Applied Economics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/raec20

# Do Spanish regions converge? A time-series approach using fractional cointegration

Mariam Kamal & Josu Arteche

**To cite this article:** Mariam Kamal & Josu Arteche (02 Jan 2024): Do Spanish regions converge? A time-series approach using fractional cointegration, Applied Economics, DOI: 10.1080/00036846.2023.2293089

**To link to this article:** https://doi.org/10.1080/00036846.2023.2293089



Published online: 02 Jan 2024.

APPLIED ECONOMICS
https://doi.org/10.1080/00036846.2023.2293089

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

# Do Spanish regions converge? A time-series approach using fractional cointegration

Mariam Kamal and Josu Arteche

Department of Quantitative Methods, University of the Basque Country, Bilbao, Spain

**ABSTRACT**
This article investigates economic convergence in terms of real income per capita between the autonomous regions of Spain over the period 1955–2020. In order to converge, the series should be cointegrated. This necessary condition is checked using two testing strategies recently proposed for fractional cointegration, finding no evidence of cointegration, which rules out the possibility of convergence between all or some of the Spanish regions. As an additional contribution, an extension of the critical values of Nielsen's (2010) test of fractional cointegration is provided for a different number of variables and sample sizes from those originally provided by the author, fitting those considered in this article.

**KEYWORDS**
Fractional integration; fractional cointegration; long memory; persistence

**JEL CLASSIFICATION**
C12; C22; C32

## I. Introduction

Economic convergence has been one of the main focal points of the empirical literature on economic growth. It implies that income gaps between countries/regions tend to disappear, hence involving convergence to a single steady state (equilibrium).

The European economy has become more integrated in recent decades, with the states following a converging path due to economic, political, and institutional factors, such as, for example, the exchange rate mechanism in 1979, and the introduction of the euro in 2001. Numerous empirical studies have provided evidence of this integration in the European Union (Beckfield 2006; Caporaso and Pelowski 1971; Martin and Ross 2004). However, this integration among states can come together with economic disparities between the different regions, which may cause non-convergence within a country. This article analyses this possibility in Spain.

Several countries have typically employed regional policies to address structural disparities among their geographical areas. Spain, for example, began implementing regional policies in the early 1960s. However, since 1986, Spanish regional policies have undergone significant changes due to its inclusion in the European Union (EU), which has been particularly important in providing regional governments with opportunities to engage in European networks, facilitating the exchange of interests, knowledge, and values. As argued by Arregui (2020), Spain is likely to be one of the members where some state restructuring has taken place, both at national and regional level. This transformation has been influenced by both European integration and the decentralization of political power. These two processes have mutually reinforced each other, and Spain's EU membership has solidified the role of Spanish Autonomous Communities in shaping and implementing policies in crucial areas such as environment, agriculture, or fishing policies (Arregui 2020).

Spanish regions are divided into 17 Autonomous Communities. Some of these regions are richer than others due to their economic or sector specialization and disaggregation according to branches of activity. The income of each Autonomous Community depends on the economic specialization of that region, with some specializations generating low incomes, while others generate significantly higher incomes. Table 1 shows the high-sector heterogeneity presented by the different Spanish regional economies. These regions span from those experiencing substantial growth driven by tourism-related

**CONTACT** Mariam Kamal mariam.kamal@ehu.eus Department of Quantitative Methods at the University of the Basque Country, Bilbao 48015, Spain
Supplemental data for this article can be accessed online at https://doi.org/10.1080/00036846.2023.2293089

Table 1. Characteristics and productive structures of the Spanish regions.

| | | | |
|---|---|---|---|
| **Andalucia** | Agri-food sector Transport and logistics sector | **Extremadura** | Agri-food sector. Livestock farming Food Industry |
| **Aragón** | Automotive industry<br>Transport and logistics sector | **Galicia** | Textile and automotive sector<br>Agri-food sector |
| **Asturias** | Metal and mining sector | **Madrid** | Biomedical and pharmaceutical companies<br>Information and Communication Technology sector (ICT)<br>Logistics and transportation<br>Aerospace industry |
| **Balearic Islands** | Tourism sector<br>Food and catering industry<br>Fashion industry | **Murcia** | Agricultural sector<br>Plastic sector |
| **Canary Islands** | Tourism Sector<br>Cultural industries<br>Logistics sector | **Navarre** | Automotive sector<br>Biomedical cluster<br>ICT sector |
| **Cantabria** | Agri-food sector<br>Automotive components<br>Biotechnology and health | **Basque Country** | Energy sector<br>Automotive and aeronautic sector<br>Maritime industry<br>ICT sector<br>Bio-health sector<br>Service sector |
| **Catalonia** | Biotechnology<br>Petrochemical sector<br>Automotive sector<br>Agricultural sector | **La Rioja** | Agri-food sector<br>Footwear sector<br>Automotive sector<br>Service sector |
| **Castilla-La Mancha** | Agri-food sector<br>Wine production | **Valencia** | Automotive and capital goods sector<br>Agri-food sector<br>ICT and services sector<br>Chemical and pharmaceutical sector<br>Plastic sector |
| **Castilla-León** | Agri-food sector<br>Chemical-Pharmaceutical sector | | |

activities (e.g. Balearic Islands, Canary Islands) to those where economic activity is still largely dependent on primary sectors (e.g. Galicia, La Rioja, Murcia, Andalucía). This heterogeneity in the regional economic structures and sources of income generates regional disparities that may prevent convergence, as each region's growth may be driven by different factors and industries.

The different ways of generating income among different regions have produced the current regional differences in productivity and income. Hence, it might be unsurprising to find disparities between the 17 Autonomous Communities. This prompts us to question whether economic convergence among the 17 Autonomous Communities could not occur, which is a question that drives the current research. In addition, we also aim to determine whether some converging subgroups can be identified, for example, among developed or less developed regions, which may be used to define more efficient regional policies delimiting their geographical impact.

As a first example of this heterogeneity across the Spanish regions, Figure 1 shows the evolution of the cross-sectional standard deviations for all the logs of per capita income in the 17 Autonomous Communities in Spain from 1955 to 2020. The dispersion begins in 1955 at around 0.91 and rises and declines over time ending at around 0.94 in 2020, confirming the absence of sigma convergence.

The large dispersion shown in the graph confirms our previous suspicion of possible heterogeneity and non-convergence. Note that the dispersions in the decade of 1970 and 2010 are the highest. However, the heterogeneity in Spain remains in the sample before and after these peaks, which may hinder regional convergence.

In the literature on economic growth, there are three main definitions of convergence: (i) beta convergence, (ii) sigma convergence and (iii) stochastic convergence based on time-series analysis. As the first two have several statistical problems (see Durlauf 2000; Friedman 1992; Quah 1993), we will focus on the time-series approach on cointegration, as suggested by Bernard and Durlauf (1995). The use of different techniques has usually led to different conclusions about the existence of convergence (see, for example, Durlauf 2000). We follow a time-series approach to test for output

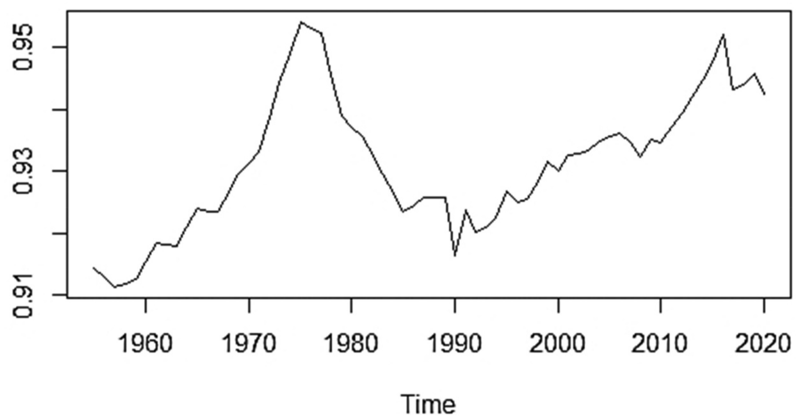


**Figure 1.** Cross sectional standard deviation of the log of per capita income.

convergence, paying particular attention to the analysis of cointegration, which provides a natural setting for testing relations between variables (see Bernard and Durlauf 1995, 1996; Durlauf 2000; Evans 1996; Quah 1993).

According to the time-series approach of Bernard and Durlauf, 1995, 1996), two series converge if the following conditions are satisfied: (i) The variables are cointegrated, (ii) the cointegrating vector is (1, −1), and (iii) the difference between the series is a stochastic variable with zero mean. Based on these conditions, the notion of convergence can be divided into strong and weak convergence (defined as catching-up in the convergence literature). If conditions (i) and (ii) are fulfilled, the series are cointegrated with cointegrating vector $[1,-1]$, but the difference between them is a stochastic variable with a mean different from zero, which suggests that the deviation between the series is expected to decrease, but not to disappear. This is weak convergence, i.e. catching up, which refers to the situation in which narrowing of the differences between the variables is observed over time, but the convergence process has yet to be complete. If all conditions (i), (ii), and (iii) are fulfilled there is strong convergence because the difference between the variables vanishes. Therefore, if there is no cointegration, convergence does not occur, neither weak nor strong.

