New revelations about unemployment persistence in Spain: time series and panel data approaches using regional data

Garcia-del-Barrio, Pedro; Gil-Alana, Luis Alberiko

NEW REVELATIONS ABOUT UNEMPLOYMENT PERSISTENCE IN SPAIN. Time series and panel data approaches using regional data.

Journal:	*Applied Economics*
Manuscript ID:	APE-06-0239.R1
Journal Selection:	Applied Economics
JEL Code:	J64 - Unemployment: Models, Duration, Incidence, and Job Search < J6 - Mobility, Unemployment, and Vacancies < J - Labor and Demographic Economics
Keywords:	Unemployment persistence, Regional unemployment

NEW REVELATIONS ABOUT UNEMPLOYMENT PERSISTENCE IN SPAIN.
(Time series and panel data approaches using regional data)

P. Garcia-del-Barrio
Universidad de Navarra, Faculty of Economics, Pamplona, Spain

and

L.A. Gil-Alana[a]
Universidad de Navarra, Faculty of Economics, Pamplona, Spain

ABSTRACT

This paper aims to re-examine the persistence of unemployment in Spain. For this purpose, we use time series and cross-section analysis. From a time series viewpoint we disaggregate unemployment by regions, and use unit root tests, AR coefficients and fractional differencing parameters as indicators of persistence. For the cross-section approach, we first estimate mean regressions of regional unemployment rates. Then, using a panel of 114 periods and 50 provinces, we estimate pooled, fixed and random effects models. Finally, following some recent developments, we implement several panel data unit root tests. Previous studies had already shown the strong persistence of Spanish unemployment. Our disaggregated analysis extends that finding to reveal that the persistence is greater in the most industrialised regions. The results also suggest that a structural break took place in 1994, implying a decline in the unemployment persistence since then.

JEL Classification: J64
Keywords: Unemployment persistence; Regional unemployment.

Correspondence author: Pedro Garcia-del-Barrio
Facultad de Ciencias Económicas
Edificio Biblioteca, Entrada Este
31080 Pamplona. SPAIN

Phone: 00 34 948 425625
Fax: 00 34 948 425626
Email: pgdelbarrio@alumni.unav.es
pedro.garcia@upf.edu

[a] The second named author gratefully acknowledges financial support from the Ministerio de Ciencia y Tecnologia (SEC2005-07657, Spain).

1. Introduction

The high persistence of unemployment in the developed countries since the first oil shock has produced an intense debate about the different theories explaining unemployment behaviour. On the one hand, some traditional theories (Friedman, 1968, Phelps, 1967, 1968) assume that fluctuations in unemployment are stationary around a natural rate. Occasionally, these fluctuations may be permanently changing the natural rate of unemployment from one level to another. On the other hand, the hysteresis hypothesis (Blanchard and Summers, 1986, 1987) assumes that shocks in unemployment have a permanent effect on the level of the variable, implying that unemployment is nonstationary.

The concept of persistence might be considered somehow a conciliating theory. Persistence implies mean reversion, even if the speed of adjustment towards the equilibrium level is slow, thereby being a special case of the natural rate hypothesis. At the same time, the hysteresis theory is also closely related to the concept of persistence, insofar as it takes long time for the shocks to disappear completely. In this context, the unemployment series is modelled by autoregressions (with the roots close to the unit circle) or by some other techniques like fractional integration.

The sources of European unemployment persistence have been examined in a number of papers. Some of them highlight the presence of persistence mechanisms that lead the current equilibrium unemployment rate (UR) to be positively related to previous realizations of unemployment (Cf. Bean, 1997; or Jimeno and Bentolila 1998 for the case of Spain). The goal of this paper is not to explain those mechanisms, but simply to examine the degree of aggregate and regional UR persistence in Spain, for the period 1976-2004.

Previous studies have examined the unemployment persistence from an aggregate approach, using time series techniques. The degree of persistence has been usually measured through the estimated coefficients of AR(I)MA models. (Blanchard and

Summers, 1986, and Barro, 1988). More recently, new developments in econometrics have studied unemployment persistence through fractionally integrated (ARFIMA) models. (See, i.e., Tschernig and Zimmermann, 1992; Crato and Rothman, 1996; Caporale and Gil-Alana, 2002; and Gil-Alana, 2001a, 2002).

The study of unemployment persistence also requires accounting for regional disparities. As stated by Blanchard et *al* (1995), the most striking fact in Spanish regional trends is not just the difference in unemployment rates across regions, but rather the persistence of these differences. In fact, regional considerations might help us to understand more deeply the nature of unemployment persistence.

The interest for revisiting the Spanish case is unquestionable given its high UR. More importantly, Spain is an obvious laboratory where to apply our regional approach, since regional discrepancies of rates between provinces are enormous. In the last quarter of 2004 various southern provinces (like Badajoz, Córdoba, Jaén and Cádiz) had UR above 20%, whereas others (Navarra, Lleida, Rioja, Soria and Teruel) presented rates of less that 5%, far below the average UR in Europe (which in 2004 was 9.2% for the EU-25 zone).

In summary, this paper aims not just to analyse thoroughly the Spanish experience, but to do it in a comprehensive way by adopting different perspectives and using a broad array of empirical devices. Should the same result be achieved from diverse approaches, the conclusion could be stated in a very consistent way. Hence, this study is challenging as far as it can be used to validate the accuracy of some recent empirical methods that we have employed. The outline of the paper is as follows: Sections 2 and 3 address the concept of persistence in time series and cross-section / panel data models respectively. Section 4 describes the data and presents the empirical results, while Section 5 concludes.

2. Time series persistence

From a time series viewpoint a classic model of persistence is the unit root case, where the effect of the shocks persists forever. To illustrate this point, consider the simplest case: the random walk model,

$$(1 - L)u_t = v_t, \qquad t = 1, 2, ..., \tag{1}$$

where u_t is the UR series we observe; L is the lag-operator ($Lu_t = u_{t-1}$) and v_t is i.i.d. Note that after recursive substitutions, (1) can be re-written as:

$$u_t = \sum_{j=1}^{t} v_j, \qquad t = 1, 2, ..., \tag{2}$$

and thus, a 1-unit shock has a 1-unit effect on the future path of the series.

Let us suppose now that v_t in (1) is an ARMA(p, q) process of form:

$$\phi_p(L)v_t = \theta_q(L)\varepsilon_t, \qquad t = 1, 2, ..., \tag{3}$$

with all the roots of $\phi_p(L)$ within the unit circle and all the roots of $\theta_q(L)$ within or in the unit circle. In such a case, v_t can also be expressed in terms of a MA(∞) representation,

$$v_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}, \qquad t = 1, 2, ..., \tag{4}$$

and the coefficients b_j will slowly (in fact, exponentially) decay to zero. These coefficients are usually associated with the short run dynamics of the series. In light of this, it is crucial to correctly determine the order of integration of the series. Thus, if it is 1 (i.e, a unit root with or without ARMA components) the series is said to be persistent as opposed to the case of simple stationary ARMA structures, where the effect of shocks disappears in the future. In the latter case the series is said to be I(0) since the order of integration is 0. Note that here persistence is measured by the impulse responses in (4) (the b_j-coefficients). Thus, i.e., if u_t is a stationary AR(1) process:

$$u_t = \phi u_{t-1} + \varepsilon_t, \qquad t = 1, 2, ..., \tag{5}$$

the model can also be expressed as

$$u_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}, \qquad t = 1, 2, ..., \tag{6}$$

and the higher is the AR coefficient ϕ, the higher is the persistent behaviour of the series, though disappearing in the long run as long as $|\phi| < 1$.

Among the many procedures for testing unit roots, the tests of Dickey and Fuller (ADF, 1979) are the most widely used. These tests involve the regression of the first-difference of the series on its lagged level and k lagged first differences. To select the value of k, the Ng and Perron (2001) procedure is applied. The starting upper bound, k_{max}, is then determined by applying the Schwert (1989) criterion. In addition to ADF tests, we can also employ a DF-GLS test, which is a modified Dickey-Fuller test in which the series has been transformed by GLS regressions. For the latter tests, we can use the interpolated critical values proposed by Elliot, Rothenberg and Stock (1996).

In view of the previous discussion the classic approach to the time series persistence consists of determining if the series is nonstationary I(1) or stationary I(0), and, in the latter case, examining the coefficients that are associated with the short run dynamics. However, in the last twenty years a new model has emerged that avoids the strong dichotomy produced by the I(0) and I(1) specifications. This is because the number of differences required to get I(0) stationarity may not necessarily be an integer value (usually 1) but is possibly a real value.

We define a I(0) process $\{v_t, t = 0, \pm 1, ...\}$ as a covariance stationary process with spectral density that is bounded and bounded away from zero at the zero frequency. In such a case, u_t is I(d) if:

$$(1 - L)^d u_t = v_t, \qquad t = 1, 2, ..., \tag{7}$$

where d can be any real value. Note that using the Binomial expansion:

$$(1 - L)^d = \sum_{j=0}^{\infty} \binom{d}{j} (-1)^j L^j = 1 - d L + \frac{d(d-1)}{2} L^2 - ....$$

(7) can be expressed as

$$u_t \;=\; \sum_{j=1}^{\infty} a_j y_{t-j} \;+\; v_t, \qquad t \;=\; 1, 2, \ldots,$$

and similarly, (if d > -0.5), using the expansion of $(1-L)^{-d}$, u_t admits a MA(∞) representation:

$$u_t \;=\; \sum_{j=0}^{\infty} c_j\, v_{t-j}, \qquad t \;=\; 1, 2, \ldots,$$

with $c_0 = 1$, and c_j decaying to zero at a hyperbolic "slow" rate (if d < 1). Therefore the parameter d also plays a crucial role in describing the time series persistence: the higher the d is, the greater is the persistent behaviour or the series. Similar to the unit root case (d = 1), v_t in (7) can be extended to allow for short run dynamics, decaying exponentially to zero.

Attempting to summarize some of the literature on hysteresis and persistence on unemployment, Gordon (1989) defined full hysteresis as the case of a unit root and persistence as stationarity AR. In that study, he concludes that no evidence of full hysteresis was found in five countries (France, Germany, USA, Japan and the UK) for the time period 1873-1986. Graafland (1991) stated that, in the 80s, the labour market in the Netherlands was characterized by high and persistent level of unemployment. Lopez et al. (1996) reported that monthly unemployment in Spain (1976m6-1994m10) was consistent with hysteresis. On the other hand, Nott (1996) did not find evidence of hysteresis in Canada, while Wilkinson (1997) did. Mitchell (1993) examines hysteresis for the OECD countries, and Leon-Ledesma (2000) analyses the case of the US states and the EU countries. Other studies on unemployment hysteresis are Song and Wu (1997), Roed et al. (1999) and Roed (1999), who proposes an alternative way of modelling unemployment hysteresis in which permanency is viewed as a continuous phenomenon.

