BANCO DE ESPAÑA
Eurosistema

ECONOMIC BULLETIN 2/2019
ANALYTICAL ARTICLES

4 June 2019

Real-time regional GDP forecasting: statistical aspects and a forecasting model

Concha Artola, María Gil, Danilo Leiva, Javier J. Pérez and Alberto Urtasun

Abstract

The monitoring of the regional economic situation takes on particular importance in highly decentralised countries, such as Spain. Against this background, this article summarises the key aspects of the BayFaR model (Bayesian Factor model for Regions), a new tool used by the Banco de España for the short-term forecasting of the GDP of the Spanish economy and of its biggest regions. The tool allows monthly indicators of activity to be combined with quarterly GDP, the benchmark variable for conjunctural analysis. In Spain's case, however, it is necessary to use alternative measures of this variable. This is because the official statistics do not provide a homogeneous measure for all the regions, nor one that is consistent with INE's national quarterly GDP.

Keywords: regional economic activity, nowcasting, dynamic factor model.

JEL codes: C32, E37, R13.

REAL-TIME REGIONAL GDP FORECASTING: STATISTICAL ASPECTS AND A FORECASTING MODEL

The authors of this article[1] are Concha Artola, María Gil, Danilo Leiva, Javier J. Pérez and Alberto Urtasun of the Directorate General Economics, Statistics and Research.

Introduction

For a central bank, the characterisation of the economy's cyclical situation is a particularly important task. In this connection, it regularly distils the latest conjunctural data, in which process econometric models are used as an aid. The Banco de España has short-term forecasting tools (typically one or two quarters ahead) for the quarterly GDP growth of the Spanish economy taken as a whole [see Álvarez *et al.* (2014), Arencibia *et al.* (2017) and Camacho and Pérez Quirós (2010 and 2011)] and in real time (i.e. as and when new data on the determining variable become known).

Against this background, this article summarises the key aspects of a new tool used by the Banco de España for the short-term forecasting of the GDP of the Spanish economy and of the biggest regions. Specifically, the model, known as the BayFaR *(*Bayesian Factor model for Regions*),* is applied to the regions of Andalusia, Catalonia, Madrid and Valencia. The monitoring of the regional economic situation takes on particular importance in countries, such as Spain, with a high level of decentralisation in respect of economic policy competences. The regional view, moreover, enables the heterogeneity of the productive structure to be characterised in a more granular fashion (see Chart 1), along with the connection between the different regional economies and their aggregate consequences.

The type of analysis that can be conducted is conditional upon data availability, given that no official measure of quarterly GDP is available for all the regions that is consistent with the national reference of the National Statistics Institute (INE). Quarterly GDP is the benchmark variable for conjunctural analysis. This situation contrasts with the availability of a high number of indicators of economic activity; these are homogeneous, for all the regions, and refer to a broad set of areas such as industrial production, consumption, services sector activity, the housing market and financial conditions (see Table 1).

However, under official statistics the quarterly GDP figures released by the regional statistics offices are available, albeit only for some regions. In view of this situation, various analysts regularly produce quarterly regional GDP estimates [see Artola *et al.* (2018)]. The most popular of these is the database published by AIReF (the Spanish Independent Authority for Fiscal Responsibility), with own estimates and time series for all the regions [see Cuevas and Quillis (2015)]. These are consistent with INE's national measure, but differ from the regional one in cases where the latter measure is available.

