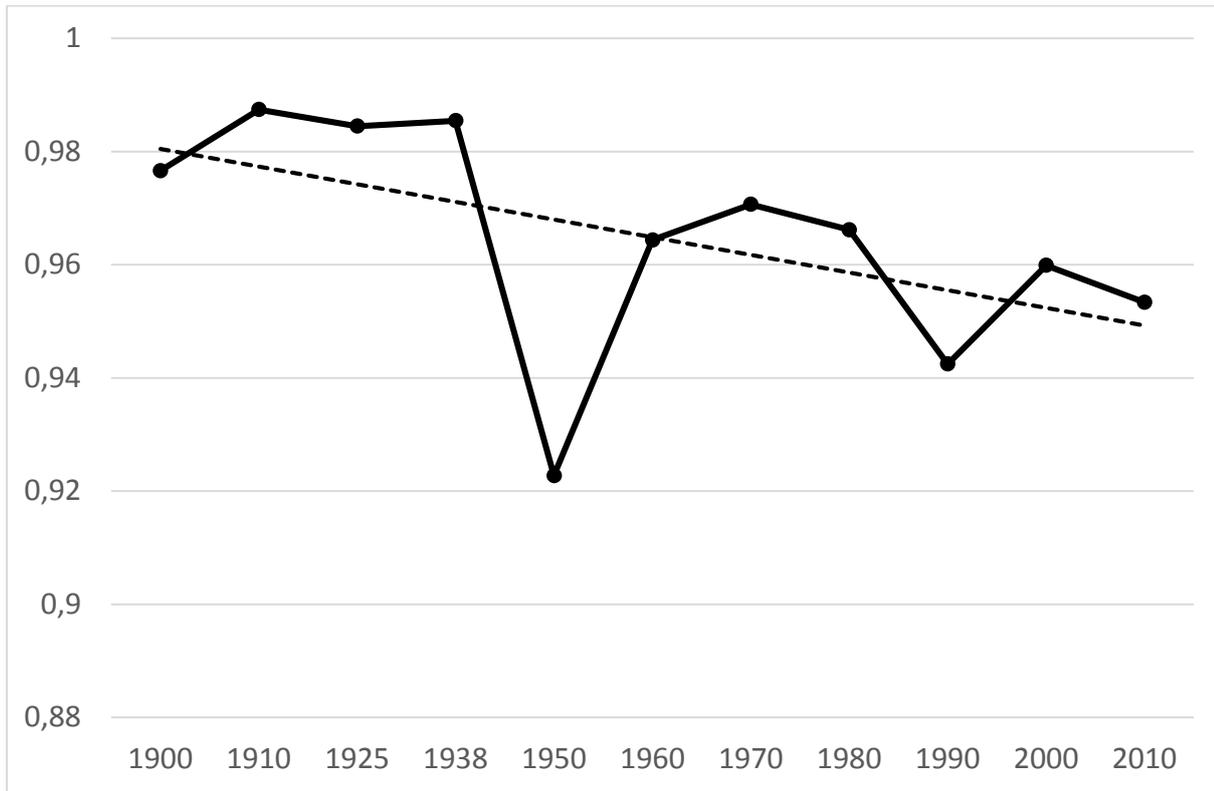
Source: see text.

4. Changing Employment Structure: structural change and localization

Before we explore the dynamics of GDP per capita, let us consider changes in employment structures across European regions over the last century. A characteristic of Europe's economic development during the 20th century – and of economic development more generally - was the continuation of structural change, with labour leaving agriculture to find employment in industry, mining and services (Broadberry 1997, Broadberry et al 2010). Figure 2 shows the evolution of average employment shares across European regions, 1900-2010 for three broad sectors agriculture, industry (incl. mining) and services. It is evident that the decline of agriculture was due not only to the expansion of industry, but already early on to an equal expansion of services. After 1960 industrial employment had reached its peak and started a rapid decline, both in absolute terms and as a share of total employment. So, by 2010, the share of industrial employment was about one fifth less than in 1900.

Figure 2: Sectoral Employment shares across 173 European regions, 1900-2010
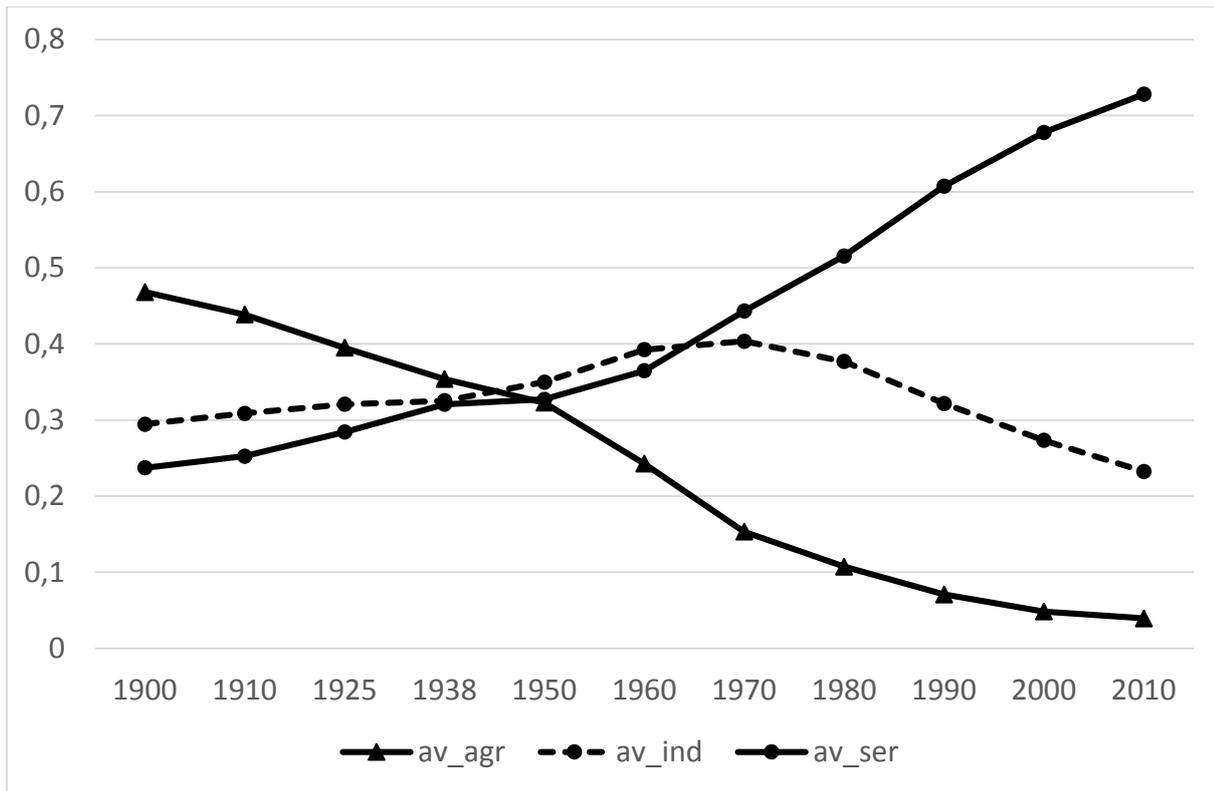
Source: see text.

However, this structural change occurred at very different speed across European regions. A way to capture the variation in sector-specific employment across regions is to use the location quotient, defined as

$$lq_i^k = \frac{x_i^k / \sum_k x_i^k}{\sum_i x_i^k / \sum_i \sum_k x_i^k} = \frac{x_i^k / \sum_i x_i^k}{\sum_k x_i^k / \sum_i \sum_k x_i^k},$$

where x_i^k is employment in region i in sector k. This can be read as either the specialization of region i in sector k, normalized by the overall share of sector k in total employment or as the concentration of employment in sector k, normalized by the overall share of region i in total employment. To summarize this evidence on "localization" for 173 regions, 11 years and three sectors, figure 3 shows the coefficient of variation over the period 1900 - 2010.

Figure 3: the coefficient of variation in localization of agriculture, industry and services, 1900-2010
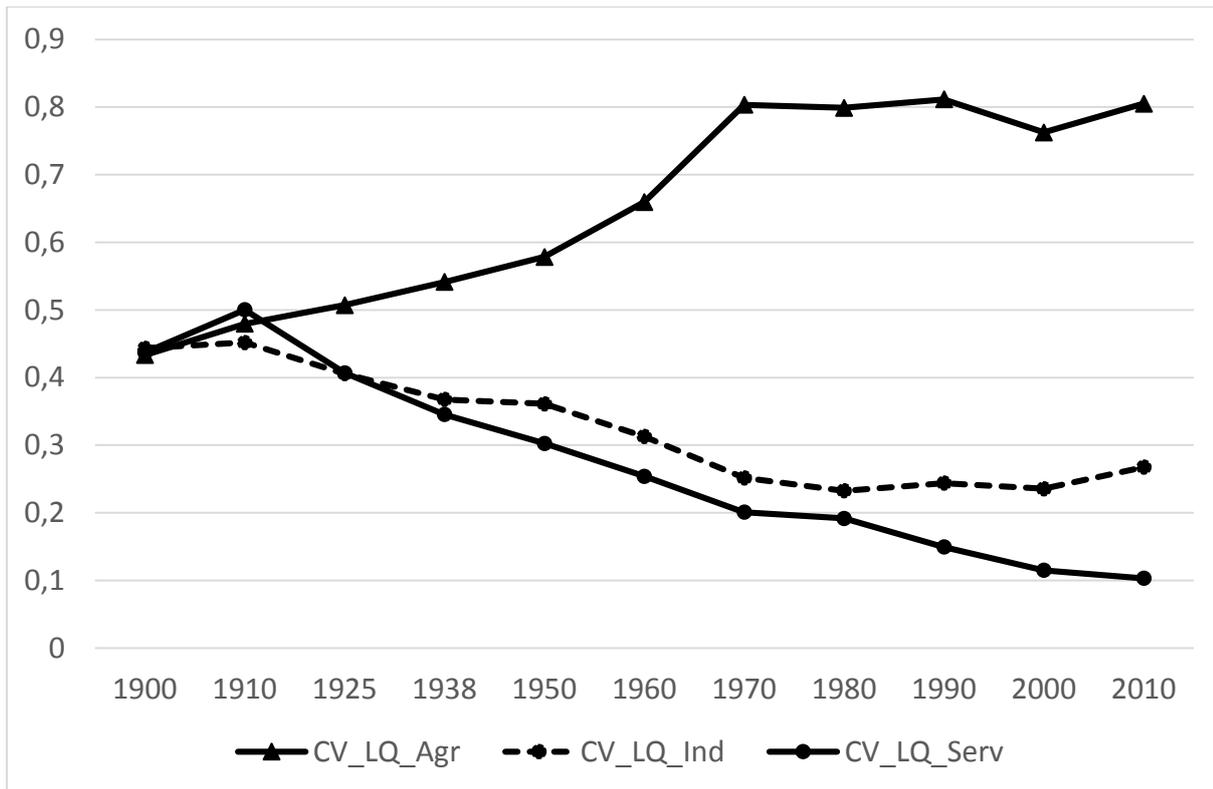
Source: see text.

Around 1900, the overall dispersion in agriculture, industry and services was still very similar: most regions have some employment in each sector, typically with the largest share in agriculture. Already the interwar period is different. We see simultaneously the spread of industry and services and a concentration of agricultural employment. This development is intensified after 1945, with increasing differences in the localization of agriculture across regions and increasing similarities in the localization of both industry and services. From about 1980 onwards, we can observe a stabilization of a new pattern of sectoral localization. A few regions are strongly specialized in agriculture, notably in southern Europe like Galicia (ES11), Alentejo (PT18), Extremadura (ES42), or Basilicata (ITF5). These regions are characterized by their overall remoteness from economic centres, their persistent backwardness in terms of GDP per capita and by the fact that they never developed a significant industrial or service sector. In contrast, many other regions by then have virtually no employment in agriculture.

The localization of industry is much less diverse. Some regions around 1980 are still strongly dominated by industrial employment, often associated with the automobile industry like Franche Comté (FR43), Thüringen (DEG0) or Stuttgart (DE11). Industry localization is markedly clustered in Germany, Eastern France and Northern Italy. A notable feature here is that some of those regions that kept a strong localization in industry after the 1970s, and which all used to be economically advanced, were falling back in terms of GDP per capita over the next decades. The correlation between income and industrial localization that had

been strongly positive for generations, started to disappear in the 1970s and turned weakly negative thereafter. A rather extreme example is the region of Hainaut (BE32) in Belgium that showed some of the highest industrial employment shares and industrial localization in Europe before the First World War and until the 1950s accompanied by high levels of GDP per capita. The region experienced a dramatic economic decline afterwards, with the lowest average annual growth rate over the century in our sample (see table 1).

The localization of services in turn was for a long time dominated by capital regions. Not only employment in public services, such as the government was concentrated in the capital, but also many private service providers such as banks had their headquarters and the bulk of their employees there. The institutional framework of the various nation states had a strong effect here, especially until 1950. As expected, capitals of more centralized nation states such as Paris in France concentrated a much higher share of service employment relative to their overall employment shares than capitals in less centralized states such as Bern in Switzerland. However, with the general increase in service employment, due to both outsourcing and growth of the public sector, we observe a spread of service employment over all regions and a strong convergence in overall services localization. To be sure, within the large and growing service sector there is a strong concentration of more specific types of services such as financial services in large urban agglomerations (Gallego and Maroto 2015, Deza and Gonzalez Lopez 2014).

5. Growth and Variation in GDP per capita

The focus of our interest is on the development of GDP per capita over time and its variation across regions, which summarizes the average level of material living standards. This indicator and especially its regional dimension is crucial for a better understanding of the European economy. It shows where income is generated and what scope there is for interregional transfers. It also shows, to what extent regional economies have become more or less similar in terms of economic potential over the last century, after wars, disintegration and the stepwise process of European integration. And not at least, the variation in income across regions over time complements our knowledge about personal income and wealth inequality. The systematic pattern of regional convergence and divergence that we document here for the first time has far-reaching implications for economic policy.

