Oxford Review of Economic Policy, Volume 37, Number 1, 2021, pp. 17–48

Regional growth and inequality in the long-run: Europe, 1900–2015

Joan R. Rosés* and Nikolaus Wolf**

Abstract In this paper we discuss regional income growth and inequality based on a new set of long-run data. The data cover 173 European regions in 16 countries, from 1900 to 2015. 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, structural change with a declining role of agriculture, the rise and fall of industry, and the long rise of services. We show how 'fundamentals' of institutions and geography affected income levels over the twentieth century, and describe how regional growth after 1945 turned from convergence and adjustment to shocks to divergence. In the long run we observe a U-shaped pattern of regional convergence followed by divergence, not unlike recent observations on personal income and wealth distributions.

Keywords: regional inequality, Europe, long run

JEL classification: D31, N1, N9, R1

I. Introduction

Over the last few generations, the European economy experienced 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 entire continent in weight and influence. During

*London School of Economics and Political Science, UK, and CEPR; e-mail: j.r.roses@lse.ac.uk

**Humboldt University Berlin, Germany, CEPR, and CESifo; e-mail: nikolaus.wolf@wiwi.hu-berlin.de

This paper is partly based on an earlier working paper, 'Regional Economic Development in Europe, 1900–2010: A Description of the Patterns' (CEPR DP 12749), which summarized a larger collaborative project. We are grateful to all contributing authors: Marc Badia-Miró, Erik Buyst, Kerstin Enflo, Emanuele Felice, Frank Geary, Jordi Guilera, Kari Anne Janisse, Peter Sandholt Jensen, Herman de Jong, Martin Henning, Alexander Klein, Julio Martínez-Galarraga, Jørgen Modalsli, Cristina Victoria Radu, M. Teresa Sanchis, Lennart Schön, Max S. Schulze, Paul Richard Sharp, Tom Stark, Dirk Stelder, Daniel A. Tirado, Ulrich Woitek, and Gabriela Wüthrich. We are also very grateful for outstanding research assistance from Iris Wohnsiedler, Kaja Rupieper, and Oliver Salzmann and helpful comments from Kalle Kappner, Max Schulze, and two anonymous referees.

doi:10.1093/oxrep/graa062

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, the European debt crisis that followed in its wake, and the Covid-19 pandemic. 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)).

Still, 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 nineteenth century. Yet this approach comes at some costs. It neglects the often considerable variation within states (sometimes larger than between states, as we will show); and it tends to attribute differences in development to differences in national institutions or policies without being able to test this. Furthermore, differences of income *per capita* (and labour productivity) within countries are larger, and sometimes more resilient, than differences across countries. In particular, income convergence across European nations was not always accompanied by 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 then from 1980 diverging back to a less equal distribution.

In this paper we reconsider the economic development of Europe since 1900 from the perspective of European regions as pioneered by Pollard (1981) and 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 purchasing power parity (PPP), 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. The developments after 1945 can be described as a story of adjustment to shocks, and convergence turning into divergence since the late 1970s. Former coal-mining regions play an important part in this. Finally, we show how the geography of activity since 1900 has changed with 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.

II. Data and methodology

Our data set contains 173 regions covering 16 European nation states at the level of NUTS-2 (as of 2010) and spans 12 benchmark years between 1900 and 2015.[1] The data-set is freely available in various formats at http://nikolauswolf.eu. 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 was 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. For more details on the data and individual countries we refer the reader to Rosés and Wolf (2019), and especially to the chapters therein on regional development within individual countries.

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; since then regional data has been provided by Eurostat, the statistical office of the European Union. For the majority of countries before 1960, we have employed the methodology suggested by 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: (i) it requires readily available data (employment by sector and region, wages by sector and region, and historical national accounts) and (ii) it 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

[1] 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 Rosés and Wolf (eds) (2019).

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 βk is a scalar that preserves the relative regional 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 were 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 benchmarks (2000, 2010, and 2015) numbers are taken directly from Eurostat databases.

To make our data homogeneous and comparable across different countries, we constructed for each country and year a regional breakdown of national GDP aggregates. We used national-level GDP estimates from the Maddison Project (Bolt and van Zanden, 2014; Bolt *et al.*, 2018), which provides data on GDP of European nation states expressed at PPP in 1990 international dollars, and alternative estimates in 2011 international dollars. This database incorporated recent updates to national GDP estimates, such as Germany (Burhop and Wolff, 2005), Sweden (Schön and Krantz, 2012) and 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 German Democratic Republic (GDR) and the Federal Republic of Germany (FRG) for the years 1950–80. 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. We note 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 the base of our regional distribution of GDP within countries. As we will show, our main results using 1990 GK$ are robust to the latest Maddison Project Database (2018) with multiple benchmarks.[2]

Our methodology also implies that we have no regional price deflators but, instead, we use national deflators, just as Eurostat and the OECD do. Hence, we need to 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 (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 homogeneous within countries, while differences in housing prices may have increased over time. However, 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 are an imperfect measure of 'welfare' differential across regions in the same country, but more likely reflect the 'true' differences in labour productivity across regions. Finally, we note that our estimates of regional shares in national GDP are based on census years that vary across nation states. To make the data comparable across countries we extra- or interpolate 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 periods of war. 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.

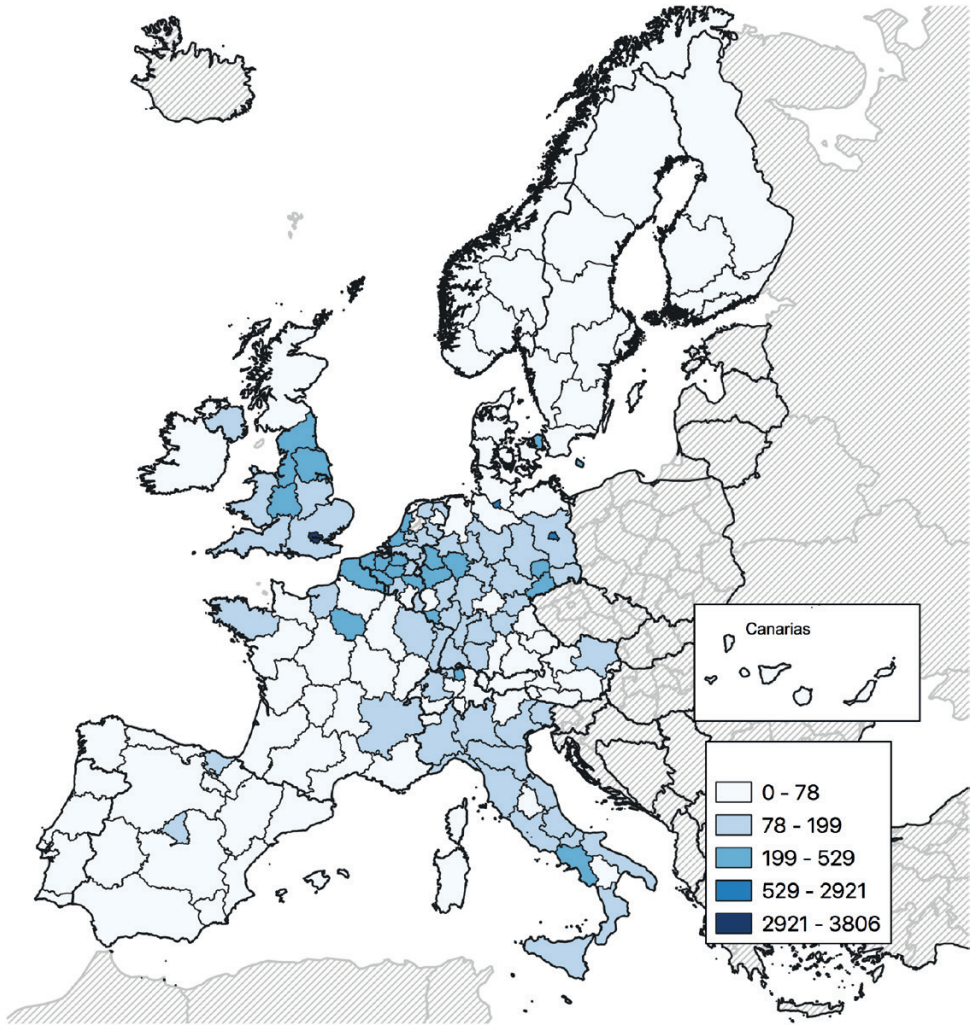
In the next parts of our paper, we first describe these data in terms of densities of population, employment, and GDP per area (section III), changes in the structure of employment (section IV), and variation in GDP *per capita* (section V). A major question is whether European regions have converged, notably whether poor regions managed to catch up or not, and which factors can explain this (section VI). We provide a summary perspective on spatial convergence and divergence 1900–2015 in section VII, before we conclude.

III. Basic facts on regional economic development: density of population, employment, and GDP (1990 international dollars)

We start with a look at population density as the simplest possible (and maybe most reliable) indicator of economic activity (see Maps 1 and 2). As expected, the density of

[2] For our baseline results, we use the Maddison Data with 1990 international dollars to ensure comparability with the bulk of research on long-run development. However, we also 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' in 2011 international dollars, based on multiple benchmark comparisons to improve historical income comparisons across countries. We discuss these results later in the paper.