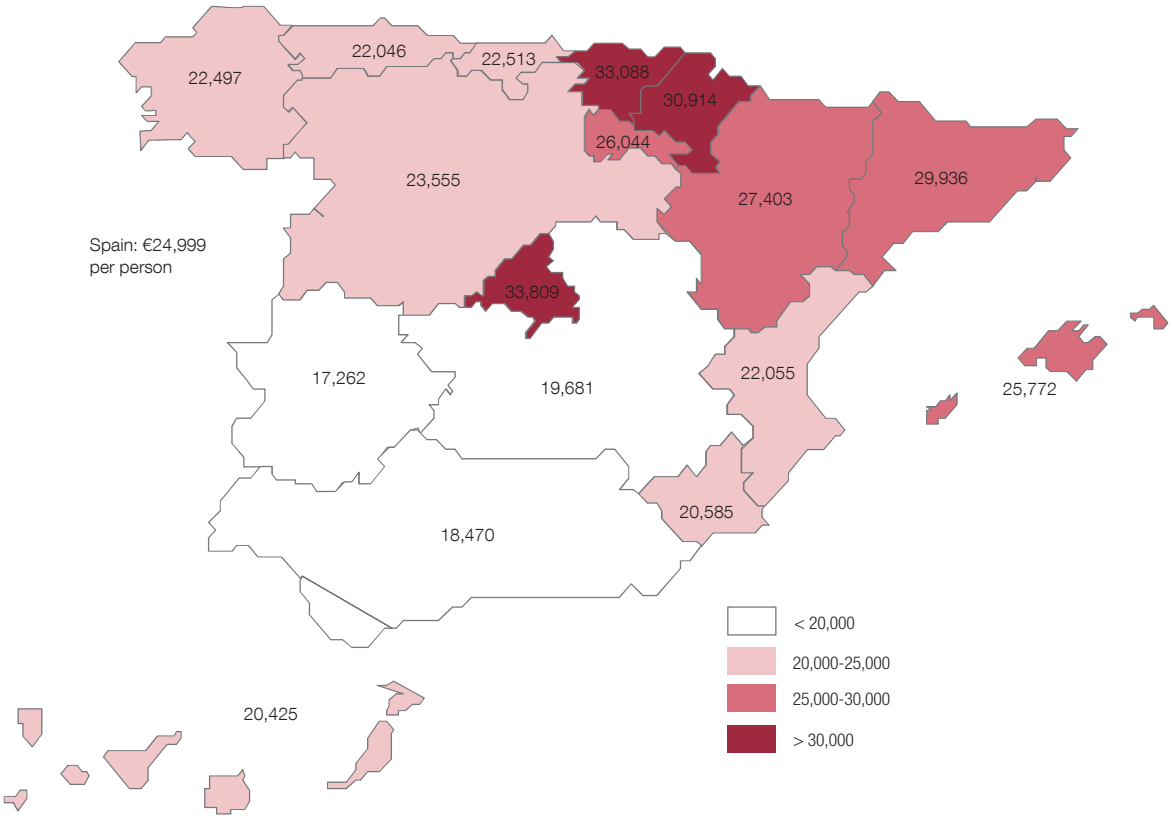
The rest of the article reviews, firstly, the national accounts sources available for the regional economies, and those used within the framework of the BayFaR model. The final section presents the main elements of this model, along with examples of its use for the real-time short-term forecasting of regional economic activity.

Availability of data on aggregate regional economic activity

INE draws up the CRA (Spanish Regional Accounts) annually. The CRA provide a systematic and homogeneous measure of economic activity in the Spanish regions, and they are integrated into the conceptual and quantitative framework of Spain's national accounts

1 This article is based on Artola et al. (2018) and Gil et al. (2019).

SPANISH PER CAPITA GDP FROM A REGIONAL PERSPECTIVE (2017) CHART 1



SOURCE: INE.

(CNE); accordingly, they are perfectly compatible with CNE (see Chart 2). However, as mentioned in the previous section, INE does not publish quarterly national accounts for the regions.

In this context, a wide set of regional public statistics institutes publish their own quarterly and, in some cases, annual accounts. Currently, quarterly accounts are available for 11 of the 17 Spanish regions[2] (see Table 2). In some cases, the estimated quarterly accounts take INE's CRA as a reference, while in others, the regional institute prepares its own annual national accounts. True, in all cases the procedure followed adheres to the general principles in the European System of Accounts (ESA-2010, in its latest version). But the methodology is not necessarily compatible with CRA in the selection of sources, operational procedures or time coverage. Chart 3 shows some examples of the discrepancies between INE's CRA and those prepared by some selected regions (Catalonia, Galicia, the Basque Country and Navarre). The differences are, generally, small; but in specific periods they may be significant and relevant. For instance, in 2015 the growth in Galicia, Navarre and the Basque Country estimated by the respective regional institutes was lower by at least 1 pp than that published by INE.

As earlier stated there are non-official sources, including most notably that published by AIReF. This institution provides quarterly estimates of GDP for all the regions, prepared with the same methodology and consistent both with the CRA and with INE's Quarterly

2 Andalusia, Aragon, the Canary Islands, Cantabria, Castile-Leon, Catalonia, Extremadura, Galicia, Navarre and the Basque Country.

