Convergence in Spanish provinces*

Antonio Montañés†, Lorena Olmos and Marcelo Reyes
University of Zaragoza, Gran Vía 2, 50005 Zaragoza (Spain)

May 2013
WORK IN PROGRESS. PLEASE DO NOT QUOTE OR CITE WITHOUT PERMISSION

Abstract

This paper analyses to what extent the Spanish provinces have shown a convergence process since 1980 to nowadays. The application of the Phillips-Sul methodology to the HDI recently developped by IVIE lead us to show that we cannot admit the exitence of a unique convergence process between them. Rather, we can observe the existence of 4 differentiated clubs. Higher school education and creativity plays determinant roles in explaining the evolution of these clubs. Improvements in these variables are vital for the provinces to move to the clubs with highest HDI values.

- JEL Codes: C22; E43
- Keywords: Stochastic Convergence; IDH; Convergence Clubs.

*The authors acknowledge the financial support of the MEC under project ECO2009-13085 and that of the excelence group ADETRE. The usual disclaimer applies.

†Corresponding author. email: amontane@posta.unizar.es

1 Introduction

The study of economic convergence has recently received a great deal of attention from applied economic researchers because it is a good way of assessing whether economies have diminished their disparities. After the seminal papers of Barro and Sala-i-Martí (1991, 1992), who based their results on the use of cross-sectional techniques, some authors such as Carlino and Mills (1993, 1996), Bernard and Durlauf (1995), Evans and Karras (1996), Loewy and Papell (1996), Nahar and Inder (2002), Strazicich et al. (2004) and Carrió-i-Silvestre and Soto (2007), amongst many others, have addressed the issue of economic convergence by studying the stochastic properties of some macroeconomic aggregates. The use of different unit root tests leads these authors to find mixed evidence about whether economies are converging and about the speed of convergence.

All these papers base their results on the use of the regional per capita GDP as the most appropriate indicator of the real situation of a particular economy. We should note, however, that the per capita GDP cannot capture some interesting aspects related to human welfare and economic potential, including environment, health, education and social integration. Consequently, it seems sensible to consider the use of multidimensional indexes in order to assess the evolution of a particular region or country. An interesting index is the Human Development Index (HDI), which has been published since 1990 by the United Nations Development Programme in its annual Human Development Report. This index provides information about the capacities of a particular economy and not about realisations as the simple GDP. Consequently, HDI can offer alternative results in the analysis of convergence between a group of economies.

Herrero et al. (2010) have calculated this index for the Spanish economy. These data have been used in Montañés and Olmos (2013) in order to study the convergence hypothesis. Two main results emerge in this paper. First, GDP data seems to offer more evidence of convergence thatn HDI data and, secondly, the evolution of the Spanish regions can be better understood as the addition of some divergent patterns of behaviour rather than the presence of a real convergence process. However, we should take into account that the use of regional data may not sometimes provided very useful insights. We are here thinking in those cases when the regional division includes infraregional territorial structures, such as provinces or cities, that behave heterogeneously. Given that the regional data as a weighter average of all of them, the conclusions based on regional data cannot be very informative. This is clear the Spanish case, whose regional structure is composed by a great variety of provinces/cities whose behaviour is far from being similar. For instance, we can consider the case of Aragon. This region has their main focus of activity in the cities of Huesca, Teruel and Zaragoza, which give the name to the three homonymous provinces. The population of the city of Zaragoza represents the 50% of the population of Aragon, wilst this percentage raises to a 72% if we consider the province of Zaragoza. Furthermore, the regional productive structure is not similar if we analyse it from a city perspective. For instance, the weight of manufacturing industry in Zaragoza almost doubles that of Huesca and Teruel. Simlarly, the contribution of Agriculture sector to the GDP is clearly higher in Huesca and, to very lesser extent, in Teruel than in Zaragoza. Therefore, the conclusions that emerge from a regional analysis cannot offer as richer insights as those obtained from a more disaggregated study, especially so far the active growth policies is concerned.

Against this background, the aim of this paper is to analyse the existence of a convergence process in Spain using the most disaggretated level of information possible. In the present case, we dispose data for the different provinces that compose the Kingdom of Spain for the HDI, but not for the GDP.

The rest of the paper is organized as follows. Section 2 describes the database.Section 3 presents the methodology that will be employed in the paper and the results obtained from the convergence analysis. Section 4 explores the sources that generate the different clubs of behaviour. The paper ends with a review of the most important insights.

2 The evolution of HDI in Spanish provinces: A brief review

As we have mentioned earlier, the aim of the paper is to analyse the degree of convergence between the Spanish regions. To that end, we prefer to use the HDI recently calculated by Herrero et al. (2010) rather than the regional per capita output, although we will employ both of them for comparative purposes. The HDI follows the United Nations recommendations and is based on the idea of Nobel Prize winner Amartya Sen of reflecting capabilities and opportunities more than realizations (see Sen, 1985, in this regard). It can be defined as follows:

$$HDI_i = \frac{1}{3}\, HI_i + \frac{1}{3}\, EI_i + \frac{1}{3}\, MWI_i, \quad i = 1, 2, ..., n$$

where HI, EI and MWI mean a Health Index, an Education Index and a Material Wellbeing Index, respectively. The HI depends on the life expectancy at birth (LE) and is defined as follows:

$$HI_i = \frac{LE_i - \min\ LE}{\max\ LE\ -\ \min\ LE}$$

the education index is obtained as follows:

$$EI_i = \frac{2}{3} LI_i + \frac{1}{3} GEI_i$$

where LI represents is the number of literates divided by the working-age population and GEI is the Gross Enrolment Rate divide by 100. Finally, the material wellbeing index can be defined as follows:

$$MWI_i = \frac{\log(GDP_i) - \log(100)}{\log(40,000) - \log(100)}$$

where GDP is the Gross Domestic Product of each province.

