

AN AUTOMATIC ALGORITHM TO DATE
THE REFERENCE CYCLE OF THE
SPANISH ECONOMY

2021

BANCO DE **ESPAÑA**
Eurosistema

Documentos de Trabajo
N.º 2139

Máximo Camacho, María Dolores Gadea
and Ana Gómez Loscos

**AN AUTOMATIC ALGORITHM TO DATE THE REFERENCE CYCLE OF THE
SPANISH ECONOMY**

AN AUTOMATIC ALGORITHM TO DATE THE REFERENCE CYCLE OF THE SPANISH ECONOMY (*)

Maximo Camacho (**)

UNIVERSITY OF MURCIA

María Dolores Gadea (***)

UNIVERSITY OF ZARAGOZA

Ana Gómez Loscos (****)

BANCO DE ESPAÑA

(*) We thank seminar participants at Banco de España seminar. M. Camacho and M. D. Gadea are grateful for the support of grants PID2019-107192GB-I00 (AEI/10.13039/501100011033) and 19884/GERM/15, and ECO2017-83255-C3-1-P and ECO2017-83255-C3-3-P (MICINU, AEI/ERDF, EU), respectively. Data and codes that replicate our results are available from the authors' websites. The views in this paper are those of the authors and do not represent the views of the Banco de España, the Eurosystem, or the Spanish Business Cycle Dating Committee.

(**) Department of Quantitative Analysis, University of Murcia. Campus de Espinardo 30100 Murcia (Spain). Tel: +34 868 887 982 and e-mail: mcamacho@um.es.

(***) Department of Applied Economics, University of Zaragoza. Gran Vía, 4, 50005 Zaragoza (Spain). Tel: +34 976 761 842 and e-mail: lgadea@unizar.es.

(****) Banco de España, Alcalá, 48, 28014 Madrid (Spain). Tel: +34 91 338 58 17 and e-mail: agomezloscos@bde.es.

The Working Paper Series seeks to disseminate original research in economics and finance. All papers have been anonymously refereed. By publishing these papers, the Banco de España aims to contribute to economic analysis and, in particular, to knowledge of the Spanish economy and its international environment.

The opinions and analyses in the Working Paper Series are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

The Banco de España disseminates its main reports and most of its publications via the Internet at the following website: <http://www.bde.es>.

Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.

© BANCO DE ESPAÑA, Madrid, 2021

ISSN: 1579-8666 (on line)

Abstract

This paper provides an accurate chronology of the Spanish reference business cycle by adapting the multiple change-point model proposed by Camacho, Gadea and Gómez Loscos (2021). In that approach, each individual pair of specific peaks and troughs from a set of indicators is viewed as a realization of a mixture of an unspecified number of separate bivariate Gaussian distributions, whose different means are the reference turning points and whose transitions are governed by a restricted Markov chain. In the empirical application, seven recessions in the period from 1970.2 to 2020.2 are identified, which are in high concordance with the timing of the turning point dates established by the Spanish Business Cycle Dating Committee (SBCDC).

Keywords: business cycles, turning points, finite mixture models, Spain.

JEL classification: E32, C22, E27.

Resumen

Este trabajo proporciona una cronología precisa del ciclo económico de referencia de la economía española. Para ello, se adapta el procedimiento de fechado, que considera la posibilidad de múltiples puntos de cambio en la fase del ciclo, propuesto por Camacho, Gadea y Gómez Loscos (2021). En dicho enfoque, cada par individual de picos y valles específicos, obtenidos a partir de un conjunto de indicadores económicos, se considera como una realización de una mixtura de un número no especificado de diferentes distribuciones gaussianas bivariantes, cuyas medias son los puntos de cambio en la fase del ciclo de referencia y cuyas transiciones se rigen por una cadena de Markov restringida. En la aplicación empírica, se identifican siete recesiones en el período comprendido entre 1970.2 y 2020.2. Este fechado es muy similar al establecido por el Comité de Fechado del Ciclo Económico Español.

Palabras clave: ciclos económicos, puntos de cambio en la fase del ciclo, modelos de mixturas finitas de distribuciones, España.

Códigos JEL: E32, C22, E27.

1 Introduction

Although it seems a truism truth, the business cycle turning points, the dates at which the economy switches from expansion to recession and vice versa, are not directly observable. Indeed, determining the reference cycle peaks and troughs is very crucial for policy makers, for businesses and for the academia because they are used to determine the causes of recessions, to design public policies, to guide investors and to test competing economic theories, among others.

Aware of this requirement, Martin Feldstein established a Business Cycle Dating Committee of National Bureau of Economic Research (NBER) scholars and gave it responsibility for business cycle dating when he took over the institution in 1978. In sum, the committee looks at various coincident indicators to make informative judgments on when to set the historical dates of the peaks and troughs of the past US business cycles. Following these guidelines, other countries have been creating similar committees, such as the Euro Area Business Cycle Dating Committee of the Center for Economic Policy Research, founded in 2002, and the Spanish Business Cycle Dating Committee (SBCDC) of the Spanish Economic Association, created in 2014.

