Combining Benchmarking and Chain-Linking for Short-Term Regional Forecasting[1]

Ángel Cuevas
Research Unit
Ministry of Industry, Tourism and Trade, Spain
acuevas@mityc.es

Enrique M. Quilis
Macroeconomic Research Department
Ministry of Economy and Finance, Spain
enrique.quilis@meh.es

Antoni Espasa
Department of Statistics and Instituto Flores de Lemus
Universidad Carlos III de Madrid, Spain
espasa@est-econ.uc3m.es

November, 2011

Abstract

In this paper we propose a methodology to estimate and forecast the GDP of the different regions of a country, providing quarterly profiles for the annual official observed data. Thus the paper offers a new instrument for short-term monitoring that allow the analysts to quantify the degree of synchronicity among regional business cycles.

Technically, we combine time series models with benchmarking methods to forecast short-term quarterly indicators and to estimate quarterly regional GDPs ensuring their temporal and transversal consistency with the National Accounts data. The methodology addresses the issue of non-additivity taking into account explicitly the transversal constraints imposed by the chain-linked volume indexes used by the National Accounts and provides an efficient combination of structural as well as short-term information.

The methodology is illustrated by an application to the Spanish economy, providing real-time quarterly GDP estimates and forecasts at the regional level (i.e., with a minimum compilation delay with respect to the national quarterly GDP).

JEL CODES: C53, C43, C82, R11.

[1] The authors thank Ana Abad for her comments. Antoni Espasa acknowledges financial support from *Ministerio de Eduación y Ciencia* project ECO 2009-08100. This paper has been presented at DIW Econometric Workshop (Berlin, 2010), International Symposium on Forecasting (Prague, 2011) and IWH-CIREQ Macroeconometric Workshop (Halle, 2011). Any views expressed herein are those of the authors and not necessarily those of the Spanish Ministry of Industry, Tourism and Trade or the Spanish Ministry of Economy and Finance.

1. INTRODUCTION

Business cycle analysis and short-term monitoring can be substantially improved if an explicit regional dimension is taken into account. In this way we can analyze in detail the diffusion of the aggregate (or national) cycle: identifying leading/coincident/lagged regions; detecting common and specific shocks; etc. The relevance of this added dimension is especially important both for countries of large or medium size as well as for countries with decentralized systems that allow specific economic policies to be applied. Of course, these quarterly forecasts are also very useful for regional governments.

The Regional Accounts (RA) are usually annual and in this context we propose in this paper a methodology to estimate and forecast quarterly Gross Domestic Product (GDP) time series at the regional level, providing a new instrument for short-term monitoring that allow us to gauge the degree of synchronicity and the identification of common and idiosyncratic shocks to different regions.

One notable feature of our methodology is that these GDP estimates aggregate[2] consistently with the national GDP provided by the Quarterly National Accounts (QNA). In the same vein, our forecasts are consistent with the corresponding national forecasts. Note also that the same principles of volume estimation using chain-linked indices have been used and the same procedures of seasonal and calendar adjustment have been applied as in the QNA.

Structural consistency is also ensured since the quarterly regional GDPs are temporally consistent with their annual Regional Accounts counterparts. The fact that both QNA and RA share the same National Accounts (NA) framework[3] is an additional source of consistency. In this way, we can use the quarterly regional estimates and forecasts to perform comparative analysis and to derive structural measures at the regional level.

The modeling approach is highly reliant on a set of regional high-frequency indicators. These indicators provide the ultimate basis used by the model to generate GDP forecasts according to time series techniques which may range from univariante ARIMA models to multivariate dynamic factor models. The set of indicators and models are homogeneous across regions, ensuring the comparability of the results.

The methodology has three main stages:

1. Processing of high frequency indicators available at the regional level.
2. Temporal disaggregation and extrapolation of annual regional GDP using the indicators processed in step 1.

[2] Aggregation is performed according to the chain-linking methodology.
[3] In particular, the 1995 European System on National Accounts (ESA-95).

3. Balancing of these initial quarterly estimates in order to ensure transversal consistency with national quarterly GDP, preserving the temporal consistency achieved in the previous stage.

It is worth to emphasize that, from an operational perspective, timely forecasts of quarterly regional GDPs may be available with a minimum delay with respect to the national quarterly GDP release. In this way, the national figure may have timely regional counterparties, enhancing the informational content of analysis carried out at the aggregate level.

The main contributions of our paper are:

- A set of quarterly estimates of GDP for all the regions in a country, derived in a consistent way with the official available data provided by the National Accounts, both RA and QNA.
- Early (or flash) estimates (forecasts) of quarterly GDP at the regional level that may be released at the same time as the national GDP.
- Short-term forecasts of quarterly GDP at the regional level by conditioning them on the projected path of the underlying short-term quarterly regional indicators.

The paper is organized as follows. The second section outlines the modeling approach, going into detail of its main steps. A case study using Spanish data appears in section three. Finally, in section four, we present our conclusions and future lines of research.

2. MODELING APPROACH

In this section we present the main steps of the proposed methodology. The modeling approach consists of three basic steps: (i) seasonal adjustment and forecasting of regional short-term raw indicators, (ii) initial quarterly estimates of regional GDP provided by benchmarking and (iii) enforcement of the transversal constraint that links the regional quarterly GDPs with their national counterpart.

This aggregation constraint must be consistent with the chain-linking procedure used to compile quarterly GDP at the national level, dealing with the non-additivity issue in an appropriate way. We now turn to examine the three stages in more detail but, to simplify the exposition, we first present the required information set.

2.1. Information set

The model requires as input three elements that vary according to their sampling frequency (annual or quarterly), their spatial coverage (regional or national) and their method of compilation (National Accounts or short-term indicators).

The variables of the system are: regional GDPs (y), national GDP (z) and regional short-term indicators in original or raw[4] form (xo). Upper-case letters refer to annual variables while lower-case letters refer to quarterly variables. Let $T=1..N$ be the annual (low-frequency) index, $t=1..4$ the quarterly index within a natural year and $j=1..M$ the regional (cross-section) index.

Hence, $Y=\{Y_{T,j}: T=1..N; j=1..M\}$ is a NxM matrix comprising the annual regional GDPs that play the role of temporal benchmarks of the system. Aggregation of the regional GDPs generates the GDP at the national level[5].

Variable z is a $nx1$ vector comprising the observed quarterly GDP provided by the QNA, being $n \geq 4N$. This figure is available more timely than the regional data and shares with them the corresponding annual GDP volume index[6]:

$$[2.1] \qquad Z_T = \frac{1}{4}\sum_{t \in T} z_{t,T}$$

Finally, xr is a nxM matrix comprising the observed raw quarterly indicators that operate as high-frequency proxies for the regional aggregates Y. The indicators should be chosen in order to show a high degree of low-frequency conformity with the regional GDP time series. For the sake of simplicity, we assume that there is only one available indicator although the system easily accommodates more than one indicator. As will be explained later, we work with the seasonally and calendar adjusted indicators (x) instead of the raw indicators (xo).

Only the indicators are observed at the three dimensions of the system: T (annual index), t (quarterly index) and j (regional index). Therefore, they provide the interpolation basis for Y (across the quarterly dimension t) and z (across the regional dimension j). In other words, our objective is to estimate and forecast y using x as interpolators and consistently with both Y and z.

The following table resumes the information set of the system:

Table 1: Information set

Variable	Frequency	Breakdown	Type of data	Source
Y	Annual	Regional	Raw	Regional Accounts: GDP
z	Quarterly	National	Seasonally adjusted	Quarterly National Accounts: GDP
xo	Quarterly	Regional	Raw	Short-term indicators
x	Quarterly	Regional	Seasonally adjusted	System estimates
y	Quarterly	Regional	Seasonally adjusted	System estimates

[4] That is, incorporating seasonal and calendar effects.
[5] Again, aggregation is performed according to the chain-linking methodology.
[6] For example, taking 2011 as a reference, the QNA released its first estimate of 2010:Q4 on February 11 while the RA release its first estimate of 2010 on March 24. Both estimates share the annual figure for 2010 implicitly provided by the QNA by means of temporal aggregation of the four quarters of 2010.

2.2. Processing short-term indicators

Typically, short-term regional economic indicators are compiled in raw form by the statistical agencies. However, the volume GDP used for short-term monitoring at the national level is calculated in two ways: using raw indicators or using seasonal and calendar adjusted indicators. Since seasonal and calendar effects could be quite different between indicators and the macromagnitudes, the second procedure for the calculation of the GDP seems more reliable. Usually these GDP figures are referred as seasonal and calendar adjusted.

In order to ensure the homogeneity between both sources of information, regional raw indicators and seasonally adjusted quarterly national GDP, we apply an ARIMA model-based correction that filter out the raw data from seasonal and calendar effects, if they are present. The procedure has been implemented using the TRAMO-SEATS program, see Gómez and Maravall (1995) and Caporello and Maravall (2004). Formally:

[2.2] $$x_{j,t,T} = V(B,F;\psi_j)\, xo_{j,t,T}$$

where $xo_{j,t,T}$ is the raw short-term indicator[7]; $V()$ is the Wiener-Kolmogorov filter symmetrically defined on the backward and forward operators B and F and ψ_j are the parameters of the filter derived consistently with those of the ARIMA model for $xo_{j,t,T}$.

At the same time, we can forecast the short-term indicators used as proxies for regional GDP using the same ARIMA models that we have used to perform seasonal adjustment.

2.3. Initial quarterly regional GDP estimation

Preliminary estimates of quarterly GDP at the regional level are compiled using benchmarking techniques, see Di Fonzo (1987, 2002) and Proietti (2006) for an in-depth exposition. These techniques play an important role in the compilation practices of Quarterly National Accounts around the world, see Eurostat (1998) and Bloem et al. (2001).

