The evolution of regional unemployment in the EU
An analysis via the Gompertz diffusion process

(Preliminary version)

XXXIII Annual Scientific Conference AISRe
Roma, 13-15 September 2012
Session:
Comparing European Labour Markets in Crisis Time:
A Regional Perspective

Stefano M. Iacus
Department of Economics, Management and Quantitative Methods
University of Milan, Italy

Giuseppe Porro
Department of Law, Economics and Culture
University of Insubria, Italy

July 27, 2012

Abstract

At the end of Nineties, Danny Quah devoted several papers to the analysis of polarization and stratification in the convergence processes of economies, creating the image of the "convergence clubs" and suggesting the importance of studying the distribution dynamics of the macroeconomic variables. As for the labor markets, Overman and Puga (2002) showed that a progressive polarization of unemployment was in fact occurring among the European regions in 1986-1996, causing a phenomenon of cross-border clusterization of the European regions. Here we propose to analyze the evolution of the unemployment rates of the EU27 regions in the last two decades assuming that the unemployment rates evolve according to a Gompertz stochastic process. The estimated parameters of the process - intrinsic growth rate, deceleration factor, volatility - represent the evolutionary path of the unemployment rate and allow for estimating the steady state of the process. A cluster analysis is performed on the steady state values of the unemployment rates. The analysis confirms the emergence of several "convergence clubs" among the European regional labor markets, which are compared to the cluster resulting from the more traditional clusterization on the current unemployment rates.

JEL Classification: R11; R23; C20

Keywords: European regions; regional labor markets; convergence; Gompertz process

1 Introduction

During the years of progressive enlargement of the European Union (EU), great attention has been devoted to the changes in the distribution of unemployment across the EU regions: are the regions where unemployment is high benefitting from increasing competition and reduction of economic barriers or are inequalities widening over time? From the study by Overman and Puga (2002), concerning the decade 1986-1996, a process of polarization of unemployment emerges: while the disparities among regions often increase within the countries, new transnational macro-regions take shape, where the unemployment rates seem to converge towards similar values. This evidence gives renewed importance to the analysis of regional inequalities and the processes of divergence/convergence among the EU regions.

In the literature on convergence of economic systems the phenomenon of polarization had been previously described in several papers by Quah (1996a, 1996b, 1997), where the inadequacies of the traditional measures of convergence are examined with respect to the evolution of per capita GDP among countries and regional economies. In particular, Quah (1996a) has pointed out that the standard empirical measures - σ-convergence and β-convergence (Barro and Sala-i-Martin 1992, 1995; Mankiw et al. 1992; Sala-i-Martin 1996) - are essentially uninformative when the distribution of the variable of interest (GDP per capita) tends to polarization (two extreme clusters form and the middle class vanishes) or stratification (more than two clusters form). In these cases - the author suggests - an empirical model for the evolution of the variable should be provided, which is able to capture the multimodal aspects of the distribution. Moreover, most of the studies on convergence suffer from two main (and connected) limitations: a) they make use of the current values of the variables of interest (GDP per capita, unemployment rates), while convergence is basically a long-run matter; b) standard measures of convergence do not distinguish between transitory and permanent components: therefore the estimated convergence is affected also by contingent changes of the variables.

Following this suggestion, in this paper we investigate the emerging of unemployment clusters among the EU regions in the last decade, representing the evolution of the unemployment rates in EU countries and regions as a Gompertz stochastic diffusion process (Capocelli and Ricciardi, 1974; Ricciardi, 1977, Gutiérrez-Jáimez et al., 2009). The Gompertz process describes the change in the unemployment rate as a continuous time process and is characterized by three parameters (intrinsic growth rate, deceleration factor, volatility), that describe the features of the path (is the variable stucturally increasing or decreasing? Is it converging slowly or rapidly to the steady state? Is it hit by frequent shocks?) and allow for estimating the steady state value. Therefore, clustering the European regions on their steady state unemployment rates means grouping the regions that exhibit a similar evolutionary process of unemployment (not simply a similar current value of the variable). In other words, we offer a definition of the "convergence clubs" defined by Quah (1997) which is both alternative and complementary to the traditional convergence

analysis.

In order to make a comparison between the two approaches, we examine the evolution of the unemployment rates in the EU27 regions over the period 1991-2008, providing both kind of clustering: one based on the current unemployment rates in 2008 and one based on the steady state values of the Gompertz process. The first approach - as in the study by Overman and Puga (2002) - make some "clubs" emerge, where the unemployment assumes a similar incidence. These groups show that the distribution of the unemployment is re-designing the inner borders of the EU, creating transnational macro-regions which are more similar - according to the unemployment profiles - than the different countries they belong to. The second approach identifies macro-regions that are following a similar path to the steady state unemployment rate: their distribution over the continent gives some additional suggestions about the long-run redistribution of unemployment in Europe.

The paper is organized as follows: in the next section some details about the characteristics of a Gompertz process are provided. Section 3 develops the analysis of the EU regional unemployment in 1991-2008, while section 4 explains how the estimated Gompertz parameters can help in interpreting the different regional dynamics; a measue of the convergence speed is also proposed. Section 5 contains a brief discussion on the interpretation and use of a Gompertz process. Final remarks are in section 6.

2 The Gompertz model

A diffusion process à la Gompertz shows the evolution of a variable as a stochastic S-shaped curve with a stationary state and a log-normal limit distribution. It has been introduced by Capocelli and Ricciardi (1974) and Ricciardi (1977) as suitable for representing demographic trends such as population growth, but it has been used to describe evolutionary phenomena in different fields, from biology to economics (Gutiérrez et al. 2005; Gutiérrez et al. 2006; Gutiérrez-Jáimez et al. 2007; Gutiérrez et al. 2009), because of its flexibility: the process is in fact characterized by a few parameters whose values can draw a wide range of convergence behaviors in continuous time.

Let us denote by $X(t) = X_t$ the unemployment rate at time $t \geq 0$. This variable evolves continuously with time but we observe it only at given times t_i. We assume X_t to have a stochastic dynamical behavior with a long-run stationary state. In particular, we assume X_t can be represented by a Gompertz diffusion model, which is a markovian process solution to the following stochastic differential equation:

$$\mathrm{d}X_t = (\alpha X_t - \beta X_t \log(X_t))\mathrm{d}t + \sigma X_t \mathrm{d}W_t \tag{1}$$

with some initial value $X_0 = x_0$, where $\{W_t, t \geq 0\}$ is a standard Brownian motion (see Iacus,

2008); $\sigma > 0$ is the *volatility* parameter and α and β represent respectively the *intrinsic growth rate* and the *deceleration factor*. When $\beta > 0$, the process is always non negative, i.e. $X_t \in [0, \infty)$ for all t, and as $t \to \infty$, the process X_t is stationary with limiting distribution

$$f(x) = \left(\frac{\pi\sigma^2}{\beta}\right)^{-\frac{1}{2}} x^{-1} \exp\left\{-\frac{\beta}{\sigma^2}\left(\log(x) - \frac{\alpha - \sigma^2/2}{\beta}\right)^2\right\}$$

i.e. the stationary distribution of X_∞ is a log-normal distribution with parameters $\mathcal{LN}((\alpha - \sigma^2/2)/\beta, \sigma^2/2\beta)$.

2.1 The S-shaped trend of the process

The mean of the Gompertz process, i.e. the expected value of X_t, is given by the following formula:

$$\mathrm{E}(X_t) = \exp\left\{e^{-\beta t}\log(x_0) + \frac{\alpha - \sigma^2/2}{\beta}\left(1 - e^{-\beta t}\right) + \frac{\sigma^2}{4\beta}\left(1 - e^{-2\beta t}\right)\right\}, \tag{2}$$

for $t \geq 0$. The S-shaped behavior of the trend of the Gompertz process is regulated by the following two parameters:

- α is the *intrinsic growth rate* of the process, because it indicates at which speed the process would go to infinity in the absence of any attrition;
- β indicates the *deceleration factor* or attrition factor: it decelerates the process, leading it to its stationary value.

The S-shape curve (2) has a point of inflection which is not symmetric, contrary to a standard logistic growth curve. Figure 1 shows several examples of the deformation of the Gompertz trend (2) as a function of α and β.

The parameter σ, on the other side, measures the *volatility* of the process around the trend. When $\sigma = 0$, i.e. no stochasticity is involved, the solution $X_t = x(t)$ to (1) is a deterministic dynamical system $x(t)$ and (2) becomes the Gompertz growth function

$$x(t) = \exp\left\{\frac{\alpha}{\beta} - \left(\frac{\alpha}{\beta} - \log(x_0)\right) e^{-\beta t}\right\}$$

and the inflection point is at $\bar{t} = \frac{1}{\beta}\log(\alpha/\beta - \log(x_0))$ and $x(\bar{t}) = e^{\frac{\alpha}{\beta}-1}$ (see Vieira and Hoffmann (1977) and Ricciardi (1997)).