Over the last generations, all regions experienced a remarkable economic development in terms of GDP per capita. Figure 4 shows the change in median GDP per capita, average GDP per capita, as well as the smallest and largest values across regions over the last century.

Figure 4: GDP per capita across regions, 1900-2010 (1000 GK$)
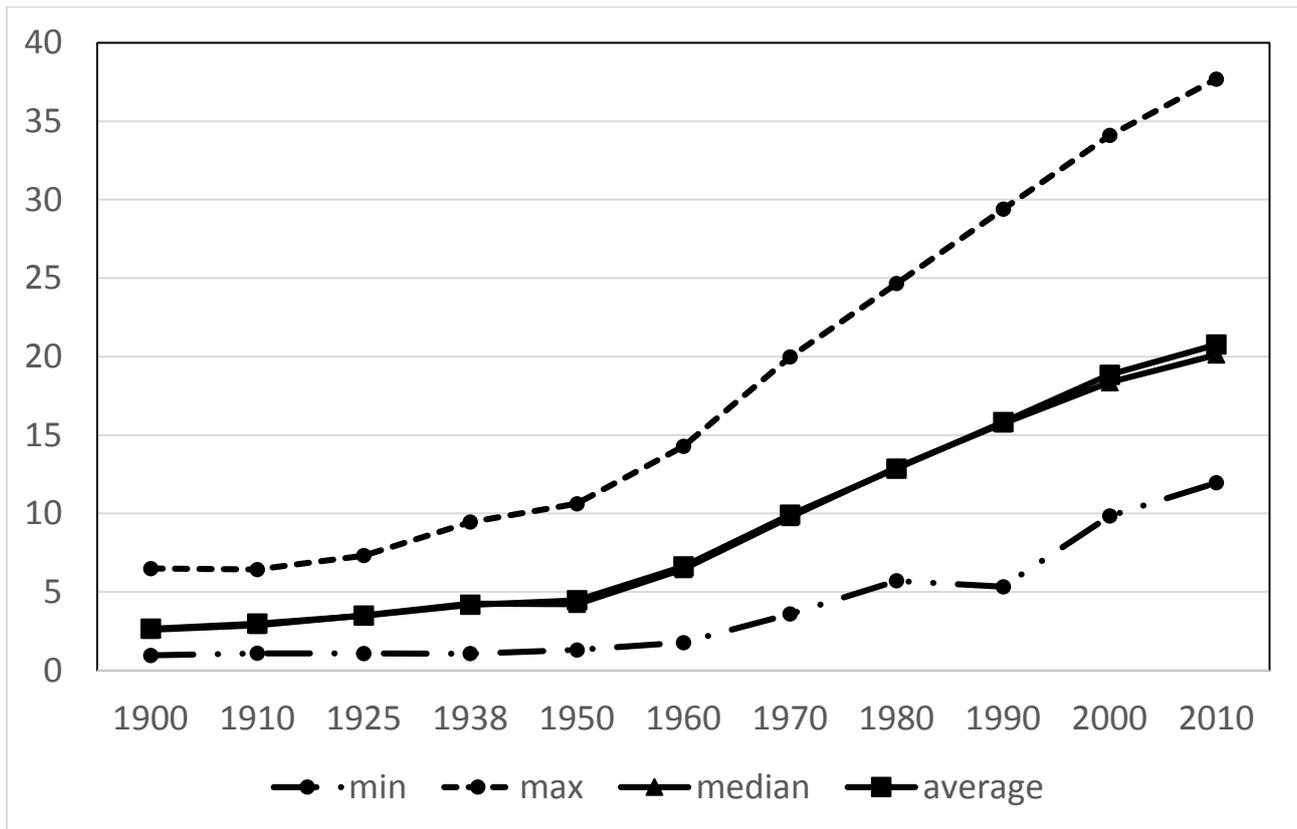
First of all, we see that the average level of GDP per capita has increased by more than 750% over the last century, in purchasing power parity, expressed in 1990 international dollars. This historically unprecedented increase in material living standards occurred mainly after the Second World War. Next, there was always substantial variation between regions, but until recently the average of GDP per capita was rather close to the median. As we see in the figure, the absolute distance between the poorest and the richest regions has sharply increased over time, but the differences have declined in relative terms. In any case it is surprising to see that in spite of wars and economic crisis the expansion in levels occurred at a very steady pace. The sudden collapse of industrial activity in regions of the former GDR is visible in the data, as the poorest region in our sample in 1990 is indeed Dresden (DED2). Afterwards, East-Germany experienced a strong recovery and convergence to West-Germany, yet it is still far behind its pre-war position relative to other parts of Europe or within Germany. On a European scale in contrast, our data show how the long-run trend of regional convergence came to an end in the 1980s. The small but growing difference between average and median is indicative of growing divergence.

A simple way to show the distribution dynamics in our data is to represent them in form of a histogram, where we divide for each year the regional GDP per capita data into evenly distributed bins as a

simple approximation of the underlying probability density function. Figure 5a shows the distribution for the years 1900, 1950 and 1980, figure 5b for 1980 and 2010.

Figure 5a: Regional Income Distribution 1900, 1950 and 1980

Source: own

Figure 5b: Regional Income Distribution 1980 and 2010
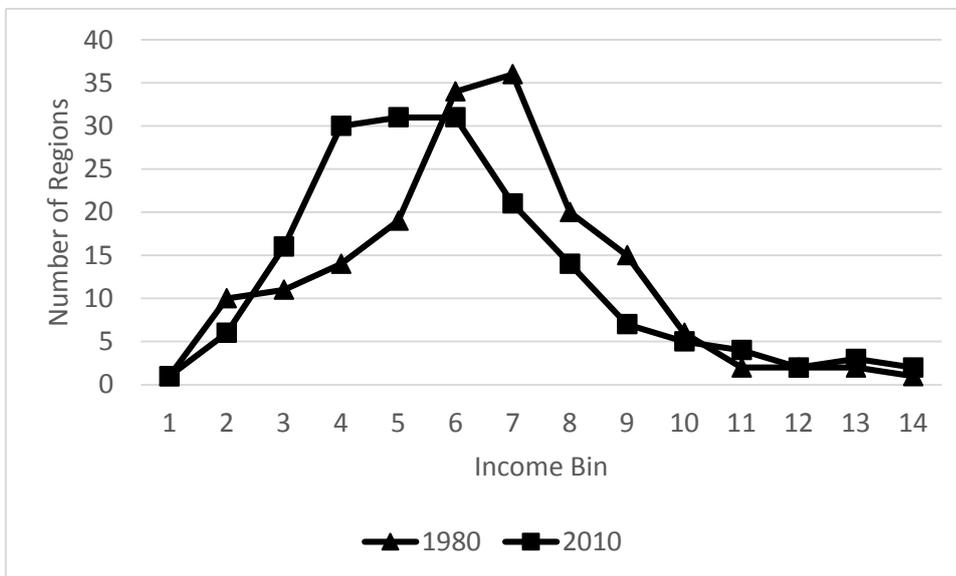
Source: own

We see in figure 5a how between 1900, 1950 and again 1950 and 1980 the distribution shifted quite systematically to the right, with a growing number of regions positioned around the median. Especially strong is the shift between 1950 and 1980, which will be mirrored in evidence about beta-convergence strong

14

enough to reduce overall dispersion in GDP per capita across regions (compare tables 3, 5 and figure 7 below). But this changed around 1980, when convergence weakened and dispersion across regions started to increase again. Figure 5b shows that between 1980 and 2010 the distribution shifted back to the left. There was an overall decline in growth rates after the 1970s, but this was clearly very uneven, where some rich regions maintained steady growth rates, while other started to fall behind. This emergence of a small club of growth centres has recently been highlighted by Iammamrino et al (2017), who link it to the rise of new economy industries and a new global division of labour.

Table 1 adds more detail to this. We see for example that some poor regions experienced above average growth rates, while some formerly rich regions such as Hainaut (BE32) in Belgium or Berlin (DE30) in Germany were falling behind. The former is due to structural change and the decline of traditional industrial regions, the latter is a result of dramatic institutional change in the wake of the Second World War, namely the division of Germany during the period 1949-1990.

Table 1: GDP per capita across regions, 1900-2010 – cross-sectional variation and growth

Levels in 1000 GK$90		1900	1950	2010
	Average GDP per Capita (in parentheses: population weighted average)	2.63 (2.92)	4.44 (4.62)	20.76 (21.39)
	Poorest Region	0.96 (Galicia , ES11)	1.31 (Extremadura, ES43)	11.95 (Calabria ITF6)
	Richest Region	6.49 (London Counties UKI)	10.62 (Zurich, CH04)	37.68 (Luxembourg LU00)
Growth Rates in %		1900-2010	1900-1950	1950-2010
	Average annual Growth Rates	1.91%	1.02%	2.66%
	Highest Growth Rates	2.62% (Galicia , ES11)	2.55% (Västsverige, SE23)	3.99% (Extremadura, ES43)
	Lowest Growth Rates	1.10% (Hainaut, BE32)	-0.01% (Berlin, DE30)	1.49% (Espace Mittelland, CH02)

Source: own.

Map 7: relative GDP per capita (EU average = 1), 1900

„Map7_relGDPpcEU_1900.png"

Source: see text.

Map 8: relative GDP per capita (EU average = 1), 2010

„Map8_relGDPpcEU_2010.png"

Source: see text.

Maps 7 and 8 show more systematically the variation in GDP per capita across regions. The first impression form these two maps is that the pattern of variation was more compact back in 1900 than in 2010. Broadly speaking, it is easy to detect a centre-periphery pattern: some macro-regions like England and North-western Europe were richer than the average, the regions of France and central Europe were close to the average, while several regions in Scandinavia and Southern Europe were poorer than average. In 2010 the picture is more complex. There are islands of prosperity, such as Paris (FR10) in France or Madrid (ES30) in Spain surrounded by regions with below average GDP per capita. We will see below that indeed the degree of spatial correlation has systematically declined over the last century, while the geographical concentration of economic activity has first declined and then again increased since the 1970s. In other words, the centre-periphery pattern prevalent in 1900 is vanishing rapidly and is being replaced by a system based on several central regions, which commonly comprise the largest metropolitan areas (see a further discussion of this phenomenon in section 8).

6. Explaining Economic Growth: adjustment and fundamentals
a) Exploring growth dynamics: convergence, structural change and reconstruction

How can we explain the enormous variation in economic activity between regions and their changes over time that is evident from section 5 above? The literature on economic growth distinguishes between factors that shape the short-to medium-run adjustment to a steady state level of growth and factors that shape economic growth in the medium- to long-run. We will use this as a guidance to describe the growth

pattern in our data, but we leave an in-depth analysis of economic growth across European regions for future work.

Let us start with a short-run perspective on growth dynamics. From the perspective of a simple Solow-growth model (Solow 1956, Barro and Sala-i-Martin 1992) we expect to find that poor regions tend to exhibit on average faster growth rates in GDP per capita. Barro and Sala-i-Martin found for the US a rate of convergence of around 2 per cent per year. The intuition behind this is that we expect a lower capital per labour ratio in poor regions and hence a higher return to investment in these regions, ceteris paribus. This is indeed suggested by table 1 above. The regions that were initially poorest (Galicia, ES11 and Extremadura, ES43) showed above average growth rates. The general approach is to regress the average annual growth rate in GDP per capita of a region over some period on the level of GDP per capita at the beginning of the period, or $\hat{y}_{i,t1-t0} = \alpha + \beta \ln(y_{i,t0}) + \varepsilon_i$. In figure 6 we plot the average annual growth rates of regions against their initial level of GDP per capita (in logs). Figure 6a shows the result if we consider the entire sample period 1900-2010, while figures 6b and 6c distinguish between the period before and after the Second World War.

Figure 6: beta-convergence over time (172 regions)

