# Vulnerable regional growth: the case of Spain


Martín Llada
Universidad de Buenos Aires
ARGENTINA

Pilar Poncela
Universidad Autónoma de Madrid
SPAIN

Eva Senra
Universidad de Alcalá
SPAIN


April 20, 2023


**Abstract**

The heterogeneity of the regions' structural characteristics induces differences in their growth and renders some more vulnerable to downside risks. Economic policy requires the assessment of vulnerability at a regional level to mitigate the effects of business cycle downturns. This task becomes more difficult in real-time, due to the need of using the available monthly information to nowcast annual GDP figures that are released with a delay of 12 months after the end of the reference period. In this paper we build a monthly indicator of economic activity for each region in Spain that poses the basis for quantile regressions that characterize real-time vulnerability to downside shocks associated to the business cycle, as well as regional and sector heterogeneity.



**Keywords**: composite indicator, regional GDP, vulnerability, quantile regressions.

# 1 Introduction

Conducting economic policy depends upon the assessment of the state of the economy in real time, a challenging task that requires both, nowcasting of macroeconomic indicators, as well as an analysis of the risks and vulnerability. The need for early assessment becomes highly relevant, especially on the irruption of any business cycle downturn. From a regional point of view, structural characteristics within the different regions might exacerbate their vulnerability to downside risks when exposed to a single global policy. The importance of regional heterogeneity has been recognized in the literature. For instance, Chamie et al. (1994) and Beine and Coulombe (2003) have shown that regional heterogeneity in a country's industrial structure can lead to a variety of responses to external shocks and changes in business cycles across regions. For the US, Owyang et al. (2005) find that those states with strong sectors in mining, construction and manufacturing face more severe recessions than the national average. This effect is not only detected on real activity variables but it also matters for monetary policy, as the transmission of shocks can be heterogeneous across regions (see, for instance, Carlino and DeFina, 1998; Fratantoni and Schuh, 2003). The situation is even more critical in the euro area where the European Central Bank (ECB) faces heterogeneity not only at the national but also regional level. In Spain, there are 17 autonomous regions with significant regional differences involving to their geographical situation, their productive structure and their local decentralized governments. In this regards, Villaverde and Maza (2009), Gadea et al. (2012), Bandrés and Gadea (2013), Camacho et al. (2019) or, more recently, Bandrés et al. (2021) amongst others, provide evidence of business cycle synchronization but also considerable regional disparities that suggest the need for their individual analysis.

On a regional basis, GDP figures are reported only on an annual basis, 12 months after the end of the reference period, and which makes it impossible to know the state of the region in real time. It is therefore, essential to produce a monthly composite indicator to assess the economic momentum and the real-time vulnerability to downside shocks. This need increases at business cycle turning points. There is related literature on regional GDP nowcasting that has been mainly addressed either on the basis of MIDAS approaches (Grant et al., 2014; Claudio et al., 2020), bridge equations (Lehmann and Wohlrabe, 2014, 2015; Henzel et al., 2015; Lehmann and Wohlrabe, 2017), or factor models (Clayton and Clayton-Matthews, 2005; Kopoin et al., 2013; Chernis et al., 2020). Our monthly indicator would fall within this last group.[1]

To measure regional economy vulnerability, one has to go beyond point spot forecasts and include

[1]See Bańbura et al. (2013) and Camacho et al. (2013) for a review and comparison of different procedures used in nowcasting.

estimations of the risks around the central scenario. González-Rivera and Ruiz (2019), Chen et al. (2021) or Adrian et al. (2019) are some recent examples of the use of the lower quantiles of GDP distributions to gain insight on vulnerabilities. Instead of targeting annual GDP, we focus on the quantiles of estimated monthly regional factors to contribute to the assessment of real-time vulnerabily in different dimensions. First, we check the sensitivity of the vulnerability response to business cycle conditions by estimating a quantile regression with the regional factor as the variable to be explained against its own past. In this regard, we confirm previous results in Adrian et al. (2019) on the evolution of the distribution that is mostly symmetric in time but left-skewed under recession conditions. Second, with the same quantile regression, we compare the vulnerability across regions, finding similar resilience in normal times but significant heterogeneity during recession episodes, depending on productive structures and the causes underlying the crisis. And third, on the basis of alternative regional quantile regressions for each variable in the factor against the lagged common factor, we also obtain the vulnerability of various economic indicators, which are proxies of economic sectors or production structures.

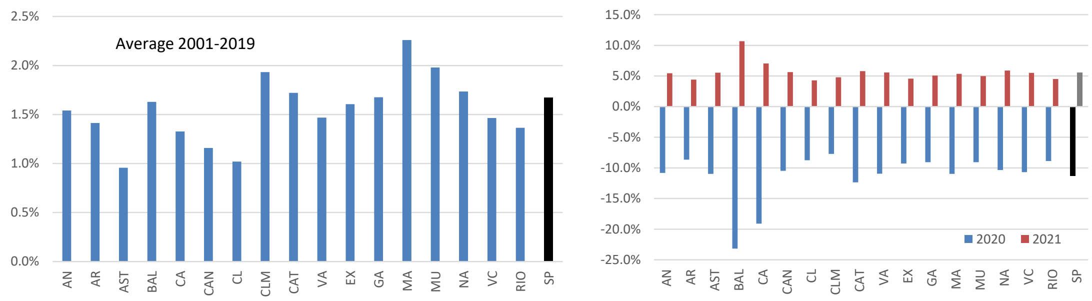
The rest of the paper is organized as follows. Section 2 describes the Spanish regional data. Section 3 presents the monthly dynamic factor model to track annual GDP. Section 4 estimates time, regional and sector vulnerability and Section 5 offers a conclusions.

## 2 Spanish Regional Data

Spain is the 4th largest country in Europe with 505,992 $km^2$ and is divided into 17 regions (NUTS-2 in the Eurostat nomenclature): ES61.Andalucía (AN), ES24.Aragón (AR), ES12.Asturias (AST), ES53.Balearic Islands (BAL), ES70.Canary Islands (CA), ES13.Cantabria (CAN), ES41.Castilla y León (CL), ES42.Castilla-La Mancha (CLM), ES51.Cataluña (CAT), ES52.Comunidad Valenciana (VA), ES43.Extremadura (EX), ES11.Galicia (GA), ES30.Comunidad de Madrid (MA), ES62.Región de Murcia (MU), ES22.Comunidad Foral de Navarra (NA), ES21.Basque Country (VC) and ES23.La Rioja (RIO).

The regions are geographically and climatologically diverse, with considerable differences between North and South, East, and West, coastal areas and inland. Also, two of the territories are insular regions, the Balearic Islands in the Mediterranean Sea, and the Canary Islands in the Atlantic Ocean. The Mediterranean regions and the islands are responsible for the Spain's leadership in world tourist competitiveness since 2015 (WEO, 2015, 2017, 2019). The inland regions base their economy on manufacturing, other services and agriculture. The left panel in Figure 1 shows the average annual growth rate of the chain-linked volume indices for the Spanish regions between 2001-2019 as published by the National Institute of Statistics. On average, Spain grew by 1.7% per year in that period, and growth throughout the regions ranged from 1% in Asturias and Castilla y León and 2.3% in Madrid.

**Figure 1**. Average annual GDP growth between 2001-2019, 2020 and 2021




The COVID-19 crisis severely affected the whole country, with GDP dropping 11.3% in 2020 with respect to 2019 (see the right panel in Figure 1). The impact was unequal, with the regions more exposed to tourism, namely the Balearic and Canary Islands, reporting the biggest downturns (-23.2 and -19.1%, respectively).

In 2021, Spain and its regions exhibited a prompt economic recovery. Spain's GDP grew by 5.5% in 2021 with respect to 2020 (see Figure 1), while the Balearic and Canary Islands reported the strongest growth rates in 2021 (10.7% and 7%, respectively).

## 2.1 Regional Indicators

Regional Spanish National Statistics are only available on an annual basis[2] and with a delay of 12 months after the end of the reference year. Regional nowcasting has been conducted mainly by means of dynamic factor models. In this regards, Cuevas and Quillis (2015) set up the nowcasting methodology behind the quarterly regional estimates of the Independent Authority for Fiscal Responsibility (AIReF). Their methodology combines time series models with benchmarking methods to process short-term monthly and quarterly indicators in order to estimate quarterly regional GDPs ensuring their temporal and transversal consistency with the National Accounts data. Gil et al. (2019) introduce the Bayesian Factor model for the Spanish regions (BayFaR), including as one of the inputs the AIReF regional forecasts.