Despite the interest in establishing and maintaining a historical chronology of the business cycle, the dating methodology of the Committees has received some criticism as their decisions represent the consensus of individuals. Thus, their dating methodology is neither transparent nor reproducible. In addition, the dating committees date the turning points with a considerable lag, which last more than two years in some cases. This reduces the interest of the committees' decisions to provide real-time assessments of the business cycle changes.

To overcome these drawbacks, this paper codifies in a systematic and transparent way a historical chronology of business cycle turning points for Spain. Following the business cycle concept of Burns and Mitchell (1946), we consider the reference cycle as a set of dates of wave-like movements occurring at about the same time and in many economic activities. Its computation requires collecting a number of coincident indicators and determining global change-points, from which the reference cycle can be extracted using average-then-date or date-then-aggregate methods. The first case consists on summarizing the coincident information in a single composite indicator from which the turning points are determined. The second case first identifies turning points in the individual indicators and looks for common turning points. Although the literature has focused primarily on average-then-date methods, the date-then-average alternative has recently proved to be successful in dating the turning points. Examples of the latter are Harding and Pagan (2006, 2016), Chauvet and Piger (2008), and Stock and Watson (2010, 2014).

Against this background, Camacho et al. (2021) recently proposed a novel date-then-average methodology that views the reference cycle as a multiple change-point mixture model. They consider that the reference cycle is a collection of increasing change-points (peak-trough dates) that segment the time span into non-overlapping episodes. With the help of a Markov-switching mixture model representation, the method classifies the dates and the specific turning points, which are viewed as realization of a mixture of Gaussian distributions.

This method has a number of important advantages over other date-then-average methods. First, the number of historical change-points are data-driven, so the estimates are not conditioned by the known occurrence of a phase shift as in Stock and Watson (2010, 2014). Second, the estimation is simple as it is estimated using standard Bayesian techniques of finite Markov mixture models. Third, the reference dates are population concepts, which allows us to make inferences in

contrast to Harding and Pagan (2006, 2016). Four, missing data for some coincident indicators is not a problem as they only imply determining some reference cycles from less observations. Five, the detection of phase changes in real time is as simple as it reduced to a classification problem. This method was successfully applied to date the US business cycle by using several monthly coincident indicators.

This paper addresses the challenge of applying this methodology to achieve the reference cycle dates in Spain in a quarterly basis. This poses us several difficulties. Our first challenge was collecting a set of coincident economic indicators with homogeneous quality information throughout the entire sample given that the Spanish economy has suffered from dramatic transformations in the last five decades, especially in the seventies and eighties. Our second challenge was adapting the method to a quarterly basis. In order to date the turning points of each quarterly indicator, we employ either the so-called BBQ algorithm -proposed by Harding and Pagan (2002), who extend the monthly algorithm of Bry and Boschan (1971) to a quarterly basis- or the parametric Markov-Switching procedure -proposed by Hamilton (1989)- depending on the characteristics of each indicator.

We evaluate the extension of the Camacho et al. (2021) algorithm developed in this paper in terms of its capacity to generate the SBCDC reference cycle for Spain. The results of this exercise suggest that our approach is capable of identifying the SBCDC turning points with reasonable accuracy. Leaving aside the period of the late 1970s and early 1980s, where some deviations occur with respect to the chronology established by the SBCDC, the method clearly identifies the crisis of the 1990s, the double dip with which Spain received the impact of the Great Recession and, finally, the current impact of the Covid-19 pandemic. Thus, we consider that this method can be very useful to complement and guide the work carried out by the SBCDC in dating the Spanish business cycles.

The rest of the paper is organized as follows. Section 2 summarizes the methodology proposed in Camacho et al. (2021) which estimation technique is described in Section 3. Section 4 presents the empirical application to date the reference cycle of the Spanish reference cycle since the 70s. Finally, Section 5 concludes.

2 Multiple change-point model

Reference cycle dates can be obtained from the estimates of a multiple change-point model with an unknown number of K breaks. In practice, we have data in the bivariate time series $\tau = \{\tau_1, \dots, \tau_N\}$ of the specific pairs of turning point dates collected from a set of coincident economic indicators. Each of these turning points, τ_i , contains a peak date, τ_i^P , and a trough date, τ_i^T , and determines the start and the end of one specific recessions for a particular coincident indicator. For estimation purposes, the observed pairs of turning point dates are stacked in ascending order, meaning that $\tau_i^P \leq \tau_{i+1}^P$ and $\tau_i^T \leq \tau_{i+1}^T$, for $i = 1, \dots, N - 1$.

The approach described in Camacho et al. (2021) assumes that each individual pair of peak and trough dates in a reference cycle k , τ_i , is a realization of a density conditioned by $\tau^{i-1} = \{\tau_1, \dots, \tau_{i-1}\}$ that depends on the two dimension vector $\mu_k = (\mu_k^P, \mu_k^T)'$ and a covariance matrix Σ_k . The reference cycle dates are viewed as the means of these distributions, around which the specific turning point dates are clustered. The means and covariances change at unknown time periods breaking the

time span of interest into segments corresponding to the periods occupied by the $k = 1, 2, \dots, K$ business cycles.