The values for the Spanish economy covers the 1980-2007 period and are available for both the regional and the provincial degree of disaggregation[1]. We should recall that Spain is divided into 17 regions[2]. Some of them are composed by a single province: Asturias (AST), Cantabria (CAB), Madrid (MAD), Región de Murcia (MUR), Navarra (NAV), La Rioja (LAR) and Islas Baleares (BAL). The rest contains different number of provincies: Galicia: La Coruña (COR), Pontevedra (PON), Orense (ORE) and Lugo (LUG); País Vasco: Alava (ALV), Vizcaya (VIZ) and Guipúzcoa (GUI); Aragón: Huesca (HUE), Teruel (TER), Zaragoza (ZAR); Cataluña; Lérida (LER), Gerona (GER), Barcelona (BAR), Tarragona (TAR); Castilla y León: León (LEO), Zamora (ZAM), Salamanca (SAL), Valladolid (VAD), Palencia (PAL), Burgos (BUR), Soria (SOR), Avila (AVI) and Segovia (SEG); Castilla-La Mancha: Guadalajara (GUA), Toledo (TOL), Albacete (ALB), Ciudad Real (CRE), Cuenca (CUE); Región de Valencia: Castellón (CAS), Valencia (VAL) and Alicante (ALC); Extremadura: Cáceres (CAC) and Badajoz (BAD); Andalucía: Almería (ALM), Cádiz (CAD), Córdoba (COR), Granada (GRA), Huelva (HUE), Jaén (JAE), Malaga (MAL) and

[1] The database was obtained from http://www.ivie.es/es/banco/desarrollo_humano.php

[2] The exact name is Autonomous comunuties, although we will use regions given that this is a much common denomination acrross the international studies.

Sevilla (SEV); Islas Canarias: Gran Canaria (GCA) and Tenerife (TEN). Additionally, there exist two autonomous cities Ceuta and Melilla, but we have preferred to omit from the study due to the small dimensions and their administrative peculiarities. Thus, we will base the study on the abovementioned 50 provinces.

Herrero et al (2010) analyze the evolution of the index for the different provinces from a descriptive point of view. They show that those provinces with low initial HDI values exhibit great growth rates. This is the case of ORE and BAD, with growth rates of 20% and 18%, respectively. This result is a necessary condition to prove the presence of convergence, with this hypothesis seeming to be reinforced by Figure 1, where we can observe the negative relationship between growth rates and intial HDI values. MOreover, if we regress both variables, what is comonly knows as a b-convergence analysis, the estimation of the slope is - 60, adding new evidence to the possible existence of convergence. However, a more careful analysis of this Figure led us to note that the provinces with high initial HDI values do not show low growth rates. For instance, only 6 provinces amongst those that the greatest initial HDI values exhibit growth rates below the average for the total Spanish economy. If we aditionally look at the Figures 2a-2c, where the Spanish map is coloured according to the values of the HDI in 1980, 1995 and 2007, new doubts raise about the convergence hypothesis, given that we cannot observe great variations between the three maps. Consequently, although the less developped provinces roughly show high growth rates, they cannot catch the most developped provinces up in that the former does not exhibit low growth rates that would balance the initial distance.This result can be interpreted by taking into acount the results of Montañés and Olmos (2013), where the evolution of the Spanish HDI seems to describe a sum of convergence/divergence forces, more than to exhibit a pure convergence episode. Thus, it seems to be sensible consider the existence of different clubs of convergence, which could explain the Spanish HDI behavior.

3 Testing for convergence: Methodology and results

This Section first discusses the methodology that we have employed for determining the presence of convergence. Most previous convergence analysis of the variable X over an objective value X* has been traditionally based on the study of the presence of a unit root in the ratio $\ln(X_{it}/X^*)$. However, we will employ the methodology recently proposed in Phillips and Sul (2007) given that if offers some advantages with respect to the standard unit-root based analysis and, furthermore, these authors develop a method for detecting the existence of convergence clubs, a quite feasible hypothesis according to the results of the previous Section. Later, we will reports the result that we have obtained.

3.1 Methodology

Following Phillips and Sul (2007), let us consider a time-varying factor representation of the variable under analysis:

$$hdi_{it} = b_{it}\ \lambda_t$$

where hdi_{it} is the log HDI of province i at time t and λ_t is a common house price factor such as common trend component in the panel. The relative share in λ_t of individual province i at time t is measured by b_{it} . We can decompose b_{it} into time-invariant mean and time-varying idiosyncratic error term and allow the variance of the error term to be time varying, which can state this as follows:

$$b_{it} = \ b_i + \ \xi_{it}\ \left[\frac{\sigma_i}{L\left(t\right)\ t^{\alpha_i}}\right]$$

where ξ_{it} is iid(0,1) across i but may be weakly dependent over t, σ_i is idiosyncratic scale parameter, and $L\left(t\right)$ is a slowly varying function.