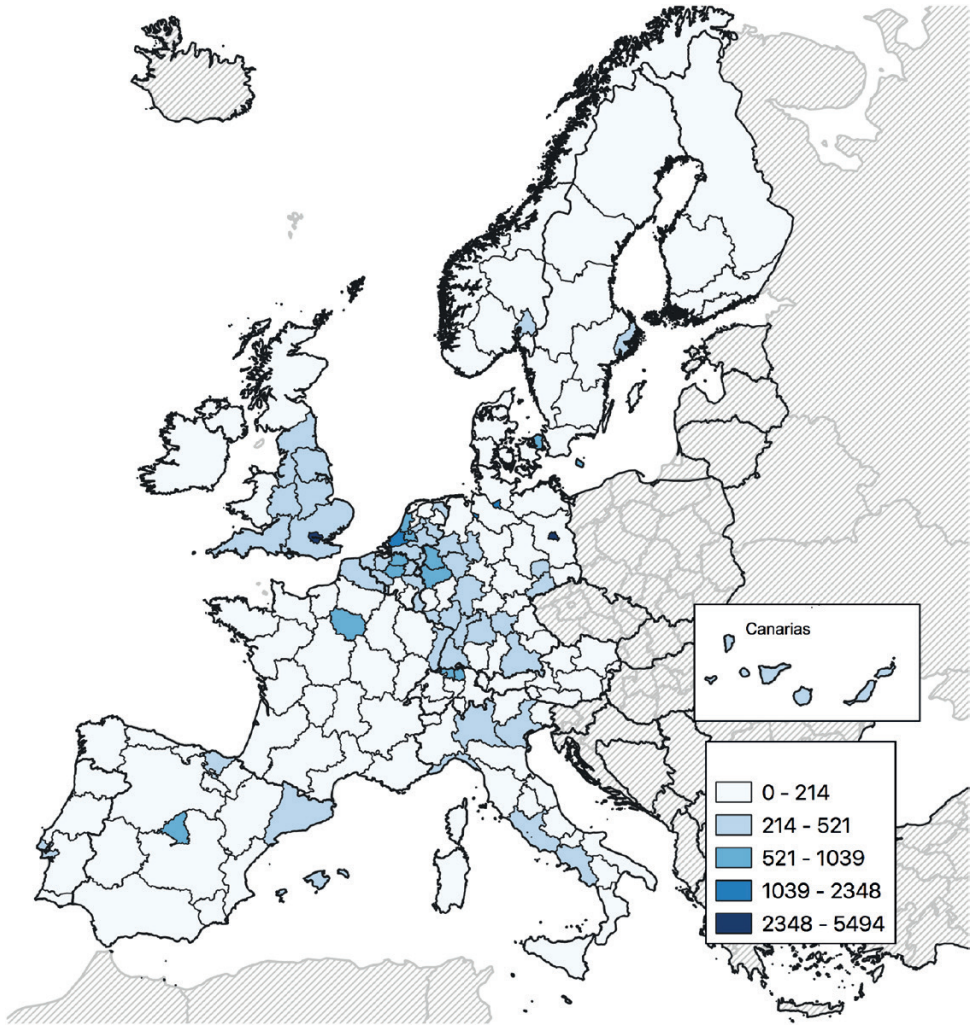
Map 1: Population density, 1900



Source: Authors' own drawing.

population measured as persons per km^2 shows considerable variation across regions and over time. The average density increased from 150 (1900) to 288 (2015), the median from 76 (1900) to 151 (2015), showing 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 economic decline over time, including Hainaut in Belgium (BE32) as well as Chemnitz (DED4) and Leipzig (DED12) in Germany, reflected in nearly stagnating population density in 2015 compared to 1900. We come

Map 2: Population density, 2015



Source: Authors' own drawing.

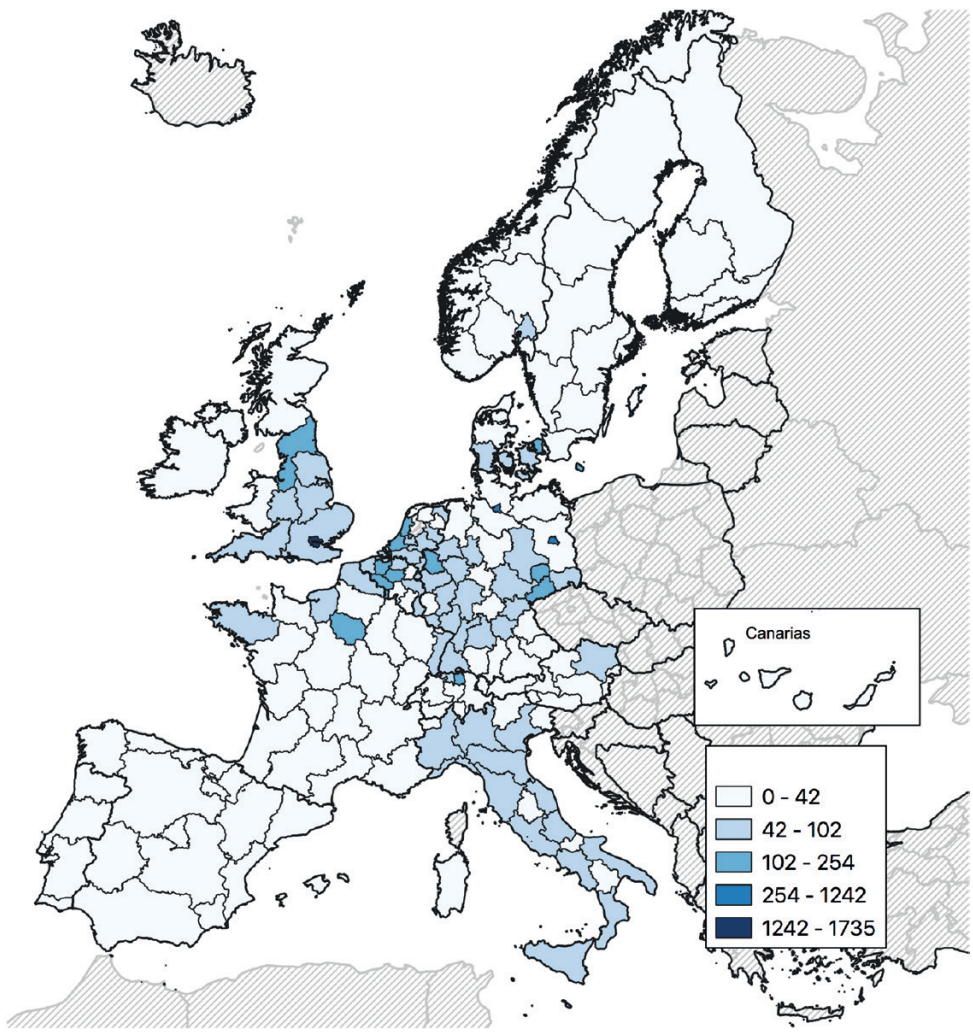
back to their destiny further below. There was more stability at the bottom of the distribution, with the least densely populated regions both then and now being located in the northern parts of Sweden, Norway, and Finland, followed by Alentejo (PT18) in Portugal and regions in central Spain, namely Castile-Leon (ES41), Castile-La Mancha (ES42), and Extremadura (ES43).

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, is, indeed, very strongly correlated with variation in population density. Moreover, the correlation between the density of population and these geographical variables in 2015 is only very slightly weaker than

it was in 1900. Apparently, the impact of natural geography on the location of population across Europe 1900–2015 is strong and persistent. We consider some of these factors in section VI below.

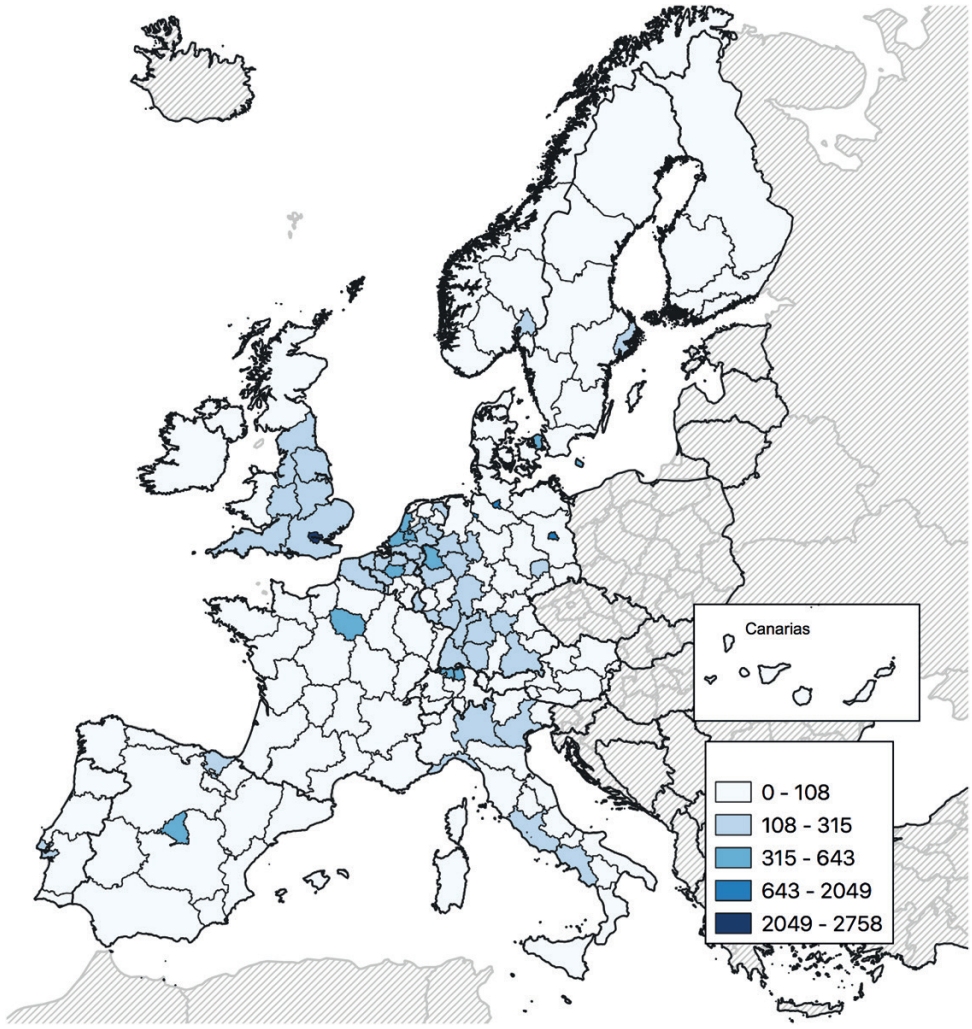
In a next step, we look at the density of economic activity as measured by total employment per km^2 (see Maps 3 and 4). The average employment density roughly doubled from 67 persons per km^2 (1900) to 146 (2015), the median increased from 38 to 68. Not surprisingly, employment density is closely correlated 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

Map 3: Employment density, 1900



Source: Authors' own drawing.