We collect the distribution parameters in the reference cycle k in $\theta_k = (\mu_k, \Sigma_k)$ and assume that these parameters remain constant within each regime and change their values when a regime change occurs. Then, given that the state is k , τ_i is drawn from the population given the conditional density

$$\tau_i | \mathcal{T}_{i-1}, \theta_k \sim p(\tau_i | \tau^{i-1}, \theta_k) \quad (1)$$

for $k = 1, \dots, K$, where $p(\tau_i | \tau^{i-1}, \theta_k)$ is the Gaussian density $N(\mu_k, \Sigma_k)$. Let us collect all the unknown distribution parameters in $\theta = (\theta_1, \dots, \theta_K)$.

This multiple change-point model can be formulated in terms of an integer-valued unobserved state variable s referred to as the state of the system and that controls the regime changes. In particular, the discrete random variable is allowed to take values in the set $\{1, \dots, K\}$ in the whole sequence of realizations, which are collected in $S = (s_1, \dots, s_N)$. Thus, $s_i = k$ indicates that τ_i is drawn from $p(\tau_i | \mathcal{T}_{i-1}, \theta_k)$.

The indicator variable s_i is modeled as a discrete time, discrete-state first-order Markov chain, which implies that the probability of a change in regime depends on the past only through the value of the most recent regime. In this case, the probability of moving from regime l to regime k at observation i conditional on past regimes and past observations τ^{i-1} is

$$\Pr(\tau_i = k | s_{i-1} = l, \dots, s_1 = w, \tau^{i-1}) = \Pr(s_i = k | s_{i-1} = l) = p_{lk}. \quad (2)$$

In this context, the transition probabilities are constrained to reflect the one-step ahead dynamics of the multiple change-point specification. In particular, the order constraints of the break point model hold when the transition probabilities hold the following restrictions

$$\Pr(s_i = k | s_{i-1} = l) = \begin{cases} p_{ll} & \text{if } k = l \neq M \\ 1 - p_{ll} & \text{if } k = l + 1 \\ 1 & \text{if } l = K \\ 0 & \text{otherwise} \end{cases}. \quad (3)$$

The restrictions imply that when the process reaches one regime, for example regime l , it remains in this regime with probability p_{ll} or moves to regime $l+1$ with a probability $1-p_{ll}$. The process starts at regime 1 and moves forward (never backward) to the next regime until it reaches regime K in which the process stays permanently. Let us collect the transition probabilities $\{p_{11}, \dots, p_{K-1K-1}\}$ in the vector π .

3 Bayesian estimation

The problem is assigning each specific turning point τ_i to an unknown reference cycle by making inference on the unobserved allocation s_i . To perform the estimation of the parameters collected in θ and π and the inference on the set of allocations S , Chib (1998) describes a Markov Chain Monte Carlo (MCMC) algorithm. From the set of observations, the algorithm is implemented using the Gibbs sampler by simulating the conditional distribution of the parameters given the states, and the conditional distribution of the states given the parameters.

3.1 Simulation of the parameters

Let us consider sampling the transition probabilities and the parameters of the Gaussian distributions given the observations and their corresponding allocations, which are assumed to be independent. Starting with the transition probabilities, Albert and Chib (1993) showed that the full conditional distribution $\pi|S, \tau$ is independent of τ given S . Following Chib (1998), we simulate the transition probabilities from $\pi|S$ using independent *Beta* priors, $p_{kk} \sim \text{Beta}(a, b)$. The posteriors remain independent and also follow *Beta* distributions

$$p_{kk}|S \sim \text{Beta}(a + n_{kk}, b + 1), \quad (4)$$

where n_{kk} is the number of periods the process stays in regime k , with $k = 1, \dots, K$.¹

To sample θ conditioned to the data and their allocations, we propose an independent normal inverse-Wishart prior. Conditional on the mean μ_k the prior of the inverse covariance matrix is $\Sigma_k^{-1} \sim W(c_0, C_0^{-1})$ and the posterior is $\Sigma_k^{-1}|S, \tau, \mu_k \sim W(c_k, C_k^{-1})$, where

$$c_k = c_0 + N_k \quad (5)$$

$$C_k = C_0 + \sum_{i:s_i=k} (\tau_i - \mu_k)(\tau_i - \mu_k)'. \quad (6)$$

In addition, conditional on the covariance matrix, the prior distribution for the means is $\mu_k \sim N(b_0, B_0)$, and its posterior is $\mu_k|S, \tau, \Sigma_k \sim N(b_k, B_k)$, where

$$B_k = (B_0^{-1} + N_k(S)\Sigma_k^{-1})^{-1} \quad (7)$$

$$b_k = B_k(S) \left(B_0^{-1}b_0 + \Sigma_k^{-1} \sum_{i:s_i=k} \tau_i \right). \quad (8)$$

Given the data, one can easily sample means and covariances from their conditional posterior distributions.