Figure 6a: 1900-2010

17
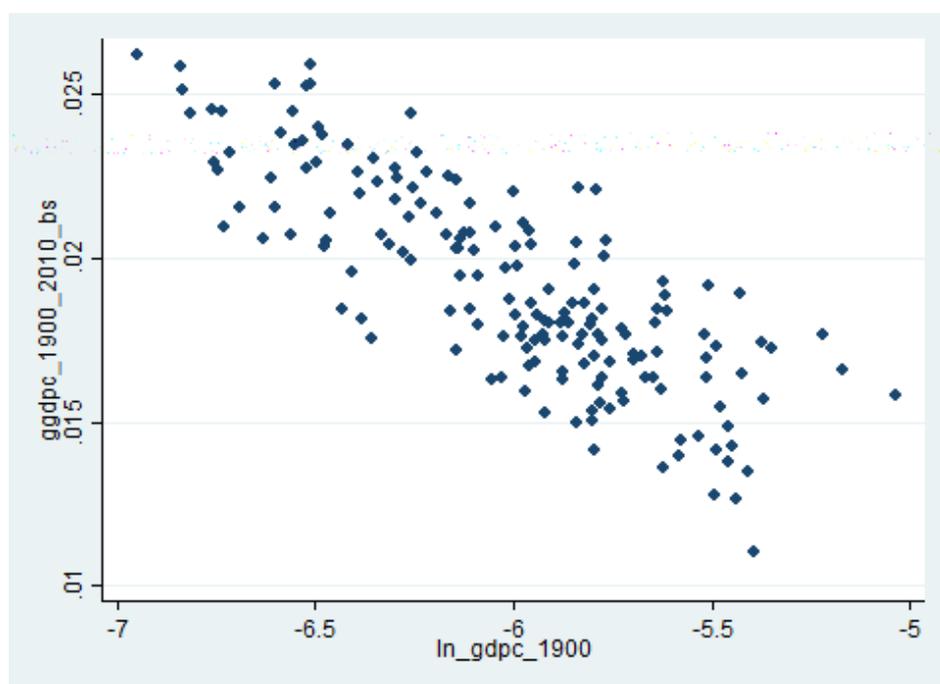
Figure 6b: 1900-1938
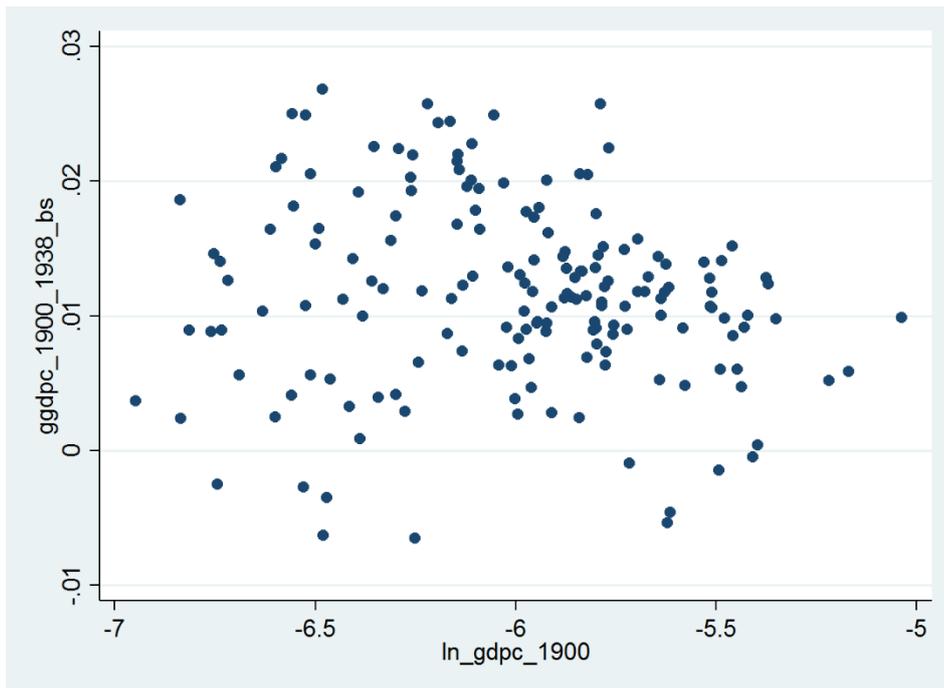
Figure 6c: 1950-2010
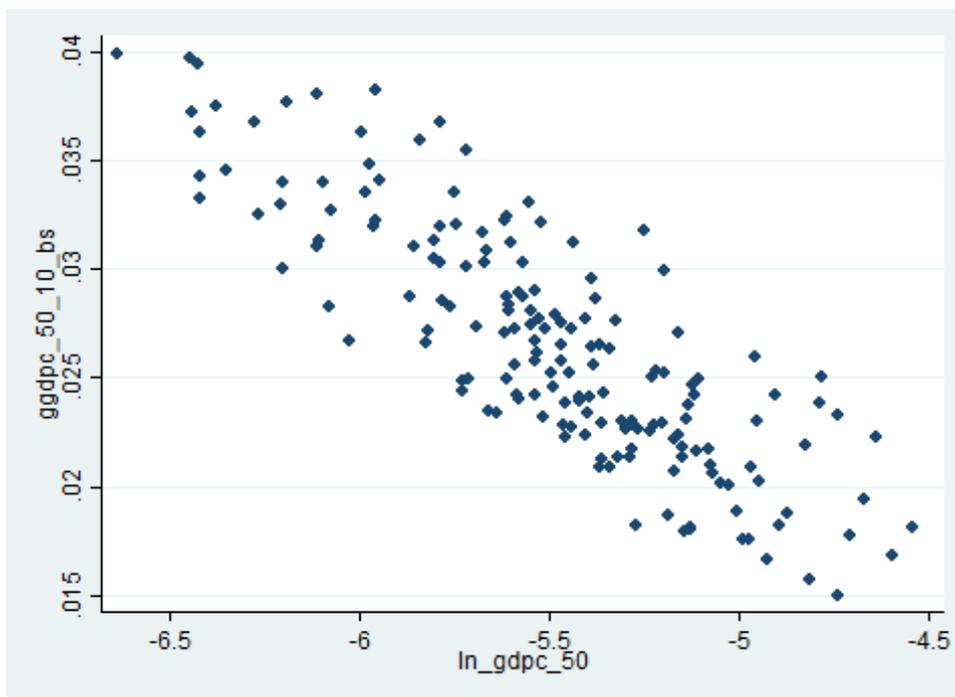
Figure 6a suggests that in the long-run regions converged in the sense that growth rates in GDP per capita were systematically higher the lower the initial level of GDP per capita. However, figures 6b and 6c indicate that this view is incomplete. Clearly, convergence was very weak before 1945 (with a beta-coefficient of -0.002, just about significant at 10%), but turned quite strong after the Second World War (beta coefficient

18

of -0.01, significant at 1%). This absence of convergence before the war is remarkable, and it is worthwhile to look into the details of this. In table 2 we show how the beta-coefficients changed over shorter periods of time, notably before and after the two World Wars. The last column shows the implied speed of convergence.

Table 2: Explaining regional growth rates: convergence over time

|	$\widehat{\beta_0}$ (t-stat)	Speed of convergence (per cent)
1900-1910	-0.009 (3.89)	0.94
1925-1938	0.002 (0.44)	-0.20
1950-1960	-0.015 (5.38)	1.63
1960-1970	-0.026 (13.37)	3.01
1970-1980	-0.024 (9.36)	2.74
1980-1990	-0.009 (1.88)	0.94
1990-2000	-0.026 (9.81)	3.01
2000-2010	-0.009 (2.93)	0.94

Source: own.

While the pattern of convergence changed substantially over time, we find that similar forces were at work in the decade before the First World War and the last decade of our series (2000-10), with low but highly significant rates of convergence (closest to one percent, half of the rate found elsewhere). In contrast, the interwar years were characterized by a divergence of regional incomes. Some regions grew strongly and many others stagnated, but there is no detectable relationship to initial levels of growth as suggested by a bare-bones Solow-model. Elsewhere we have argued that this very likely reflects the political tensions between European states after 1918 that limited market forces such as trade, migration and capital flows as well as technology diffusion (Roses and Wolf 2010). After 1945 there is ample evidence for strong convergence. This was particularly fast during the 1960s and 1970s (well above 2 percent), before it slowed down during the 1980s, coming back to the fastest rates during the 1990s.

Given that we consider regions and nation states in their modern borders, our data includes also the regions of the former socialist GDR as part of modern Germany. One might suspect that this should affect the estimated beta-coefficients, because the post-war convergence may have been absent in these regions, while the same regions experienced first a collapse followed by an unusual period of catch-up growth after the fall of the Iron Curtain in 1989 and German unification in 1990. What is more, all GDP estimates for socialist states are questionable, and even more so any regional breakdown of such estimates. An exclusion of the GDR regions only affects the results for the decades 1980-2000 somewhat, but does not change the

overall pattern. We find that regions converged on average at a rate of between below 1% (1900-1910 and 2000-2010) and around 2.5% per year during the1960s and 1970s (not shown here).

Barro and Sala-i-Martin (1995) suggested to extend the simple Solow-framework to take aggregate shocks into account. Along these lines, Temin (2002) argued to consider the period between 1950 and 1980, often termed the "Golden Age of Growth" (Crafts and Toniolo 1996), as a period of disequilibrium. The evidence from table 2 suggested that much of this growth was due to a process convergence, where poorer parts of Europe were catching up to the rich. However, some regions had suffered strong destruction during the war, alongside with massive population movements, while economic reintegration after 1945 allowed for increased technology diffusion and structural change. From this perspective, some regions in Germany, but also in Austria, Italy, Belgium, the Netherlands and Eastern France should have seen economic growth driven by reconstruction and to some extent structural change rather than convergence as suggested in a Solow-framework. We apply this idea to our data and construct a variable "gap" defined as the log-difference in GDP per capita 1938 and 1950 for a region. For example, a region that was strongly negatively affected by the war due to bombing, immigration of refugees or a combination thereof, such as Hamburg (DE6) would show a positive gap. We would expect that such a region would grow faster after the war due to efforts to reconstruct housing and infrastructure. Related, we add a control for the share of agriculture in total employment of a region in 1950 to test for the idea of "arrested development" (Temin 2002) – the economic reintegration and trade after 1945 might have facilitated structural change and freed up labour in agriculture to move to more productive employment in industry or services. Hence, we re-estimate the relationship between the average annual growth rate of GDP per capita, the initial level of GDP per capita (in logs), controlling for the "gap" between 1938 and 1950 and the share of agriculture in total employment of a region in 1950, or $\hat{y}_{i,t1-t0} = \alpha + \beta_0 \ln(y_{i,t0}) + \beta_1 gap_{i,1938\_50} + \beta_2 ag\_sh_{i,1950} + \varepsilon_i$.

We vary the period under consideration, starting with average annual growth rates in 1950-60, and extending the period until 1950-90. Table 3 shows the results (where we exclude the regions of the former GDR to avoid the distortions caused by economic planning).

Table 3: explaining regional growth rates: convergence, reconstruction and structural change, 1950-1990

	$\widehat{\beta_0}$ (t-stat) initial GDP per capita$_{1950}$	$\widehat{\beta_1}$(t-stat) "gap"$_{1938\_50}$	$\widehat{\beta_2}$(t-stat) ag_sh$_{1950}$	Adj. R^2	Observations
1950-60	-0.007 (-1.77)	0.045 (7.92)	0.009 (1.01)	0.37	164
1950-70	-0.017 (-7.25)	0.016 (5.00)	0.002 (0.47)	0.57	164
1950-80	-0.016 (-9.77)	0.014 (6.08)	0.003 (0.87)	0.72	164
1950-90	-0.014 (-11.89)	0.009 (5.75)	0.002 (0.71)	0.77	164

Source: own.

Overall, these results are in line with those of Vonyó (2008), who used panel data techniques on a sample of 21 nation states, including Japan, Australia, Canada, New Zealand and the United States. Convergence mattered, but it was strongly affected by post-war reconstruction. All results are robust to a modification, where we restrict "gap" to those cases where GDP per capita in 1938 was actually above the level in 1950 (not shown here). Instead, the contribution of structural change to growth after 1945 is less clear. A closer inspection of the regional data suggests that for regions outside of West-Germany there was some positive contribution of structural change to growth beyond reconstruction and convergence (not shown here). We note however, that all these results are little more than correlations, because the initial level of income in 1950, the share of agriculture in 1950 as well as the "gap" are all related to each other and would require a more detailed analysis, which is beyond the scope of this paper.

b) A long-run perspective: institutions and geography

Let us now adopt a long-run perspective with regard to our time span of 110 years of regional development. Which regional characteristics can account for the large variation in growth rates since 1900? Following the literature, we can distinguish broadly between institutional factors and geographical factors, which shape the incentives to invest and adopt new technologies and hence economic growth (Acemoglu 2009, ch. 1). With regard to the development of European regions a distinction between national and supra-national institutions suggests itself, to see how for example the process of European integration after 1945 affected growth rates. Specifically, we control for a set of national dummies, varying over time in order to capture very broadly institutional differences between nation states as well as European dummies, whenever a state entered the EEC, the EU or the Eurozone. To capture at least some of the institutional variation within nation states, we add a dummy for capital regions.

When it comes to geographical factors, it is common to distinguish between first and second nature characteristics: the former are factors that can be considered to be exogenous or given by "nature" (at least

over the time horizon considered here) such as climate (captured by extreme values in average temperature and rainfall), soil quality (captured by the average caloric value of crops per hectare), access to coal fields (measured as the proximity to rock strata from the carboniferous era) or the location of main sea ports (captured by the distance to the nearest deep water seaport). Instead, second nature geography refers to factors that result from human intervention, notably the accessibility of markets which depends on both the economic size of neighbouring regions and access to them (Fujita et al. 1999, Redding and Venables 2004), In a simple first step we can test for the respective role of each of these factors to explain variation in the level of GDP per capita across our regions and over time. Table 4 shows the results of a pooled regression that uses all our data. Note that we use the sum of inverted distances to other regions as an instrument for market access.

Table 4: Pooled GLS regression with Random Effects, IV, Dep Variable: Ln(GDP per capita), 1900-2010

	Coefficients (t-stat)
European Community	0.112 (6.16)
European Union	0.141 (2.42)
Euro-Zone	0.225 (4.43)
Capital Region	0.305 (3.70)
National Dummies (time-varying)	Yes
Market Access (Instrumented)	0.280 (1.77)
Distance to Deepwater Port (ln)	-0.016 (-0.73)
Extreme Rainfalls	-0.073 (-0.89)
Extreme Temperatures	-0.172 (-2.04)
Soil Quality	-0.069 (-1.68)
Coal Potential	0.290 (2.33)
Constant	-7.66 (-7.25)
Time Dummies	Yes
Observations	1886
Groups	172
Adj. R^2 within, between, overall	0.96, 0.48, 0.89

Source: own. Note: the country-time dummies allow for country-specific linear time-trends.

We have added random effects, clustered standard errors at the regional level, and allowed for common time effects. With this rather naïve approach, where we simply pool all data over regions and time, we find that regions with good access to coalfields tend to have higher income levels in the long-run, while good soil quality and extreme temperatures are associated with lower income levels. In turn, a good accessibility of markets seems to exert a strong positive effect on income levels (note that this variable has

been instrumented in order to deal with endogeneity). After taking geographical factors into account we also find support for the role of institutions: controlling for unobserved factors at the national level we find that capital regions have always higher levels of income and that membership in the various European institutions made a positive contribution to income levels, notably after controlling for a full set of time effects. Obviously, this is only a first explorative glance at the data but it suggests some regularities in line with the idea that a combination of institutional and geographical factors has shaped the economic geography of Europe in the long-run. Note also, that all results are robust to excluding the regions, which were part of the GDR during the period of Germany's division (1949-1989). By closer inspection we can see that these effects are indeed quite stable over our period 1900-2010. An interesting exception to this rule is the effect of access to coalfields, which is becoming much weaker after 1945.