In order to perform the estimation of quarterly regional GDPs we start assuming that there is a linear model that links the (observable) high-frequency indicators with the (unobservable) QRA variable. The model assumes a static relationship between y and x, additively perturbed:

[2.3] $$y_j = x_j \beta_j + u_j \quad j = 1..m$$

The innovation u follows a stationary AR(1) process (therefore, assuming cointegration between y and x) or a non-stationary I(1), random walk process (therefore, excluding cointegration between y and x):

[7] If calendar effects (e.g., Easter effect or trading day effect) are present, a preliminary correction is also performed. Without loss of generality, we will continue to call the possibly calendar-corrected data as raw data.

$$[2.4] \qquad \begin{aligned} & u_j \sim N(0, v_j) \\ & v_j = \sigma_j^2 \left[(I_n + \rho_j \Xi)'(I_n + \rho_j \Xi) \right]^{-1} \quad -1 < \rho_j \leq 1 \end{aligned}$$

The auxiliary matrix Ξ used in the previous equation is defined as:

$$\Xi = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ -1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -1 & 0 \end{bmatrix}$$

Initial conditions depend on the value of ρ_j:

$$[2.5] \qquad u_{j,0} = \begin{cases} N(0, \sigma_j^2 / (1 - \rho_j^2)) & \text{if } -1 < \rho_j < 1 \\ 0 & \text{if } \rho_j = 1 \end{cases}$$

The model and the estimation procedure may be easily extended to include a richer dynamic structure for the innovations. This is the case for the method proposed by Litterman (1983). This author assumes an ARI(1,1) model for the innovations instead of the AR(1) or I(1) model used above. However, the empirical and Monte Carlo evidence show that its performance is sometimes disappointing. This fact is due to the flatness of the implied likelihood profile and, therefore, the corresponding observational equivalence in a wide range of values for ρ_j, see Proietti (2006).

The model includes a temporal constraint that makes y_j quantitatively consistent with its annual counterpart Y_j:

$$[2.6] \qquad Y_j = C y_j$$

with a temporal aggregation-extrapolation matrix defined as:

$$[2.7] \qquad C = (I_N \otimes c \mid O_{N,n-sN})$$

where N is the number of low-frequency observations, $\otimes$ stands for the Kronecker product, c is a row vector of size s which defines the type of temporal aggregation and s is the number of high frequency data points for each low frequency data point. If $c=[1,1,...,1]$ we would be in the case of the temporal aggregation of a flow, if $c= [1/s,1/s,...,1/s]$ in the case of the average of an index and, if $c=[0,0,...,1]$, an interpolation would be obtained. In out case, $s=4$.

Extrapolation arises when $n>sN$. In this case, the problem can easily be solved by simply extending the temporal aggregation matrix by considering new columns of zeroes which do not distort the temporal aggregation relationship and that do not pose any difficulty to the inclusion of the last $n-sN$ data points of the indicators in the estimation process of y.

The estimation of y_j according to the model [2.3]-[2.5] and satisfying the constraint [2.6]-[2.7] is performed by means of the Chow-Lin (1971) or Fernández (1981) procedures. The proposed Best[8] Linear Unbiased Estimator (BLUE) estimator adopts the form:

[8] In the sense of minimizing the mean squared error.

[2.8] $$\hat{y}_j = x_j\hat{\beta}_j + v_j C' V_j^{-1}\hat{U}_j = x_j\hat{\beta}_j + L_j\hat{U}_j$$

with $V_j=C'v_jC$. The annual disturbance term is defined as:

[2.9] $$\hat{U}_j = Y_j - X_j\hat{\beta}_j$$

with $X_j=Cx_j$.

Equation [2.8] expresses the estimator as the combination of a term linearly linked to the indicator and a temporally disaggregated residual series. The main feature of the estimator is the dependency of the temporal disaggregation filter L_j on the form adopted by the quarterly model, and in particular, on the dynamic structure of its disturbance term u_j. The Generalized Least Squares (GLS) estimator of β_j is:

[2.10] $$\hat{\beta}_j = (X_j' V_j^{-1'} X_j)^{-1}(X_j' V_j^{-1'} Y_j)$$

An important advantage of this method is that it generates confidence intervals for the quarterly estimates from the corresponding variance-covariance matrix:

[2.11] $$\Sigma_{\hat{y}_j} = (I_n - L_jC)v_j + (x_j - L_jX_j)\Sigma_{\hat{\beta}_j}(x_j - L_jX_j)'$$

The last equation implies that the uncertainty associated with the quarterly estimates is tied to two sources: one related to the variability of the quarterly stochastic disturbance term u and the other linked to the imprecision in which we incur when estimating β_j.

Expressions [2.9] to [2.12], which fully define the Chow-Lin method, require for its implementation prior knowledge of the variance-covariance matrix v_j of the quarterly disturbance term u_j which depends on ρ_j, see [2.5]. The estimation of this parameter is accomplished by means of the evaluation of the implied log-likelihood function of the low-frequency model. The function is (dropping subindex j):

[2.12] $$\ell(\beta,\sigma_a^2 \mid \bar{\rho}) = -\frac{N}{2}ln(2\pi\sigma^2) - \frac{1}{2}ln(|Cv(\bar{\rho})C'|) - \\ -\frac{1}{2\sigma^2}(Y - X\beta)'(Cv(\bar{\rho})C')^{-1}(Y - X\beta)$$

This optimization is performed by means of a grid search on the stationary domain of ρ, and pinning down the values of β_j and σ_j that maximizes [2.13] conditioned on the selected value for ρ_j, see Bournay and Laroque (1979).

The estimation is simpler in the case $\rho=1$ (Fernández) than in the case $-1<\rho<1$ (Chow-Lin), becoming:

[2.13] $$\hat{\beta} = \left[x'C'\left[C(D'D)^{-1}C'\right]^{-1}Cx\right]^{-1}x'C'\left[C(D'D)^{-1}C'\right]Y \\ \hat{y} = x\hat{\beta} + (D'D)^{-1}C'\left[C(D'D)^{-1}C'\right]^{-1}(Y - Cx\hat{\beta}) \\ D = I_n + \Xi$$

Several authors have proposed the use of dynamic models to perform temporal disaggregation, see Gregoir (1994), Salazar et al. (1994), Santos and Cardoso (2001), Di Fonzo (2003), Proietti (2006), among others. These models generalize and sometimes encompass the standard Chow-Lin model (including its extensions due to Fernández and Litterman).

In practice, the specification of these models must be restricted in order to take into account the lack of information due to the temporal aggregation loss[9] and to the need of not introducing "noise" in the estimates (e.g., in the AR(p) case, oscillations due to parameters generating complex roots). These constraints are specially important in the case of short samples and temporal aggregation from quarterly to annual frequency, as is the case of our application, and explain why we have confined our models to be of the Chow-Lin or Fernández type.

2.4. Balancing in a chain-linking setting

The estimates derived in the previous step do not verify the transversal constraint that should relate them to the national quarterly GDP, satisfying the same type of relationship that links annual regional GDPs and annual national GDP. We solve the problem applying a multivariate balancing procedure, in particular a multivariate extension of the Denton (1971) method. This extension can be expressed in matrix form, as in Di Fonzo (1994) and Di Fonzo and Marini (2003), as well as in state space form, see Proietti (2011a). In this paper we have adopted the former approach, using the functions written in MATLAB by Abad and Quilis (2005).

This balancing method depends on the formulation of additive constraints. However, volume indexes compiled according to the chain-linking methodology are non-additive, see Bloem et al. (2001) and Abad et al. (2007). Fortunately, we can transform the chain-linked measures in order to write them in an additive form and then use the powerful machinery of balancing procedures to ensure transversal and temporal consistency. Finally, we can express the results in the initial chain-linked format by reversing the transformation.

The constraint that links regional and national quarterly volume GDP is:

[2.14] $$z_{t,T} = \left(\sum_j W_{j,T-1} \frac{y_{j,t,T}}{Y_{j,T-1}} \right) Z_{T-1}$$

Where $z_{t,T}$ is the national quarterly volume GDP, $W_{j,T-1}$ is the weight of region j in year $T-1$ and $y_{j,t,T}$ is the quarterly volume GDP of the j-th region[10]. Finally, Z_T and $Y_{j,T}$ are the annual counterparts $z_{t,T}$ of and $y_{j,t,T}$.

After some algebraic manipulations, we can express the constraint in additive form:

[9] Wei and Stram (1986) provide and in-depth analysis of the indeterminacies that arise when performing temporal aggregation from the standpoint of ARIMA modeling.

[10] Weights are computed using GDPs valued at current prices, see Abad et al. (2007) for a complete derivation.

[2.15] $$\underbrace{\frac{z_{t,T}}{Z_{T-1}}}_{r_{t,T}} = \sum_j \underbrace{W_{j,T-1}\frac{y_{j,t,T}}{Y_{j,T-1}}}_{wr_{j,t,T}} = \sum_j wr_{j,t,T}$$

In [2.15] the relationship between the national ratio $r_{t,T}$ and the weighted regional ratios $wr_{j,t,T}$ is additive.

Plugging the initial estimates derived according to [2.8] or [2.13] into [2.15] we obtain the preliminary, unbalanced estimates:

[2.16] $$wr^{*}_{j,t,T} = \sum_j W_{j,T-1}\frac{\hat{y}_{j,t,T}}{Y_{j,T-1}}$$

The balanced and temporally consistent time series $wr^{**}_{j,t,T}$ are the output from the following constrained quadratic optimization program:

[2.16] $$\underset{wr^{*}}{MIN} \quad (wr^{**} - wr^{*})' D' D (wr^{**} - wr^{*}) \qquad s.t. \quad H wr^{**} = R_e$$

being:

$$H = \begin{bmatrix} 1_M' \otimes I_n \\ I_M \otimes C \end{bmatrix} \quad \text{and} \quad R_e = \begin{bmatrix} z \\ WR \end{bmatrix}$$

Where 1_M is a column vector of ones and WR is the annual counterpart of the weighted regional ratios written in matrix form.

In the program [2.16] the objective function reflects the volatility of the discrepancies between the quarter-to-quarter growth rates of the balanced series and those of the unbalanced ones. After some mathematical manipulation, an explicit expression can be derived:

[2.17] $$wr^{**} = wr^{*} + (D'D)^{-1}H'\left[H(D'D)^{-1}H'\right]^{-1}(R_e - Hwr^{*})$$

The interpretation of equation [2.17] is straightforward: the quarterly balanced series are the result of adding up a correction factor to the unbalanced series. This correction factor originates from the distribution of the discrepancy between the preliminary unbalanced estimates and the constraint series R_e.