3. Cross-section and Panel data persistence

Cross-sectional techniques may help to examine if regional UR are persistent of not. In the context of regional convergence, unemployment persistence could be evaluated by computing mean regressions of regional UR. In the regressions, the dependent variable is obtained after taking differences of the UR, where the length of the spell for which the differences are taken changes. This procedure was employed by Blanchard and Katz (1992) and Bertola and Ichino (1995). The estimated models are of the form:

$$\Delta_t u^i = \alpha + \beta u^i_q + \varepsilon \qquad (8)$$

where u^i_q is the UR in region i at the starting period q, and $\Delta_t u^i$ is the change in the UR in that region over the following t periods. If the cross-sectional coefficient of the initial UR (β) becomes more negative when the length of the spell increases, it means that regional UR are not persistent (since their reversion towards the aggregate mean is stronger over longer intervals).

Regarding panel data methods, some developments can also be implemented so that we may benefit from the better statistical properties associated with large samples. Panel data analysis permits us extracting greater information from the data and, more importantly, taking into account the possible influence of individual heterogeneity components. In particular, by isolating the unobservable individual heterogeneity, the fixed and random effects models filter the impact that these unobservable elements have upon the dependent variable.

Given the temporal dimension of the data and the nature of unemployment persistence itself, it seems imperative to specify and estimate the dynamic structure of the model. Accordingly, we also implement the technique proposed by Arellano and Bond

(1991), which is advised to deal with dynamic panel data.[1] The GMM estimators use variables in differences to eliminate unobservable individual effects, and include lagged values (in levels) as instruments to correct for simultaneity.

Finally, we also consider unit root tests for panel data, following the lines of Levin, Lin and Chu (LLC, 2002) and Im, Pesaran and Shin (IPS, 2003). The former method provides an alternative approach to weigh up the results obtained when computing ADF tests for time series data. LLC approach is appropriated for panels of moderate size (say, 10 to 250 individuals, with 25-250 time series observations) and is particularly useful wherever the (time) dimension of the panel is not large enough, so that unit root test procedures performed separately to each individual in the panel would not be sufficiently powerful. The IPS (2003) test is prescribed for panels where a higher degree of heterogeneity exists in the cross-section dynamics. Unlike LLC (2002), which assumed that all series are stationary under the alternative, IPS (2003) is consistent under the alternative that only a fraction of the series is stationary.

4. Data and results

Data on unemployment were obtained from the EPA (Encuesta de Población Activa. INE, Spain). The sample consists of quarterly (seasonally unadjusted) data for the time period 1976q3 – 2004q4. Although there were some methodological changes over this period, the data series provided by the INE has been conveniently homogenized according to the methodological guidelines given by EPA-2002.[2] The time series

[1] Arellano and Bond (1991) derived a Generalized Method of Moments (GMM) estimator using lagged levels of the dependent variable and the predetermined variables and differences of the strictly exogenous variables. This methodology assumes that there is no autocorrelation in the error term.

[2] The ongoing nature of the EPA has allowed the INE, since the 3rd quarter of 1976, to compute homogeneous data on the main characteristics of the labour market. In order to ensure such homogeneity in the results, it was required to carry out retrospective adaptations (from 1994 untill the 2nd quarter of 1987) as well as to incorporate to the survey some improvements applied since 1999. For a detailed description of changes affecting EPA-2002, see: http://www.ine.es/daco/daco42/daco4211/notamet.htm.

analysis uses aggregate UR for the whole country, as well as separate UR for each region.[3] Panel data analysis is carried out using a set of 5,700 observations for 114 quarterly periods and 50 provinces.

To measure unemployment persistence at the regional level, the analysis can be based on absolute regional UR or, alternatively, on regional relative UR. (The latter uses differentials; that is, the regional UR minus the national UR rate). Given our interest for examining unemployment persistence in Spain as a whole, we have chosen the first route. Other results based on the two approaches can be found in Jimeno and Bentolila (1998).

4.1 Results based on time series approaches

According to the classic (time series) perspective we first determine if the unemployment series is stationary or not. A visual inspection at the series, displayed in Figure 1, does not produce clear conclusions. The first 50 sample autocorrelation values are also displayed in Figure 1, and they show a very slow decay to zero. Moreover, the periodogram presents the highest value at the lowest frequency. This may be an indication that first differences are required, implying thus nonstationarity I(1). Nevertheless, these features might also be the result of a stationary AR process with the roots close to the unit circle.

[Insert Figure 1 about here]

Several methods have been proposed for testing the hypothesis of a unit root. Table 1 displays the results of the ADF and DF-GLS tests for the whole country as well as for each of the 17 regions. The latter were computed using Elliott-Rothenberg-Stock (1996) efficient test for an autoregressive unit root, which has a better overall statistical

[3] Spain is divided into 50 provinces, which are gathered in 17 regions, with different cultural traditions and political autonomy: Andalucía, Aragón, Asturias, Baleares, Canarias, Cantabria, Castilla-Mancha, Castilla-León, Cataluña, Valencia, Extremadura, Galicia, Madrid, Murcia, Navarra, País Vasco and Rioja.

performance. The procedures were carried out for the model with and without trend, and they include as many lagged first differences as indicated by k in columns (1) and (4). The tests provide no evidence against the unit root hypothesis: the unit root null cannot be rejected for any of the 17 regions at the 1%, and not even at the 5% significance level.

[Insert Table 1 about here]

However, the method presented above has very low power if the alternative is an AR process with the roots close to the unit circle (Campbell and Perron, 1991; DeJong et al., 1992), but also if they are of a fractional form (Diebold and Rudebusch, 1991; Hassler and Wolters, 1994; Lee and Schmidt, 1996).[4] Hence, we also use as a measure of persistence the fractional differencing parameter d, as described in Section 2.

The first two values in the first row in Table 2 display the estimates of the AR parameter assuming that d = 0 and that y_t follows a non-seasonal and a seasonal AR(1) process.[5] We observe that in both cases the coefficients are very close to 1: it is 0.9944 in the non-seasonal case, and 0.9518 with a seasonal AR(1) process. In the right hand side of the table we display the estimates of d based on a parametric approach of Robinson (1994). (See Appendix A for a full description of this procedure). The latter method uses a Whittle function, which is an approximation to the likelihood function. Since it is parametric, we have to specify the functional form of the I(0) disturbances u_t in (7). We try with white noise, Bloomfield and seasonal AR(1) v_t. The model of Bloomfield (1973) is a non-parametric way of modelling v_t that produces autocorrelations decaying exponentially as in the AR case (See, e.g., Gil-Alana, 2001b). We observe that for the three cases the values of d are much higher than 1, ranging

[4] See also Cook (2002) for spurious rejections using Dickey-Fuller tests.

[5] A seasonal AR(1) process can be defined as $(1-\phi L^s)u_t = \varepsilon_t$, where s is the number of time periods within a year.

between 1.318 and 1.381. Moreover, the unit root hypothesis (d = 1) is rejected in the three cases in favour of higher orders of integration.

[Insert Table 2 and Figure 2 about here]

Figure 2 displays the estimates of d based on a semiparametric approach (Robinson, 1995; See Appendix B) along with the 95% confidence interval corresponding to the I(1) hypothesis. It is semiparametric in the sense that we do not specify any functional form for the I(0) disturbances v_t. We see that practically all values are above the I(1) interval, implying orders of integration higher than 1, which is consistent with the results based on the parametric approach.

Next, we wonder if the unemployment persistence in Spain has experienced any substantial alteration along the last three decades. In other words, we are concerned with the possibility of a structural break in the data, since a number of institutional changes have taken place during this period. Various techniques can be applied in order to test structural shifts in regional UR (see, e.g., Baddeley, 1998, who studies the impact of the 1980s recession on regional unemployment in Britain). We employ here a methodology suggested by Hsu and Kuan (2004), testing simultaneously the order of integration (d) and the time of the break, which, in this context, is a single mean shift. (Appendix C).

[Insert Figure 3 about here]

The upper part of Figure 3 displays the estimates of the time break across the bandwidth number m. We see that in all except two cases, the break takes place at the fourth quarter in 1993. The lower part of the figure refers to the order of integration of the series and the results are to some extent ambiguous. If the bandwidth number m is smaller than T/4, the order of integration is strictly higher than 1, however, if m > T/4, the values are within the I(1) interval. Thus, it seems that once the break is taken into account the evidence of orders of integration higher than 1 is not as clear as before. In any case, from the analysis of the results presented so far it seems that unemployment, at an aggregate level, is nonstationary, with an order of integration equal to or larger than 1, supporting thus the hysteresis hypothesis in the unemployment series for Spain.

Next, we examine the persistence of UR at a regional level. Starting with the AR coefficients, we see (Table 2) that all values are very close to 1. The lowest values are those referring to Baleares, with the coefficients smaller than 0.90 in the two cases. At the other end, Cataluña and PaísVasco are two of the areas with the highest coefficients. Performing the parametric procedure of Robinson (1994) (in the right hand side of the table) the most noticeable features are first the fact that for Baleares, the unit root null hypothesis is rejected in favour of smaller orders of integration. This may be related to the tourism sector, which is highly predominant in this region. On the opposite side we find that Cataluña, País Vasco and Madrid, precisely among the most industrialized areas of Spain, are the regions where d is found to be higher than 1.[6]

[Insert Figure 4 about here]

Figure 4 displays the estimates of d based on the Whittle semiparametric approach of Robinson (1995). The results here can be grouped into three different categories: Baleares, with values smaller than one in most of the cases; Cataluña, País Vasco, Madrid, Aragón and Valencia, with values of d strictly above 1 in many cases; and the remaining regions, with values of d within the unit root interval in most of the cases. Thus, the results based on the semiparametric approach are completely in line with the parametric method: Baleares presents the smallest degree of UR persistence, while the most industrialized areas are those with the highest degree of persistence.