Map 4: Employment density, 2015

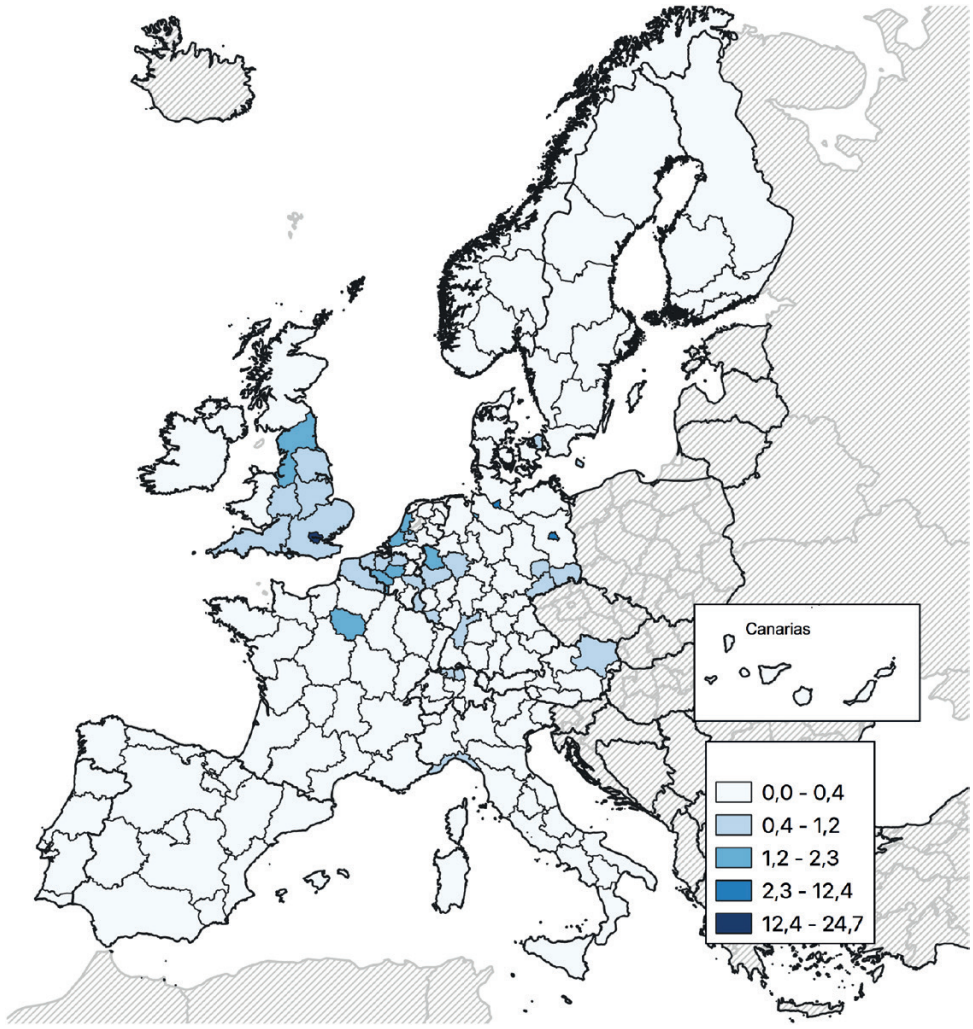


Source: Authors' own drawing.

people tend to have lower employment). Next, female participation rates could differ and regions with lower female participation rates tend to have a lower overall share of their working-age population employed. 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 they 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 has certainly changed over time. Maps 3 and 4 show how employment density developed between 1900 and 2015.

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

Map 5: GDP density, 1900 (1990 GK$)

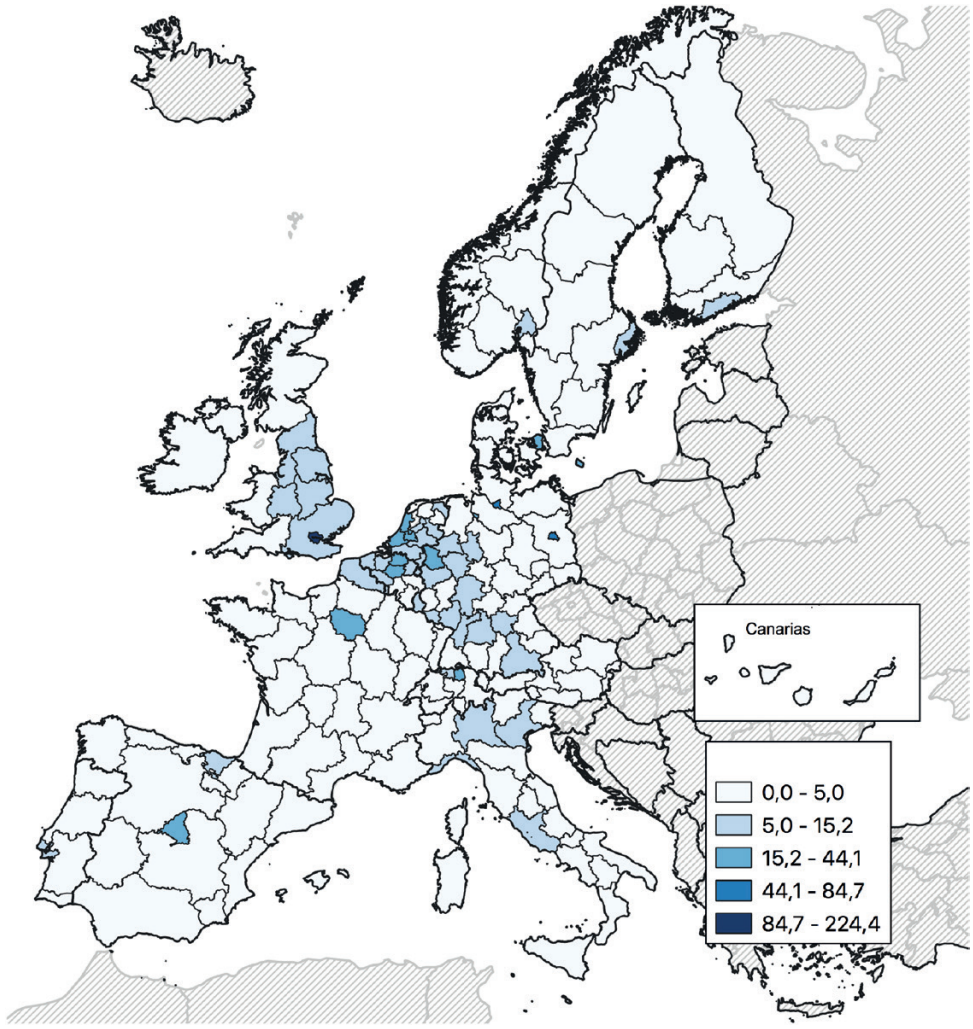


Source: Authors' own drawing.

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.

The cross-sectional correlation between population density and GDP density over time is always very strong (both using 1990 and 2011 PPPs). However, after both wars the correlation weakened, especially after 1945. In 1950, several densely populated regions had quite low GDP *per capita*, partly due to destruction such as bombing, and partly due to migration of the working population. Already by 1960, this effect has largely disappeared, which confirms the resilience of economic activity in certain regions. Another drop in this correlation can be found in 1990, due to the collapse of the GDR economy in the wake of German unification.

Map 6: GDP density, 2015 (1990 GK$)



Source: Authors' own drawing.

IV. Structural change at the regional level: Europe, 1900–2015

Let us first consider the changing employment structure of Europe. A characteristic of Europe's economic development during the twentieth 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 1 shows the evolution of average employment shares across European regions, 1900–2015, for three broad sectors: agriculture, industry (including 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. By 2015, the share of industrial employment was about one-fifth less than in 1900.
While this structural change occurred at different speeds across European regions, the decline of agricultural employment was a universal phenomenon and nearly everywhere most pronounced in the two decades after 1945. A way to capture the variation in sector-specific employment across regions is to use the location quotient, defined as

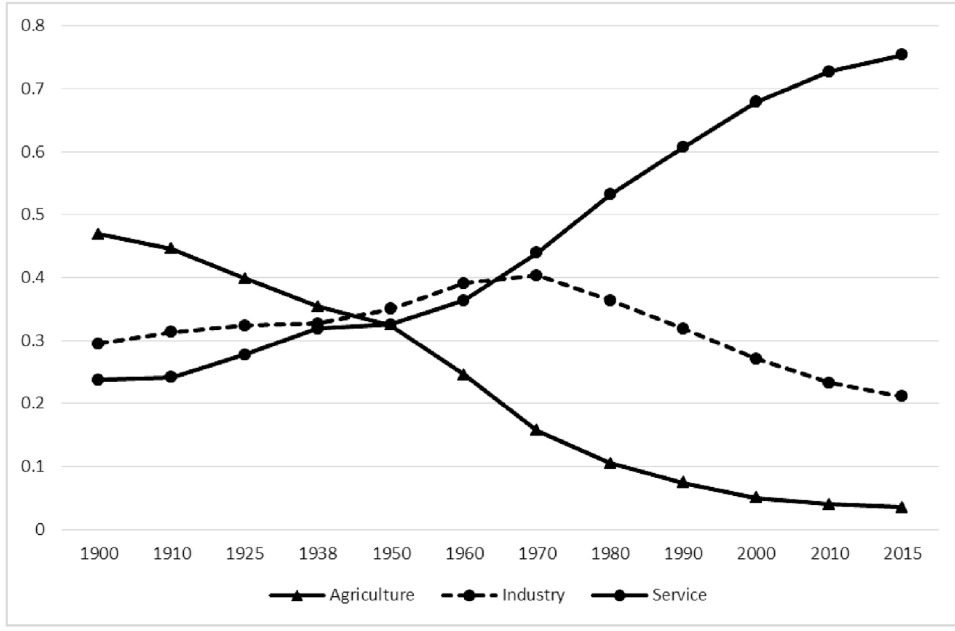
$$lq_i^k = \frac{x_i^k / \sum_k x_i^k}{\sum_i x_i^k / \sum_i \sum_k x_i^k} = \frac{x_i^k / \sum_i x_i^k}{\sum_k x_i^k / \sum_i \sum_k x_i^k},$$

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, 12 years, and three sectors, Figure 2 shows the coefficient of variation over the period 1900–2015.

Around 1900, the overall dispersion in agriculture, industry, and services was remarkably similar (which is no statistical artefact): 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 growing 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 now strongly specialized in agriculture, notably in southern Europe such as 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.

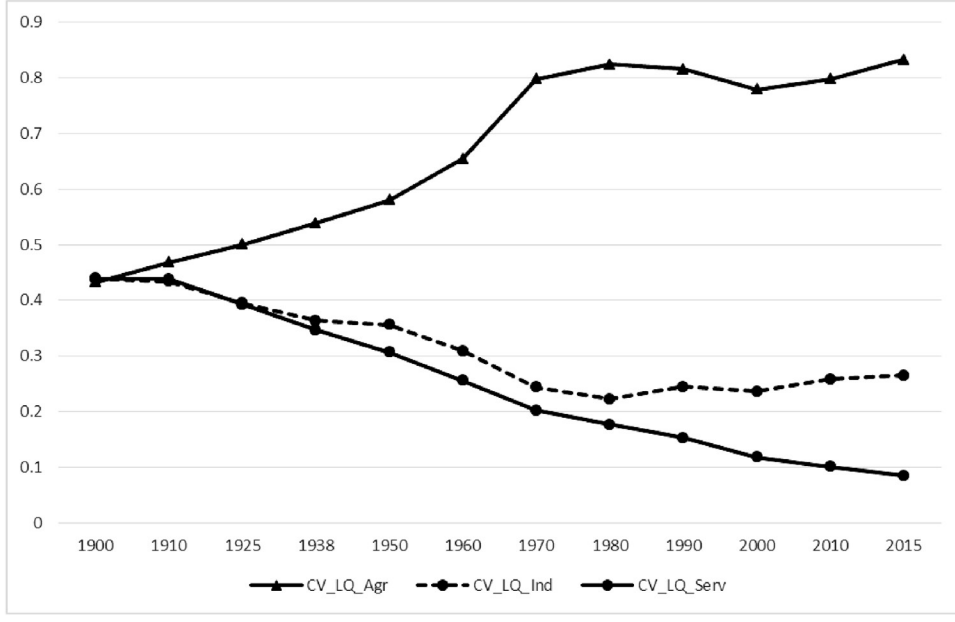
The localization of industry is much less diverse. Some regions around 1980 are still strongly dominated by industrial employment, often associated with the automobile industry, such as Franche Comté (FR43), Thüringen (DEG0), or Stuttgart (DE11). Industry is strongly clustered in Germany, eastern France, 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, which 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 largest urban agglomerations (like London, Paris, or Berlin) all experienced a rapid de-industrialization, which had already started before the First World War. Other work has shown in more detail how

Figure 1: Sectoral employment shares across 173 European regions, 1900–2015



Source: See text.

