CONTRIBUTI DI RICERCA CRENOS

UNIVERSITÀ DI CAGLIARI UNIVERSITÀ DI SASSARI
CENTRE FOR NORTH SOUTH ECONOMIC RESEARCH
CRENoS
CENTRO RICERCHE ECONOMICHE NORD SUD

ASSESSING THE EUROPEAN CONVERGENCE MACHINE: DO COUNTRIES CONVERGE AND REGIONS DO NOT?

Maria Grazia Pittau
Andrea Vaiano
Roberto Zelli

WORKING PAPERS

2025/07

arkadia

CENTRO RICERCHE ECONOMICHE NORD SUD
(CRENOS)
UNIVERSITÀ DI CAGLIARI
UNIVERSITÀ DI SASSARI

CRENOS was set up in 1993 with the purpose of organising the joint research effort of economists from the two Sardinian universities (Cagliari and Sassari) investigating dualism at the international and regional level. CRENoS' primary aim is to improve knowledge on the economic gap between areas and to provide useful information for policy intervention. Particular attention is paid to the role of institutions, technological progress and diffusion of innovation in the process of convergence or divergence between economic areas. To carry out its research, CRENoS collaborates with research centres and universities at both national and international level. The centre is also active in the field of scientific dissemination, organizing conferences and workshops along with other activities such as seminars and summer schools.
CRENoS creates and manages several databases of various socio-economic variables on Italy and Sardinia. At the local level, CRENoS promotes and participates to projects impacting on the most relevant issues in the Sardinian economy, such as tourism, environment, transports and macroeconomic forecasts.

www.crenos.unica.it
crenos@unica.it

CRENOS – CAGLIARI
VIA SAN GIORGIO 12, I-09124 CAGLIARI, ITALIA
TEL. +39-070-6756397

CRENOS - SASSARI
VIA MURONI 25, I-07100 SASSARI, ITALIA
TEL. +39-079-213511

Title: ASSESSING THE EUROPEAN CONVERGENCE MACHINE: DO COUNTRIES CONVERGE AND REGIONS DO NOT?

Prima Edizione: Marzo 2025
ISBN: 978 88 6851 576 8

Viale Bonaria 98 - 09125 Cagliari
Tel. 070/6848663 - info@arkadiaeditore.it
www.arkadiaeditore.it

Assessing the European convergence machine: do countries converge and regions do not?

Maria Grazia Pittau
Sapienza University of Rome & CRENoS

Andrea Vaiano
Sapienza University of Rome

Roberto Zelli
Sapienza University of Rome

March 25, 2025

Abstract

According to the World Bank report (Gill & Raiser, 2012), the EU has become the modern world's greatest "convergence machine". While the process of convergence has been acknowledged at country level, results at regional level are still unclear. Using the most advanced techniques, we assess convergence across European NUTS2 regions over forty years. The distributional dynamic approach unveils different perspectives that traditional methods have overlooked. We conclude that a process of catching-up between low- and middle-income regions has been in progress, while wealthy regions have been drifting away.

Keywords: EU regions, Club convergence, Mixture models.
JEL Codes: C14, O11, O47.

1 Introduction

The European Union (EU) was not conceived to be a small league of privileged Member States, but rather to extend favorable conditions of living and trading to as many countries as possible. Starting from the Treaty of Rome and throughout its policy, the EU has taken measures to help the less developed areas catch up with more developed ones. The goal of "harmonious development" required budgetary instruments to support regional growth: the creation of a European Social Fund (ESF) was destined to improve employment opportunities, facilitate mobility and raise living standards of workers. The first concrete fund for regional distribution, the European Regional Development Fund (ERDF), was established in 1975 with the purpose of contributing to reducing disparities between the levels of development of European regions. It was designed to strengthen economic, social and territorial cohesion in the EU.

Although a small fund when created, it set the basis for further reforms leading to an institutionalised Cohesion Policy in the late 1980s (European Commission, 2024b). The Mediterranean enlargements of 1981 (Greece) and 1986 (Portugal and Spain) resulted in widening disparities, since those countries were facing multiple socioeconomic challenges after liberalisation from dictatorship. The Single European Act (SEA) came into effect in 1987, providing juridical basis to the ERDF and establishing the legislation for Cohesion Policy. The Maastricht Treaty on the European Union marked the foundation of the EU as we know it today; attention to regional cohesion was required to mitigate the possibly negative effects of market unification. During the 1994–1999 period, the Cohesion Fund (CF) was introduced for the poorest countries in the Union. With the accession of Austria, Finland and Sweden in 1995, support was also shared to sparsely populated regions in the northernmost regions. The Eastern Enlargement in 2004 widened the geographical scope of Cohesion Policy considerably. Entering the EU often implies undergoing significant institutional and infrastructural transformations, hence candidate governments also received preliminary funds through the Instrument for Pre-Accession Assistance (IPA). Central and Eastern European (CEE) countries were indeed granted financial support with the specific goal of enhancing their productive systems and opening their economies. The Lisbon Treaty in 2007 reinforced the Cohesion Policy by recognising, for the first time, territorial cohesion as a fundamental objective of the EU. The Lisbon Strategy, along with the place-based regional policy approach of the Report by Barca (2009), highlighted the importance of economic support of sub-national levels (namely NUTS 2 or NUTS 3) that would later be iterated by the "Europe 2020" Strategy.

The purpose of reaching structural and economic homogeneity is even more important in the context of a progressively expanding community. When the ERDF was officially established in 1975, the EU consisted of nine Member States; the latest installment (covering from 2021 to 2027) addresses twenty-seven countries. The landscape of involved territories continues to evolve, encompassing diverse realities and stages of development. The European Commission (EC) recognises three categories of regions: *less developed* regions, whose per capita GDP is below 75% of the Union average; *transition* regions, whose per capita GDP is between 75% and 90% of the Union average; and *more developed* regions, whose per capita GDP is above 90% of the Union average. Less developed and transition regions together are sometimes referred to as "cohesion regions" (European Investment Bank, 2022). The designation is used to determine the allocation of funds: regions receive access to specific resources and are expected to comply with different objectives according to their status. The operating framework is continuously updated to guarantee punctual interventions and deal with crises; thematic areas concern agriculture, sustainability, infrastructure, education, institutional capacity building, job creation and more. This financial tool includes the ERDF, the CF and the ESF, including the Youth Employment Initiative. NUTS 2 regions are considered as the preferred geographical classification, offering an ideal balance between detailed precision and broad generalization. They serve as the reference entities for the allocation structural funds (European Commission, 2024b; Fischer & Stumpner, 2008; Monfort, 2020).

Convergence in per capita GDP has been observed in Europe since the 1980s. In particular, CEE countries have experienced unprecedented growth since joining the EU, reducing disparities at the national level (European Commission, 2024b). However, development has not spurred equally at the regional level: capital regions have been agglomerating job opportunities, services, economic activity and innovations (European Investment Bank, 2022). The Cohesion Policy aims to sustain regional cooperation across and within countries, by stimulating investments that can amend unbalances. Gill and Raiser (2012) praised the ability of creating a thriving environment for firms and national economies to grow extraordinarily, igniting a "convergence machine". Regional disparities have been declining through the 1960s and 1970s, while income inequality remained generally unchanged during the 1980s (Paci, 1997). At the turn of the century, contrasting evidence started to question the effectiveness of cohesion policies on reducing disparities. Boldrin and Canova (2001) criticised the current interventions and questioned the effectiveness of their role in actively fostering economic growth. Fischer and Stumpner (2008) and Pittau and Zelli (2006) observed a slow process of catching up of the poorest regions to a distinctive middle-income class, with a small cluster of very rich regions sliding away. This empirical evidence differs from the results of López-Bazo et al. (1999) and Quah (1996b), but the analyses built on different samples and data pre-processing. Becker et al. (2018) found positive but short-lived effects of the former *Convergence Objective* transfers, arguing for the contribution of other trends to the increasing disparities. Merler (2016) observed convergence at the regional level; Nagy and Šiljak (2022) confirmed convergence of CEE countries towards Western Europe and highlighted the consistently positive effect of EU membership on economic growth. In general, the results observed during the whole accession process and beyond testify the impact of institutions as a fundamental cause of long-run growth (Acemoglu et al., 2005).

Assessing income distribution is still equally relevant today in the context of a new Union, characterised by an expanded backbone, numerous candidate members and crises which call for innovative policy approaches. The 2020s have been a time of unprecedented challenges, but also exceptional solidarity. Overall, positive results have been balanced by growing disparities. The European Commission (2024b) suggested to upgrade the system of resources destined to regional development and cohesion objectives. The available data provide adequate context to appraise trends in the long run and partake to the persisting debate around convergence. We provide an overview of growth and convergence across the EU over forty years. We then complement the analysis by introducing mixture models to formally explore the distribution of income in motion. The adoption of a dynamic approach addresses the need for multiple perspectives on the evolution of regional income distribution over time to uncover the presence of a balanced path towards greater prosperity. This examination leads to a comprehensive evaluation of the processes shaping the socioeconomic landscape after four decades of European policy efforts, delivering a makeshift perspective on the topic.

The aim of the analysis is to assess two main questions, that will be answered throughout the text: is there overall convergence of European regions or do countries converge but regions do not? Can we observe actual harmonic development –in terms of poor regions catching up to wealthy regions– or has there rather been a scattered growth pattern? Section 2 describes the dataset and the choice of regions to analyze. In Section 3, we explore absolute β-convergence. Results for different specifications are scrutinised together with strengths and limitations of the method. Section 4 introduces a distributional approach, according to non-parametric (4.1) and semi-parametric specifications to approximate income distributions with mixtures of normal component densities (4.2). Finally, a flexible classification of regions into sub-populations with distinctive economic characteristics is proposed (4.3). Section 5 summarises the findings related to each specific method, providing a comprehensive conclusion to the analysis.

2 Data and territorial units

EUROSTAT is the reference agency for data concerning Europe in general and the European Union in particular. ARDECO is the Annual Regional Database of the European Commission's Directorate General for Regional and Urban Policy, and it contains time-series variables and indicators at various statistical scales. We refer to ARDECO for data on GDP and population of regions in the EU and in some partner or candidate countries (Eurostat, 2024a). It must be noted that previous studies often employed different data one from the other, due to availability issues before the introduction of ARDECO and in relation to the specific research questions. For example, data from Quah (1996b) excluded developing regions from Greece and Portugal. The choice of including some regions can influence the conclusions and induce discordance among results.