Looking at formula (2) it is possible to see that the long-run mean of X_t, i.e. the mean of X_∞, is given by

$$\mathrm{E}(X_\infty) = e^{\frac{\alpha}{\beta} - \frac{\sigma^2}{4\beta}}. \tag{3}$$

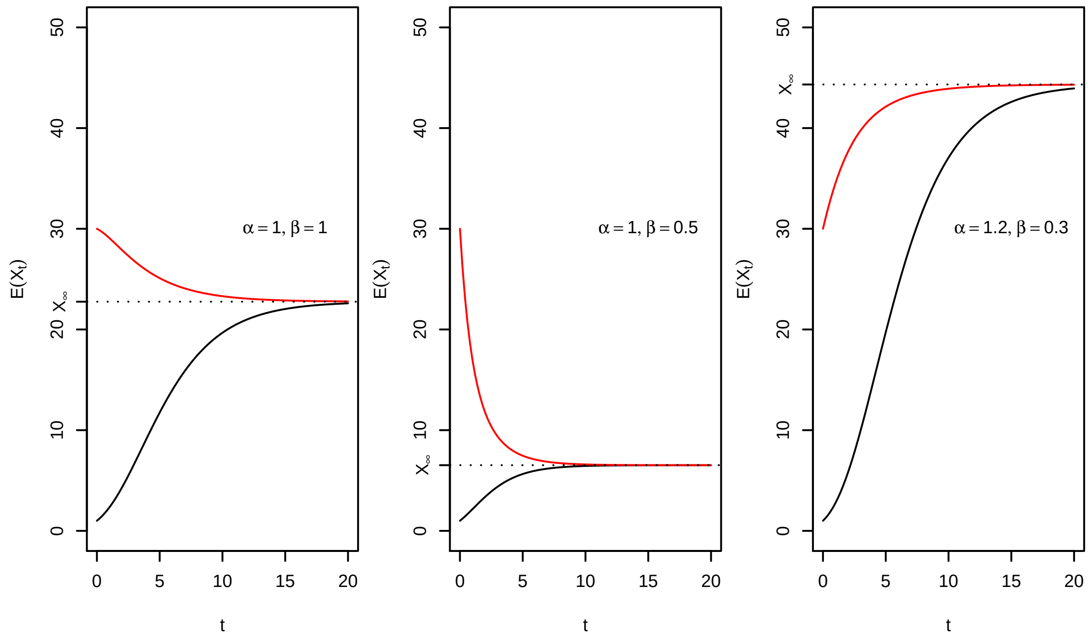

Figure 1: The behaviour of $E(X_t)$ as a function of α, β and initial value x_0 ($x_0 = 1$ or $x_0 = 30$). The dotted line represents the long run steady state value $X_\infty = \exp\left(\frac{\alpha}{\beta} - \frac{\sigma^2}{4\beta}\right)$. In all plots $\sigma = 0.5$.

This limiting mean value is an increasing function of α and decreasing in σ. If β increases, the steady state value decreases (increases) when $4\alpha > (<)\sigma^2$.

Non-linear combinations of α and β and σ define both the long run steady state of the process and the convergence speed of $\mathrm{E}(X_t)$ to the stationary value $\mathrm{E}(X_\infty)$. Virtually, any kind of convergence path (linear, exponential, increasing or decreasing, etc.) to the long run limiting trend $\mathrm{E}(X_\infty)$, can be described by a Gompertz process with appropriate values of α, β and σ. See, for instance, Figure 1: the two processes in the three panels have the same starting points x_0 ($x_0 = 1$ and $x_0 = 30$ respectively). In the first panel $\alpha = 1$ and $\beta = 1$; in the second panel β is lower, while α is unchanged: as a consequence, the steady state value $\mathrm{E}(X_\infty)$ is reduced, the points of inflection are moved and the curvature of the converging paths is changed.

Although the Gompertz solution X_t theoretically takes values in $\in [0, \infty)$, infinite values are taken only for non stationary sequences. In most applications, the steady state value is always finite, so one can reasonably assume a finite support for X_t as in our application in Section 3. The very few exceptions should then be dropped from the analysis, because for these data one does not have a steady state solution anyway.

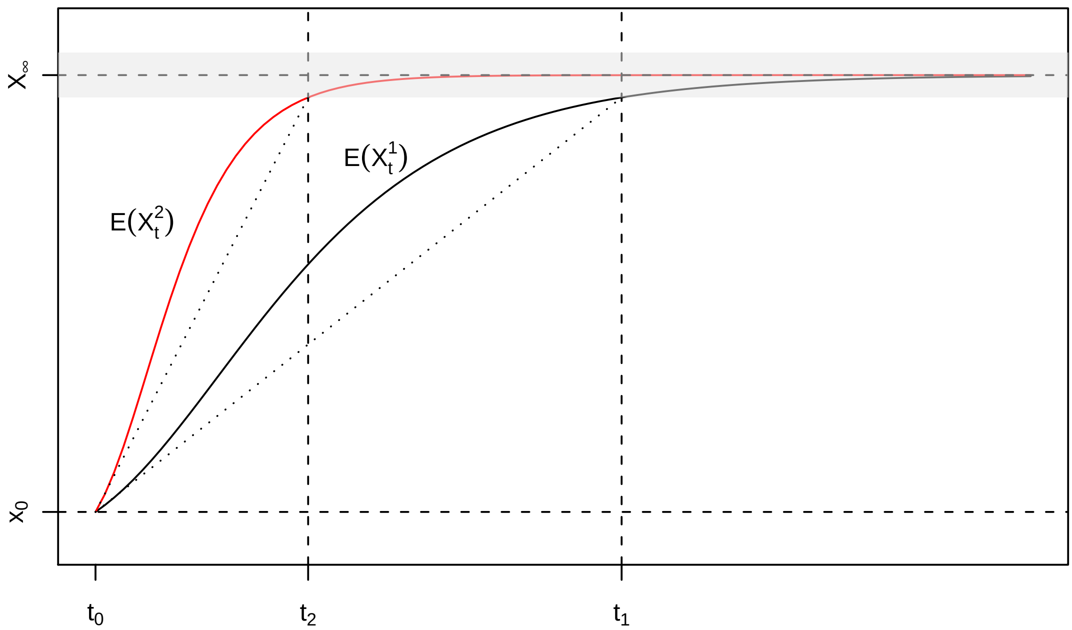

Figure 2: The speed of convergence of the trend s, $E(X_t^1)$ and $E(X_t^2)$, of two Gompertz processes with same limiting value X_∞ and initial value x_0. The speed of convergence $s_1(b)$ and $s_2(b)$ are the slopes of the dotted lines respectively for $E(X_t^1)$ and $E(X_t^2)$.

2.2 Speed of convergence to the steady state

The speed of convergence of $\mathrm{E}(X_t)$ to the steady state $\mathrm{E}(X_\infty)$ is obtained as the derivative of $\mathrm{E}(X_t)$ with respect to t. This quantity depends on t and it is not easy to interpret.

Therefore, although the process X_t is in a steady state only at $t = \infty$, one can ask the following question: let μ be some reference value and assume $E(X_t) > \mu$, then when is $E(X_t)$ sufficiently close to μ, i.e. for which t^* do we have that $E(X_{t^*}) - \mu < b_1$? More precisely:

$$t^* = \arg\min_{t>t_0} E(X_t) < \mu + b_1, \quad b_1 > 0.$$

Conversely, if $E(X_t) < \mu$, then we search for t^* such that

$$t^* = \arg\min_{t>t_0} E(X_t) > \mu - b_2, \quad b_2 > 0,$$

where b_1 and b_2 do not necessarily coincide. In both cases, the value of t^* always exists because $E(X_t)$ is monotonic.

Then, for an initial t_0, when $E(X_t) > \mu$ the speed of convergence of $E(X_t)$ around the reference value μ by an error of b_1 is defined as follows:

$$s(b_1) = \frac{(\mu + b_1) - E(X_{t_0})}{t^* - t_0} < 0 \tag{4}$$

and, when $E(X_t) < \mu$, we have:

$$s(b_2) = \frac{(\mu - b_2) - E(X_{t_0})}{t^* - t_0} > 0. \tag{5}$$

Figure 2 illustrates the trend functions of two Gompertz processes, say X_t^1 and X_t^2, with the same limiting steady state value X_∞, but different sets of parameters (α, β, σ). Let us choose $\mu = X_\infty$ and $b_2 > 0$, then

$$t_i = \arg\min_{t > t_0} E(X_t^i) > X_\infty - b_2, \quad i = 1, 2.$$

Therefore, the speed of convergence of $E(X_t^i)$ to $\mu = X_\infty$ is calculated as

$$s_1(b_2) = \frac{(X_\infty - b_2) - E(X_{t_0}^1)}{t_1 - t_0}, \quad s_2(b_2) = \frac{(X_\infty - b_2) - E(X_{t_0}^2)}{t_2 - t_0}$$

which correspond to the slopes of the two dotted lines in Figure 2.

In Section 4.2 we will construct clusters of regions and choose μ as the average steady state value of the cluster, where $\mu + b_1$ and $\mu - b_2$ are, respectively, the highest and the lowest steady state values of the cluster itself.

3 EU27 unemployment over the period 1991-2008

We assume that the unemployment rates in the EU regions evolve according to a Gompertz process. This allows for estimating the long run stationary value of the unemployment rate for each region and examining the tendency of the unemployment distribution among regions.

Consider the unemployment rates of the EU27 regions in the last three decades (1991-2008): we try to discover the emergence of possible "convergence clubs", as theorized by Quah (1997).

3.1 Steps of the analysis

The evolution of the unemployment distribution across the EU regions along the decades is examined according to the following steps:

1. we estimate the Gompertz process for the unemployment rate of every region in the dataset;

2. using the estimated parameters (α, β, σ) of the Gompertz process, we calculate the long-run value of the unemployment rate X_∞;

3. applying a cluster analysis algorithm, we group the regions according to their long-run values of unemployment rates X_∞;

4. we evaluate the variance of the unemployment rates within clusters and between clusters and compare it to the expected long run regional variance of unemployment rates within the countries;

5. we calculate the convergence speed and the time of convergence of each region to the average long-run unemployment rate of their cluster.

3.2 Data description

The analysis is made on the NUTS2 EU27 regions whose data are made available for the period 1991-2008 by Cambridge Econometrics[1]. Data on unemployment rates are available for 255 regions in 21 countries[2]. For three regions (HU33: Dél-Alföl - Hungary; PT16: Centro - Portugal; UKF3: Lincolnshire - Uk) the estimated Gompertz processes do not converge to a steady state, due to a sequence of increasing unemployment rates that prevents form estimating the deceleration factor of the process: the regions have been dropped from the subsequent analysis. Therefore, the number of regions examined is 252. Estimation of the parameters of the Gompertz process is done via maximum likelihood estimation using the R package `yuima` (YUIMA Project Team, 2012).