SELECTED SHORT-TERM INDICATORS OF REGIONAL ECONOMIC ACTIVITY

TABLE 1

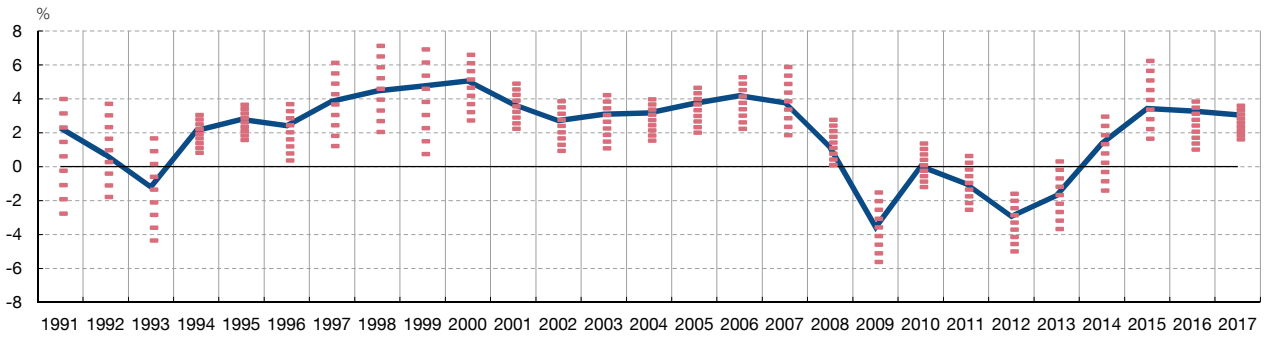
Indicator (a)	Frequency	Period	Publication (b)	Downloaded from website of:
Deposits	Quarterly	Since 1986 Q3	t+90	Banco de España
Total credit	Quarterly	Since 1987 Q1	t+90	Banco de España
Freight transport	Quarterly	Since 1993 Q1	t+58	Ministerio de Fomento
Consumer Price Index (CPI)	Monthly	Since Jan 1980	t+14	INE
Social Security registrations	Monthly	Since Jan 1982	t+3	Ministerio de Empleo y Seguridad Social
Public works procurement	Monthly	Since Jan 1989	t+53	Ministerio de Fomento
Industrial Production Index	Monthly	Since Oct 1991	t+37	INE
Car registrations	Monthly	Since Apr 1992	t+15	Ministerio del Interior (DGT)
Commercial motor vehicle registrations	Monthly	Since Apr 1992	t+15	Ministerio del Interior (DGT)
Mortgages on dwellings	Monthly	Since Jan 1994	t+57	INE
House sales and purchases	Monthly	Since Jan 2007	t+57	INE
Exports	Monthly	Since Jan 1995	t+50	Ministerio de Industria, Comercio y Turismo
Imports	Monthly	Since Jan 1995	t+50	Ministerio de Industria, Comercio y Turismo
Social Security-registered businesses	Monthly	Since Apr 1996	t+33	Ministerio de Empleo y Seguridad Social
Registered unemployment	Monthly	Since Jan 1996	t+3	Ministerio de Empleo y Seguridad Social
Gas, oil and fuel consumption	Monthly	Since Jan 1997	t+42	CORES (c)
Overnight stays	Monthly	Since Jan 1999	t+22	INE
Travellers	Monthly	Since Jan 1999	t+22	INE
Retail trade index	Monthly	Since Jan 2003	t+28	INE
Services Sector Activity Indicator (IASS)	Monthly	Since Jan 2002	t+50	INE
Industrial New Orders Indices	Monthly	Since Jan 2002	t+50	INE
Industrial Turnover Indices	Monthly	Since Jan 2002	t+55	INE

SOURCES: INE, Ministerio de Fomento, Ministerio de Empleo y Seguridad Social, Ministerio del Interior (DGT), Ministerio de Industria, Empleo y Turismo and Cores.

a For a detalied description of each indicator, see Annex I.
b Refers to the lag (in days) of indicator release date.
c CORES: "Corporación de Reservas Estratégicas de Productos Petrolíferos", Public Law Corporation of the Spanish Ministry for the Ecological Transition.

SPANISH REAL GDP GROWTH AND REGIONAL HETEROGENEITY (a)

CHART 2



SOURCES: INE and Spanish Regional Accounts.

a The vertical dotted lines indicate the distance between the maximum and minimum growth rates of each of the 17 regions.