Once we have obtained the consistent weighted ratios, we can reverse the transformation [2.15] to derive the final estimates of the quarterly regional GDP in volume terms:

[2.18] $$y^{**}_{j,t,T} = wr^{**}_{j,t,T}\frac{Y_{j,T-1}}{W_{j,T-1}}$$

In this way, the estimates of quarterly GDP derived in the previous equation are quantitatively consistent in their time dimension (taking as benchmark their annual regional counterparts) and in their cross-section dimension (generating the GDP provided by the QNA by regional aggregation). We should also emphasize that the consistency extends to the methodological dimension too, since the chain-linking procedures in current use by the National Accounts have been properly taken into account. Finally, using time series methods to project the basic short-term indicators we can derive nowcasts (or flash estimates) and forecasts of regional quarterly GDP in a timely way.

2.4. Comparison with other approaches

The following table compares our methodology with related approaches along six dimensions: econometric modeling, role of constraints (temporal and transversal), explicit consideration of chain-linking, mixing data frequencies (e.g., annual and quarterly data) and main objectives.

Table 2: Comparison with other methodological approaches

	MIDAS	**Factor models**	**Di Fonzo (1990)**	**Di Fonzo & Marini (2005)**	**Proietti (2011)**	**Ours**
Econometric modeling	Dynamic regression	Dynamic factor models	Static regressions + I(1) innovations	Unspecified	Dynamic factor models	Static regressions + AR(1)/I(1) innovations
Temporal constraints	No	No	Yes	Yes	Yes	Yes
Transversal constraints	No	No	Yes	Yes	Yes	Yes
Chain-linking constraints	No	No	No	No	Yes	Yes
Mixing frequencies	Yes	Optional	Yes	Yes	Yes	Yes
Main objectives	Forecasting	Nowcasting	Benchmarking	Benchmarking and balancing	Nowcasting and benchmarking	Nowcasting and benchmarking

In our paper the transversal and temporal constraints play a critical role that is absent in other approaches that also combine data observed at different frequencies, such is the case of the MIDAS[11] methodology, see Guérin and Marcellino (2010) and the references cited therein. Our approach and MIDAS share the use of high-frequency indicators as basic inputs but our methodology aims at forecasting an unobservable variable (quarterly regional GDP) satisfying temporal and transversal constraints rather than forecasting an observable variable. In our case, the MIDAS approach would combine annual GDP (provided by the RA) and a quarterly indicator and the result would be forecasts of the former, missing the link with the national quarterly GDP and the possibility to obtain estimates of quarterly regional GDPs.

Our approach and dynamic factor models share the use of many indicators but our approach links them via temporal, transversal and chain-linking constraints rather than using common unobservable factors as in the case of dynamic factor models. In addition, mixing frequencies is a hallmark of our approach but it is only optional in the case of dynamic factor models.

Di Fonzo (1990) presents a methodology closely related to ours. We have expanded his approach to cope with the issue of chain-linking and focus the results to nowcasting and benchmarking. Di Fonzo and Marini (2005) may be considered as a variant of Di Fonzo (1990) in which balancing plays also a critical role.

[11] MIDAS stands for "MIxed DAta Sampling".

Finally, Proietti (2011) is our closest reference. He uses dynamic factor models, mixing quarterly and monthly data and the computational approach relies on Kalman filtering. Our approach mixes annual and quarterly frequencies and the computational approach is matrix-oriented rather than recursive.

3. CASE STUDY: A SYSTEM OF FLASH REGIONAL QUARTERLY GDP ESTIMATES FOR SPAIN

In this section we present the main results of a system of regional quarterly GDP nowcasts and forecasts for the Spanish economy, following the modeling approach previously outlined.

Regional annual GDPs in chained volume indices are provided by the Regional Accounts (RA) according to ESA-95 conventions and are available in the time span 1995-2010. The cross-section dimension includes 17 regions (*Comunidades Autónomas*) plus two autonomous cities that will be jointly considered, rendering M=18, a NUTS-2 regional breakdown according to Eurostat's classification.

The set of regional high-frequency indicators is the total employment, as provided by the register of Social Security Contributors (SSC). It is a timely and reliable indicator of the overall performance of the labor market at the regional level and, due to its close linkage to real output, of the overall economic conditions. More sophisticated models may consider a wider set of short-term indicators but our main objective now is to check the feasibility of our modeling approach, preserving this extension for future research.

Finally, the quarterly transversal constraint is the Spanish quarterly volume GDP provided by the QNA. This variable is seasonally and calendar adjusted.

The following table shows the timetable for publication of the aforementioned statistics:

Table 2: Release calendar

Release date in year T:	RA	QNA	SSC
4th january			M12 year T-1
2th february			M1 year T
11th february		Q4 year T-1	
2th march			M2 year T
30th march	year T-1		
4th april			M3 year T
4th may			M4 year T
12th may		Q1 year T	
2th june			M5 year T
2th july			M6 year T
2th august			
13th august		Q2 year T	
2th september			M8 year T
4th october			M9 year T
3th november			M10 year T
11th november		Q3 year T	
2th december			M11 year T
30th december	year T-1 (revision)		

In relation to this table it is worth mentioning three aspects:

- The release date of QNA data refers to the flash estimate. We have used the forecasts provided by *Instituto Flores de Lemus* for the interval 2011:Q3 – 2011:Q4.

- In August the Annual National Accounts publishes a revision of the series of annual data. The QNA data is immediately adapted to them from the second quarter, but the RA does not adapt this data until December 30th, so there is a period between August and December where QNA and RA are not consistent, affecting to the Q2 and Q3 releases. To avoid this peculiarity we realize the example making a prior adjustment of annual regional data to national data, which for simplicity is a simple proportional adjustment to the previously published data.

- The SSC data, in order to be consistent with the QNA data, have to be seasonally and calendar adjusted and added to a quarterly frequency. They have been forecasted by means of their ARIMA representations.

The following tables show the estimates of quarterly GDP by regions, consistent with the GDP growth figure up to the second quarter of 2011 (and by extension with the entire series of GDP). In the Annex 1 you can find graphs and maps of the complete results.

Table 3: Quarterly GDP (annual growth rates)

GDP growth by regions Annual growth rates												
	2008	2009	2010	2011	2010				2011			
					T I	T II	T III	T IV	T I	T II	T III	T IV
Spain	**0,9**	**-3,7**	**-0,1**	**0,7**	**-1,4**	**0,0**	**0,2**	**0,6**	**0,9**	**0,7**	**0,6**	**0,7**
Andalucía (AND)	0,6	-3,6	-0,9	0,5	-1,9	-0,9	-0,9	-0,2	0,3	0,4	0,8	0,5
Aragón (ARA)	0,9	-4,5	-0,5	0,1	-1,3	-0,2	0,1	-0,4	-0,1	0,1	0,0	0,2
Asturias (AST)	1,1	-4,0	0,3	0,9	-1,0	1,0	0,7	0,5	1,0	1,1	0,4	0,9
Baleares (BAL)	1,3	-3,8	-0,3	1,4	-2,0	-0,3	0,4	0,8	0,9	1,6	1,7	1,3
Canarias (CAN)	0,3	-4,2	-0,8	2,1	-2,2	-0,7	-0,7	0,3	1,5	2,2	2,4	2,2
Cantabria (CANT)	1,1	-3,5	0,1	0,5	-1,2	0,2	0,7	1,0	1,1	0,6	0,2	0,2
Castilla La Mancha (CLM)	1,5	-3,2	-0,9	-0,6	-1,5	-0,6	-1,0	-0,6	-0,3	-0,4	-0,7	-1,0
Castilla León (CYL)	0,9	-3,3	0,8	1,2	-0,6	0,9	1,2	1,6	1,5	1,3	0,9	1,0
Cataluña (CAT)	0,2	-4,2	0,1	0,6	-1,9	0,0	1,0	1,4	1,2	0,8	0,2	0,2
Extremadura (EXT)	1,8	-2,2	0,0	1,5	-0,5	0,5	0,0	-0,1	1,2	1,1	1,6	2,1
Galicia (GAL)	1,7	-3,1	0,1	-0,2	-1,2	0,6	0,7	0,2	0,3	-0,5	-0,9	0,2
Madrid (MAD)	1,0	-3,3	0,0	0,8	-0,3	0,2	-0,1	0,3	0,8	0,5	0,9	1,0
Murcia (MUR)	1,6	-3,4	-0,6	0,8	-2,0	-0,6	-0,1	0,5	1,0	0,8	0,5	1,1
Navarra (NAV)	1,9	-2,5	1,2	1,4	0,0	1,6	1,5	1,8	1,8	1,4	1,3	0,9
País Vasco (PV)	1,4	-3,7	0,8	1,4	-1,5	1,1	1,7	1,9	2,2	1,3	1,0	1,1
La Rioja (RIO)	1,5	-3,4	-0,3	0,7	-1,0	-0,6	-0,5	0,7	0,5	1,0	1,2	0,2
Valencia (VAL)	0,8	-4,4	-0,6	0,9	-2,1	-0,5	-0,2	0,3	0,7	0,7	0,9	1,2
Ceuta y Melilla (CyM)	2,8	-1,4	-0,1	2,5	-2,1	-1,7	1,6	1,9	1,2	3,8	2,5	2,4

Higher growth that Spain

Table 4: Quarterly GDP (quarterly growth rates)