3.3 Steady state analysis

The range of the current unemployment rates of the EU regions in 1991 is [2.50, 26.90]. The range of the estimated steady state values by region is [0.69, 20.99]. The dispersion of the unemployment rates, measured by the standard deviation, decreases from 5.18 to 3.98, showing a progressive reduction and concentration of the unemployment values (see Figure 3). This may be read as a general convergence to a lower average unemployment rate. Nevertheless, this is not the whole story the data reveal: in fact, even if the distributions of unemployment rates do not show multimodality, groups of regions are directed to similar long-run stationary unemployment rates and these groups constitute, in some cases, local macro-regions, whose borders can change over time. This seems to be the point raised by Quah (1996a).

Applying an agglomerative algorithm to the steady state unemployment rates and by inspection of the dendrogram (see Figure 4), it seems reasonable to consider a six-clusters partition of the regions[3].

[1] For details on the dataset see: *http://www.camecon.com/Europe/Regional_Local_Cities/KnowledgeBase*. For some regions the dataset provides also the decade 1980-1990 and we use this additional information to estimate the steady state of these regions in order to get more reliable estimates of the parameters of the Gompertz model.

[2] See the Appendix for a list of the regions.

[3] A finer-grain clusterization leads, in some cases, to a fragmentation of small clusters, while the largest ones remain substantially unchanged. Progressively (nine clusters and more) the largest groups split up, but the differences among clusters become less sharp.

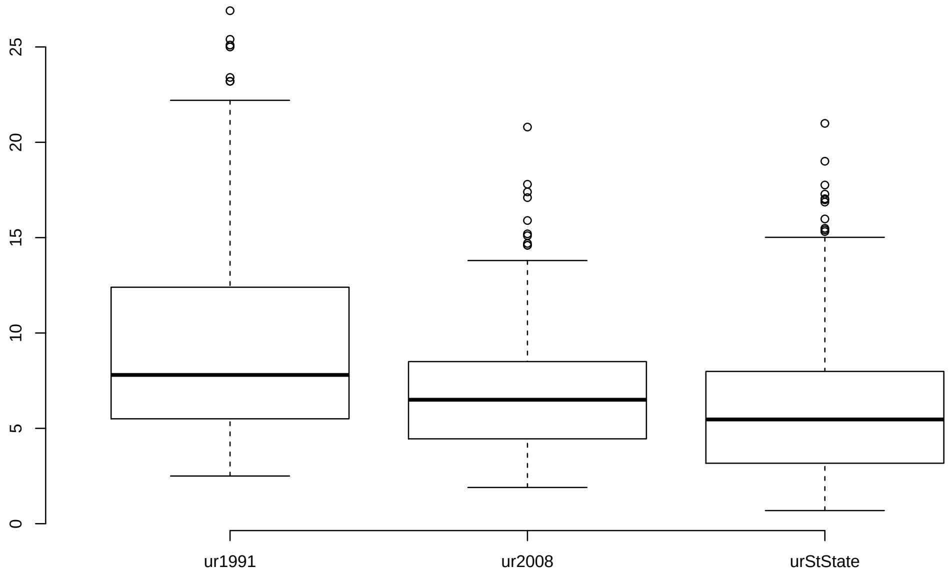

Figure 3: Distributions of unemployment rates in 1991, 2088 and in steady state.

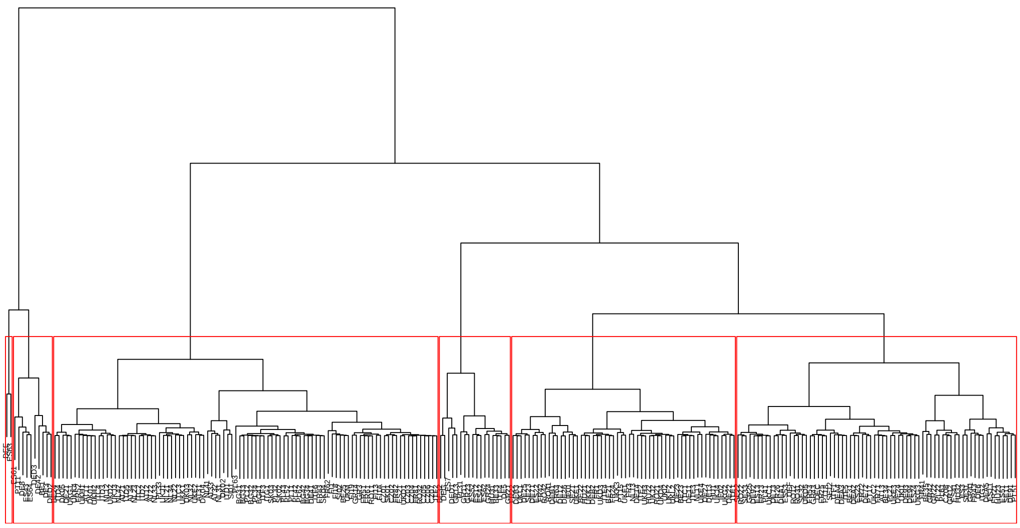

Figure 4: Cluster dendrogram of steady state unemployment rates.

The regions belonging to each group are listed in the Appendix and the clusters[4] are represented in the map in Figure 5.

	min X_∞	med X_∞	max X_∞	α	β	σ	n
CL1	0.69	1.68	4.32	0.01	0.02	0.14	96
CL4	4.47	5.36	6.68	0.46	0.28	0.13	56
CL2	6.80	7.92	10.26	0.45	0.22	0.15	70
CL5	10.65	11.47	13.62	0.55	0.21	0.19	18
CL3	15.02	16.42	17.76	0.56	0.20	0.21	10
CL6	19.01	20.00	20.99	1.02	0.33	0.33	2

Table 1: Distribution of steady state values X_∞ and estimated Gompertz parameters by cluster.

The groups are sharply separated (see Table 1) in terms of long-run unemployment rates: the median values go from 1.68 (cluster n.1) to 16.42 (cluster n.3)[5]. Visual inspection of Figure 5 shows that the clusters do not follow the national borders. Roughly, we can describe the clusters as follows:

1. the largest group of regions has the lowest median unemployment rate, occupies a macro-region in Central-Eastern Europe (from Northeastern Italy to Austria, Czech Republic and Poland) and extends to the Netherlands, Belgium and Denmark. The cluster also includes Bulgaria, Finland, Southwestern France, several regions in Southern Italy and a large portion of Great Britain;

2. Central Europe is occupied by two clusters with medium steady state unemployment (median rates are 5.36 and 7.92 respectively); the area also includes Central Italy, Sverige and Ireland;

3. the clusters that exhibit the highest steady state unemployment rates (11.47 and 16.42 as median values) include regions in Greece, Spain and Northestern Germany.

Table 2 also shows the sectoral composition of the GDP, the share of the employment in the main economic sectors of the six clusters, and an indicator of per capita GDP in each cluster with respect to the EU27 average.

To avoid misinterpretation of the results, it is worth noting once more that the regions are not grouped according to a forecast of their unemployment rates, but on a value that summarizes the evolutionary path of the unemployment rate of the last two decades. Italy is a good example of this apparent paradox. In fact, regions from both Northern and Southern Italy belong to the lowest-unemployment cluster but for structurally different reasons: in Northern regions the

[4]The color indicates the average value of the steady state rates of the cluster.

[5]Cluster n. 6 is made by two regions only, with median long-run unemployment rate 20%. The clusters are disjoined by construction.

	agr	nrg	mkt	oagr	ocos	odis	ofin	other	onrg	gdp
CL1	0.08	0.46	0.46	0.05	0.08	0.25	0.13	0.27	0.18	102.50
CL4	0.06	0.47	0.47	0.03	0.07	0.25	0.14	0.30	0.17	106.50
CL2	0.07	0.47	0.47	0.03	0.07	0.25	0.12	0.31	0.17	95.00
CL5	0.07	0.47	0.47	0.06	0.09	0.25	0.11	0.27	0.16	78.50
CL3	0.09	0.45	0.45	0.03	0.08	0.24	0.14	0.35	0.11	86.00
CL6	0.18	0.41	0.41	0.01	0.09	0.22	0.10	0.48	0.10	87.50

Table 2: Sectoral composition of the GDP: `agr`: % Agriculture; `ngr`: % Energy and Manufacturing, Construction; `mkt`: % Market and Non-Market Services. Employment by sector: `oagr`: % Agriculture; `ongr`: % Energy and Manufacturing; `ocos`: % Construction; `odis`: % Distribution, Hotel & Restaurants, Transport, Storage and Communications; `ofin`: % Financial Intermediation, Real Estate, Renting and Business Activities; `other`: % Non-Market Services. `gdp`: indicator of per-capita GDP. Average cluster values.

(relatively low) unemployment rates show a high stabiliy and only a slight decline in the last years of in the period examined; on the contrary, the high unemployment rates of the Southern regions have shown a significant decreasing trend in the same period.

3.4 Unemployment 1991 and clusters 2008

The initial distribution of unemployment is depicted in Figure 7, where the 1991 unemployment rate is assigned to each region. Note that the partiton of the range of unemployment rates is the same as in Figure 5, which is deduced from the steady state clusterization: therefore the colors indicate the 'steady state' cluster each region would belong to.

As a term of comparison to the steady state situation, we have grouped in six clusters the EU27 regions according to their unemployment rate in the last year of the period (2008). The result is represented in Figure 6. Note that each group falls in a different segment of the partition defined in Figure 5, which has been maintained also in this map: the two pictures, hence, can be easily compared by visual inspection on the basis of the colors of the clusters.

Is there anything new emerging from the comparison of Figure 6 and Figure 5? In other words, what is the added value of the estimation and clusterization on the steady state values, compared to, e.g., the Overman and Puga (2002) approach? As one can see, the convergence process which is pointed out by Figure 5 can in fact been deduced observing the evolution of the current unemployment: a comparison between Figure 7 and Figure 6 shows that a macro-region characterized by a low unemployment rate is growing in Central-Eastern Europe, and also includes Denmark, Belgium and the Netherlands. The area is surrounded by medium-unemployment regions, while the periphery of the continent (Greece, Spain, Southern Italy, together with Northestern Germany) shows higher rates of unemployment. The extension of the analysis to the steady state allows for emphasizing the long-run tendency of these 'peripherical'

regions: as Figure 5 shows, in fact, while Southern Italy is aggregated to the cluster with the lowest steady state unemployment rate - as a result of an incidence of unemployment which is decreasing over time - both Spain and, particularly, Greece converge to higher unemployment clusters. The different economic condition and policies adopted in these countries in the period examined (particulartly with respect to the labor market) must make the reader suspicious of any attempt of a comprehensive interpretation of this evolution. Nevertheless, the evidence seems to suggest that these critical areas have faced the 2008 economic crisis with a different background - and hence difference perspectives - at least from a labor market viewpoint.