Given this factor structure, a convergence analysis using panel unit root tests has no advantages over standard time-series unit root tests if the common component has a unit root because the nonstationary behavior is driven by the same univariate time series. Rather, we can employ the concept of relative convergence studied in Phillips and Sul (2007), which refers to the idea that the cross-sectional dispersion of HDI should display a tendency to decrease over time. The, relative convergence can be defined as

$$\lim_{T\to\infty}\frac{\log\ IDH_{it}}{\log\ IDH_{jt}} = 1 \text{ for all i and j}$$

The existence of convergence can be then analysed by testing the following null hypothesis:

$$H_o :\ b_i = b \ and \ \alpha_i \geq 0$$

To this purpose, Phillips ans Sul (2007) employ the quadratic distance measure, which can be defined as follows:

$$H_t = N^{-1} \sum_{i=1}^{N} (h_{it} - 1)^2$$

where h_{it} means the relative transition coefficient

$$h_{it} = \frac{hdi_{it}}{N^{-1} \sum_{i=1}^{N} hdi_{it}} = \frac{b_{it}}{N^{-1} \sum_{i=1}^{N} b_{it}}$$

which measures the transition element for individual i relative to the cross-section arithmetic mean, and can be obtained directly from the data. When there is a common transition behavior across individuals, $h_{it} = h_t$ across i. In particular, when there is ultimate convergence, $hdi \to 1$ for all i, as $t \to \infty$. Alternatively, under the convergence, $H_t \to 0$ as $t \to \infty$. On the other hand, when the convergence does not hold, the distance remain positive as $t \to \infty$.

Under the null hypothesis, the transition distance measure H_t has the limiting form of $H_t\tilde{}\ A\ \left[L(t)^2\ t^{2\alpha_i}\right]^{-1}$ 1 for some constant $A > 0$ as $t \to \infty$. Following Phillips and Sul (2007), we consider $L(t) = \log\ t$ and, consequently, the empirical log t regression model takes the form

$$\log \frac{H_1}{H_t} - 2\ \log\ (\log t)\ = a + \gamma\ \log t + \ \varepsilon_t \ \text{ for } t = rT,\ rT+1,\ ..., T \quad (1)$$

where $r \in (0.2, 0.3)$. Therefore, the null and alternative hypothesis can be transformed into $H_o : \gamma \geq 0$ and $H_A : \gamma < 0$.

We should note that the value of γ under the null hypothesis is the scaled speed of convergence parameter 2α, where $\alpha = \ \min\ (\alpha_i)$. Since the point estimate of γ converges to zero regardless of the true value of a under the alternative hypothesis, but the corresponding t-statistic diverges to negative infinity, a one-sided t-test of $\alpha \geq 0$ is used for the $log - t$ convergence test. For example, if the point estimate, $\hat{\gamma}$, is significantly less than zero, the null hypothesis is convergence is rejected at some significance level. Secondly, the term $2\ \log\ (\log t)$ acts like a penalty function. Without having this term in the regression model, under the alternative, the least squares estimator of γ becomes biased upward which impedes the test results in a finite sample. Finally, since the regression errors are serially correlated, the heteroskedasticity and autocorrelation consistent (HAC) estimator for the covariance of γ must be used to compute the t-statistic.

3.2 Results

This Section reports the results obtained when the $log-t$ statistic is employed to test for convergence between the Spanish provinces. These results are reported in Table 1, where we can be observed that the convergence hypothesis is strongly rejected for the several values of r taken into account. However, it is possible that the dynamics of the transtition coeffcient may exhibit some kind of convergence between related groups of provinces. Consequently, the presence of some clusters should be considered. a question which emerge from descpritive analysis as well as from previous papers, as we have already commented. In this regard, we can use the methodology described in Phillips and Sul (2007) which employs the that includes the stepwise application of $log-t$ regression tests since they has a discriminatory power against club convergence alternatives. This clustering procedure differs from conventional clustering methods, such as Durlauf and Johnson (1995) and Hobijn and Franses (2000), in the sense that it focuses on how idiosyncratic transitions behave over time in relation to the common factor component. The results of the application of this cluster methodology is reported in Table 2.

Table 2 shows that the data support the existence of 4 different clubs. The first club includes the provinces with the highest values of the HDI: ALV, GUI, VIZ, which form the Pais Vasco region, and MAD, the capital of Spain. These provinces increased their HDI in around 10 points during the sample size, showing a similar behaviour across the whole sample. The second group in composed by BAR, BUR, CRÑ, LAR, NAV, SAL, SEG, SOR, VAD and ZAR. These provinces are all placed in the North of Spain. LAR and NAV are regions composed by a single province, whilst BAR, VAD and ZAR are the capital of three regions (Cataluña, Castilla-León and Aragón, respectively). The provinces includes in this second club show high values of the HDI, although slightly lower than those of the provinces in the first club. The values of the HDI has increased 10 points on average, although the increment for the 1980-1994 period (6.4 points) has been greater than the one of the 1995-2007 period (3.8). The third club groups the provinces of ALB, AVI, BAD, GER, GUA, HSC, LEO, LER, LUG, ORE, AST, PAL, PON, CAN, TAR, TER, VAL and ZAM. They are mostly located in the North of Spain, with the exception of BAD and MUR, which are commonly considered as Southern provinces. VAL is the capital of the Comunidad Valenciana region and AST and CAN are single-province regions. The values of the HDI are lower of the provinces included in the previous clubs. They show an average increase of 10.8 points, standing out the cases of ORE and BAD which improves these values in 15.7 and 14.0 points, respectively.