This motivates us to consider in a next step the short- to medium term dynamics of growth rates as a process of adjustment, conditional on more persistent geographical and institutional factors. Table 5 gives the results of pooled regressions, where we regress the average annual growth rates of regions each decade on the levels of GDP per capita at the beginning of each decade, the share of agriculture at the beginning of each decade, and the "gap" defined by the difference in GDP per capita in each region before and after a war. Here we distinguish between destruction during the first and the second world war and restrict the gap to affect only the immediate decade after the war. In column 1 we repeat the exercise of table 3 for the years 1950-2010, but now exploiting the panel structure of our data. In column 2 we extend the analysis to all decades 1900-2010. In column 3 we include the geographical and institutional factors. In each case we exclude the regions of the GDR.

Table 5: Pooled GLS regression with Random Effects, IV, Dep. Variable: Ln(Growth Rates), 1900-2010

	Coefficients (t-stat)	Coefficients (t-stat)	Coefficients (t-stat)
Ln(GDP per capita), t_0	-0.014 (-5.78)	-0.0146 (-7.37)	-0.019 (-9.48)
"gap" (1910-1925)	-	0.035 (3.18)	0.032 (3.40)
"gap" (1938-1950)	0.044 (7.26)	0.044 (3.69)	0.044 (7.58)
Share of agriculture in employment, t_0	0.042 (0.66)	-0.012 (-3.18)	-0.010 (-2.83)
European Community	-	-	0.007 (4.03)
European Union	-	-	0.006 (2.17)
Euro-Zone	-	-	-0.004 (-1.88)
Capital Region	-	-	0.004 (2.96)
National Dummies (time-varying)	yes	yes	Yes
Ln(Market Access), Instrumented	-	-	0.009 (2.71)
Ln(Distance to Deepwater Port)	-	-	0.001 (1.99)
Extreme Rainfalls	-	-	-0.002 (-1.44)
Extreme Temperatures	-	-	0.003 (0.32)
Ln(Soil Quality)	-	-	-0.002 (-4.66)
Ln(Coal Potential)	-	-	0.001 (0.38)
Constant	0.000 (0.10)	-0.072 (-6.76)	-0.143 (-6.92)
Time Dummies	yes	Yes	Yes
Observations	994	1650	1650
Groups	166	165	165
Years	1950-2010	1900-2010	1900-2010
Adj. R^2 within, between, overall	0.66, 0.65, 0.66	0.59, 0.36, 0.57	

Source: own. Note: the national dummies allow for country-specific linear time-trends

The results in table 5, column 1 are quite similar to those from table 3 above. For the period after the Second World War, regions did show beta-convergence. This result is virtually unchanged if we take various conditioning factors into account, notably reconstruction after wars, which apparently mattered both after 1918 and 1945, and several geographical and institutional factors. Results in column 3 of table 5 show that

most of our findings for the long-run determinants of regional income carry through to a panel analysis of regional growth rates, but with some revealing exceptions.

Regions with a high share of employment in agriculture in time t_0 tend to grow systematically less, and having above average soil quality tends to be harmful. An explanation for this could be specialization along differences in endowments. Interestingly, the positive effect of coal potential that we found on levels of GDP per capita (table 4) does not show up in a growth framework. Another result is that membership in the Euro was not beneficial for growth, in contrast to strong positive growth "effects" of the earlier steps of European integration. However, this should be interpreted with caution, because our time frame might be too short to test this (our data ends in 2010) and for a proper analysis we would have to address issues of selection bias, among other things (see Persson 2001, Ritschl and Wolf 2011).

7. Regional Rankings over time

Another way to look at our data is in terms of rankings (compare figures 5a and 5b on distribution dynamics). Which regions were on top of the league, which at the bottom? And especially, which regions gained relative to others and which ones were falling behind? This can help to interpret the abstract estimation results and relate them to other evidence on particular regions. Table 6a shows the ten richest regions at four points of time: 1900, 1938, 1950 and 2010, each time with the GDP per capita estimate in parentheses. Table 6b repeats the same exercise for the ten poorest regions. In all cases we exclude Flevoland (NL23), for which we have data only from 1970 onwards.

Table 6a: Ten Richest Regions, GDP per Capita ((1990 Int. GK$))

Rank of 172	1900	1938	1950	2010
1	London Counties (UKI) 6.489$	London Counties (UKI) 9.448$	Zürich (CH04) 10.618$	Luxembourg (LU00) 37.683$
2	Île de France (FR10) (5.688$	Stockholm (SE11) 8.132$	Nordwestschweiz (CH03) 10.100$	London Counties (UKI) 36.844$
3	Luxembourg (LU00) 5.412$	Berlin (DE30) 7.576$	London Counties (UKI) 9.672$	Île de France (FR10) 35.371$
4	Helsinki-Uusimaa (FI1B) 4.743$	Zürich (CH04) 7.549$	Hovedstaden (DK01) 9.365$	Hamburg (DE60) 35.123$
5	SouthEast England (UKH+UKJ) 4.653$	SouthEast England UKH+UKJ 7.441$	Région lémanique (CH01) 9.006$	Stockholm (SE11) 34.725$
6	Zürich (CH04) 4.630$	Oslo og Akershus (NO01) 7.351$	Île de France (FR10) 8.740$	Brabant (BE10 + BE24) 33.393$
7	Hainaut (BE32) 4.537$	Île de France (FR10) 7.106$	Espace Mittelland (CH02) 8.739$	Groningen (NL11) 33.337$
8	North England (UKC+UKD) 4.482$	Nordwestschweiz (CH03) 7-083$	Luxembourg (LU00) (8.393$	Helsinki-Uusimaa (FI1B) 31.705$
9	Bremen (DE50) 4.420$	Helsinki-Uusimaa (FI1B) 6.871$	Stockholm (SE11) (8.319$	Zürich (CH04) 31.420$
10	Hamburg (DE60) 4.386$	West Midlands (UKG) 6.747$	Ostschweiz (CH05) 8.112$	Hovedstaden (DK01) 30.085$

Source: own.

Table 6b: Ten Poorest Regions, GDP per Capita ((1990 Int. GK$)

Rank of 172	1900	1938	1950	2010
163	Abruzzo (ITF1) 1.210$	Región de Murcia (ES62) 1.541$	Canarias (ES70) 1.879$	Sardegna (ITG2) 14.240$
164	Calabria (ITF6) 1.191$	Centro (PT16) 1.539$	Norte (PT11) 1.741$	Mecklenburg-Vorpommern (DE80) 14.221$
165	Corse (FR83) 1.187$	La Rioja (ES23) 1.507$	Región de Murcia (ES62) 1.695$	Thüringen (DEG0) 14.063$
166	Extremadura (ES43) 1.177$	Alentejo (PT18) 1.496$	Calabria (ITF6) 1.631$	Norte (PT11) 13.827$
167	Molise (ITF2) 1.167$	Andalucía (ES61) 1.357$	Castilla-La Mancha (ES42) 1.631$	Centro (PT16) 13.283$
168	Algarve (PT15) 1.159$	Castilla y León (ES41) 1.319$	Basilicata (ITF5) 1.628$	Basilicata (ITF5) 12.702$
169	Región de Murcia (ES62) 1.097$	Castilla-La Mancha (ES42) 1.209$	Algarve (PT15) 1.618$	Puglia (ITF4) 12.458$
170	Canarias (ES70) 1.075$	Canarias (ES70) 1.178$	Andalucía (ES61) 1.589$	Sicilia (ITG1) 12.249$
171	Nord-Norge (NO07) 1.072$	Galicia (ES11) 1.105$	Galicia (ES11) 1.586$	Campania (ITF3) 11.971$
172	Galicia (ES11) 0.959$	Extremadura (ES43) 1.071$	Extremadura (ES43) 1.306$	Calabria (ITF6) 11.951$

Source: own.

Obviously, all of these results need to be interpreted with caution, because the underlying data for both the historical and more recent periods are incomplete and to some extent uncertain. For example, the difference between Bremen (DE50) and Hamburg (DE60) around 1900 is easily within the margin of error. Nevertheless, there is a clear group of regions that is systematically ahead of the others: London (UKI), Paris - Île de France (FR10) and Zurich (CH04) are always among the top-10 regions, Luxembourg (LU00), Stockholm (SE11) and Helsinki (FI1B) in three of the four years. But there are some remarkable changes. Berlin (DE30) raced up the table between 1900 (position 14) and 1938 (position 3), before it would start a long decline due to the war and its consequences. Hainaut (BE32) in Belgium in contrast, was a rich region in 1900 (position 7) and still in 1910 (position 9), before it started its decline after the Great War to position 70 (1938), and further down to 153 (2010). While London (UKI) kept its strong position over time, other parts of England and indeed Wales (UKL) declined in the ranking very considerably.

Table 6b shows the bottom end of the distribution. As expected, nearly all of these regions are in southern Europe (but note that we have no data for Greece), and the differences between them are often too small to have a meaningful interpretation. Our earlier findings suggest that a combination of poor access to markets, poor climatic conditions together with weak national institutions contributed to their relative underdevelopment. If we compare the situation of 1950 to that in 2010, we see that the Spanish regions did relatively better than others in the periphery and that some parts of the former GDR in Germany have fallen behind in relative terms to levels similar to those in Portugal and Southern Italy.

Table 7 explores these changes in position more explicitly. The first column shows the "winners", ten regions in our sample that improved their position most strongly, the second column shows those that lost most in the ranking over time.

Table 7: Ten Biggest Winners and Losers, 1900-2010

| Winners		Losers	
Improvement in Rank	Name	Decline in Rank	Name
124	Comunidad Foral de Navarra (ES22)	146	Hainaut (BE32)
114	Provincia Autonoma di Bolzano/Bozen-Provincia Autonoma di Trento (ITH1+ITH2)	123	Namur (BE35)
113	Agder og Rogaland (NO04)	119	Luxembourg (BE34)
111	Åland (FI20)	115	Liege (BE33)
103	Vestlandet (NO05)	107	Chemnitz (DED4)
92	Tyrol (AT33)	99	Wales (UKL)
88	Stuttgart, Regierungsbezirk (DE11)	94	Leipzig (DED5)
85	Valle d'Aosta/Vallée d'Aoste (ITC2)	89	Haute-Normandie (FR23)
84	Emilia-Romagna (ITH5)	88	Nord - Pas-de-Calais (FR30)
84	Västsverige (SE23)	88	Dresden (DED2)

Source: own.

The results are revealing and given the extent of changes we see, they are very likely economically and statistically significant. From the first column we see that regions in many different countries have

improved their position relative to others in Europe. Regions in Italy, but also Scandinavian countries, as well as Stuttgart (DE11) in Germany and Navarra (ES22) in Spain are in this group. Among the biggest winners are regions that have some autonomy within their country. It would be interesting to analyse if this is cause or consequence of improvement. On the other side, the second column shows how dramatically some European regions were falling back over the last century. The formerly highly industrialised regions of Wallonia in Belgium, but also Wales (UKL) in the UK and Nord-Pas-de-Calais (FR30), Haute-Normandie (FR23) in France that had been reliant on abundant coal resources and to some extent on textile industry were declining with the depletion of resources and changing technology and global competition. The case is different for the three German regions which experienced extreme decline, all of them located in Saxony. While these regions were around 1900 and still in 1938 among the richest in Germany, with a strong industrial sector but also substantial employment in services, they suffered from the combination of division of Germany after 1945 and massive deindustrialisation in the wake of unification.

Overall, we see that many regions in Europe do not fit easily into a picture of growth and convergence. Some stayed at the top of the league over more than 100 years, others grew much faster than average, while a third group was quite systematically losing out. Among those who did poorly are several of the former coal-mining regions.

In contrast, the capital regions of Europe did typically rather well, even though their share declined during the "Golden Age" of growth and convergence. Figure 7 shows how the share of capital regions in total sample GDP developed over time.

Figure 7: Share of Capital Regions in Total Sample GDP
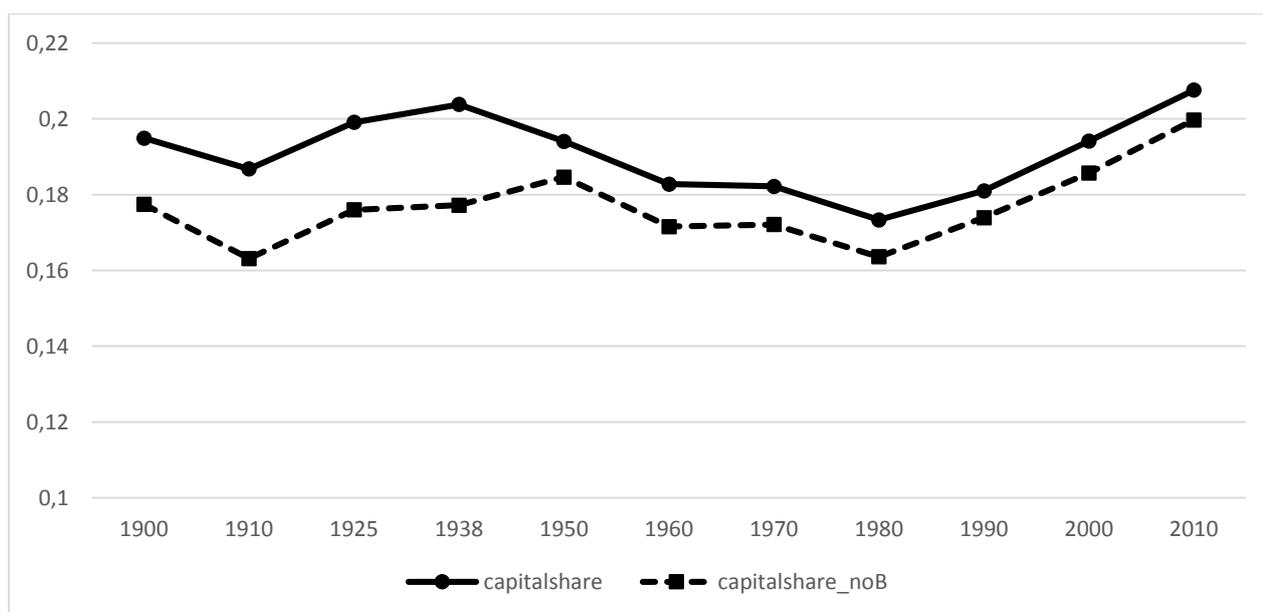
We see that the share of capital regions in the GDP of Europe was around 20 percent, which declined after 1945 in line with our findings of convergence after the World War Two. The findings are unchanged if

we exclude Berlin, which is a rather special case due to the division of Germany and the city itself and the decision to declare Bonn as temporal capital of West-Germany. Around 1980 things start to change, and the capital regions seem to increase their economic size relative to others – a development that is also reflected in population and employment data (not shown). In the next section we will look more closely into such geographical changes. The decline and rise of capital regions is reflected in broader trends of spatial divergence and concentration over the last century.