3.5 Convergence clubs: clusters or countries?

We compare now the estimated convergence of unemployment rates within the clusters previously defined and within the countries. Being the clusters defined on the basis of a similar long-run stationary value, it is almost tautological to expect a higher convergence within the clusters: this notwithstanding, the exercise provides a measure of the significance of the 'new' EU inner borders compared to the 'traditional' ones.

To this aim, we estimate, via the Gompertz process, the average steady state unemployment rate for each cluster, each nation and the whole EU. Then, we evaluate the variance of the long-run regional unemployment around the continental rate (σ^2_{EU}) and decompose it, in the spirit of ANOVA, into the *within groups* and *between groups* components, using as groups, alternatively, the nations and the clusters. More formally, looking at the decomposition by cluster: let F_i be the labor force of region i, $i = 1, \ldots, N$, $F_{EU} = \sum_{i=1}^{N} F_i$ the continental labor force and $f_i = F_i/F_{EU}$ the proportion of labor force of region i. Let u_i the steady state unemployment rate of region i; $\bar{u}_{EU}$ be the average steady state unemployment rate for the whole EU; $\bar{u}_c$ be the average steady state unemployment rate for the cluster c; σ^2_c be the variance of the steady state unemployment rate inside cluster c, then:

$$\bar{u}_{EU} = \sum_{i=1}^{N} u_i f_i, \quad \sigma^2_{EU} = \sum_{i=1}^{N} (u_i - \bar{u}_{EU})^2 f_i$$

Now let f'_c the sum of the f_i over cluster c. Then

$$\sigma^2_{EU} = \sum_{c=1}^{n_c} \sigma^2_c f'_c + \sum_{c=1}^{n_c} (\bar{u}_c - \bar{u}_{EU})^2 f'_c = \sigma^2_{\text{within}} + \sigma^2_{\text{between}}$$

with n_c the number of clusters and f_c the labor force in each cluster c.

In the line of Overman and Puga (2002), we also calculate the same decomposition of the variance using the current values of unemployment rates in 1991 and 2008: a comparison of the two values with the steady state situation helps in reading the evolution of the unemployment distribution over time, pointing out differences between contingent and long-run components of

the phenomenon.

3.5.1 Results of ANOVA by cluster.

We evaluate the ANOVA decomposition by cluster using unemployment rates in 1991, 2008 and steady state. The significance of the decomposition by cluster is performed using a simple regression analysis on the unemployment rates u_i, weighted by f_i with cluster dummies in the model.

The coefficients (see Table 5 in the Appendix) represent the average unemployment rate of each cluster. The overall significance is tested using the Fisher-F test: the increase in the F-statistics (from 192.65 in the 1991 model, to 474.33 in the 2008 model, to 2109.56 in the steady state regression) shows what was largely expected, i.e. that the clusters have a strong role in explaining the variability of the steady state unemployment, but also reveals that the convergence can be partially read also through the evolution of the current unemployment rates. In other words, the *between groups* component $\sigma^2_{\text{between}}$ of the variance σ^2_{EU} conveys a similar information: it raises from less than 20% of the total variance σ^2_{EU} using the 1991 unemployment rates to over 90% with the steady state rates, but it is already around 55% when is calculated with the 2008 rates.

3.5.2 Results of ANOVA by country.

How far from the traditional borders of the European countries are the clusters that emerge from the long-run trends of the European unemployment? First of all, when we estimate the linear regressions of the unemployment rates on the country dummy variables (see Table 6 in the Appendix), we find that the F-statistics does not show significant changes over time: the value raises from 75.96 in 1991 to 85.63 in 2008, but it falls down to 49.44 when using the steady state rates. The *within* and *between* countries components of the variance of continental unemployment rate indicate that the explicative power of the country variable - the *between* component - falls from 42% in 1991 to 35% in 2008. In a long-run perspective the situation is only sligthly different: the *between* countries share of variance is less than 41%. As expectable, the country variable is still significant in explaining the evolution of unemployment, but - in a period when the distribution of the regional unemployment rates show a decline on average and in variance (see Figure 3) - the decomposition of the continental variance of unemployment does not exhibit a clear pattern, suggesting that regional inequalities, evaluated on a country basis, are not necessarily going to decrease.

4 What do we learn from the Gompertz model?

As illustrated in Section 2, the parameters of the Gompertz process both determine the value of the steady state and control the convergence speed.

4.1 The parameters of the Gompertz model.

In particular, α, as the intrinsic growth rate, measures the speed the process would go to infinity in absence of attrition, while β is the attrition (i.e., the deceleration) factor. Therefore, reading Table 1 three different patterns of convergence of the clusters to the steady state emerge, clearly indicated by the average values of the parameters[6] and illustrated in Figure 8:

- cluster n.1 with the lowest long-run unemployment rate has low values both for α and β. The two parameters contribute to define a low steady state unemployment. Besides, the virtual absence of attrition make the convergence rapid. This allow for two possible paths: on one side, regions whose initial value of current unemployment is not far from the steady state show a constant unemployment rate over time; on the other side, regions whose initial unemployment rate is far from the steady state rapidly fall to long-run values that are even lower than the custer median. As discussed in Section 3.3, this is the reason why both Northern and Southern Italian regions belong to the same cluster;

- two clusters (n. 4 and 2) with intermediate values of the steady state unemployment both show a higher intrinsic growth rate ($\alpha = 0.45$ and 0.46 respectively). The attrition factor regulates the median long-run unemployment: where β is higher (0.28), the steady state value is lower, and viceversa;

- a similar behavior characterizes the clusters with the highest long-run unemployment (n. 5 and 3). In their case, however, higher intrinsic growth rates ($\alpha = 0.55$ and 0.56 respectively) lead to higher steady state unemployment. The slight differences in α and β between the two groups account for the different long-run median rates.

The volatility component (σ) is quite similar for all the clusters (between 0.13 and 0.21), hence it does not play a determinant role in diversifying the steady state values.

4.2 The speed of convergence

If an inspection of the Gompertz parameters can easily give information about the convergence values, this is not the case for the convergence speed. As argued in Section 2.2, the convergence speed estimated as the derivative of $\mathrm{E}(X_t)$ with respect to t is a function of t and of the Gompertz parameters which cannot provide an intuitive interpretation. For instance, as we noticed in

[6]A fourth, trivial, behaviour is shown by cluster n. 6, made by two regions only.

cluster n.1, in a group of regions characterized by similar parameters we may observe different convergence speed, depending on the distance between the initial point of the process and the steady state value. That is why we propose an explicit evaluation of the convergence speed as defined in equations (4) and (5). For each cluster $\mu + b_1$ is the highest regional unemployment steady state rate and $\mu - b_2$ is the lowest, where μ is the average steady state of the cluster. Clearly, the interval $[\mu - b_2, \mu + b_1]$ is not symmetric around μ. Therefore, the $s(b_1)$ and $s(b_2)$ of (4) and (5) denote the speed at which the estimated regional trend $E(X_t)$ enters the cluster bands. Figure 8 illustrates the convergence bands for each cluster and the trajectories of the single regions.

Table 3 reports the absolute average speed of convergence by cluster and the average time of entrance into the bands. The most rapid average convergence is registered by cluster n.4 and n.2 (around 5 years). Clusters n.1 and n.5 converge, on average, in a period five time higher. In cluster n.3, on the other hand, average convergence is ten times lower than in clusters n.4 and n.2 . The last column shows that in these two clusters (n.4 and n.2) more than 50% of the regions lie inside the convergence bands from the beginning of the observed period: this, of course, reduces the average convergence time of the clusters. The third and fourth columns report the minimum and maximum regional speed of convergence for each cluster: positive or negative values - as illustrated in Section 2.2 - indicate a convergence “from below” or “from above” respectively. The average of the absolute values of the convergence speed (second column) clarifies that there is no necessary correlation between the time to convergence and the convergence speed: given the time to enter the cluster bands, in fact, the speed of convergence is determined by the initial value of the unemployment rate. Therefore, we may have - as for cluster n.3 - both a high time to convergence and a high average convergence speed, or - as for cluster n.4 - low values for both variables.

Table 4 summarizes, as an example, the content of what is graphically represented in the top-right panel of Figure 8, i.e. the behaviour of estimated trends of the regions belonging to cluster n.3. the trajectories of four regions are included into the cluster bands from 1991.

An estimate of the β-convergence. Thanks to the evaluation of the speed of convergence, we can also estimate a quantity equivalent to the traditional β-convergence (Barro and Sala-i-Martin 1995; Sala-i-Martin 1996). A positive correlation between the absolute convergence speed and the distance (in absolute value) of the initial unemployment rate from the steady state indicates, in fact, that the regions that are more distant from their steady state are converging more rapidly.

The average value of β-convergence - calculated in such a way - among the European regions is 0.42. It is worth noting that - with the exception of cluster n.3, which has β-convergence equal to 0.31 - the average values of the correlation calculated by cluster is significantly higher (from 0.57 to 0.83). The distribution of the correlation values by country, on the contrary, has

a larger variance and the correlation takes, in one case, a negative value.

	avg time to conv	avg. abs speed	min speed	max speed	% to conv
CL1	24.88	0.21	-0.60	-0.005	0.68
CL4	5.67	0.18	-0.60	0.07	0.43
CL2	5.05	0.27	-0.91	0.34	0.46
CL5	25.22	0.31	-0.76	0.18	0.67
CL3	49.05	0.33	-0.90	0.15	0.60
CL6	3.00	0.93	-1.10	0.75	1.00

Table 3: Distribution of time to convergence, convergence speed and % of regions not initially in cluster bands.