Figure 2: The coefficient of variation in localization of agriculture, industry, and services, 1900–2015



Source: See text.

Table 1: GDP *per capita* across regions, 1900–2015: cross-sectional variation and growth[3]

Levels in 1,000 GK$1990		**1900**	**1950**	**1980**	**2015**
	Average GDP per capita	2,668	4,154	13,032	20,826
	Poorest region	Galicia (ES11)	Extremadura (ES43)	Extremadura (ES43)	Calabria (ITF6)
	Richest region	London Counties (UKI)	Zurich (CH04)	Hamburg (DE60)	London Counties (UKI)
Growth rates in % per year		**1900–2015**	**1900–50**	**1950–80**	**1980–2015**
	Average growth rates	1.82	1.03	3.72	1.33
	Highest growth rates	2.68 (Nord-Norge, NO07)	2.55 Västsverige (SE23)	5.94 (Basilicata, ITF5)	3.27 Ireland (IE0)
	Lowest growth rates	1.05 (Hainaut, BE32)	–0.16 (Berlin, DE30)	1.68 (West Midlands, UKG)	0.38 (Umbria, ITI2)

Source: Authors' own calculations.

regional development was affected by the rise and fall of particular industries (for example, Marshall (1987) with long-run evidence on the UK). While our data cover many parts of Europe since 1900, we still lack a further disaggregation into particular industries to explore such ideas at the European level.

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 being 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 (Deza and López, 2014; Gallego and Maroto, 2015).

V. Growth and variation in GDP *per capita* (1990 and 2011 PPP)

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.

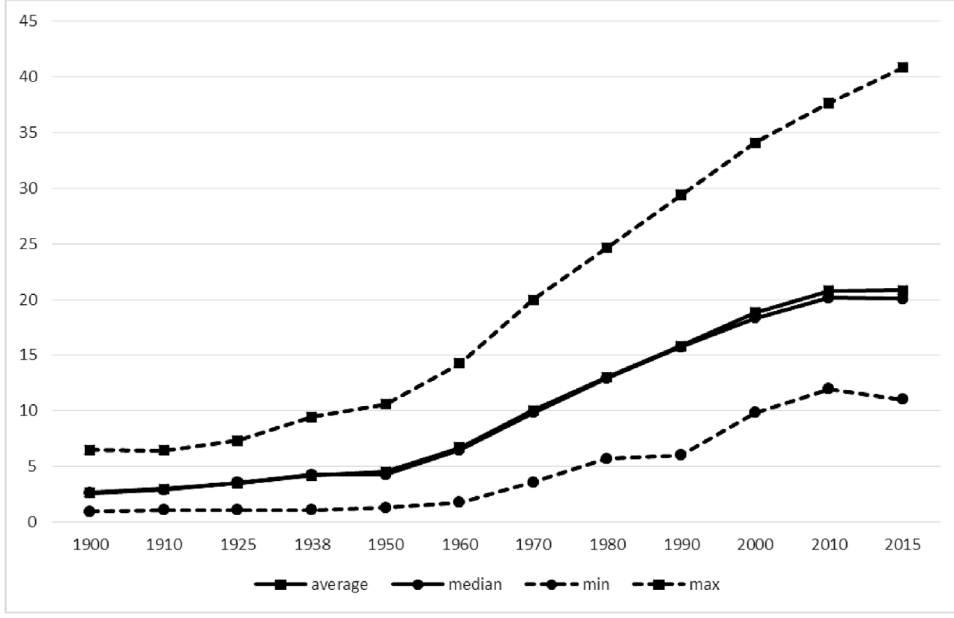
[3] We show the data based on 1990 GK$. The results would remain largely unchanged if we would use 2011 US$ instead.

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. Not 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 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 3 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.[4]

First, we see that the average level of GDP *per capita* has increased by more than 750 per cent over the last century in PPP, expressed in 1990 international dollars. This historically unprecedented increase in material living standards occurred mainly in a period of growth-acceleration 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 difference between the poorest and the richest regions has sharply increased over time, but the differences have declined in relative terms until about 1990. 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

Figure 3: GDP *per capita* across regions, 1900–2015 (1,000s, 1990GK$)



[4] This holds equally if we compare regions in terms of 2011 US$ PPPs.

pace. The 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 Chemnitz in Saxony (DED4).[5] 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, as we show in more detail below.

A simple way to show the distribution dynamics in our data is to represent them in form of a histogram, where we split the regional GDP *per capita* data for each year into evenly distributed bins as a simple approximation of the underlying probability density function. Figure 4 shows the distribution for the years 1900, 1950, and 1980; and also for 1980 and 2015.

We see in Figure 4(*a*) how between 1900 and 1950, and again between 1950 and 1980 the distribution shifted, with a growing number of regions positioned around the median. Especially strong is the shift between 1950 and 1980, which is mirrored in evidence about beta-convergence strong enough to reduce overall dispersion in GDP *per capita* across regions (compare Tables 3 and 4 and Figure 6, below). But this changed around 1980, when convergence weakened and dispersion across regions started to increase again. Figure 4(*b*) shows that between 1980 and 2010 the distribution shifted back, with a growing tail at the right and less mass around the median. 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 Iammarino *et al.* (2017), who link it to the rise of new economy industries and a new global division of labour. If we use the GDP data in 2011 US$ the development looks very similar, with an even more pronounced shift after 1980.

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 1949–90. Again, we note that the picture remains qualitatively unchanged if we use the GDP data in 2011 PPP, except that in 1980 Oslo (NO01) moves to the top of the distribution. For further details on Norway we refer the reader to the chapter by J. Modalsli in Rosés and Wolf (2019).

Maps 7(*a*, *b*) and 8(*a*, *b*) show more systematically the variation in GDP *per capita* across regions (in 1990 and 2011 PPPs, respectively). The first impression is that the pattern of variation was more compact back in 1900 than in 2015. Broadly speaking, we see a centre-periphery pattern: in 1900 some macro-regions, like England and north-western Europe, but also Switzerland, 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 2015, the picture is more complex. The relative decline of the UK is very visible, as is the impressive rise of the Republic of Ireland (mostly driven by Dublin). On the map there are 'islands'

[5] We always exclude Flevoland (NL23), because the region did not exist before the 1950s.

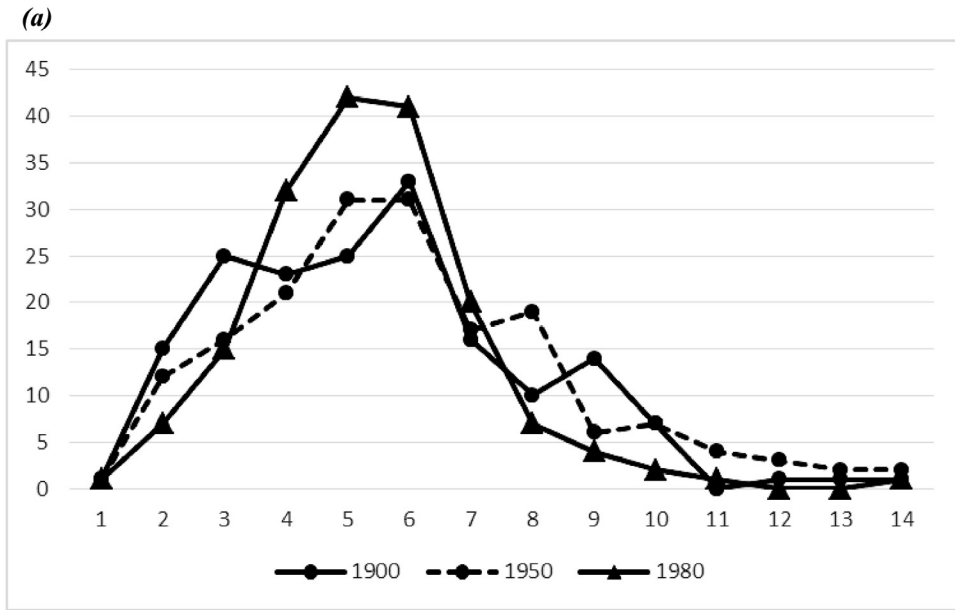
of prosperity, such as Paris (FR10) in France or Madrid (ES30) in Spain, surrounded by regions with below average GDP *per capita*. In addition, large parts of Scandinavia are now richer than average. Indeed, a measure like Moran's I shows clearly that the degree of spatial correlation has systematically 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 being replaced by a system based on several central regions, which commonly comprise the largest metropolitan areas.

VI. Explaining economic growth: fundamentals and shocks

(i) A long-run perspective: geography and institutions since 1900

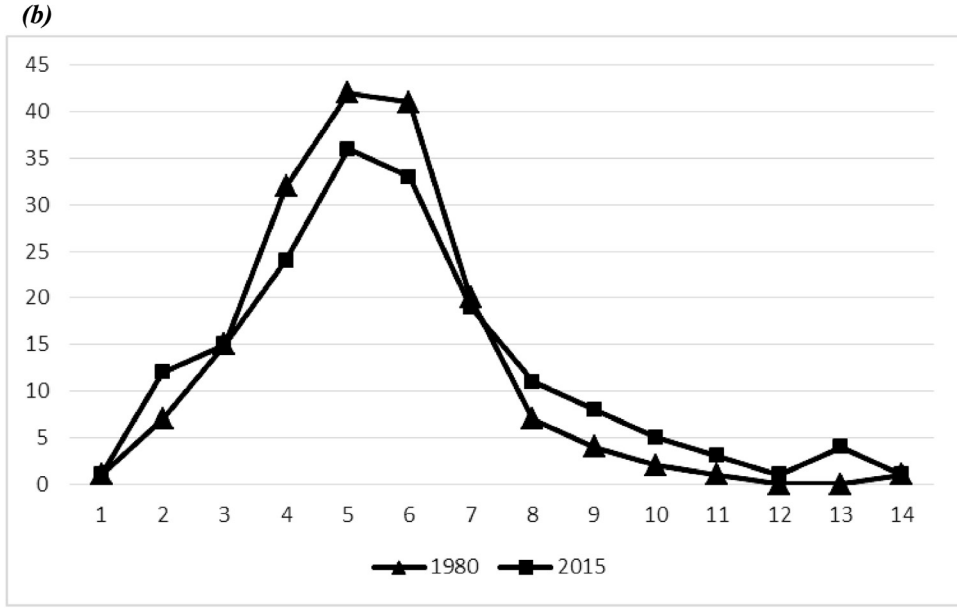
Let us first adopt a long-run perspective with regard to our time span of 115 years of regional development. Which factors can account for the large variation in growth rates across regions and over time since 1900? Following the literature, we can distinguish

Figure 4: Regional income distribution



Note: (Panel A) For the years 1900, 1950, and 1980 (1990 GK$) and (Panel B) for the years 1980 and 2015 (1990 GK$).
Source: Authors' own calculations.