GDP growth by regions Quarterly growth rates												
	2009				2010				2011			
	T I	T II	T III	T IV	T I	T II	T III	T IV	T I	T II	T III	T IV
Spain	**-1,6**	**-1,1**	**-0,3**	**-0,2**	**0,1**	**0,3**	**0,0**	**0,2**	**0,4**	**0,2**	**-0,1**	**0,3**
Andalucía (AND)	-1,3	-1,1	-0,4	-0,2	-0,2	-0,1	-0,4	0,5	0,3	0,0	0,0	0,3
Aragón (ARA)	-1,6	-1,2	-0,3	0,4	-0,1	-0,2	0,0	-0,1	0,3	0,0	-0,2	0,0
Asturias (AST)	-2,2	-1,4	0,8	0,0	-0,4	0,5	0,5	-0,1	0,1	0,6	-0,2	0,3
Baleares (BAL)	-1,2	-1,5	-0,6	-0,3	0,5	0,1	0,2	0,1	0,6	0,8	0,2	-0,3
Canarias (CAN)	-1,9	-1,4	-0,3	-0,4	0,0	0,0	-0,2	0,6	1,1	0,8	-0,1	0,4
Cantabria (CANT)	-1,4	-1,0	-0,4	-0,2	0,3	0,4	0,2	0,1	0,4	-0,1	-0,3	0,1
Castilla La Mancha (CLM)	-1,3	-1,2	-0,1	-0,2	-0,1	-0,2	-0,5	0,2	0,2	-0,3	-0,8	-0,2
Castilla León (CYL)	-1,5	-1,1	-0,2	0,0	0,7	0,5	0,1	0,3	0,6	0,2	-0,2	0,4
Cataluña (CAT)	-1,7	-1,4	-0,8	-0,2	0,5	0,5	0,2	0,3	0,3	0,1	-0,5	0,3
Extremadura (EXT)	-1,2	-0,4	0,3	-0,1	-0,3	0,5	-0,2	-0,1	1,0	0,5	0,3	0,4
Galicia (GAL)	-1,4	-1,2	-0,2	-0,2	0,4	0,6	-0,1	-0,7	0,5	-0,2	-0,5	0,4
Madrid (MAD)	-1,7	0,1	0,3	-0,1	-0,6	0,6	0,0	0,2	-0,1	0,3	0,4	0,4
Murcia (MUR)	-1,4	-1,1	-0,4	-0,5	0,0	0,3	0,1	0,1	0,5	0,1	-0,1	0,6
Navarra (NAV)	-1,3	-1,1	0,2	0,3	0,6	0,5	0,1	0,6	0,7	0,1	-0,1	0,2
País Vasco (PV)	-1,8	-1,6	-0,4	0,0	0,5	0,9	0,2	0,3	0,8	0,0	-0,2	0,4
La Rioja (RIO)	-1,8	-0,9	0,2	-0,6	0,4	-0,5	0,3	0,6	0,2	0,0	0,4	-0,3
Valencia (VAL)	-1,9	-1,5	-0,4	-0,4	0,2	0,1	-0,2	0,2	0,6	0,1	0,0	0,4
Ceuta y Melilla (CyM)	1,3	-3,5	-2,4	0,8	3,1	-3,1	0,8	1,2	2,4	-0,7	-0,4	1,0

Higher growth that Spain

These forecasts, as well as being consistent with the corresponding quarterly aggregate, are determined by the behavior of indicator considered. This can be proved by calculating the correlation between employment (SSC) and quarterly GDP estimated for each region:

Table 5: Indicator and GDP estimate: correlations

	Levels	Annual growth rate	Quarterly growth rate
Andalucía	0,991	0,919	0,833
Aragón	0,992	0,917	0,849
Asturias	0,983	0,811	0,642
Baleares	0,978	0,877	0,823
Canarias	0,954	0,928	0,837
Cantabria	0,993	0,934	0,833
Castilla La Mancha	0,993	0,860	0,768
Castilla León	0,985	0,875	0,740
Cataluña	0,980	0,949	0,858
Extremadura	0,987	0,858	0,777
Galicia	0,975	0,899	0,778
Madrid	0,991	0,881	0,704
Murcia	0,979	0,908	0,824
Navarra	0,982	0,913	0,820
País Vasco	0,988	0,908	0,808
La Rioja	0,982	0,891	0,822
Valencia	0,952	0,939	0,849
Ceuta y Melilla	0,996	0,771	0,919

It can be observed the high correlation between the indicator and the balanced estimates for all regions. These results are even good for smaller regions, such as Ceuta and Melilla, which are usually conditioned for its small size and special economic structure (see Annex 2).

In order to analyze both the duration and the date of entry and exit of the recession in each region, the following table presents the evolution of the estimated annual rates in the quarterly frequency:

Table 6: Dating recession in quarterly GDP (annual growth rates)

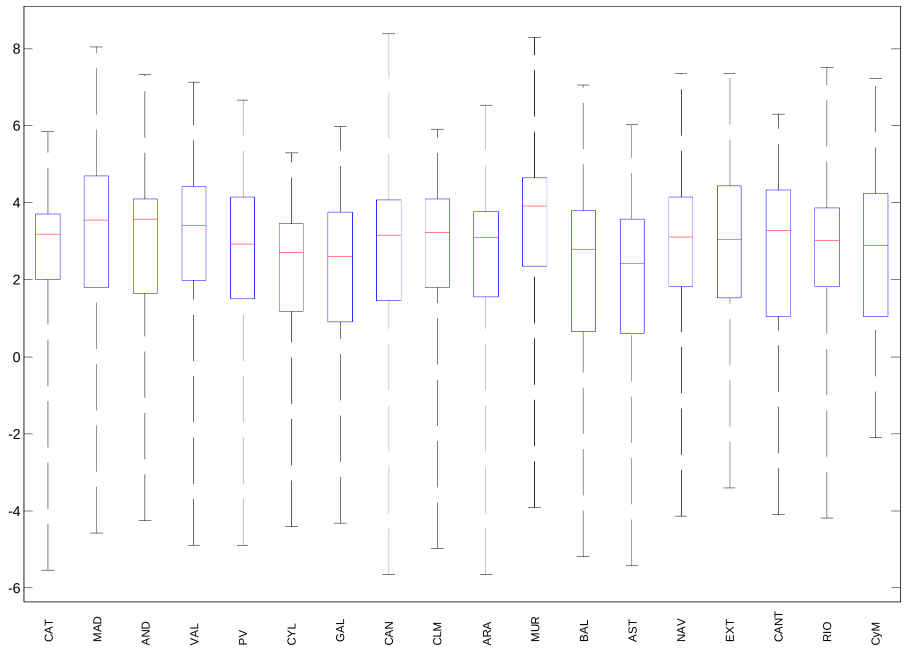
	2008				2009				2010				2011			
	T I	T II	T III	T IV	T I	T II	T III	T IV	T I	T II	T III	T IV	T I	T II	T III	T IV
Spain	2,7	1,9	0,3	-1,4	-3,5	**-4,4**	-3,9	-3,0	-1,4	0,0	0,2	0,6	0,9	0,7	0,6	0,7
Andalucía	2,7	1,9	-0,3	-1,8	-3,4	**-4,4**	-3,7	-3,0	-1,9	-0,9	-0,9	-0,2	0,3	0,4	0,8	0,5
Aragón	2,8	2,2	0,8	-2,3	-4,2	**-5,8**	-5,3	-2,8	-1,3	-0,2	0,1	-0,4	-0,1	0,1	0,0	0,2
Asturias	3,3	1,7	0,6	-1,1	-4,5	**-5,1**	-3,8	-2,8	-1,0	1,0	0,7	0,5	1,0	1,1	0,4	0,9
Baleares	3,1	2,5	0,4	-0,9	-2,9	**-4,5**	-4,4	-3,5	-2,0	-0,3	0,4	0,8	0,9	1,6	1,7	1,3
Canarias	1,9	1,1	-0,3	-1,3	-3,5	**-5,0**	-4,4	-3,9	-2,2	-0,7	-0,7	0,3	1,5	2,2	2,4	2,2
Cantabria	2,9	1,9	0,7	-1,2	-3,1	**-4,0**	-3,9	-2,9	-1,2	0,2	0,7	1,0	1,1	0,6	0,2	0,2
Castilla La Mancha	3,6	2,5	0,9	-1,0	-2,7	**-4,0**	-3,4	-2,7	-1,5	-0,6	-1,0	-0,6	-0,3	-0,4	-0,7	-1,0
Castilla León	2,8	1,8	0,1	-1,2	-2,9	**-4,0**	-3,5	-2,7	-0,6	0,9	1,2	1,6	1,5	1,3	0,9	1,0
Cataluña	1,8	0,8	-0,3	-1,6	-3,4	**-4,8**	-4,8	-4,0	-1,9	0,0	1,0	1,4	1,2	0,8	0,2	0,2
Extremadura	3,6	3,0	1,2	-0,4	-2,4	**-2,8**	-2,1	-1,4	-0,5	0,5	0,0	-0,1	1,2	1,1	1,6	2,1
Galicia	3,2	2,4	1,1	0,0	-2,2	**-3,7**	-3,5	-3,0	-1,2	0,6	0,7	0,2	0,3	-0,5	-0,9	0,2
Madrid	3,6	2,7	0,4	-2,6	**-4,9**	-4,2	-2,8	-1,4	-0,3	0,2	-0,1	0,3	0,8	0,5	0,9	1,0
Murcia	3,6	2,5	0,9	-0,6	-2,7	**-3,8**	-3,8	-3,5	-2,0	-0,6	-0,1	0,5	1,0	0,8	0,5	1,1
Navarra	3,2	2,6	1,6	0,2	-1,8	**-3,4**	-3,0	-1,8	0,0	1,6	1,5	1,8	1,8	1,4	1,3	0,9
País Vasco	2,3	2,0	1,1	0,0	-2,2	-4,3	**-4,5**	-3,7	-1,5	1,1	1,7	1,9	2,2	1,3	1,0	1,1
La Rioja	3,2	2,6	0,7	-0,4	-3,1	**-4,4**	-3,1	-3,1	-1,0	-0,6	-0,5	0,7	0,5	1,0	1,2	0,2
Valencia	2,4	1,7	0,2	-1,2	-3,6	**-5,1**	-4,8	-4,1	-2,1	-0,5	-0,2	0,3	0,7	0,7	0,9	1,2
Ceuta y Melilla	1,4	3,2	3,6	2,9	4,2	-0,8	**-5,1**	-3,8	-2,1	-1,7	1,6	1,9	1,2	3,8	2,5	2,4

Negative rates
Minimum rate
Rate greater than 1

The table shows how the crisis has affected regions unevenly. For example, we can place the bulk of the recession between the fourth quarter of 2008 and the first quarter of 2010. Most of the regions fell into recession at the same time but not all of them left it out simultaneously, this is the case of regions such as Andalucía, Aragon or Castilla La Mancha where the contractionary period will be particularly long.