Download

PUBLICATION OF QUARTERLY NATIONAL ACCOUNTS BY REGIONAL STATISTICS INSTITUTES

TABLE 2

	Regional Statistics Institute (link to website)	Availability	Consistency with Spanish Regional Annual Accounts (SRA) prepared by INE (a)
Andalusia	IECA	Yes	No
Aragon	IAEST	Yes	No
Asturias	SADEI	No	—
Balearic Islands	IBESTAT	No	—
Canary Islands	ISTAC	Yes	Yes
Cantabria	ICANE	Yes	No
Castile-Leon	ECL	Yes	No
Castile-La Mancha	IES	No	—
Catalonia	IDESCAT	Yes	No
Valencia Region	PEGV	No	—
Extremadura	IEEX	Yes	Yes
Galicia	IGE	Yes	No
Madrid	IEM	Yes	No
Murcia	CREM	No	—
Navarre	NASTAT	Yes	No
Basque Country	EUSTAT	Yes	No
La Rioja	IER	No	—

SOURCE: Artola *et al.* (29018).

a Consistency refers to the cut-off date of the information for Artola et al. (15 July 2018).

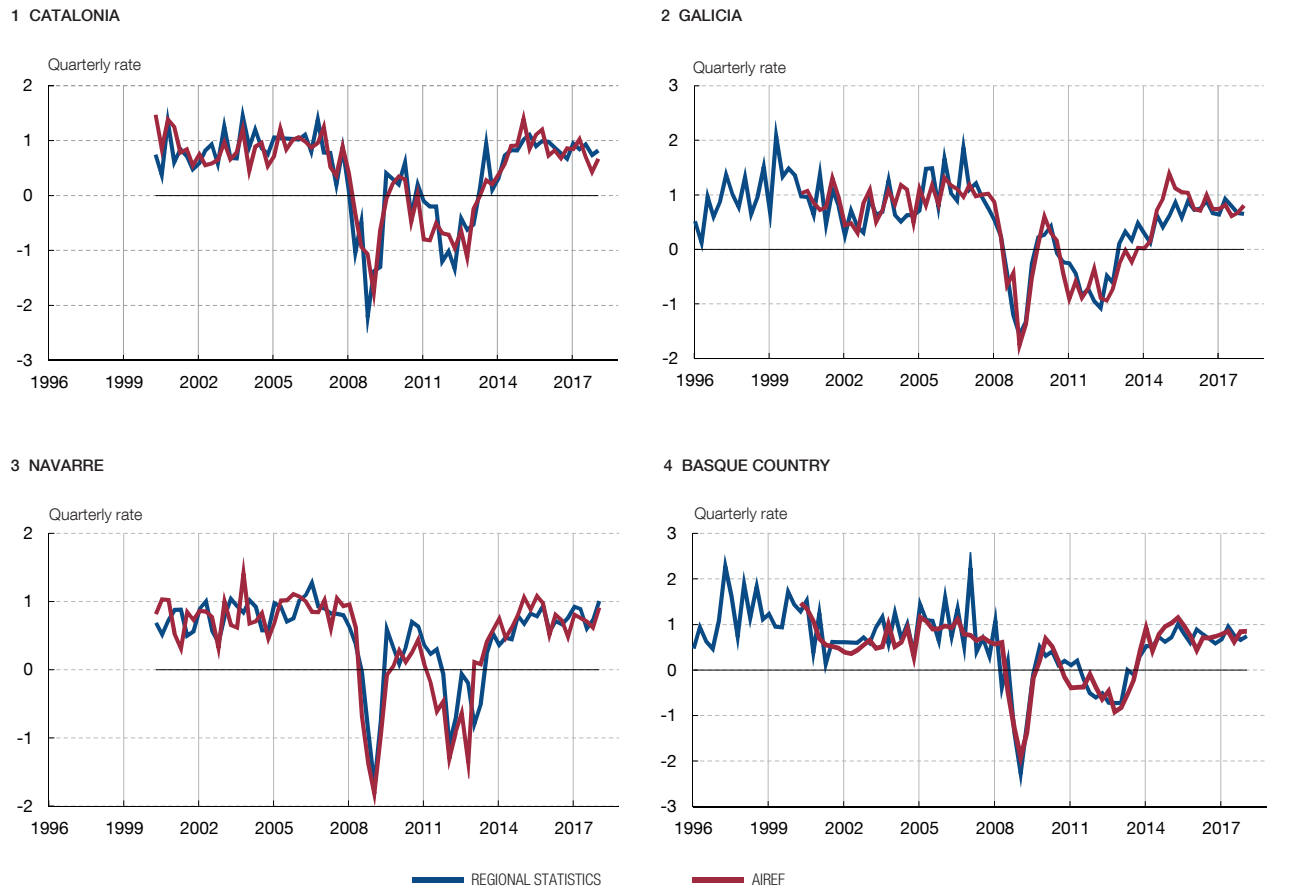
ANNUAL REGIONAL REAL GDP ACCORDING TO INE AND REGIONAL STATISTICS INSTITUTES' FIGURES

CHART 3

SOURCES: INE and regional statistics institutes.