The rest of the article is organized as follows. *Section II* provides a literature review on convergence in Spain. *Section III* explains the methodologies used in our analysis. *Section IV* contains the data and presents the results, and finally, *Section 5* presents the conclusions.

## II. Literature review

There are few studies that specifically examine output convergence between Spanish regions. The majority of them use cross-regional analysis approaches to estimate beta convergence. Meanwhile, those using a time-series approach opt for unit root test as the Augmented Dickey Fuller test in a non-fractionally integrated context, which lacks the power and flexibility needed for a comprehensive analysis.

Some authors have found results indicating non-convergence in the Spanish regions in agreement with our results. Martínez-Argüelles and Rubiera-Morollón (1998), focusing solely on the service sector, identify distinct regional growth patterns within this sector using integer cointegration techniques. Cuadrado-Roura et al. (1999) investigate the evolution of regional differences in Spain and use an analysis of beta and sigma convergence to conclude that the primary source of convergence in observed productivity is the alignment of regional sectorial structures. Lamo (2000) examines output convergence across Spanish regions using cross-sectional distribution dynamics. She finds no evidence of income convergence. Maza (2006) examines the phenomenon of regional convergence in per capita income in Spain and studies what factors influence migration patterns within these regions. Using a beta convergence analysis, he concludes that there is no convergence among Spanish regions because migrants tend to move towards regions with higher per capita income, inducing a slower pace of regional convergence in Spain. Arroyo et al. (2013) examine the pairwise convergence hypothesis among the 17 Spanish regions using the Augmented Dickey and Fuller (1979) test for unit roots. The findings reveal incomplete catching-up in many instances, with only four converging regions (Andalucia, Extremadura, Castilla-La Mancha, and Galicia) and just one (Baleares) converging with the European Union. Puente (2017) uses traditional growth regressions to analyse the existence of beta-convergence. The results indicate that labour productivity convergence stands out as the primary driver in narrowing regional income disparities. On the contrary, labour market variables such as employment and unemployment, as well as total factor productivity, do not have a substantial impact on diminishing regional disparities during the period under analysis.

Other studies have focused on the convergence among provinces rather than Autonomous Communities. Dolado et al. (1994) examine the growth and disparities across Spanish provinces. They use traditional cross-regional analysis to estimate beta convergence, and find evidence of provincial convergence, although with some signs of instability in the speed of convergence during specific subperiods. Gardeazábal (1996) analyses the dynamic evolution of income distribution among Spanish provinces. Using Markov processes, he concludes that per capita incomes among Spanish provinces converge towards equilibrium. Villaverde (2005) examines the existence of beta convergence in labour productivity in the provinces of Spain. He concludes that Spanish provinces with low (high) relative productivity tend to be geographically close to each other, indicating a concentration of productivity. The convergence process does occur, but at a slightly slower pace than in the classical model, and there is a gap that separates the provinces from their steady state. Hierro and Maza (2010) investigate the role played by internal migration of foreign individuals in the income convergence of provinces in Spain between 1996 and 2005. Their results refute the hypothesis that internal migration of the foreign-born influences income convergence. Montañés et al. (2018) investigate convergence between Spanish provinces, with a particular focus on the impact of the recent international crisis. Their results indicate the formation of several convergence clubs, the patterns of which were altered by the 2007 crisis. Tapia and Galarraga (2020) investigate the empirical connection between economic growth and inequality, quantifying the disparities between Spanish provinces for various reference years spanning from 1860 to 1930 and concluding that the growth of income did not directly lead to a reduction in inequality.

This article analyses convergence in annual real output per capita of the 17 Autonomous Communities in Spain from 1955 to 2020. We contribute to the existing empirical literature in three main dimensions:

- We test cointegration in an economic framework of convergence that follows the Bernard and Durlauf (1995, 1996) definition of time-series convergence.
- We use semiparametric and nonparametric techniques, which have never been used before to analyse regional convergence, to test for fractional cointegration: the strategies proposed by Robinson (2008), Hualde (2012) and Nielsen (2010). These fractional integration and cointegration techniques are used to avoid the low power of traditional unit root and cointegration tests against fractional alternatives and are more reliable to explore economic convergence.
- We complement Nielsen (2010) test with a new set of critical values of independent interest for practitioners. In particular, we provide critical values for up to 17 variables and three different sample sizes $T = 66$, $T = 150$ *and* $T = 1000$ for two different values of the memory parameter of the original series: $d = 1$ and $d = 1.4$. The latter corresponds to the values found in the series analysed here, while the former (in the supplementary material) corresponds to the traditional unit root case.

## III. Methodology

The methodology used in this article is based on the concepts of fractional integration and fractional cointegration.

### Fractional integration and cointegration

The idea of fractional integration was introduced by Granger and Joyeux (1980), Granger (1980, 1981) and Hosking (1981) allowing a continuous transition from non-unit to unit root behaviours, offering a more flexible context for the modelling of long-run persistence. A time-series $\{y_t, t = 1, 2, 3, \ldots\}$ is (fractionally) integrated of order d, $I(d)$, if it satisfies:

$$(1 - L)^d y_t = u_t, t = 0, \pm 1, \ldots, \quad (1)$$

where d is the memory parameter and $u_t \sim I(0)$, meaning that $u_t$has a finite variance and a spectral density function $f(w)$, satisfying $0 < f(w) < \infty$. If$d = 0, y_t = u_t$ and $y_t$ is short memory; if $0 < d < \frac{1}{2}$, $y_t$ is said to be long memory. Finally, if $-\frac{1}{2} < d < 0$, $y_t$presents anti-persistence. Also, if $d < 0.5$, $y_t$ is covariance stationary. However, a value $d \geq 0.5$ implies non-stationarity, but if $d < 1$, the series is mean reverting. In addition, if $d = 1$, the series has a unit root. If $d < 1$ the effects of the shocks disappear in the long-run and if $d \geq 1$ the shocks persist indefinitely.

Note that $u_t$ in (1) may include some type of weak dependence in the form of, for example, a stationary and invertible autoregressive moving average (ARMA) process:

$$\Phi(L)u_t = \theta(L)\varepsilon_t, t = 0, \pm 1, \ldots, \quad (2)$$

where $\varepsilon_t$ is an independent and identically distributed (*iid*) sequence. In this case, $y_t$ in (1) is an Auto-Regressive Fractionally Integrated Moving Average (ARFIMA) process:

$$\Phi(L)(1 - L)^d y_t = \theta(L)\varepsilon_t, t = 0, \pm 1, \quad (3)$$

Engle and Granger (1987) defined cointegration as follows: "*A vector $y_t$ is said to be co-integrated of order d, b, denoted $y_t \tilde{C}I(d, b)$, if the components of $y_t$ are $I(d)$ and there exists a vector $\alpha(\neq 0)$ such that $z_t = \alpha' y_t \tilde{I}(d - b), b > 0$. The vector $\alpha$ is called the co-integrating vector and b denotes the degree of cointegration*".

The original testing strategies proposed for cointegration were only suitable for bivariate settings, and they thus could only identify one cointegration vector. Johansen (1988, 1991, 1995) developed a maximum likelihood approach for testing cointegration in a multivariate setting, allowing for several relations and determining the rank of cointegration. Following these pioneering authors, other standard techniques were developed by Phillips and Ouliaris (1990), Harris (1997), Bierens (1997), and Breitung (2002), among others. The generalization of the traditional Johansen test to a fractional context was proposed by Johansen (2008) and Johansen and Nielsen (2010, 2012, 2014) with the fractionally cointegrated vector autoregressive (FCVAR) model.

Standard traditional cointegration is just one particular case of fractional cointegration where the memory parameters $d$ and the degree of cointegration $b$ are restricted to be integer values. Fractional values of $d$ and $b$ allow more flexibility and are good alternatives because many economic series are known to exhibit non-stationary behaviours that may not be exactly I(1), and there is also no need to assume that the equilibrium relation is exactly I(0).

### Testing for fractional cointegration

The strategy we follow is based on the estimation of the cointegration rank in a fractional setting using two different and flexible techniques with good asymptotic properties under mild conditions. First, the methodology proposed by Nielsen (2010) has the following advantages over other cointegration tests: (i) The test statistic is computed without prior knowledge of the order of integration of the series. (ii) Since the test is nonparametric, it does not require specification of a particular model and is invariant to short-run dynamics. This is important because mis-specified short-run dynamics may lead to inconsistent estimation and hence to erroneous inference regarding the cointegration rank in other parametric techniques. (iii) The proposed test has good power for large and small samples.