3.2 Simulation of the states

Let us consider now the question of sampling the states from the distribution $S|\theta, \pi, \tau$. Using the properties of the first-order Markov chain, this distribution can be stated as

$$\Pr(S|\theta, \pi, \tau) = \Pr(s_N|\theta, \pi, \tau^N) \prod_{i=1}^{N-1} \Pr(s_i|s_{i+1}, \theta, \pi, \tau^i), \quad (9)$$

where the last of these distributions is degenerated because we assume that the process starts at $s_1 = 1$.

Chib (1996) showed that each of the probabilities in expression (9) holds

$$\Pr(s_i|s_{i+1}, \theta, \pi, \tau^i) \propto \Pr(s_{i+1}|s_i) \Pr(s_i|\theta, \pi, \tau^i). \quad (10)$$

The first ingredient in this expression refers to the transition probabilities, which are simulated from the distribution.

¹Note that $N_{KK+1} = 1$.

The second ingredient in (10), $\Pr(s_i|\theta, \pi, \tau^i)$, refers to the filtered probabilities and requires a recursive calculation. Starting from an initial value $\Pr(s_0 = k|\theta, \pi, \tau^0)$, with $k = 1, \dots, K$, the following two steps are carried out recursively for $i = 1, \dots, N$.² First, the one-step ahead prediction with the information up to turning point $i - 1$

$$\Pr(s_i = k|\theta, \pi, \tau^{i-1}) = \sum_{l=1}^K p_{lk} \Pr(s_{i-1} = l|\theta, \pi, \tau^{i-1}) \quad (11)$$

is computed. Second, when the i -th turning point is added, the filtered probability is updated as follows

$$\Pr(s_i = k|\theta, \pi, \tau^i) = \frac{\Pr(s_i = k|\theta, \pi, \tau^{i-1}) p(\tau_i|\theta_k, \tau^{i-1})}{p(\tau_i|\theta, \pi, \tau^{i-1})}, \quad (12)$$

where $p(\tau_i|\theta, \pi, \tau^{i-1}) = \sum_{k=1}^K \Pr(s_i = k|\theta, \pi, \tau^{i-1}) p(\tau_i|\theta_k, \tau^{i-1})$. Now, one can obtain the conditional distribution of the states, $\Pr(s_i|s_{i+1}, \theta, \pi, \tau^i)$, by backward recursion.

The unconstrained MCMC sampler described above could have identifiability problems because it is subject to potential label switching issues. To identify the model, we use the identifiability constraint that the draws must imply a segmentation of the time span into K non-overlapping episodes, i.e., $\mu_k^P < \mu_k^T < \mu_{k+1}^P$ for all $k = 1, \dots, K$. To ensure that the restrictions apply, we employ rejection sampling, which is achieved by discarding the draws that do not satisfy the constraints and the sampler is implemented again until the condition is satisfied.³

3.3 The number of clusters of turning point dates

We have assumed so far that the number of components, K , was known. However, in practice, one needs to infer the number of groups of specific turning point dates that are cohesive and form a distinct cluster separate from other clusters of specific business cycle turning point dates. Notably, despite the huge effort made in this area, the problem of choosing the number of clusters is still unsolved. So, with the aim of robustness, we follow several different approaches in the empirical analysis.

Among the likelihood-based methods, the simplest case is to choose the model with the number of components K that reaches the highest marginal likelihood, $\log p(\tau|\hat{\theta}_K, M_K)$, from a set of potential values of $\{1, \dots, K^*\}$, where the upper bound K^* is specified by the user, M_K is a mixture model of K components, and $\hat{\theta}_K$ is the d_K -dimensional vector of its maximum likelihood estimated parameters.

Since this method tends to choose models with a large number of components, we also consider selecting criteria that introduce an explicit penalty term for model complexity. The Akaike model selection procedure is based on choosing the value of K for which $AIC_K = -2 \log p(y|\hat{\theta}_K) + 2d_K$ reaches a minimum. Alternatively, one can choose the model that minimizes Schwarz's criterion, which is often formulated in terms of minimizing $BIC_K = -2 \log p(y|\hat{\theta}_K) + d_K \log(N)$.⁴

We also decide the number of components by choosing the model with the number of components that maximizes the quality of the classification. For this purpose, we define the entropy as

²In the applications, we set $Pr(s_0 = 1|\theta, \pi, \tau^0) = 1$, and $Pr(s_0 = k|\theta, \pi, \tau^0) = 0$ for $k = 2, \dots, K$.

³The identifiability problem is discussed extensively in Frühwirth-Schnatter (2006). It should be noted that the multiple change-point model can also be interpreted as a model of finite mixtures.

⁴In practice, Akaike's criterion tends to favor more complex models than Schwarz's criterion since the latter penalizes over-fitted models more heavily.

$$EN_K = - \sum_{i=1}^N \sum_{k=1}^K p(s_i = k | \tau_i, \theta) \log p(s_i = k | \tau_i, \theta), \quad (13)$$

which measures how well the data are classified given a mixture distribution. The entropy takes the value of 0 for a perfect partition of the data and a positive number that increases as the quality of the classification deteriorates.