[2]There are several regional agencies in Spain that produce their own quarterly figures for GDP: Andalucía, Aragón, Canary Islands, Cantabria, Castilla León, Catalonia, Extremadura, Galicia, Madrid, Navarre and the Basque Country. However, they are produced in an independent way with different methodologies, sources, time coverage, and do not need to show any consistency with national figures.

Based on the previous analyses and the structural characteristics of the Spanish regions, we have selected a set of nine monthly indicators representing of different economic sectors to build a factor model for each region. For comparative purposes, we use the same set of indicators for each region although the weight given to each indicator in the common factor points out the different structural characteristics across regions. We also use annual GDP chain-linked volume growth rates and show that the monthly common factor follows the annual regional figures. These indices are provided by the Spanish Regional Accounts (SRA) and are available for the time span 2000–2021. Subsequently, we use the regional monthly indicators to assess the vulnerability of growth in each region.

The set of nine monthly indicators covers the sample 2004:01 to 2022:11 and comprises the following: Overnight Stays in all kind of accommodation (OVERNIGHT), Car Registrations (CAR), Passenger Traffic at Spanish Airports (AIR PASSENGER), Number of affiliations to the Spanish Social Security System (EMPLOYMENT), Apparent Cement Consumption (CEMENT), Petroleum Product Consumption (PETROLEUM), Services Sector Activity Indicator (SERVICES), Retail Trade Index (RETAIL), and Industrial Production Index (INDUSTRY). The variable selection process was carried out under the premise that the indicators must be available in a timely and homogeneous manner and must provide a synthetic measure of the economy in each autonomous region. These indicators are used to track the annual GDP growth figures. In Appendix A and B, we provide more information on the selected indicators and their graphs transformed into index numbers with 2004=100.

The COVID-19 pandemic had a striking effect on all indicators, requiring some pretreatment of the data. First, trend and seasonality were greatly hit by the effects of the lockdown and seasonal adjustment, a tool frequently used in real-time monitoring of the state of the economy, became a difficult task generating considerable distress regarding the reliability of alternative estimations, see Bógalo et al. (2022). To avoid deseasonalizing we have considered the log annual difference of the indicators. Second, the initial drops as a consequence of strict lockdown measures, that even sent some indicators to 0 or close to 0, where followed by enormous increases as a rebound effect. To deal with extreme observations, we have identified an observation as an outlier if the log annual difference of the indicator at that moment deviates $+/-$ 5 standard deviations in the transformation of the variable.[3] Outliers are mainly concentrated in April 2020 and 2021 and, to a lesser extent, in May 2020 and only in 3 series in May 2021. Regarding variables, OVERNIGHT, CAR and AIR PASSENGER suffer the most extreme observations, followed by PETROLEUM, RETAIL and SERVICES. CEMENT shows no outliers and very few EMPLOYMENT and INDUSTRY very few.

[3]We define this value considering the extraordinary event associated with COVID-19 and previous works. In this sense, Stock and Watson (2002) considered an observation that exceeding 10 times the interquartile range from the median as outlier. Lewis et al. (2022) define outliers as observations for which the magnitude of the first difference is greater than three scaled median absolute deviations.

In a second step, we treat the identified outliers as missing values and impute them as in Cahan et al. (2022). Based on factor models, the procedure imputes the missing observation by using the information from the time and cross-section dimensions in the data set.

Figure 2 shows the indicators using the sample until 2019. The indicators in 2020-2021 are shown in a different Figure owing to their very different scale (see Figure 3).

**Figure 2**. Evolution of the log annual difference of the regional monthly economic indicators pre-COVID-19

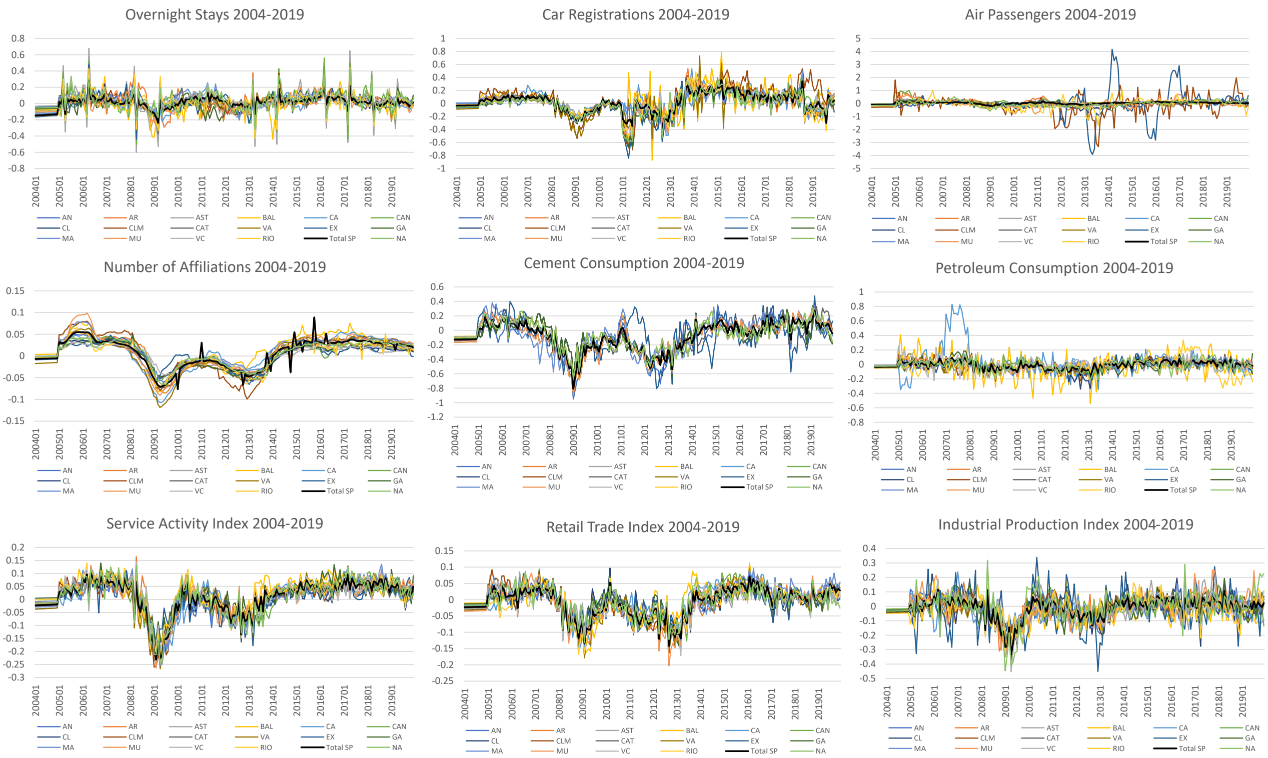
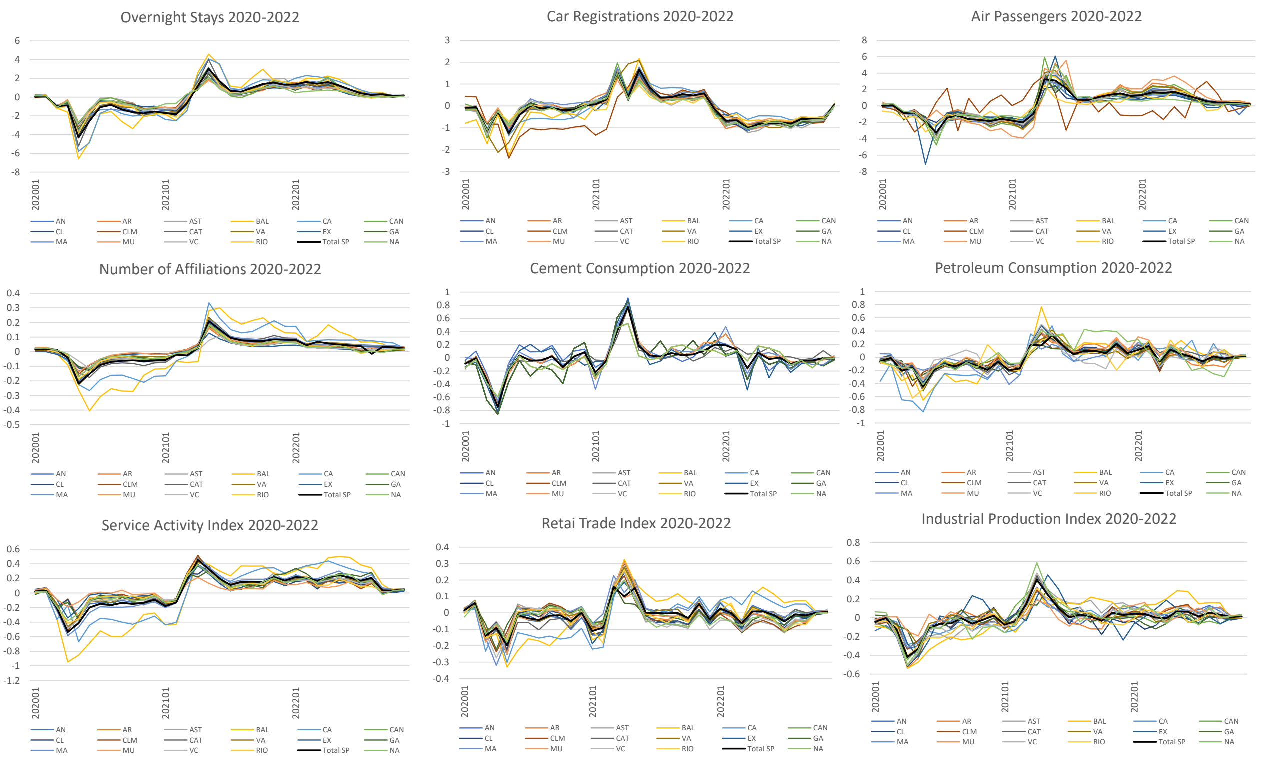