Second, the methodology offered by Hualde (2012), together with the testing strategy used by Robinson (2008), is characterized by the following benefits: (i) The testing strategies in Robinson (2008) do not require estimation of any cointegrating relations or prior selection of any tuning numbers beyond one bandwidth parameter. (ii) Hualde (2012) proposes a procedure to estimate the rank of cointegration in multivariate fractional series, and therefore this can be implemented together with the procedure in Robinson (2008) to infer the dimension of the possible cointegrating subspaces. (iii) The combination of both techniques allows for precise detection of the common trends.

### *Robinson (2008) and Hualde's (2012) approaches*

We first consider the test proposed by Robinson (2008) combined with the strategy in Hualde (2012). Hualde's procedure has the advantage over other fractional cointegration approaches of providing an automatic method for inferring cointegrating relationships without any prior information about the variables. He defines the possibility of cointegration as a situation in which a linear combination of fractional processes is integrated of a strictly smaller order than the maximum order of the elements of the linear combination. For example, if one of the variables has an integration order that is strictly greater than the rest of the variables, then any linear combination with zero weight on this particular variable is considered to be a trivial cointegrating relation. Under this definition, the variables can have different integration orders. However, when all the variables have the same integration order, this definition coincides with that originally provided by Engle and Granger (1987).

The proposal by Hualde (2012) is based on an estimator of the cointegrating rank, $r$, obtained by applying sequentially the procedure discussed in his Theorem 1, which we rewrite here:

**Theorem 1** (Hualde 2012). $y_t$ has cointegrating rank $r \in \{1, \ldots, p-1\}$ where $p$ is the number of variables in the vector $y_t$, if (i) and (ii) are satisfied, where: (i) There exists a $(p-r)$ dimensional subvector of $y_t$, denoted as $y_{(b)t}$, whose individual components are common trends, denoted as CT. (ii) All subvectors of $y_t$ of dimension larger than $p-r$ containing $y_{(b)t}$ cointegrate.

The procedure to estimate the rank of cointegration r is based on the following steps. First, the estimates of the integration orders (memory parameters), $\hat{d}_i, i = 1, \ldots, p$ are obtained to define the CT as the series with the highest order of integration. Then, the following hypothesis is tested:

$$H_{(j_1,\ldots,j_k)} : \{y_{j_1t}, y_{j_2t}, \ldots, y_{j_kt} \text{ are not cointegrated}\};$$

$$\text{against } \bar{H}_{(j_1,\ldots,j_k)} : H_{(j_1,\ldots,j_k)} \text{ is not true,}$$

where $j_1 \ldots, j_k \in \{1, \ldots, p\}, k \leq p$, is sequentially tested. In order to estimate the memory parameters, the univariate local Whittle estimator $\hat{d}_i$, proposed by Robinson (1995a) is used. Next, the hypotheses are tested using the statistic $X^*$ proposed by Robinson (2008), defined as:

$$X^* = ms^*(\tilde{d})^2 / \left\{p^2 tr\left(\hat{R}^* A \hat{R}^* A\right) - p\right\} \quad (4)$$

where

$m$ is the bandwidth

$$s^*\left(\tilde{d}\right) = tr\left\{\hat{G}^*\left(\tilde{d}\right)^{-1} \hat{H}^*\left(\tilde{d}\right)\right\}$$

$$\hat{G}^*(d) = \frac{1}{m}\sum_{j=1}^{m} I_y(\lambda_j)\lambda_j^{2d}$$

$$\hat{H}^*(d) = \frac{1}{m}\sum_{j=1}^{m} v_j I_y(\lambda_j)\lambda_j^{2d}$$

$$v_j = \log j - \frac{1}{m}\sum_{i=1}^{m}\log i$$

$$\hat{R}^* = \hat{D}^{1/2}\hat{G}^*\left(\tilde{d}\right)\hat{D}^{-1/2}$$

$\hat{D} = diag\left\{\hat{g}_{11}, \ldots, \hat{g}_{pp}\right\}$, where $\hat{g}_{ii}$ is the ith diagonal element of $\hat{G}^*\left(\tilde{d}\right)$

$$A = diag\{a_1, \ldots, a_p\}$$

where $I_y(\lambda_j)$ is the periodogram matrix of $y$ al frequency $\lambda_j$, $\tilde{d} = \sum_{i=1}^{p} a_i \hat{d}_i$ and the $a_i$ are arbitrarily chosen weights satisfying that $\sum_{i=1}^{p} a_i = 1$. For instance, Robinson (2008) takes $a_i \equiv 1/p$, so the arithmetic mean of the $\hat{d}_i$ is used. Another option is using $a_j = 1, a_i = 0, i \neq j$ some j. In our case, we use the first option as recommended by Robinson (2008).

Under the null hypothesis of non-cointegration and stationarity of the series, which implies that all the memory parameters are smaller than 0.5,

$$X^* \xrightarrow{d} X_1^2 \qquad \text{as } T \to \infty. \tag{5}$$

The methodology to estimate the cointegration rank r is characterized by the following steps:
**Step 1**. Estimate the individual integration orders, $d_i$, by $\hat{d}_i, i = 1, \ldots, p$. Then choose a possible CT $y_{c1t}$ as the variable with the highest estimated order, such that $c_1 \in \{1, \ldots, p\}$. Next, reorder the variables in $y_t$ so that $y_{pt} = y_{c1t}$ in the new ordering. Finally, given the possible CT i.e. $y_{pt}$, we test the following hypotheses:
$H(1) : \bigcup_{i=1}^{p-1} H_{p,i}$ versus $\bar{H}(1) : \bigcap_{i=1}^{p-1} \bar{H}_{p,i}$ Note that H(1) means non-cointegration in pairs of each variable with the CT, and $\bar{H}(1)$ means that H(1) is not true. The process ends if $H(1)$ is rejected, and so it is concluded that $\hat{r} = p - 1$. Otherwise, it is not rejected, and the process continues to Step 2. Consequently, following Theorem 1, the hypotheses are equivalent to $r < p - 1$ and $r = p - 1$ respectively.

**Step 2**. If $H(1)$ is not rejected, choose a second possible CT as the variable with the smallest statistic $X^*$ i.e. $y_{c2t}, c_2 \in \{1, \ldots, p-1\}$. There will be two possible CTs altogether. These CTs are denoted as $y_{pt}$ and $y_{c2t}$ respectively. Then, we reorder again the variables so that $y_{pt} = y_{c1t}$ and $y_{p-1,t} = y_{c2t}$ in the new ordering. Finally, given the possible CTs, i.e. $y_{pt}, y_{p-1,t}$, we test the following hypotheses: $H(2) : \bigcup_{i=1}^{p-2} H_{p,p-1,i}$ versus $\bar{H}(2) : \bigcap_{i=1}^{p-2} \bar{H}_{p,p-1,i}$. Note that H(2) means non-cointegration for any set of three variables containing the CTs $y_{pt}, y_{p-1,t}$, and $\bar{H}(2)$ means that H(2) is not true. Then, the hypotheses are equivalent to $r < p - 2$ and $r = p - 2$ respectively and the process ends if H(2) is rejected.

**Step** $k$ (*for* $k = 2, \ldots, p-1$). If $H(k-1)$ is not rejected, choose $c_k$. Sort the variables so that $y_{pt} = y_{c1t}, \ldots, y_{p-k+2,t} = y_{c_{k-1},t}$ and choose the possible CTs, as previously.

Finally, test the following hypothesis:
H(k): $\cup_{i\text{-}1}^{p\text{-}k} H_{p,p-1,\ldots,p-k+1,i}$ versus $\bar{H}$(k): $\cap_{i\text{-}1}^{p\text{-}k} H_{p,p-1,\ldots,p-k,i}$ and if there is cointegration, the estimation will be $\hat{r} = p - k$. However, if we reach the last step $k = p - 1$ this means that $\hat{r} = 0$ and $H(i), i = 1, 2, 3, \ldots, p-1$, are not rejected.
The testing procedure based on the statistic $X^*$ has low power with small sample sizes, which can significantly influence the results obtained when analysing the possibility of cointegration in the Spanish regions. To complement the results obtained we also consider the test proposed by Nielsen (2010), which has higher power for small samples (see Nielsen's Monte Carlo).