In relation to the variance of these results the following graph shows the different boxplot of the annual rates in the quarterly frequency for the different regions:

Graph 1: Box-plot: annual growth rates by region in quarterly frequency, sorted according to weight on Spanish GDP

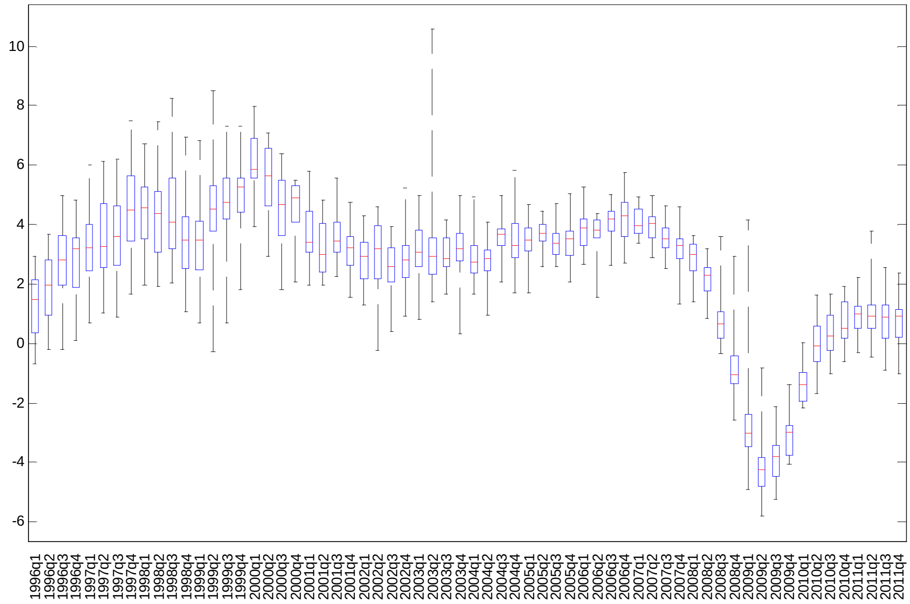


Note: Red line stands for median values, the box represents 50% of the central part of the data and the whiskers are the minimum and maximum of the data.

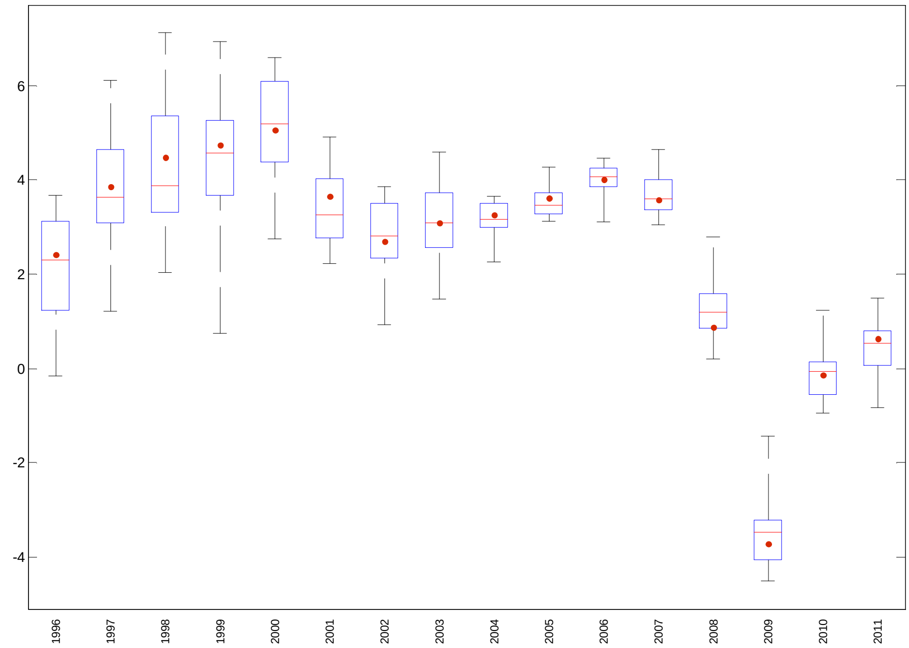
We observe a greater presence of outliers in periods of recession than in expansion. This is partly due to the longer duration of the latter, rendering the median less representative for recessionary quarters. At the same time, the highest rate variability is not linked to the larger size (GDP weight) in the region.

From the temporal dimension of the data, we can appreciate a reduction in volatility after 2003 (see graph 2), although this is a property inherited from the annual data published by the QRA (see graph 3):

Graph 2: Box-plot: annual growth rates by quarter



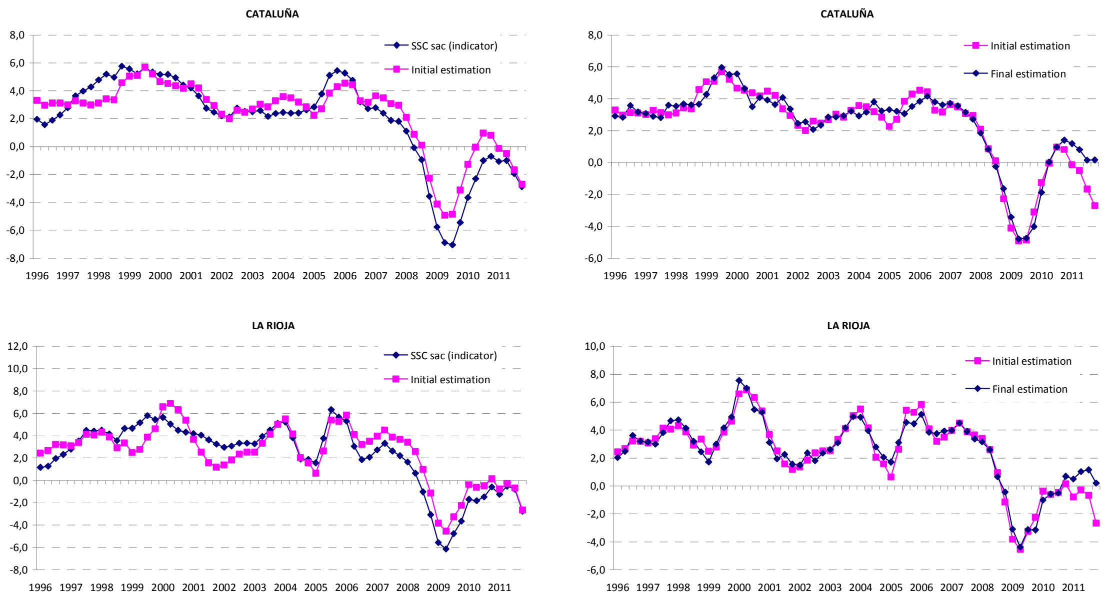
Graph 3: Box-plot: annual growth rates by year



Note: Red dot is the data of Spain

Finally, in order to clarify the importance of the balancing procedure on the final estimate, it has been carried out an exercise on two regions: one with a large size (Cataluña) and other with a small size (La Rioja). The following charts show, firstly, the initial quarterly regional GDP estimation (distribution of annual regional GDP according to the indicator) against the evolution of the indicator and, second, the initial quarterly estimation against the final quarterly GDP.

Graph 4: Importance of the first estimate and the balancing procedure (annual growth rates)



It is easy to see how the first step of estimating quarterly GDP depending on the evolution of the indicator is even more crucial to the subsequent balancing procedure. This fact shows the robustness of the latter, revealing that the variability in the final estimate is driven by the variability of the selected indicator.

4. COMMON TRENDS: AN EXPLORATORY ANALYSIS

In this section we explore the existence of a common trend in the quarterly regional GDP data compiled according to the procedures previously outlined. The existence of a common trend means that the basic drivers of the long run growth are shared by all the regions impliying that real convergence has been achieved. On the contrary, the absence of a common trend implies deep differences in the mechanics of growth and the existence of regions evolving at different speeds.

The dimension of the system precludes the use of a full-system approach and we have followed an exploratory approach akin to the one used by Mayo and Espasa (2011). The main steps of our approach are:

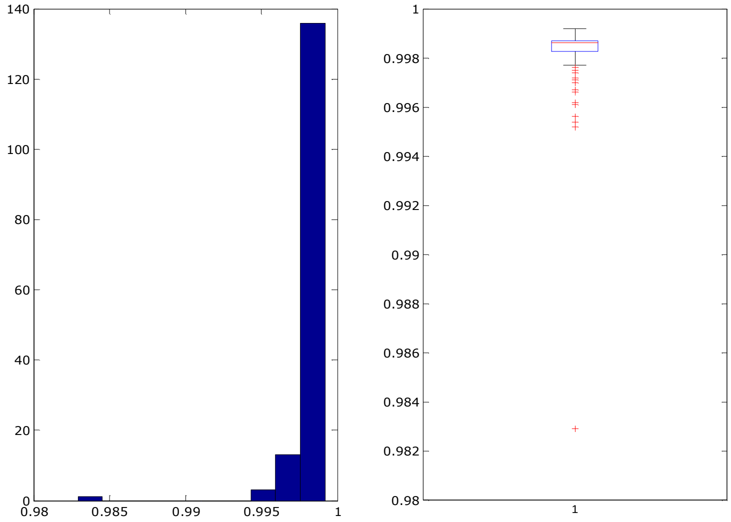
1. For each pair of time series we estimate a bivariate VARMA model. We use a VARMA model to take into account the moving average terms induced by the applicaction of seasonal adjustment procedures, see Maravall (1993) and Lüthkepohl and Claessen (1997). To keep

the models simple but still general we have used a balanced VARMA(1,1) model, see Tiao (2001) for a complete introduction to VARMA modeling.

2. The estimated VARMA(1,1) model allows us to perform the canonical analysis proposed by Box and Tiao (1977). This analysis may be applied to VARMA as well as VAR models and do not require specific distributional assumptions (e.g., Gaussian innovations). Its main drawback is the absence of a formal testing procedure due to the difficulties to derive the distribution of the Box-Tiao canonical eigenvalues, see Bewley and Yang (1994), Bewley et al. (1995).
3. For each pair the Box-Tiao analysis allows us to identify non-stationay patterns, depending on the values of the maximum canonical eigenvalue. If it is close to unity it indicates non-stationarity (trend). If, by contrast, it is close to zero the bivariate system should be regarded as stationary and cointegration does not exist.
4. Excluding all the stationary pairs, we form a distance matrix using the smallest eigenvalue of each pair ($\lambda_{i,j}$). This matrix can be interpreted as a metric for cointegration. Values close to 0 indicate that the series *i* and *j* are cointegrated, thus sharing a common trend. By contrast, values close to 1 exclude this possibility, indicating that the two series are far apart with respect to this feature and that each one has a specific, idyosincratic trend.
5. An exploratory, tentative identification of common trends for the whole set of regions is performed applying cluster analysis to the distance matrix computed in the previous step.