Recalling the distinction proposed by Milanovic (2002), we focus on *inter-national* inequality. This is the more poignant and relevant concept at the macro level, as we are investigating convergence of territories rather than individual levels of income or consumption. The unit of interest is not the single European citizen but rather each region. Controversies around the use of NUTS 2 have been extensively documented (Boldrin & Canova, 2001). One sensible topic is the role of commuting workers, highly concentrated in some particular contexts, which produce income in a region but dispose of it somewhere else. A combination of NUTS 1 and NUTS 2 has been proposed as a tentative compromise (Paci, 1997; Pittau & Zelli, 2006). Notwithstanding the disputes associated with these administrative characterisations, they still remain the reference unit of analysis in the EU. Therefore, in order to maintain coherence with the benchmark, we employ the NUTS 2 standard for all countries. Constraining the analysis to European regions allows for a common framework of definitions and data management: information on relative price levels is available in the form of Purchasing Power Parities (PPPs). These are calculated at country level, without further regional breakdowns (Eurostat, 2024b). PPPs enable to compute prices into the Purchasing Power Standard (PPS), a fictitious currency corresponding to equivalent baskets of goods and services across countries. The conversion of per capita GDP at current prices in PPS returns a cross-sectional comparable measure of the volume of an economy. Data expressed in the common artificial currency enable comparability across countries by removing the effects of differences in price levels. To further control for the common growth and business cycles, income can be divided by the population weighted average per capita GDP of EU regions in a given year.[1] To achieve temporal comparability, we account for inflation and other factors by applying deflators for each year. Per capita GDP is reported from 1980 to 2025; as measurements from 2023 onwards are forecasted (Eurostat, 2024b), the window of interest is set from 1980 to 2022.

The full dataset comprises 334 European regions at NUTS 2 level, including prospective and former EU Members as well as partner states, for a total of 37 distinct countries.[2] We include all countries for which complete data exist because they constitute a concrete group of interest in the socioeconomic and political context that we set out to explore. The panel is composed of countries for which data are available from 1980 to 2022: the EU-15 (Austria, Belgium, Denmark, France, Finland, Germany[3], Greece, Ireland, Italy, Luxembourg, The Netherlands, Portugal, Spain, Sweden, United Kingdom), Norway, Switzerland and Turkey.[4] Data for other countries do not span the whole period at this level of disaggregation, for example in the case of former Yugoslavia and Soviet Republics. A total of 80 new regions entered the dataset between 1990 and 2000, which then remained substantially unaltered.[5]

The enlargement process is a central aspect in the EU, but the evolving composition

[1]Several alternative approaches to standardisation can be applied to obtain a measure of relative income. We divide by the average of EU regions according to the configuration in each year.

[2]No data are reported for either Liechtenstein or Iceland, hence 332 regions from 35 countries are effectively represented in the data.

[3]Only 30 out of 38 German regions actually report data over the entire period: observations for East Germany and Berlin (which constitutes its own entity at NUTS 2 level) are available from 1991.

[4]Both Norway and Switzerland actively participate in funding activities that aim to reduce economic and social disparities. Turkey is a long-term IPA recipient and benefited from a €3.533 million indicative allocation for 2014-2020 –not including the allocation for Cross-border Cooperation– tangible sign of EU support to programs of regional development and territorial cooperation (European Commission, 2024a).

[5]The only exception is the Norwegian statistical designation *Svalbard og Jan Mayen*, for which data exist starting from 2008.

of the dataset within an ever-expanding community poses a methodological challenge to conducting a consistent assessment of catching up. Finally, the region of *Inner London - West* represents a clear outlier – showing double the income of the second richest territory – and is excluded from the analysis. Indeed, extreme values can hinder estimation by introducing a fictitious tweak in the upper tail of the distribution and distorting the results of both convergence regressions and mixtures. After removing this observation from the data, the definitive panel represents 250 European regions at the NUTS 2 level of aggregation.

Significant movements in the distribution do not tend to happen overnight, leading to the choice of analysing selected years representing a time period. Specific reference points were chosen in alignment with the dates of the strategic frameworks of the Cohesion Policy. The Union programming efforts represent an ideal benchmark for monitoring the evolution of income distribution. The years of interest are 1986, 1993, 1999, 2006, 2013, 2019 and 2022.[6] The choice of these points of view allows to encompass the distributional analysis into the wider frame of EU development strategies in the context of regional growth-oriented policies and initiatives.

3 The convergence hypothesis

The converge theory predicts that poor regions are going to catch up to richer regions in the long run, driven by a faster growth of their GDP. *Absolute convergence* embodies a neoclassical perspective, premised on the assumption of diminishing returns to capital (Barro, 1991). This model builds on a simple regression of per capita GDP at time $t+\Delta t$ on the initial per capita GDP, as reported in Equation 1, where ε represents an error term. The sign and magnitude of the estimated β coefficient determine, respectively, direction and intensity of the phenomenon: under the ordinary least squares (OLS) estimation, a negative $\hat{\beta}_t$ indicates convergence, while a positive $\hat{\beta}_t$ suggests divergence.

$$\log(GDPpc_{i,t+\Delta t}) - \log(GDPpc_{i,t}) = \beta_t \log(GDPpc_{i,t}) + \varepsilon_{i,t} \tag{1}$$

The concept of half-life (τ) refers to the time required to halve the gap between the current income level and the steady state. Clearly, τ is inversely related to β: larger values of the convergence coefficient result in shorter catch-up times.

[6]The year 2022 is taken as a state-of-the-art gauge, being the last available year in the data. We do express caution in evaluating trends over the latter interval, since notably narrower than its predecessor and intertwined with the struggles of post-pandemic recovery.

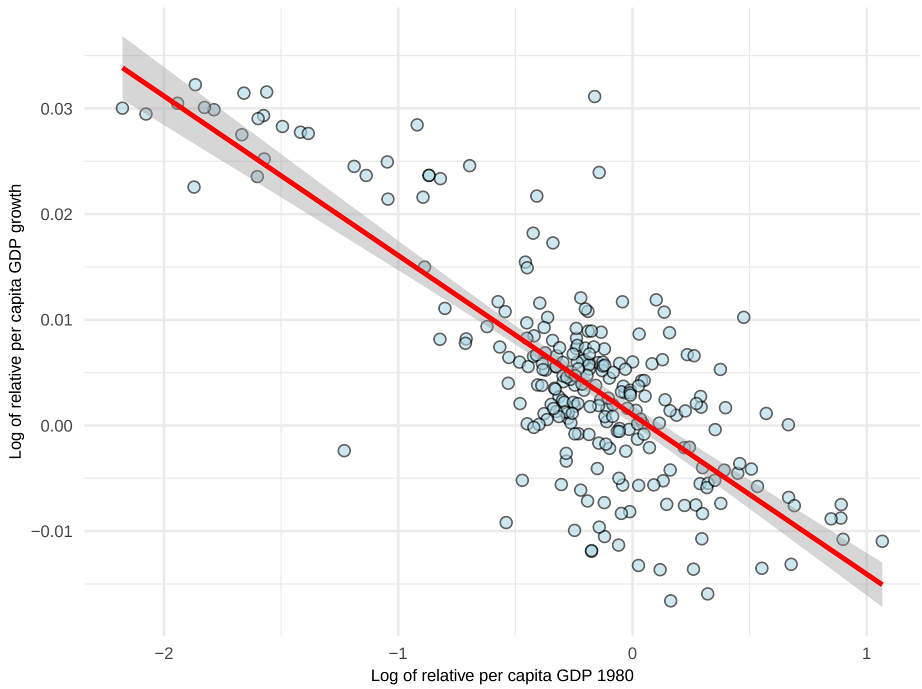


Figure 1: β-convergence in the period 1980-2022 for the panel of 250 European regions.

Income distribution across European regions has been extensively documented in the first decade of the century. Results have shown mild processes of convergence, indicated by statistically significant β coefficients with very small negative value. Over the four decades covered by the data, we observe that statistical evidence of β convergence is characterized by $\hat{\beta} = -0.01508$ (Figure 1). The coefficient is statistically significant but lower than the stable uniform rate of 2% heralded in the reference literature (Barro, 1991; Barro & Sala-i-Martin, 1992; Quah, 1996a), yielding half-life time $\tau \approx 46$ years.

Figure 2 represents the contribution of each quartile to the slope of the regression line: the segment associated to each group enables to compare distinct behaviours. Interestingly, the second quartile exhibited almost null β convergence, while the third quartile produced the steepest segment. These trends would suggest that the upper middle-income regions of 1980 exhibited the highest growth in this period. The opportunity of having access to financial support from the EU paired with a solid socioeconomic background therefore appears as the ideal setting for growing. The status of middle income regions stayed substantially unchanged. This decomposition is a first attempt at exploring income dynamics of subsets characterised by independent endevours to provide a dimension of movement within the distribution.

Following the approach delineated in Kremer et al. (2022), we fix Δ_t and change the starting year to explore trends over different time frames. Results are proposed for

Figure 2: β-convergence in the period 1980-2022, with the LOESS regression fragmented by quartile, for the panel of 250 European regions.

$\Delta t = \{1, 5, 10\}$ in Figure 3. We observe consistent patterns across the intervals. The coefficients computed for the 1 year window are characterised by greater fluctuations and the highest absolute values. A first phase of convergence peaks in the early 1980s, is exhausted by the second half of the decade and becomes effectively overturned by the 1990s. The latter diverging trend is evident in all three specifications and is significantly protracted for $\Delta t = 10$. Convergence is observed again at the beginning of the new century, culminating in the late 2000s. The values of $\hat{\beta}$ initially match the magnitude of the 1980s, but slowly deteriorate after 2010. It is difficult to comment current patterns, as convergence should be appreciated on the medium- to long-term (Islam, 1995).