[Insert Table 3 about here]

[6] Though not reported in the paper we also examined the orders of integration of unemployment disaggregating the series by sectors. It was found that the lowest degree of integration was obtained for the Services sector. On the other extreme, Industry had the highest levels of persistence. Tolvi (2003) found that unemployment in Finland is less persistent for females and young people. We leave for future research the analysis of persistence for groups of age and sex in Spain.

In view of the potential presence of a break taking place at 1993q4, we compute again, in Table 3, the parametric approaches but using now data ending at 1993q4. We see here that values of d higher than 1 are once more obtained for Cataluña, Valencia, Aragon, Madrid and País Vasco. However, for the aggregate as well as for most of the regions, the order of integration is higher in this period compared with the whole sample, meaning that persistence might have been reduced since 1993q4 onwards.

We can summarize the previous results by saying that unemployment in Spain is highly persistent, a result that is inferred from the ADF and DF-GLS tests. More significantly, when using autoregressions, the coefficients are very close to 1, while when adopting fractionally integrated approaches, the values of d are higher than 1 in all cases. Besides, if a break is taken into account, it seems to take place at 1993q4, and the level of persistence declines from that period onwards. From a regional perspective, Baleares presents the lowest degree of persistence, due possibly to the influence of the tourism sector, predominant in this region. On the contrary, some of the most industrialized areas (Cataluña, Madrid and País Vasco) are those with the highest levels of persistence.

Our results are so far consistent with other previous studies on unemployment in Spain (e.g. Jimeno and Bentolilla, 1998), finding evidence of high levels of persistence at both aggregate and regional level. One explanation to this evidence might be the low geographical labour mobility observed in Spain during the considered period (e.g. Garcia-Rubiales, 2005). This idea is congruent with the results for Baleares, which is presumably the region with the highest labour mobility and the lowest UR persistence.

4.2 Results based on cross-section and panel data approaches

The analysis made so far is based exclusively on time series approaches. Yet, the nature of unemployment persistence can be better understood exploiting the information contained in a panel data set. Applying the empirical methods that were described in Section 3, we examine, from different perspectives, the persistence of Spanish UR. The sample includes quarterly data for 114 periods and 50 provinces.

Once the temporal dimension comes into the scene, we need to pay attention to the periods in which institutional features have taken place in Spain. For this reason, and for the sake of simplicity, we have arranged the cross-sectional analysis considering homogeneous periods of 6 years each (starting at 1976 and attending specially to 1982, 1988, 1994 and 2000). Note that Spain joined the European Union (EU) in 1986 and the European Monetary System (EMS) in 1989. After some initial transitional years, the Socialist Party (PSOE) formed the government from 1982 until 1996. Since then, and almost until the end of the sample period, the conservative (right-wing) Popular Party (PP) has been driving the country.[7]

Following the lines of previous studies (Jimeno and Bentolila, 1998; Wu, 2003; etc.), we offer an initial flavour of Spanish unemployment persistence in Figure 5. This scatter graph represents provincial UR (annual averages from quarterly data), for different years, against the corresponding records for the year 1976. The positive slope in all the graphs of Figure 5 means that the ranking of provinces according to their UR has remained noticeable stable over the years. This fact suggests a high degree of persistence of regional unemployment differentials, which is noticeable broader than expected within an integrated economy.

[Insert Figure 5 and Table 4 about here]

To get a more accurate perception of this issue, Table 4 gathers the estimated coefficients for several regressions between provincial UR for two different years. In all cases, the significance of the estimated coefficients is remarkably strong as well as the

[7] Some major economic and political events in Spain are summarized by Bentolila and Blanchard (1990), p.238. The reforms of the Spanish labour market are described in European Commission (2005), p.49.

explanatory power of the regressions. The results in the first row in Table 4 seem to indicate increasing divergence of unemployment persistence in Spain until 1994. From that year onwards the trend of regional UR persistence has been reverted, inasmuch as the estimated coefficients (the slopes) experience a decline after that year. From inspection of the whole table, we can conclude that regional UR persistence in Spain increased between 1982 and 1988, remained high and stable between 1988 and 1994, and declined consistently since 1994. Note that the decline experienced since 1994 is congruent with the above results.

In addition, we also compute mean regressions of regional UR. The results vary depending both on the length of the spell for which the differences were taken (indicated in rows) and on the starting period (in columns). The results in Table 5 report the point estimates and t-statistics of β from cross-sectional regressions starting at different time periods. The estimations were computed for the annual averages obtained from quarterly data. The constant is not shown and absorbs the effects of aggregate UR changes.

[Insert Table 5 about here]

The results manifest an interesting contrast. If we take 1976 as the starting year, the values are positive and statistically significant up to the point in which we have a spell of length 24 years. (In fact, we have to wait 28 years to find a negative sign). It means that extremely high persistence exists, since reversion towards the aggregate mean takes long time to occur. The analysis of column (2) suggests that strong persistence still exists, at least, until 1996, whereas the negative signs right from the first row of columns (3), (4) and (5) indicates a change of behaviour in the unemployment persistence. It is certainly the case for the starting year 1994, in which reversion towards the aggregate mean is much faster than in previous periods. When considering the starting years 1988 and 1994, we realise that the negative coefficients become significant after 10 and 4 years respectively, suggesting that regional unemployment persistence is lower since 1998 onwards. This result is consistent with the mentioned structural break and could be explained by the labour market reforms introduced in Spain in the years 1994 and 1997.[8]

In summary, the previous results allow us to conclude that the dispersion of regional UR has persistently increased in Spain during the 80s, but it has experienced a reverted trend at the mid of the 90s. It has declined perhaps since 1994, and undeniably since 1998.

Up to now, we have studied unemployment persistence either adopting time series or cross-sectional methodologies. Next, we focus on panel data results. Table 6 collects the OLS and GLS estimates for the pooled model. Given that the estimations may have problems of heteroskedasticity, the former estimates were computed for robust standard errors, while the GLS estimations were performed specifying a heteroskedastic error structure.

[Insert Table 6 about here]

The results in Table 6, columns (3) and (6), include 17 dummies corresponding to the regions. The magnitude of the AR(1) coefficient, close to one in all the models, corroborates the strong persistence of Spanish unemployment, which is slightly lessen whenever regional dummies are present. More interestingly, the temporal dummies, that were included to explore breaks in unemployment behaviour, suggest a positive variation in 1982, whereas they manifest a significant decrease in 1994 and 2000. Note that this is again consistent with a break taking place in the last quarter of 1993, as time series analysis pointed out. It is also consistent with the conclusions reached from the cross-sectional results of Table 5.

In any case, the information contained in the data has not yet been fully exploited, and could be enhanced by applying the fixed and random effects models. Table 7 collects, in columns (7) and (8), the results for the fixed effects model, while columns (9) and (10)

[8] First, in 1994 a number of changes took place. Economic circumstances were included in the reasons justifying individual dismissal procedures, which made it easier to occur. Unemployment benefit would come to be considered as taxable income and a range of issues on working conditions were no longer regulated. In that year the principle of causality was also re-established as a general rule for fixed-term contracts. Then, in 1997 a new permanent contract with lower dismissals costs was approved.

refer to the random effects model. The results are basically consistent with the others already obtained, meaning that high persistence unemployment exists even after controlling for the elements of heterogeneity of the Spanish provinces.

[Insert Table 7 about here]

Table 7 also shows, in Columns (11) and (12), the estimates obtained through the GMM technique of Arellano and Bond (1991) described in Section 3. The results were computed using robust standard errors to avoid heteroskedasticity.[9] Given that we deal with quarterly data, we have included in the model, as instruments, four lags of the dependent variable. We also display in Table 7 the statistics for the first- and second-order autocorrelation of the residuals (m_1 and m_2, respectively). As expected, the residuals of the estimated equation show first- but not second-order autocorrelation.[10]

In summary, panel data analysis enables us to state that UR persistence in Spain has remained remarkably high in the last decades: the AR(1) coefficient is hovering around 0.95 for the majority of the estimated panel data models. It also suggests that the persistence might have experienced a substantial decline in 1994. A drawback of applying Arellano and Bond's (1991) procedure to our data set is that it is designed under the "classical" panel setting, where T is small and N is large, which is not the case in the present work. Moreover, it imposes stationarity, while previous results presented in this work indicated that unemployment was nonstationary.[11]

Finally, we perform panel unit root tests, which provide dramatic improvements in power compared to performing separate unit root tests for each individual series. In Table 8

[9] On the other hand, the Sargan test from the one-step homoskedastic estimator rejected the null hypothesis that the over-identifying restrictions are valid, but this could be due to heteroskedasticity.

[10] The presence of first-order autocorrelation in the differenced residuals does not imply that the estimates are inconsistent, while it would be the case if the second-order autocorrelation were found. See Arellano and Bond, 1991, pp. 281-2.

[11] Note, however, that even with the novel techniques used here, it is still statistically difficult to distinguish between nonstationarity (with a unit root) and stationarity with AR coefficients close to the unit circle.

we report the results of computing LLC (2002) procedure, while Table 9 shows the IPS (2003) unit root tests. The estimated coefficients of the intercepts, in the single-equation models in Table 7, exhibited important differences between regions and, therefore, a common constant for the whole panel should not be included. Instead, the estimations allow for heterogeneous intercepts.

[Insert Table 8 about here]

Recall that the ADF and DF-GLS tests for time series failed to reject the unit root null for all regions. However, the low power of these tests against the stationary alternative, when the process is near-integrated, is a well-known problem (see, for instance, León-Ledesma and McAdam, 2004). Now, using a panel data set, we apply more powerful procedures and benefit from a reasonable large panel, collecting 50 provinces and 114 time periods. The degree of persistence for each individual regression error is allowed to vary freely across the 50 provinces.

The LLC (2002) paper proves that their t-star statistic is standard normal under the null of nonstationarity. We have included in equation (10) as many lagged first differences, for each province, as indicated by k in Table 7. Since Table 7 was computed for the regions, we assume here the same k for all the provinces comprised in the same region.[12] The results in Table 8 indicate that, once again, we cannot reject the unit root null for aggregate Spanish unemployment, implying that the question about unemployment persistence in Spain finds again an affirmative answer. Next, we make up panels gathering all the provinces which belong to the same region,[13] even if most of them do not reach to the 8 or 10 individuals,

[12] The estimations based on different k for each individual (instead of enforcing the same k for all the provinces within the same region) yielded very similar results and are not shown.