Region	time to convergence	convergence speed	included
BE1	0.00		TRUE
DE3	0.00		TRUE
DE42	64.06	0.07	
DE8	0.00		TRUE
DED2	1.00	-0.34	
DED3	0.00		TRUE
ES43	14.51	-0.51	
ES61	335.84	-0.02	
ES64	4.50	-0.90	
PT11	70.57	0.15	

Table 4: Cluster n.3. Time to convergence and convergence speed by region. Four regions are initially included in the cluster bands. Cluster avgs: mean time = 49.05, abs. speed = 0.33.

5 What is a Gompertz model useful for (and what is not)?

We adopt a Gompertz process to show the evolution of the unemployment rate basically because it is a flexible convergence process: in other words, it allows for representing several different paths of convergence to a long run steady state value. Moreover, it allows to distinguish between the trend component of the process - i.e. the structural part of the convergence path - and the volatile component; on the contrary, it is not possible to disentangle the two components when we estimate the convergence through the current values of the variable.

On the other hand, the estimation of a Gompertz process must be correctly interpreted, not to be misleading: the estimated process is an adequate description of the evolution of the variable (the unemployment rate, in our case) on the basis of the current and past values of the variable itself. In other words, the parameters of the process summarize the path of the variable in the period when the variable has been observed and indicate the value where the

process is directed to under the assumption of stationarity of the process. But we know that the tendency of the unemployment rate to its long run value is perturbed by exogenous shocks and conditioned by several covariates of the economic system, and hence we cannot rely on the steady state value as an estimate of the prospective value of the unemployment rate.

Therefore, the representation of the unemployment rate as a stochastic Gompertz process is an useful tool for:

- describing the direction assumed by the variable: is the unemployment rate structurally increasing or decreasing?
- estimating the speed of convergence to the steady state: is it a rapid or slow evolution?
- evaluating the volatility of the process around its trend: is the process subject to frequent and significant shocks?

The estimation of a Gompertz process can be useful also for forcasting purposes but, if this is the case, the examination of a single variable cannot be enough and the interdependence among the features of the system must be taken into account. Therefore:

- a *conditional* Gompertz process should be estimated, where other covariates appear, conditioning the evolution of the unemployment rate;
- or a model should be proposed where the evolution process of all the relevant variables is explicitly represented and estimated.

These extensions are left as topics for further research.

6 Concluding remarks

The European unemployment rate has shown a decrease between 1991 and 2008 and the distribution of the regional rates has registered a reduction in variance: this can be read as a general convergence of the continent towards a lower average unemployment. A part of the literature on convergence empirics, however, argues that the actual dynamics can be more complex, and suggests to check for phenomena of polarization, stratification, overtaking and divergence.

In this paper we have examined to what extent the distribution of unemployment among the European region has followed a clustering pattern. The clusterization of EU regions in terms of unemployment is related to the idea of "convergence clubs" proposed in Quah (1996a): in that paper the usual concepts of σ-convergence and β-convergence are considered unable to capture "the different convergence dynamics" which are 'generated, depending on the initial distribution of characteristics across countries' (p.1368).

Assuming that the evolution of the unemployment rates can be represented by a Gompertz stochastic diffusion process, we have estimated the steady state unemployment rates of the European regions, trying to identify the "convergence clubs" that emerge from the examination of the 1991-2008 period. In the analysis, the assumption of the Gompertz process has shown a twofold utility:

- the estimation of the expected value of the process has allowed for a clusterization of the European regions which distinguishes the structural from the volatile component of the evolution of unemployment. Moreover, the clusterization on the steady state value of the variable of interest, instead of the current values only, has added interesting hints to the interpretation of the phenomenon. From this first step of the analysis five main clusters emerge (plus a minor one);
- the parameters of the Gompertz process - which measure the intrinsic growth, the attrition rate and the volatility - help to characterize the 'different convergence dynamics' - using the words of Quah (1996a) - followed by regions and clusters: in this second step, the inspection of the estimated α, β and σ allows for identifying three main patterns of convergence of the clusters to the steady state.

A method for calculating the speed of convergence of each region to its cluster has also been proposed, which allows for an evaluation of the more traditional β-convergence.

Three further remarks must be made:

1. The choice of the unemployment rate as the clusterization variable has been made in order to replicate and extend the study by Overman and Puga (2002). The same exercise can be obvioulsy replicated - with comparative purposes and following the suggestion of a large part of the literature on convergence - using the GDP or the per capita GDP as the criterion to identify the convergence or clusterization process.
2. Our results show some regularities in the distribution of unemployment over time among the EU regions. This does not mean we provide a unitary explanation of the evolution of unemployment in the continent. Each cluster is composed by different processes of convergence, as the examination of the parameters and the time and speed of convergence may confirm, and also similar processes of convergence are likely the result of different policies in different economic contexts. Therefore, several structural models are required to explain the trends and agglomeration that our estimates have pointed out.
3. As illustrated in Figure 3, in the period 1991-2008 the European regional unemployment rates show, on average, a decline and a reduction in variance. One may wonder how the number and the borders of the clusters are going to change during and after the current

economic crisis, that is causing a general increase in unemployment. The answer depends on how deep and how long the crisis is going to be: the first years of deviation of the unemployment rates from the trend, in fact, will be interpreted as a part of the volatility component of the process, and will not substantially change the value of the parameters (with the exception of an increase in σ); a long lasting crisis - as the current one seems to be - on the contrary, will generate a deviation from the trend of the process that can be interpreted as permanent, and hence imply changes in the estimation of the parameters and of the steady state values of the process, and a possible redraw of the clusterization trends.

Appendix

```
CLUSTERS ON 2008 UNEMPLOYMENT RATES
===================================

cluster n.1:
 [1] "AT11" "AT12" "AT21" "AT22" "AT31" "AT32" "AT33" "AT34"
 [9] "BE23" "BE24" "BE25" "BG41" "CZ01" "CZ02" "CZ03" "CZ05"
[17] "CZ06" "DE11" "DE13" "DE14" "DE21" "DE22" "DE23" "DE27"
[25] "DK01" "DK02" "DK03" "DK04" "DK05" "ITC2" "ITC4" "ITD1"
[33] "ITD2" "ITD3" "ITD4" "ITD5" "NL11" "NL12" "NL13" "NL21"
[41] "NL22" "NL23" "NL31" "NL32" "NL33" "NL34" "NL41" "NL42"
[49] "RO11" "RO32" "SI02" "SK01" "UKD1" "UKE2" "UKG1" "UKJ1"
[57] "UKJ2" "UKJ3" "UKK1" "UKK4"

cluster n.2:
 [1] "AT13" "BE21" "BE22" "BE31" "BG34" "BG42" "CZ07" "DE12"
 [9] "DE24" "DE25" "DE26" "DE71" "DE72" "DE93" "DE94" "DEA3"
[17] "DEB1" "DEB2" "DEB3" "DEF"  "ES21" "ES22" "FI1"  "FI18"
[25] "FI19" "FR24" "FR25" "FR26" "FR42" "FR51" "FR52" "FR62"
[33] "FR63" "FR71" "FR72" "GR3"  "GR41" "GR43" "HU1"  "HU21"
[41] "HU22" "IE02" "ITC1" "ITC3" "ITE1" "ITE2" "ITE3" "ITF1"
[49] "PL11" "PL12" "PL21" "PL22" "PL34" "PL41" "PL43" "PL52"
[57] "PL63" "RO21" "RO31" "RO41" "RO42" "SE11" "SE21" "SE23"
[65] "SI01" "SK02" "UKD2" "UKD4" "UKE1" "UKE4" "UKF1" "UKF2"
[73] "UKG2" "UKH1" "UKH2" "UKH3" "UKI2" "UKJ4" "UKK2" "UKK3"
[81] "UKL1" "UKL2" "UKM"

cluster n.3:
[1] "BE1"  "DE3"  "DE8"  "DED3" "DEE"  "ES43" "ES61" "ES63"
[9] "ES7"

cluster n.4:
 [1] "BE32" "DE41" "DED1" "DED2" "ES42" "ES52" "ES62" "FR3"
 [9] "GR13" "HU31" "HU32" "ITF3" "ITF4" "ITF5" "ITF6" "ITG1"
[17] "ITG2" "SK03" "SK04"

cluster n.5:
 [1] "BE33" "BE34" "BE35" "BG31" "BG32" "BG33" "CZ04" "CZ08"
```

```
 [9] "DE42" "DE5"  "DE6"  "DE73" "DE91" "DE92" "DEA1" "DEA2"
[17] "DEA4" "DEA5" "DEC"  "DEG"  "ES11" "ES12" "ES13" "ES23"
[25] "ES24" "ES3"  "ES41" "ES51" "ES53" "FI13" "FI1A" "FR1"
[33] "FR21" "FR22" "FR23" "FR41" "FR43" "FR53" "FR61" "FR81"
[41] "FR82" "FR83" "GR11" "GR12" "GR14" "GR21" "GR22" "GR23"
[49] "GR24" "GR25" "GR42" "HU23" "IE01" "ITE4" "ITF2" "PL31"
[57] "PL32" "PL33" "PL42" "PL51" "PL61" "PL62" "PT11" "PT15"
[65] "PT17" "PT18" "RO12" "RO22" "SE12" "SE22" "SE31" "SE32"
[73] "SE33" "UKC1" "UKC2" "UKD3" "UKD5" "UKE3" "UKG3" "UKI1"

cluster n.6:
[1] "ES64"

CLUSTERS ON STEADY STATE UNEMPLOYMENT RATES
===========================================

cluster n.1:
 [1] "AT11" "AT12" "AT21" "AT22" "AT31" "AT32" "AT33" "AT34"
 [9] "BE24" "BE25" "BG31" "BG32" "BG33" "BG34" "BG41" "BG42"
[17] "CZ01" "CZ02" "CZ03" "CZ04" "CZ05" "CZ06" "CZ07" "DE21"
[25] "DE41" "DK01" "DK02" "DK03" "DK04" "DK05" "ES21" "FI1"
[33] "FI13" "FI18" "FI19" "FI1A" "FR51" "FR52" "FR61" "FR62"
[41] "FR81" "FR82" "FR83" "GR14" "GR3"  "ITC2" "ITC3" "ITC4"
[49] "ITD1" "ITD2" "ITD3" "ITD4" "ITD5" "ITF2" "ITF3" "ITF6"
[57] "ITG1" "NL12" "NL21" "NL22" "NL23" "NL31" "NL32" "NL33"
[65] "NL34" "NL41" "NL42" "PL11" "PL12" "PL32" "PL33" "PL34"
[73] "PL42" "PL43" "PL52" "PL62" "PL63" "RO11" "RO32" "SI01"
[81] "SI02" "SK01" "SK02" "SK03" "SK04" "UKD1" "UKD2" "UKE2"
[89] "UKG1" "UKH1" "UKH2" "UKJ1" "UKJ3" "UKK1" "UKK4" "UKM"

cluster n.2:
 [1] "AT13" "BE31" "BE34" "BE35" "CZ08" "DE24" "DE6"  "DE72"
 [9] "DE73" "DE91" "DE92" "DE93" "DE94" "DEA1" "DEA2" "DEA3"
[17] "DEA4" "DEA5" "DEC"  "DEF"  "ES11" "ES13" "ES23" "ES24"
[25] "ES3"  "ES41" "ES51" "ES53" "FR21" "FR23" "FR25" "FR26"
[33] "FR41" "FR43" "FR53" "FR71" "FR72" "GR12" "GR22" "GR25"
[41] "GR41" "GR42" "GR43" "HU23" "HU32" "ITF1" "ITF5" "ITG2"
[49] "PL21" "PL22" "PL41" "PL51" "PL61" "PT17" "PT18" "RO12"
[57] "RO22" "RO31" "SE11" "SE12" "SE22" "SE31" "SE32" "SE33"
[65] "UKC1" "UKC2" "UKD5" "UKE3" "UKG3" "UKI1"

cluster n.3:
 [1] "BE1"  "DE3"  "DE42" "DE8"  "DED2" "DED3" "ES43" "ES61"
 [9] "ES64" "PT11"

cluster n.4:
 [1] "BE21" "BE22" "BE23" "DE11" "DE12" "DE13" "DE14" "DE22"
 [9] "DE23" "DE25" "DE26" "DE27" "DE71" "DEB1" "DEB2" "DEB3"
[17] "ES22" "FR1"  "FR22" "FR24" "FR3"  "FR42" "FR63" "HU1"
[25] "HU21" "HU22" "IE01" "IE02" "ITC1" "ITE1" "ITE2" "ITE3"
[33] "ITE4" "NL11" "NL13" "PT15" "RO21" "RO41" "RO42" "SE21"
[41] "SE23" "UKD3" "UKD4" "UKE1" "UKE4" "UKF1" "UKF2" "UKG2"
[49] "UKH3" "UKI2" "UKJ2" "UKJ4" "UKK2" "UKK3" "UKL1" "UKL2"
```

```
cluster n.5:
 [1] "BE32" "BE33" "DE5"  "DED1" "DEG"  "ES12" "ES42" "ES52"
 [9] "ES62" "ES7"  "GR11" "GR13" "GR21" "GR23" "GR24" "HU31"
[17] "ITF4" "PL31"

cluster n.6:
[1] "DEE"  "ES63"
```