One interesting option is to combine the aim of selecting a model with an optimal number of components (as in the case of the likelihood-based methods) with the aim of obtaining a model with a good partition of the data (as proposed by model selection criteria based on entropy measures). For this purpose, we also consider choosing a model whose number of components minimizes $BIC_K + EN_K$, which is a metric that penalizes not only model complexity but also misclassification.

In this context, Kass and Raftery (1995) show that a useful way to compare two models with different numbers of clusters K_i and K_j , is to compute twice the natural logarithm of the Bayes factor B_{ij} , which is the ratio of the two integrated likelihoods that correspond to the models with K_i and K_j clusters, respectively. Fraley and Raftery (2002) point out that this measure can be approximated by the difference between the two corresponding BIC

$$2 \log(B_{ij}) \approx BIC(K_i) - BIC(K_j). \quad (14)$$

This quantity provides a measure of whether the data increases or decreases the odds of a model with K_i clusters relative to a model with K_j clusters. Kass and Raftery (1995) propose that values of $2 \log(B_{ij})$ less than 2 correspond to weak evidence in favor of the model with K_j clusters, values between 2 and 6 to positive evidence, between 6 and 10 to strong evidence, and greater than 10 to very strong evidence.

3.4 Data problems

The application of finite Markov mixture models to turning point data presents two data challenges. The first is that finite mixtures of Gaussian distributions are defined for continuous data. However, the turning point dates are obtained from coincident indicators that are usually sampled at monthly or quarterly frequencies and this generates a discontinuity challenge.

For example, when the coincident indicators are sampled at quarterly frequencies, turning points are dates in the format $YYYY.q$, which refers to quarter q of year $YYYY$. In this case, the distance between the third month and the fourth month of a year is lower than that between the fourth quarter of the same year and the first quarter of the following year. To overcome this drawback, we propose the transformation of the turning points to dates $YYYY.d$, where $d = 1/4(q - 1)$, which can obviously be changed back to recover the results in the standard quarterly calendar dates.⁵

The second challenge of this approach is that the individual turning points are collected from sets of disaggregated economic indicators and some of them might not be available for the full span. In practice, some indicators are reported with lags exhibiting missing observations in the latest months while others are available only for a diminished time span because they are not available at the beginning of the sample. In our context, this is rarely a problem since it only implies that the probability distributions of turning points at early and end dates must be estimated from a relatively lower number of observations due to missing data.

⁵In a similar fashion, when the coincident indicators are sampled at monthly frequency, the dates $YYYY.mm$ are transformed into dates $YYYY.d$, where $d = 1/12(mm - 1)$.

4 Empirical application

In this section we apply the proposed automatic algorithm to generate the reference cycle of the Spanish economy. Using the multiple change-point model described above to date the reference cycle turning point dates requires several steps in practice. To serve as an example to other application, we review these steps in this section.

4.1 Collect a set of business cycle indicators

The national Business Cycle Dating Committees of the National Bureau of Economic Research (NBER) examines the behavior of a broad set of economic indicators. Because a recession influences the economy broadly and is not confined to one sector, the committees emphasize economy-wide measures of economic activity. Typically, the committees view real gross domestic product (GDP) as the single best measure of aggregate economic activity.

Figure 1 shows that GDP does not increase every single year in Spain. Instead, there are particular identifiable episodes during which GDP sharply falls. The downturns occur at around the periods designated as recessions by the SBCDC, which are marked as shaded areas in the graph.⁶ The committee identified six recessions since the 1970s: the impact of the two oil crises in the seventies and early eighties, the brief crisis of the nineties, the double dip caused by the global financial crisis in 2008 and the European sovereign debt crisis in 2010 and, finally, the economic impact of Covid-19 pandemic. To determine when a particular recession begins and ends, the committee uses judgments and their decisions are, therefore, difficult to replicate.

To overcome this inconvenient, in this paper we date peaks and troughs from economic indicators using the so-called BBQ mechanical algorithm, which is an implementation of the methodology outlined in Harding and Pagan (2002). In short, the nonparametric algorithm provides peaks and troughs of a time series by following three steps: (i) it estimates the possible turning points in the series by picking the local maxima and minima; (ii) it ensures alternating the troughs and the peaks; and (iii) it applies a set of censoring of rules that meet pre-determined criteria of the duration and amplitudes of phases and complete cycles.⁷ To ease of comparison, the peaks and trough dates identified with BBQ are plotted in Figure 1 as red vertical lines. As can be seen in the figure, there is no instance in which a SBCDC recession is not identified by the BBQ algorithm, which picks up an additional recession at the begging of the eighties.