[13] The devise of grouping "ad hoc" sub-panels for particular purposes, has been applied even for the case of just a few individuals (Cf. Culver and Papell, 1997). The shorter is the panel, the more difficult is that the cross-section variation helps to reject the unit root null, but whenever it happens, the result is very strong. Note also that LLC (2002) test is to some extent accurate for small panels, since "the normal distribution provides a good approximation to the empirical distribution of the test statistic in relatively small samples".

which is a minimum required to make an accurate implementation of the test. Then, we gather the provinces in groups of 8 or 9, attending mainly to the value of k in Table 1.

It is meaningful the fact that there are some regional panels for which the unit root null is not rejected.[14] Not surprisingly, they correspond precisely to those regions for which fractional integration methods reported a greater degree of UR persistence: Cataluña, País Vasco and Madrid. The corresponding P-value, for the model without time trend in column (3), indicates that there is a certain significant probability of failure of rejecting the unit root null, which is consistent with the finding that these three regions present the strongest unemployment persistence. More significantly, the P-value for the panel gathering these three regions together is as high as [0.031] for the model without trend.

Note, on one hand, that LLC (2002) test has been criticised for assuming, under the alternative of stationarity, that all cross-sections converge to the equilibrium at the same speed of adjustment. That should not be a serious problem in the present paper as far as we are precisely interested in the cases in which we fail to reject the null hypothesis of unit roots. On the other hand, if some degree of heterogeneity between panels exists, the prescribed test procedure is the IPS (2003) unit root test. Even though the properties of both types of tests depend upon the independence assumption across individuals, we venture that they may be of some help in some applied studies like this.[15] The results of these tests, shown in Table 9, yield basically the same conclusions than those in Table 8. In all the models we have included 8 lagged first differences.[16] The tests clearly fail to reject the unit

[14] That is not the case for the majority of the regions, implying that the individual unemployment series included in those regions should be considered I(0). This panel-based test could not be carried out for the case of regions containing a number of provinces smaller that 3.

[15] As stressed by Strauss (2003), subtracting cross-sectional means to remove common time specific effects is a quite common procedure but, if there is heterogeneity in the cross sectional correlation, it will only partially reduce the problem. Regarding the LLC test, it requires independence across individuals, assumption that can be somewhat relaxed to allow for some degree of dependence by including time-specific intercepts, which does not affect the limiting distributions of the test (Cf.: LLC, 2002, p. 13).

[16] Note that IPS (2003) pointed out that the tests perform better if the orders of the underlying ADF regressions are correctly chosen or over-estimated.

root null both in the case of the national aggregate UR and for the panel that gathers the three mentioned regions. Furthermore, this result occurs regardless whether the trend is included or not. By comparing the last two rows of Table 9, IPS (2003) tests teach us that the unit root for Valencia can be discarded.

5. Concluding comments

In this paper we have examined the persistence of Spanish unemployment in the last three decades, paying especial attention to the regional unemployment persistence. From a time series viewpoint, we first computed ADF and DF-GLS unit root tests. Then, we also measure UR persistence through the AR coefficients and by fractional differencing parameters. The cross-section analysis consisted basically of running mean regressions of regional UR. For the panel data methodology, we have employed a panel of 114 periods and 50 provinces. We run regressions for the pooled model as well as for the fixed and random effects models. The Arellano and Bond method for estimating dynamic panels was applied too. Finally, we performed LLC (2002) and IPS (2003) panel data unit root tests.

Among other findings, this empirical analysis has taught us three major lessons. First, that UR persistence (in aggregate as well as in regional terms) has been very strong in Spain throughout the last 25 years. Second, that it experienced a decline around 1994. Finally, that some of the most industrialized areas, like Cataluña, Madrid and País Vasco, endure the greatest unemployment persistence across the sample. These three regions are also the areas that present the smallest share of the labour force in agriculture and where the proportion of population living in large cities is among the greatest across the country.

Regarding the first conclusion, the high level or persistence at both aggregate and regional level in Spain is probably similar to the Italian experience, whereas it is in sharp contrast with other European countries or with the US (Cf. Jimeno and Bentolila, 1998). This feature might suggest that there are some types of idiosyncratic factors (common to other Southern European countries) that explain the high persistence of unemployment. Even considering the reforms carried out in the Spanish labour market in 1994 and 1997, rigidities might have prevailed due to the inertia of legislation that has only recently been annulled.

However, other authors consider that institutional factors are not the main sources provoking high unemployment. Bertola and Ichino (1995) argue that, since institutional rigidities already existed in the past, other factors ought to be invoked to explain the high UR experienced in Europe since the middle 1980s. On their view, both the high unemployment in Europe and the increasing wage dispersion in the US must stem from the same cause. Then, institutional rigidities would have only triggered the process leading to high UR, but other forces (like a more rapid idiosyncratic shocks between sectors, suggested by Lilien, 1982) must be the ultimate cause of it.

For the case of Spain, geographic labour mobility should certainly be taken into account. Since the time in which Spain became a democratic country in 1975, a persistent characteristic of its labour market has been the low rate of interregional immigration, despite of the disparities in the UR across regions (Jimeno and Bentolila, 1998). Garcia-Rubiales (2005) moves along the same position when states that Spanish unemployment is more closely linked to the lack of labour mobility than to institutional rigidities.[17] We venture that the lack of labour mobility across Spanish regions could be the root of not just of high UR level but also of high unemployment persistence. Note that they are different concepts, since high persistence in a region can be experienced

[17] This can be related to the number of temporary jobs too. Spain is the EU member with the highest level of temporary employment, which affects the duration distribution of unemployment (Cf. Guell, 2003).

accompanied by low UR. And yet, the fact is that high UR become more harmful as long as it coexists with high UR persistence.

Other meaningful finding of this study is the lower unemployment persistence experienced in Spain from the middle 1990s onwards. Why did such a decline in unemployment persistence occur around 1994? We have already mentioned that important labour market reforms were introduced precisely in that year. Besides that, by 1994 Spain had already had time enough to accommodate its integration into the EU (1986) and the EMS (1989). Besides, note that other important institutional changes took placed at that time, linked to government changes coming from the European, regional and general elections of 1994, 1995 and 1996, respectively.

According to the report for Spain of the European Commission (2005, p. 5): "*High and persistent unemployment prompted a labour market reform in 1984, which was instrumental to job creation. During the economic boom in the second half of the eighties, the UR went down to around 16% in 1991. Labour rigidities, which enhanced the persistence of adverse shocks (Dolado and Jimeno, 1997) coupled with the incorporation of the baby boom cohorts and women into the labour market in the late 80s and the early 90s, pushed the UR above 20% in the 1993 slowdown. Since the second half of the 90s the UR fell significantly below 15% to nearly 11% in 2003 (…). Moreover, the resilience of employment growth has been particularly impressive during the last ten years. Since 1995, more than 25% of the total net job creation in the euro area was registered in Spain. Such (un)employment performance could point to a major structural change with respect to the past*".

The previous comments are basically consistent with our empirical findings. Since 1995, the Spanish employment rate has increased briskly by more than 13%, standing now close to 60%. This might be the main reason for the structural break that we have found around 1994. Moreover, the UR in Spain, although still the highest in the Euro-area, has been steadily falling from its peak of around 23% in 1994 to slightly above 11% in 2003. The positive change in labour market outcomes with respect to past trends might also be explained by: (i) the stability-oriented policies associated to the EMU membership; (ii) the labour market reforms as well as the moderation in nominal unit labour costs; and (iii) the expansion of some labour-intensive sectors (construction and some services), which have significantly contributed to sustain employment growth.

Finally, we have found greater persistence in some regions, namely: Cataluña, Madrid and País Vasco (and, in a minor extent, in Valencia and Aragón). The reasons for this evidence are not straightforward. In any case, the issue must be placed into the context of regional unemployment divergence between regions, since there are large disparity in the geographical distribution of UR. Andalucía and Extremadura recorded the highest UR in 2004, around 18%; whereas the opposite geographical and socio-economic extremes, Navarra and La Rioja, registered rates about 5%. Madrid exhibited a rate of 6.5% and Cataluña and País Vasco of about 9%.

The study accomplished by the European Commission (2005) justifies the disparities across regions on the bases of low geographical mobility of the labour force, which in turn is amplified by the characteristics of the housing market (poor development of rental housing, etc.). However, regardless of the origin of such persistence in regional unemployment dispersion, it does not explain why Madrid, Cataluña and País Vasco are the regions with greater unemployment persistence. The first one is by far the region in which a bigger proportion of the population lives in large cities, while the other two are among the most industrialized regions in Spain. These are also the three regions in which the smallest percentage of labour force is enrolled in the agriculture sector. Further study must be done to explain the causes behind this facts.

Appendix A

The LM test of Robinson (1994) for testing H_o: d = d_o in a model given by

$$u_t \;=\; \beta' z_t \;+\; x_t, \qquad t \;=\; 1, 2, ... \tag{A1}$$

and (8) is:

$$\hat{r} \;=\; \frac{T^{1/2}}{\hat{\sigma}^2} \hat{A}^{-1/2} \hat{a},$$

where T is the sample size and:

$$\hat{a} \;=\; \frac{-2\pi}{T} \sum_{j=1}^{T-1} \psi(\lambda_j) g(\lambda_j; \hat{\tau})^{-1} I(\lambda_j); \qquad \hat{\sigma}^2 \;=\; \sigma^2(\hat{\tau}) \;=\; \frac{2\pi}{T} \sum_{j=1}^{T-1} g(\lambda_j; \hat{\tau})^{-1} I(\lambda_j);$$

$$\hat{A} \;=\; \frac{2}{T} \left(\sum_{j=1}^{T-1} \psi(\lambda_j)^2 \;-\; \sum_{j=1}^{T-1} \psi(\lambda_j) \hat{\varepsilon}(\lambda_j)' \times \left(\sum_{j=1}^{T-1} \hat{\varepsilon}(\lambda_j) \hat{\varepsilon}(\lambda_j)' \right)^{-1} \times \sum_{j=1}^{T-1} \hat{\varepsilon}(\lambda_j) \psi(\lambda_j) \right)$$

$$\psi(\lambda_j) = \log \left| 2 \sin \frac{\lambda_j}{2} \right|; \qquad \hat{\varepsilon}(\lambda_j) \;=\; \frac{\partial}{\partial \tau} \log g(\lambda_j; \hat{\tau}); \qquad \lambda_j \;=\; \frac{2\pi\, j}{T}; \qquad \hat{\tau} \;=\; \arg\min \sigma^2(\tau).$$