Figure 4: Continued



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, pp. 3–25). 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 was related to 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.

In terms of geographical factors, we can distinguish between 'first nature' 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 here by the proximity to rock strata from the carboniferous era), and 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. The most important factor here is accessibility of markets that depends on both the economic size of neighbouring regions and access to them (Fujita *et al.*, 1999). We follow the literature (e.g. Redding and Venables, 2004) and approximate this for each point in time with the sum of GDP of all neighbours of a region, weighted by their geographical distance. If we want to regress this on the region's own GDP *per capita*, this clearly involves a problem of reverse causality. To limit this problem, we first exclude the region's own

GDP (that is access to the region's own market) and second use the sum of inverted distances to other regions as a simple instrument.

To what extent can these factors 'explain' variation in the level of GDP *per capita* across our regions and over time? Clearly, our results will show correlations, and need not reflect causality. Table 2 shows the results of a pooled regression that uses all our data. In column 1 we only use the largely exogenous 'first nature' variables; in column 2 we add time-varying national dummies, to capture differences due to national-level institutions. In columns 3 and 4, we add more variables capturing institutional variation within and between states, as well as market access.

Throughout, we added random effects, clustered standard errors at the regional level, and allowed for common time effects. Coefficients that are statistically significant at least at the 5 per cent level are shown in bold font. Consider the first column: 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 (significantly) associated with lower income levels. There are no significant long-run correlations between distance to ports and rainfalls with income levels. In column 2, we allow for variation at the national level, including different time trends at the level of nation-states. We add a separate set of dummies for regions located in the former GDR. Reassuringly, this leaves our results largely unchanged. In column 3, we allow for more institutional variation and add market access to capture the idea of 'second nature'. A good accessibility of markets is strongly positively correlated with income levels (note that this variable has been instrumented in order to deal at least partly with reverse causality). We also find support for the role of institutions: unsurprisingly, we find that capital regions have higher levels of income. Moreover, membership in the various European institutions was positively correlated with income levels, particularly membership in the Eurozone. Arguably, the most problematic variable here is our proxy for 'market access'. While our finding that good access to markets is in turn positively correlated with a region's income level is intuitive, this can also be quite misleading. For one, the instrument used here (and elsewhere in the literature) is likely to violate the exclusion restriction. More challenging is the fact that market access is very strongly correlated with proximity to coal. As shown in Pollard (1981), Fernihough and O'Rourke (2020), and others, since the nineteenth century European coalfields attracted migrants and industry and hence increased the size of regional markets. In our data, the cross-sectional correlation between market access and proximity to coal (proximity to rock strata from the carboniferous era) is above 0.85 for each year in the sample. Hence, in the last column of Table 2, we rerun the regression excluding market access. We see that this leaves our results qualitatively unchanged but doubles the coefficient estimated for coal.

(ii) Exploring growth dynamics: convergence and adjustment to shocks

Between 1900 and 2015, Europe experienced two World Wars, the Great Depression, and the fall of the Iron Curtain. All of this must have had strong effects on aggregate economic growth, visible at the regional level. To explore this in more detail, let us now adopt a more short-run growth perspective, starting with a simple Solow growth model (Solow, 1956; Barro and Sala-i-Martin, 1992). With no variation in terms of

Figure 5: Beta-convergence over time (172 regions)

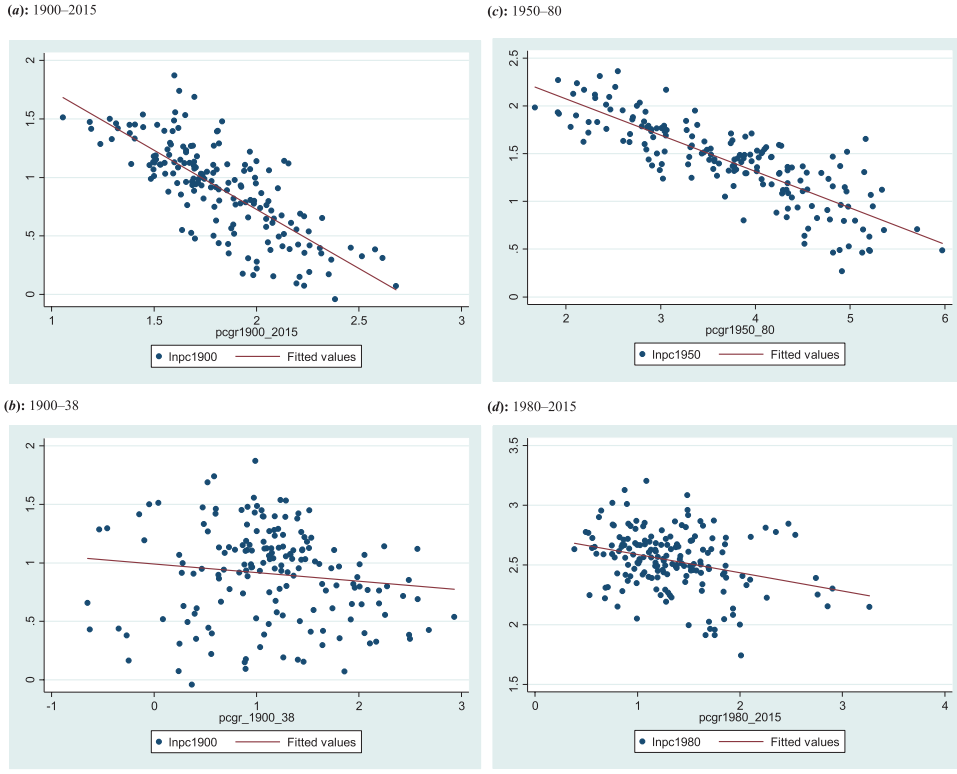
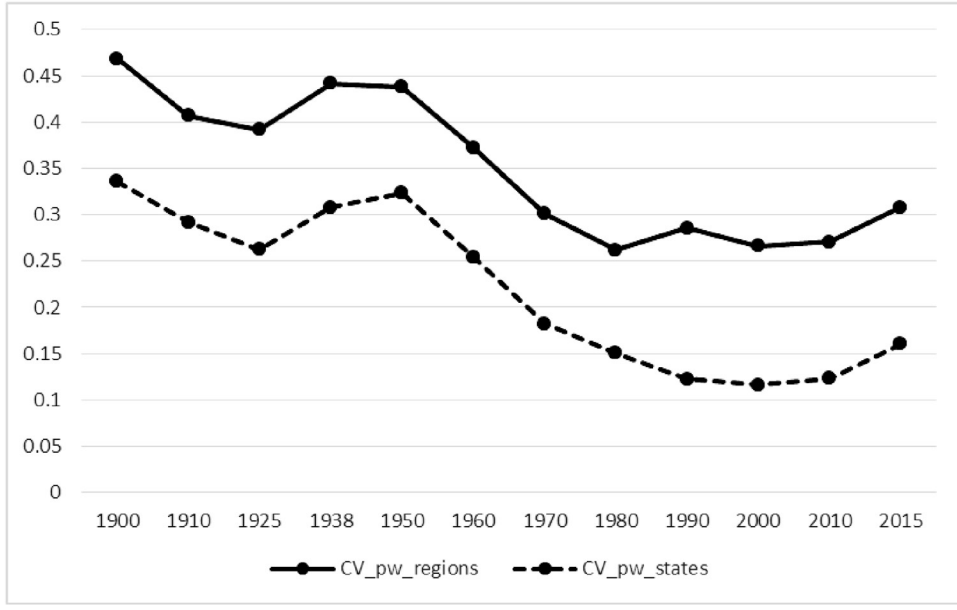


Figure 6: Sigma-convergence (population weighted coefficient of variation), 1900–2010



Note: The figure remains qualitatively unchanged if we use regional GDP in 2011 PPP instead.
Source: Authors' own calculations.

Table 2: Geography, institutions, and income levels; pooled GLS regression with random effects, IV, dep. variable: Ln (GDP *per capita*), 1900–2015

	First nature	First nature with national dummies	Geography and institutions	Geography and institutions with national dummies	Excluding market access
Distance to deepwater port (ln)	–0.016 (0.024)	–0.038 (0.028)	0.009 (0.019)	–0.009 (0.022)	–0.027 (0.023)
Extreme rainfalls	–0.069 (0.072)	–0.104 (0.096)	–0.041 (0.067)	–0.073 (0.081)	–0.085 (0.084)
Extreme temperatures	**–0.168** (0.087)	**–0.245** (0.097)	–0.127 (0.076)	–0.159 (0.081)	**–0.214** (0.767)
Soil quality	**–0.072** (0.035)	**–0.083** (0.043)	**–0.092** (0.037)	–0.070 (0.039)	–0.075 (0.037)
Coal potential	**0.260** (0.067)	**0.469** (0.083)	0.127 (0.082)	**0.256** (0.116)	**0.499** (0.069)
European Community	–	–	0.009 (0.018)	**0.092** (0.018)	**0.108** (0.020)
European Union	–	–	–0.070 (0.078)	0.096 (0.063)	**0.129** (0.064)
Eurozone	–	–	**0.129** (0.033)	**0.241** (0.051)	**0.233** (0.056)
Capital region	–	–	**0.292** (0.089)	**0.292** (0.081)	**0.297** (0.084)
Market access (instrumented)	–	–	**0.247** (0.093)	**0.337** (0.146)	–
National dummies (time-varying)	No	Yes	No	Yes	Yes
Constant	1.360 (0.425)	1.107 (0.506)	–0.127 (0.673)	–1.108 (1.001)	**0.920** (0.418)
Time dummies	Yes	Yes	Yes	Yes	Yes
Observations	2,058	2,058	2,058	2,058	2,058
Groups	172	172	172	172	172
R^2 within, between, overall	0.95, 0.30, 0.87	0.96, 0.18, 0.86	0.95, 0.46, 0.89	0.97, 0.47, 0.91	0.96, 0.40, 0.89

Note: The country-time dummies allow for country-specific linear time-trends. Standard errors in parentheses.
Source: Authors' own calculations.