However, this is not a time homogeneous behaviour, in that these provices has agumented 6.5 point during the 1980-1994 period and only 4.3 points for 1995-2007. The rest of the provinces are included in the club 4. Most of them are again placed in the South of the Iberic Peninsula, the exception is CAS, plus the provinces of the Canary and Balearic Islands. SEV and TOL are the capitals of Andalucia and Castilla-La Mancha regions and MUR is a single-province region. The HDI values are the lowest of all the provinces included in the sample and their evolution is similar to that of club 3, with a gain of 10.6 points on average for all the sample and time heteregenous behaviour in the sense that they grow 6.2 point during the 1980-1994 period and only 4.2 points during 1995-2007 period.

In order to better appreciate the differences between the different clubs, we have constructed a new index for each club based on the use of the arithmetical mean of the raw values of the provinces included in each club. These new indices are presented in Fig. 2. We can see as all the new indices grow across the sample, which implies that there have been a clear improvement in deveopment in Spain. The rythm of this increase seems to be similar, when a global perspective is used.. We can see it if we take into account the the average growth is 0.42, 0.41. 0.45 and 0.45, for clubs 1-4, respectively. However. there are clear differences if we split the sample in two periods. The growth rates for the 1980-1994 are 0.43, 0.51, 0.53 and 0.53, whilst are 0.41, 0.31, 0.36 and 0.38. This differences explain why we can appreciate that Club 2 reduces the distance with respect Club 1 during the period 1980-1994, becoming the distance of a single point in 1994, a half of the initial value. Nevertheless, this inital convergence is not maintained across the time and the distance is 2.3 points at the end of sample, bigger than the initial one. A similar pattern of behaviour can be observe if Club 3 and Club 4 are compared to Club 1. The intial distance is of 4.2 and 6.5 points, respectively. This is reduced during first part of the sample, getting minimum values in 1994 (3.2 and 5.7, respectively), whilst slightly increase in the second part, being 3.9 and 6.3 at the end of the sample. However, this reduction cannot allow us to admit the presence of a global convergence behaviour between club 3 and club 4 with respect club 1, given that the distance has maintained almost unaltered. Similar conclusion can be drawn if club 3 and club 4 are compared to club 2. In this case, the HDI are getting closer, with a reduction of 0.5 points with respect club 2. Crece en la primera parte de la muestra, pero en la segunda se reduce. Finally, the distance between club 3 and club 4 remains unaltered across the sample.

Thus, the conclusion that clearly emerges from this analysis is that there is an absence of convergence process between the HDI of the Spanish provinces. However, the results could be quite different if the 1980-1994 period was considered. We can appreciate a clear approach of the HDI to the values of Club 1 and, therefore, some degree of convergence could be found. However, the picture radically changes from 1995 and a divergent process can be seen, given that the growth speed of clubs 2, 3 and 4 slows, whilst club 1 maintains 1.

4 Forces that drive the clusters

In this study we have shown that the Spanish provinces are not converging . Rather, there exist some groups of provinces that exhibit a common pattern of behaviour, which allow us to consider the existence of several clubs. Clearly, it is of interest to investigate the characteristics of convergent and divergent subgroups as well as possible factors driving the convergence clubs. To that end, the first question to solve is the selection of those variables which have some potential to explain these forces.

Past research has demonstrated that climate is a significant and important determinant of labor supply (Rappaport, 2007). Assuming a pleasant climate is a normal good, increasing income over time should be making this amenity increasingly valuable. We use percent of daylight hours that are sunny as our measure of nice weather. To measure increasing labor demand in a city, we use the percent of those in 1990 over age 25 with a college degree from the 1990 Census. This variable has been found to be the most reliable determinant of future labor demand growth (Glaeser and Saiz, 2003; Moretti, 2004).

Table 4 summarizes the average values of the explanative variables for all the provinces included in each cluster. As we can see, the provinces in Club 4 are characterised by a high number of sunny days, combined with the lowest values of the percentage of population with high school studies and number of books edited per inhabitant. Clubs 2 and 3 present similar values of the number of sunny days, although are sligthly lower that those of Club 4. These two groups exhibit great differences in both the creativity and in the education indexes, being the average values of Club 2 clearly greater than those of Club 3. Finally, Club 1 shows the lowest value of sunny days, but the highest values of the rest of the variables.

To analyze a the complex interaction between explanative variables and club membership, a orderded model is used to predict how provincial characteristics affect the likelihood that any given city would be found to be a member of each convergence club. To explain the structure of the model, we should note that the dependent variable y_i has 4 possible outcomes, each of them related to each club we have obtained. Consequently, we have that:

$$y_i = m \text{ for } m = 1, ..., 4$$

The observed values are assumed to derive from some unobservable latent variable y_i^* where

$$y_i^* = \; x_i^{'}\beta + u_i \;, i = 1, 2, ..., 17$$

for some $k \times 1$ parameter vector β and (univariate) stochastic disturbance term u_i.