Figure 4 presents a heat map showing the size of estimated β coefficients for all possible combinations of year intervals. The horizontal blocks correspond to a fixed starting year, while vertical blocks to fixed end years; diagonal sections represent constant time windows of different lengths. The heat map confirms previous evidence and highlights two key frames of convergence as well as the substantial divergence in the middle of the period. The highest intensity of negative coefficients is registered for

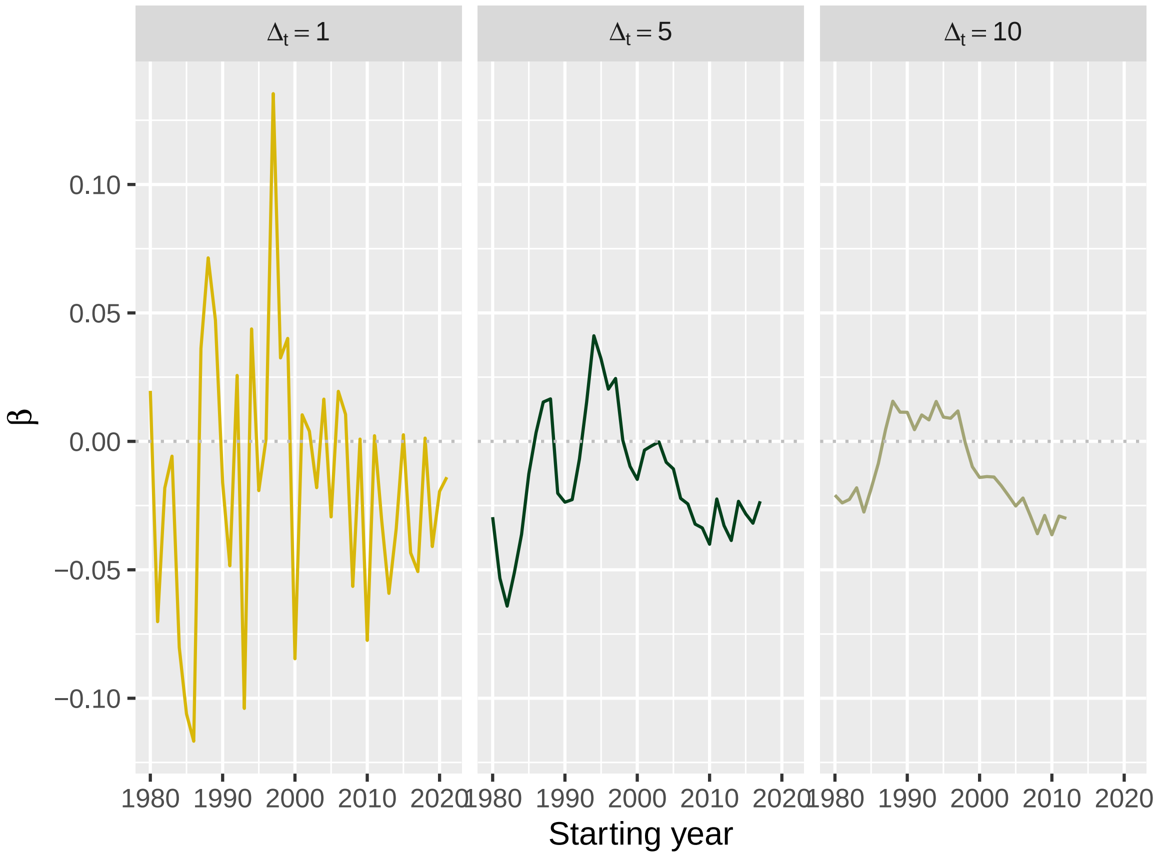


Figure 3: β-convergence coefficient for $\Delta t = 1, 5, 10$ years intervals for the panel of 250 European regions.

different periods in the 1980s and in 2010, peaking at $\hat{\beta} = -0.1167$ for 1986-1987. On the other hand, we observe two clear moments of divergence, and the maximum positive value $\hat{\beta} = 0.1353$ appears in 1997-1998. Values on the hypotenuse represent one-year spikes and are the greatest in absolute terms, as expected; the bursts are often confirmed in the subsequent span but tend to flatten in the long run. Convergence is overall predominant and more sustained over time: stronger in the initial period, slow but persistent throughout the 2000s. Divergence, conversely, appears more concentrated in the short term but dilutes rapidly. Overall, the impression of general bland intensity is corroborated: approximately half of the coefficients lie within a symmetric interval around zero ($[-0.015, 0.015]$). The empirical analysis is mainly coherent with previous literature. Figure 4 highlights the persistence of modest negative coefficients but also reveals the turnover between convergence and divergence, hinting at a global trend distinguished by alternating patterns.

Sigma (σ-) convergence is used to quantify dispersion in the sample (Dalgaard & Vastrup, 2001; Young et al., 2008). There is no restriction as to which measure

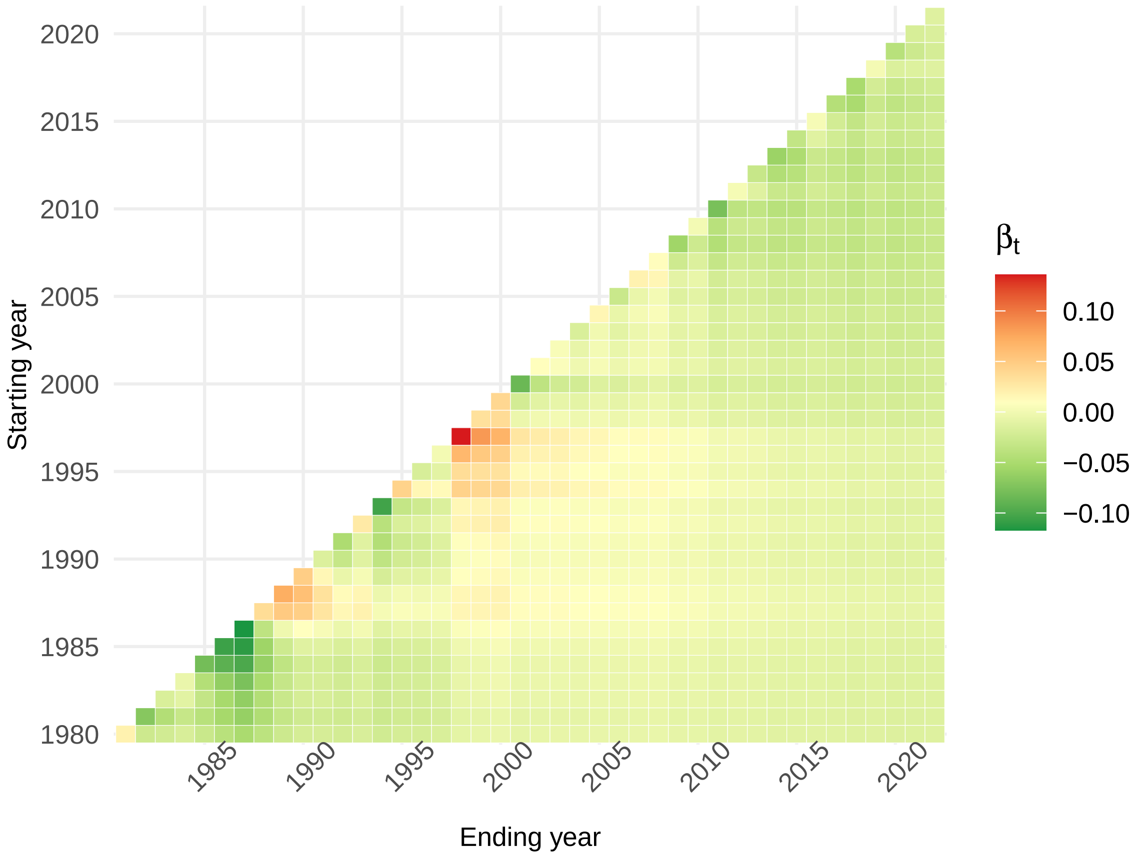


Figure 4: β-convergence coefficient for all possible Δt for the panel of 250 European regions.

to employ: both standard deviation and coefficient of variation are widely accepted and used. Lower dispersion is associated with greater similarity between regions, as more concentration mirrors increasing homogeneity. The trend of yearly standard deviation in normalized per capita GDP is represented in Figure 5. Two periods of σ-convergence emerge: through the late 1980s, and from the early 2010s. After a phase of fluctuations, we notice an upswing in divergence in the late 1990s. Overall, there is lower dispersion than in 1980, reflecting the more compact density. Increasing or decreasing concentration, however, does not reveal any insight on the shape of the distribution (Quah, 1996c). Both σ- and β-convergence present this limitation, as fluctuations in the metrics may depend on several patterns taking place the data. Different methodological techniques are needed to understand this aspect in greater detail and, ultimately, answer the real question of interest. In general, these measures of inequality do not unveil any intra-distribution churning.

In order to account for diversity in steady state levels of structural characteristics, the notion of *conditional convergence* has been proposed and widely investigated by

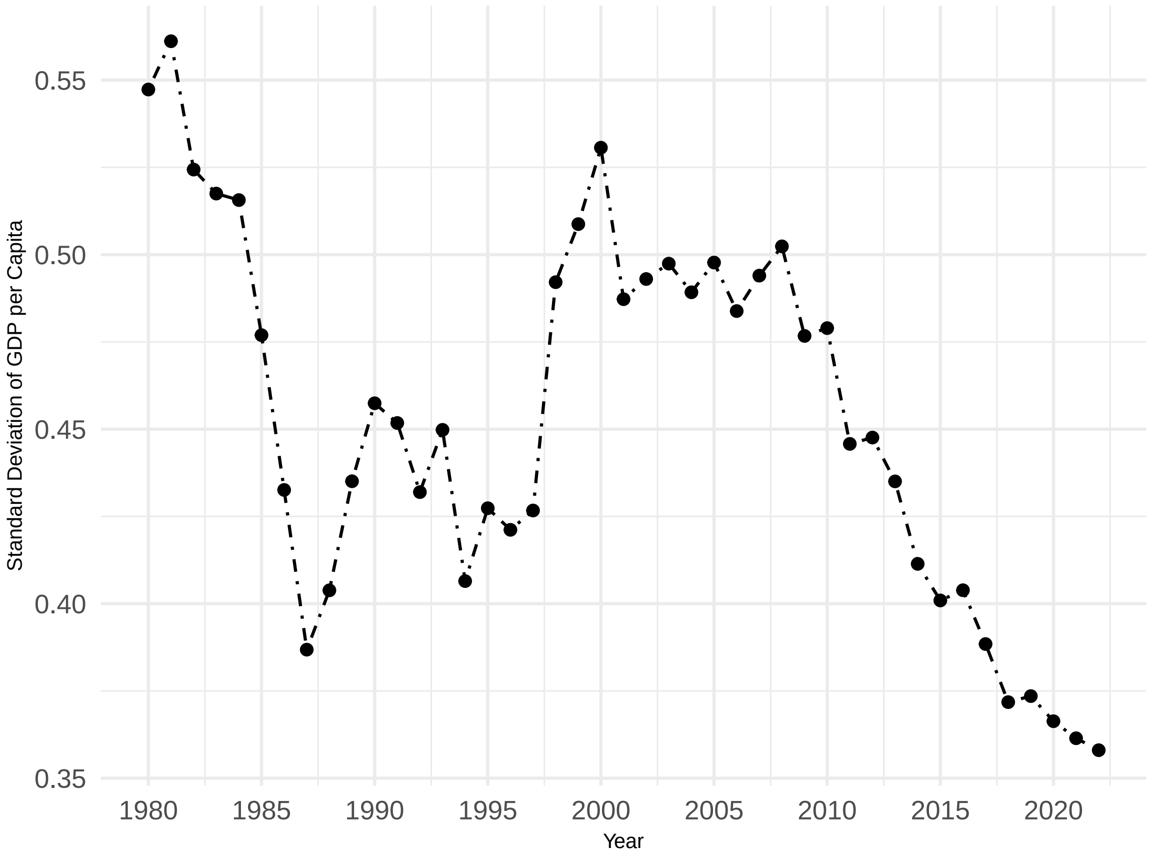


Figure 5: σ-convergence of per capita relative GDP in PPS over the period 1980–2022 for the panel of 250 European regions.

some authors. The extended regression in Equation (2) includes control variables such as human capital levels, investment rates, government policies, technological innovation, and demographic trends. The coefficients γ_j indicate the effect of each conditioning variable on the convergence rate, allowing for a more nuanced understanding of the process; μ_t represents a time-specific effect; ε is the usual error term.

$$\log(GDPpc_{i,t+\Delta t}) - \log(GDPpc_{i,t}) = \beta_t \log(GDPpc_{i,t}) + \sum_{j=1}^{k} \gamma_j X_{i,j,t} + \mu_t + \varepsilon_{i,t} \quad (2)$$

As an econometric tool, Equation 2 introduces a more refined approach to modeling convergence; however, Cho (1996) finds that the coefficient on income per capita is very sensitive to the choice of control variables. Moreover, expecting economic growth from a predetermined steady state without fully addressing the intricacies of catching up efforts oversimplifies the diverse paths economies take. The assumption that countries can achieve parity at the technological frontier overlooks the significant challenges in transferring or acquiring technical knowledge, human capital and resources swiftly and effectively.