ANOVA results

	ur1991	ur2008	urSt.State
clu: 1	9.327***	5.409***	2.318***
	(0.464)	(0.229)	(0.105)
clu: 2	9.269***	7.702***	8.089***
	(0.536)	(0.263)	(0.121)
clu: 3	15.843***	14.386***	16.684***
	(1.301)	(0.618)	(0.285)
clu: 4	7.045***	5.741***	5.352***
	(0.561)	(0.277)	(0.128)
clu: 5	13.610***	11.612***	11.799***
	(1.246)	(0.583)	(0.269)
clu: 6	17.628***	14.661***	20.942***
	(3.445)	(1.853)	(0.854)
R-squared	0.825	0.920	0.981
adj. R-squared	0.820	0.918	0.980
F	192.653	474.332	2109.563

Table 5: ANOVA by cluster

	ur1991	ur2008	urSt.State
country: AT	3.760*	3.803**	4.056**
	(1.878)	(1.265)	(1.506)
country: BE	8.531***	6.966***	7.535***
	(1.773)	(1.193)	(1.420)
country: BG	15.501***	5.593***	1.036
	(1.846)	(1.382)	(1.646)
country: CZ	7.918***	4.390***	2.351
	(1.571)	(1.140)	(1.357)
country: DE	8.779***	7.477***	8.364***
	(0.576)	(0.402)	(0.479)
country: DK	3.676	3.324*	3.439
	(2.119)	(1.518)	(1.807)
country: ES	14.883***	11.331***	10.771***
	(0.910)	(0.546)	(0.650)
country: FI	10.134***	6.618***	1.605
	(1.622)	(1.147)	(1.365)
country: FR	11.930***	7.359***	5.062***
	(0.726)	(0.492)	(0.586)
country: GR	11.272***	7.664***	6.269***
	(1.811)	(1.174)	(1.397)
country: HU	6.716***	7.687***	7.030***
	(1.800)	(1.361)	(1.620)
country: IE	5.866	5.962***	5.169*
	(3.096)	(1.745)	(2.077)
country: IT	10.501***	6.732***	4.382***
	(0.746)	(0.521)	(0.620)
country: NL	3.661**	2.747**	3.363**
	(1.365)	(0.878)	(1.045)
country: PL	12.412***	7.128***	4.841***
	(0.875)	(0.632)	(0.753)
country: PT	6.871***	8.455***	12.118***
	(1.883)	(1.302)	(1.550)
country: RO	6.881***	5.787***	5.557***
	(1.036)	(0.827)	(0.985)
country: SE	7.704***	6.873***	7.675***
	(1.689)	(1.238)	(1.473)
country: SI	5.597	4.362	2.007
	(3.675)	(2.555)	(3.042)
country: SK	16.183***	9.505***	1.104
	(2.344)	(1.591)	(1.893)
country: UK	5.517***	5.434***	4.723***
	(0.665)	(0.467)	(0.556)
R-squared	0.874	0.886	0.818
adj. R-squared	0.862	0.876	0.801
F	75.957	85.626	49.435

Table 6: ANOVA by country

Region	Name
at11	Burgenland (A)
at12	Niederösterreich
at13	Wien
at21	Kärnten
at22	Steiermark
at31	Oberösterreich
at32	Salzburg
at33	Tirol
at34	Vorarlberg
be1	Région de Bruxelles-Capitale
be21	Prov. Antwerpen
be22	Prov. Limburg (B)
be23	Prov. Oost-Vlaanderen
be24	Prov. Vlaams Brabant
be25	Prov. West-Vlaanderen
be31	Prov. Brabant Wallon
be32	Prov. Hainaut
be33	Prov. Liège
be34	Prov. Luxembourg (B)
be35	Prov. Namur
bg31	Severozapaden
bg32	Severen tsentralen
bg33	Severoiztochen
bg34	Yugoiztochen
bg41	Yugozapaden
bg42	Yuzhen tsentralen
cz01	Praha
cz02	Strední Cechy
cz03	Jihozápad
cz04	Severozápad
cz05	Severovýchod
cz06	Jihovýchod
cz07	Strední Morava
cz08	Moravskoslezsko
de11	Stuttgart
de12	Karlsruhe
de13	Freiburg
de14	Tübingen
de21	Oberbayern
de22	Niederbayern
de23	Oberpfalz
de24	Oberfranken
de25	Mittelfranken
de26	Unterfranken
de27	Schwaben
de3	Berlin
de41	Brandenburg - Nordost
de42	Brandenburg - Südwest
de5	Bremen
de6	Hamburg
de71	Darmstadt

Region	Name
de72	Gießen
de73	Kassel
de8	Mecklenburg-Vorpommern
de91	Braunschweig
de92	Hannover
de93	Lüneburg
de94	Weser-Ems
dea1	Düsseldorf
dea2	Köln
dea3	Münster
dea4	Detmold
dea5	Arnsberg
deb1	Koblenz
deb2	Trier
deb3	Rheinhessen-Pfalz
dec	Saarland
ded1	Chemnitz
ded2	Dresden
ded3	Leipzig
def	Schleswig-Holstein
deg	Thüringen
dk01	Hovedstaden
dk02	Sjælland
dk03	Syddanmark
dk04	Midtjylland
dk05	Nordjylland
es11	Galicia
es12	Principado de Asturias
es13	Cantabria
es21	Pais Vasco
es22	Comunidad Foral de Navarra
es23	La Rioja
es24	Aragón
es3	Comunidad de Madrid
es41	Castilla y León
es42	Castilla-la Mancha
es43	Extremadura
es51	Cataluña
es52	Comunidad Valenciana
es53	Illes Balears
es61	Andalucia
es62	Región de Murcia
es63	Ciudad Autónoma de Ceuta (ES)
es64	Ciudad Autónoma de Melilla (ES)
es7	Canarias (ES)
fi1	Manner-Suomi
fi13	Itä-Suomi
fi18	Etelä-Suomi
fi19	Länsi-Suomi
fi1a	Pohjois-Suomi
fr1	Île de France