We could interpret that the different m values does imply a preference or an ordenation of the clubs. However, it is clear that provinces would prefer to belong to $m+1$ club, better than to m club, in that this implies a higher level of development. Therefore, the possible outcomes suggest a natural ordering. Thus, the observed variable y_i are assumed to be related to the latent variable through the following observability criterion:

$$y_i = m, \text{ if } \alpha_{m-1} \leq y_i^* \leq \alpha_m$$

for a set of parameters α_0 to α_M , where $\alpha_0 < \alpha_1 < \alpha_2 ... < \alpha_M$, with $\alpha_0 = -\infty$ and $\alpha_M = \infty$. Then, the conditional probability of observing the $m - th$ category can be written as:

$$\begin{aligned} \Pr\left(y_i = m \; / \; x_i\right) &= \Pr\left(\alpha_{m-1} \leq y_i^* \leq \alpha_m\right) \; = \Pr\left(\alpha_{m-1} \leq x_i^{'}\beta + u_i \leq \alpha_m\right) = \\ &= \Pr\left(u_i \leq \alpha_m - x_i^{'}\beta\right) - \Pr\left(u_i \leq \alpha_{m-1} - x_i^{'}\beta\right) \end{aligned}$$

for $m = 1, .., 4$. We can now see the importance of the relationship between the $\alpha' s$ in order for this probability to be strictly positive. To evaluate the conditional probability, we are required to make some distributional assumption for the disturbance term u_i. In this present case, we assume a logistic distribution, yielding the Ordered Logit model. Table 5 reports the results of the estimation of this model.

We can observe that the estimated model includes the sunny days, the creativity index and the education index. Furhermore, we can see that the education index has the highest estimated value and, consequently, this variable show the biggest influence on the creation of the Clubs. The second most influent variable is the creativity index, whilst the sunny days is significant, but the estimated value of this parameter is low and so is its relevance.

In order to better appreciate the magnitude of the estimated effects on the HDI, we have graphed the predicted probability of being in each club for an average city[3] allowing each explanatory variable to vary in turn. The graphical results of these simulations are shown in Figures 3-5. For Fig. 3 we created a city with average creatitivy and sunny days and then allowed education to vary across the range we see in the data (roughly, 9–30% of adults with college studies). Given the comparatively big estimated coefficient of this variable, it is not surprising to see that even small changes in the percent of adults with college studies affects club membership. Thus, we can conclude that this variable is playing a crucial role in determining the club membership. For instance, it is required a percentage of adults with college studies greater than 22% (1,5 times the average value of the sample) for the probability of an average city of being in club 1 to exceed the 10%. We can also conclude that an increase of a 1% in this percentage, diminished almost a 10% the probability of being in club 4, whilst an increase slightly greater than a 3% reduces this probability below 50%. If we also analyze Figures 4 and 5, we can appreciate the probability of being in club 1 or in club 4 to be very small for an average city. Thus, the changes in the explanatory variables mainly affect to the probability of being in clubs 2 and 3. In this regard, we should note that an augment in creativity increases the probability of being in club 2, abandoning club 3. By contrast, the higher the number of sunny days, the lower the probability of being in club 2.

Taken as a whole, the analysis of HDI clubs reveals that the education variable is playing a really key role in determining the asignation of a province of a particular club. Consequently, policies to improve the development of the Spanish provinces should base in a proper education policy, specially devoted to promote the realisation of college studies. Alternative measures could help to enhance the relative position, but will not imply a significative upgrade of this province.

5 Conclusions

This paper studies the convergence hypothesis for the Spanish provinces by way of the use of the HDI. This index takes into account the evolution of the per capita GDP, the education level and the health level. Consequently, we

[3]The estimated probability of this average city of being in club 3 is 0.7, a 17% for club 2 and 13% for club 3.

consider it to better represent the evolution of a particular zone than the per capita GDP, the most employed measure in previous convergence analysis.

The convergence analysis has been carried out by the use of the methodology recently proposed in Phillips and Sul (2007, 2009). This method allows us to determine the existence of convergence, but also to detect the possible existence of convergence clubs, whenever the existence of a unique convergence process is rejected. This is our case, given that we can reject the null hypothesis of convergence, but we can admit the existence of 4 differentiated clubs, each of them showing a different pattern of behaviour. Club 1 includes the provinces with highest values of the HDI. Club 2 reflects provinces with high values of the HDI, most of them placed around the Ebro River Corridor, and also characterised by a high values of education and creativity. Club 4 engloba most of the Southern provinces, all of them showing very low values of HDI. The rest of the provinces belongs to the Club 3. We can also observe that these clubs may exhibit some kind of convergence during the 1980-1994 period, in the sense that the HDI values get closer. However, the differences clearly increase after this period.

We have also found that some few variables can help us to better undestand the creation of these Clubs and, consequently, lead provinces to move to upper clubs. The most relevant vairable is the percentage of population with high studies, followed by a creativity index. Improvements in both variable increases the probabiilty of beloging to the clubs with highest values of the HDI. Consequently, a education and creativity are the key factor in the development of the Spanish provinces.

References

[1] Bernard AB, Durlauf SN (1995) Convergence in international output. J Appl Econ 10:97–108

[2] Carlino GA, Mills L (1993) Are U.S. regional incomes converging? A time series analysis. J Mon Econ, 32:335–346

[3] Carlino GA, Mills L (1996) Convergence and the U.S. States: a time-series analysis. J Reg Sci 36(4):,597–616

[4] Carrion-i-Silvestre JL, German-Soto V (2007) Stochastic convergence amongst Mexican states. Reg Stud, 41(4):531–541

[5] Carrion-i Silvestre, J.L., D. Kim, and P. Perron (2009). "GLS-Based Unit Root Tests with Multiple Structural Breaks Under Both the Null and the Alternative Hypotheses", Econometric Theory, 25(6): 1754–1792.

[6] Dickey DA, FullerWA (1979) Distribution of the estimators for autoregressive time series with a unit root., J Am Stat Assoc 74:427–431

[7] Evans P, Karras G (1996) Convergence revisited. J Mon Econ 37:249–265

[8] Herrero, C., Martínez, R. and Villar, A. (2010a). "Improving the Measurement of Human Development", Human Development Research Papers, New York: United Nations Development Program.