Note that z_t in (A1) is a (kx1) vector of deterministic variables that might include an intercept (i.e. z_t = 1) or an intercept and a linear time trend. Also, $\hat{a}$ and $\hat{A}$ in the above expressions are obtained through the first and second derivatives of the log-likelihood function with respect to d (see Robinson, 1994, page 1422, for further details). $I(\lambda_j)$ is the periodogram of u_t evaluated under the null, i.e.:

$$\hat{v}_t \;=\; (1 \;-\; L)^{d_o} u_t \;-\; \hat{\beta}' w_t; \quad \hat{\beta} \;=\; \left(\sum_{t=1}^{T} w_t\, w_t' \right)^{-1} \sum_{t=1}^{T} w_t\, (1 \;-\; L)^{d_o} u_t; \qquad w_t \;=\; (1 \;-\; L)^{d_o} z_t,$$

and g is a known function related to the spectral density function of u_t:

$$f(\lambda; \sigma^2; \tau) \;=\; \frac{\sigma^2}{2\pi} g(\lambda; \tau), \qquad -\pi \;<\; \lambda \;\leq\; \pi.$$

Appendix B

The Whittle estimate of Robinson (1995) is implicitly defined by:

$$\hat{d} = \arg\min{}_d \left(\log \overline{C(d)} \;-\; 2\,d\, \frac{1}{m} \sum_{j=1}^{m} \log \lambda_j \right), \tag{B1}$$

$$for \;\; d \in (-1/2\,,\; 1/2)\,; \quad \overline{C(d)} \;=\; \frac{1}{m} \sum_{j=1}^{m} I(\lambda_j)\, \lambda_j^{2d}, \quad \lambda_j = \frac{2\,\pi\, j}{T}, \quad \frac{m}{T} \to 0.$$

where m is a bandwidth parameter number.

Appendix C

The starting model in Hsu and Kuan (2004) is:

$$u_t = \begin{cases} \mu_1 + x_t, & t = 1, 2, \dots, T_b \\ \mu_2 + x_t, & t = T_b + 1, \dots, T \end{cases}, \tag{C1}$$

where T_b is unknown and x_t is given by (1). The procedure is simple. For each hypothetical change point (T_b) we estimate first μ_1 and μ_2 in (C1) from the pre- and post-change observations, i.e

$$\mu_1 = \frac{1}{T_b} \sum_{t=1}^{T_b} u_t; \text{ and } \mu_2 = \frac{1}{T - T_b} \sum_{t=T_b+1}^{T} u_t.$$

The residuals are

$$\hat{x}_t = \begin{cases} u_t - \hat{\mu}_1, & t = 1, 2, \dots, T_b \\ u_t - \hat{\mu}_2, & t = T_b + 1, \dots, T \end{cases},$$

and the periodogram of $\hat{x}_t$, (evaluated at the discrete Fourier frequencies, $\lambda_j = 2\pi j/T$) is given by:

$$I_{\hat{x}}(\lambda_j) = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} \hat{x}_t \, e^{i \lambda_j t} \right|^2.$$

Then, for each change point T_b, an estimate of d can be obtained by minimizing the Whittle function in the frequency domain, using a band of frequencies that degenerates to zero, (see, e.g., Robinson, 1995), i.e.,

$$\hat{d}(T_b) = \arg\min_d L(d(T_b), T_b) = \arg\min_d \left(\log \overline{C(d)} - 2\, d \, \frac{1}{m} \sum_{j=1}^{m} \log \lambda_j \right),$$

$$for \;\; d \in (-1/2, 1/2); \;\; \overline{C(d)} = \frac{1}{m} \sum_{j=1}^{m} I_{\hat{x}}(\lambda_j)\, \lambda_j^{2d}, \;\; \lambda_j = \frac{2\pi j}{T}, \;\; \frac{m}{T} \to 0,$$

where m is a bandwith parameter number. The change point estimator $\tilde{T}_b$ is then

$$\arg\min_{T_b} L(\hat{d}(T_b); T_b)$$

where $T_b \in [T_1, T_2] \subseteq (1, T)$, and the estimator for d is $\tilde{d}(\tilde{T}_b)$. Hsu and Kuan (1998) showed that under very mild regularity conditions $\tilde{T}_b$ is a consistent estimate for T_b and thus, following Robinson (1995),

$$\sqrt{m}\,(\hat{d} - d_o) \to_d N(0, 1/4) \quad as \;\; T \to \infty.$$

References

Arellano, M., S. Bond, 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, The Review of Economic Studies 58, 277-297.

Baddeley, M., 1998. The impact of the early 1980s recession on British regional unemployment Applied Economics 30, 19-30.

Barro, R. J., 1988. The persistence of unemployment, American Economic Review 78, 32-37.

Bean, C., 1997. The role of demand management policies in reducing unemployment, Chapter 4 in Snower DJ and G de la Dehesa (1997) (eds). Unemployment Policy: Government options for the labour market, CEPR Conference Vol, Cambridge University Press, 83-111.

Bentolila, S. and O. Blanchard, 1990. Spanish unemployment, Economic Policy, 233-280.

Bertola, G. and A. Ichino, 1995. Wage Inequality and Unemployment: US vs Europe, NBER Macroeconomics Annual, MA, 13-54.

Blanchard, O.J. and L.H. Summers, 1986. Hysteresis and the European unemployment problem, NBER Macroeconomics Annual, Cambridge, MA, 15-78.

Blanchard, O.J. and L.H. Summers, 1987. Hysteresis in Unemployment, European Economic Review, 31 (1-2), 288-295.

Blanchard, O. J. and L. F. Katz, 1992. Regional evolutions. Brookings Papers on Economic Activity 1, 1-74.

Blanchard, O. J. et *al*, 1995. Spanish unemployment: Is there a solution?, London: Centre for Economic Policy Research.

Bloomfield, P., 1973. An exponential model in the spectrum of a scalar time series, Biometrika 60, 217-226.

Campbell, J.Y. and P. Perron, 1991, Pitfalls and opportunities. What macroeconomists should known about unit roots, NBER Macroeconomic Annual, 1, 141-201.

Caporale, G.M. and L.A. Gil-Alana, 2002. Unemployment and input prices. A fractional cointegration approach, Applied Economics Letters 9, 347-351.

Culver, S. and D.H. Papell, 1997. Is there a unit root in the inflation rate? Evidence from sequential break and panel data models, Journal of Applied Econometrics 12, 435-444.

Crato, N. and P. Rothman, 1996. Measuring hysteresis in unemployment rates with long memory models, Working Paper, East Carolina University, Department of Economics.

Cook, S., 2002, Spurious rejection using Dickey-Fuller tests, Applied Economics Letters 9, 557-562.

Cook, R. and S. Weisberg, 1983. Diagnostics for heteroscedasticity in regression, Biometrika 38: 159-178.

DeJong, D.N., J. Nankervis, N.E. Savin and C.H. Whiteman, 1992. The power problems of unit root tests in time series with autoregressive errors, Journal of Econometrics, 53, 323-343.

Dickey, D.A. and W.A. Fuller, 1979. Distributions of the estimators for autoregressive time series with a unit root, Journal of the American Statistical Association, 74, 427-431.

Diebold, F.X. and G. Rudebusch, 1991. On the power of Dickey-Fuller tests against fractional alternatives, Economic Letters 35, 155-160.

Dolado, J.J and J.F. Jimeno, 1997. The causes of Spanish unemployment; a structural VAR approach, European Economic Review 41, 1281-1307.

Elliot, G., T. Rothenberg and J. Stock, 1996. Efficient tests for an autoregressive unit root, Econometrica, 64, 813-36.

European Commission, 2005. Country of study: Spain in EMU: a virtuous long-lasting cycle? Occasional Papers nº 14, February 2005.

Friedman, 1968. The Role of Monetary Policy, American Economic Review, 58 (1), 1-17.

Garcia-Rubiales, V., 2005, Unemployment in Spain: An analysis of labor mobility and young adult unemployment. Working paper, Stanford University.

Gil-Alana, L.A., 2001a. A Fractionally Integrated Exponential Model for UK Unemployment, Journal of Forecasting 20, 329-340.

Gil-Alana, L.A., 2001b. The Persistence of Unemployment in the USA and Europe in Terms of Fractionally ARIMA Models, Applied Economics, 33 (10), 1263-1269.

Gil-Alana, L.A., 2002. Modelling the Persistence of Unemployment in Canada, International Review of Applied Economics, 16 (4), 465-477.

Gordon, R.J., 1989. Hysteresis in history. Was there ever a Phillips curve? American Economic Review, Papers and Proceedings 79, 220-225.

Graafland, J.J., 1991. On the causes of hysteresis in long term unemployment in the Netherlands, Oxford Bulletin of Economics and Statistics 53, 155-170.

Guell, M., 2003. Fixed-Term Contracts and the Duration Distribution of Unemployment, IZA Discussion Paper, No. 791 (www.iza.org)

Hassler, U. and J. Wolters, 1994. On the power of unit root tests against fractional alternatives, Economic Letters 45, 1-5.

Hsu, C.C. and C.M. Kuan, 1998, Change-point estimation of fractionally integrated processes, Journal of Time Series Analysis 19, 693-708.

Hsu, C.C. and C.M. Kuan, 2004. Long memory or structural change: Testing method and empirical examination, Mimeo, National Central University of Taiwan, Paper presented at the Econometric Society 2000 World Congress, Seattle.

Im, K., H. Pesaran and Y. Shin, 2003, Testing for unit root in heterogeneous panels, Journal of Econometrics 115, 53-74.

Jimeno, J.F. and S. Bentolila, 1998. Regional unemployment persistence (Spain, 1976-1994), Labour Economics 5, 25-51.

Lee, D. and P. Schmidt, 1996. On the power of the KPSS test of stationarity against fractionally integrated alternatives, Journal of Econometrics 73, 285-302.

Leon-Ledesma, M., 2000. Unemployment hysteresis in the US states and the EU. A panel approach, International Review of Applied Economics 14, 55-69.

Leon-Ledesma, M.A. and P. McAdam, 2004. Unemployment, hysteresis and transition, Scottish Journal of Political Economy,

Levin, A., C. Lin and C.J. Chu, 2002. Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties, Journal of Econometrics, 108 (1), 1-24.

Lilien, D. M, 1982. Sectoral Shifts and Cyclical Unemployment, Journal of Political Economy, 90 (4), 777-793.