## *Nielsen's (2010) approach and critical values extension*

The test statistic is defined as follows:

$$\Lambda_{p,r}(d_1) = T^{2d_1} \sum\nolimits_{j=1}^{p-r} \vartheta_j, \quad r = 0, \ldots, p-1 \tag{6}$$

where $\vartheta_j, j = 1, \ldots, p$, are the eigenvalues of $|\vartheta B_T - A_T| = 0$, for $A_T = \sum_{t=1}^{T} Z_t Z_t{}', B_T = \sum_{t=1}^{T} \tilde{Z}_t \tilde{Z}_t{}'$, $B_T = \sum_{t=1}^{T} \tilde{Z}_t \tilde{Z}_t'$, $\tilde{Z}_t = \Delta_1^{-d_1} Z_t$ with $d_1 > 0, t = 1, 2, \ldots, T$, and $Z_t$ is the p-vector of time series under analysis (perhaps after extracting deterministic terms), which is fractionally integrated of order d, where d is a vector containing the individual orders of integration of the elements in $Z_t$, which possibly differ from each other. Note that (6) defines a family of tests indexed by the fractional integration parameter, $d_1$. Nielsen (2010) argues in favour of using $d_1 = 0.1$ based on an asymptotic local power analysis and on simulations. For this reason, we use this value in the empirical application. Large values of $\Lambda_{p,r_0}(d_1)$ are associated with the rejection of the null hypothesis $H_0 : r = r_0$ versus $H_1 : r > r_0$.

Nielsen's procedure has the advantage of not requiring knowledge of the fractional integration and cointegration orders $d$ and $b$ as long as the series are non-stationary, implying memory parameters greater than 0.5. However, its asymptotic distribution is non-standard, but Nielsen (2010) simulated critical values for $p < 8$ variables with a sample size of 1000 observations to facilitate its application. We complement Nielsen's (2010) tables by providing more critical values to cover up to 17 variables for all models and three different sample sizes.

The observed time series $\{Y_t\}_{t=1}^{T}$considered by Nielsen (2010) are generated by

$$Y_t = \alpha^{'}\delta_t + Z_t, \qquad t = 1, 2, \ldots \tag{7}$$

where $\delta_{\mathrm{t}}$ may contain deterministic terms. Three different cases are analysed: $\delta_{\mathrm{t}} = 0$ when there are no deterministic terms, $\delta_{\mathrm{t}} = 1$ when there is a non-zero mean, and $\delta_{\mathrm{t}} = [1, \mathrm{t}]'$ when there is a deterministic trend. The critical values in Nielsen (2010) are here extended for these three cases up to 17 variables, sample sizes $T = 1000$, 150 and 66, and two different values of the memory parameter of the original series $d = 1$ and $d = 1.4$, the latter corresponding to the values found in the series here analysed. All tables are based on 100,000 replications. Tables 2, 3 and 4 show these new critical values for a memory parameter equal to $d = 1.4$. Tables 24, 25 and 26 in the *Supplementary Material* show the same set of critical values for $d = 1$, which are of independent interest for practitioners.

## IV. Empirical analysis

### *Preliminary analysis of the variables*

The data analysed are the logarithms of the annual real GDP per capita (constant 2010 €) of the 17 Autonomous Communities in Spain (omitting the autonomous cities) in thousands of euros from 1955 to 2020 for a total of $T = 66$ observations. In order to apply the semiparametric cointegration analysis by Robinson (2008), the series are differenced to obtain growth rates that are stationary, whereas for the non-parametric cointegration analysis by Nielsen (2010), the series are raw series (non-stationary). The variables in logarithms are denoted by the name of the Autonomous Community, and the growth rates are denoted by the abbreviation of these Autonomous Communities (in parenthesis the notation of the growth rates): Andalucia (andal), Aragon (ara), Asturias (ast), Balearic Islands (bal), Canary Islands (can), Cantabria (cant), Catalonia (cat), Castilla-LaMancha (clm), Castilla-Leon (cyl), Extremadura (ext), Galicia (gal), Madrid (mad), Murcia (mur), Navarre (nav), Basque Country (pv), Rioja (rio) and Valencia (val). All data were provided by FEDEA (Foundation for the Study of Applied Economics) and INE (Spanish National Statistics Institute). Figures 2 and 3 show these series.

To shed more light on the persistence of the series, the Exact Local Whittle (ELW) estimator proposed by Shimotsu, et al., 2005, which is consistent and asymptotically normal for any value of $d$, was applied. The ELW estimates and their 95% confidence intervals are shown in Table 5. They confirm the non-stationarity of the series, which is a requirement for the applicability of Nielsen's procedure. Note also that a value of $\mathrm{d} = 1.4$ is not rejected for any of the series, falling within all the confidence intervals, which justifies the use of this value in the construction of the statistic for Nielsen's test.

The estimation of the memory parameter of the growth rates, required for the application of Robinson's (2008) test is, however, obtained using the Local Whittle (LW) estimator of Robinson (Robinson 1995) as suggested in Robinson (2008), which is consistent for $\mathrm{d} < 1$ and asymptotically normal for $\mathrm{d} < 0.75$. All the LW estimates, shown in Table 6, are between 0.2 and 0.5, indicating that the growth rate series can be considered stationary ($d < 0.5$).

### *Robinson (2008) and Hualde (2012) cointegration results*

In order to analyse the robustness of the results to the selection of the bandwidth, the entire analysis has been implemented using three different bandwidths, $m = 18$, $m = 23$ and $m = 28$. According to the results of the LW estimated memory parameters, the possible CT in Step 1 (variable with the largest estimated d) is *Murcia (mur)* for all

**Table 2.** Simulated critical values for large sample ($T = 1000$).

| | $p-r$ | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\xi$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| $CV_{\xi,p-r}(1.4, 0.1)$ *for* $\delta_t = 0$ | | | | | | | | | | | | | | | | | |
| 0.10 | 1.55 | 3.18 | 5.00 | 6.95 | 9.01 | 11.16 | 13.38 | 15.68 | 18.04 | 20.45 | 22.92 | 25.44 | 28.00 | 30.61 | 33.26 | 35.94 | 38.66 |
| 0.05 | 1.65 | 3.27 | 5.10 | 7.05 | 9.11 | 11.26 | 13.49 | 15.79 | 18.16 | 20.57 | 23.04 | 25.56 | 28.12 | 30.74 | 33.38 | 36.07 | 38.79 |
| 0.01 | 1.83 | 3.46 | 5.28 | 7.24 | 9.31 | 11.47 | 13.71 | 16.00 | 18.37 | 20.79 | 23.26 | 25.79 | 28.36 | 30.98 | 33.63 | 36.31 | 39.04 |
| $CV_{\xi,p-r}(1.4, 0.1)$ *for* $\delta_t$= 1 | | | | | | | | | | | | | | | | | |
| 0.10 | 1.71 | 3.46 | 5.34 | 7.33 | 9.42 | 11.58 | 13.82 | 16.13 | 18.49 | 20.91 | 23.39 | 25.91 | 28.47 | 31.09 | 33.73 | 36.42 | 39.14 |
| 0.05 | 1.79 | 3.54 | 5.43 | 7.43 | 9.52 | 11.68 | 13.93 | 16.23 | 18.61 | 21.03 | 23.50 | 26.03 | 28.60 | 31.21 | 33.86 | 36.55 | 39.27 |
| 0.01 | 1.93 | 3.71 | 5.60 | 7.61 | 9.71 | 11.88 | 14.14 | 16.44 | 18.82 | 21.25 | 23.72 | 26.26 | 28.83 | 31.45 | 34.10 | 36.79 | 39.51 |
| $CV_{\xi,p-r}(1.4, 0.1)$ *for* $\delta_t$= [1, t] | | | | | | | | | | | | | | | | | |
| 0.10 | 1.94 | 3.85 | 5.83 | 7.89 | 10.03 | 12.24 | 14.51 | 16.83 | 19.22 | 21.65 | 23.13 | 26.66 | 29.23 | 31.84 | 34.49 | 31.18 | 39.90 |
| 0.05 | 2.01 | 3.93 | 5.92 | 7.99 | 10.13 | 12.34 | 14.61 | 16.94 | 19.33 | 21.77 | 24.25 | 26.78 | 29.36 | 31.97 | 34.62 | 37.31 | 40.03 |
| 0.01 | 2.13 | 4.08 | 6.09 | 8.17 | 10.32 | 12.53 | 14.82 | 17.15 | 19.55 | 21.99 | 24.47 | 27.01 | 29.58 | 32.21 | 34.87 | 37.55 | 40.27 |

NOTE: The simulated critical values are based on 100,000 replications for up to 17 series.