Club convergence represents yet a different approach to the topic, aiming to underscore the heterogeneity between countries while avoiding the restrictive effects of conditioning. The hypothesis that different economies obey different linear models proves viable both conceptually and in practice (Chatterji, 1992; Durlauf & Johnson, 1995). Club convergence can therefore be defined as the tendency of units to be distributed in clusters, polarising over basins of attraction. In terms of world income, Quah (1996c) observed a "twin-peaks" (or, plainly, bimodal) conformation. Elsewhere, a greater number of poles has been proposed for different contexts. Phillips and Sul (2009) find evidence for the existence of four convergence clubs and a fifth divergent group; Lee and Lee (2016) suggest increasing similarities and the subsequent broadening of convergence clubs; Bandyopadhyay (2011) explored polarisation trends across different Indian states; Canova (2004) reported four poles of European regions at the NUTS 2 level. This framework allows to explore the transition dynamics and evolving membership that shape the groups, for a more detailed examination of how specific factors –including technological advancements, policy choices, and institutional environments– influence the movements within and across these clubs.

4 Distribution dynamics of EU regions

Absolute β-convergence synthesises information in a single measure, providing punctual evidence of a flowing phenomenon. However, it does not exhaustively reveal whether lower-income regions have been catching up to higher-income ones. Indeed, Johnson and Papageorgiou (2020) insist that "the notion of convergence is only a theoretical construction that characterizes part of the broader dynamic growth process across countries". Furthermore, we ignore the relative dimension: each region should be compared with the others, rather than its own previous performance. The introduction of inequality measures still cannot paint a clear picture of the processes in motion: σ-convergence is similarly unable to recover relative movements of individual economies. Different formal tools are needed. A *model of explicit distribution dynamics* is used recover two features of pressing interest: shape dynamics and intra-distribution mobility.

4.1 A closer look at the shape of the distribution over time

Among others, Anderson (2004), Anderson et al. (2016), and Henderson et al. (2012), have proposed non-parametric or semi-parametric approaches. Kernel density estimation is unconstrained by any specification, allowing a visual exploration and understanding of the data. Finite mixture models are a flexible and more rigorous method to determine the presence of multiple components in the density. In this context, underlying distributions can immediately be related to sub-populations of regions with precise socioeconomic characteristics.

We employ a Gaussian kernel with the Sheather and Jones plug-in bandwidth (Sheather & Jones, 1991). The graphical interpretation represents a crucial starting point to inform the inferential analysis. Figure 6 illustrates the remarkable evolution of the distribution over the forty years span. We immediately notice a shift of the density towards the right, an increased concentration in the main mode and the erosion of the left peak. Finally, the wealthiest regions have been sliding away in recent years. In Figure 7 we overlay the estimated densities for relative income. The distribution in 1980 shows a clear multimodal outline. By 2000, three groups are quite distinctively outlined in the estimated density: the overall structure remains roughly unchanged, mirroring the presence of a low-income cluster to the left of the main group and a pole of high-income regions. In particular, the conglomeration of poor regions having an income lower than 50% of the European average persists. In 2022, however, the left peak is almost completely smoothed out and integrated into the distribution.

Kernel density estimation yields interesting results on the presence of groups within the distribution. Multimodality is coherent with the theory of convergence clubs, although neither a sufficient nor necessary condition for the existence of underlying clusters. We highlight the change in shape of the distribution and the substantial shift towards the right. The visual analysis suggests a successful momentum of catching up and a general increase in mean income over the set of EU regions.

4.2 A semi-parametric approach

Heterogeneity of units grouped together can often be observed in biological, physical, and social sciences. Mixture models combine the flexibility of non parametric approach and the formality of parametric estimation. One significant challenge is that complexity increases with the number of mixture components; to deal with that, maximum likelihood estimation via the expectation-maximisation (EM) algorithm (Dempster et

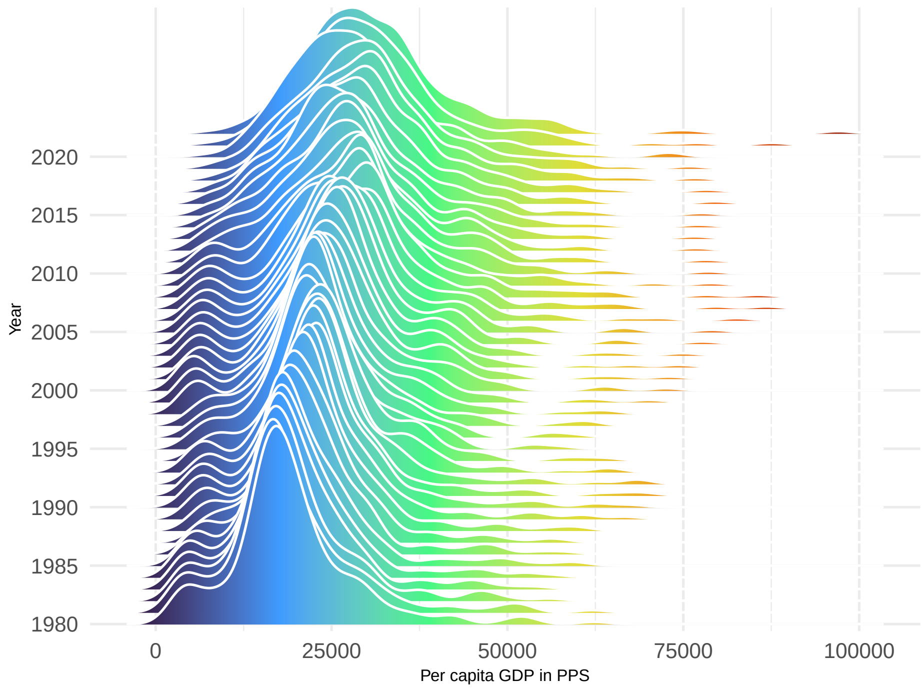


Figure 6: Evolution of the density between 1980 and 2022 for the panel of 250 European regions.

al., 2018) has become the standard procedure (Mclachlan et al., 2019). Modern approaches exploit computationally intensive methods that can rely on multiple software implementations (Chassagnol et al., 2023; Leisch, 2004) and the great popularity of the method is manifested by the ever increasing contributions to the literature (Mclachlan et al., 2019; McLahlan & Peel, 2000; Melnykov & Maitra, 2010, for an overview).

Finite mixtures of univariate Gaussian distributions with unequal variances are among the most popular application when modelling income distributions. Consider the canonical functional form $f_k(x;\boldsymbol{\theta_k}) \equiv \mathcal{N}(x;\mu_k,\sigma_k^2)$ with parameter vector $\boldsymbol{\theta}_k = (\mu_k,\sigma_k^2)$ for each component $k = 1,...,K$. The corresponding mixture density function can be written as:

$$f(x;\boldsymbol{\psi}) = \sum_{k=1}^{K} \pi_k \phi_k(x;\mu_k,\sigma_k^2), \tag{3}$$

where μ_k intuitively represents the group mean per capita income, and σ_k^2 is the associated within-component variation. The *mixing proportions* or *weights* of the mixture $\pi_1,\pi_2,...,\pi_K$ are non negative quantities that sum to one, representing the prior probability that an observation originated from to the corresponding *component density* $f_k(x)$. In our application, the mixing proportions give the prior probability that a territory belongs to a specific sub-population of high-, low- or middle-income regions.

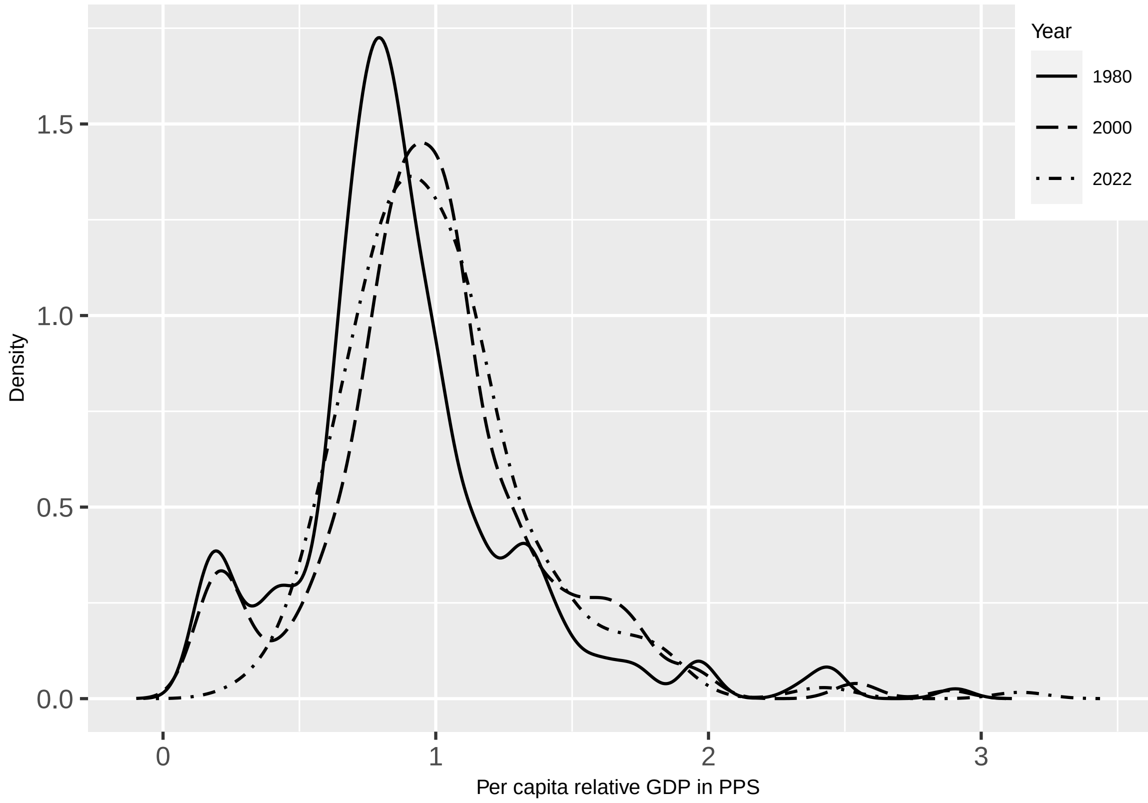


Figure 7: Kernel density estimations of per capita income in the balanced panel of 250 NUTS 2 regions regions at twenty years intervals. We avoid representing 2020, as the consequences of the pandemic crisis are preponderant for that year, and choose 2022 as benchmark instead.