**Figure 3**. Evolution of the log annual difference of the regional monthly economic indicators post-COVID-19




## 3 Monthly dynamic factor model to track annual GDP

We build a monthly index of economic activity for each region using dynamic factor models. The model for region $i, i = 1, \cdots, G$ and time $t, t = 2004:01, ..., 2022:11$, is as follows: we assume that a vector of $N$ stationary economic indicators, $\mathbf{y}_t^i = (y_{1,t}^i, ..., y_{N,t}^i)'$, moves contemporaneously with the overall economic conditions. We also assume that the fluctuations for each variable can be broken down as the sum of two components. The first of these, the common component, accounts for the co-movements within each region. The second component is the $N \times 1$ time series vector $\mathbf{u}_t^i$, that represents the idiosyncratic movements in the series within each region. As our aim is to look for the specific vulnerabilities within each region, we built a separate dynamic factor model for each region. We also built a factor model for the Spanish aggregate. This suggests the following formulation for the set of indicators for each region:

$$\mathbf{y}_t^i \quad = \quad \mathbf{\Lambda}^i \quad f_t^i \quad + \quad \mathbf{u}_t^i, \tag{1}$$

where $\mathbf{\Lambda}^i = (\lambda_1^i, \lambda_2^i, ..., \lambda_N^i)'$ is the vector of factor loadings. The common factor follows an autoregression of order $q_i$ for each region,

$$f_t^i = \phi_1^i f_{t-1}^i + ... + \phi_{q_j}^i f_{t-q_j}^i + a_t^i, \tag{2}$$

where $a_t^i$ is assumed to be white noise $(0, (\sigma_a^2)^i)$. We also assume that each component of the specific or idiosyncratic error term follows an autoregression of order $P_j^i, j = 1, ..., N, i = 1, ..., G$,

$$u_{j,t}^i = \psi_{\iota,1}^i u_{t-1}^i + ... + \psi_{\iota,p_j}^i u_{t-p_j}^i + \epsilon_{j,t}^i, \tag{3}$$

being $\left\{\epsilon_{j,t}^i\right\}$ a white noise process with variance $(\sigma_j^2)^i$. In matrix form, we can write the equation for the idiosyncratic or specific error as

$$\mathbf{u}_t^i = \mathbf{\Psi}_1^i \mathbf{u}_t^i + ... + \mathbf{\Psi}_P^i \mathbf{u}_{t-P}^i + \epsilon_t^i, \tag{4}$$

being $\mathbf{\Psi}_j^i = diag(\psi_{1,j}^i, ..., \psi_{N,j}^i)$, $P_i = \max(p_1^i, ..., p_N^i)$ and $\text{var}(\epsilon_t^i) = diag((\sigma_1^2)^i, ..., (\sigma_N^2)^i)$. As stated the model is not identified as we can multiply the unobserved common factor and factor loadings by a non-zero constant $H$ and its inverse $H^{-1}$ to obtain the same observed time series in terms of a new set of loadings $\mathbf{\Lambda}^{i*}$ and common factor $f_t^{i*}$,

$$\mathbf{y}_t^i = \mathbf{\Lambda}^i H^{-1} H f_t^i + \mathbf{u}_t^i = \mathbf{\Lambda}^{i*} f_t^{i*} + \mathbf{u}_t^i \tag{5}$$

where $f_t^{i*} = Hf_t$ and $\mathbf{\Lambda}^{i*} = \mathbf{\Lambda}^i H^{-1}$. Therefore, we will need to impose identification restrictions in order to estimate the model.[4]

### 3.1 Estimation

We standardized the data and use several methods to estimate the model: principal components and the Kalman filter and smoother with and without estimating the idiosyncratic components. See Ruiz and Poncela (2022) for the comparison of the three procedures and the identification conditions imposed for each estimation procedure. We compute a first estimation of the factor model by principal components. As the number of series per region $N = 9$ is not large, we just use the common factors extracted by principal components to estimate by least squares the autoregressive parameters that will govern the dynamics of the common factors. The autocorrelation analysis of the first principal

[4]In case of models with $r$ common factors, $H$ is any $r \times r$ non-singular matrix and $r^2$ identification restrictions are needed. See Ruiz and Poncela (2022) for different identification restrictions according to the estimation procedure.

component for each region and the national aggregate indicates that $q_i = 3$ is appropriate for all data sets. The initial conditions for the state vector in order to run the Kalman filter are also derived from the pre-estimation of the common factor $f_0^i = (v^i)'\mathbf{y}_0^i$ where $v^i$ is the first eigenvector for the variance-covariance matrix of the variables for region $i, i = 1, ..., G$. The initial variance is given by the unconditional variance of an autoregressive process of order $q_i = 3$. The first eigenvector $v^i$ is also used to estimate the loadings for each region $i = 1, ..., G$. The specific components are estimated by the difference $\hat{\mathbf{u}}_t^i = \mathbf{y}_t^i - v^i(v^i)'\mathbf{y}_t^i$. Even though the small number of time series would preclude us from estimating the model with techniques that are only asymptotically appropriate, we have checked that there is high correlation between the common factors estimated with alternative methods. As Poncela et al. (2020) find for commodity prices, better out-of-sample forecasting results are obtained when using common factors without modeling and forecasting the idiosyncratic component. Therefore, from now on we report the results with the estimated factor from the Kalman filter without modeling the idiosyncratic component.

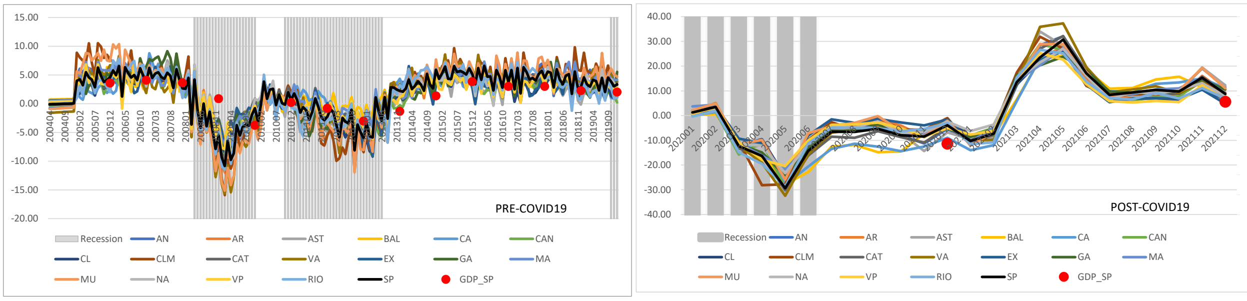
### 3.2 Regional Factors

The first component is, as usual, a weighted average of the variables. The different weights of the variables in the regions are associated to the differences in their productive structure.