**Table 3.** Simulated critical values for small sample size ($T = 150$).

| | $p-r$ | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\xi$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| $CV_{\xi,p-r}(1.4, d_1)$ *for* $\delta_t = 0$ | | | | | | | | | | | | | | | | | |
| 0.10 | 1.53 | 3.06 | 4.75 | 6.55 | 8.43 | 10.40 | 12.42 | 14.49 | 16.61 | 18.77 | 20.97 | 23.20 | 25.45 | 27.76 | 30.08 | 32.43 | 34.81 |
| 0.05 | 1.61 | 3.14 | 4.83 | 6.63 | 8.51 | 10.47 | 12.50 | 14.57 | 16.69 | 18.85 | 21.05 | 23.28 | 25.55 | 27.84 | 30.17 | 32.51 | 34.89 |
| 0.01 | 1.74 | 3.31 | 4.99 | 6.79 | 8.67 | 10.63 | 12.66 | 14.72 | 16.84 | 19.01 | 21.21 | 23.43 | 25.70 | 28.00 | 30.33 | 32.65 | 35.03 |
| $CV_{\xi,p-r}(1.4, d_1)$ *for* $\delta_t = 1$ | | | | | | | | | | | | | | | | | |
| 0.10 | 1.75 | 3.47 | 5.28 | 7.18 | 9.13 | 11.15 | 13.23 | 15.34 | 17.50 | 19.71 | 21.94 | 24.20 | 26.51 | 28.83 | 31.18 | 33.55 | 35.95 |
| 0.05 | 1.80 | 3.53 | 5.36 | 7.25 | 9.21 | 11.23 | 13.30 | 15.42 | 17.58 | 19.79 | 22.01 | 24.29 | 26.58 | 28.91 | 31.26 | 33.63 | 36.03 |
| 0.01 | 1.91 | 3.67 | 5.49 | 7.39 | 9.34 | 11.37 | 13.45 | 15.58 | 17.74 | 19.94 | 22.16 | 24.43 | 26.73 | 29.05 | 31.40 | 33.77 | 36.17 |
| $CV_{\xi,p-r}(1.4, d_1)$ *for* $\delta_t$= [1, t] | | | | | | | | | | | | | | | | | |
| 0.10 | 1.91 | 3.77 | 5.69 | 7.67 | 9.70 | 11.77 | 13.89 | 16.06 | 18.26 | 20.49 | 22.76 | 25.06 | 27.39 | 29.74 | 32.11 | 34.50 | 36.92 |
| 0.05 | 1.96 | 3.83 | 5.77 | 7.74 | 9.77 | 11.85 | 13.97 | 16.14 | 18.34 | 20.58 | 22.84 | 25.13 | 27.46 | 29.82 | 32.18 | 34.58 | 37.00 |
| 0.01 | 2.05 | 3.95 | 5.90 | 7.88 | 9.91 | 12.00 | 14.11 | 16.29 | 18.49 | 20.73 | 22.99 | 25.29 | 27.61 | 29.96 | 32.33 | 34.72 | 37.14 |

NOTE: The simulated critical values are based on 100,000 replications for up to 17 series.

**Table 4.** Simulated critical values for small sample size ($T = 66$).

| | $p-r$ | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\xi$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| $CV_{\xi,p-r}(1.4, 0.1)$ *for* $\delta_t = 0$ | | | | | | | | | | | | | | | | | |
| 0.10 | 1.53 | 3.04 | 4.71 | 6.49 | 8.35 | 10.27 | 12.25 | 14.27 | 16.33 | 18.43 | 20.56 | 22.71 | 24.90 | 27.11 | 29.33 | 31.58 | 33.85 |
| 0.05 | 1.60 | 3.12 | 4.78 | 6.57 | 8.43 | 10.34 | 12.33 | 14.34 | 16.40 | 18.51 | 20.62 | 22.78 | 24.97 | 27.17 | 29.40 | 31.65 | 33.91 |
| 0.01 | 1.74 | 3.30 | 4.95 | 6.72 | 8.58 | 10.49 | 12.46 | 14.49 | 16.53 | 18.63 | 20.76 | 22.91 | 25.09 | 27.29 | 29.51 | 31.76 | 34.04 |
| $CV_{\xi,p-r}(1.4, 0.1)$ *for* $\delta_t = 1$ | | | | | | | | | | | | | | | | | |
| 0.10 | 1.74 | 3.44 | 5.23 | 7.10 | 9.03 | 11.00 | 13.03 | 15.10 | 17.19 | 19.33 | 21.49 | 23.68 | 25.89 | 28.11 | 30.37 | 32.64 | 34.92 |
| 0.05 | 1.80 | 3.51 | 5.30 | 7.17 | 9.10 | 11.07 | 13.10 | 15.17 | 17.26 | 19.40 | 21.56 | 23.74 | 25.95 | 28.17 | 30.42 | 32.69 | 34.99 |
| 0.01 | 1.91 | 3.66 | 5.44 | 7.31 | 9.24 | 11.20 | 13.24 | 15.30 | 17.40 | 19.52 | 21.66 | 23.86 | 26.06 | 28.29 | 30.54 | 32.81 | 35.09 |
| $CV_{\xi,p-r}(1.4, 0.1)$ *for* $\delta_t$= [1, t]′ | | | | | | | | | | | | | | | | | |
| 0.10 | 1.89 | 3.74 | 5.62 | 7.57 | 9.56 | 11.58 | 13.66 | 15.76 | 17.89 | 20.05 | 22.24 | 24.45 | 26.68 | 28.93 | 31.20 | 33.48 | 35.78 |
| 0.05 | 1.94 | 3.80 | 5.69 | 7.64 | 9.63 | 11.65 | 13.73 | 15.82 | 17.95 | 20.11 | 22.30 | 24.51 | 26.74 | 28.99 | 31.25 | 33.54 | 35.84 |
| 0.01 | 2.02 | 3.92 | 5.82 | 7.76 | 9.75 | 11.77 | 13.85 | 15.96 | 18.07 | 20.24 | 22.42 | 24.62 | 26.85 | 29.10 | 31.37 | 33.66 | 35.95 |

NOTE: The simulated critical values are based on 100,000 replications for up to 17 series.