Lopez, H., E. Ortega and A. Ubide, 1996. Explaining the dynamics of Spanish unemployment, Working Paper in Economics, no.96/14, European University Institute.

Mitchell, W.F., 1993, Testing for unit roots and persistence in OECD unemployment rates, Applied Economics 25, 1489-1501.

Ng, S. and P. Perron, 2001. LAG Length Selection and the Construction of Unit Root Tests with Good Size and Power, Econometrica 69 (6), 1519-54.

Nott, L., 1996. Hysteresis in the Canadian labour market. Evidence from the 1990s, Research Report No. 9605, Department of Economics, University of Western Ontario.

Phelps, 1967. Philips Curves, Expectations of Inflation and Optimal Unemployment, Economica, 34 (3), 254-281.

Phelps, 1968. Money-Wage Dynamics and Labor-Market Equilibrium, Journal of Political Economy, 76 (4), 678-711.

Robinson, P.M., 1994. Efficient tests of nonstationary hypotheses, Journal of the American Statistical Association 89, 1420-1437.

Robinson, P.M., 1995. Gaussian semi-parametric estimation of long range dependence, Annals of Statistics 23, 1630-1661.

Roed, K., 1999. A note on the macroeconomic modelling of unemployment, Applied Economic Letters 9, 255-258.

Roed, K., Raaum, O. and Goldstein, M., 1999. Does unemployment cause unemployment? Micro evidence from Norway, Applied Economics 31, 1207-1218.

Schwert, G.W., 1989. Test for unit roots; A Monte Carlo investigation, Journal of Business and Economics Statistics 7, 147-160.

Song, F.M. and Y. Wu, 1997. Hysteresis in unemployment. Evidence from 48 states, Economic Inquiry 25, 235-244.

Strauss, J. and T. Yigit, 2003. Shortfalls of Panel Unit Root Testing; Economic Letters 81, 309-13.

Tolvi, J., 2003. Unemployment Persistence of Different Labour Force Groups in Finland; Applied Economcis Letters 10 (8), 455-58.

Tschernig, R. and K.F. Zimmermann, 1992. Illusive persitence in German unemployment, Recherches Economiques du Lovaine 58, 441-453.

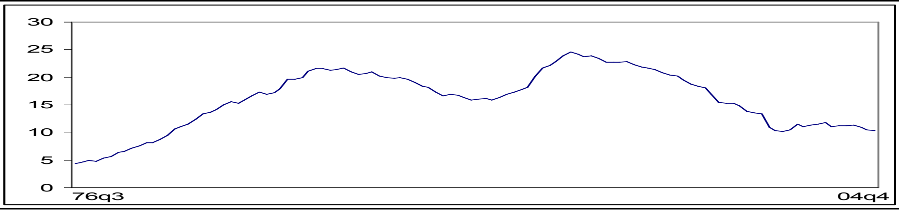
Wilkinson, G., 1997. A micro approach to the issue of hysteresis in unemployment; evidence from the 1988-1990 labour market activity survey", Bank of Canada, Working Papers 97-12.

Wu, Z., 2003. The persistence of regional unemployment: evidence from China, Applied Economics 35, 1417-1421.

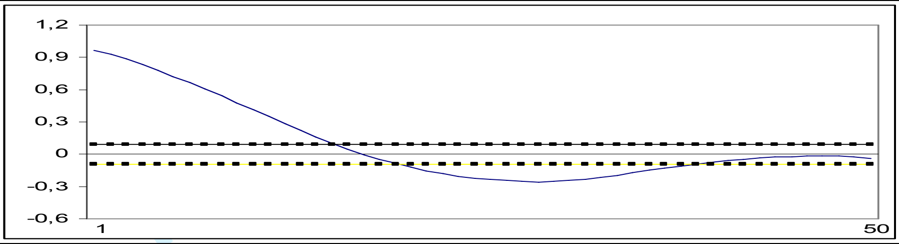
FIGURE 1

Unemployment rate in Spain, with its corresponding correlogram and periodogram

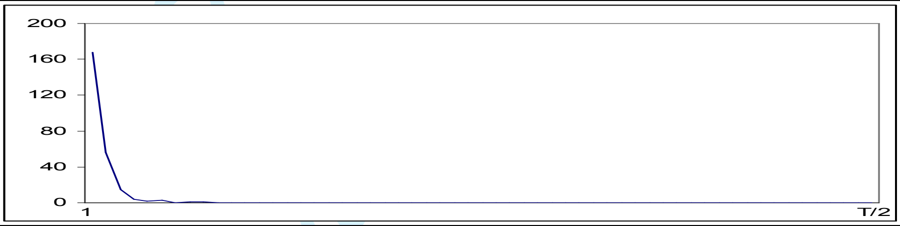
a) Original time series



b) Correlogram



c) Periodogram



The large sample standard error under the null hypothesis of no autocorrelation is $1/\sqrt{T}$. The periodogram was computed based on the discrete Fourier frequencies $\lambda_j = 2\pi j/T$.

TABLE 1

ADF unit root test (1976q3:2004q4)

	Without trend			With trend		
Critical values	**1%**	**5%**	**10%**	**1%**	**5%**	**10%**
ADF [a]	–3.51	–2.89	–2.58	–4.04	–3.45	–3.15
DF-GLS (ERS) [b]	–2.60	–1.95	–1.61	–3.56	–3.02	–2.73
	(1)	(2)	(3)	(4)	(5)	(6)
	k	t_α **(ADF)**	t_α **(dfgls)** [c]	***k***	t_α **(ADF)**	t_α **(dfgls)** [c]
TOTAL	10	–1.99	–0.88	8	–2.30	–1.52
ANDALUCIA	4	–1.71	–0.76	5	–1.23	–0.91
ARAGON	9	–1.71	–0.98	9	–1.94	–1.08
ASTURIAS	0	–2.07	–0.70	9	–1.56	–0.96
BALEARES	12	**–2.87**	–1.10	12	–2.89	–1.63
CANARIAS	12	–1.98	–1.15	12	–2.06	–1.36
CANTABRIA	9	**–2.60**	–0.95	11	–2.36	–1.82
CASTILLA-LEON	4	–2.08	–0.82	12	–1.44	–1.16
CASTILLA-MANCHA	12	–2.57	–1.02	12	–2.31	–1.70
CATALUÑA	8	**–2.81**	–1.40	8	**–3.17**	–1.86
COM.VALENCIANA	7	–2.35	–0.90	7	–2.14	–1.19
EXTREMADURA	10	–1.88	–0.53	10	–1.10	–0.70
GALICIA	10	–2.37	–0.46	10	–1.12	–1.02
MADRID	8	–2.29	–1.37	8	–2.39	–1.75
MURCIA	8	–2.37	–1.10	8	–2.05	–1.43
NAVARRA	8	–1.88	–1.27	8	**–3.29**	–1.37
PAIS VASCO	4	–2.27	–0.92	6	–2.27	–1.31
RIOJA	8	–2.30	–1.14	8	–2.19	–1.37

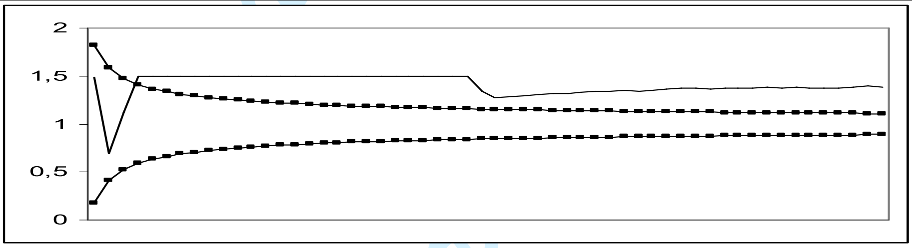
a ADF: interpolated Dickey-Fuller critical values.
b ERS: represent interpolated critical values from Elliot, Rothenberg, and Stock.
c Modified Dickey-Fuller *t*-test for a unit root in which the series has been transformed by a GLS regression. These are appropriate to be confronted with the ERS critical values.

TABLE 2					
Unemployment persistence in Spain. Aggregated and disaggregated by regions (1976q3:2004q4)					
	d = 0		Estimates of d		
	AR (1)	Seasonal AR(1)	White noise	Bloomfield (1)	Seasonal AR(1)
TOTAL	0.9944	0.9518	**1.381** [a]	**1.318** [a]	**1.321** [a]
ANDALUCIA	0.9901	0.9493	**1.103** [a]	1.136	1.041
ARAGON	0.9839	0.9157	**1.104** [a]	**1.250** [a]	1.053
ASTURIAS	0.9802	0.9146	1.014	1.050	0.999
BALEARES	0.8360	0.8555	0.668 [b]	0.314 [b]	0.699 [b]
CANARIAS	0.9869	0.9382	1.078	**1.128** [a]	1.068
CANTABRIA	0.9845	0.9281	1.059	1.092	1.045
CASTILLA-LEON	0.9905	0.9575	1.092 [a]	1.047	0.996
CASTILLA-MANCHA	0.9798	0.9356	1.016	0.969	0.939
CATALUÑA	0.9907	0.9241	**1.315** [a]	**1.363** [a]	**1.239** [a]
COM.VALENCIANA	0.9904	0.9397	**1.166** [a]	**1.496** [a]	1.088
EXTREMADURA	0.9694	0.9141	0.881	0.811 [b]	0.849
GALICIA	0.9888	0.9617	1.035	0.897	0.979
MADRID	0.9875	0.9180	**1.205** [a]	**1.289** [a]	**1.165** [a]
MURCIA	0.9761	0.9214	0.942	1.125	0.899
NAVARRA	0.9710	0.9280	0.932	1.025	0.827 [b]
PAIS VASCO	0.9935	0.9580	**1.201** [a]	**1.408** [a]	**1.137** [a]
RIOJA	0.9706	0.8890	0.984	1.060	0.956

a Rejections of the null hypothesis d = 1 in favour of d > 1. (Also in bold)
b Rejections of the null hypothesis d = 1 in favour of d < 1.