As our goal is to use the common factor as a proxy of monthly GDP growth rate for each region, we need to express it in meaningful terms. To do so, we transform the factor of each region into units of GDP. In this sense, we define $\tilde{f}_t^i = f_t^i * \sigma_{GDP_i} + \mu_{GDP_i}$, where $\sigma_{GDP_i}$ and $\mu_{GDP_i}$ are the sample standard deviation and the sample mean of the GDP growth rate of region $i$, respectively. Figure 4 shows the estimated factors for each of the regions as well as for the whole of Spain (the bold black line). The shadowed areas are the quarterly dating of the Spanish recessions (Association, 2021) and the red dots are the growth rate of Spanish GDP. To avoid problems arising from the scale of the data, the panel on the left shows pre-COVID-19 (2004-2019) data and the panel on the right the post-COVID-19 (2020-2022) data[5].

[5]Figures C.1 and C.2 in the Appendix C show more detailed information on the regional factors in units of their GDP $\tilde{f}_t^i$, as well as the annual GDP growth, in the period previous and posterior to COVID-19, respectively.

**Figure 4**. Common factor in each region and observed Spanish GDP growth




Despite the volatility of monthly figures, the factor for the different regions is in line with the evolution of the annual GDP growth, both pre- and post COVID-19. Therefore, the monthly factor serves as an indicator to assess the state of the economy in real time and likewise provides relevant information about the timing of the crisis. Unlike other crisis, COVID-19 is a global recession episode for which the starting point is dated in March 2020 in Spain. The lockdown ceased activities in many sectors but was most devastating in tourism, one of the main sources of income in many regions of Spain, at the precise moment when the high season was to begin. Regional factor estimations in March 2020, plummeted to negative values between -9% and -15% and in two months these negative values doubled. GDP growth in 2020 ranged from a decrease of 7.69% in Castilla-La Mancha and around around 20% in the islands regions (BA and CA). A similar pattern was obtained with the estimation of the factor in December 2020. The recovery has been unequal, due to the regions' different productive structure, but also to the variety of confinement measures to restrain the pandemic (Ghirelli et al., 2021).

## 4 Vulnerable Growth: Time, Regional and Sector Skewness

In this section we go beyond the point forecasts provided by the factor model in section 3 and analyze the full distribution to study the characteristics of the different quantiles and assess vulnerability from three alternative view points. The first examines the evolution of the estimated distributions of the regional factors over time, finding that in line with the literature, recessions are associated with left-skewed distributions while, during expansions, the conditional distribution is closer to being symmetric. The second approach, compares vulnerability across regions finding that its magnitude depends on the causes behind the crisis. And the third, reports on sector vulnerability by region, linking the evolution of each variable with the economic activity in the region.

### 4.1 Quantile Regression

To set the framework for quantile regression, let us call the dependent variable $y_t^i$ and the vector of explanatory variables $x_{t-1}^i$ in the regression for region $i$, including a constant and their lags. In a quantile regression of $y_t^i$ on $x_{t-1}^i$, the regression slope $\beta_\tau$ is choosen to minimize the quantile weighted absolute value of the residuals:

$$\hat{\beta}_\tau = \arg\min_{\beta_\tau \in \mathbb{R}^k} \sum_{t=1}^{T} (\tau \cdot \mathbb{1}_{(y_t^i \geq x_{t-1}^i \beta)} \, |y_t^i - x_{t-1}^i \beta| + (1-\tau) \cdot \mathbb{1}_{(y_t^i < x_{t-1}^i \beta)} \, |y_t^i - x_{t-1}^i \beta|) \tag{6}$$

where $\mathbb{1}_{(\cdot)}$ expresses the indicator function. To assess the time and regional vulnerability we explained $y_t^i$ in terms of its own past by means of an AR(p) model as in the previous section. The predicted value from that regression, which exploits only the the available information, is the quantile of $y_t^i$ conditional on $x_{t-1}^i$,

$$\hat{Q}_{y_t^i | x_{t-1}^i}(\tau | x_{t-1}^i) \; = \; x_{t-1}^i \hat{\beta}_\tau \tag{7}$$

Once we have $\hat{Q}_{(\cdot)}$ for $\tau \in \{0.05, ..., 0.95\}$, for each date between 01/2010 and 11/2022, we assess vulnerability in different dimensions, across time, regions and sectors.

### 4.2 Vulnerability Across Time

The first analysis involves describing the vulnerability over time. To this end, we apply the equations (6) and (7) with $y_t^i = f_t^i$ in terms of its own past. We exploit the predicted value from the quantile regression $\hat{Q}_{(\cdot)}$, as mentioned above in subsection 4.1. We smooth the estimated quantile distribution every month by estimating a kernel density. This allows us to transform the empirical quantile distribution into an estimated conditional distribution of the regional factors. Figure 5 and 6 show the results. In line with Adrian et al. (2019), we evidence not only that the entire distribution evolves over time but also that the median and the left tail of the distribution exhibit the strongest time series variation. During the second quarter of 2020, the worst period of the pandemic, the entire conditional distribution shifted to the left, which reflects the higher level of vulnerability and uncertainty during this period.