[9] Herrero, C. Martínez, R. and Villar, A. (2010b). "Multidimensional social evaluation: an application to the measurement of Human Development", Review of Income and Wealth 56, 483-497.

[10] Herrero, C., Soler, Á. and Villar A. (2010). Desarrollo Humano en España (1980-2007), Valencia: Ivie-Fundación Bancaja,.

[11] Loewy MB, Papell DH (1996) Are U.S. regional incomes converging? Some further evidence. JMon Econ, 38:587–598

[12] Lumsdaine RL, Papell DH (1997) Multiple trend breaks and the unit-root hypothesis. Rev Econ Stat 79:, 212–218

[13] Mazumdar, K. (2002): A Note on Cross-Country Divergence in Standard of Living, Applied Economics Letters, vol. 9, pp. 87-90.

[14] Montañés A, Olloqui I, Calvo E (2005) Selection of the break in the Perron-type test. J Econom 129:41–64

[15] Nahar S, Inder B (2002) Testing convergence in economic growth for OECD countries. Appl Econ 34: 2011–2022

[16] Nelson CR, Plosser CI (1982) Trends and random walks in macroeconomic time series. Some evidence and implications. J Mon Econ 10:139–162

[17] Ng S, Perron P (1995) Unit root test in ARMA models with data-dependent methods for the selection of the truncation lag. J Am Stat Assoc 90:268–281

[18] Perron P (1989) The Great Crash, the oil price shock and the unit root hypothesis. Econometrica 57:1361–1401

[19] Perron P (1997) Further evidence on breaking trend functions in macroeconomic variables. J Econom 80:355–385

[20] Perron P, Vogelsang T (1992) Nonstationarity and level shifts with an application to Purchasing Power Parity. J Bus Econ Stat 10:301–320

[21] Perron P, Yabu T (2009) Testing for shifts in trend with an integrated or stationary noise component. J Bus Econ Stat (in press)

[22] Pesaran MH (2007) A pair-wise approach to testing for output and growth convergence. J Econom 138: 312–355

[23] Strazicich MC, Lee J, Day E (2004) Are incomes converging among OECDcountries? Time series evidence with two structural breaks. J Macroecon 26:131–145

[24] Zivot E, Andrews DWK (1992) Further evidence on the Great Crash, the oil price shock, and the unit-root hypothesis. J Bus Econ Stat 10:251–270

Table 1. Testing for convergence

	$r = 0.20$	$r = 0.23$	$r = 0.26$	$r = 0.30$
$\hat{\gamma}$	-1.72	-1.77	-1.82	-1.87
$t_{\hat{\gamma}}$	-49.72^{a}	-47.71^{a}	-46.19^{a}	-44.96^{a}

This table reflects the values of the log-t statistic when regression (1) is estimated for the Spanish

provinces for several values of the parameter r .

[a] denotes rejection at 1% significance level.

Table 2. Convergence club classification

Club	$\hat{\gamma}$	$t_{\hat{\gamma}}$	Provinces
Club 1 [4]	0.02	0.18	ALV DON MAD BIL
Club 2 [10]	0.09	0.67	BAR BUR CRÑ LOG PAM SAL SEG SOR VAD ZAR
Club 3 [18]	0.16	1.57	ALB AVI BAD GIR GUA HSC LEO LLE LUG OUR OVI PA
Club 4 [18]	-0.11	-1.49	ALC ALM PMA CAC CAD CAS CDR COR CUE GRA HU

Table 3. Testing for club merging

Merging Clubs	$\hat{\gamma}$	$t_{\hat{\gamma}}$
Club 1+2	-2.44	-18.93^{a}
Club 2+3	-1.50	-81.83^{a}
Club 3+4	-1.28	-77.65^{a}

Table 4. Summary statistics of key club determinants

Club	Climate	Beds	Creativity	Education
Club 1	$2,026$	42.7	5.2	24.5
Club 2	$2,508$	41.3	3.6	18.3
Club 3	$2,465$	43.5	0.8	14.0
Club 4	$2,871$	31.4	0.5	12.9

Table 5. Ordered Logit Estimation

Explanative Variables	$\hat{\beta}$	$t_{\hat{\beta}}$
Education	-0.61	-4.6
Creativity	-0.21	-2.5
Climate	0.003	2.4