8. Back to space: dispersion, spatial correlation, and concentration over time

Given the evidence on beta-convergence after 1945 from tables 3 and 5, one might expect to find that the overall dispersion of GDP per capita across regions should have declined. A simple measure to capture this is the coefficient of variation (the cross-sectional standard deviation divided by the mean), often referred to a "sigma-convergence". Following Williamson (1965) we use a weighted version of this, where each region enters with a weight according to its population share. Figure 8 shows the result of this for all regions and for regions aggregated to nation states. The difference highlights how much of the variation is typically missed if researchers use national instead of regional data.

Figure 8: sigma-convergence (population weighted coefficient of variation), 1900-2010
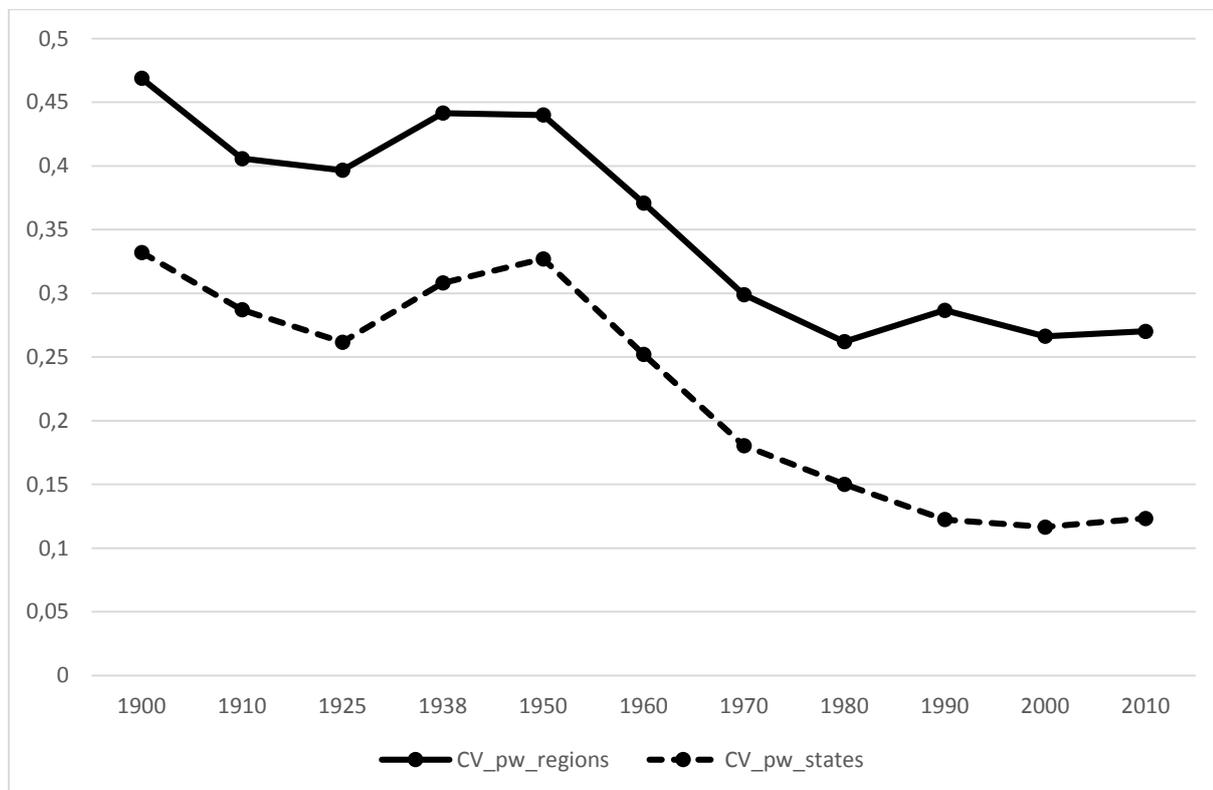

Source: own.

According to this measure, dispersion in GDP per capita has declined, at least until about 1990 and stagnated thereafter. It is evident that measures based on regional data show more dispersion, but we also observe some differences in trends. While on average for the period 1900-1960, measures based on national figures alone would capture around 70% of all underlying variation, this figure has declined to 60% in 1970 and below 50% after 1990. One reason is the very high degree of variation within Germany, which increased between 1950 and 1980, increased again with unification, and declined only slowly thereafter. Another, more general reason was visible from our maps on GDP per capita for 1900 and 2010. In spite of strong overall growth and a systematic convergence of backward regions during the "Golden Age" period, the coefficient of variation hides a process of declining spatial correlation, namely that many neighbouring regions are becoming actually less similar over time. We have seen above that there was substantial turbulence, in the sense that some regions were improving their relative position, while others were falling behind, sometimes even within the same country.

A relatively simple way to capture similarity between neighbours is Moran's I. Applied to GDP per capita as our variable of interest, this statistic measures the sum of differences between all pairs of regions in terms of deviations from average GDP per capita, weighted by their distances. For large samples, the expected value under the null hypothesis of no spatial autocorrelation across regions approaches zero. If this measure is large (compared to the expected value) we say that a region shows positive spatial autocorrelation, indicating that nearby regions tend to have similar values of GDP per capita. Instead, if the measure is small (compared to the expected value) this is evidence for negative spatial autocorrelation, indicating that nearby regions tend to have dissimilar values of GDP per capita. Finally, we may find that the measure is not significantly different from the expected value under the assumption of no spatial autocorrelation. In figure 9 we plot a global version of Moran's I across our regions over time.

Figure 9: Global Moran's I, 1900-2010 (z-scores)
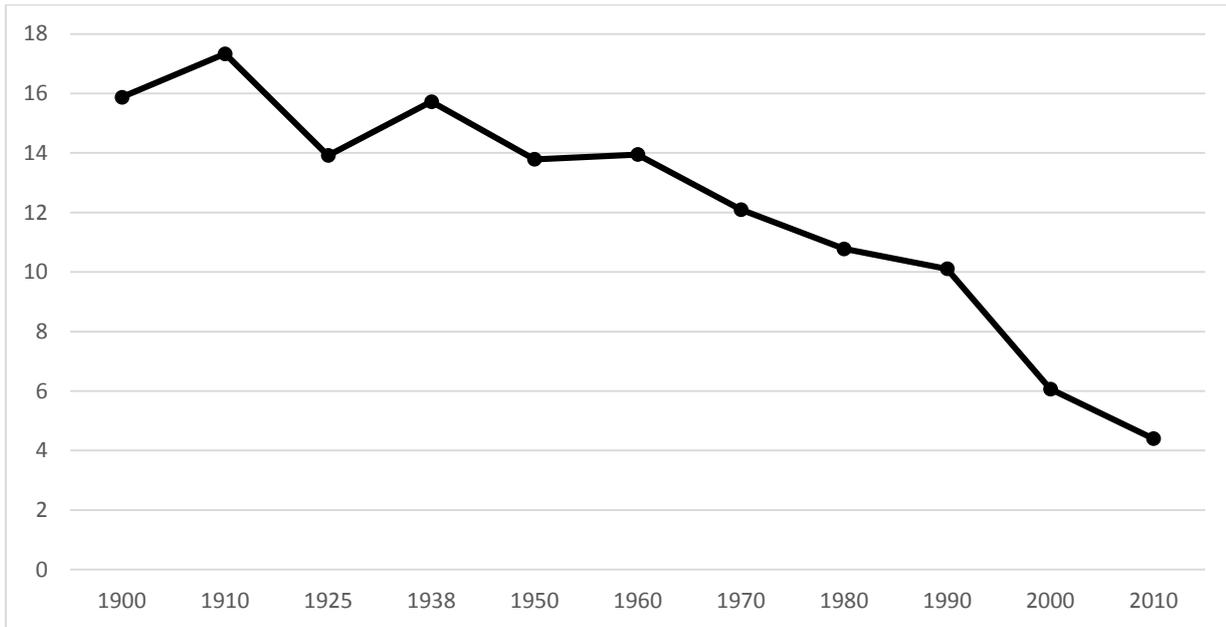
Source: own.

The figure shows standardized values of the measure (z-scores), all of which were strongly significant. We see that there is a systematic decline in average positive spatial autocorrelation, especially after 1960 and again after 1990. This shows something that was not visible in the simple measures of sigma-convergence: apparently the large clusters of neighbouring regions with similar levels of development have started to disappear. Another perspective on this can be gained from maps 9 and 10, where we show local Moran's I scores, distinguishing between positive autocorrelation and negative autocorrelation and excluding the large number of insignificant values.

Map 9: Local Moran's I (z-scores), 1900

„Map9_MoransIz_1900.png"

Source: own.

Map 10: Local Moran's I (z-scores), 2010

„Map10_MoransIz_2010.png"

Source: own.

We see that the old clusters are indeed shrinking and most regions do exhibit neither positive nor negative spatial autocorrelation with their neighbours. Note that "neighbours" here includes all other regions in the sample, weighted by their distance as well as national data from Poland, Czech Republic and Hungary.

Finally, let us return to the question of the dispersion of economic activity across regions from the perspective of geographical distribution. In figures 5a and 5b we could see that the distribution of GDP per capita across regions had become more equal between 1900 and 1980, but that this process was reversed between 1980 and 2010. Underneath this change is a geographical re-concentration of economic activity, both in terms of population but more so in terms of GDP. The slight but steady decline in the correlation between population density and GDP density (figure 1) suggested that the latter is increasingly driven by regional variation in productivity. Following Krugman (1991) we constructed a locational Gini-coefficient, based on the share of each region in total population, respectively total GDP. As usual, this coefficient is bounded between zero (all regions have equal shares) and one (all activity is concentrated in one region). Figure 10 shows the Gini for population and GDP over time.

Figure 10: Locational Gini-Coefficients for Population and GDP, 1900-2010
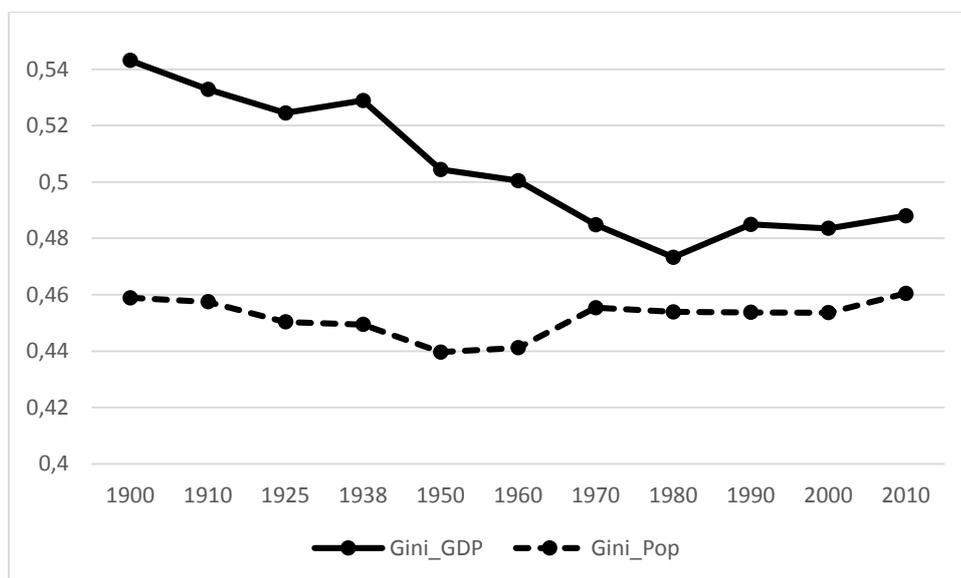
Source: own.

We find that the concentration of activity measured in terms of population and GDP followed similar trends. The concentration of population declines slowly after 1900 until 1960 followed by an increase in the 1970s. The pattern is similar but more pronounced for GDP, where we find that concentration declined in the long-run, but that this trend was stopped and even reversed around 1980. Alternatively, a simple Herfindahl-Index of Concentration would show a very similar picture, with a decline in the concentration of economic activity until 1980, followed by an increase until 2010. If we would weight each region by area (which is constant over time), we would again find similar pictures.

Other recent research suggests that this phenomenon continues and is not limited to Europe but applies to other OECD countries as well (see OECD regions at a glance, 2016). In fact, this trend of increasing

spatial concentration from around 1980 onwards can also be found for the US. The evidence on sigma-convergence is quite similar for our set of European regions and US states. Starting from a high level, there is no clear change during the interwar years, but a substantial decline in dispersion until about 1980. After this, we find for both the US states and European regions a trend reversal with rising regional inequality, due to strong growth in densely populated metropolitan areas (see fig. 3.7 on capital regions).

It is remarkable that the emerging picture on regional inequality in the long-run is also similar to the pattern of inequality in terms of personal income and wealth distributions, which has been extensively documented and analyzed (Piketty and Saez 2003, Piketty 2014). Regional inequality declined since 1900 but started to increase again around 1980, very much at the same moment when personal income inequality started its dramatic rise. These trends seem to be related to each other, suggesting that we need to rethink the driving forces behind both, as well as their consequences for economic development and political stability. Technological change and a new, deeper type of global market integration, favouring high-skilled labour and specific types of services at the expense of traditional, low-skill and often resource intensive industries have probably contributed to these changes. The combination of rising personal and regional inequality is a major challenge to economic policy for the decades to come.