**Figure 5**. Estimated conditional distribution of the regional factors

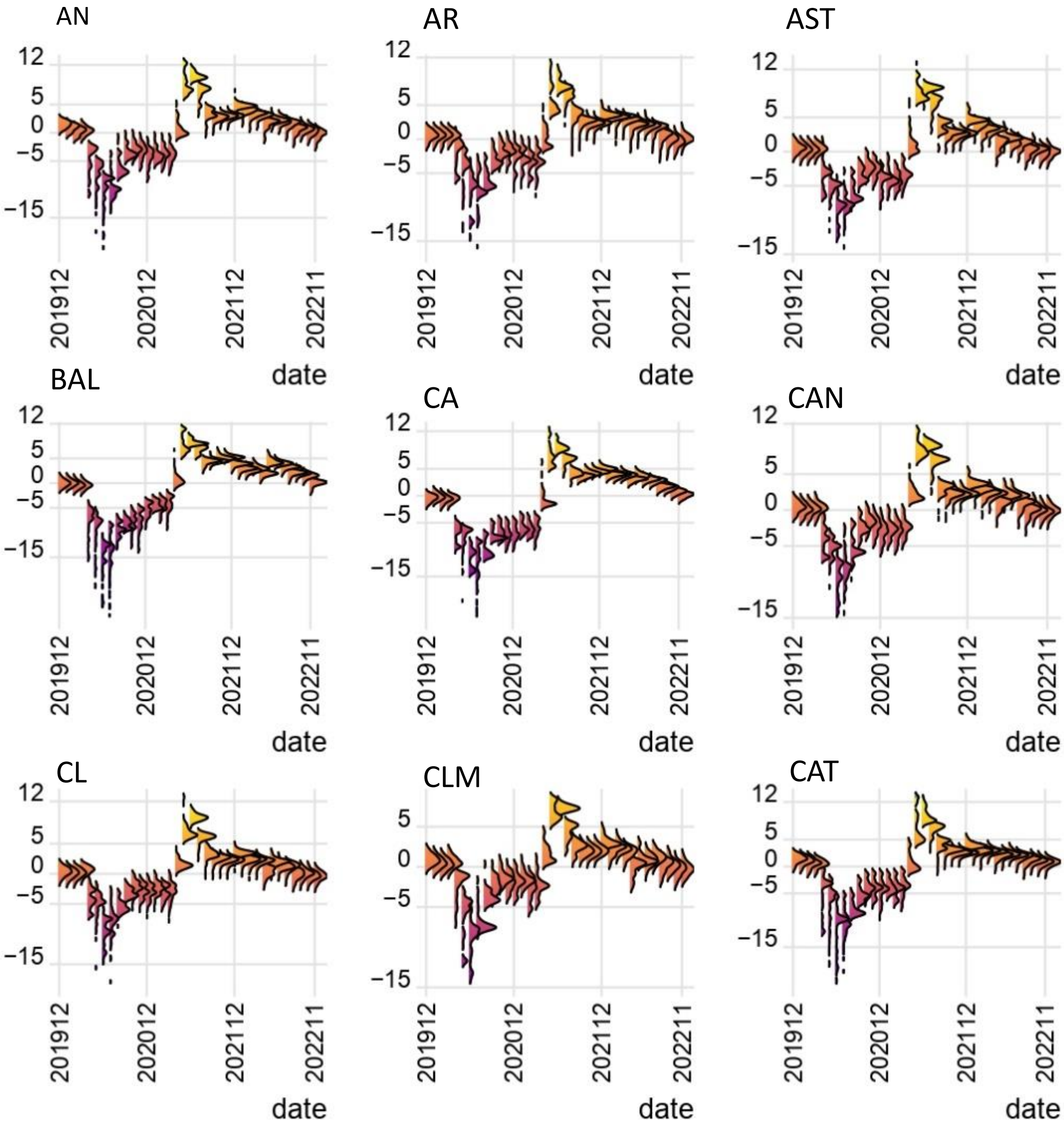
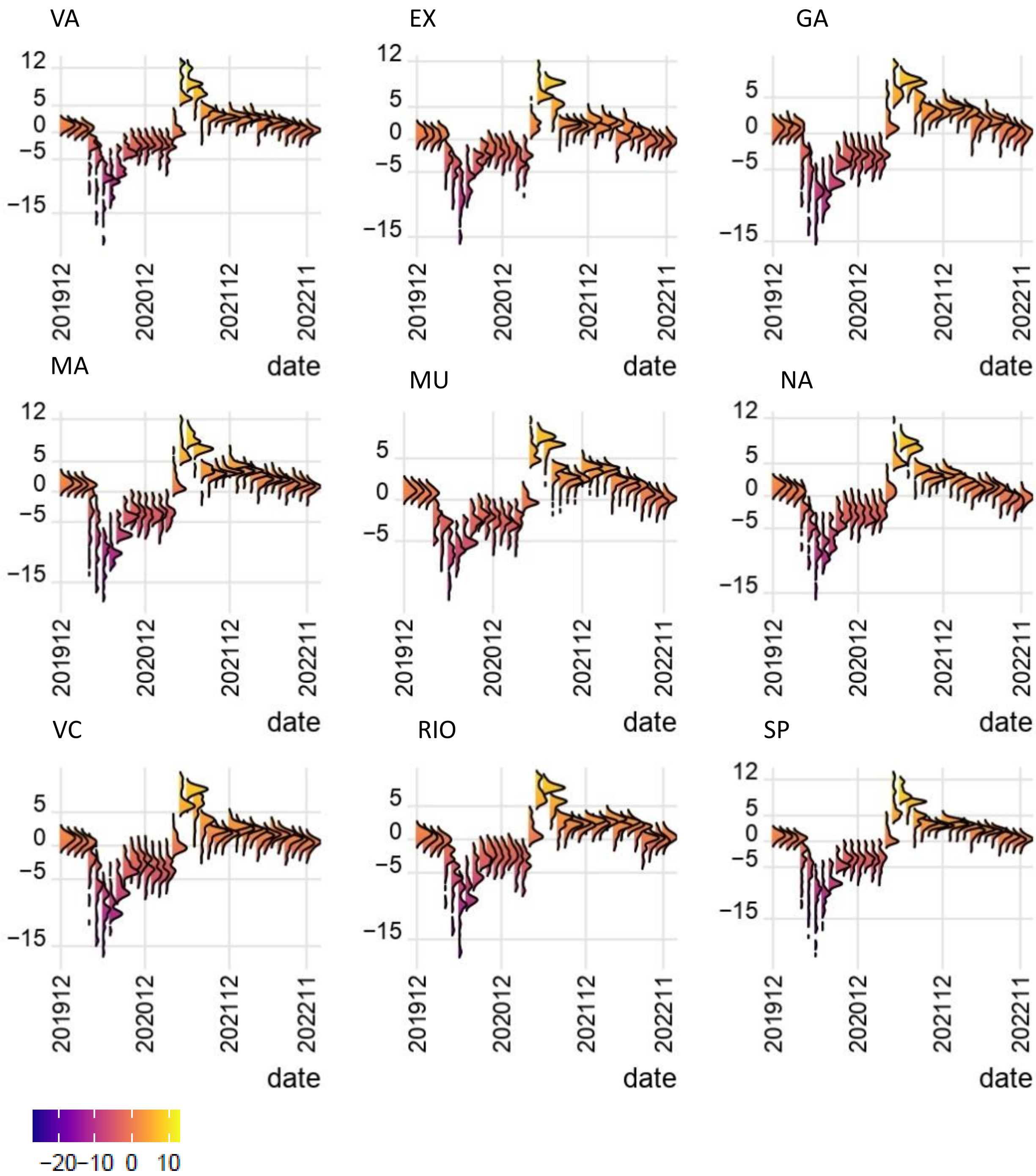


**Figure 6**. Estimated conditional distribution of the regional factors (cont.)




One stylized fact of vulnerability is that recessions are associated with left-skewed distributions, while in expansions the conditional distribution is closer to being symmetrical, see for instance Adrian et al. (2019). In fact, this can be observed in the plotted densities in Figure 6, in the moments when the exposure to COVID-19 was greatest in year 2020.

### 4.3 Regional Vulnerability

The second analysis compares vulnerability across regions.

Although the regional common factor estimated in section 3 provides information about the magnitude of the pandemic's effect on regional economies, analysis of the lower distribution quantiles identifies the worst case scenarios. In this sense, higher vulnerability is identified by larger distances between the median and one of the lowest percentiles. In this section, we consider as a measure of vulnerability the distance between the median and percentile 5 ($Q_{0.50} - Q_{0.05}$). The biggest this distance, the more vulnerable the region.

To better discriminate the regions, we performed a hierarchical cluster analysis on the data for post-COVID period, that is, December 2019 to November 2022. The hierarchical clustering can be visualized using a dendogram whose y-axis depicts the distances between objects. To compute distances we used the Euclidean norm. Figure 7 shows the dendogram that classifies Spanish regions by their vulnerability. It can be seen that there are two well-differentiated groups. One of these includes the most vulnerable regions during the time of COVID-19 due to the strong impact of the lockdown on tourism and the slow return to normality. This first group includes the two insular regions (BA and CA) and also most of the regions on the Mediterranean coast (CAT, VA, AN). The second group was less affected by COVID-19 due to the smaller weight of tourism in these regions. Figure 8 locates on the map both groups of regions.

**Figure 7**. Cross-section vulnerability. Dendogram

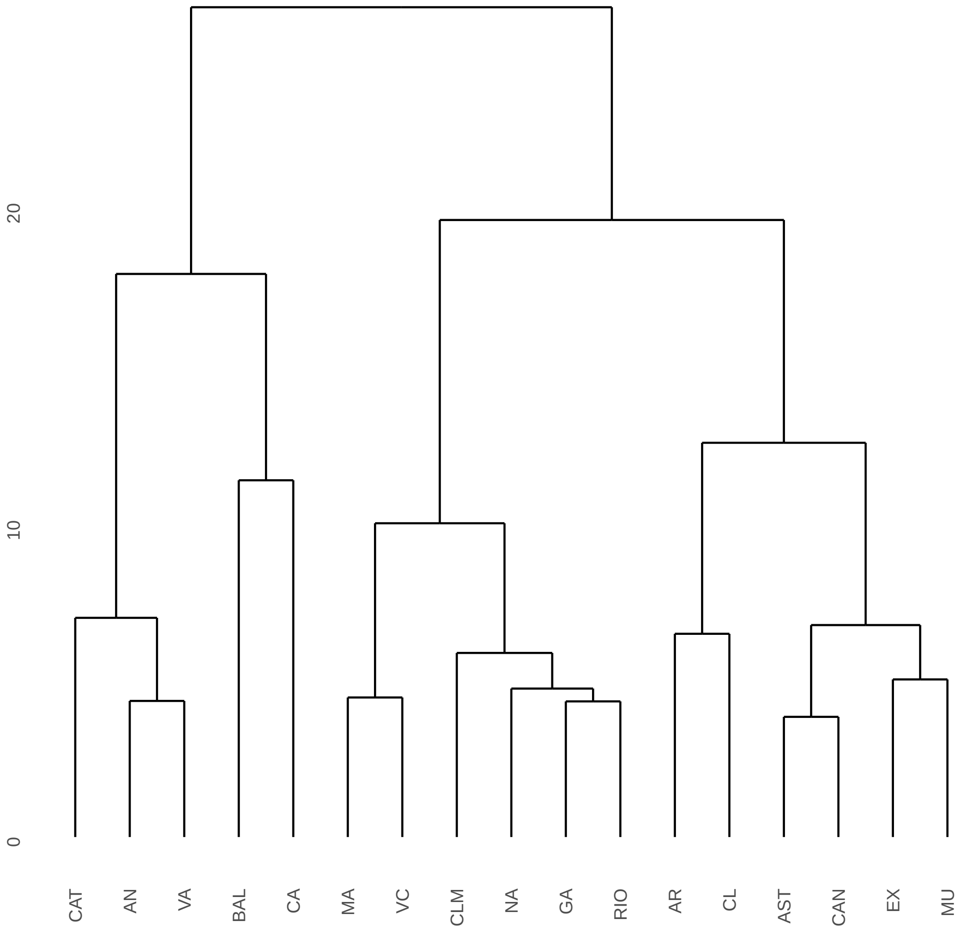
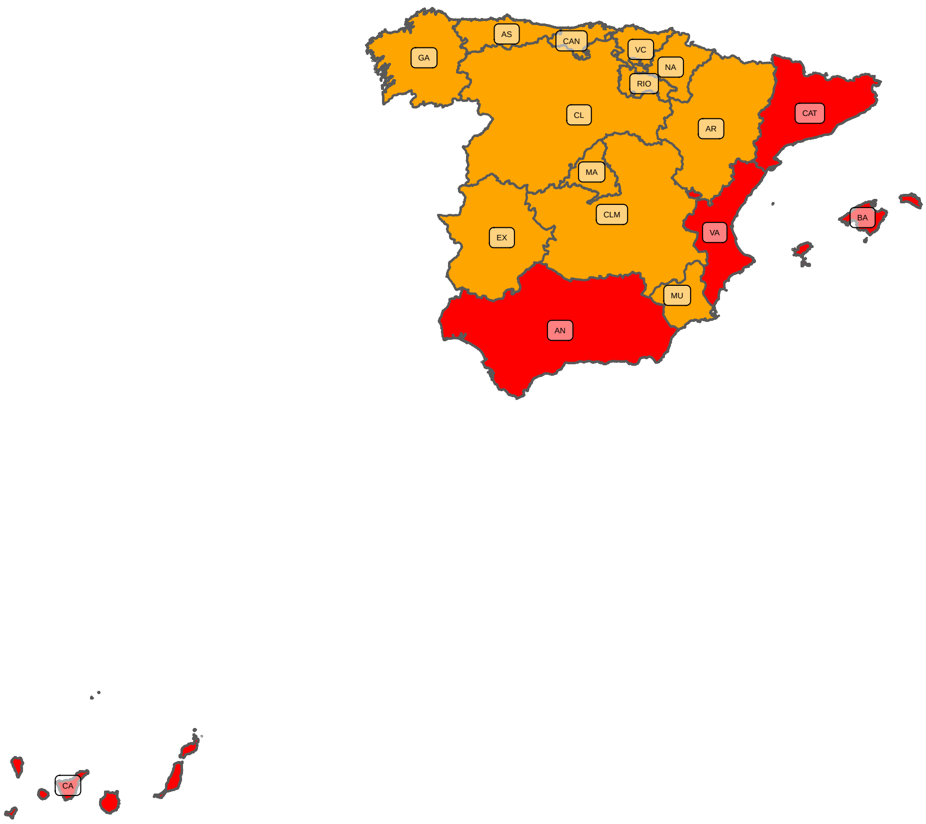