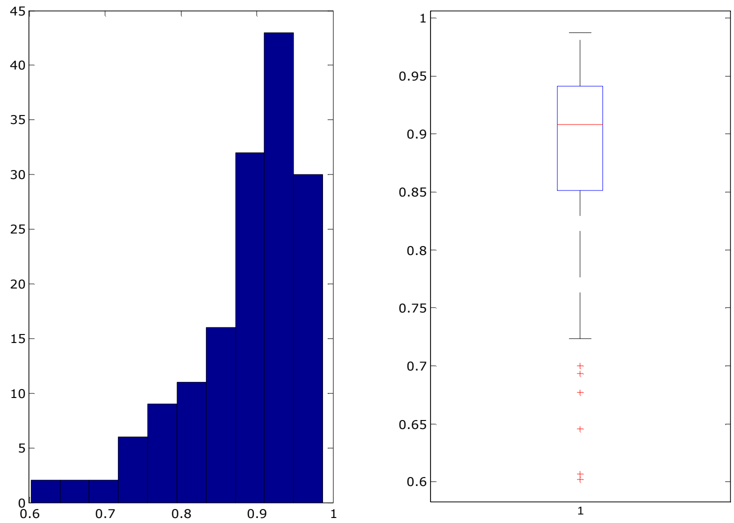
Let us now examine the application of the algorithm to the quarterly regional GDP data. Firstly, the analysis of the largest eigenvalue leads to the conclusion that all pairs of series have at least one non-stationary component, making germane to consider the issue of cointegration. The following graph shows the distribution of these eigenvalues indicating how close are all of them to 1.

Graph 5: Distribution of the largest Box-Tiao canonical eigenvalue



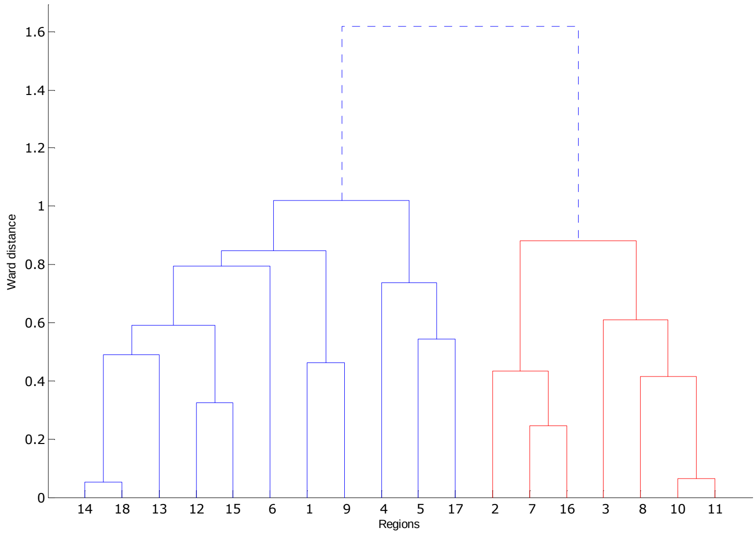
By contrast, the distribution of the smallest Box-Tiao eigenvalue shows much more disperion than the one corresponding to the maximum eigenvalue and indicates that in most pairs the no cointegration is the more likely state.

Graph 6: Distribution of the smallest Box-Tiao canonical eigenvalue



These eigenvalues allow us to perform an exploratory analysis by means of clustering techniques (Everitt, 1993). Confirmatory analysis using formal statistical procedures is left for future reseach. Using the Ward (1963) criterion[12] we get the following hierachical graph:

Graph 7: Dendrogram of Spanish regions according to their pairwise smallest Box-Tiao canonical eigenvalue.



Considering two clusters the corresponding groups are:

Table 8: Two groups clustering

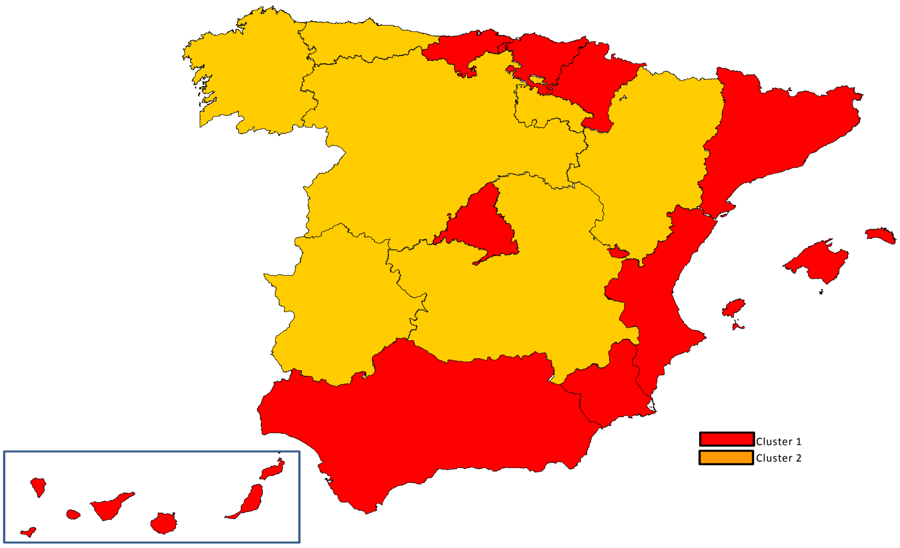
Region	Cluster	Weight on Spanish GDP(2010)	Mean growth (2006-2010)
Navarra	1	1,75	3,07
Murcia	1	2,57	3,27
Madrid	1	17,93	3,13
País Vasco	1	6,30	2,67
Cantabria	1	1,28	2,55
Andalucía	1	13,50	2,74
Cataluña	1	18,64	2,44
Baleares	1	2,51	2,42
Canarias	1	3,89	2,66
Valencia	1	9,61	2,73
Sum / Average		**77,98**	**2,77**
Aragón	2	3,08	2,35
Castilla La Mancha	2	3,38	2,53
Rioja	2	0,74	2,58
Asturias	2	2,18	1,95
Castilla León	2	5,39	2,16
Extremadura	2	1,71	2,89
Galicia	2	5,24	2,11
Sum / Average		**21,72**	**2,37**

[12] This procedure considers an explicit objective function that depends both on the within group sum of squares and on the between groups sum of squares. Additionally, this objective function if affine to the one used by the k-means algorithm, the most widely used non-hierarchical clustering technique.

The first cluster represents most of the Spanish and with a more dynamic stance, according to their average growth rates, above the Spanish average (2.64%).

The following map shows a very significative pattern of clustering along the Mediterranean coast, the Basque Country and, isolated in the center, Madrid. The remaining regions (except Canary Islands) belong to the second cluster.

Figure 1: Geographical two groups clustering



In order to check the robustness of the grouping to the clustering technique we have used the k-means algorithm (Faber, 1994) assuming two groups. Only one region changes its classification (Rioja), moving from cluster 2 to cluster 1.

5. CONCLUSIONS

In this paper we have presented a feasible way to add a regional dimension to short-term macroeconomic forecasting, satisfying the temporal and cross-section constraints imposed by the National Accounts. Our procedure generates results comparable across regions, are based on meaningful short-term information and may be updated at the same time as the flash estimates, providing a solid basis for specific regional forecasts.

Summarizing, the major outcomes of the model are:

- It solves the lack of quarterly GDP at the regional level, providing estimates consistent with the official available data published by the National Accounts, (RA and QNA). These estimates are a stand-alone product that may be used as input in regional econometric models.
- A regional breakdown of the early estimates of the quarterly national volume GDP, that may be released simultaneously providing flash estimates at the regional level.

- Short-term forecasts of quarterly GDP at the regional level by conditioning them on the projected path of the underlying short-term quarterly regional indicators.

There are several promising lines of research that may widen the scope of the paper. Additional indicators, combined using dynamic factor models, should provide a complete description of the state of the business cycle at the regional level. In this sense, the approach proposed by Proietti (2011b) is extremely promising although hampered by its computational complexity.

At the same time, more sophisticated univariante benchmarking techniques may improve the quantitative results. In this vein, an avenue worth to pursue is the implementation of a dynamic extension of the van der Ploeg (1982) procedure. Finally, additional breakdowns (e.g., by industry) may offer a high granularity image of the business cycle, taking into account both its regional dimension and its sectoral dimension.

REFERENCES

Abad, A. and Quilis, E.M. (2007) "Software to perform temporal disaggregation of economic time series", Eurostat, Working Papers and Series.

Abad, A., Cuevas, A. and Quilis, E.M. (2007) "Chain-linked volume indexes: a practical guide", Universidad Carlos III de Madrid, Instituto Flores de Lemus, *Boletín de Inflación y Análisis Macroeconómico*, n 157, p. 72-85.

Bewley, R., Orden, D., Yang, M. and Fisher, L.A. (1994) "Comparison of Box-Tiao and Johansen canonical estimators of cointegrating vectors in VEC(1) models", *Journal of Econometrics*, n. 64, p. 3-27.

Bewley, R. y Yang, M. (1995) "Tests for cointegration based on canonical correlation analysis", *Journal of the American Statistical Association*, vol. 90, n. 431, p. 990-996.

Bloem, A.M., Dippelsman, R.J., and Mæhle, N.O. (2001) *Quarterly National Accounts Manual. Concepts, data sources*, and compilation, International Monetary Fund.

Bournay, J. and Laroque, G. (1979) "Réflexions sur la méthode d'elaboration des comptes trimestriels", *Annales de l'INSEE*, n. 36, p. 3-30.

Box, G.E.P. and Tiao, G.C. (1977) "A canonical analysis of multiple time series", *Biometrika*, vol. 64, n. 2, p. 355-365.

Cancelo, J.R. (2004) "El ciclo del empleo en España: un análisis regional", *CLM.Economía*, n. 4, p. 107-138.

Caporello, G. and Maravall, A. (2004) "Program TSW. Revised manual", Bank of Spain, Ocassional Paper n. 0408.

Chow, G. and Lin, A.L. (1971) "Best linear unbiased distribution and extrapolation of economic time series by related series", *Review of Economic and Statistics*, vol. 53, n. 4, p. 372-375.

Denton, F.T. (1971) "Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization", *Journal of the American Statistical Society*, vol. 66, n. 333, p. 99-102.

Di Fonzo, T. (1987) *La stima indiretta di serie economiche trimestrali*, Cleup Editore.

Di Fonzo, T. (1994) "Temporal disaggregation of a system of time series when the aggregate is known", in Eurostat (Ed.) *Workshop on Quarterly National Accounts*.

Di Fonzo, T. (2002) "Temporal disaggregation of economic time series: towards a dynamic extension", Dipartimento di Scienze Statistiche, Università di Padova, Working Paper n. 2002-17.