9. Conclusion

How did European regions do over the last 110 years? We have constructed a new dataset at the level of 172 (173) regions to trace their economic development in terms of employment structure and GDP at purchasing power parity in the longer run. The broad trends of growth and stagnation as well as of convergence after 1945 that earlier studies have found are confirmed by our data. We saw that some regions stayed specialized in agriculture, while employment in industry first spread before it became more concentrated again. The level and growth of GDP per capita can be rather well explained in terms of conditional convergence, taking geographical and institutional factors into account.

But the long-run data allows us to see something more. In the long-run there is remarkable variation within states and some very deep changes that took place from about 1980 onwards. The share of overall variation in GDP per capita that is due to within-country differences has grown from around 30% in 1900 to above 50% in 2010, notably from 1980 onwards. Also, we found a growing disconnection between regions, with the emergence of islands of prosperity out of sync with their hinterland. Most important maybe is the observation that the pattern of regional inequality over the last 110 years follows a U-shape, just like the pattern of personal income inequality as documented by Piketty and Saez (2003) and others: after 1900 we find a spread of economic activity across regions and convergence between until about 1980, and divergence as well as geographical re-concentration thereafter.

There is a lot of diversity, but also remarkable similarities between European countries. The major changes in the economy of European regions occurred around the same time, notably the changes in growth rates, convergence and geographical concentration. To what extent are new technologies, and recently services, especially financial and business services driving these changes? What role is there for international trade and capital flows within and beyond Europe? How did European policies affect the dynamics, and what does this imply for issues like regional cohesion (see Becker et al 2012)? We hope that our new data will help to analyse these and other questions in a long-run perspective.

References

Acemoglu, D. (2009) *Introduction to Modern Economic Growth,* Princeton (NJ): Princeton University Press: chapter 1, pp. 3-25.

Baffigi, A. (2013) 'National Accounts, 1861-2011', in G. Toniolo (ed), *The Oxford Handbook of the Italian Economy since Unification*, Part II Sources of Growth and Welfare, Oxford: Oxford University Press, pp. 157-86.

Barro, R. J., & Sala-i-Martin, X. (1992). Convergence. *Journal of political Economy*, *100*(2), 223-251

Barro, R. J. and Sala-i-Martin, X. (1995) *Economic growth*, New York: McGraw Hill.

Becker, S. O., Egger, P. H. and von Ehrlich, M. (2012) 'Too much of a good thing? On the growth effects of the EU's regional policy', *European Economic Review*, Vol. 56 (4): 648-68.

Berend, I.T. (2016) *An economic history of twentieth-century Europe: economic regimes from laissez-faire to globalization*, Cambridge: Cambridge University Press.

Bolt, J. and Zanden, J.L. (2014) 'The Maddison Project: collaborative research on historical national accounts', *The Economic History Review*, Vol. 67 (3): 627-51.

Broadberry, S., Federico, G. and Klein, A. (2010) 'Sectoral developments, 1870–1914', in S. Broadberry and K.H. O'Rourke (eds), *Cambridge Economic History of Modern Europe, Vol. 2*, pp. 59-83.

Broadberry, S. and Klein, A. (2012) 'Aggregate and per capita GDP in Europe, 1870–2000: continental, regional and national data with changing boundaries', *Scandinavian Economic History Review,* Vol. 60 (1): 79-107.

Broadberry, S. (1997) 'Anglo-German productivity differences 1870–1990: A sectoral analysis', *European Review of Economic History*, Vol. 1 (2): 247-67.

Burhop, C. and Wolff, G.B. (2005) 'A compromise estimate of German net national product, 1851–1913, and its implications for growth and business cycles', *The Journal of Economic History*, Vol. 65 (3): 613-57.

Crafts, N. and Toniolo, G. (eds.) (1996) *Economic growth in Europe since 1945*, Cambridge: Cambridge University Press.

Deza, X.V. and López, M.G. (2014) 'Regional concentration of knowledge-intensive business services in Europe', *Environment and Planning C: Government and Policy*, Vol. 32 (6): 1036-58.

Eichengreen, B. (2007) 'The real exchange rate and economic growth', *Social and Economic Studies*, Vol. 56 (4): 7-20.

Fujita, M., Paul Krugman and Anthony J. Venables (1999), *The Spatial Economy*, MIT Press.

Gallego, J. and Maroto, A. (2015) 'The Specialization in Knowledge-Intensive Business Services (KIBS) across Europe: Permanent Co-Localization to Debate', *Regional Studies*, Vol. 49 (4): 644-64.

Geary, F. and Stark, T. (2002) 'Examining Ireland's post-famine economic growth performance', *The Economic Journal*, Vol. 112 (482): 919-35.

Iammarino, S., A. Ridríguez-Pose and M. Storper (2017), 'Why Regional Development matters for Europe's Economic Future', European Commission Working Papers, WP 07/ 2017.

James, H. (2017), 'Deglobalization as a Global Challenge', CIGI Paper 135, Centre for International Governance Innovation, Waterloo, Ontario.

Krugman, P. (1991) 'Increasing returns and economic geography', *Journal of political economy*, Vol. 99 (3): 483-99.

Moretti, E. (2013) 'Real wage inequality', *American Economic Journal: Applied Economics*, Vol. 5 (1): 65-103.

Persson, T. (2001) 'Currency unions and trade: how large is the treatment effect?', *Economic Policy*, Vol. 16 (33): 434-48.

Pollard, S. (1981) *Peaceful conquest: the industrialization of Europe, 1760-1970*, Oxford: Oxford University Press.

Redding, Stephen and Anthony Venables (2004), "Economic geography and international inequality", *Journal of International Economics*, 62: 53 – 82.

Rosés, J.R., Wolf, N. (2010) 'Aggregate Growth, 1913–1950', in S. Broadberry, K.H. O'Rourke (eds.), *The Cambridge economic history of modern Europe, Volume 2,* Cambridge: Cambridge University Press, pp. 181-207.

Schön, L. and Krantz, O. (2012) 'The Swedish economy in the early modern period: constructing historical national accounts', *European Review of Economic History*, Vol. 16 (4): 529-49.

Solow, R.M. (1956) 'A contribution to the theory of economic growth', *The quarterly journal of economics*, Vol. 70 (1): 65-94.

Temin, P. (2002) 'The golden age of European growth reconsidered', *European Review of Economic History*, Vol. 6 (1): 3-22.

Vonyó, T. (2008) 'Post-war reconstruction and the Golden Age of economic growth', *European Review of Economic History*, Vol. 12 (2): 221-41.

Williamson, J.G. (1965) 'Regional inequality and the process of national development: a description of the patterns', *Economic development and cultural change*, Vol. 13 (4, Part 2): 1-84.

Wolf, N. and Ritschl, A.O. (2011) 'Endogeneity of currency areas and trade blocs: evidence from a natural experiment', *Kyklos*, Vol. 64 (2): 291-312.

Map 1: population density 1900
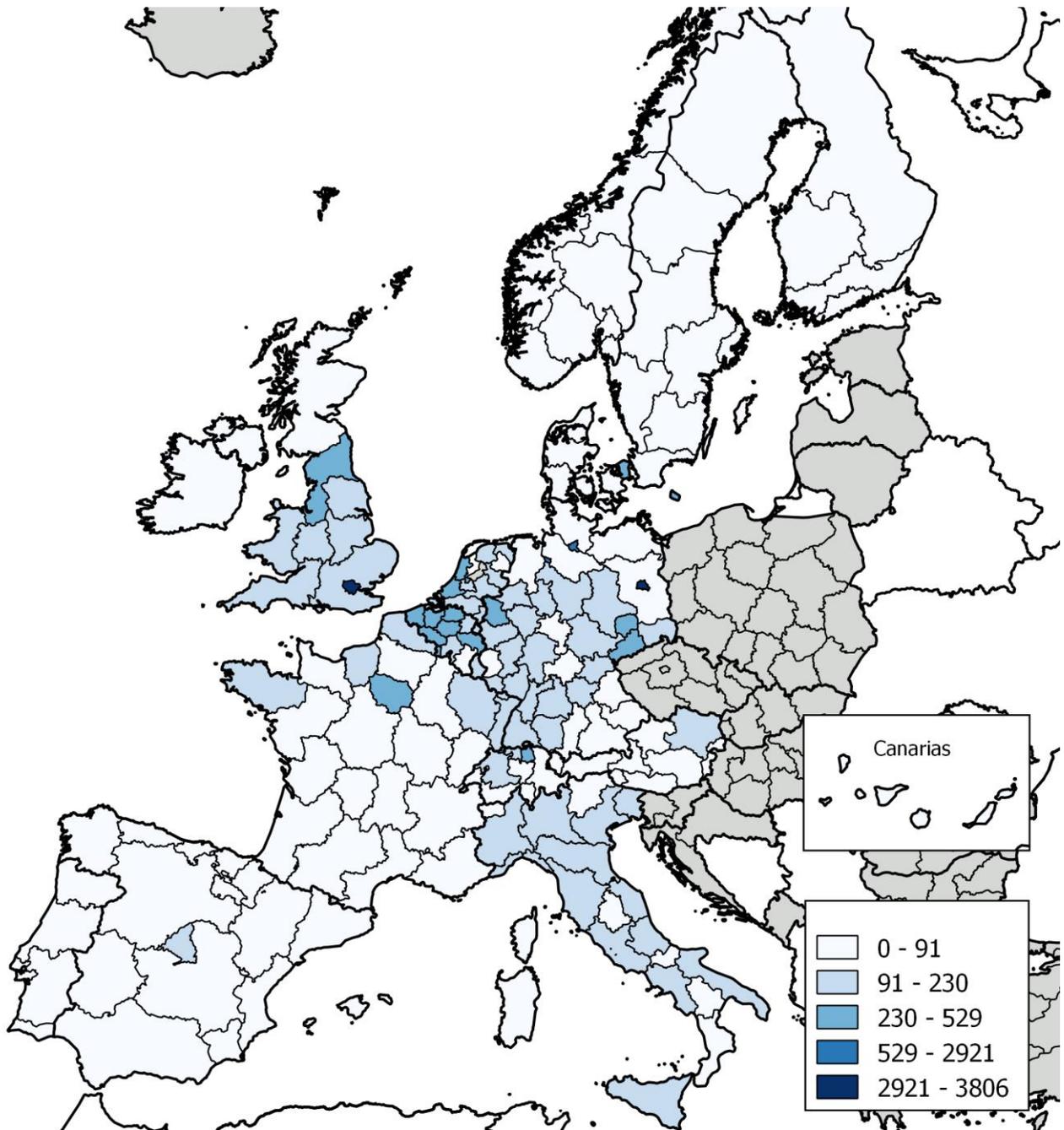


Source: own.

Map 2: population density 2010
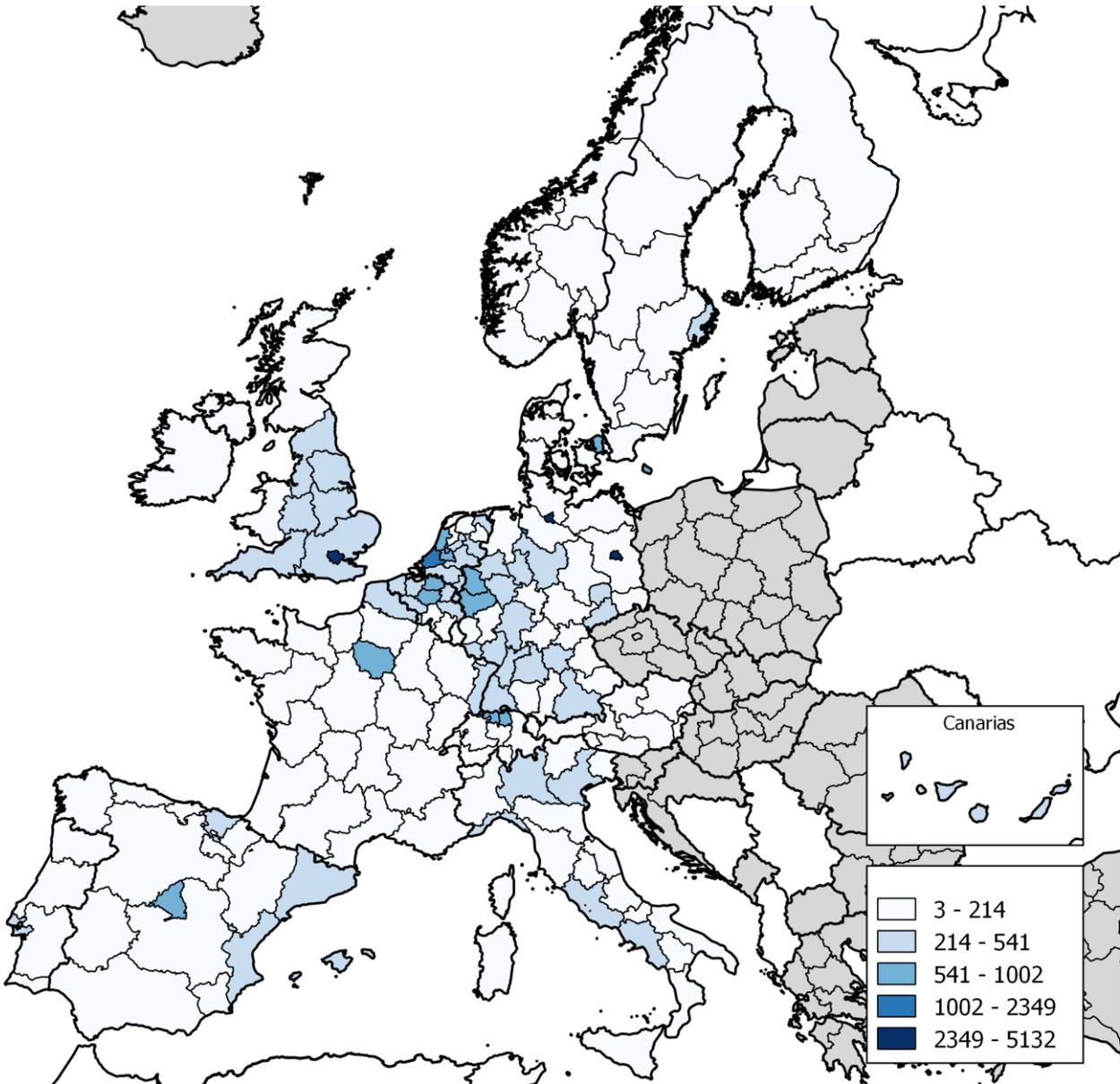
Source: own.