Region	Name
fr21	Champagne-Ardenne
fr22	Picardie
fr23	Haute-Normandie
fr24	Centre
fr25	Basse-Normandie
fr26	Bourgogne
fr3	Nord - Pas-de-Calais
fr41	Lorraine
fr42	Alsace
fr43	Franche-Comté
fr51	Pays de la Loire
fr52	Bretagne
fr53	Poitou-Charentes
fr61	Aquitaine
fr62	Midi-Pyrénées
fr63	Limousin
fr71	Rhône-Alpes
fr72	Auvergne
fr81	Languedoc-Roussillon
fr82	Provence-Alpes-Côte d'Azur
fr83	Corse
gr11	Anatoliki Makedonia, Thraki
gr12	Kentriki Makedonia
gr13	Dytiki Makedonia
gr14	Thessalia
gr21	Ipeiros
gr22	Ionia Nisia
gr23	Dytiki Ellada
gr24	Sterea Ellada
gr25	Peloponnisos
gr3	Attiki
gr41	Voreio Aigaio
gr42	Notio Aigaio
gr43	Kriti
hu1	Közép-Magyarország
hu21	Közép-Dunántúl
hu22	Nyugat-Dunántúl
hu23	Dél-Dunántúl
hu31	Észak-Magyarország
hu32	Észak-Alföld
hu33	Dél-Alföld
ie01	Border, Midlands and Western
ie02	Southern and Eastern
itc1	Piemonte
itc2	Valle d'Aosta/Vallée d'Aoste
itc3	Liguria
itc4	Lombardia
itd1	Provincia Autonoma Bolzano-Bozen
itd2	Provincia Autonoma Trento
itd3	Veneto
itd4	Friuli-Venezia Giulia

Region	Name
itd5	Emilia-Romagna
ite1	Toscana
ite2	Umbria
ite3	Marche
ite4	Lazio
itf1	Abruzzo
itf2	Molise
itf3	Campania
itf4	Puglia
itf5	Basilicata
itf6	Calabria
itg1	Sicilia
itg2	Sardegna
nl1	Noord-Nederland
nl12	Friesland (NL)
nl13	Drenthe
nl21	Overijssel
nl22	Gelderland
nl23	Flevoland
nl31	Utrecht
nl32	Noord-Holland
nl33	Zuid-Holland
nl34	Zeeland
nl41	Noord-Brabant
nl42	Limburg (NL)
pl11	Lódzkie
pl12	Mazowieckie
pl21	Malopolskie
pl22	Slaskie
pl31	Lubelskie
pl32	Podkarpackie
pl33	Swietokrzyskie
pl34	Podlaskie
pl41	Wielkopolskie
pl42	Zachodniopomorskie
pl43	Lubuskie
pl51	Dolnoslaskie
pl52	Opolskie
pl61	Kujawsko-Pomorskie
pl62	Warminsko-Mazurskie
pl63	Pomorskie
pt11	Norte
pt15	Algarve
pt16	Centro (PT)
pt17	Lisboa
pt18	Alentejo
ro11	Nord-Vest
ro12	Centru
ro21	Nord-Est
ro22	Sud-Est
ro31	Sud - Muntenia

Region	Name
ro32	Bucuresti - Ilfov
ro41	Sud-Vest Oltenia
ro42	Vest
se11	Stockholm
se12	Östra Mellansverige
se21	Småland med öarna
se22	Sydsverige
se23	Västsverige
se31	Norra Mellansverige
se32	Mellersta Norrland
se33	Övre Norrland
si01	Vzhodna Slovenija
si02	Zahodna Slovenija
sk01	Bratislavský kraj
sk02	Západné Slovensko
sk03	Stredné Slovensko
sk04	Východné Slovensko
ukc1	Tees Valley and Durham
ukc2	Northumberland, Tyne and Wear
ukd1	Cumbria
ukd2	Cheshire
ukd3	Greater Manchester
ukd4	Lancashire
ukd5	Merseyside
uke1	East Yorkshire and Northern Lincolnshire
uke2	North Yorkshire
uke3	South Yorkshire
uke4	West Yorkshire
ukf1	Derbyshire and Nottinghamshire
ukf2	Leicestershire, Rutland and Northants
ukf3	Lincolnshire
ukg1	Herefordshire, Worcestershire and Warks
ukg2	Shropshire and Staffordshire
ukg3	West Midlands
ukh1	East Anglia
ukh2	Bedfordshire, Hertfordshire
ukh3	Essex
uki1	Inner London
uki2	Outer London
ukj1	Berkshire, Bucks and Oxfordshire
ukj2	Surrey, East and West Sussex
ukj3	Hampshire and Isle of Wight
ukj4	Kent
ukk1	Gloucestershire, Wiltshire and Bristol/Bath area
ukk2	Dorset and Somerset
ukk3	Cornwall and Isles of Scilly
ukk4	Devon
ukl1	West Wales and The Valleys
ukl2	East Wales
ukm	Scotland