Although GDP is the most prominent indicator, the committees emphasize that a recession involves a significant decline in economic activity that is widespread across the economy. Thus, they consider a variety of indicators to determine turning points. In this paper, we have collected a broad set of 51 specific indicators of both aggregate activity and employment, and others of a sectoral nature, which includes both hard and soft indicators. Table A1 in the Appendix shows the details of the definitions and the acronyms that will be used in the rest of the work as well as the sources and samples. Overall, the sample begins in 1970.2 and ends in 2020.3, although each series has a different length and combines monthly and quarterly frequencies. The monthly series are transformed into quarterly by taking the average of the corresponding 3 months, although all series have been previously seasonally adjusted to their original frequency.

⁶Information about the Committee and the Spanish turning point dates can be found at <http://asesec.org/CFCweb/en/>

⁷For example, a complete cycle must last four quarters at least. In addition, the minimum duration of a recession is two quarters.

The specific business cycle turning points have been dated by applying either the BBQ algorithm, when the series have a trend, or the Markov-Switching procedure, when they are represented by a composite index or growth rates. Figure 2 plots series-specific recession episodes, where a series-specific recession (in red) is defined to be the period from a BBQ peak to a BBQ trough or periods with a probability of being in recession above 0.5. In addition, Figure 3 displays the estimated probability that all indicators are in recession in each quarter, adding the SBCDC turning points as red vertical lines.

The SBCDC recessions are clearly visible in these figures. Dating the two oil crises is quite challenging due to the scarce number and low quality of available series in the seventies when Spain was under the late years of Franco's dictatorship. The first impact of the oil price rise in the fall of 1973 only became noticeable in the Spanish economy in 1975, and the implementation of several accommodative policies delayed its impact on the real sector until 1977. In 1979, the economy suffered from serious structural problems, while it was hit by the second rise in the oil price, plunging the country into a recession and a long period of weakness that lasted until well into the next decade.

The SBCDC identifies two peaks: in 1974.4, signaling a recession between 1975.1 and 1975.2, and in 1978.3, resulting in a recession between 1978.4 and 1979.2. The application of the BBQ algorithm, delays the dating of each of these recessions by a quarter. The algorithm also identifies a third recession between 1981.2-1981.3, which captures the slow recovery of the Spanish economy in the face of the second oil price hike. This behavior differs from most of the rest of the European countries, which underwent a recovery at the beginning of the eighties.

The crisis of the nineties took place between 1992.2 and 1993.2 according to the SBCDC and is delayed by a quarter if we apply the BBQ algorithm to GDP. It was a short intense recession after years of sound economic growth. The recession was synchronized with the cyclical phase of most European countries. A high percentage of series point to a recession: 71% of the 51 series detect the crisis in the quarters spanning from 1992.3 to 1993.1 and 54% lead or lag its effect.

The global financial crisis and the European debt crisis had a strong impact on the Spanish economy, which involved two waves of recession separated by a brief interval of recovery. Most of the series accurately pinpoint the first hit of the crisis (Figures 2 and 3). Specifically, between 2008.3 and 2009.1 (official SBCDC dates) more than two thirds of the series identified the crisis, although 69% had already detected it in 2008.2. The second impact, dated between 2011.1 and 2013.2 according to SBCDC, is confirmed by 70% of the series analyzed. A non-negligible set of series, around 50%, did not detect the intermediate recovery period between both impacts.

At the beginning of 2020, the world was suddenly hit by a global pandemic due to the coronavirus disease. The pandemic triggered the most severe global economic crisis since World War II in terms of speed and intensity. To mitigate the propagation of the virus, governments implemented unprecedented measures, such as restrictions on movement and travel, population lockdowns, and social distancing in some targeted activities. Official institutions place the recession peak in the last quarter of 2019 or in February of 2020 at the monthly frequency. The SBCDC places the peak in the fourth quarter of 2019 and, therefore, the beginning of the crisis in the first quarter of 2020. More than 70% of the series detect the crisis, although there is a significant percentage of series, around 50%, that already pointed to a slowdown in the economy that was obviously not related to Covid-19 pandemic which, being an exogenous shock to the economy, was impossible to predict.

Summing up, around the dates that the SBCDC identifies recessions there is a wide set of specific indicators that show a recession. However, there are indicators that either advance (usually related to the industrial sector), delay or extend the duration of recessions (for example, those associated with the labor market). Furthermore, not all the specific indicators have the same behavior in each recession.

4.2 Select a set of highly coincident indicators

Although Figures 2 and 3 show that the set of indicators would be useful for ascertaining whether a turning point has occurred, it is evident that there is considerable dispersion of the specific-series turning points around the beginning and end of the recessions. As the multiple change-point method assumes that the series being dated have been chosen to be coincident indicators, we strongly recommend testing for this in advance by checking how synchronized the cycles are.

For this purpose, we address the issue of synchronization between each indicator and the SBCDC chronology by computing the index of concordance advocated by Harding and Pagan (2002), which measures the percentage of time units spent in the same phase. The index, whose values are plotted in Figure 4, confirms that there is a relatively high degree of synchronization within the SBCDC business cycles. To ensure strong synchronization, we pick only those indicators for which the concordance index exceeds 90%. Thus, the final set of indicators that exhibit the highest correlation with the SBCDC chronology is formed by six time series: *SSAI*, *SAI*, *ESIS*, *ESIC*, *GDP* and *PC*. This selection, which combines aggregated, sectorial, hard and soft indicators, is not very different from the one commonly used by the NBER. Employment indicators are not capable of pinpointing the Spanish cycle since they have a lagging behavior during expansions.