Map 3: employment density 1900

Canarias

	0 - 35
	35 - 88
	88 - 239
	239 - 1242
	1242 - 1735
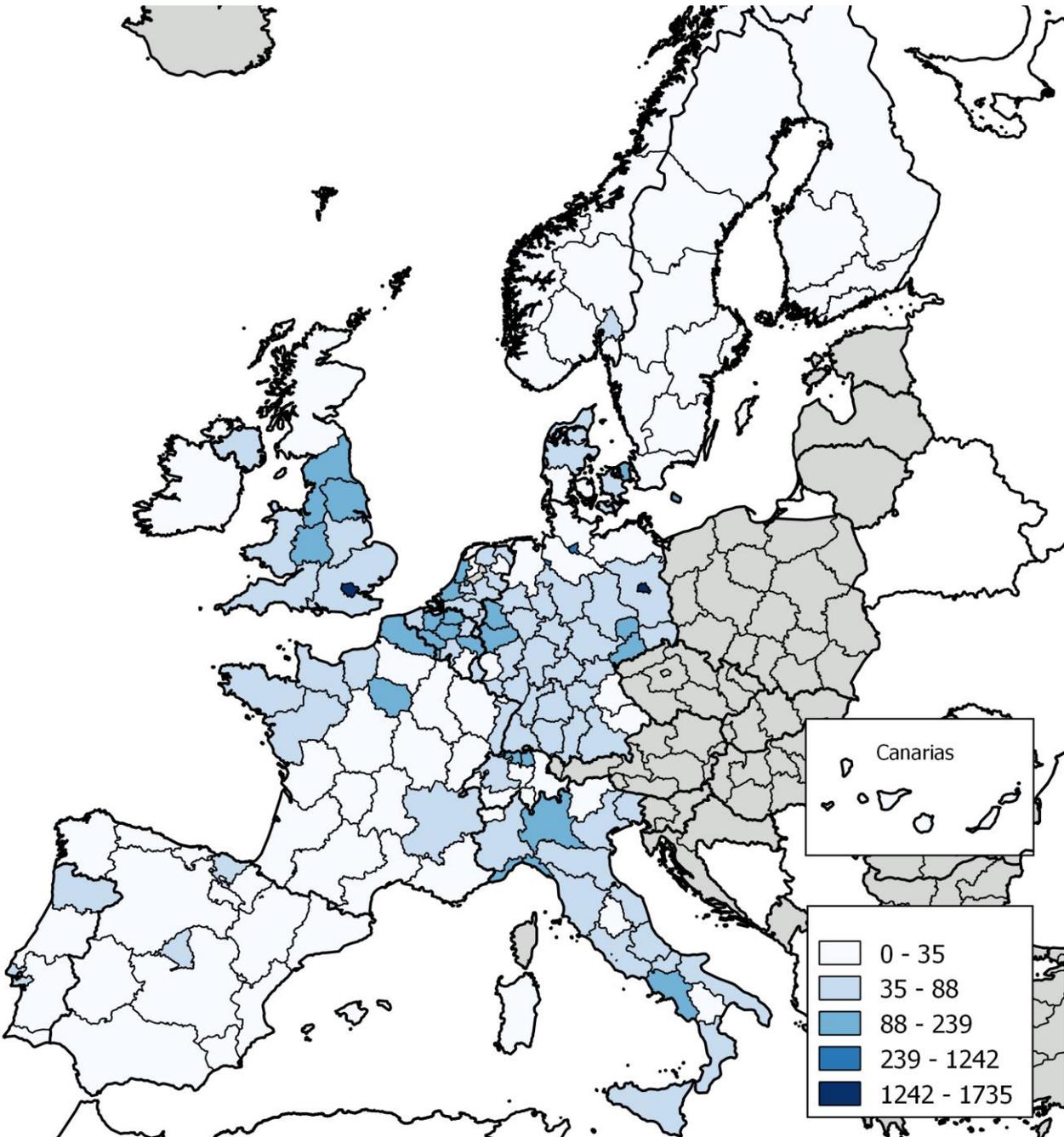
Source: own.

Map 4: employment density 2010

Canarias

| 0 - 60 |
| 60 - 140 |
| 140 - 300 |
| 300 - 960 |
| 960 - 1890 |
| 1890 - 2500 |
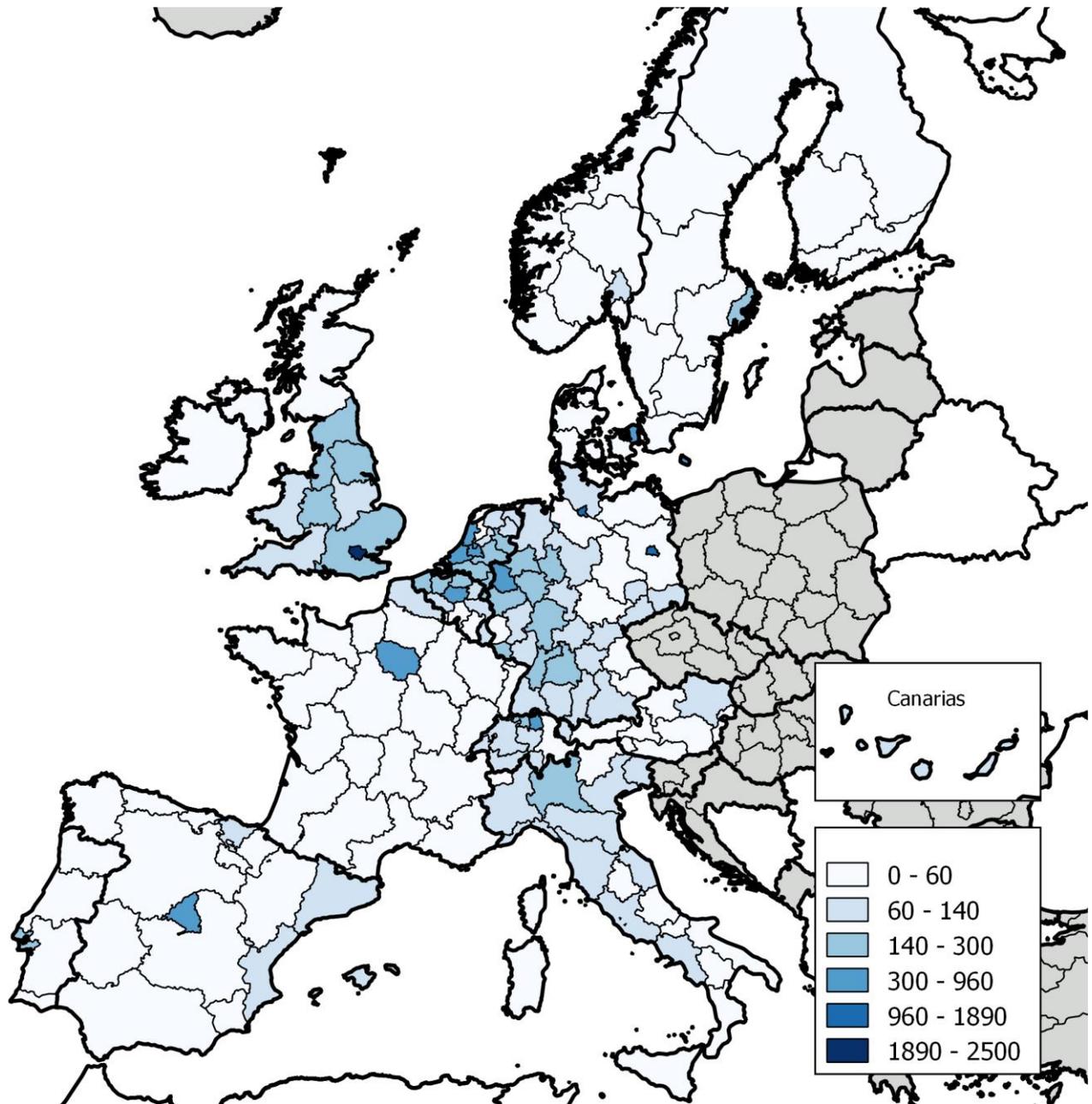
Source: own.

Map 5: GDP density 1900

Canarias

0.00 - 0.19
0.19 - 0.49
0.49 - 1.07
1.07 - 2.34
2.34 - 12.44
12.44 - 24.70
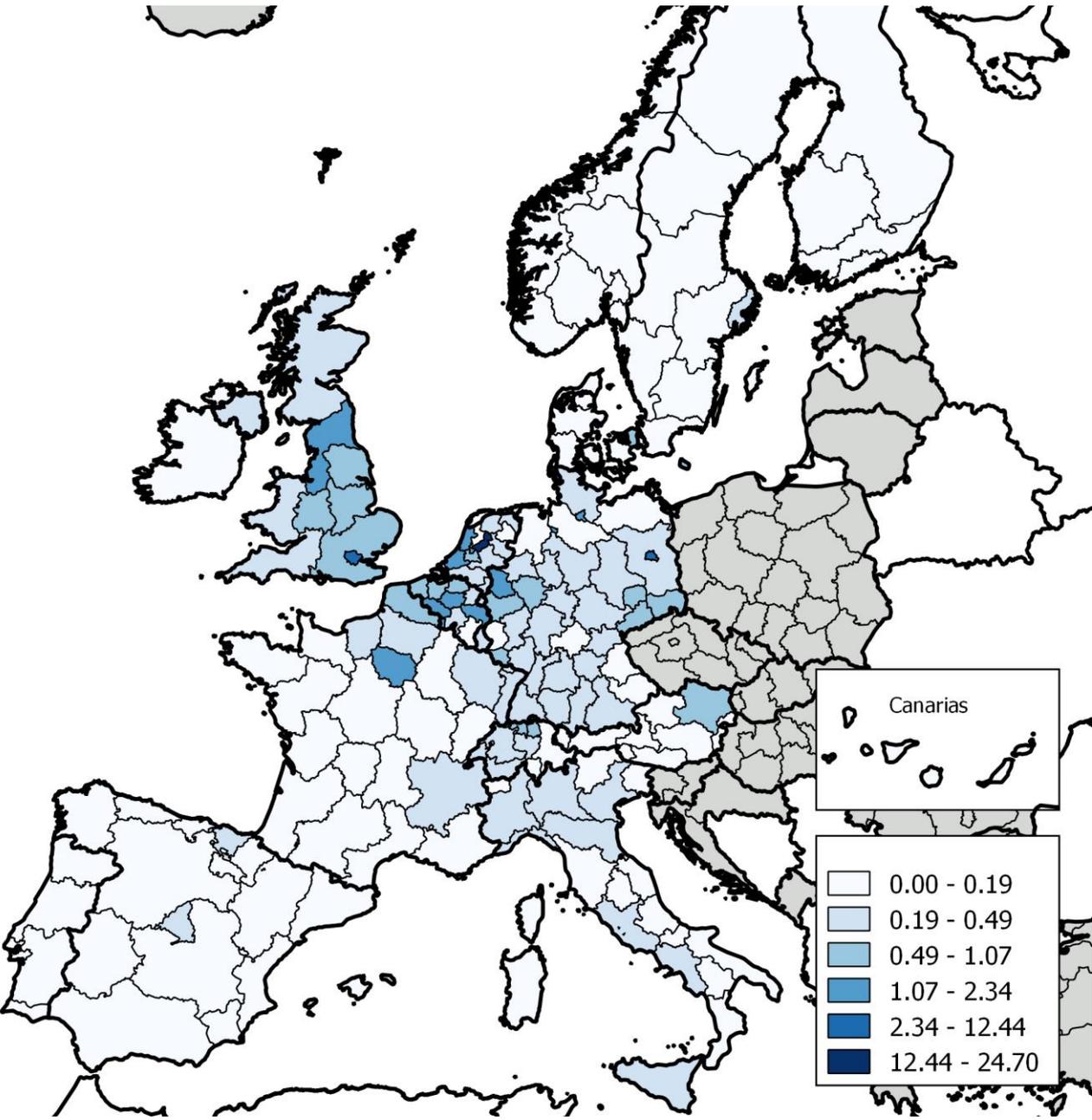
Source: own.