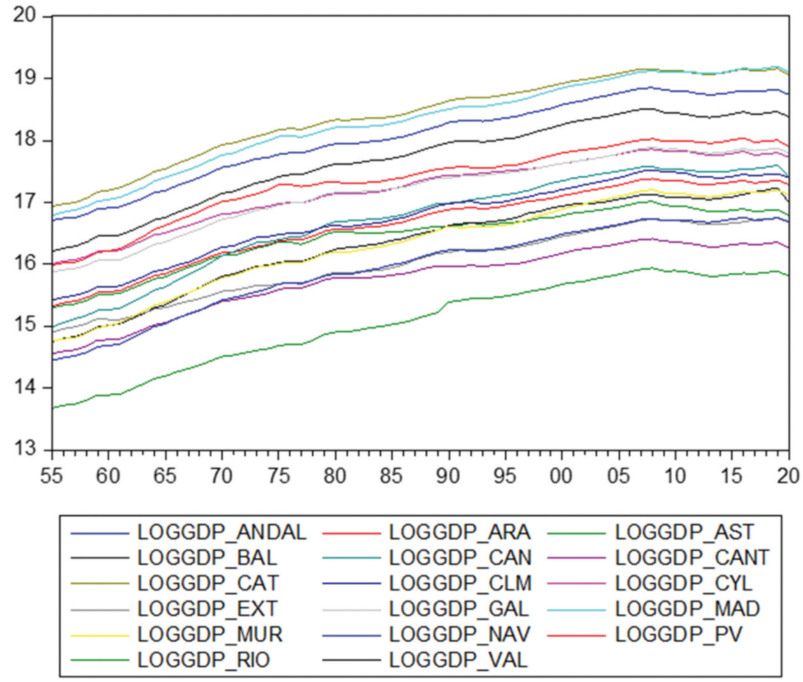


**Figure 2.** Time-series plot of all log real GDP variables.

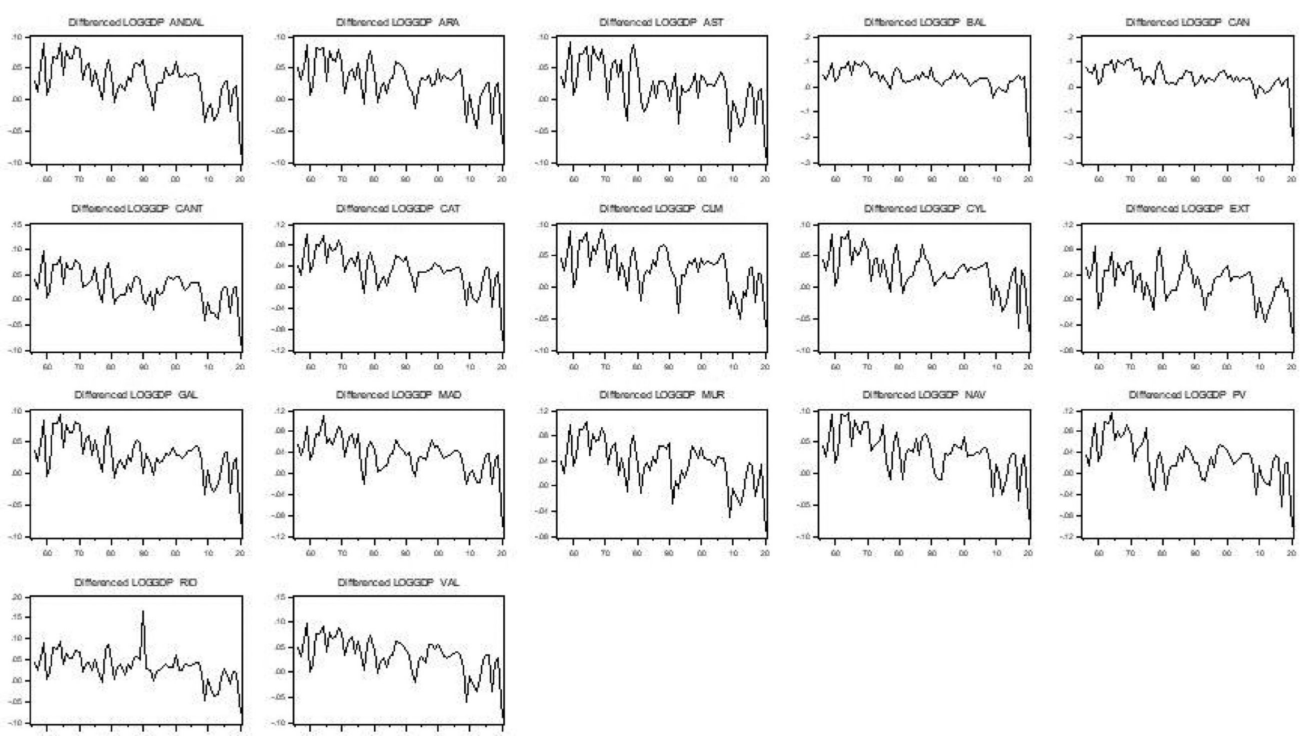


**Figure 3.** Time-series plots of the GDP growth rate.

the bandwidths. In subsequent steps, i.*e.* *2, 3, ... , p-1 = 16*, the possible CTs are chosen as the series with the smallest statistic X* for each particular bandwidth.

To understand how Hualde's (2012) procedure works, we will describe in more detail some of its steps. Step 1 checks for cointegration, two by two. For instance, for a bandwidth

**Table 5.** ELW estimates in the original series.

Exact Local Whittle IC in []

| Series\bandwidth | m = 18 | m = 23 | m = 28 |
|---|---|---|---|
| Andalucia | 1.41 [1.17; 1.64] | 1.40 [1.19; 1.60] | 1.40 [1.28; 1.58] |
| Aragon | 1.31 [1.07; 1.54] | 1.25 [1.04; 1.45] | 1.27 [1.08; 1.45] |
| Asturias | 1.26 [1.02; 1.49] | 1.23 [1.02; 1.43] | 1.27 [1.08; 1.45] |
| Balearic Islands | 1.34 [1.10;1.57] | 1.32 [1.11; 1.52] | 1.32 [1.13; 1.50] |
| Canary Islands | 1.33 [1.09; 1.56] | 1.25 [1.04; 1.45] | 1.28 [1.09; 1.46] |
| Cantabria | 1.30 [1.06; 1.53] | 1.28 [1.07; 1.48] | 1.32 [1.13; 1.50] |
| Catalonia | 1.32 [1.08; 1.55] | 1.29 [1.08; 1.49] | 1.32 [1.13; 1.50] |
| Castilla-LaMancha | 1.41 [1.17; 1.64] | 1.31 [1.10; 1.51] | 1.33 [1.14; 1.51] |
| Castilla-León | 1.31 [1.07; 1.54] | 1.26 [1.05; 1.46] | 1.29 [1.10; 1.47] |
| Extremadura | 1.36 [1.12; 1.59] | 1.27 [1.06; 1.47] | 1.34 [1.15; 1.52] |
| Galicia | 1.36 [1.12; 1.59] | 1.27 [1.06; 1.47] | 1.33 [1.14; 1.51] |
| Madrid | 1.29 [1.05; 1.52] | 1.28 [1.07; 1.48] | 1.31 [1.12; 1.49] |
| Murcia | 1.35 [1.11; 1.58] | 1.34 [1.13; 1.54] | 1.36 [1.17; 1.54] |
| Navarre | 1.29 [1.05; 1.52] | 1.29 [1.08; 1.49] | 1.31 [1.12; 1.49] |
| País Vasco | 1.30 [1.06; 1.53] | 1.31 [1.10; 1.51] | 1.35 [1.16; 1.53] |
| Rioja | 1.30 [1.06; 1.53] | 1.34 [1.13; 1.54] | 1.35 [1.16; 1.53] |
| Valencia | 1.32 [1.08; 1.55] | 1.29 [1.08; 1.49] | 1.33 [1.14; 1.51] |

**Table 6.** LW estimates in growth rate.

Local Whittle

| | m:18 | m:23 | m:28 |
|---|---|---|---|
| Andal | 0.491 | 0.448 | 0.337 |
| Ara | 0.394 | 0.339 | 0.268 |
| Ast | 0.361 | 0.319 | 0.274 |
| Bal | 0.341 | 0.323 | 0.218 |
| Can | 0.432 | 0.332 | 0.259 |
| Cant | 0.397 | 0.378 | 0.310 |
| Cat | 0.448 | 0.410 | 0.326 |
| Clm | 0.496 | 0.383 | 0.323 |
| Cyl | 0.389 | 0.310 | 0.270 |
| Ext | 0.355 | 0.345 | 0.260 |
| Gal | 0.436 | 0.357 | 0.296 |
| Mad | 0.418 | 0.383 | 0.288 |
| Mur | 0.497 | 0.454 | 0.373 |
| Nav | 0.380 | 0.385 | 0.314 |
| PV | 0.429 | 0.414 | 0.357 |
| Rio | 0.346 | 0.316 | 0.281 |
| Val | 0.447 | 0.387 | 0.311 |

$m = 18$, the results indicate that $mur_t$ is the possible common trend. Therefore, according to this choice of bandwidth, the first step in this procedure is to test:

$$\mathrm{H}(1) : \mathrm{H}_{mur,andal} \cup \mathrm{H}_{mur,ara} \cup \mathrm{H}_{mur,ast} \cup \mathrm{H}_{mur,bal} \cup \mathrm{H}_{mur,cant} \cup \mathrm{H}_{mur,cat} \cup \mathrm{H}_{mur,clm} \cup \mathrm{H}_{mur,cyl} \cup \mathrm{H}_{mur,ext} \cup \mathrm{H}_{mur,gal} \cup \mathrm{H}_{mur,mad} \cup \mathrm{H}_{mur,can} \cup \mathrm{H}_{mur,nav} \cup \mathrm{H}_{mur,pv} \cup \mathrm{H}_{mur,rio} \cup \mathrm{H}_{mur,val};$$

against

$$\bar{\mathrm{H}}(1) : \bar{\mathrm{H}}_{mur,andal} \cap \bar{\mathrm{H}}_{mur,ara} \cap \bar{\mathrm{H}}_{mur,ast} \cap \bar{\mathrm{H}}_{mur,bal} \cap \bar{\mathrm{H}}_{mur,cant} \cap \bar{\mathrm{H}}_{mur,cat} \cap \bar{\mathrm{H}}_{mur,clm} \cap \bar{\mathrm{H}}_{mur,cyl} \cap \bar{\mathrm{H}}_{mur,ext} \cap \bar{\mathrm{H}}_{mur,gal} \cap \bar{\mathrm{H}}_{mur,mad} \cap \bar{\mathrm{H}}_{mur,can} \cap \bar{\mathrm{H}}_{mur,nav} \cap \bar{\mathrm{H}}_{mur,pv} \cap \bar{\mathrm{H}}_{mur,rio} \cap \bar{\mathrm{H}}_{mur,val}$$

where for

$$a_t = andal_t, ara_t, ast_t, bal_t, cant_t, cat_t, clm_t, cyl_t, ext_t, gal_t, mad_t, can_t, nav_t, pv_t, rio_t, val_t$$

(all variables without the CT)

$$\mathrm{H}_{mur,a} : mur_t, a_t \text{ are not cointegrated}$$

$$\bar{\mathrm{H}}_{mur,a} : \mathrm{H}_{mur,a} \text{is not true}$$

The above hypotheses are equivalent to $\mathrm{H}(1) : r < 16$ and $\bar{\mathrm{H}}(1) : r = 16$. According to Hualde's (2012) procedure, if $mur_t$ is a true CT, then $r < 16$ if and only if $\mathrm{H}(1)$holds, and $r = 16$ if and only if $\bar{\mathrm{H}}(1)$ holds.