When K is known, only $\boldsymbol{\psi}$ needs to be estimated. In numerous applications, however, the number of components is also unknown and has to be inferred from the data.

The functional form of the of the log-likelihood suggests an iterative procedure fitting ML estimates for the parameter vector to replicate until convergence (according to a specific criterion). The introduction of the EM algorithm (Dempster et al., 2018) guaranteed formal substantiation to the technique, introducing a straightforward method with generally vast applications. In the context of mixtures, the labels associating each observation to the originating component are unobserved (missing). The posterior probability τ_{ik} that observation x_i originated from component k is defined as:

$$\tau_{ik} = \frac{\pi_k \phi_k(x_i; \mu_k, \sigma_k)}{\sum_{h=1}^{K} \pi_h \phi_h(x_i; \mu_h, \sigma_h)}. \tag{4}$$

In the specific case of normal mixture models, some additional considerations can be made. First, the conditional expectation of the complete data log-likelihood can be reduced to the computation of posterior (membership) probabilities. The other elements in the equation are not unknown, and only the update of posterior probabilities is required to finalise the estimation. Additionally, the two sets of parameters in $\boldsymbol{\psi}$ can be maximised separately since the mixing proportions and the Gaussian parameters appear in different linear terms.

The updated posterior probabilities $\tau_{ik}^{(s)}$ given the current estimate for the parameter vector $\boldsymbol{\psi}^{(s-1)}$ ultimately become:

$$\tau_{ik}^{(s)} = \frac{\pi_k^{(s-1)}\phi_k(x_i; \mu_k^{(s-1)}, \sigma_k^{2(s-1)})}{\sum_{h=1}^{K} \pi_h^{(s-1)}\phi_h(x_i; \mu_h^{(s-1)}, \sigma_h^{2(s-1)}))}. \tag{5}$$

The EM algorithm is a local method, which proves notoriously challenging in the context of a multimodal likelihood function. As a consequence, results are strongly sensitive to the choice of initializing values. Several possible procedures have been proposed, but there is no uniformly ideal solution to this matter (Biernacki et al., 2003; Figueiredo & Jain, 2002; Maitra, 2009; Melnykov & Maitra, 2010). [7]

4.2.1 Testing for the number of components

The number of components K is usually to be estimated. The underlying group structure in the population is commonly unknown and often represents the information of interest to be recovered. It must be remembered that the presence of several modes in the distribution does not necessarily imply distinct underlying groups in the population. The mixture could even be unimodal when the components are not sufficiently far apart. There is again no definite consensus over the most reliable procedure for such a task. Parsimony-based methods include Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), but suffer from a main limitation: it is unclear how to evaluate discrepancies in the scores associated to different models. Testing-based method cannot rely on asymptotic properties in this context, hence the distribution of the test statistic under H_0 is estimated by computing bootstrap replicates ($\hat{\theta}^*$). The bootstrapped Likelihood Ratio Test (LRT) relies on the test statistic $-2\log\lambda$ for the hypotheses $H_0 : K = K_0$ vs $H_1 : K = K_1$, where $K_1 > K_0$:

$$-2\log\lambda = 2\{\log L(\hat{\boldsymbol{\psi}}_1) - \log L(\hat{\boldsymbol{\psi}}_0)\}. \tag{6}$$

Tables 1 and 2 report values of AIC, BIC and $\hat{\theta}^*$ for selected years of interest. For 1980, a solution with four component is supported by the LRT and AIC, while the BIC

[7]For a stable solution, multiple random starts, clustering algorithms, two-stages EM (*emEM*) can be employed.

Table 1: Values of AIC and BIC for selected years

	1980		1986		1993		1999		2006		2013		2019		2022	
k	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
1	319.07	326.12	226.9	233.94	226.01	233.05	256.36	263.41	251.13	258.17	223.94	230.98	154.61	161.65	160.42	167.47
2	255.02	272.63	179.85	197.49	178.63	196.24	215.97	233.58	212.56	230.17	207.66	225.27	127.65	145.25	107.4	125.01
3	255.01	272.62	179.82	197.42	178.63	196.24	215.96	233.57	212.56	230.16	213.65	238.35	133.65	161.82	107.4	125.01
4	246.04	274.21	157.3	196.04	184.62	212.71	215.96	233.57	218.56	230.22	210.48	249.04	139.65	178.38	107.4	125.01

Table 2: Values of the bootstrapped LRT statistic and related p-value for selected years

	1980		1986		1993		1999		2006		2013		2019		2022	
k	$\hat{\theta}^*$	p	$\hat{\theta}^*$	p	$\hat{\theta}^*$	p	$\hat{\theta}^*$	p	$\hat{\theta}^*$	p	$\hat{\theta}^*$	p	$\hat{\theta}^*$	p	$\hat{\theta}^*$	p
1	58.34	0	57.42	0	53.38	0	46.4	0	44.57	0	22.29	0.01	32.97	0	59.03	0
2	27.07	0	11.09	0.08	9.96	0.08	29.44	0.01	19.33	0.02	6.36	0.28	8.19	0.09	8.93	0.09
3	15.95	0.04	-	-	-	-	0.9	0.91	9.75	0.14	-	-	-	-	-	-
4	2.85	0.58	-	-	-	-	-	-	-	-	-	-	-	-	-	-

also provide evidence for two or three components. For 1986 and 1993, the LRT rejects the hypothesis of two components at the 5% level of significance but not at the 10% level. For 1993, AIC and BIC support the presence of both two and three components, and are in favour of even a four components solution for 1986 data. Almost equivalent values of BIC are observed for 1999 and 2006 for $k = 2, 3, 4$, while the LRT more decisively suggests a three components mixture. For 2013, evidence from AIC, BIC and LRT consistently points out a two components mixture. For 2019 and 2022, the hypothesis of $k = 2$ is again rejected at the 5% level of significance but not at the 10% level. AIC and BIC refer support for $k = 2$ in 2019 and indicate stronger preference for any non-unimodal solution in 2022.

Model selection assesses the trade off between interpretability and simplicity. Overfitting the data could provide a mixture model that is apparently more adherent to the density but not meaningful for any conclusion. On the contrary, the risk of underfitting is losing out crucial information. The partition should be informative and improve our comprehension of the phenomenon, providing economic significance together with statistical significance: an optimal solution with little practical meaning bears no interest in the applied context (Pittau et al., 2010). Therefore, the final decision is supported by the statistical tests but also strongly informed by the overall knowledge and understanding of the data and their context. Previous work can also help educating the final choice: empirical evidence suggests that the distribution of income in Europe can be approximated by two to four groups (Canova, 2004; López-Bazo et al., 1999; Pittau & Zelli, 2006). Our approach follows a twofold investigation: first, we fix K; then, we allow K to vary.

4.3 Is there evidence of club convergence?

The European Commission identifies "less developed", "transition" and "more developed" regions. A threefold classification appears intuitively natural and coherent with the academic records; it is also widely supported by AIC and BIC. Moreover, the official thresholds proposed by the EC can be used as reference starting points to initialise the EM algorithm. We present a mixture model with $K = 3$ fixed groups spanning from 1980 to 2022. Figure 8 follows the three clusters over the period of interest, representing the evolution of group mean, standard deviation and posterior weights in the mixture. The information summarised in the plot allows to describe the regional scenario and highlight trends in the configuration of groups. The low-income component shows an overall steady increase in mean relative income, marking extraordinary progress from 2010 onwards: we observe a catching up process towards the above group. Encouragingly, this is the less populated group: we can argue for a migration of regions towards the middle component. This cluster is quite consistently stable throughout the decades, exhibiting a slight improvement in its average while keeping constant size. The associated probabilities range from 60% to 90% of units, averaging 75%. The majority of region therefore safely belong to the central group, which is itself an indicator of cohesion, or at least similarity, in the set. The high-income component presents mean income well above the EU average and quite constant size (around 20%). Following the oil crisis, three years present anomalous behaviour: former rich regions "fell" into the second group, isolating a few extremely high-income observations. The 2008 financial crisis hindered growth: regions recovered to prior levels of income but struggled to restore the ascending trend.

Consider now the case of varying number of components in the mixture model.

The fitted mixture density for 1980 consists of four components that appear quite distinct. The first component represents a set of poor regions, defined by a mean value of €$4,235$PPS, accounting for 7% of the population. The middle-income group describes the most relevant share of the population, 74%, a core of regions with mean equal to €$17,384$PPS. Finally, we can identify an upper-middle cluster of regions (7%), with mean income €$28,759$PPS, and the tail of rich territories (12%) characterised by a mean value of €$34,047$PPS. The conditional probabilities τ_{ik} enable to asses the degree of uncertainty according to which regions can be allocated to each component. Most of the assignments are quite well defined by a value greater than 0.75, but partial overlap must be noted (8% of units). The majority of poor regions are in Turkey, while