**Figure 8**. Cross-section vulnerability. Map

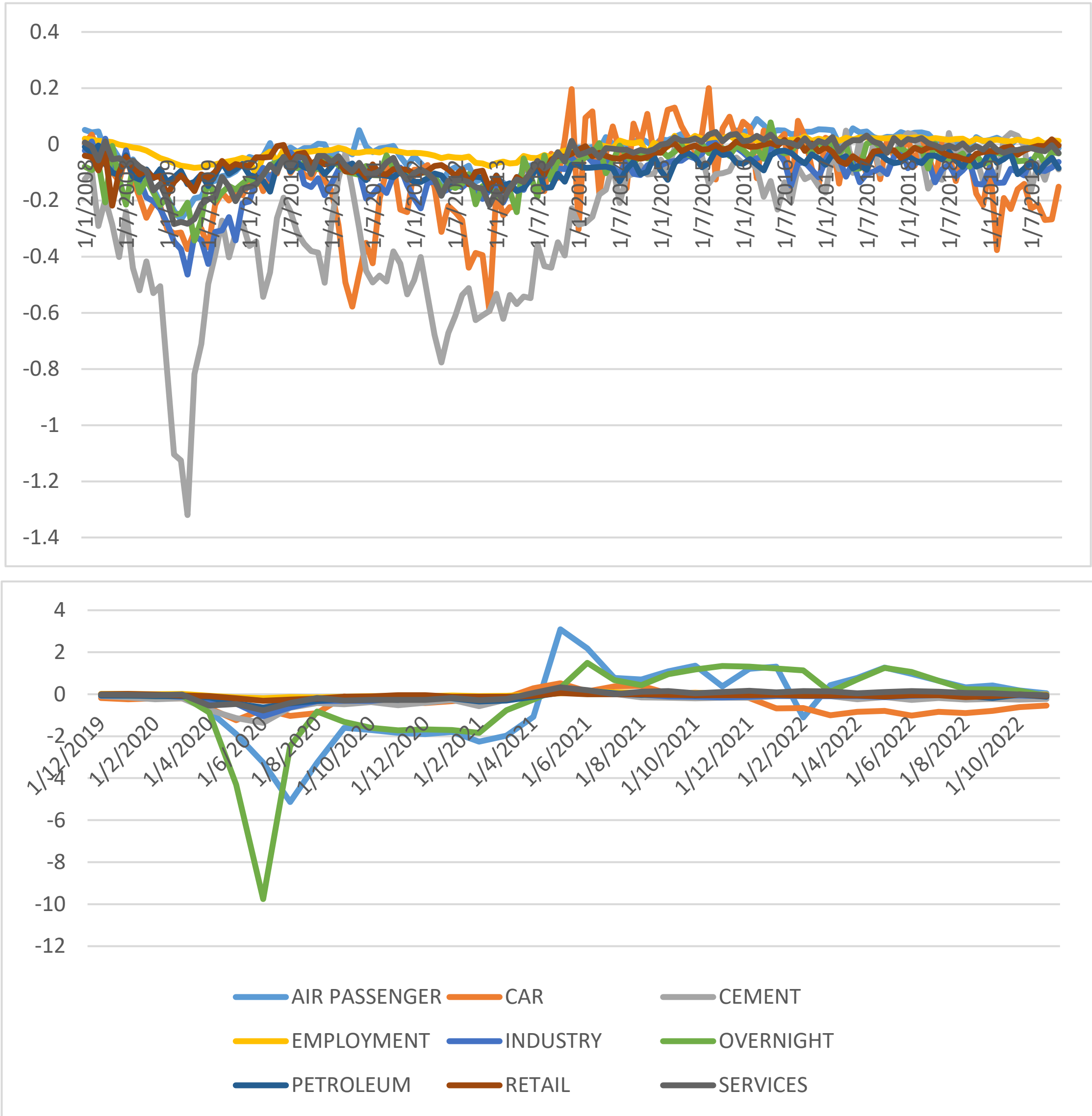



### 4.4 Sector Vulnerability

To check the vulnerability of sectors to the different nature of the crisis, we take the economic indicators used to build the common factor as their proxies. We estimated a quantile regression as in Subsection 4.1 between each economic indicator and its lags and the lags of the common factor. In order to illustrate the results, we present the outcome for Spain as a whole. Once we have the predicted value from the quantile regression, we estimate the conditional distribution of each specific variable. This analysis allows us to estimate the effect of changes in economic conditions on each quantile of a specific variable. In this case, the 5th percentile will show the vulnerability of the sector represented by the variable to the downside risks in the general economic conditions given by the common factor.

Figure 9 shows the estimated results for each variable at two moments in time corresponding to crisis of a very different nature. In this regard, during the financial crisis in the period 2010-2013, we can see that CEMENT, CAR REGISTRATION and INDUSTRY appeared as the most affected, highlighting the highest vulnerablity of the construction and manufacturing sectors. By contrast, during the COVID-19 recession, AIR PASSENGER and OVERNIGHT were the variables that exhibited the biggest falls due to lockdowns and their effect on global movements worldwide that mainly paralyzed the tourism sector.

**Figure 9**. Sectorial vulnerability: cross-variable and time-variant



## 5 Conclusions

Real-time assessment of the vulnerability of a region is highly relevant to correctly applying economic policy. Heterogeneity in the structural composition of the productive sector causes different regions' responses to the crisis to be different.

In this paper we have used regional factor models and quantile regressions to assess on three dimensions of vulnerability. First we have checked that the vulnerability, defined as a greater probability of the left tail of the distribution, is more pronounced in times of crisis. Second, the degree of vulnerability of the regions, defined as the biggest probability of being below the median, also shows great heterogeneity. The differences are always related to their exposure to the nature of the crisis. Finally, sectors are also heterogeneous and respond differently to the nature of the crisis.