-8.46

-3.42

0.03

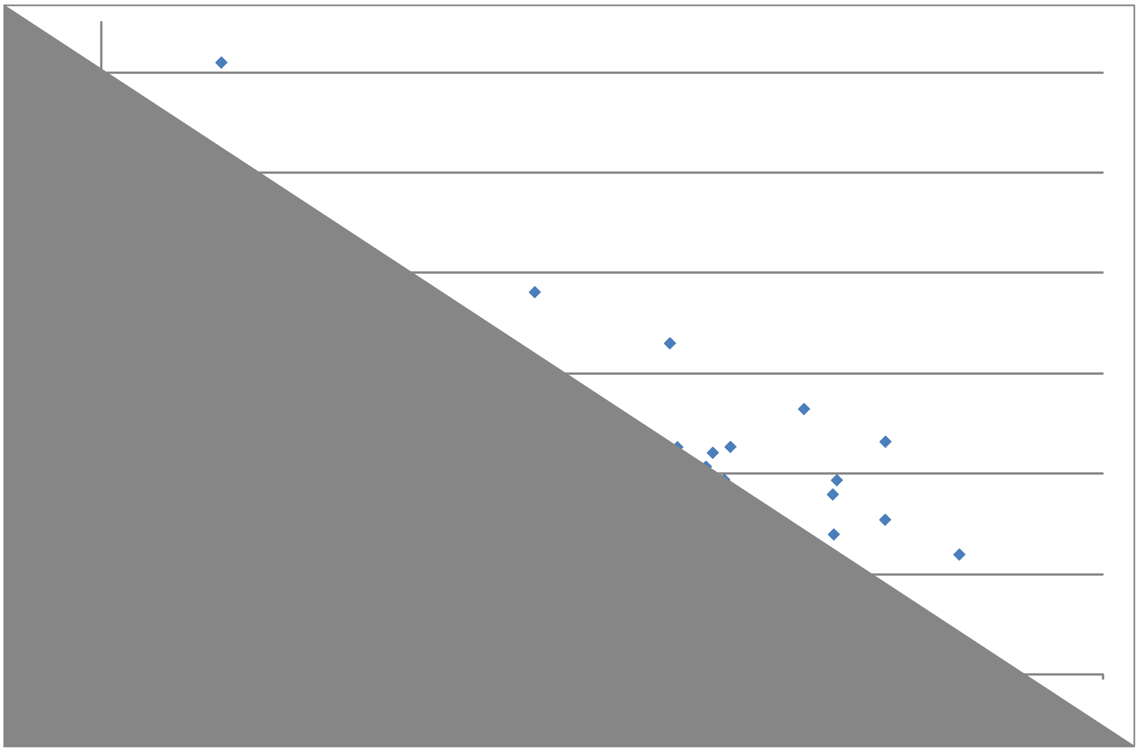


Figure 1. B-convergence

Figure 2a. HDI distribution. 1980

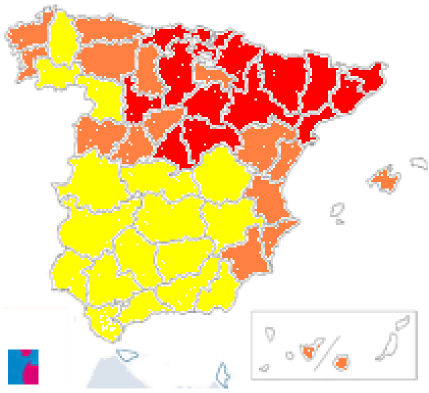


Figure 2a. HDI distribution. 1995

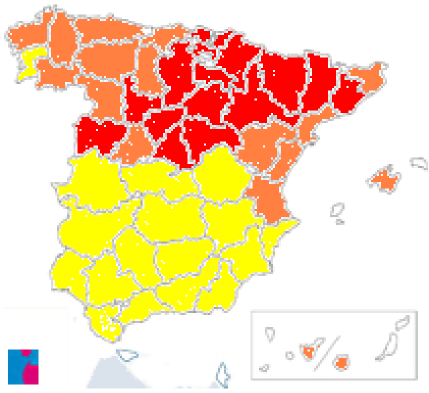
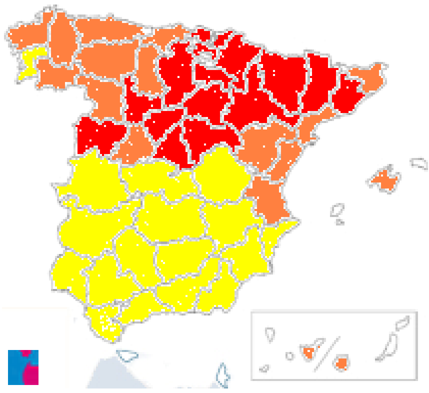