Di Fonzo, T. and Marini, M. (2003) "Benchmarking systems of seasonally adjusted time series according to Denton's movement preservation principle", Dipartimento di Scienze Statistiche, Università di Padova, Working Paper n. 2003-09.

Eurostat (1998) *Handbook of Quarterly National Accounts*, Eurostat, Statistical Office of the EC.

Everitt, B.S. (1993) *Cluster analysis*, Arnold Press.

Faber, V. (1994) "Clustering and the continuous k-means algorithm", *Los Alamos Science*, n. 22, p. 138-144.

Fernández, R.B. (1981) "Methodological note on the estimation of time series", *Review of Economic and Statistics*, vol. 63, n. 3, p. 471-478.

Gómez, V. and Maravall, A. (1996) "Programs TRAMO and SEATS", Bank of Spain, Working Paper n. 9628.

Gregoir, S. (1994) "Propositions pour une désagrégation temporelle basée sur des modèles dynamiques simples", in Eurostat (Ed.) *Workshop on Quarterly National Accounts*.

Guérin, P. and Marcellino, M. (2010) "Markov-switching MIDAS models", 6th Eurostat Colloquium on "Modern Tools for Business Cycle Analysis", Luxembourg, 26th – 29th September.

Litterman, R.B. (1983) "A random walk, Markov model for the distribution of time series", *Journal of Business and Economic Statistics*, vol. 1, n. 2, p. 169-173.

Lütkepohl, H. and Claessen, H. (1997) "Analysis of cointegrated VARMA processes", *Journal of Econometrics*, vol. 80, p. 223-239.

Maravall, A. (1993) "Stochastic linear trends. Models and estimators", *Journal of Econometrics*, n. 56, p. 5-37.

Mayo, I. and Espasa, A. (2011) "Forecasting aggregate and disaggregates with common features", Department of Statistics, Universidad Carlos III de Madrid, Working Paper.

Pinheiro, M. and Coimbra, C. (1993) "Distribution and extrapolation of time series by related series using logarithms and smoothing penalties", *Economia*, vol. 17, p. 359-374.

Proietti, T. (2006) "Temporal disaggregation by state space methods: dynamic regression methods revisited", *Econometrics Journal*, vol. 9, p. 357-372.

Proietti, T. (2011a) "Multivariate temporal disaggregation with cross-sectional constraints", *Journal of Applied Statistics*, vol. 38, n. 7, p. 1455–1466.

Proietti, T. (2011b) " Estimation of common factors under cross-sectional and temporal aggregation constraints: nowcasting monthly GDP and its main components ", COMPSTAT, vol. 15, p. 547-558.

Salazar, E., Smith, R., Wright, S. and Weale, M. (1994) "Indicators of monthly national accounts", in Eurostat (Ed.) *Workshop on Quarterly National Accounts*.

Santos Silva, J.M.C. and Cardoso, F. (2001) "The Chow-Lin method using dynamic models", *Economic Modelling*, vol. 18, p. 269-280.

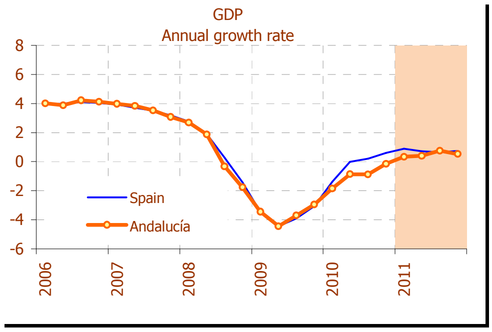
Tiao, G. C. (2001) "Vector ARMA models". In *A course in Time Series Analysis*. Peña, D., Tiao, G. C., and Tsay, R. S. (Eds.) John Wiley and Sons.

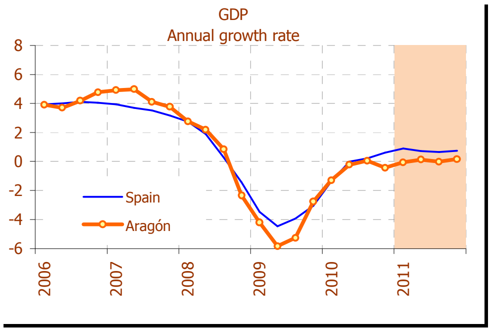
van der Ploeg, F. (1982) "Reliability and the adjustment of large economic accounting matrices", *Journal of the Royal Statistical Society*, series A, vol. 145, part 2, p. 169-194.

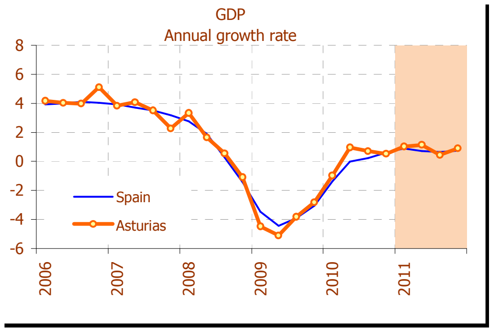
Ward, J.H. (1963) "Hierarchical grouping to optimize an objective function", *Journal of the American Statistical Association*, vol. 53, p. 236-244.

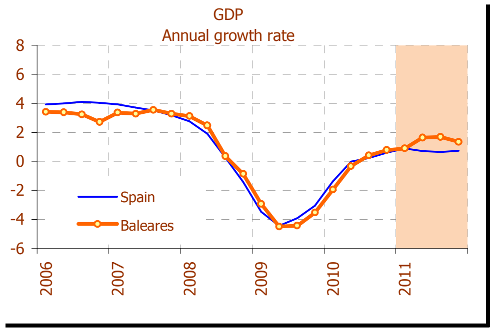
Wei, W.S.W. and Stram, D.O. (1986), ''Temporal aggregation in the ARIMA process'', *Journal of Time Series Analysis*, vol. 7, p. 279-292.

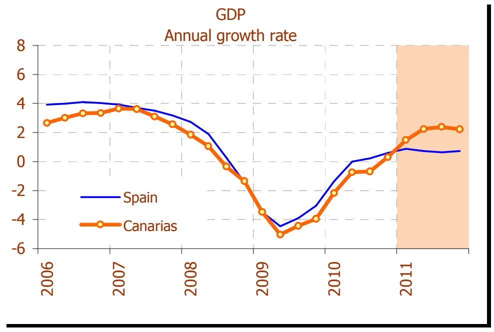
ANNEX 1: GDP growth (Annual rates)