We use the SBCDC chronology to select the coincident indicators, given the high coincidence of the modes detected in the set of indicators and the dating of the committee. Nevertheless, it should be noted that our methodology does not require any prior information or the existence of any dating committee. Alternatively, we could select the optimal set of coincident indicators with a search algorithm applied to different combinations of indicators, where the optimal combination could be selected on the basis of the number of clusters, the entropy and the variance of the confidence intervals. This procedure could allow us to discard leading or lagging indicators and to choose the combination of indicators that produces the largest modes. Since this option is computationally more expensive, external information of experts, such as that of the SBCDC, can be used to speed up the process.

It should be noted that, in any case, the choice of coinciding indicators, a key element for the success of the procedure, requires the good judgment of the expert who must evaluate the characteristics of each country, such as its productive structure, the quality and methodological changes in the series or any other circumstance. Likewise, a qualitative assessment of the resulting chronology would also be helpful. This is the reason to establish expert committees in different countries.

Figure 5 provides a preliminary inspection of the individual chronologies of turning points by plotting the kernel density estimator of the bivariate distribution of the resulting pairs of specific peaks and troughs computed with the BBQ algorithm or the Markov-Switching procedure. The distribution of turning points is multimodal, exhibiting various modes that cluster the individual turning points around periods of SBCDC-referenced recoveries and declines. In particular, the

figure exhibits several modes around the seventies and the beginning of the eighties, the early nineties, the first decade of the new century and, finally, a large concentrated mode at the end of the sample, in 2020. The figure also shows a larger variance of the turning point dates in the seventies and early eighties, probably due to the lower quality of the indicators in the late years of Franco’s dictatorship and the unstable political transition towards democracy. By contrast, this figure shows a lower dispersion of the specific turning point dates in the nineties, around the three modes of the new century related with the double-dip, and in 2020, and the specific turning points shrink toward the modes.

4.3 The number of clusters

Prior to determining the number of clusters, we need to specify whether the first phase of reference cycle starts in a recession or an expansion. In our case, we determine that the Spanish economy was in expansion in 1970.2 because the first turning point for all the indicators that were available in this period was a peak. In addition, the last turning point in the six indicators is a peak, which implies that the pairs of turning points are not complete. In this case, we create artificial troughs by adding the average duration of the preceding recessions to the peaks.

The modes of the kernel density of pairs of specific peaks and troughs displayed in Figure 5 suggest that the tentative number of different clusters in the reference cycle could be six or seven, which correspond with the distinct local maxima. The main discrepancy between the modes and the SBCDC referenced turning points appears at the end of the seventies and the early eighties. This period seems to include an additional cluster.

To formally determine the number of clusters, we estimate a set of models M_K for $K = 1, \dots, K^*$, with $K^* = 8$, and compute the measures described in Section 3.3 for each K , whose estimates are reported in Table 1. The first column of the table shows that the log of the marginal likelihood increases uninterruptedly when the number of clusters increases from $K = 1$ to $K = 7$, where it reaches a peak at 42.61, which corresponds with $K = 7$. The algorithm is not able to generate a valid result for $K = 8$ as there are not sufficient numbers of draws to achieve identification. Although this suggests choosing $K = 7$, the marginal likelihood does not take into account the number of parameters in the model selection and tends to overestimate the number of clusters.

Regarding model selection criteria that introduce penalties in model selection, the reported values in Table 1 of AIC , BIC and BIC corrected by misclassification reach their minimums of -3.22 , 134.01 , and 134.01 , respectively, for $K = 7$. This indicates that the Spanish reference cycle requires seven separate cycles. In addition, the entropy of the mixture model with seven clusters is zero, showing that the model produces a clear segmentation of the reference cycle.⁸ Finally, as in Camacho et al. (2021), we follow Chib (1998) and compute the Bayes factors sequentially. The sequence of (twice the log of) Bayes factors that compare two models with $K - 1$ and K different numbers of clusters, with $K = 1, \dots, 8$, also points to $K = 7$. The Bayes factor that compares the model with $K = 7$ clusters with the model with $K = 6$ clusters favors the extra cluster. Although this result does not require prior knowledge of the number of clusters, it is worth noting that the

⁸The procedure also takes into account some indicators of the quality of the results. More specifically, we count the percentage of draws in which the algorithm has been unable to generate a bivariate vector of μ that complies with the partition constraints of the reference cycle; we check if the estimated μ values are within the range of data values allowing a margin of 1 year at the beginning and end; and, finally, we count the percentage of classifications that leave an empty cluster.