FIGURE 2

Estimates of d based on the Whittle semiparametric approach of Robinson (1995)

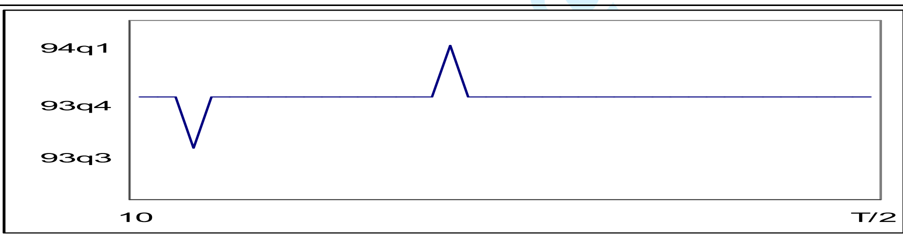


The horizontal axe refers to the bandwidth number while the vertical one corresponds to the estimates of d.

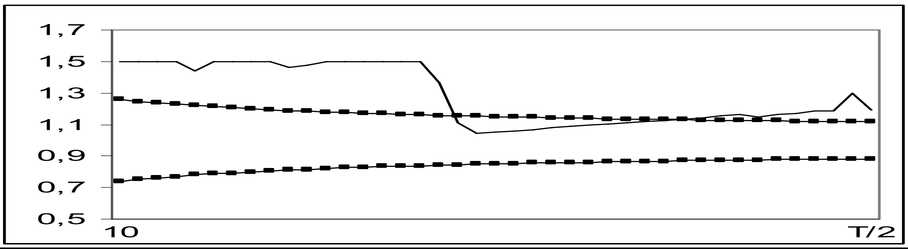
FIGURE 3

Estimates of the time break and the fractional differencing parameter

a) Estimates of the time break



b) Estimates of the fractional differencing parameter

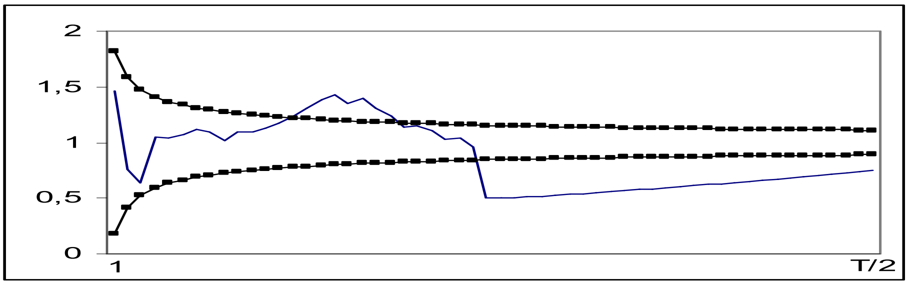


The horizontal axe refers to the bandwidth number m while the vertical is the estimated time break in the upper plot and the estimated values of d in the lower plot. In the latter, we also include the 95% confidence interval of the I(1) hypothesis.

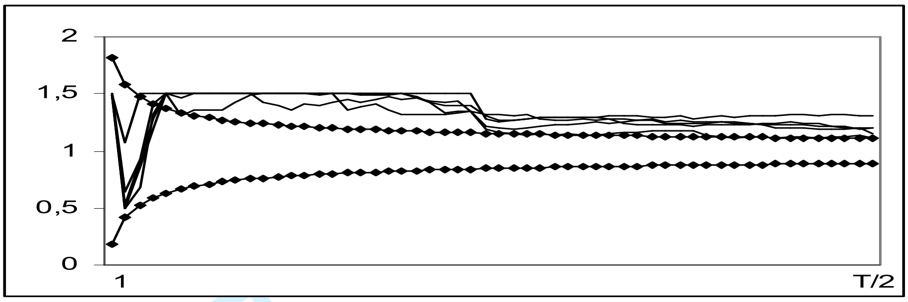
FIGURE 4

Estimates of d based on the semiparametric approach of Robinson (1995)

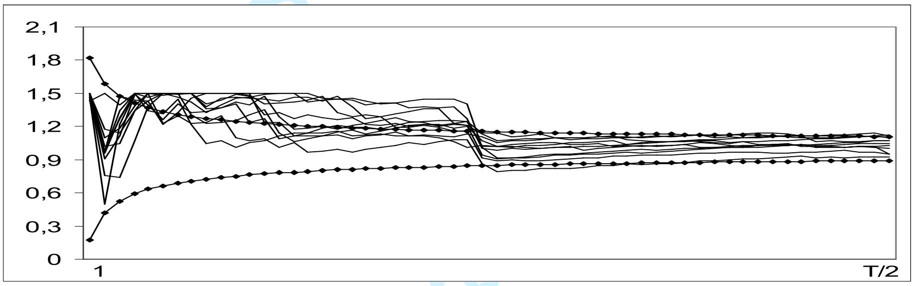
a) Baleares



b) Cataluña; País Vasco; Madrid; Valencia and Aragón



c) Remaining regions: Andalucía; Asturias; Canarias; Cantabria; Castilla-León; Castilla-Mancha; Extremadura; Galicia; Murcia; Navarra and Rioja.



The horizontal axe refers to the bandwidth number while the vertical one corresponds to the estimates of d.

TABLE 3

Unemployment persistence: Aggregated and disaggregated by regions with data ending at 1993q4

	d = 0		Estimates of d		
	AR (1)	Seasonal AR(1)	White noise	Bloomfield (1)	Seasonal AR(1)
TOTAL	0.9976	0.9668	**1.402**	**1.342**	**1.336**
ANDALUCIA	0.9945	0.9629	1.02	**1.25**	0.97
ARAGON	0.9880	0.9204	1.06	**1.30**	1.00
ASTURIAS	0.9937	0.9631	1.00	0.90	0.97
BALEARES	0.8728	0.8888	0.55	0.29	0.56
CANARIAS	0.9925	0.9640	0.97	0.95	0.93
CANTABRIA	0.9926	0.9560	0.94	1.00	0.94
CASTILLA-LEON	0.9946	0.9747	0.97	0.73	0.98
CASTILLA-MANCHA	0.9883	0.9480	0.97	0.82	0.91
CATALUÑA	0.9939	0.9237	**1.38**	**1.51**	**1.29**
COM.VALENCIANA	0.9957	0.9588	**1.19**	1.48	**1.12**
EXTREMADURA	0.9788	0.9424	0.67	0.53	0.64
GALICIA	0.9964	0.9777	0.99	0.76	0.91
MADRID	0.9884	0.8967	**1.20**	**1.33**	1.14
MURCIA	0.9891	0.9449	0.91	0.83	**0.91**
NAVARRA	0.9631	0.9138	0.89	0.98	0.78
PAIS VASCO	0.9962	0.9691	1.12	**1.49**	1.04
RIOJA	0.9759	0.8994	0.93	**1.20**	0.87

In bold those values which are higher than in Table 3 (based on whole sample size).

FIGURE 5

Provincial Unemployment Rates for 1988, 1994, 2000 and 2004 with respect to 1976

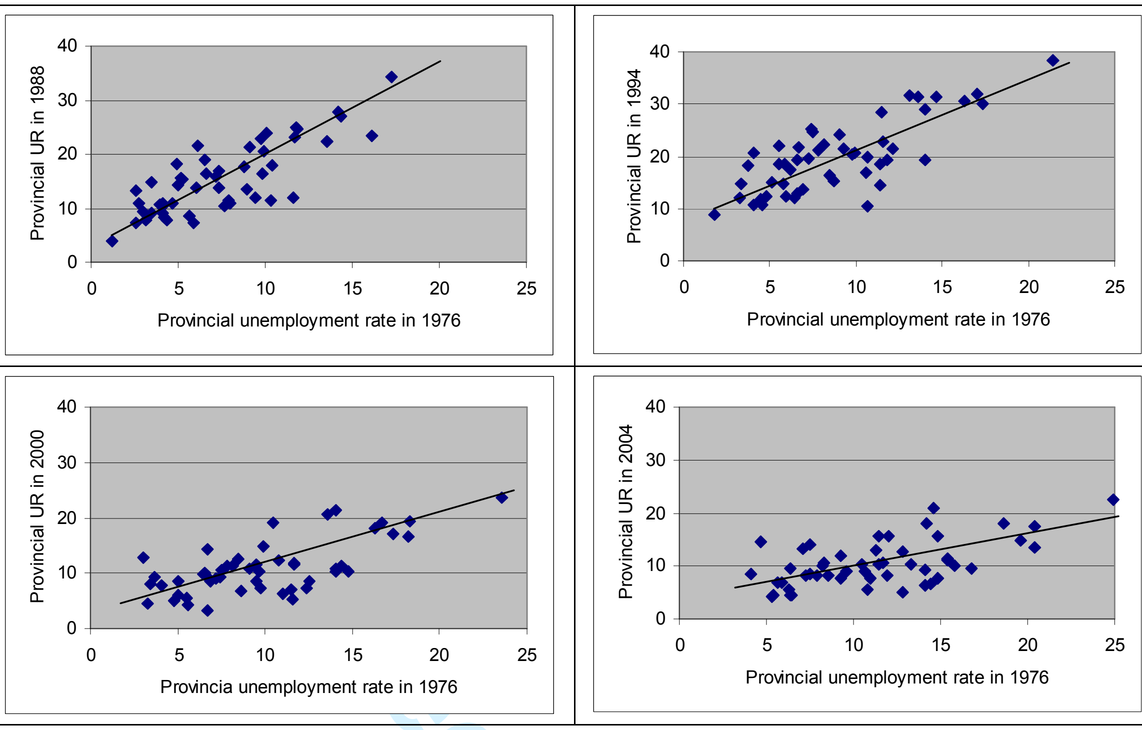


TABLE 4

Regressions of Unemployment Rates for provinces with respect to previous years

	Dependent variable: UR (annual averages for provinces)				
	1982	**1988**	**1994**	**2000**	**2004**
1976	**1.288** (8.44)	**1.914** (10.5)	**1.925** (8.93)	**1.449** (6.60)	**0.960** (5.89)
	$R^2 = 0.597$	$R^2 = 0.696$	$R^2 = 0.624$	$R^2 = 0.475$	$R^2 = 0.419$
1982		**1.176** (11.4)	**1.146** (8.77)	**0.715** (4.78)	**0.474** (4.36)
		$R^2 = 0.731$	$R^2 = 0.615$	$R^2 = 0.322$	$R^2 = 0.284$
1988			**0.978** (16.3)	**0.751** (9.91)	**0.487** (7.94)
			$R^2 = 0.847$	$R^2 = 0.671$	$R^2 = 0.567$
1994				**0.744** (11.8)	**0.516** (11.0)
				$R^2 = 0.745$	$R^2 = 0.719$
2000					**0.650** (16.5)
					$R^2 = 0.850$

Estimations computed for the annual average unemployment rates. Coefficients in bold at the 95% level. (t-*statistic*) in parenthesis.