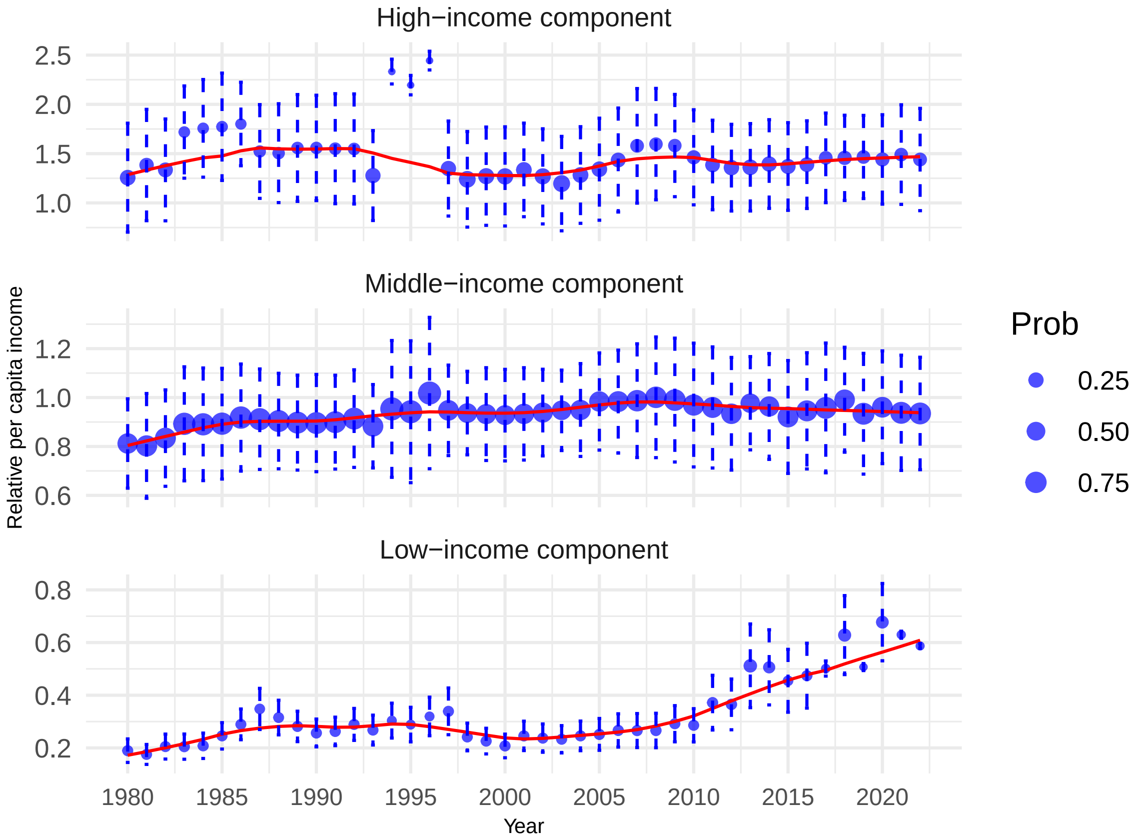
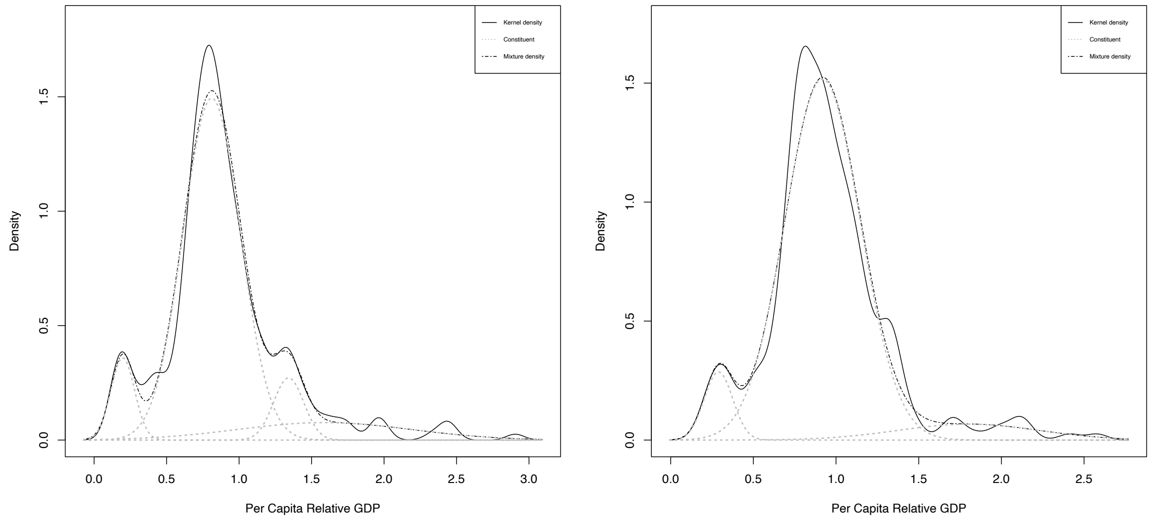


Figure 8: Estimates of mean, standard deviation and mixing proportion for each component of the mixture fit at each time point, Euros in PPS.

the rich group is composed of regions from Switzerland, Northern Italy, Germany, The Netherlands, Finland and a few administrative and service districts (Hamburg, Wien, Bruxelles, Luxembourg). The majority of the set belongs to the middle-income group with very high probability, demonstrating the presence of a solid basin of regions with coherent characteristics.

In 1986, the mixture describes a well defined partition into three latent classes representing low-, middle- and high-income regions. The clusters present low to minimum overlap: only a few units (2%) show uncertain assignment to a component, illustrated by conditional probabilities in the range $0.35 - 0.66$. The high-income group is characterised by greater mean per capita GDP (€$40,485$PPS) and lower dispersion, describing a pool of notably rich domains. The dominant component accounts for 87% of the population and represents wealthier middle-income regions than 1980. Indeed, GDP increased (€$20,645$PPS) as this group started to merge with the previously observed "upper-middle" class.

Results for 1993 confirm again the presence of three components, however charac-



(a) 1980, four components

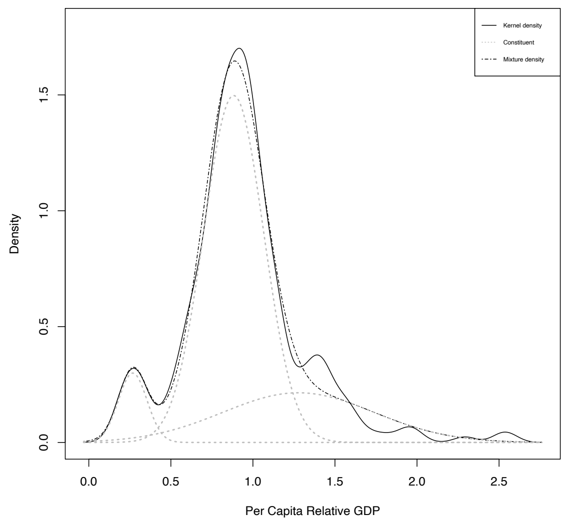
(b) 1986, three components

Figure 9: Kernel density estimation and the mixture model fit.

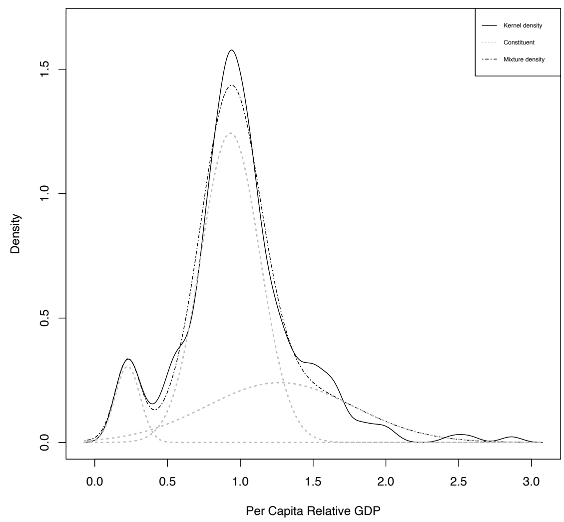
terised by a different structure. The most striking discrepancy with previous years is the conformation of the second and third clusters. The group of high-income begins to enlarge and accounts for 25% of the population, with a mean GDP of €32, 366PPS but great dispersion. For instance, a few regions from Italy, Germany and the UK belong to this component with conditional probability greater than 80%. The main component encompasses 69% of the regions, exhibiting yet again an increase in the mean income of the cluster (€22, 360PPS), coherent with the population average. The conditional probabilities show increased overlap in the assignment of units, mainly between middle and rich groups.

In 1999, the fitted mixture highlights the distinction in three components reflecting the ongoing expansion of the high-income cluster (30% of the population), while the poor regions appear substantially immobile. The distribution for 2006 is also well approximated by three components. The mean per capita GDP is notably higher than the previous reference years for all groups, but the second component regains predominance. The high-income group only accounts for 22% of the regions in the sample, but there is greater overlap than before.

The mixture model exhibits a different behaviour for 2013, where two components fit the distribution appropriately. The most eminent feature is the erosion of the left peak, manifested by the absence of a low-income cluster. We observe a main component accounting for 83% of the population, therefore comprising the vast majority of regions.

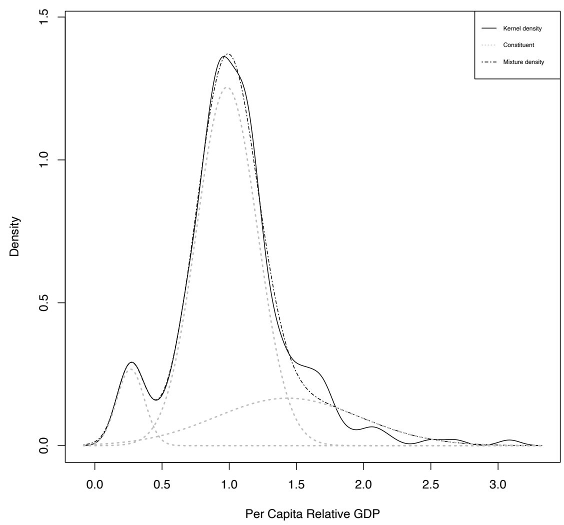


(a) 1993, three components

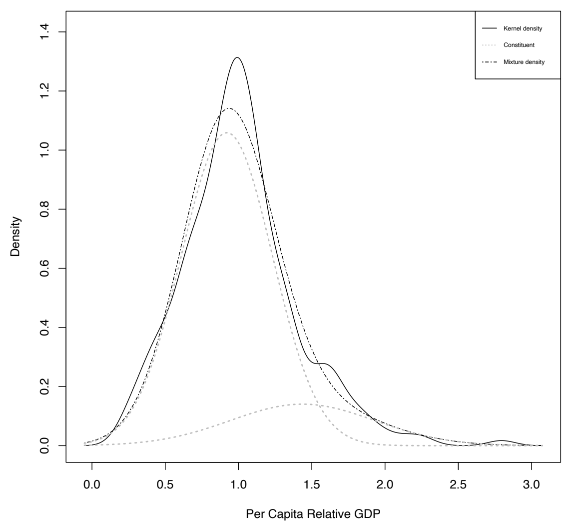


(b) 1999, three components

Figure 10: Kernel density estimation and the mixture model fit.