Download

REGIONAL REAL GDP ACCORDING TO THE REGIONAL STATISTICS INSTITUTES AND AIReF CHART 4



SOURCES: INE, AIReF and Regional Statistics Institutes (or similar).

Download

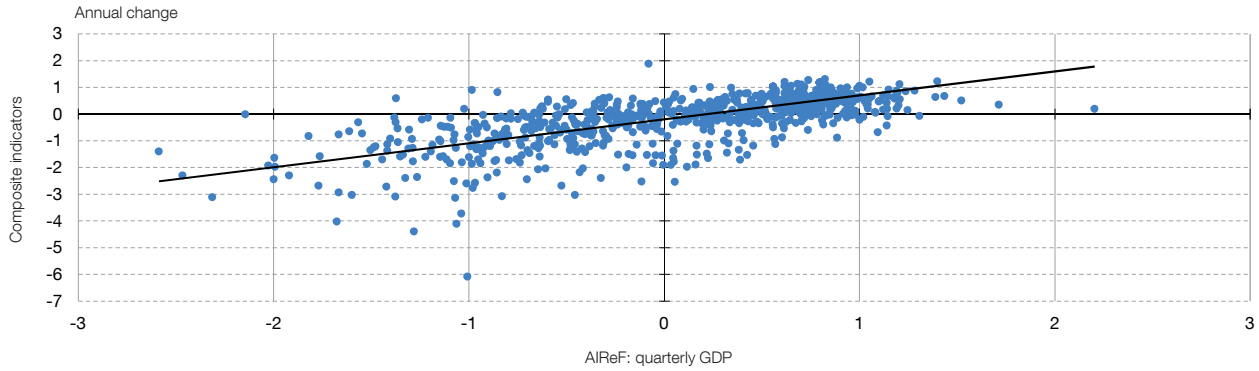
National Accounts. Chart 4 shows all the differences between AIReF's quarterly figures and those of the institutes of the regions selected as examples. The patterns are, generally, very similar. But – as was the case with the annual comparison – there are possibly significant differences at certain specific times, which may distort the real-time signals available for conjunctural analysis.

On the basis of the conjunctural indicators available, composite indicators of activity can be constructed as alternatives to the use of national accounts-based metrics. As could be seen in Table 1, Spain has a broad set of monthly and quarterly indicators of different aspects of economic activity, homogeneously prepared by different national institutions for the 17 regions. One simple approach to the conjunctural monitoring of regional economic activity is to construct monthly composite indicators based on a set of indicators selected by the analyst. Chart 5 shows a composite indicator obtained as the first principal component of 8 reference indicators (Social Security registrations, retail trade index, industrial production index, new car registrations, services sector activity indicator, exports, imports and overnight hotel stays, both by residents and foreigners) that approximates to the quarterly GDP of AIReF and of the regional institutes. In fact, for the period covered by the charts (2007 Q4-2017 Q4), the simple correlation between the composite indicator and the two measures of GDP is 0.71 and 0.72, respectively.

GDP GROWTH IN THE DIFFERENT REGIONS ESTIMATED BY AIREF, COMPARED WITH THE ESTIMATES OF THE COMPOSITE INDICATORS

CHART 5

1 AIREF ESTIMATES COMPARED WITH THE COMPOSITE INDICATORS



2 ESTIMATES BY THE REGIONAL STATISTICS INSTITUTES OR SIMILAR, COMPARED WITH THE COMPOSITE INDICATORS

SOURCES: INE, regional statistics institutes or similar and AIReF.