Figure 2a. HDI distribution. 2007



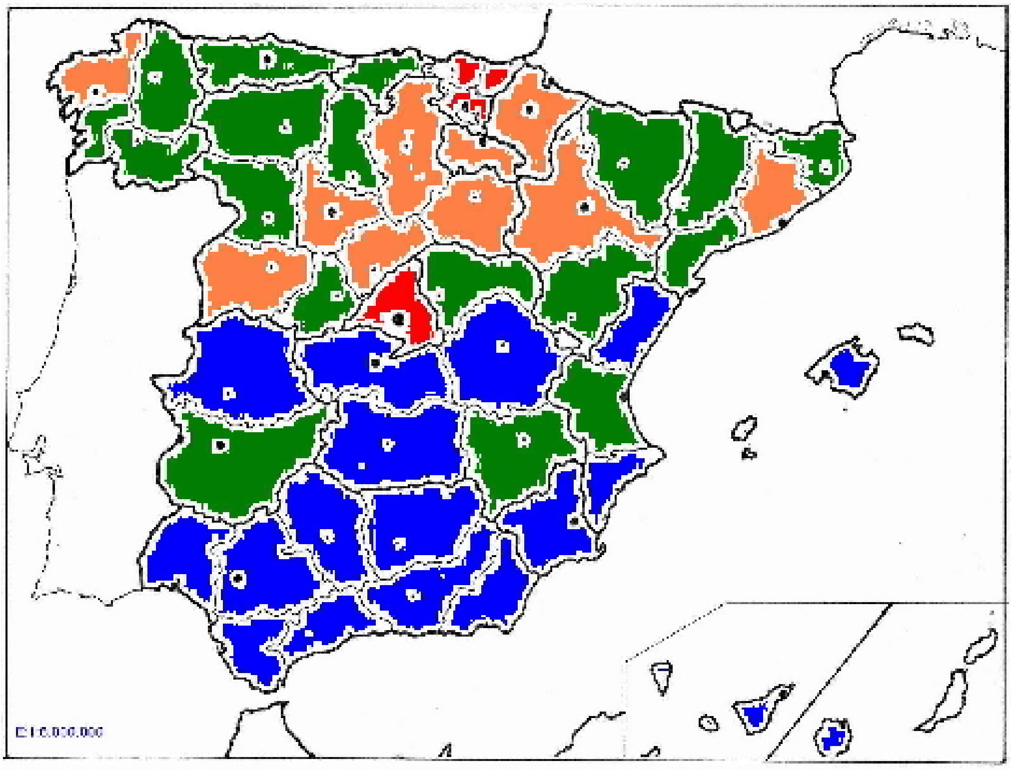
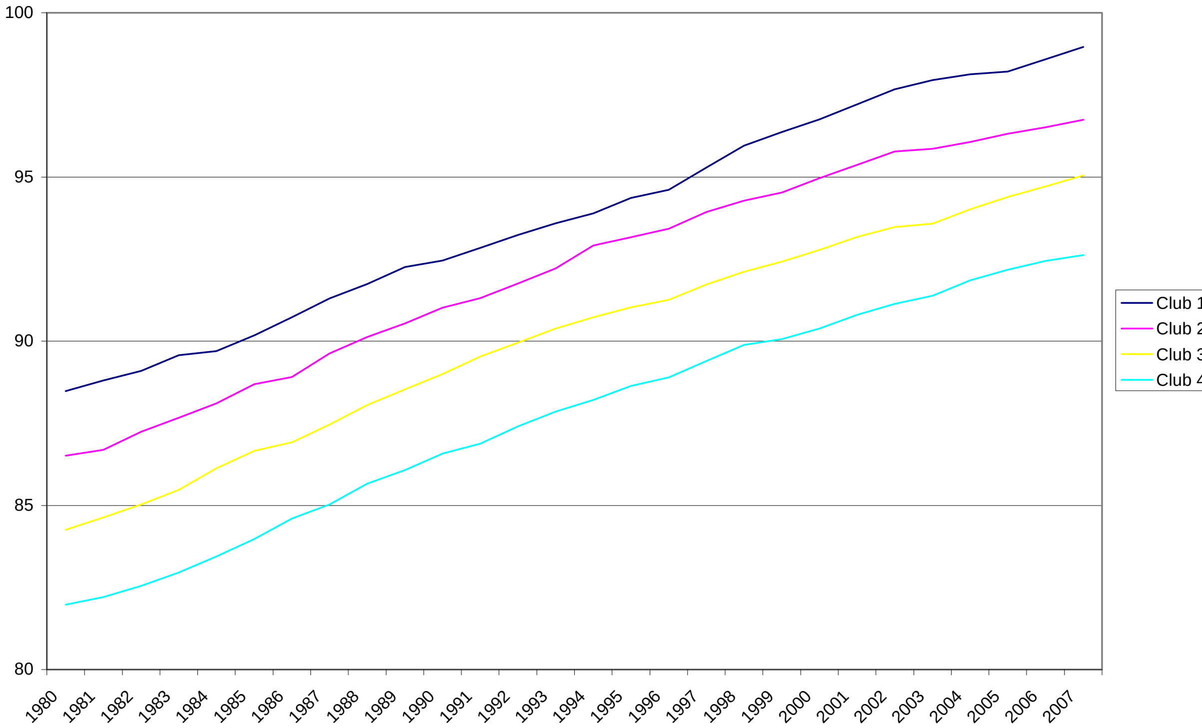


Fig 2. Club Indices



Note: This figure plots arithmetic mean of HDI of the provinces included in the respetive cl

Figure 3. Simulated effect of a change in percentage of adult with college studies on the like

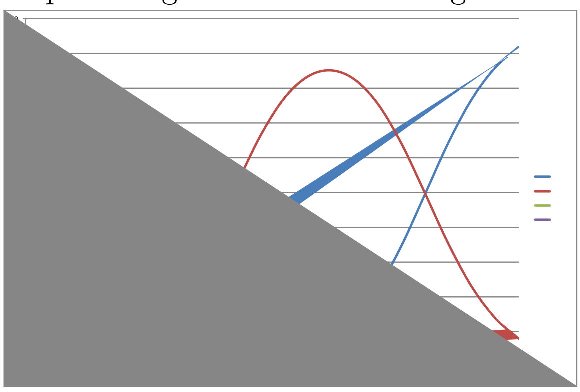


Figure 4. Simulated effect of a change in creativity on the likelihood of club membership

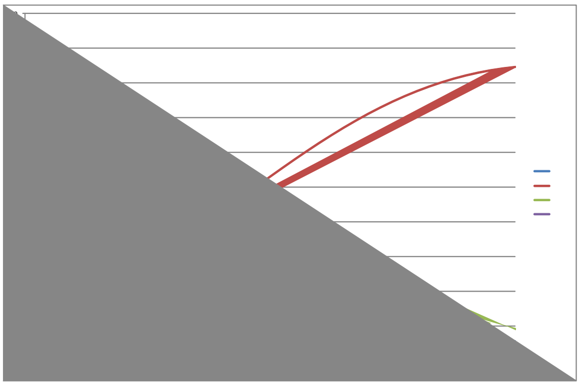


Figure 5. Simulated effect of a change in the number of sunny days on the likelihood of clu