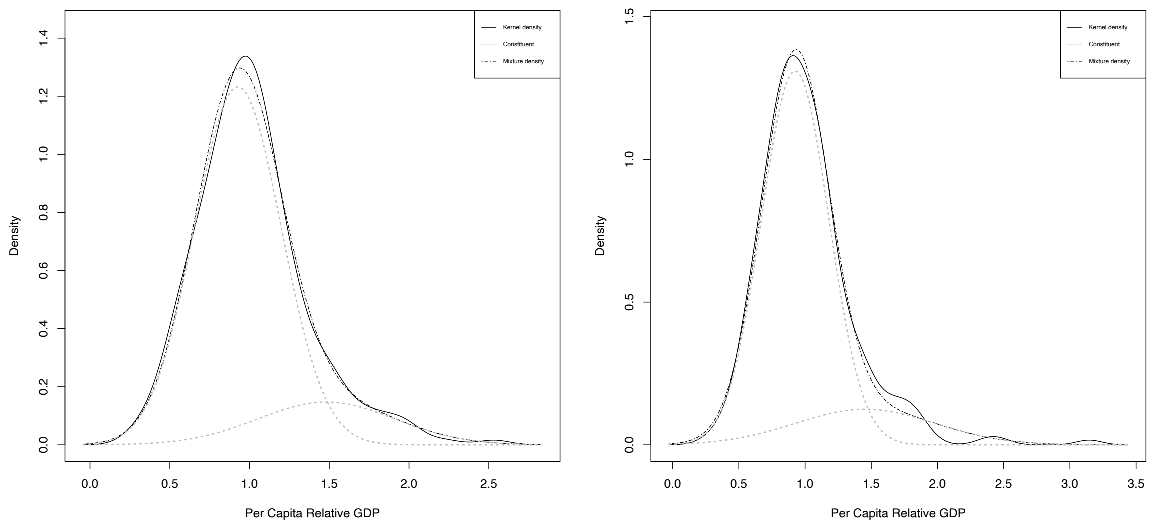


(a) 2006, three components



(b) 2013, two components

Figure 11: Kernel density estimation and the three- or two-component mixture model fit.



(a) 2019, two components

(b) 2022, two components

Figure 12: Kernel density estimation and the two-component mixture model fit.

The mean of this group is 22.1% higher than the main component observed in 1986. The remaining group (17%) is defined by mean GDP equal to €$39,349PPS$ and appears to be slowly drifting away from the rest of the distribution. The two components slightly overlap, as 5% of the units show uncertain assignment. In 2019, only two components can be discerned again into a group of middle and high-income regions, reflecting approximately the same overall structure. An increase in mean GDP can be observed for both groups, 10% and 13% respectively, consistently with the pattern previously recognised. We notice a broader central group (85% of the population), but also highlight the distinguished club of rich regions and their unique composition, revealing a recognisable pattern of divergence from the rest of the sample. Finally, the two components structure is also valid for the 2022 data. The increase in mean values illustrates post-pandemic recovery, while the mixing proportions and overlapping remain essentially unchanged.

Three patterns of interest emerge from the analysis. There is a dominant component observed in the two-, three- and four-components mixtures that attracts the majority of regions into its domain. This central mass assumes a broader aspect throughout the years, encompassing an increasing number of regions and progressively increasing the mean value of per capita GDP in PPS. It corresponds to a wide range of regions with different characteristics, but sharing a common solidity in terms of economic well-being. The 1990s witnessed the approach of a large number of upper middle-income

Table 3: Means, standard deviations (expressed in per capita GDP, PPS) and mixing proportions of the fitted mixture models

Years	Mean				Standard deviation				Mixing proportion			
	k1	*k2*	*k3*	*k4*	*k1*	*k2*	*k3*	*k4*	*k1*	*k2*	*k3*	*k4*
1980	4,235	17,384	28,759	34,047	1,155	4,025	1,801	12,764	0.074	0.74	0.071	0.115
1986	6,508	20,645	40,485	-	1,311	4,923	9,526	-	0.062	0.866	0.072	-
1993	6,782	22,360	32,366	-	1,436	4,317	11,554	-	0.066	0.687	0.248	-
1999	5,530	22,815	31,109	-	1,192	4,641	12,202	-	0.066	0.631	0.303	-
2006	7,130	26,313	38,406	-	1,691	5,632	14,117	-	0.069	0.709	0.222	-
2013	25,211	39,349	-	-	8,142	13,247	-	-	0.827	0.173	-	-
2019	27,762	44,392	-	-	7,592	12,530	-	-	0.841	0.159	-	-
2022	28,538	44,859	-	-	7,267	15,918	-	-	0.836	0.164	-	-

regions to the right-tail club, affecting the relevance of this cluster in terms of group membership probabilities. Successive decades, however, confirm the tendency for the richest regions to diverge and constitute their own league. The left peak of the initial distribution disappears in the 2010s, consistently with a process of catching up of the poorest regions to the middle-income group. This empirical evidence suggests a positive impact of cohesion policies and regional integration on levelling per capita incomes across Europe. Conversely, it confirms the presence of persistently wealthier regions, often corresponding to metropolitan conglomerations of service centers.

The mixture model provides a further layer of interpretation to the analysis of convergence presented above. A few authors report that in the EU the process stopped in the 1980s, consistently with the pattern observed from the data. Considering 5- and 10-years time frames, initial convergence made way to a period of slight divergence over the 1990s. The fitted models generally reflect this behaviour in the persisting distance between poor and rich regions. The starting point of the distributional analysis was a four components mixture, characterised by fairly distinct poles. This conformation mutates quite suddenly when the fitted model displays a three components structure, indicating a reasonably stable partition into a classic low-, middle- and high-income group conformation. From the empirical results, we therefore observe a mostly unchanged distribution for two decades which does not necessarily contradict the β convergence analysis. The divergence process could be reflected in the increasing distance between the unchanging groups of poor and middle-income regions and the enlarging third component, departing from the rest of the population. A new process of convergence apparently sprung in the 2000s, peaking in the 2010s and gradually slowing down by the beginning of the 2020s, cementing a new equilibrium. A distinctive mutation of the mixture fit is observed when the two components represent a comprehensive cluster of middle-income regions and the relatively small club of high-income regions. This outline is consolidated in the following decade and still appears vividly distinguishable in the most recent data.

5 Conclusions

The aim of social, economic and territorial cohesion is embedded in the foundation of the European Union. The EU has been dubbed a "convergence machine" and the economic growth observed across the whole territory mirrors the longstanding political commitment to regional integration. Convergence of new member countries towards the older members has been observed in the literature, highlighting that entering the EU has a consistently positive effect on income growth. While we cannot ascribe all developments to policy alone, the facilitating role in creating a supporting environment for these countries to operate and evolve their economies must not be overlooked.

We assessed the evolution of income distribution across forty years for a selected panel of regions. The classical absolute β-convergence analysis reported mild indications of overall catching up, with specific fluctuations observed across different periods. Dispersion measured through σ-convergence also presents substantially lower dispersion in recent years when compared to the 1980s. Our results are coherent with the references, showing a general trend of cohesion that might not be outstanding in magnitude but is surely still relevant. However, these measures are not truly able to recover distributional dynamics in terms of shape, location and internal mobility of regional economies. We complement the traditional literature on β- and σ- convergence by introducing the distributional analysis. The main tool of interest is the Gaussian mixture model, a semi-parametric method with clustering properties to recover unobserved heterogeneity in the data.

Following the literature and reference official work from the European Commission, we initially explore cross-sectional partitions of the distribution into three groups. Intuitively, these represent the *less developed, transition* and *more developed* regions, respectively. In facts, the density can be approximated by a variable number of components: K becomes a parameter to be estimated as well. This approach provides greater flexibility in modelling the shape of the distribution. We find a four components structure in 1980 that evolves into three components throughout the following decades. The early 2010s mark another significant shift, in that only two clusters are clearly discernible: the group of lower-income regions is smoothly integrated into the middle cluster. Overall, the "middle-income" group increased mean per capita GDP and grew to comprise 83% of units in 2022. The club of wealthy regions represents an autonomous and persistent bundle whose extremely high-income stems directly from the agglomeration of financial, administrative and logistic services. We conclude that a process of catching up has actually been in place, manifested by the disappearance of consistently poor regions and the consequent enlargement of the middle-income group. On the other hand, we cannot overlook the presence of a more prosperous pole drifting away from the rest of the distribution. The key takeaway is the existence of three pivotal junctures, corresponding to phases of partial convergence in a scenario of otherwise fundamental stability. In the 1980s, three well-defined clusters emerge; in the 1990s, wealthy regions start to diverge (a potential polarisation trend); in the 2010s, poor regions are incorporated in the middle-income class.

References

Acemoglu, D., Johnson, S., & Robinson, J. A. (2005). Chapter 6: Institutions as a fundamental cause of long-run growth. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of economic growth* (pp. 385–472, Vol. 1). Elsevier. https://doi.org/https://doi.org/10.1016/S1574-0684(05)01006-3

Anderson, G. (2004). Making inferences about the polarization, welfare and poverty of nations: A study of 101 countries 1970–1995. *Journal of Applied Econometrics*, *19*(5), 537–550.

Anderson, G., Pittau, M. G., & Zelli, R. (2016). Assessing the convergence and mobility of nations without artificially specified class boundaries. *Journal of Economic Growth*, *21*, 283–304.

Bandyopadhyay, S. (2011). Rich states, poor states: Convergence and polarisation in india. *Scottish Journal of Political Economy*, *58*(3), 414–436.

Barca, F. (2009). *An agenda for a reformed cohesion policy a place-based approach to meeting european union challenges and expectations* (tech. rep.). Economics and Econometrics Research Institute (EERI), Brussels.

Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *The Quarterly Journal of Economics*, *106*(2), 407–443. https://doi.org/https://doi.org/10.2307/2937943

Barro, R. J., & Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy*, *100*(2), 223–251. https://doi.org/10.1086/261816

Becker, S. O., Egger, P. H., & von Ehrlich, M. (2018). Effects of EU Regional Policy: 1989-2013. *Regional Science and Urban Economics*, *69*, 143–152. https://doi.org/10.1016/j.regsciurbeco.2017.12.001

Biernacki, C., Celeux, G., & Govaert, G. Ã. E. (2003). Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models. *Computational Statistics & Data Analysis*, *41*, 561–575. www.elsevier.com/locate/csda

Boldrin, M., & Canova, F. (2001). Inequality and convergence in Europe's regions: reconsidering European regional policies. *Economic Policy*, *16*, 206–253. https://doi.org/10.1111/1468-0327.00074

Canova, F. (2004). Testing for convergence clubs in income per capita: A predictive density approach. *International Economic Review*, *45*(1), 49–77. https://doi.org/10.1111/j.1468-2354.2004.00117.x

Chassagnol, B., Bichat, A., Boudjeniba, C., Wuillemin, P.-H., Guedj, M., Gohel, D., Nuel, G., & Becht, E. (2023). Gaussian Mixture Models in R. *The R Journal*, *15*, 56–76. https://journal.r-project.org/articles/RJ-2023-043/RJ-2023-043.pdf

Chatterji, M. (1992). Convergence clubs and endogenous growth. *Oxford Review of Economic Policy*, *8*, 57–69. https://doi.org/10.1093/oxrep/8.4.57

Cho, D. (1996). An Alternative Interpretation of Conditional Convergence Results. *Journal of Money, Credit and Banking*, *28*(4), 669–681. Retrieved April 6, 2024, from http://www.jstor.org/stable/2078077

Dalgaard, C.-J., & Vastrup, J. (2001). On the measurement of σ-convergence. *Economics Letters*, *70*(2), 283–287.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (2018). Maximum Likelihood from Incomplete Data Via the EM Algorithm. *Journal of the Royal Statistical Society:*