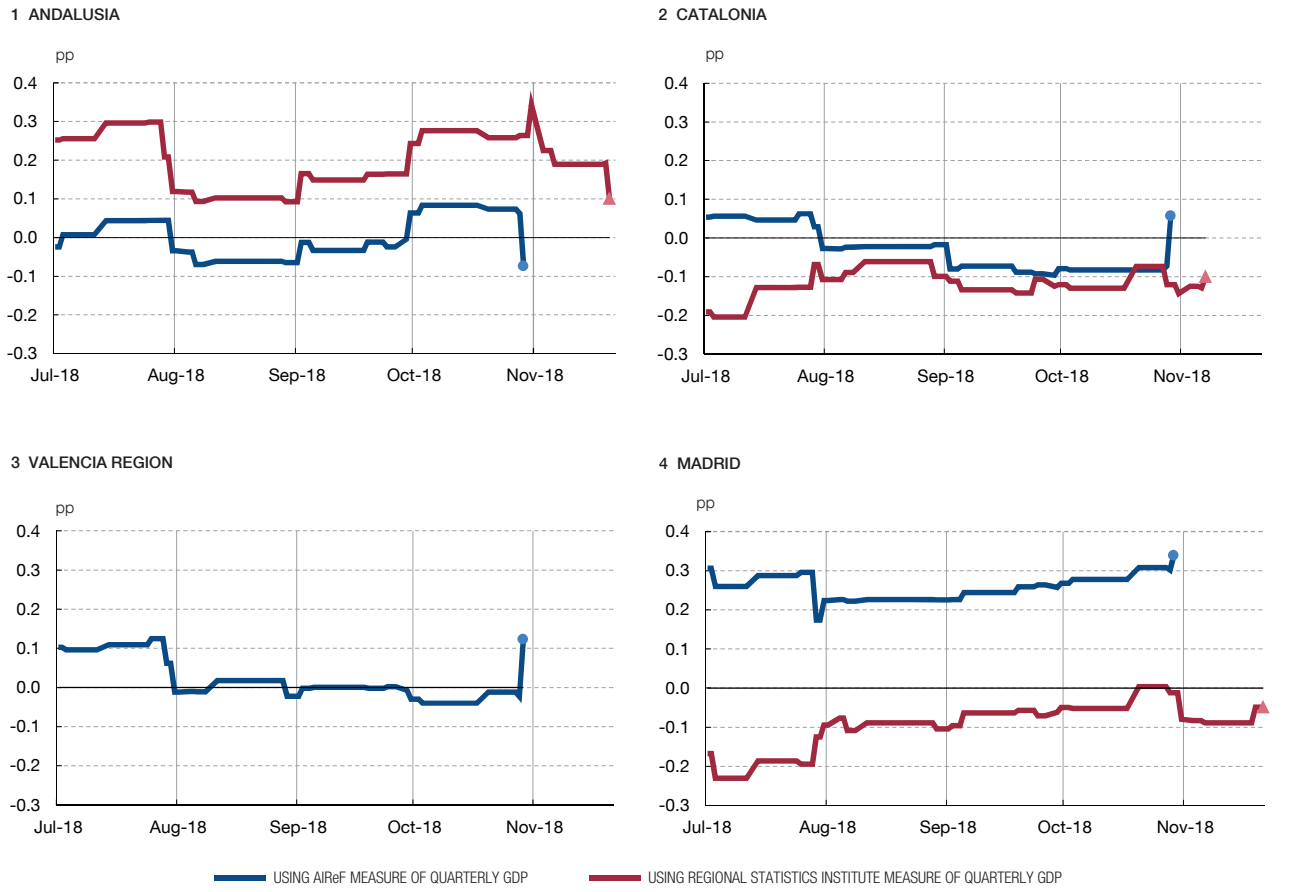
Download

BayFaR: econometric methodology and some results

The BayFaR model allows monthly indicators of activity to be combined with measures of quarterly GDP, the benchmark variable for conjunctural analysis. All the methodological details can be consulted in Gil *et al.* (2019). What is involved here is a dynamic factor model that allows for the parsimonious inclusion of the relevant information available at the time the estimates are made, at different frequencies (quarterly GDP and monthly indicators). In keeping with the habitual methodology, the monthly indicators selected are: Social Security registrations, retail trade index, industrial production index, services sector activity indicator and new commercial vehicle registrations. The authors adopt the Bayesian perspective because the latest work in macroeconomic forecasting indicates that this type of model enables more robust estimates to be obtained than those from classical forecasting models. Gil *et al.* (2019) independently model the quarterly GDP of Andalusia, Catalonia, Madrid and the Valencia region. In the first three cases are models that take the quarterly GDP of AIReF as a reference, on one hand, and that of the related regional statistics agency, on the other. In the case of Valencia, the regional statistical institute does not publish the region's quarterly GDP. The same model is also applied to obtain the GDP nowcast of the Spanish economy as a whole, so as to have an aggregate reference available.