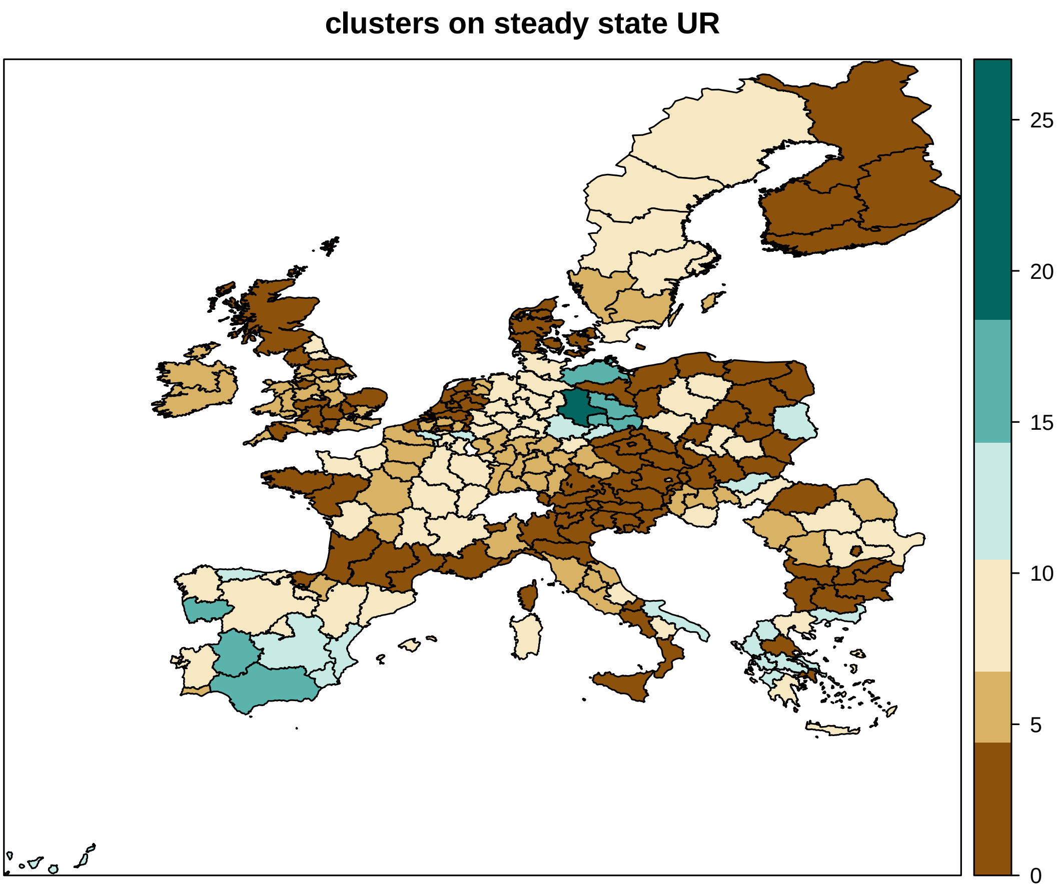

Figure 5: Steady State

clusters on 2008 UR

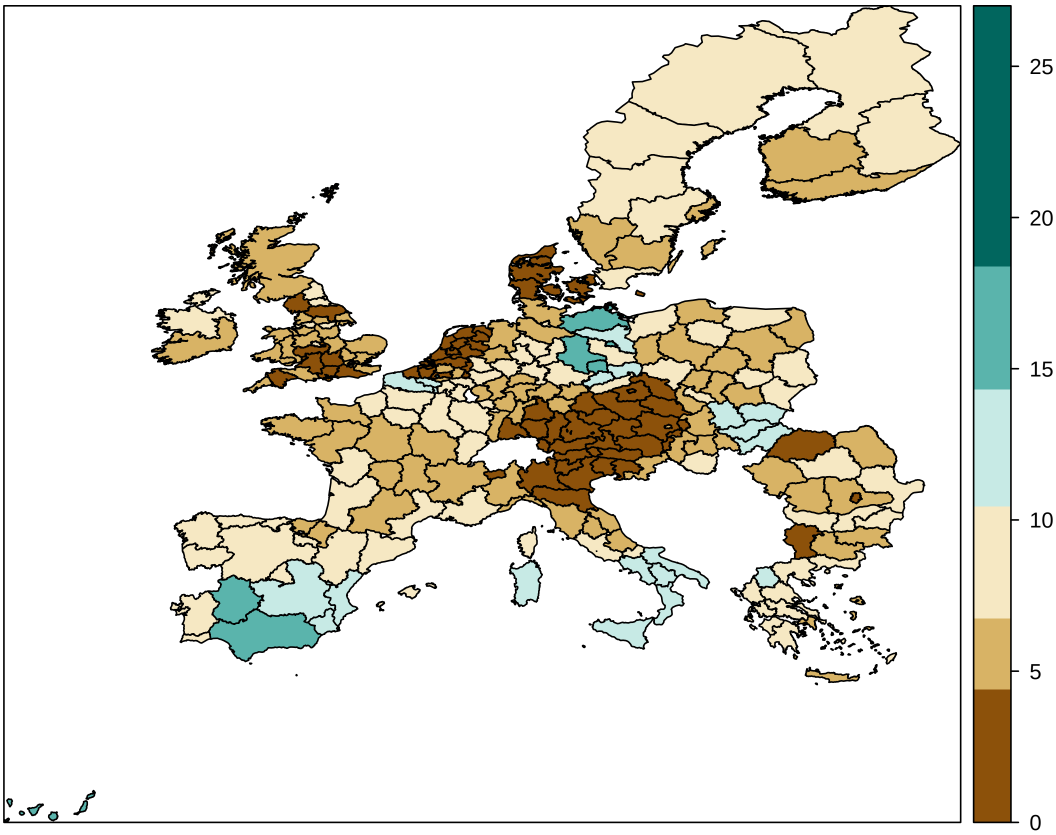

Figure 6: 2008

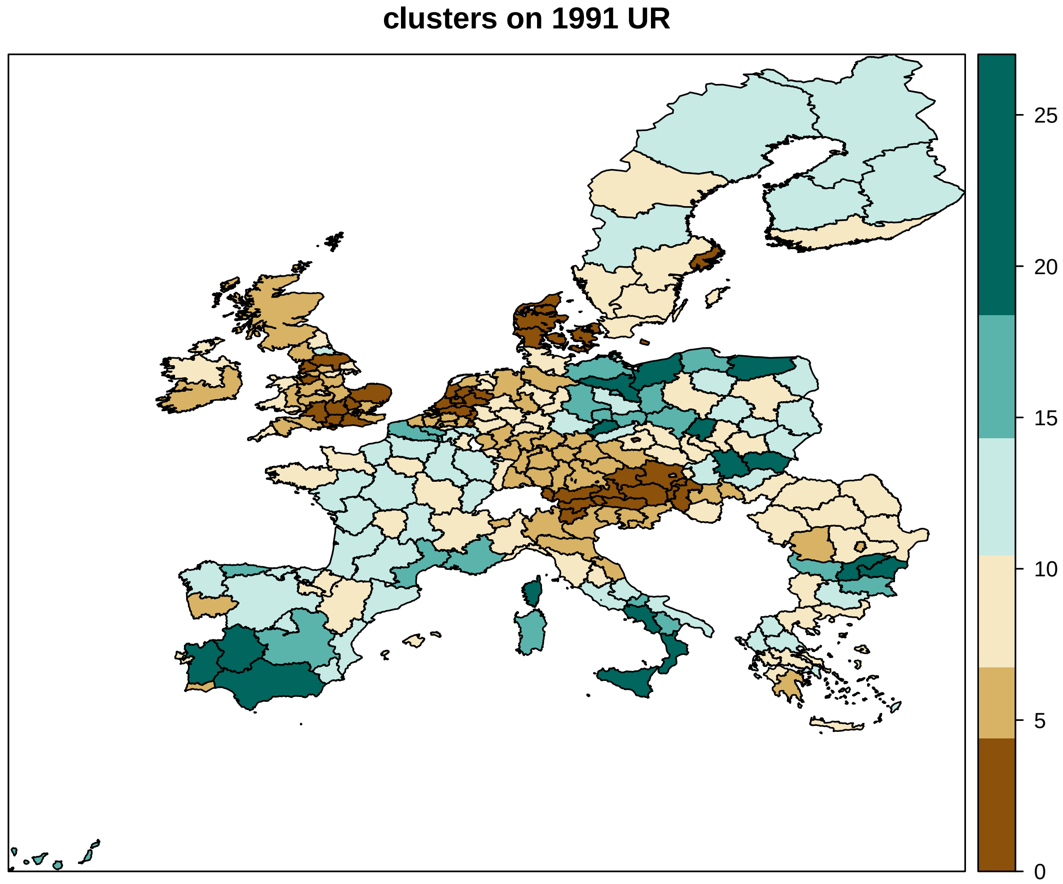

Figure 7: 1991

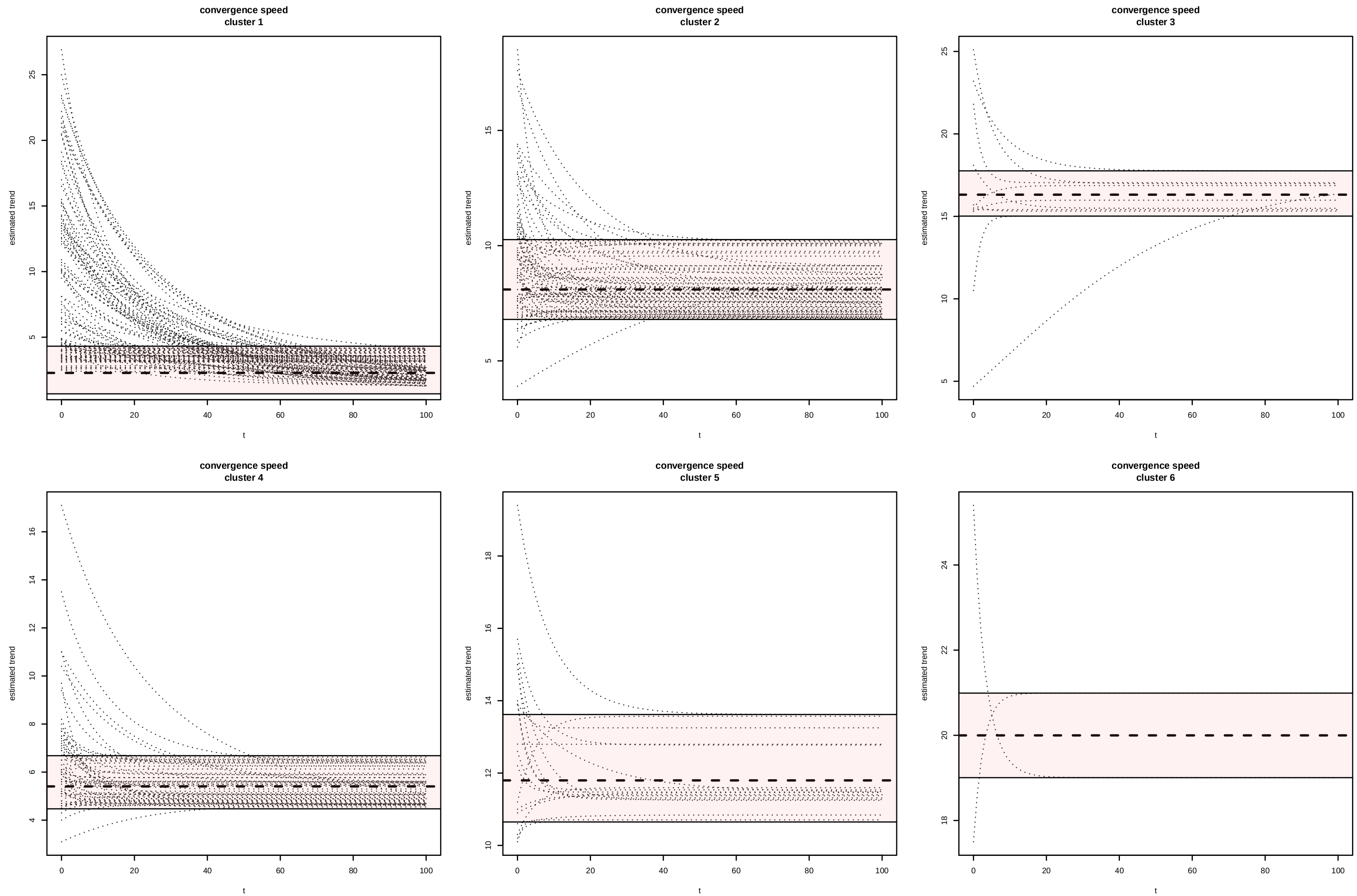

Figure 8: Convergence speed and clusters

References

[1] Barro R.J. - Sala-i-Martin X.X. (1992), “Convergence”, *Journal of Political Economy*, 100, 223-251.

[2] Barro R.J. - Sala-i-Martin X.X. (1995), *Economic Growth*, MIT Press, Cambridge.

[3] Capocelli, R.M. - Ricciardi, L.M. (1974) Growth with regulation in random environment, *Kybernetik*, 15, 147–157.

[4] Gutiérrez, R. - Nafidi A. - Gutiérrez-Sáchez R. (2005), “Forecasting total natural gas consumption in Spain by using the stochastic Gompertz innovation diffusion model”, *Applied Energy*, 80(2), 115-124.

[5] Gutiérrez, R. - Gutiérrez-Sáchez R. - Nafidi A. (2006), “Electricity consumption in Morocco: stochastic Gompertz diffusion analysis with exogenous factors”, *Applied Energy*, 83(10), 1139-1151.

[6] Gutiérrez, R. - Gutiérrez-Sáchez R. - Nafidi A. (2009), “Modelling and forecasting vehicle stocks using the trends of stochastic Gompertz diffusion models: the case of Spain”, *Applied Stochastic Models in Business and Industry*, 25, 385-405.

[7] Gutiérrez-Jáimez - Román P. - Romero D. - Serrano J.J. - Torres F. (2009), “A new Gompertz-type diffusion process with application to random growth”, *Mathematical Biosciences*, 208, 147-165.

[8] Iacus, S.M. (2008) *Simulation and Inference for Stochastic Differential Equations: with R examples*, Springer Series in Statistics, Springer NY, 300 pages, ISBN: 978-0-387-75838-1.

[9] Mankiw N.G. - Romer D.H. - Weil D.N. (1992), “A contribution to the empirics of Economic growth”, *Quarterly Journal of Economics*, 107, 407-437.

[10] Overman H.G - Puga D. (2002), “Unemployment clusters across Europe’s regions and countries”, *Economic Policy*, 34, 115-147.

[11] Quah D.T. (1996a), “Empirics for economic growth and convergence”, *European Economic Review*, 40, 1353-1375.

[12] Quah D.T. (1996b), “Regional convergence clusters across Europe”, *European Economic Review*, 40, 951-958.

[13] Quah D.T. (1997), “Empirics for growth and distribution: stratification, polarization, and convergence clubs”, *Journal of Economic Growth*, 2, 27-59.

[14] Ricciardi, L.M. (1977) *Diffusion processes and related topics in Biology*, Lecture Notes in Biomathematics, Springer-Verlag, Berlin.

[15] Sala-i-Martin X.X. (1996), "Regional cohesion: evidence and theories of regional growth and convergence", *European Economic Review*, 40, 1325-1352.

[16] Vieira, S. - Hoffmann, R. (1977), "Comparison of the logistic and the Gompertz growth functions considering additive and multiplicative error terms"., *Applied Statistics*, 26(2), 143–148.

[17] YUIMA Project Team (2012), "yuima: The YUIMA Project package", R package version 0.1.195. http://R-Forge.R-project.org/projects/yuima/