Table 3: Convergence over time

	Beta	Standard error
1900–10	–0.847	0.226
1925–38	0.344	0.337
1950–60	–1.505	0.288
1960–70	–2.585	0.192
1970–80	–2.386	0.204
1980–90	–0.504	0.475
1990–2000	–2.584	0.265
2000–10	–0.922	0.318
2010–15	0.725	0.624

Source: Authors' own calculations.

Table 4: Convergence and adjustment to shocks, 1950–2015

	Constant	$<\hat{\beta}_0$ (standard errors) initial GDP *per capita*$_{1950}$	$\hat{\beta}_1$ (standard errors) 'gap'$_{1938\_50}$	$\hat{\beta}_2$ (standard errors) ag_sh$_{1950}$	$\hat{\beta}_3$ (standard errors) Ln(coal)	Adj. R2
1950–60	**4.93** (0.90)	**–0.88** (0.42)	**–4.82** (0.61)	1.01 (0.95)	0.34 (0.25)	0.38
1950–70	**6.29** (0.48)	**–1.74** (0.23)	**–1.68** (0.33)	0.38 (0.51)	**0.28** (0.13)	0.59
1950–80	**5.78** (0.33)	**–1.64** (0.15)	**–1.22** (0.21)	0.49 (0.34)	**0.16** (0.09)	0.74
1950–90	**5.23** (0.25)	**–1.36** (0.11)	**–0.81** (0.16)	0.22 (0.26)	0.03 (0.07)	0.77
1950–2000	**4.90** (0.21)	**–1.25** (0.09)	**–0.44** (0.14)	–0.08 (0.22)	–0.03 (0.06)	0.76
1950–2015	**4.16** (0.18)	**–0.95** (0.08)	–0.13 (0.12)	**–0.42** (0.19)	**–0.15** (0.05)	0.67

Source: Authors' own calculations.

institutions or geography, and in the absence of shocks, we expect to find that poor regions tend to exhibit above average growth rates in GDP *per capita*, which implies a process of convergence. 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*.

It might surprise in the light of a European history of shocks and disasters, but this is indeed suggested by Table 1 above. Some of the regions that were initially poorest (Extremadura, Nord-Norge, or Basilicata) showed growth rates far above the average. To test for this idea more systematically, we 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 5 we plot the average annual growth rates of regions (in per cent) against their initial level of GDP *per capita* (in logs). Figure 5(*a*) shows the result if we consider the entire sample period, 1900–2015. In Figures 5(*b*), 5(*c*), and 5(*d*) we distinguish between the periods 1900–38, 1950–80, and 1980–2015 (we exclude the region of Flevoland throughout to ensure comparability over time).

Figure 5(*a*) suggests that, in the long run, European regions converged in the sense that growth rates in GDP *per capita* were systematically higher than the initial level of GDP *per capita* (here 1900). Indeed, Figure 5(*a*) looks like a beautiful example of long-run convergence. However, we see in Figures 5(*b*), (*c*), and (*d*) that this is largely due to the period 1950–80: convergence was very weak before 1945, turned quite strong after

the Second World War, before it weakened again after about the 1980s. This absence of convergence before the war and the weakening of convergence since the 1980s is quite remarkable, and it is worthwhile to look into the details of this. Table 3 shows how the estimated beta-coefficients change over sub-periods (varying the initial year).

The pattern of convergence changed substantially over time. We find some evidence for convergence already before 1914, followed by a lack of convergence during the interwar period. Between 1925 and 1938 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 (Rosés and Wolf, 2010). After 1945, there is ample evidence *fo*r strong convergence. This was particularly fast during the 1960s and 1970s (well above 2 per cent), before it slowed down during the 1980s. Convergence returned briefly in the 1990s,[6] before it faded again in recent years. Still, the remarkable picture of convergence in the long run (based on 1900 as the initial year) remains somewhat puzzling, and motivates a more systematic analysis.

Barro and Sala-i-Martin (1995) suggested extending the simple Solow framework to take aggregate shocks into account. Along these lines, Temin (2002) argued for considering the period between 1950 and 1980, often termed the 'Golden Age of Growth' (Crafts and Toniolo, 1996), as a period of disequilibrium. The basic intuition developed by Temin (2002) is that the remarkable growth experience after 1945 was due to a combination of three factors: reconstruction after the war, structural change out of agriculture, and Solow-type convergence. Temin had to rely on estimates of GDP *per capita* at the level of nations and work with a very small data-set. Clearly, he missed the huge variation within European states as described above. Here, we redo Temin's exercise with our regional data. Moreover, our discussion on geographical factors showed that proximity to coal mattered for regional income levels. It is likely that the role of coal changed over time, from being a blessing to becoming a curse in many parts of Europe. Let us shortly revisit the basic intuition behind Temin's (2002) approach, and explain how to add coal to the picture.

First, some regions had suffered large-scale destruction during the war, alongside 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 after the war, rather than convergence as suggested in a Solow framework. We apply this idea to our data and construct a variable 'gap', simply defined as the log-difference in GDP *per capita* for 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,

[6] Given that we consider regions and nation states in their modern borders, our data include 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. Indeed, if we exclude Germany, the 'beta-coefficient' for the 1990s declines (from –2.58 to –1.0), but stays strongly significant. The other results from Table 3 stay largely unchanged.

such as Hamburg (DE60), 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. Second, Temin suggested that the disintegration of the European economy in the wake of wars and the Great Depression might have 'arrested' development, notably due to oversized agricultural sectors. The economic re-integration after 1945 should have facilitated structural change and freed up labour to move into more productive employment (in industry or services). A high share of agriculture in employment at the beginning of a period should be related to more scope for structural change and hence higher growth. However, structural change after 1950 was often related to coal. In the 1950 and 1960s, coal-abundant regions again attracted migrants and new industry. We hypothesize that proximity to coal was beneficial for regional growth during the industrialization, and again after 1945, when the European economy expanded and coal was in high demand. However, technological change, oil and gas findings, and the globalization of energy markets all contributed to a crisis in coal mining and regions that depended on it. Therefore, we suggest adding our measure of proximity to coal. We expect that its role for growth changed over time from positive to negative. Finally, the stable institutional environment after 1945 should have fostered catch-up growth in backward regions, due to a diffusion of technology and organisational change, such that the convergence suggested by Solow (1956) would become visible.

We now re-estimate the relationship between the average annual growth rate of GDP *per capita* (in per cent) and the initial level of GDP *per capita* (in logs), augmented by three variables: the 'gap' between 1938 and 1950, the share of agriculture in total employment of a region in 1950, and a region's proximity to coal, or $\hat{y}_{i,\ t1-t0} = \alpha + \beta_0 \ln(y_{i,\ t0}) + \beta_1 gap_{i,1938\_50} + \beta_2 ag\_sh_{i,1950} + \beta_3 \ln(cp_i) + \varepsilon_i$. Obviously, this is an explorative exercise, and we cannot interpret the estimated coefficients as reflecting causal mechanisms. For example, neither destruction during the war nor the share of agricultural employment is exogenous to income levels.

Intuitively, the effect of these factors should vary over time. Reconstruction effects should be strongest in the immediate post-war period and fading afterwards, similar to the scope for structural change away from agriculture. Also, the role of coal should change over time, as argued above. Hence, we vary the period under consideration, starting with average annual growth rates in 1950–60, and extending the period until 1950–90. This allows us to keep the initial year of income constant at 1950. Table 4 shows the results (where we exclude Flevoland and the regions of the former GDR to avoid the distortions). Coefficients that are significant at least at the 5 per cent level are shown in bold font.

The results from Table 4 are intriguing. They tell a story of convergence and adjustment to shocks that stretches over two generations. In the first decade after the Second World War, reconstruction dominates the picture. Regions that had suffered an absolute loss in income (due to war destruction, population change, or a combination thereof) grew much faster than others. In part, this can be explained by the reconstruction of electricity grids, road and rail infrastructure, and housing construction (see Vonyó (2008, 2012) on the case of West Germany). There is also some evidence for 'catch-up' growth during the 1950s, while we do not find support for the idea of 'arrested development' or coal regions growing much faster. This starts to change in the 1960s, when the effect of reconstruction growth weakens, while convergence gets stronger. The coefficient on the share of agriculture has the expected positive sign, but

is never significant. However, we see that during the 1960s regions with good access to coal grew faster than average, notably at a time of very high average growth rates. While there are clear signs of a 'coal crisis', on average European regions with good access to coal continue to grow fast until the end of the 1970s. Afterwards, convergence gets weaker and the effect of post-war reconstruction is disappearing. What is more, the benefits of coal abundance start to turn into a burden. Once we extend this until 2015, we find that coal regions are growing below average, similar to regions with a traditionally high share of employment in agriculture. The most dramatic example for a declining coal region is Hainaut (BE 32; see Table 1), but other former coal-mining regions in Germany, the UK, and France experienced a similar fate. Importantly, this helps to understand why the long-run picture of annual growth rates plotted against initial income in 1900 looks like beautiful convergence. In fact, this seems to be the result of a short period of massive catch-up by poor regions during the 'golden age', accompanied by downward convergence with growth rates far below the average in formerly rich regions since the 1970s. An important question, that we cannot address here, is why coal regions often failed to adjust. Possibly, a lack of human capital combined with barriers to mobility contributed to their decline (as has been suggested for the US by Glaeser and Saiz (2004)).

VII. Spatial convergence and divergence, 1900–2015

The evidence on beta-convergence after 1945 from Tables 2 and 3 seems inconclusive. Was this convergence strong enough to make regions more similar over time? A simple summary measure is the coefficient of variation (the cross-sectional standard deviation divided by the mean), often referred to as '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 6 shows the result of this for all regions and for regions aggregated to nation states. The difference between the two indicates how much of the variation is typically missed if researchers use national instead of regional data.

According to this measure, dispersion in GDP *per capita* has until about 1990, stagnated but started to increase after 2000. It is evident that measures based on regional data show more dispersion, but since 1980 we also observe some differences in trends. While on average, for the period 1900–60, measures based on national figures alone would capture around 70 per cent of all underlying variation, this figure has declined to 60 per cent in 1970 and below 50 per cent 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 2015. 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. Moreover, some formerly rich regions started to fall behind, as discussed in section VI above, while capital regions show above average growth rates (their share in total sample GDP increased since 1980).