Map 6: GDP density 2010

Canarias

0.1 - 2.8
2.8 - 6.7
6.7 - 16.6
16.6 - 42.6
42.6 - 82.5
82.5 - 189.1
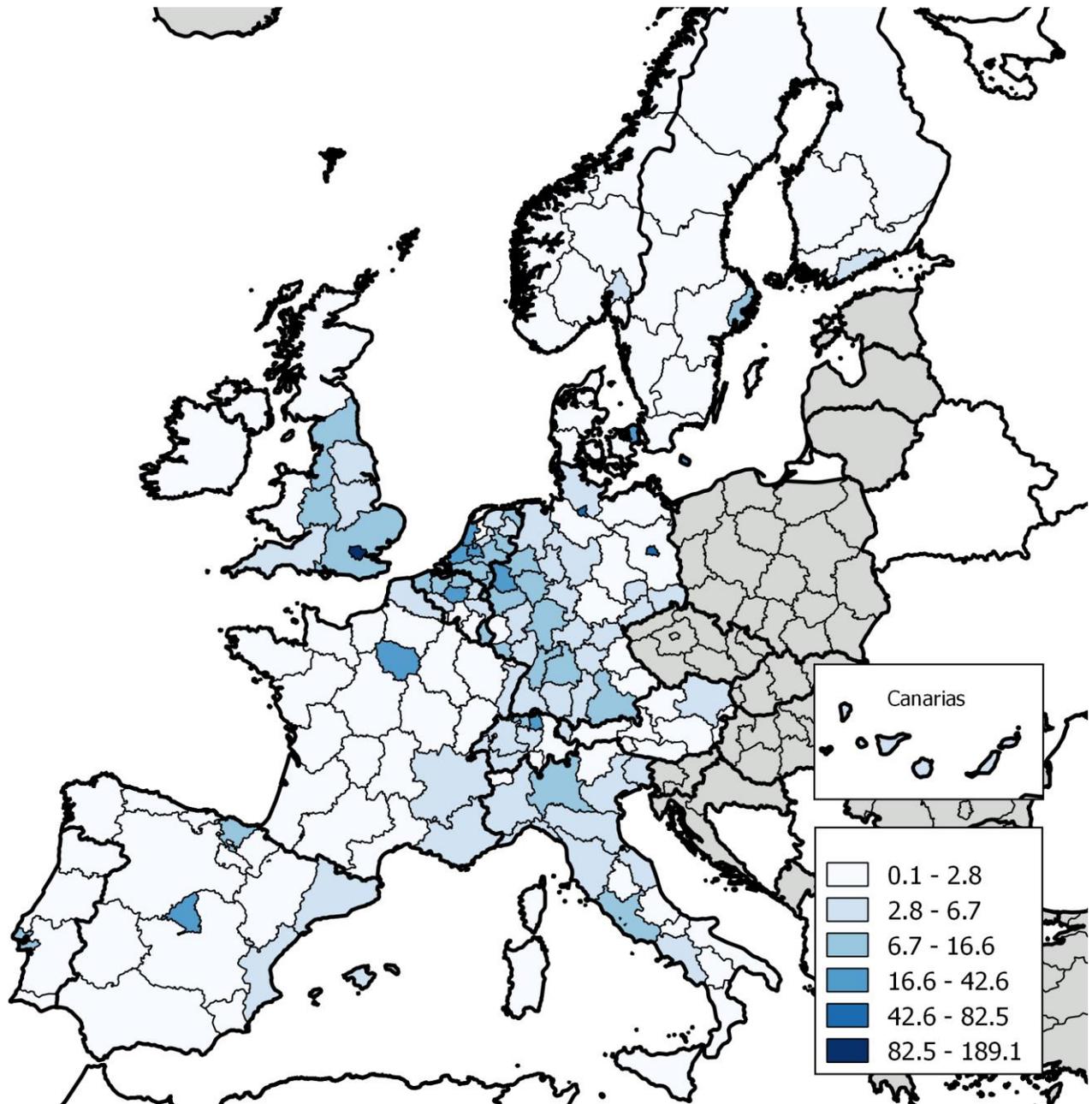
Source: own.

Map 7: GDP per capita, 1900
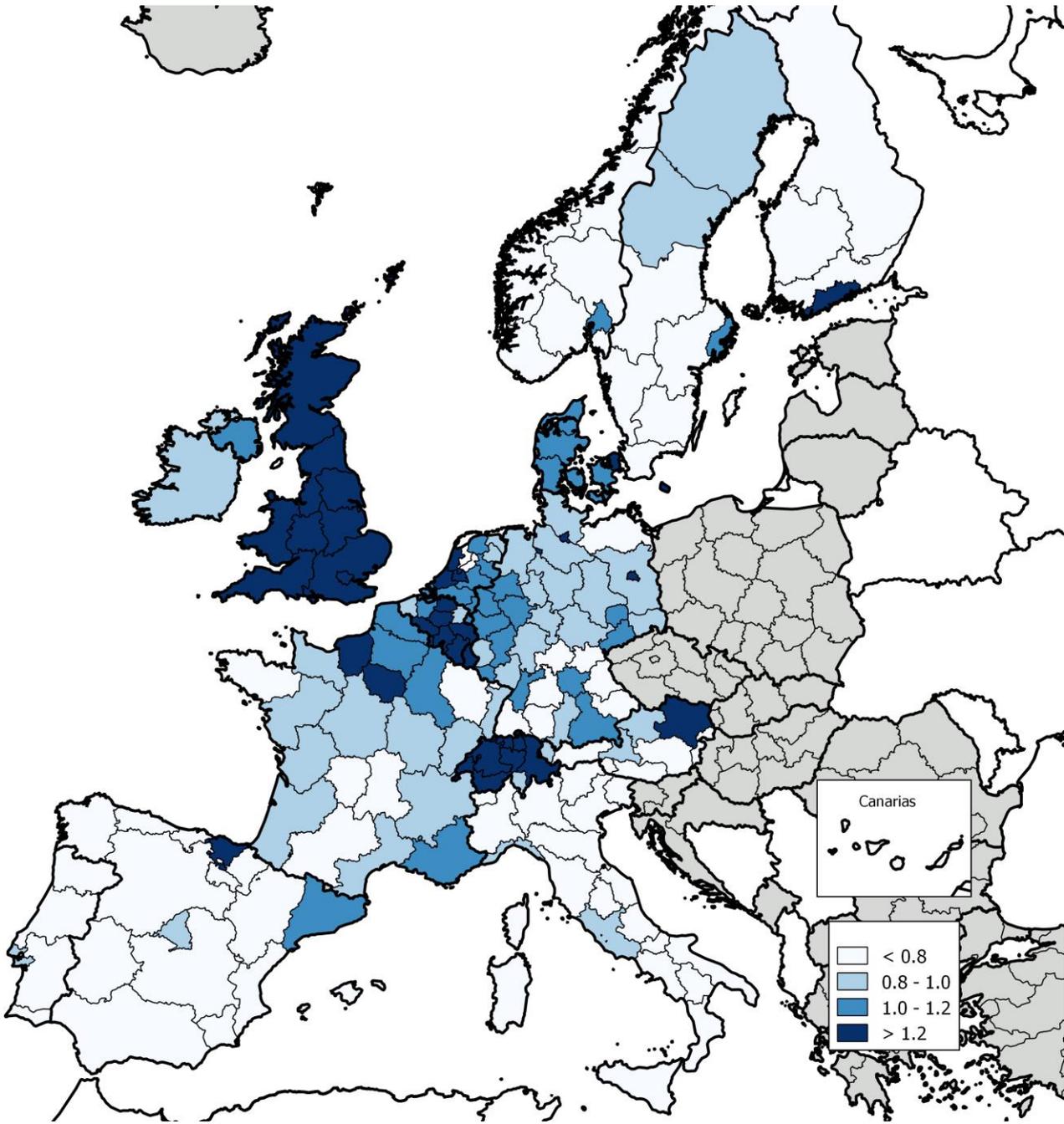
Source: own.

Map 8: GDP per capita 2010
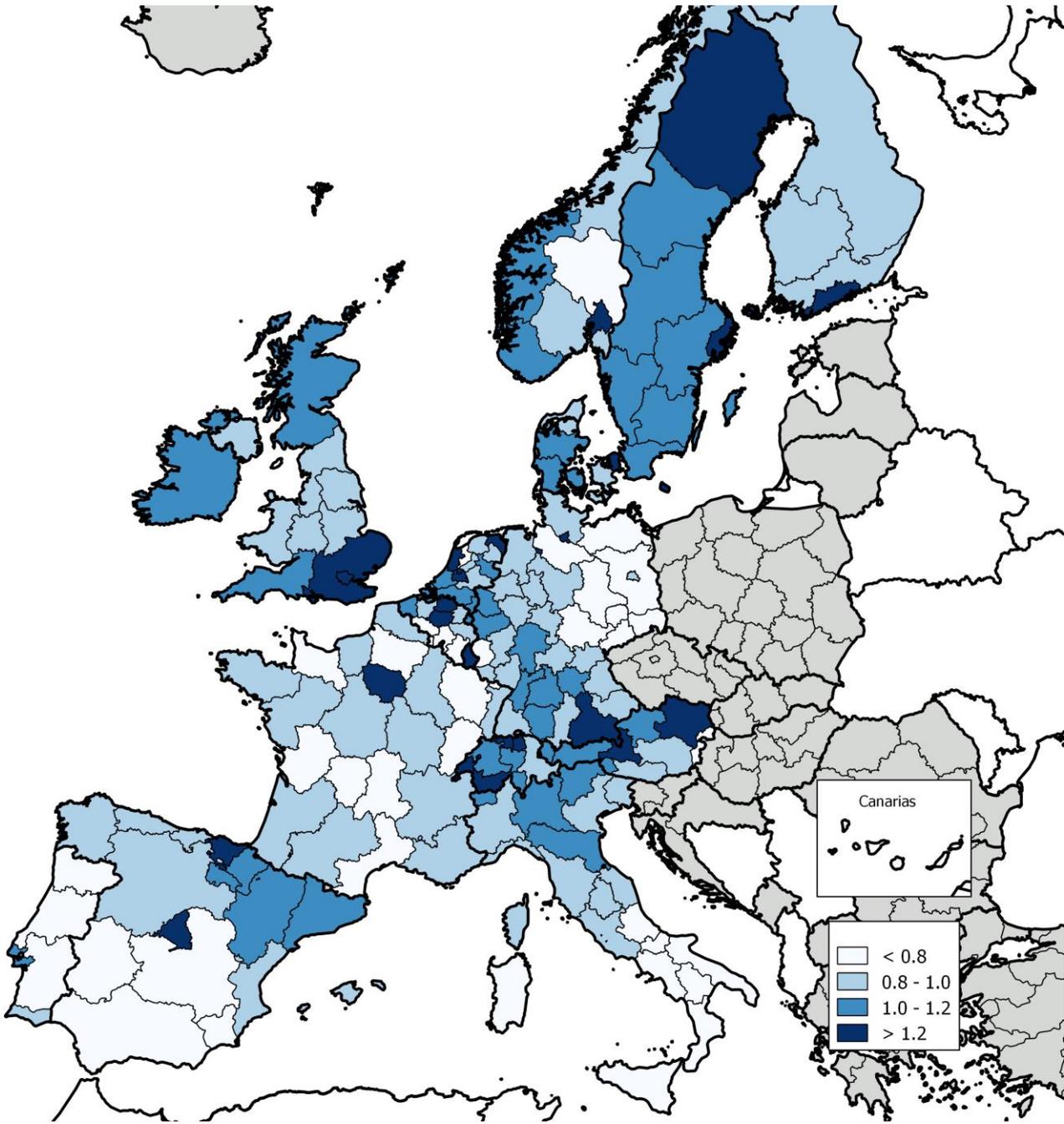
Source: own.

Map 9: Moran's I, 1900
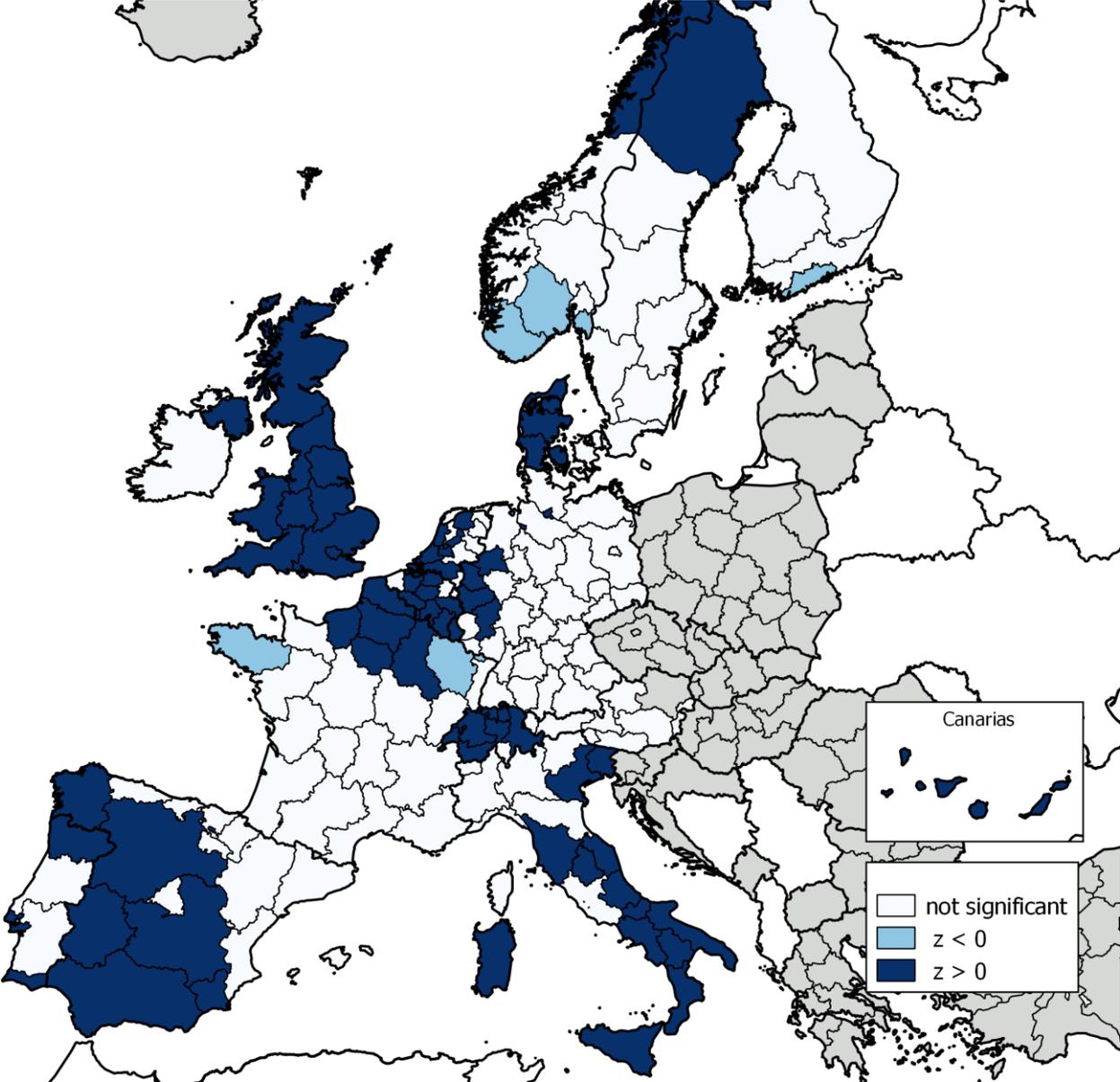
Source: own.

Map 10: Moran's I, 2010

Source: own.