From an economic policy point of view the paper clearly states the need to go beyond national aggregates and attend to time, regional and sector analysis. When exiting a crisis, policy makers need to have the full picture of the economy in real time. To have the monthly real-time regional distribution of the evolution of the economy is a key tool to gaining awareness of the increase in vulnerability, its causes and the areas that need to be supported.

## Acknowledgements

This paper was written while the second author was visiting the Instituto Universitario de Investigación en Estudios Latinoamericanos (IELAT) of the University of Alcalá (Spain). He thanks the financial support received from Santander Bank and University of Alcalá. The remaining authors also thanks the financial support provided by the Spanish Ministry of Science and Innovation, contract grants: MINECO/FEDER PID2019-107161GB-C32 and PID2019-108079GB-C22 and the Spanish Ministry of Economic Affairs and Digital Transformation.

## References

Adrian, T., Boyarchenko, N., and Giannone, D. (2019). Vulnerable growth. *American Economic Review*, 109(4):1263–89.

Association, S. E. (2021). Chronology of the spanish business cycle.

Bandrés, E. and Gadea, M. D. (2013). Crisis económica y ciclos regionales en españa. *Papeles de Economía Española*, 138.

Bandrés, E., Gadea Rivas, M. D., and Gómez Loscos, A. (2021). Dating and synchronisation of regional business cycles in spain. Documentos Ocasionales 2107, Banco de España.

Bańbura, M., Giannone, D., Modugno, M., and Reichlin, L. (2013). Chapter 4 - now-casting and the real-time data flow. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2 of *Handbook of Economic Forecasting*, pages 195–237. Elsevier.

Beine, M. and Coulombe, S. (2003). Regional perspectives on dollarization in canada. *Journal of Regional Science*, 43(3):541–570.

Bógalo, J., Llada, M., P. P., and Senra, E. (2022). Seasonality in covid-19 times. 211:110206.

Cahan, E., Bai, J., and Ng, S. (in press 2022). Factor-based imputation of missing values and covariances in panel data of large dimensions. *Journal of Econometrics*.

Camacho, M., Pacce, M., and Ulloa, C. (2019). Regional business cycle phases in spain. *Estudios de Economía Aplicada*, 36(3):875–896.

Camacho, M., Perez-Quiros, G., and Poncela, P. (2013). Short-term forecasting for empirical economists: A survey of the recently proposed algorithms. *Foundations and Trends® in Econometrics*, 6(2):101–161.

Carlino, G. and DeFina, R. (1998). The differential regional effects of monetary policy. *Review of Economics and Statistics*, 80(4):572–587.

Chamie, N., DeSerres, A., and Lalonde, R. (1994). Optimum currency areas and shock asymmetry: a comparison of europe and the united states. Working paper.

Chen, L., Dolado, J. J., and Gonzalo, J. (2021). Quantile factor models. *Econometrica*, 89(2):875–910.

Chernis, T., Cheung, C., and Velasco, G. (2020). A three-frequency dynamic factor model for nowcasting Canadian provincial GDP growth. *International Journal of Forecasting*, 36:851–872.

Claudio, J. C., Heinisch, K., and Holtemöller, O. (2020). Nowcasting east german gdp growth: a midas approach. *Empirical Economics*, 58(1):29–54.

Clayton, T. M. and Clayton-Matthews, A. (2005). Consistent economic indexes for the 50 states. *The Review of Economics and Statistics*, LXXXVII:593–603.

Cuevas, A. and Quillis, E. (2015). Quarterly regional gdp flash estimates for the spanish economy (metcap model). Working Paper 2015/3, AIReF.

Fratantoni, M. and Schuh, S. (2003). Monetary policy, housing, and heterogeneous regional markets. *Journal of Money, Credit and Banking*, 35(4):557–589.

Gadea, M. D., Gómez-Loscos, A., and Montañés, A. (2012). Cycles inside cycles. spanish regional aggregation. *SERIEs*, 3(4):423–456.

Ghirelli, C., Gil, M., Hurtado, S., and Urtasun, A. (2021). The relationship between pandemic containment measures, mobility and economic activity. Documentos Ocasionales 2109, Bank of Spain.

Gil, M., Leiva-Leon, D., Pérez, J. J., and Urtasun, A. (2019). An application of dynamic factor models to nowcast regional economic activity in spain. Documentos Ocasionales 1904, Bank of Spain.

González-Rivera, G., M. J. and Ruiz, E. (2019). Growth in stress. *International Journal of Forecasting*, 35(3):948–966.

Grant, A., Koop, G., McIntyre, S., and Smith, P. (2014). Nowcasting Scotish GDP Growth. RCEA Working Papers 41_14, Rimini Centre for Economic Analysis (RCEA).

Henzel, S. R., Lehmann, R., and Wohlrabe, K. (2015). Nowcasting regional gdp: The case of the free state of saxony. *Review of Economics*, 66(1):71–98.

Kopoin, A., Moran, K., and Paré, J. (2013). Forecasting regional GDP with factor models: How useful are national and international data? 121:267–270.

Lehmann, R. and Wohlrabe, K. (2014). Forecasting gross value-added at the regional level: Are sectoral disaggregated predictions superior to direct ones? *Rev Reg Res Jahrbuch für Regionalwissenschaft*, 34(1):61–90.

Lehmann, R. and Wohlrabe, K. (2015). Forecasting gdp at the regional level with many predictors. *German Economic Review*, 16(2):226–254.

Lehmann, R. and Wohlrabe, K. (2017). Boosting and regional economic forecasting: the case of germany. *Letters in Spatial and Resource Sciences*, 10(2):161–175.

Lewis, D. J., Mertens, K., Stock, J. H., and Trivedi, M. (2022). Measuring real activity using a weekly economic index. *Journal of Applied Econometrics*, 37:667–687.

Owyang, M. T., Piger, J., and Wall, H. J. (2005). Business cycle phases in u.s. states. *The Review of Economics and Statistics*, 87(4):604–616.

Poncela, P., Senra, E., and Sierra, L. P. (2020). Global vs sectoral factors and the impact of the financialization in commodity price changes. *Open Economies Review*, 31:859–879.

Ruiz, E. and Poncela, P. (in press 2022). Factor extraction in dynamic factor models: Kalman filter versus principal components. *Foundations and Trends in Econometrics.*

Stock, J. H. and Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162.

Villaverde, J. and Maza, A. (2009). The robustness of okun's law in spain, 1980–2004: Regional evidence. *Journal of Policy Modeling*, 31(2):289–297.

WEO (2015). The travel & tourism competitiveness report 2015. `http://www3.weforum.org/docs/TT15/WEF_Global_Travel&-Tourism_Report_2015.pdf`.

WEO (2017). The travel & tourism competitiveness report 2017. `http://www3.weforum.org/docs/WEF_TTCR_2017_web_0401.pdf`.

WEO (2019). The travel & tourism competitiveness report 2019. `http://www3.weforum.org/docs/WEF_TTCR_2019.pdf`.

# Appendix A Regional Indicators

- **GDP**: Chain-linked volume indices.
  - Units: Index number.
  - Source: Spanish Regional Accounts(SRA)[6].
  - Starting date: 2000.
- **AIR PASSENGER**: Passenger traffic in Spanish airports[7].
  - Units: Numbers of passenger in Spanish airports.
  - Source: Aena[8].
  - Starting date: 2004:01.
- **CEMENT CONSUMPTION**: Apparent cement consumption.

[6] `https://www.ine.es/dyngs//INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736167628&menu=resultados&idp=1254735576581#`

[7] Include: Adolfo Suárez Madrid-Barajas and Madrid-Cuatro Vientos airports

[8] `https://wwwssl.aena.es/csee/Satellite?c=Page&cid=1113582476721&pagename=Estadisticas%2FEstadisticas`

- Units: Tonnes.
- Source: General Secretariat of Industry and Small and Medium Enterprises[9].
- Starting date: 1992:01.

- **PETROLEUM CONSUMPTION**: Petroleum Product Consumption.
  - Units: Thousands of tons.
  - Source: CORES[10].
  - Starting date: 1992:01.

- **EMPLOYMENT**: Affiliated to the Spanish Social Security System.
  - Units: Average affiliated.
  - Source: General Treasury of the Social Security[11].
  - Starting date: 2001:01.

- **OVERNIGHT**: Overnight stays (all kind of accommodation[12]).
  - Units: Number of overnight stays.
  - Source: Hotel Occupancy Survey[13], National Statistical Institute (Instituto Nacional de Estadística, INE).
  - Starting date: 2001:01.

- **SERVICES**: Service Activity Index.
  - Units: Index number.
  - Source: National Statistical Institute (Instituto Nacional de Estadística, INE)[14].
  - Starting date: 2005:01.