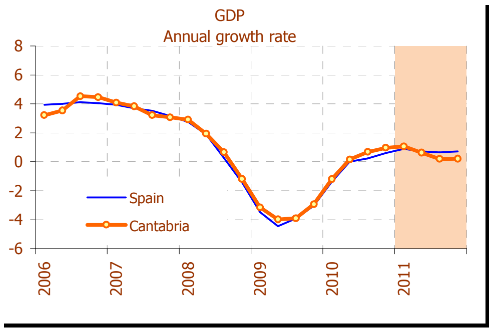


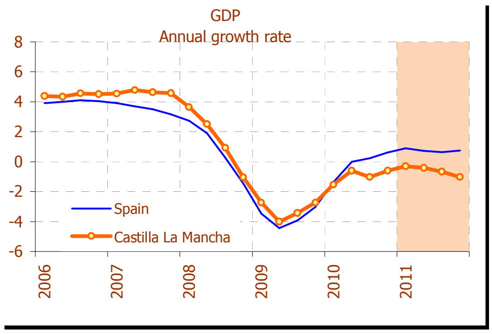


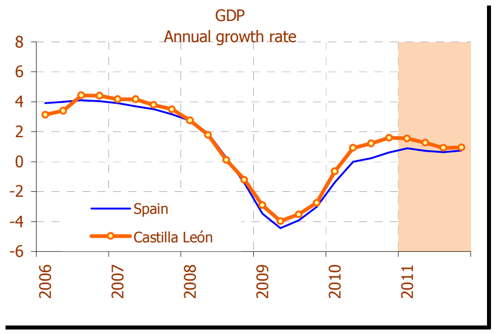


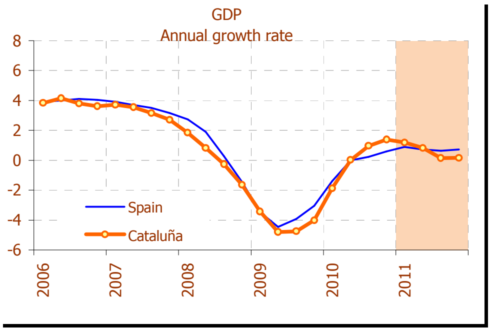


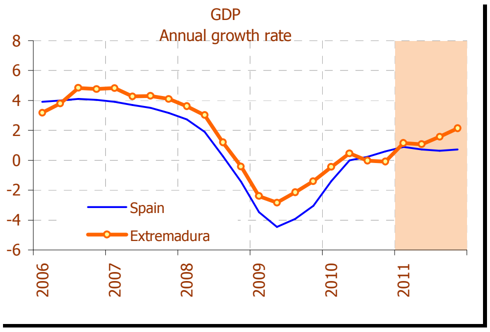


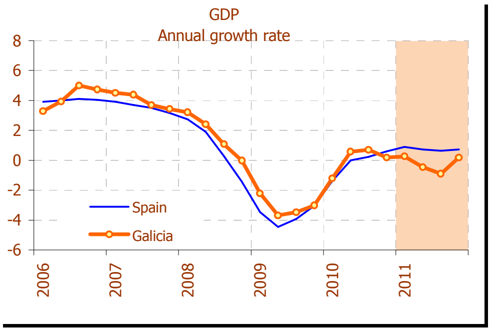


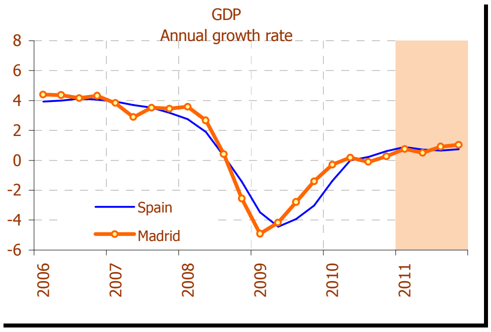


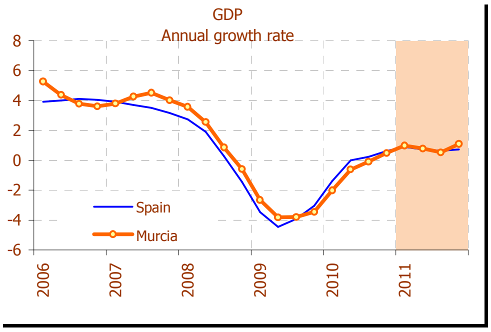


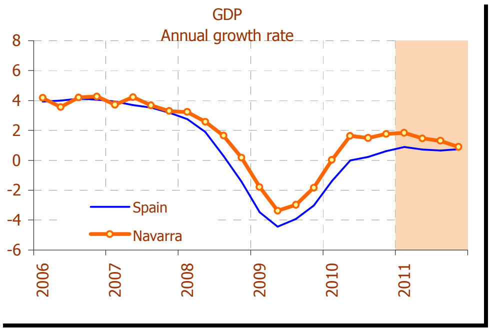

GDP
Annual growth rate
Spain
Cataluña

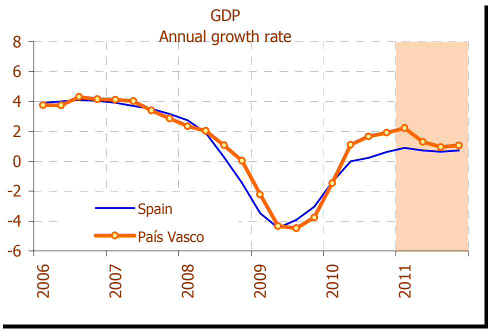

GDP
Annual growth rate
Spain
Extremadura

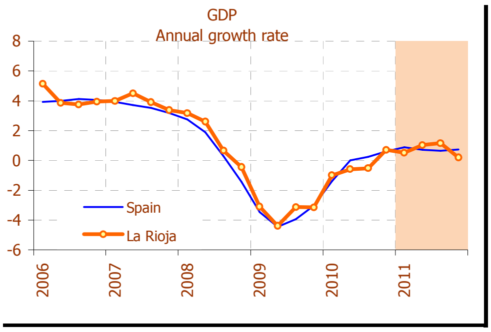

GDP
Annual growth rate
Spain
Galicia

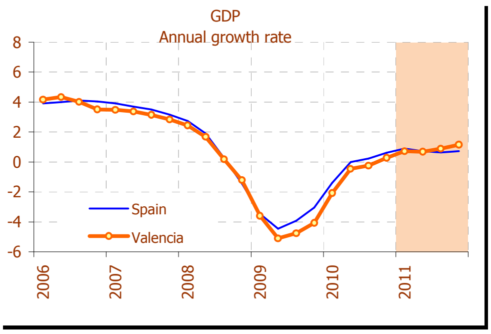

GDP
Annual growth rate
Spain
Madrid

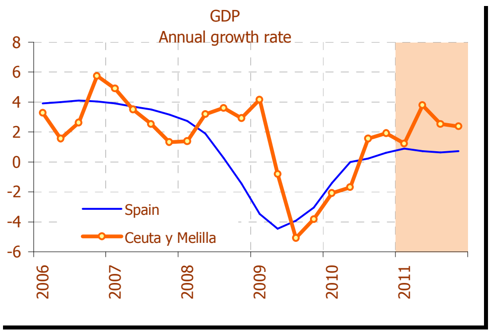

GDP
Annual growth rate
Spain
Murcia


GDP
Annual growth rate
Spain
Navarra

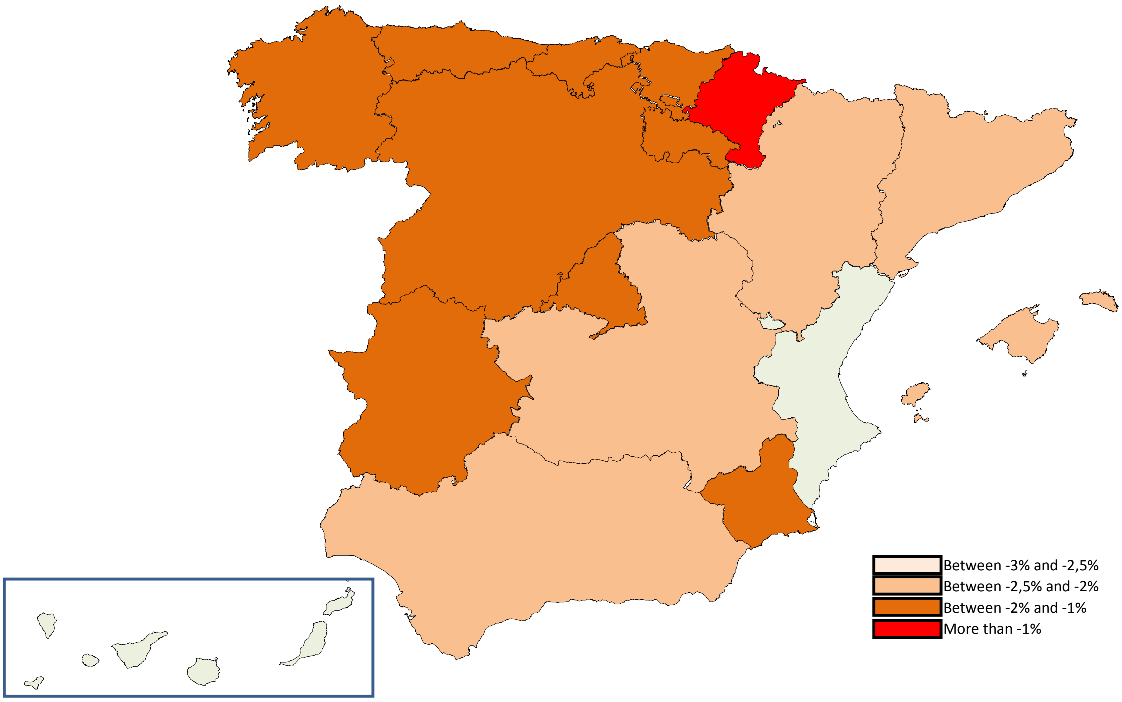

GDP
Annual growth rate
Spain
País Vasco


GDP
Annual growth rate
Spain
La Rioja

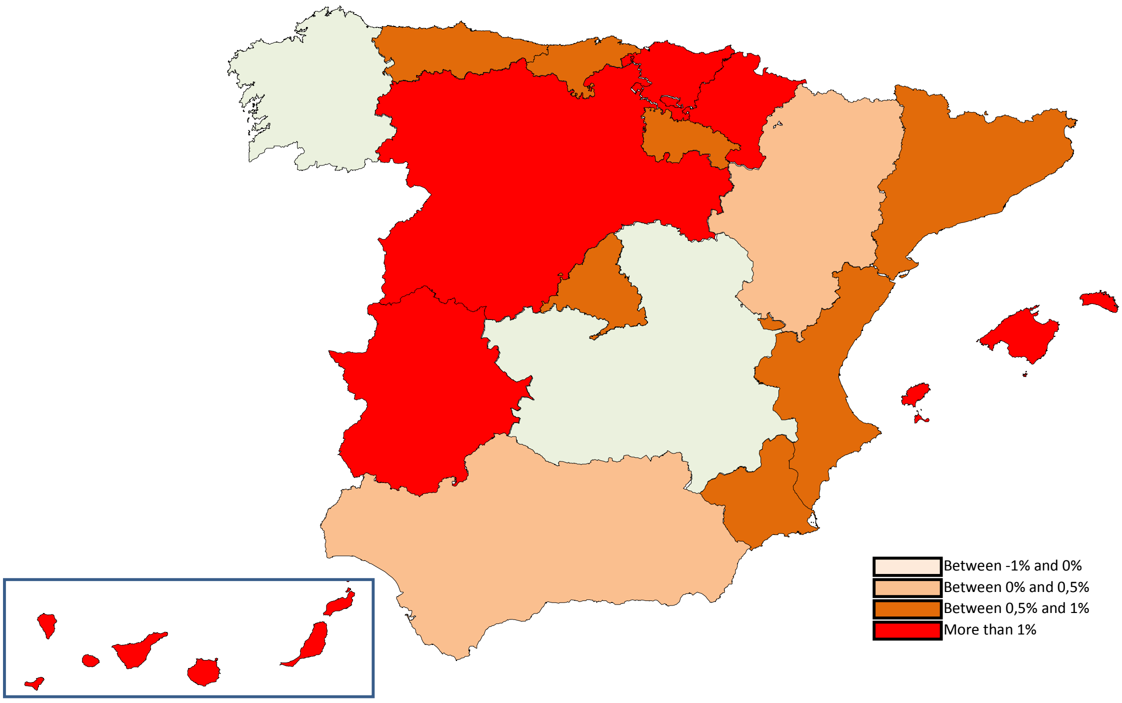

GDP
Annual growth rate
Spain
Valencia


GDP
Annual growth rate
Spain
Ceuta y Melilla

GDP growth by region 2009-2010 (annual rate)



GDP growth by region 2011 (annual rate)



ANNEX 2: Features of Spanish regions (2010)

	Population (thousand)	Population weight	GDP weight	Employment weight
Andalucía	8.238,8	17,9%	13,5%	14,6%
Aragón	1.313,2	2,9%	3,1%	3,1%
Asturias	1.057,1	2,3%	2,2%	2,1%
Baleares	1.080,1	2,3%	2,5%	2,5%
Canarias	2.092,4	4,5%	3,9%	4,0%
Cantabria	579,1	1,3%	1,3%	1,3%
Castilla La Mancha	2.039,5	4,4%	3,4%	3,8%
Castilla León	2.495,0	5,4%	5,4%	5,5%
Cataluña	7.321,1	15,9%	18,6%	17,8%
Extremadura	1.082,4	2,3%	1,7%	2,0%
Galicia	2.736,6	5,9%	5,2%	5,7%
Madrid	6.358,6	13,8%	17,9%	16,6%
Murcia	1.465,8	3,2%	2,6%	2,8%
Navarra	620,7	1,3%	1,8%	1,7%
País Vasco	2.137,9	4,6%	6,3%	5,4%
La Rioja	314,7	0,7%	0,7%	0,7%
Valencia	4.990,6	10,8%	9,6%	10,0%
Ceuta y Melilla	149,2	0,3%	0,3%	0,3%
Spain	46.072,8	100,0%	100,0%	100,0%