TABLE 5

Mean regression of UR for provinces in Spain: $\Delta_t u^i = \alpha + \beta u^i_q + \varepsilon$

	(1)	(2)	(3)	(4)	(5)
After:	***q* = 1976**	***q* = 1982**	***q* = 1988**	***q* = 1994**	***q* = 2000**
1 year	0.087 (1.64)	0.035 (0.83)	−0.025 (−0.66)	−0.013 (−0.37)	**−0.317** (−6.74)
2 years	0.071 (0.93)	**0.303** (3.62)	−0.087 (−1.63)	−0.062 (−1.49)	**−0.217** (−5.61)
4 years	0.247 (1.88)	**0.254** (2.74)	−0.093 (−1.64)	**−0.143** (−2.35)	**−0.349** (−8.87)
6 years	0.288 (1.89)	0.176 (1.71)	−0.021 (−0.37)	**−0.255** (−4.06)	
8 years	**0.916** (4.50)	0.069 (0.58)	−0.067 (−1.00)	**−0.380** (−7.45)	
10 years	**0.982** (5.38)	0.011 (0.08)	−0.115 (−1.61)	**−0.483** (−10.3)	
12 years	**0.914** (5.01)	0.146 (1.12)	**−0.248** (−3.28)		
14 years	**0.760** (3.73)	0.015 (0.10)	**−0.389** (−6.00)		
16 years	**0.816** (4.16)	−0.061 (−0.41)	**−0.512** (−8.35)		
18 years	**0.925** (4.29)	−0.284 (−1.90)			
20 years	**0.793** (3.48)	**−0.380** (−3.14)			
22 years	**0.774** (3.56)	**−0.525** (−4.84)			
24 years	**0.449** (2.05)				
26 years	0.243 (1.39)				
28 years	−0.039 (−0.24)				

Estimations computed for the annual average unemployment rates. Coefficients in bold at the 95% level. (*t*-statistic) in parenthesis. *i* denotes province, u^i_q is the UR in the starting period *q*, and $\Delta_t u^i$ is the variation of the UR over the following *t* periods.

TABLE 6

Spanish unemployment persistence. Panel data estimations (1976q3-2004q4)

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	**OLS**	**OLS**	**OLS**	**GLS**	**GLS**	**GLS**
AR(1)	**0.9663** (269.2)	**0.9655** (217.5)	**0.9285** (152.4)	**0.9735** (373.8)	**0.9776** (308.4)	**0.9607** (230.7)
Break1982	–	0.0628 (0.86)	**0.4108** (4.72)	–	−0.0698 (−1.11)	0.0887 (1.31)
Break1988	–	0.0149 (0.19)	0.0205 (0.27)	–	−0.0075 (−0.14)	−0.0123 (−0.22)
Break1994	–	**−0.4628** (−5.92)	**−0.3832** (−4.88)	–	**−0.4933** (−9.04)	**−0.4575** (−8.28)
Break2000	–	**−0.1892** (−2.11)	**−0.0494** (−5.12)	–	−0.0267 (−0.42)	**−0.1675** (−2.50)
Constant	**0.5572** (11.70)	**0.7171** (15.02)	**0.6255** (5.24)	**0.4133** (10.29)	**0.6114** (13.21)	**0.5553** (4.67)
Reg.dummies	–	–	Yes	–	–	Yes
R^2	0.9456	0.9467	0.9476	–	–	–
Nº periods	113	113	113	113	113	113
Nº provinces	50	50	50	50	50	50
Nº	5650	5650	5650	5650	5650	5650

Coefficients in bold at the 95% level. (t-*statistic*) in parenthesis. [P-*value*] in brackets.

TABLE 7						
Spanish unemployment persistence. Panel data estimations (1976q3-2004q4)						
	(7)	(8)	(9)	(10)	(11)	(12)
Variables	**Fixed Effects**	**Fixed Effects**	**Random Effects. ML**	**Random Effects ML**	**GMM**	**GMM**
AR(1)	**0.9341** (221.3)	**0.8839** (135.5)	**0.9515** (182.4)	**0.9060** (121.6)	**0.9470** (67.17)	**0.8577** (20.92)
Break1982	–	**0.8294** (8.69)	–	**0.6219** (6.14)	–	0.4662 (1.33)
Break1988	–	0.0271 (0.38)	–	0.0238 (0.33)	–	**–0.5710** (–4.78)
Break1994	–	**–0.2874** (–3.94)	–	**–0.3349** (–4.57)	–	**–0.8634** (–7.97)
Break2000	–	**–0.8623** (–9.25)	–	**–0.6801** (–6.95)	–	**–1.6972** (–4.10)
Constant	**1.0380** (15.41)	**1.3577** (18.45)	**0.7771** (8.95)	**1.1842** (11.32)	**–0.0016** (–1.76)	**0.0278** (–8.08)
R^2	0.8974	0.9013	–	–	–	–
Nº time periods	113	113	113	113	112	112
Nº provinces	50	50	50	50	50	50
Nº observations	5650	5650	5650	5650	5600	5600
Individual effects	[0.000] [a]	[0.000] [a]	[0.001] [b]	[0.000] [b]	–	–
m_1					[0.003]	[0.001]
m_2					[0.783]	[0.767]

Coefficients in bold at the 95% level. (t-*statistic*) in parenthesis. [P-*value*] in brackets.
a: F-test that all $u_i = 0$ (H_0: no fixed effects). b: likelihood ratio test of $\sigma_u = 0$ (H_0: no random effects).
m_i: (i-order correlation test): Arellano-Bond test that average autocovariance in residuals of order i is 0. (H_0: no autocorrelation).

TABLE 8					
Levin-Lin-Chu (2002) Panel unit root test (1976q3:2004q4)					
		Constant		**Constant & trend**	
	(1)	(2)	(3)	(4)	(5)
	N [a]	***t*-star** [b]	***P*-value**	***t*-star** [b]	***P*-value**
SPAIN	50	**–0.288**	**[0.386]**	**0.571**	**[0.715]**
SPAIN (1976q3:1993q4)	50	**–2.242**	**[0.012]**	**0.962**	**[0.832]**
ANDALUCIA	8	–11.35	[0.000]	–15.57	[0.000]
ARAGON	3	–3.001	[0.001]	–4.770	[0.000]
ASTURIAS	1	–	–	–	–
BALEARES	1	–	–	–	–
CANARIAS	2	–	–	–	–
CANTABRIA	1	–	–	–	–
CASTILLA-LEON	9	–7.195	[0.000]	–9.976	[0.000]
CASTILLA-MANCHA	5	–3.119	[0.000]	–5.208	[0.000]
CATALUÑA	4	**–1.831**	**[0.033]**	–4.405	[0.000]
COM.VALENCIANA	3	–2.838	[0.002]	–3.269	[0.000]
EXTREMADURA	2	–	–	–	–
GALICIA	4	–4.235	[0.000]	–5.094	[0.000]
MADRID	1	–	–	–	–
MURCIA	1	–	–	–	–
NAVARRA	1	–	–	–	–
PAIS VASCO	3	**–2.302**	**[0.011]**	–4.902	[0.000]
RIOJA	1	–	–	–	–
BALEARES & CANARIAS & CAST-MANCHA	8	–3.128	[0.000]	–4.900	[0.000]
EXTREMADURA & GALICIA & ARAGON	9	–4.394	[0.000]	–8.373	[0.000]
MURCIA & NAVARRA & RIOJA & ASTURIAS & CANTABRIA & ARAGON	8	–4.494	[0.000]	–6.506	[0.000]
CATALUÑA & MADRID & VALENCIA	8	–2.874	[0.002]	–5.208	[0.000]
CATALUÑA & MADRID & PAIS VASCO	8	**–1.858**	**[0.031]**	–4.801	[0.000]

a Indicates the number of provinces included in each panel. The total number of observations is $N \cdot (114 - k)$.
b Adjusted *t*-statistic, derived by LLC (2002), that obeys asymptotically the standard normal distribution.

TABLE 9					
Im-Pesaran-Sin (2003) Panel unit root test (1976q3:2004q4)					
		Constant		**Constant & trend**	
	(1)	(2)	(3)	(4)	(5)
	N [a]	***t*-bar** [b]	***P*-value**	***t*-bar** [b]	***P*-value**
SPAIN	50	**–2.173**	**[0.015]**	**0.571**	**[0.715]**
SPAIN (1976q3:1993q4)	50	**–0.205**	**[0.419]**	**–0.590**	**[0.277]**
ANDALUCIA	8	–11.24	[0.000]	–11.57	[0.000]
ARAGON	3	–2.400	[0.001]	–2.990	[0.001]
ASTURIAS	1	–	–	–	–
BALEARES	1	–	–	–	–
CANARIAS	2	–	–	–	–
CANTABRIA	1	–	–	–	–
CASTILLA-LEON	9	–6.126	[0.000]	–6.147	[0.000]
CASTILLA-MANCHA	5	–2.697	[0.004]	–3.050	[0.001]
CATALUÑA	4	**–2.570**	**[0.005]**	–2.809	[0.002]
COM.VALENCIANA	3	–2.693	[0.004]	**–1.395**	**[0.081]**
EXTREMADURA	2	–	–	–	–
GALICIA	4	–3.964	[0.000]	–2.978	[0.001]
MADRID	1	–	–	–	–
MURCIA	1	–	–	–	–
NAVARRA	1	–	–	–	–
PAIS VASCO	3	**–2.302**	**[0.011]**	–4.902	[0.000]
RIOJA	1	–	–	–	–
BALEARES & CANARIAS & CAST-MANCHA	8	–3.056	[0.001]	**–2.247**	**[0.012]**
EXTREMADURA & GALICIA & ARAGON	9	–3.959	[0.000]	–5.922	[0.000]
MURCIA & NAVARRA & RIOJA & ASTURIAS & CANTABRIA & ARAGON	8	–4.366	[0.000]	–3.956	[0.000]
CATALUÑA & MADRID & VALENCIA	8	–4.380	[0.000]	–3.532	[0.000]
CATALUÑA & MADRID & PAIS VASCO	8	**–2.242**	**[0.012]**	**–1.599**	**[0.055]**

a Indicates the number of provinces included in each panel.
b The *t*- bar statistic, derived by Im-Pesaran-Sin (2003), converges in probability to a standard normal distribution.