The projections path obtained with the BayFaR for each of the four regions provides for a comparative assessment of the impact of the latest information on regional economic

ACCELERATION (+) / DECELERATION (-) IN REAL GDP FOR 2018 Q3 ACCORDING TO THE BAYFAR MODEL (a) CHART 6



SOURCE: Gil et al. (2019).

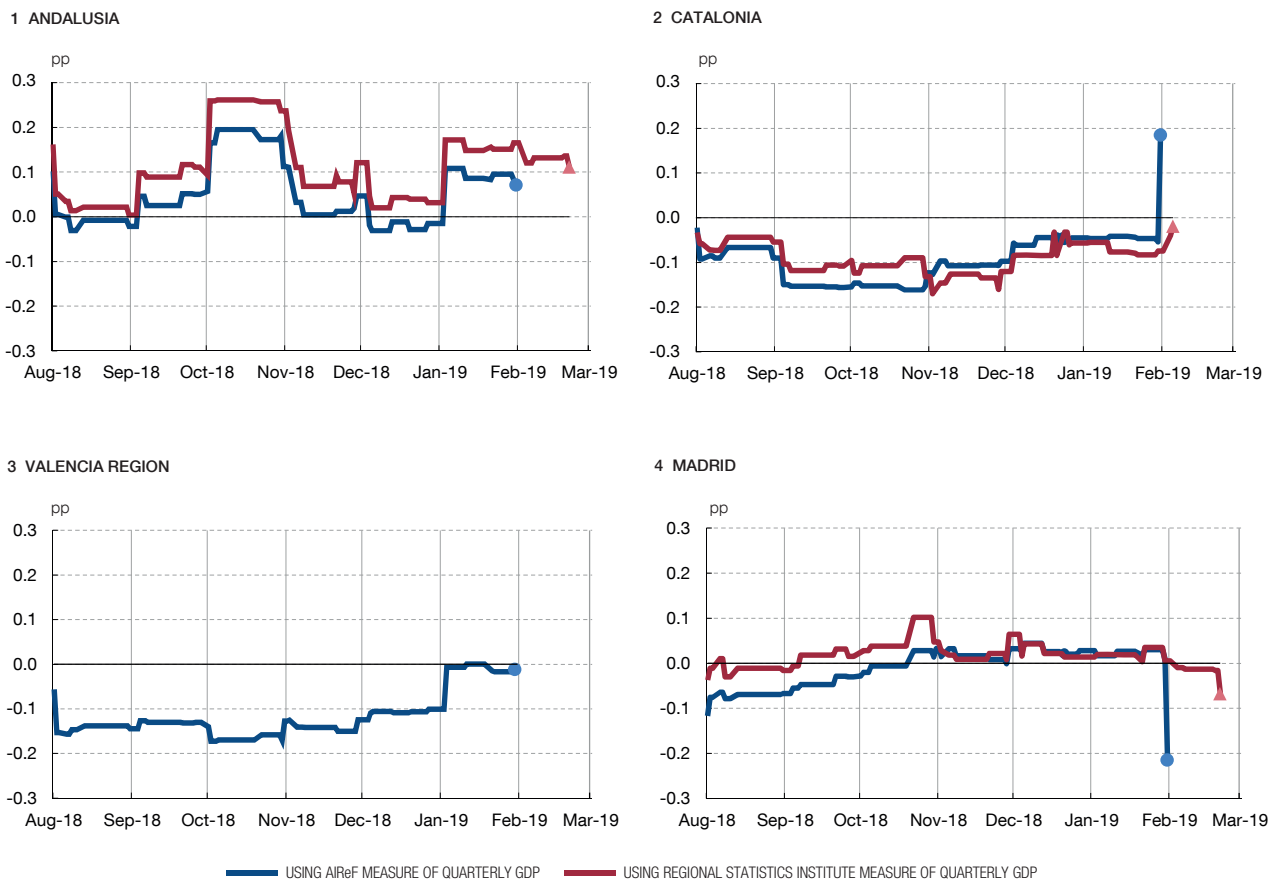
a Calculated as the difference between the point estimate in real time for 2018 Q3 and the same agency's estimate for 2018 Q2.

Download

activity. Notably, in interpreting the results of these models, greater significance is usually given to changes in the projections over time than to the point estimates themselves, which should not generally be interpreted as the best available projection on GDP growth [see Banco de España (2019)]. The models generate forecasts for each of the indicators they include and, therefore, the GDP growth forecast will only be revised if the new information differs from that expected by the model. Hence, it is the positive (negative) surprises in the data that impact positively (negatively) on the GDP estimate.