Finally, let us return to the question of the dispersion of economic activity across regions from the perspective of geographical distribution. In Figures 5(*a*) and 5(*b*) we

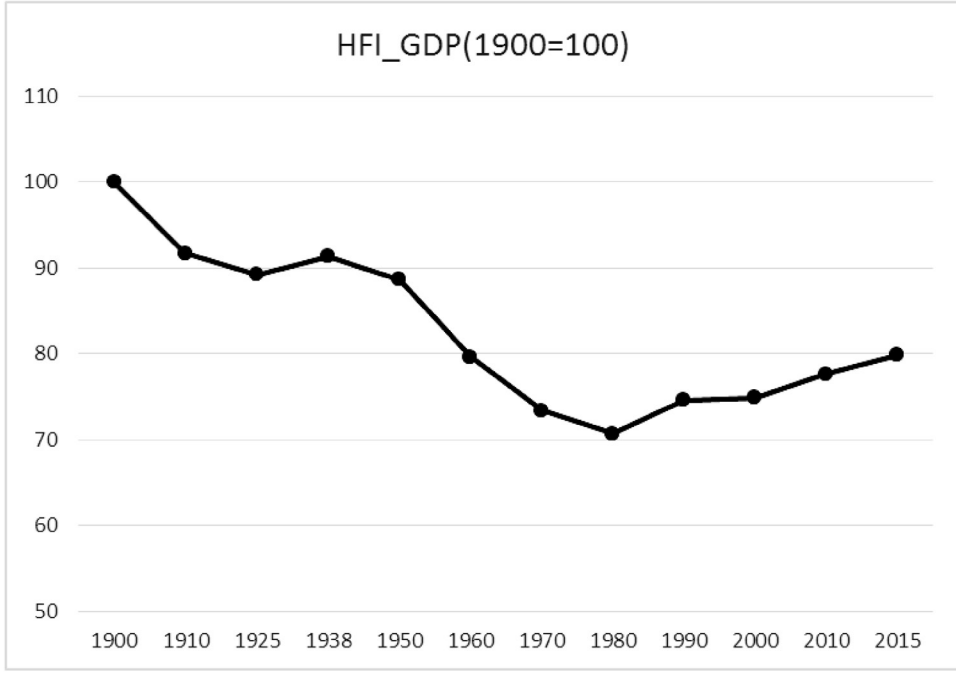
could see that the distribution of GDP *per capita* across regions had become more equal between 1900 and 1980, but that this process was reversed between 1980 and 2015. Underneath this change is a geographical re-concentration of economic activity, in terms of population and employment, but more so in terms of GDP. A very simple way to look at this is to construct a Herfindahl index of regional GDP for each year, indexed to 1900=100 (see Figure 7).

We find a picture very similar to that of sigma-convergence. The concentration of GDP across regions declined in the long run, but this trend was stopped and even reversed around 1980. If we would use GDP in 2011 PPP instead, we would find a more pronounced decline in concentration after 1950, again stagnation since 1980, followed by a recent increase.

Other recent studies suggest that this phenomenon continues and is not limited to Europe but applies to other OECD countries as well (see OECD, 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.

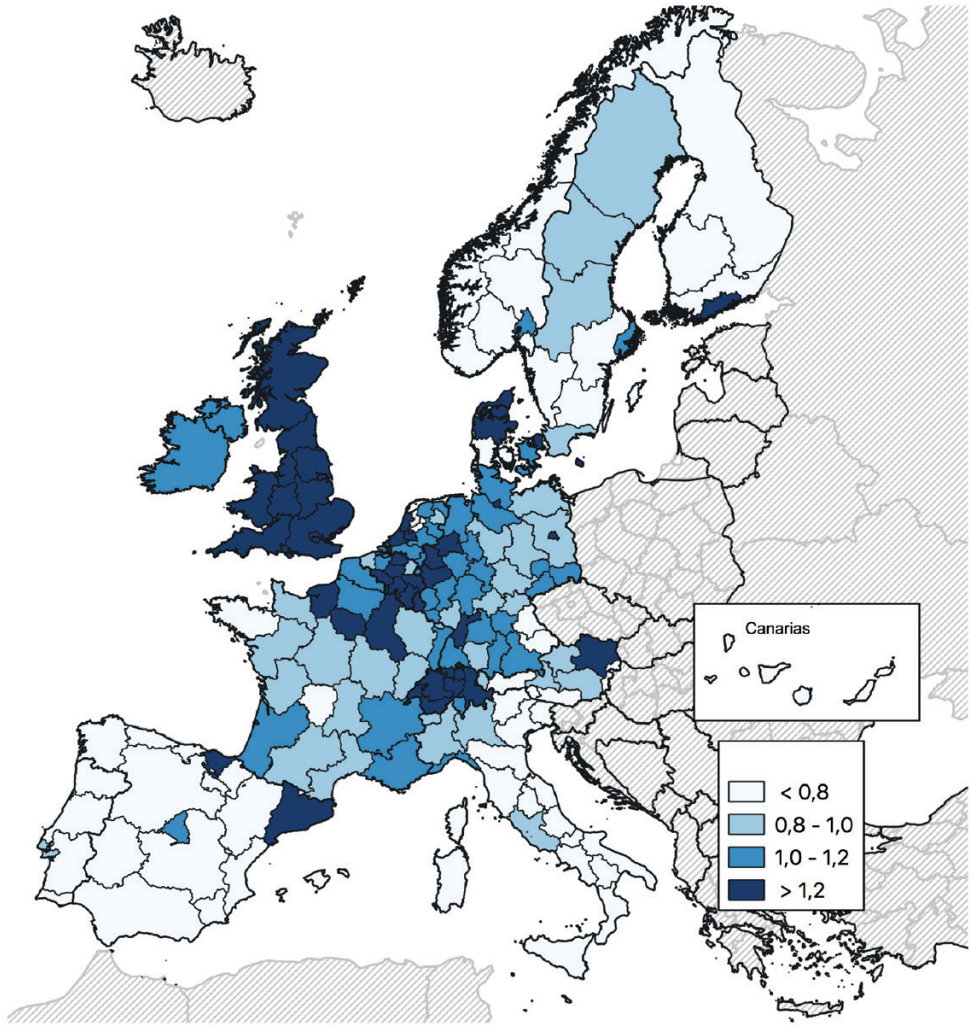
It is remarkable that the emerging picture on regional inequality in the long run is also similar to the pattern of inequality in terms of personal income and wealth

Figure 7: Herfindahl index of regional GDP, 1900–2015



Source: Authors' own calculations.

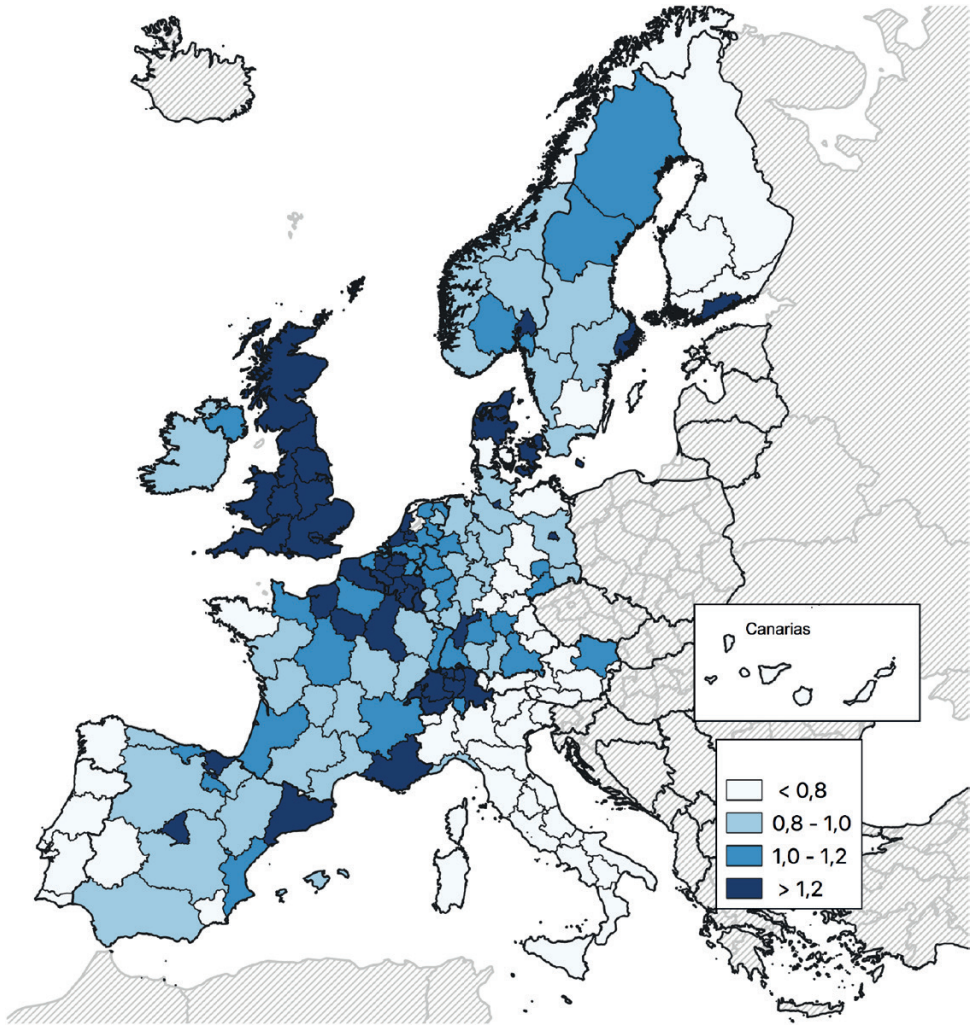
Map 7(*a*): Relative GDP *per capita*, 1900 (1990 GK$)



Source: Authors' own drawing.

distributions, which has been extensively documented and analysed (Piketty and Saez, 2003; Piketty, 2014). Regional inequality declined since 1900 but started to increase again around 1980, very much at the same moment as 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-skilled, 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.