The process ends when H(k) for $k = 1, 2, \ldots, 16$ is rejected. In this case, the process continues to Step 2 because the null hypothesis of no cointegration cannot be rejected for any pair of variables and for any bandwidth choice because none of the X* statistics are greater than the critical value of the $\chi_1^{\,2}$, that is, 3.84 at 5% level of significance (see Table 7). The second possible *CT* is chosen as the series leading to the smallest X* statistics obtained in Step 1 (see also Table 7), that is$rio_t$(La Rioja). The null hypothesis of no cointegration is not rejected in either of the steps. The last step tests if there is any long-run relationship between the 17 variables, i.e., the hypotheses are equivalent to $r < 1$ and $r = 1$ that is:

$$\mathrm{H}(16) : \mathrm{H}_{can,cyl,pv,rio,cant,ext,gal,cat,clm,bal,andal,val,ast,ara,nav,mad,mur}$$

against

$$\mathrm{H}(16) : \mathrm{H}_{can,cyl,pv,rio,cant,ext,gal,cat,clm,bal,andal,val,ast,ara,nav,mad,mur}$$

where

$$\mathrm{H}_{can,cyl,pv,rio,cant,ext,gal,cat,clm,bal,andal,val,ast,ara,nav,mad,mur} :$$

$$\mathrm{can}_t, \mathrm{cyl}_t, \mathrm{pv}_t, \mathrm{rio}_t, \mathrm{cant}_t, \mathrm{ext}_t, \mathrm{gal}_t, \mathrm{cat}_t, \mathrm{clm}_t, \mathrm{bal}_t, \mathrm{andal}_t, \mathrm{val}_t, \mathrm{ast}_t, \mathrm{ara}_t, \mathrm{nav}_t, \mathrm{mad}_t, \mathrm{mur}_t \text{ are not cointegrated.}$$

$$\bar{\mathrm{H}}_{can,cyl,pv,rio,cant,ext,gal,cat,clm,bal,andal,val,ast,ara,nav,mad,mur} :$$

$\mathrm{H}_{can,cyl,pv,rio,cant,ext,gal,cat,clm,bal,andal,val,ast,ara,nav,mad,mur}$ is not true.

**Table 7.** Cointegration rank test, step I.

| Variables\m | m:18 | m:23 | m:28 |
|---|---|---|---|
| Step 1 | | | |
| Andal | 2.63 | 2.44 | 1.47 |
| Ara | 1.58 | 0.879 | 0.24 |
| Ast | 0.81 | 1.20 | 0.35 |
| Bal | 1.30 | 0.622 | 0.06 |
| Can | 1.06 | 1.23 | 0.62 |
| Cant | 1.03 | 1.75 | 0.85 |
| Cat | 1.57 | 1.32 | 0.86 |
| Clm | 2.32 | 1.88 | 1.29 |
| Cyl | 0.48 | 0.242 | 0.16 |
| Ext | 0.91 | 0.111 | 0.046 |
| Gal | 1.51 | 1.42 | 1.32 |
| Mad | 1.69 | 2.26 | 1.56 |
| Mur | CT | CT | CT |
| Nav | 1.58 | 1.52 | 0.73 |
| Pv | 0.39 | 1.21 | 0.61 |
| Rio | 0.12 | 0.00 | 0.24 |
| Val | 1.49 | 1.79 | 1.19 |

*Note: The series are the growth rate per capita.*

**Table 8.** Cointegration rank test, step XVI.

| Variables\m | m:18 | m:23 | m:28 |
|---|---|---|---|
| Step 16 | | | |
| Andal | CT | CT | CT |
| Ara | CT | CT | CT |
| Ast | CT | 0.092 | 0.0014 |
| Bal | CT | CT | CT |
| Can | CT | CT | CT |
| Cant | CT | CT | CT |
| Cat | CT | CT | CT |
| Clm | 0.0019 | CT | CT |
| Cyl | CT | CT | CT |
| Ext | CT | CT | CT |
| Gal | CT | CT | CT |
| Mad | CT | CT | CT |
| Mur | CT | CT | CT |
| Nav | CT | CT | CT |
| Pv | CT | CT | CT |
| Rio | CT | CT | CT |
| Val | CT | CT | CT |

**Table 9.** Variance ratio cointegration rank test.

| $p-r$ | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| 1.50 | 3.10 | 4.82 | 6.59 | 8.46 | 10.38 | 12.35 | 14.38 | 16.47 | 18.66 | 20.88 | 23.14 | 25.43 | 27.75 | 30.09 | 32.50 | 34.96 |

Panel reports the variance ratio cointegration test statistic $\Lambda_{p,r}(0.1)$.

The results of the test in this step for various bandwidths are shown in Table 8. There is no evidence of cointegration for any of the bandwidths considered. Therefore, there is no long-run relationship between any of the 17 Autonomous Communities of Spain, implying that the growth rates of the output in these regions do not converge and there does not exist any group of regions converging. Full details of all the steps in the procedure can be found in Tables 10 to 23 in the supplementary material of this article. To further support this conclusion, the nonparametric method proposed by Nielsen (2010) is also implemented.

### *Nielsen's (2010) cointegration results*

Unlike the cointegration test proposed by Robinson (2008), the procedure in Nielsen (2010) begins with $\mathrm{H}_0 : \mathrm{r} = \mathrm{r}_0$ versus$\mathrm{H}_1 : \mathrm{r} > \mathrm{r}_0$ and the testing sequence ends when the null hypothesis is not rejected. The variance ratio rank test is based on the statistic $\Lambda_{p,r}(0.1)$, defined in the previous section, for $p = 1$ to 17 and applied to the detrended non-stationary series. Denoting $CV_{\xi,p-r}(d, 0.1)$, the critical values, where $d$ represents the memory parameter of the series, the null hypothesis is rejected if $\Lambda_{p,r}(0.1) > CV_{\xi,p-r}(d, 0.1)$. In that case, the procedure continues to the next step.

The results of this variance ratio test applied to our sample are shown in Table 9, which shows the variance ratio test statistics for the case of constant and trend (selected in view of the behaviour of the series in Figure 2). Comparing the statistics with the critical values obtained with $\delta_t = 1$ in Table 4, we have that $\Lambda_{17,0}(0.1) = 34.96 < CV_{0.05,17}(1.4, 0.1) = 34.99$, concluding that the null hypothesis $r = 0$ cannot be rejected at 5% significance level. This result indicates no cointegration among the per capita gross domestic product of the Autonomous Communities in Spain, reinforcing the lack of convergence found with the strategy designed by Hualde (2012) and Robinson (2008).

## V. Conclusions

In this work, we have tested for the existence of convergence of real per capita GDP in 17 Spanish

Autonomous Communities from 1955 to 2020, applying advanced and up-to-date time-series methods in a context of fractional integration and cointegration. Our analyses yielded no evidence of long-run equilibrium relationships, thus ruling out the possibility of convergence according to the definitions of Bernard and Durlauf (1995). The results are robust, confirmed with different techniques used to test for cointegration, and are also consistent with a cross-sectional analysis based on sigma-convergence.

Our empirical findings provide us with an important insight into this framework: There is no overall economic convergence between the Spanish Autonomous Communities, and no convergent subgroup of regions has been identified. The complete lack of convergence among the 17 Autonomous Communities could be explained by the persistent economic differences between them, which the convergence process has not been able to offset in the period analysed. This may be due to the existence of heterogeneous economic structures, different levels of human capital, and regional differences in the quality of institutions, among other factors that affect the growth rates of the regions differently. The different economic structures, mentioned in the Introduction, lead to disparities in the economic growth. Some regions rely more heavily on certain economic sectors, such agriculture that experience slower economic growth. On the other hand, regions with high concentration of industries with high levels of innovation and technological advancement may grow faster than those based on traditional sectors with lower productivity levels.

The absence of economic convergence can have substantial consequences, including inequalities in access to employment, education, and public services, as well as political and social tensions. Consequently, it represents a challenge that Spanish governments need to address through regional development policies and other measures aimed at promoting economic convergence and reducing disparities across regions, implementing policies that address the specific needs and challenges of each Autonomous Community.

## Acknowledgments

Mariam wishes to express her heartfelt gratitude to Oscar for his extremely valuable and profound help.

## Disclosure statement

The author(s) declare that they have no conflict of interest.

## Funding

Mariam acknowledges financial support from project PRE_2018_1_0088 awarded by the Department of Education, Linguistic Policy and Culture of the Basque Government. Josu acknowledges the Spanish Ministry of Science and Innovation grant PID2019-105183GB-I00 and UPV/EHU Econometrics Research Group (Basque Government grant IT1359-19).

## ORCID

Mariam Kamal http://orcid.org/0000-0001-6122-7602
Josu Arteche http://orcid.org/0000-0002-6700-3615

## References

Arregui, J. 2020. *Europeanization of Political Structures and Public Policies*. Oxford. Handbook of Spanish Politics. Oxford:Oxford University Press132–150.