The exercises performed with the models highlight their usefulness for monitoring regional activity in real time, and the need to use different GDP measures. Charts 6 and 7 present the results of the application in real time, using the tool described, in the last two quarters of 2018. The dates on the x-axis reflect the moments at which a new figure is published for each of the five monthly indicators selected, during each quarter, until the moment at which the quarterly GDP figure is released, and the differential impact generated by its inclusion in the GDP forecast for the quarter. Let us take as an example the first panel in Chart 6, for Andalusia. The flow of new information between 9 August and 2 November 2018 led, according to the red line [model that includes as a reference the GDP of the IEA (Andalusian Statistics Institute)], to estimates of a greater acceleration in Andalusian GDP, of the order of 0.1-0.3 pp of growth, whereas the data published from that time to the GDP

ACCELERATION (+) / DECELERATION (-) IN REAL GDP FOR 2018 Q4 ACCORDING TO THE BAYFAR MODEL (a) CHART 7



SOURCE: Gil et al. (2019).

a Calculated as the difference between the point estimate in real time for 2018 Q4 and the same agency's estimate for 2018 Q3.

Download

release data (22 November) were interpreted by the model as more negative, prompting downward revisions, more in line with the figure published by the IEA.

The signals extracted in real time taking as a reference the two GDP measures used in this article occasionally do not match; accordingly, it would seem appropriate to monitor all the information available, given the strengths of both. Charts 6 and 7 are presented in terms of accelerations and decelerations, which provides for control of the fact that the GDP growth rates of the two sources differ. Notwithstanding, different signals are discerned when conjunctural monitoring is in real time. Taking the Madrid region as an example, in Chart 6 (2018 Q3) the data, filtered by the model with the AIRef GDP, would be indicating an acceleration in GDP of between 0.2 and 0.3 pp between 15 July and 30 October, which provides a very similar GDP estimate to the first estimate published by AIReF. However, if the regional institute's GDP is taken as a reference, the signal would be one of slowdown or maintenance of growth, in line with the official estimate. In the case of Q4 (see Chart 7), although the signals associated with the new data were similar (maintenance of growth), the final estimate was a significant downward surprise in the case of the AIReF GDP-based model (publication showed a slowdown of 0.22 pp compared with the previous day's forecast of 0.01 pp), and substantially less in the other case (-0.07 pp, compared with the expected figure of -0-01 pp). As earlier discussed, the figures published by the regional institutes have the advantage of being official statistics. However, they are not constructed so that they are consistent, in aggregate terms, with INE's national GDP. And nor are they

available for all the regions. The main advantages of AIReF's quarterly GDP are that it is available for all the regions and is fully consistent with INE's GDP for Spain as a whole.

4.6.2019.

REFERENCES

ÁLVAREZ, L. J., A. CABRERO and A. URTASUN (2014). "A procedure for short-term GDP forecasting", *Economic Bulletin,* October, Banco de España, pp. 73-80.

ARENCIBIA, A., A. GÓMEZ LOSCOS, M. DE LUIS and G. PÉREZ QUIRÓS (2017). "A short-term forecasting model for GDP and its demand components", Analytical Articles, *Economic Bulletin,* 4/2017, Banco de España.

ARTOLA, C., A. FIORITO, M. GIL, J. PÉREZ, A. URTASUN and D. VILA (2018). *Monitoring the Spanish Economy from a Regional Perspective: Main Elements of Analysis,* Occasional Papers, no. 1809, Banco de España.

BANCO DE ESPAÑA (2019), "A comparison of recent economic developments in Spain and in the euro area", Box 6 of the "Quarterly report on the Spanish economy", *Economic Bulletin* 1/2019.

CAMACHO, M., and G. PÉREZ-QUIRÓS (2010). "Introducing the Euro-Sting: Short-Term Indicator of Euro Area Growth", *Journal of Applied Econometrics,* 25(4), pp. 663-694.

— (2011). "Spain-STING: Spain Short Term INdicator of Growth", *The Manchester School,* 79, pp. 594-616.

CUEVAS, A., and E. QUILLIS (2015). *Quarterly Regional GDP Flash Estimates for the Spanish Economy (METCAP model),* AIReF Working Paper 2015/3.

GIL, M., D. LEIVA, J. J. PÉREZ and A. URTASUN (2019). *An Application of Dynamic Factor Models to Nowcast Regional Economic Activity in Spain,* Occasional Papers, no. 1904, Banco de España.