Map 7(*b*): Relative GDP *per capita*, 1900 (2011 US$)

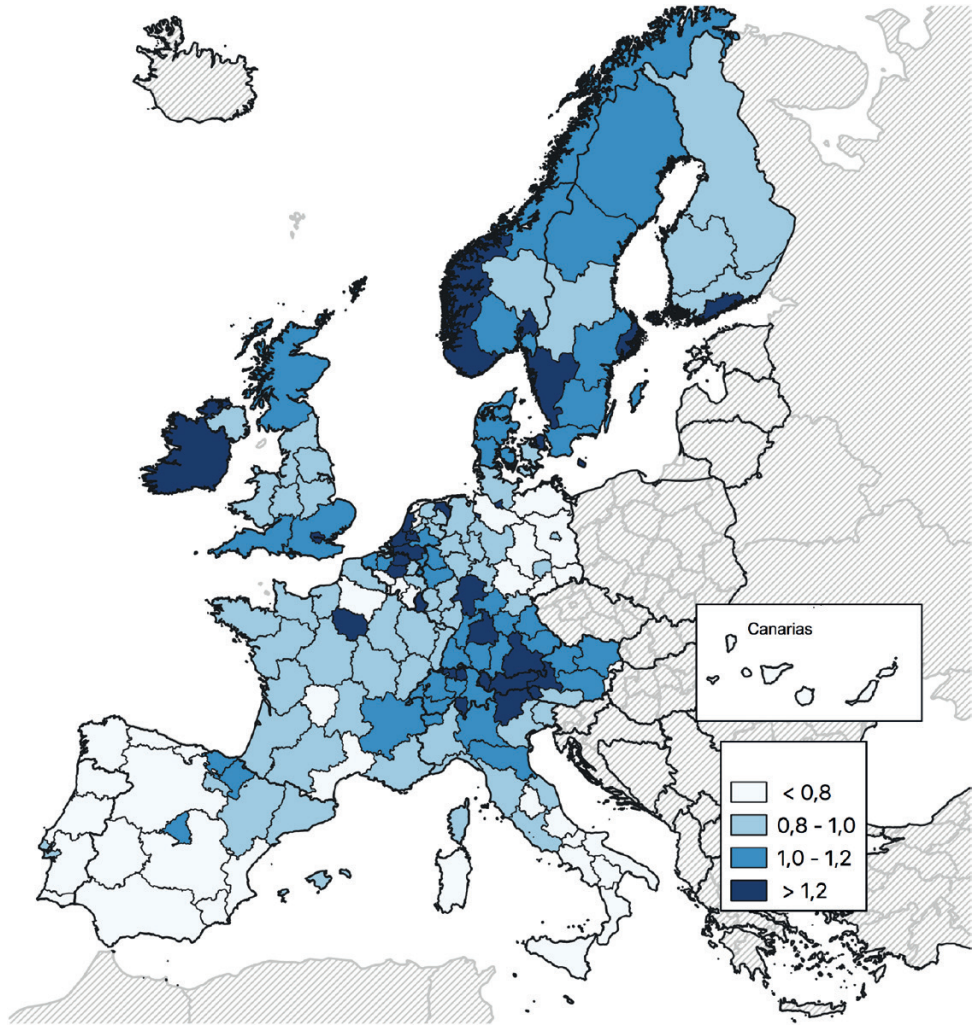


Source: Authors' own drawing.

VIII. Conclusion

How did European regions do over the last 115 years? We have constructed a new dataset at the level of 173 regions to trace their economic development in terms of employment structure and GDP at PPP 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. Some regions kept substantial shares of employment in agriculture, while employment in industry first spread before it became more concentrated again. Variation in the levels of GDP *per capita* is systematically related to 'fundamentals', and we discussed how changes in growth rates after 1945 were affected by reconstruction after the war, convergence, and a changing role of coal.

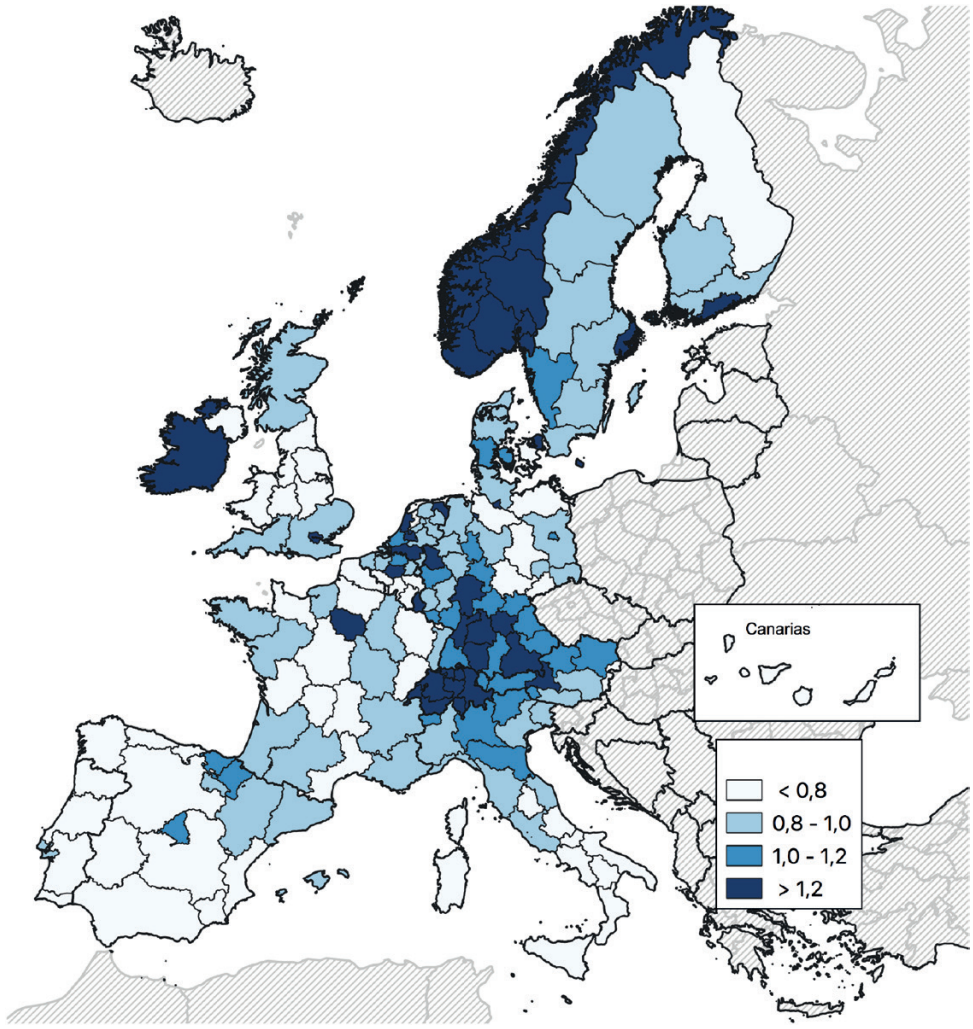
Map 8(*a*): Relative GDP *per capita*, 2015 (1990 GK$)



Source: Authors' own drawing.

But the long-run data allow us to see something more. 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 per cent in 1900 to above 50 per cent in 2015, notably from 1980 onwards. Formerly rich, industrialized regions failed to adjust to structural change, such as the decline of coal-mining. Most important, perhaps, is the observation that the pattern of regional inequality over the last 115 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 them until about 1980, and divergence as well as geographical re-concentration thereafter.

Map 8(*b*): Relative GDP *per capita*, 2015 (2011 US$)



Source: Authors' own drawing.

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 such as 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.

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, 157–86.

Barro, R. J., and Sala-i-Martin, X. (1992), 'Convergence', *Journal of Political Economy,* **100**(2), 223–51.

— — (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*, **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 van Zanden, J. L. (2014), 'The Maddison Project: Collaborative Research on Historical National Accounts', *The Economic History Review*, **67**(3), 627–51.

— Inklaar, R., de Jong, H., and van Zanden, J. L. (2018), 'Rebasing "Maddison": New Income Comparisons and the Shape of Long-run Economic Development', Maddison Project Working Paper 10, University of Groningen.

Broadberry, S. (1997), 'Anglo-German Productivity Differences 1870–1990: A Sectoral Analysis', *European Review of Economic History*, **1**(2), 247–67.

— Klein, A. (2012), 'Aggregate and Per Capita GDP in Europe, 1870–2000: Continental, Regional and National Data with Changing Boundaries', *Scandinavian Economic History Review*, **60**(1), 79–107.

— 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**, 59–83.

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*, **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*, **32**(6), 1036–58.

Eichengreen, B. (2007), 'The Real Exchange Rate and Economic Growth', *Social and Economic Studies*, **56**(4), 7–20.

Fernihough, A., and O'Rourke, K. H. (2020), 'Coal and the European Industrial Revolution', *The Economic Journal*, ueaa117. doi:10.1093/ej/ueaa117.

Fujita, M., Krugman, P., and Venables, A. J. (1999), *The Spatial Economy*, Cambridge, MA, 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*, **49**(4), 644–64.

Geary, F., and Stark, T. (2002), 'Examining Ireland's Post-famine Economic Growth Performance', *The Economic Journal*, **112**(482), 919–35.

Glaeser, E. L., and Saiz, A. (2004), 'The Rise of the Skilled City', *Brookings–Wharton Papers on Urban Affairs 2004*, Washington, DC, Brookings Institution Press, 47–105.

Iammarino, S., Rodríguez-Pose, A., and Storper, M. (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.

Marshall, M. (1987), *Long Waves of Regional Development*, London, Palgrave.

Moretti, E. (2013), 'Real Wage Inequality', *American Economic Journal: Applied Economics*, **5**(1), 65–103.

OECD (2016), *OECD Regions at a Glance 2016*, Paris, OECD Publishing.

Piketty, T. (2014), *Capital in the Twenty-First Century*, Cambridge, MA, Harvard University Press.

Piketty, T., and Saez, E. (2003), 'Income Inequality in the United States, 1913–1998', *Quarterly Journal of Economics*, **118**(1), 1–39.

Pollard, S. (1981), *Peaceful Conquest: The Industrialization of Europe, 1760–1970*, Oxford, Oxford University Press.

Redding, S., and Venables, A. J. (2004), 'Economic Geography and International Inequality', *Journal of International Economics*, **62**, 53–82.

Rosés, J. R., and Wolf, N. (2010), 'Aggregate Growth, 1913–1950', in S. Broadberry and K. H. O'Rourke (eds), *The Cambridge Economic History of Modern Europe*, *Vol.* **2**, Cambridge, Cambridge University Press, 181–207.

— — (eds) (2019), *The Economic Development of Europe's Regions: A Quantitative History Since 1900*, Routledge Explorations in Economic History.

Schön, L., and Krantz, O. (2012), 'The Swedish Economy in the Early Modern Period: Constructing Historical National Accounts', *European Review of Economic History*, **16**(4), 529–49.

Solow, R. M. (1956), 'A Contribution to the Theory of Economic Growth', *The Quarterly Journal of Economics*, **70**(1), 65–94.

Temin, P. (2002), 'The Golden Age of European Growth Reconsidered', *European Review of Economic History*, **6**(1), 3–22.

van Zanden, J. L. (1987), 'Economische groei in Nederland in de negentiende eeuw: enkele nieuwe resultaten', *Economisch- en Sociaal-Historisch Jaarboek*, Vol. **50**, 51–76.

Vonyó, T. (2008), 'Post-war Reconstruction and the Golden Age of Economic Growth', *European Review of Economic History*, **12**(2), 221–41.

— (2012), 'The Bombing of Germany: The Economic Geography of War-induced Dislocation in West German Industry', *European Review of Economic History*, **16**(1), 97–118.

Williamson, J. G. (1965), 'Regional Inequality and the Process of National Development: A Description of the Patterns', *Economic Development and Cultural Change*, **13**(4, Part 2), 1–84.