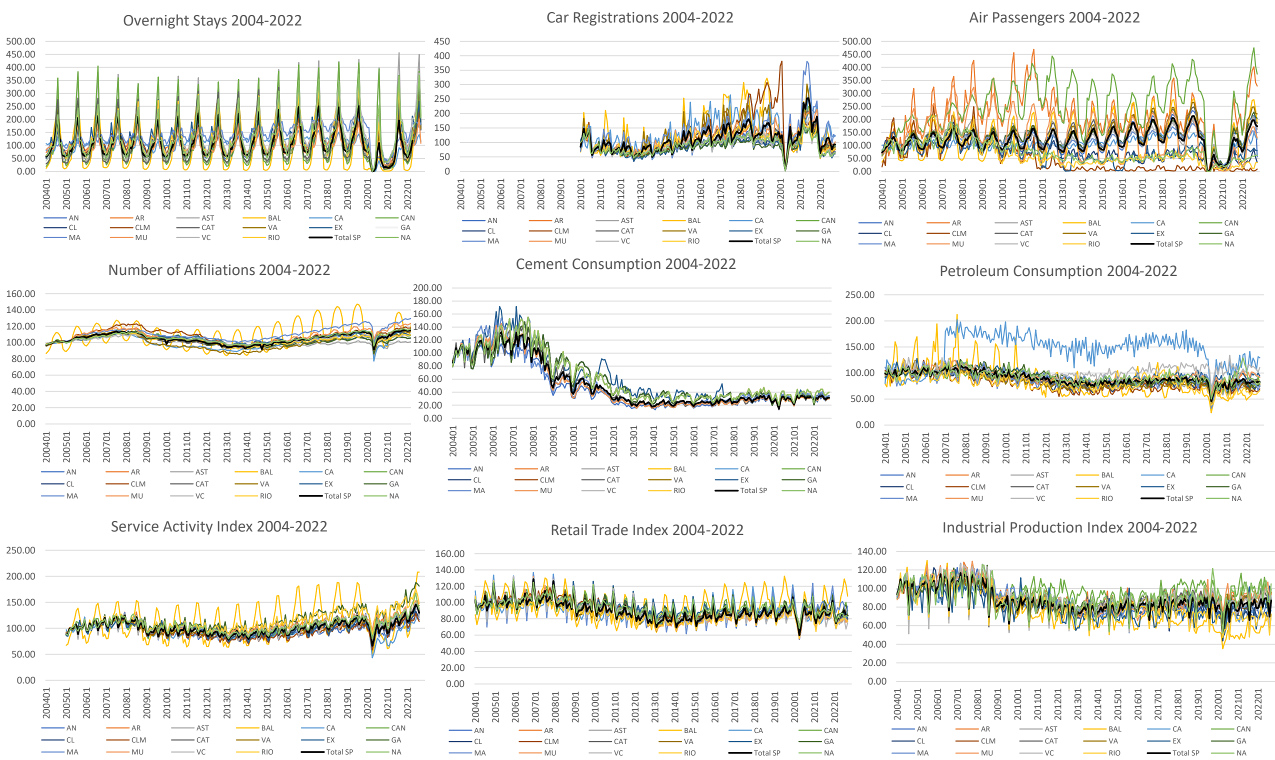
- **RETAIL**: Retail Trade Index.
  - Units: Index number.
  - Source: National Statistical Institute (Instituto Nacional de Estadística, INE)[15].
  - Starting date: 2002:01.

[9] https://industria.gob.es/es-es/estadisticas/Paginas/Estadistica-Cemento.aspx
[10] https://www.cores.es/en/estadisticas
[11] https://w6.seg-social.es/PXWeb/pxweb/es/Afiliados%20en%20alta%20laboral/
[12] Include: Hotel establishments, Holiday campsites, Holiday dwellings and Rural tourism accommodation.
[13] https://www.ine.es/jaxiT3/Datos.htm?t=2941
[14] https://www.ine.es/jaxiT3/Tabla.htm?t=25897
[15] https://www.ine.es/jaxiT3/Datos.htm?t=25992

- **INDUSTRY**: Industrial Production Index.
  - Units: Index number.
  - Source: National Statistical Institute (Instituto Nacional de Estadística, INE)[16].
  - Starting date: 2002:01.

# Appendix B Regional indicators graphs

**Figure B.1**. Regional monthly economic indicators, index number (2004) = 100)




[16] https://www.ine.es/jaxiT3/Tabla.htm?t=26061

# Appendix C Regional monthly factors

**Figure C.1**. Common factor in each region and observed GDP growth (2004-2019)

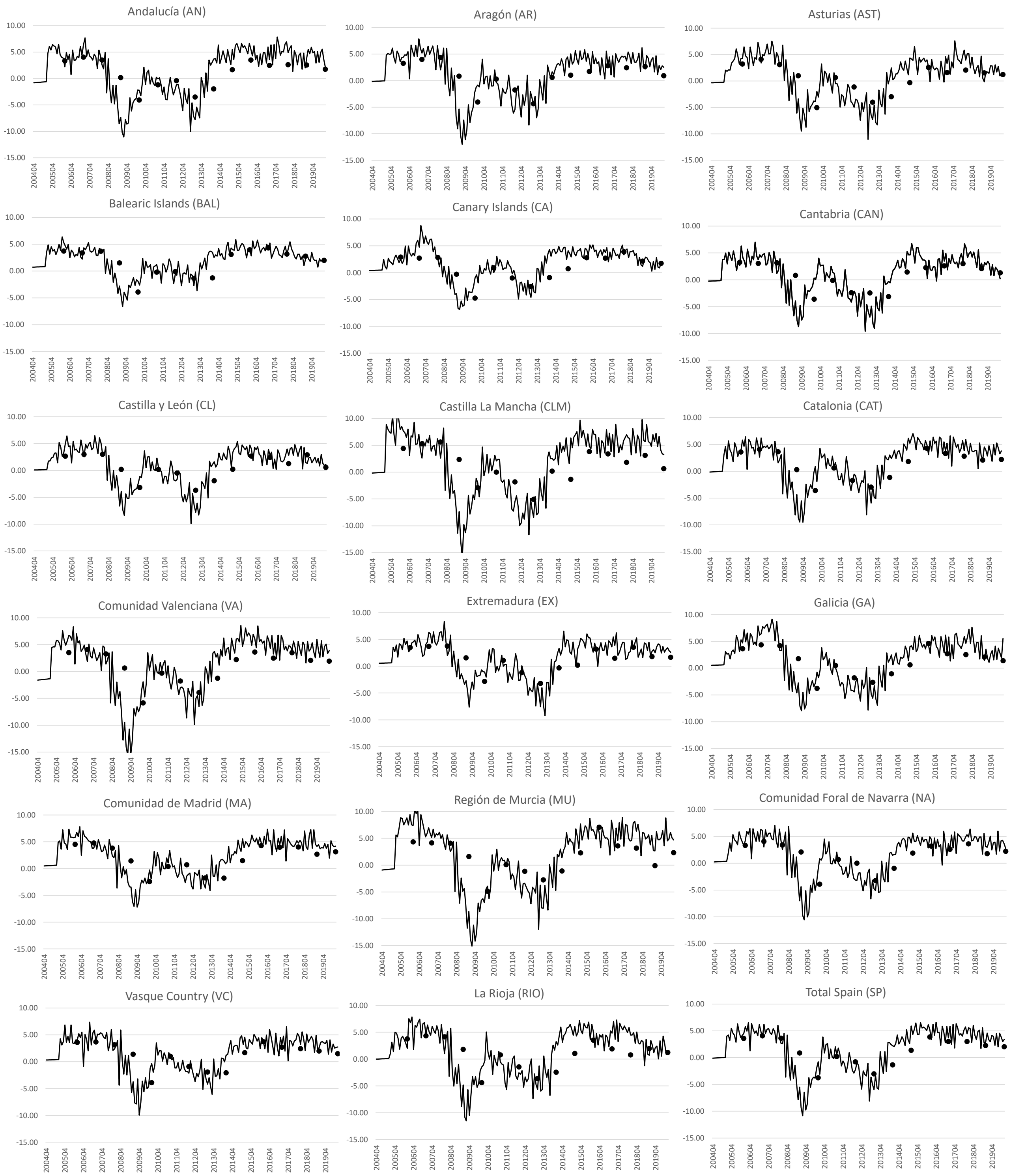


**Figure C.2**. Common factor in each region and observed GDP growth (2020-2022)