Series B (Methodological), *39*(1), 1–22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x
Durlauf, S. N., & Johnson, P. A. (1995). Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics*, *10*(4), 365–384. https://doi.org/https://doi.org/10.1002/jae.3950100404
European Commission. (2024a). Türkiye - financial assistance under ipa. %5Curl%7Bhttps://neighbourhood-enlargement.ec.europa.eu/enlargement-policy/overview-instrument-pre-accession-assistance/turkiye-financial-assistance-under-ipa_en%7D
European Commission. (2024b, March). Ninth report on economic, social and territorial cohesion. https://doi.org/10.2776/585966
European Investment Bank. (2022). Regional cohesion in europe 2021–2022: Evidence from the EIB investment survey. https://doi.org/doi:10.2867/09399
Eurostat. (2024a). *GDP per capita in EU regions, Data from ARDECO Explorer tool* [Accessed on 27 March 2024]. https://urban.jrc.ec.europa.eu/ardeco/explorer?lng=en
Eurostat. (2024b, February). *Purchasing power parities. Reference Metadata in Euro SDMX Metadata Structure (ESMS)* [Accessed on 27 March 2024]. https://ec.europa.eu/eurostat/cache/metadata/en/reg_eco10_esms.htm
Figueiredo, M. A. T., & Jain, A. K. (2002). Unsupervised learning of finite mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*, 381–396. https://doi.org/10.1109/34.990138
Fischer, M. M., & Stumpner, P. (2008). Income distribution dynamics and cross-region convergence in Europe. *Journal of Geographical Systems*, *10*, 109–139. https://doi.org/10.1007/s10109-008-0060-x
Gill, I. S., & Raiser, M. (2012). *Golden growth: Restoring the lustre of the european economic model*. World Bank Publications.
Henderson, D. J., Papageorgiou, C., & Parmeter, C. F. (2012). Growth empirics without parameters. *The Economic Journal*, *122*(559), 125–154.
Islam, N. (1995). Growth empirics: A panel data approach. *The Quarterly Journal of Economics*, *110*, 1127–1170. https://doi.org/10.2307/2946651
Johnson, P., & Papageorgiou, C. (2020). What remains of cross-country convergence? *Journal of Economic Literature*, *58*(1), 129–75. https://doi.org/10.1257/jel.20181207
Kremer, M., Willis, J., & You, Y. (2022). Converging to Convergence. *NBER Macroeconomics Annual*, *36*, 337–412. https://doi.org/10.1086/718672
Lee, J.-W., & Lee, H. (2016). Human capital in the long run. *Journal of Development Economics*, *122*, 147–169. https://doi.org/https://doi.org/10.1016/j.jdeveco.2016.05.006
Leisch, F. (2004). Flexmix: A general framework for finite mixture models and latent glass regression in r. *Journal of Statistical Software*, *11*(8), 1–18. https://ro.uow.edu.au/buspapers/487/
López-Bazo, E., Vayá, E., Mora, A. J., & Ach, J. S. (1999). Regional economic dynamics and convergence in the European Union. *The Annals of Regional Science*, *33*, 343–370. https://doi.org/https://doi.org/10.1007/s00168005010

Maitra, R. (2009). Initializing partition-optimization algorithms. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *6*, 144–157. https://doi.org/10.1109/TCBB.2007.70244

Mclachlan, G. J., Lee, S. X., & Rathnayake, S. I. (2019). Finite mixture models. *Annual Review of Statistics and Its Application*. https://doi.org/10.1146/annurev-statistics

McLahlan, G., & Peel, D. (2000). *Finite mixture models*. John Wiley; Sons. https://doi.org/10.1002/0471721182

Melnykov, V., & Maitra, R. (2010). Finite mixture models and model-based clustering. *Statistics Surveys*, *4*, 80–116. https://doi.org/10.1214/09-SS053

Merler, S. (2016). *Income convergence during the crisis: Did eu funds provide a buffer?* (Bruegel Working Paper No. 2016/06). Bruegel. Brussels. https://hdl.handle.net/10419/173142

Milanovic, B. (2002). Worlds apart: Inter-national and world inequality 1950-2000. http://www.%20worldbank.%20org/research/inequality/pdf/Maksense4.%20pdf

Monfort, P. (2020). *Convergence of EU regions redux: Recent trends in regional disparities*. http://www.europa.eu

Nagy, S. G., & Šiljak, D. (2022). Is the European Union still a convergence machine? *Acta Oeconomica*, *72*, 47–63. https://doi.org/10.1556/032.2022.00003

Paci, R. (1997). More similar and less equal: economic growth in the European regions. *Review of World Economics*, *133*, 609–634. https://doi.org/10.1007/BF02707405

Phillips, P. C. B., & Sul, D. (2009). Economic transition and growth. *Journal of Applied Econometrics*, *24*(7), 1153–1185. https://doi.org/https://doi.org/10.1002/jae.1080

Pittau, M. G., & Zelli, R. (2006). Empirical evidence of income dynamics across EU regions. *Journal of Applied Econometrics*, *21*, 605–628. https://doi.org/10.1002/jae.855

Pittau, M. G., Zelli, R., & Johnson, P. A. (2010). Mixture models, convergence clubs, and polarization. *Review of Income and Wealth*, *56*(1), 102–122. https://doi.org/https://doi.org/10.1111/j.1475-4991.2009.00365.x

Quah, D. T. (1996a). Empirics for economic growth and convergence. *European Economic Review*, *40*, 1353–1375. https://doi.org/https://doi.org/10.1016/0014-2921(95)00051-8

Quah, D. T. (1996b). Regional convergence clusters across europe. *European Economic Review*, *40*(3-5), 951–958. http://www.worldbank.org/html/emc/Welcome,html

Quah, D. T. (1996c). Twin peaks: Growth and convergence in models of distribution dynamics. *The Economic Journal*, *106*, 1045–1055. https://doi.org/10.2307/2235377

Sheather, S. J., & Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society: Series B (Methodological)*, *53*(3), 683–690. https://www.jstor.org/stable/2345597

Young, A. T., Higgins, M. J., & Levy, D. (2008). Sigma convergence versus beta convergence: Evidence from us county-level data. *Journal of Money, Credit and Banking*, *40*(5), 1083–1093.

Appendix

Table A1: Summary of regional denominations for each country represented in the data. Note: * Non-administrative units

Country code	Country name	NUTS 2	n
AT	Austria	Länder	9
BE	Belgium	Provincies/ Provinces	11
CH	Switzerland	Grossregionen/ Grandes régions/ Grandi regioni	7
DE	Germany	Regierungs-bezirke*	38
DK	Denmark	Regioner	5
EL	Greece	Periferies	13
ES	Spain	Comunidades Autónomas + Ciudades Autónomas	19
FI	Finland	Suuralueet/ Storområden*	5
FR	France	Régions + Collectivités territoriales*	27
IE	Ireland	Regions*	3
IT	Italy	Regioni	21
LU	Luxembourg	-	1
NL	The Netherlands	Provincies	12
NO	Norway	Landsdeler	7
PT	Portugal	Grupos de Entidades Intermunicipais + Regiões Autónomas*	9
SE	Sweden	Riksområden*	8
TR	Türkiye	Alt bölgeler	26
UK	United Kingdom	Regions	41

Ultimi Contributi di Ricerca CRENoS

I Paper sono disponibili in: http://www.crenos.unica.it

- **25/06** *Luciano Mauro, Francesco Pigliaru, Gaetano Carmeci,* "Government Size, Civic Capital and Economic Performance: An O-ring approach"
- **25/05** *Fabio Cerina, Luca G. Deidda, Simone Nobili,* "Skill-biased remote work and incentives"
- **25/04** *Federico Atzori, Gabriele Ballicu, Vittorio Pelligra,* "Under Under Big Gods' Eyes. The Impact of Religious Beliefs on Altruistic Punishment During COVID-19 Pandemic"
- **25/03** *Fabio Angei,* "Silent alarms: workplace injuries under-reporting in Italy"
- **25/02** *Silvia Balia, Rinaldo Brau, Sara Pau,* "One plus one makes less than two? Consolidation policies and mortality in the Italian NHS"
- **25/01** *Vania Licio, Anna Maria Pinna,* "Accessibility across Italy: A grid cell approach"
- **24/26** *Carlo Caporali, Federico Crudu, Claudio Detotto,* "Terror and Turnout: The Impact of Terrorism During the Years of Lead"
- **24/25** *Andrea Caria, Adriana Di Liberto, Sara Pau,* "Remote learning during the Covid-19 pandemic: evidence from a three-level survey of Italian schools"
- **24/24** *Emanuela Marrocu, Raffaele Paci,* "The influence of regional contextual factors on public university efficiency in Italy"
- **24/23** *Leonardo Becchetti, Irene Maria Buso, Luca Corazzini, Vittorio Pelligra,* "The taste for Generativity"
- **24/22** *Giommaria Spano,* "Financial frictions and market power accumulation"
- **24/21** *Giommaria Spano,* "Common ownership and uncertainty"
- **24/20** *Luca Deidda, Ettore Panetti,* "Bank Recovery and Resolution Planning, Liquidity Management and Fragility"
- **24/19** *Steven B. Caudill, Laura Ciucci, Claudio Detotto,* "Bargaining power: Do individuals outperform groups?"
- **24/18** *Marta Meleddu, Dimitri Paolini, Valeria Posadas,* "Determinants of Tourists' Expenditure: Systematic Literature Review, Meta-analysis, and Econometric Analysis"
- **24/17** *Pasqualina Arca, Gianfranco Atzeni and Luca Deidda,* "Collateral requirements, cost of credit, and firms' discouragement from applying for bank loans"
- **24/16** *Giampiero M. Gallo, Ostap Okhrin, Giuseppe Storti,* "Dynamic tail risk forecasting: what do realized skewness and kurtosis add?"
- **24/15** *Stefano Borgioli, Giampiero M. Gallo, Chiara Ongari,* "Financial Returns, Sentiment and Market Volatility: a Dynamic Assessment"
- **24/14** *Vincenzo Candila, Oguzhan Cepni, Giampiero M. Gallo, Rangan Gupta,* "Influence of Local and Global Economic Policy Uncertainty on the volatility of US state-level equity returns: Evidence from a GARCH-MIDAS approach with Shrinkage and Cluster Analysis"
- **24/13** *Ester Rita Lucia Sanna, Marta Meleddu,* "Integrating justice and ecological economics: a theoretical framework and indicator toolkit for analysing conflict in protected areas"
- **24/12** *Federico Aresu, Emanuela Marrocu, Raffaele Paci,* "EU funds and TFP growth: how the impact changed over time and space"
- **24/11** *Ivan Etzo, Raffaele Paci, Cristian Usala,* "Brain gain vs. brain drain. The effects of universities' mobile students on territorial inequalities"
- **24/10** *Giommaria Spano,* "Financial frictions, common ownership and firms' market power in a general equilibrium model"

www.crenos.unica.it

ISBN 978-88-6851-576-8

9 788868 515768