Arroyo, A. S. M., de Juan, and A. de Juan. 2013. "A. Spanish Regions Catching-Up to Europe: An Analysis Based on the IB-MPI Unit Root Procedure." *Empirical Economics* 45 (2): 715–733. https://doi.org/10.1007/s00181-012-0634-9.

Beckfield, J. 2006. "European Integration and Income Inequality." *American Sociological Review* 71 (6): 964–985. https://doi.org/10.1177/000312240607100605.

Bernard, A., and S. Durlauf. 1995. "Convergence in International Output." *Journal of Applied Econometrics* 10:97–108. https://doi.org/10.1002/jae.3950100202.

Bernard, A., and S. Durlauf. 1996. "Interpreting Tests of the Convergence Hypothesis." *Journal of Econometrics* 71 (1–2): 161–173. https://doi.org/10.1016/0304-4076(94)01699-2.

Bierens, H. J. 1997. "Nonparametric Cointegration Analysis." *Journal of Econometrics* 77 (2): 379–404. https://doi.org/10.1016/S0304-4076(96)01820-9.

Breitung, J. 2002. "Nonparametric Tests for Unit Roots and Cointegration." *Journal of Econometrics* 108 (2): 343–363. https://doi.org/10.1016/S0304-4076(01)00139-7.

Caporaso, J., and A Pelowski. 1971. "Economic and Political Integration in Europe: A Time-Series Quasi-Experimental Analysis." *The American Political Science Review* 65 (2): 418–433. https://doi.org/10.2307/1954458.

Cuadrado-Roura, J. R., B. García-Greciano, and J. L. Raymond. 1999. "Regional Convergence in Productivity and Productive Structure: The Spanish Case." *International Regional Science*

*Review* 22 (1): 35–53. https://doi.org/10.1177/016001799761012190.

Dickey, D., and W. Fuller. 1979. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association* 74 (366): 427–431. https://doi.org/10.1080/01621459.1979.10482531.

Dolado, J., J. M. González-Páramo, and J. M. Roldán. 1994. "Convergencia económica entre las provincias españolas." *Moneda y Crédito* 199:- 81–118.

Durlauf, S., (2000). The Convergence Hypothesis After 10 Years. Working papers 6, Wisconsin Madison - Social Systems Research Institute.

Engle, R., and C. Granger. 1987. "Cointegration and Error Correction Representation, Estimation and Testing." *Econometrica* 55 (2): 251–276. https://doi.org/10.2307/1913236.

Evans, P. 1996. "Using Cross-Country Variances to Evaluate Growth Theories." *Journal of Economic Dynamics and Control* 20 (6–7): 1027–1049. https://doi.org/10.1016/0165-1889(95)00888-8.

Friedman, M. J. 1992. "Do Old Fallacies Ever Die?" *Journal of Economic Literature* 30:2129–32.

Gardeazábal, J. 1996. "Provincial Income Distribution Dynamics: Spain 1967-1991." *Investigaciones Economicas* XX:263–69.

Granger, C. W. J. 1980. "Long Memory Relationships and the Aggregation of Dynamic Models." *Journal of Econometrics* 14 (2): 227–238. https://doi.org/10.1016/0304-4076(80)90092-5.

Granger, C. W. J. 1981. "Some Properties of Time Series Data and Their Use in Econometric Model Specification." *Journal of Econometrics* 16 (1): 121–130. https://doi.org/10.1016/0304-4076(81)90079-8.

Granger, C. W. J., and R. Joyeux. 1980. "An Introduction to Long-Memory Time Series Models and Fractional Differencing." *Journal of Time Series Analysis* 1 (1): 15–29. https://doi.org/10.1111/j.1467-9892.1980.tb00297.x.

Harris, D. 1997. "Principal Components Analysis of Cointegrated Time Series." *Econometric Theory* 13 (4): 529–557. https://doi.org/10.1017/S0266466600005995.

Hierro, M., and A. Maza, (2010). Per Capita Income Convergence and Internal Migration in Spain: Are Foreign-Born Migrants Playing an Important Role? Papers in Regional Science, Wiley Blackwell, vol. 89(1), 89–107.

Hosking, J. R. M. 1981. "Fractional Differencing." *Biometrika* 68 (1): 165–176. https://doi.org/10.1093/biomet/68.1.165.

Hualde, J. 2012. *Estimation of the cointegrating rank in fractional cointegration. Documentos de Trabajo - Lan Gaiak Departamento de Economía - Universidad Pública de Navarra 1205.* Departamento de Economía - Universidad Pública de Navarra.

Johansen, S. 1988. "Statistical Analysis of Cointegration Vectors." *Journal of Economic Dynamics and Control* 12 (2–3): 231–254. https://doi.org/10.1016/0165-1889(88)90041-3.

Johansen, S. 1991. "Estimation and Hypothesis Testing of Cointegration Vectors in GaussianVector Autoregressive Models." *Econometrica* 59 (6): 1551–1580. https://doi.org/10.2307/2938278.

Johansen, S. 1995. *Likelihood Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

Johansen, S. 2008. "A Representation Theory for a Class of Vector Autoregressive Models for Fractional Processes." *Econometric Theory* 24 (3): 651–676. https://doi.org/10.1017/S0266466608080274.

Johansen, S., and M. Nielsen. 2010. "Likelihood Inference for a Nonstationary Fractional Autoregressive Model." *Journal of Econometrics* 158 (1): 51–66. https://doi.org/10.1016/j.jeconom.2010.03.006.

Johansen, S., and M. Nielsen. 2012. "Likelihood Inference for a Fractionally Cointegrated Vector Autoregressive Model." *Econometrica* 80 6: 2667–2732.

Johansen, S., and M. Nielsen 2014. The Role of Initial Values in Nonstationary Fractional Time Series Models (Working Paper No. 1300). Kingston, Ontario, Canada: Queen's University.

Lamo, A. 2000. "On Convergence Empirics: Some Evidence for Spanish Regions." *Investigaciones Económicas* XXIV (3): 681–707.

Martínez-Argüelles, S., and F Rubiera-Morollón. 1998. "Patrones de convergencia regional en los servicios de la economía española." *Ekonomiaz* 42:146–171.

Martin, A, and G. Ross. 2004. *Euros and European: Monetary Integration and the European Social Model*, 309–330. Cambridge: Cambridge University Press.

Maza, A. 2006. "Migrations and Regional Convergence: The Case of Spain." *Jahrbuch Für Regionalwissenschaft* 26 (2): 191–202. https://doi.org/10.1007/s10037-006-0009-9.

Montañés, A., L. Olmos, and M. Reyes. 2018. "Has the Great Recession Affected the Convergence Process? The Case of Spanish Provinces." *Economic Modelling* 68:360–371. https://doi.org/10.1016/j.econmod.2017.08.006.

Nielsen, M. 2010. "Nonparametric Cointegration Analysis of Fractional Systems with Unknown Integration Orders." *Journal of Econometrics* 155 (2): 170–187. https://doi.org/10.1016/j.jeconom.2009.10.002.

Phillips, P. C. B., and S. Ouliaris. 1990. "Asymptotic Properties of Residual-Based Tests for Cointegration." *Econometrica* 58 (1): 165–193. https://doi.org/10.2307/2938339.

Puente, S. 2017. Regional convergence in Spain (2017): 19802015. Economic Bulletin – Banco de España, 3/2017, Analytical Articles.

Quah, D. T. 1993. "Galton's Fallacy and Tests of the Convergence Hypothesis." *The Scandinavian Journal of Economics* 95:427–443. https://doi.org/10.2307/3440905. 4

Robinson, P. M. 1995. "Gaussian Semiparametric Estimation of Long Range Dependence." *The Annals of Statistics* 23 (5): 1630–1661. https://doi.org/10.1214/aos/1176324317.

Robinson, P. M. 2008. "Diagnostic testing for cointegration." *Journal of Econometrics* 143 (1): 206–225. https://doi.org/10.1016/j.jeconom.2007.08.015.

Shimotsu, K., and P. C. B. Phillips. 2005. "Exact Local Whittle Estimation of Fractional Integration." *Annals of Statistics* 33 (4): 1890–1933. https://doi.org/10.1214/009053605000000309.

Tapia, B, and J. Galarraga. 2020. "Inequality and Growth in a Developing Economy: Evidence from Regional Data (Spain, 1860–1930)." *Social Science History, Cambridge University Press* 44 (1): 169–192. https://doi.org/10.1017/ssh.2019.44.

Villaverde, J. 2005. "Provincial convergence in Spain: a spatial econometric approach." *Applied Economics Letters* 12 (11): 697–700. https://doi.org/10.1080/13504850500190030.