SBCDC establishes $K = 6$ clusters, considering the last peak as an incomplete pair. The difference is found in the late seventies and early eighties, a complex period with scarce information, in which the SBCDC chronology establishes 2 recessions, while the specific indicators detect 3 as shown in Figure 5.

4.4 Estimation of turning points

Table 2 shows the results of evaluating our proposal in terms of its ability to capture the turning point dates established by the Spanish Business Cycle Dating Committee. The columns labeled SBCDC and MCPM report the reference cycle dates as determined by the Spanish Business Cycle Dating Committee and our multiple change-point model. The distinct means are estimated from the posterior distributions with the help of the Gibbs rejection sampler algorithm for the mixture model.⁹ Using the outputs of the MCMC algorithm, this table also reports, in brackets, the range of values of the posterior probability distributions that include 95% of probability.

The figures displayed in the table show that there are no instances in which a SBCDC turning point is not matched by a quite similar date produced by our proposal and that all credible intervals contain the SBCDC dates in all episodes. Notably, our proposal replicates the SBCDC peaks and troughs very accurately, especially for the turning points dated since the nineties. However, our method provides an extra pair of turning points at the beginning of the eighties. Finally, the credible intervals, which range from 0 to 1 quarters for peaks and from 0 to 3 quarters for troughs, reveal that the method provides a precise estimate of the reference dates.

In particular, the short recession of the nineties is detected with a delay of a quarter in both the peak and the trough, although the confidence intervals include the exact date provided by the SBCDC. Regarding the global financial crisis, the first impact is identified early in the peak and late in the trough, and something similar occurs with the second blow of the crisis. Finally, the peak that marks the beginning of the crisis caused by the Covid-19 pandemic fully matches that announced by the SBCDC. As expected, the uncertainty in determining the reference dates is a bit larger in the seventies. In fact, the Committee acknowledges that “Business cycle dating for the period 1970.1 (the start of the sample based on available national income accounts at quarterly frequency), until about 1986.4 is especially challenging”.

Figures 6 and 7 display some technical aspects of the estimation process. In particular, Figure 6 plots the sampling representations from the rejection sampler described in Section 2, which is a very useful tool for visualizing the posterior mixture distribution. For each of the seven clusters, the figure displays the two-dimensional scatter plots of the MCMC draws for $(\mu_k^P, \Sigma_{11,k})$ and $(\mu_k^T, \Sigma_{22,k})$ in Panel A and for (μ_k^P, μ_k^T) in Panel B. From these scatter plots, it is obvious that the component parameters differ mainly in the means, which present the highest ability to divide the draws into the 7 separate groups, whereas the variances are quite similar for many groups. In addition, Panel B suggests that the clusters of turning points alternate sequentially, supporting the non ergodic behavior imposed in the transition matrix to capture the transitions across the distinct business cycles, as a multiple-change model, which roughly occur about at the SBCDC-designated turning points.

Figure 7 shows diagnostic plots of the post burn-in draws from the conditional distributions of μ_k^P , μ_k^T , $\log(|\Sigma_k|)$ and $p_{k,k}$ and $p_{k,k+1}$ for each of the $K = 7$ clusters. These panels help us to

⁹The BBQ algorithm imposes that a recession should last at least 2 quarters and an expansion, at least 4 quarters.

detect potential convergence problems or label switching, which arise when the mixture likelihood function is not invariant to relabeling the components of a mixture model. The paths of the draws show that the rejection sampler that imposes, $\mu_k^{P(m)} < \mu_k^{T(m)} < \mu_{k+1}^{P(m)}$, for all $k = 1, \dots, K$, is useful to prevent label switching because the sampler stays within the modal regions corresponding to the initial labeling. So, these regions are well separated from the others leading to a unique labeling.

In addition, we have computed the Gelman-Rubin scale reduction factor diagnostic corrected by accounting for sampling variability, usually known as R_c . After conducting an experiment that allows the priors to move randomly over a range of values, all the Gelman-Rubin diagnostics were 1 (or very close to one). More importantly, the maximum R_c was less than 1.1, which indicated that no convergence issues were detected. This test also assesses the sensitivity of the results to the choice of different priors used in each MCMC. The variability of the μ_k^P and μ_k^T values is zero. Therefore, their precision in the estimation is very high, with the only exception of the peak that marks the beginning of the 90's crisis, which ranges between 1992.1 and 1992.2.

To provide the results with additional economic meaning, it is worth examining the ability of the multiple change-point model to provide a posterior classification of the specific turning point dates into the different business cycles of the Spanish reference cycle. With this aim, we estimate the posterior classification probabilities $\Pr(s_i = k|\theta)$, with $k = 1, \dots, 7$ and $i = 1, \dots, N$, from the MCMC draws through the corresponding relative frequency of the retained state draws, as described in Section 2. Figure 8, which plots the estimated classification probabilities for each cluster, shows that the model clearly classifies each specific date, whose classification probabilities are always close to either 1 or 0. In addition, the probabilities agree with the multiple change-point behavior of the model, with a large persistence of each state and probabilities that increase nearly to unity after each structural break that occurs about the turning point dates identified by the SBCDC. Thus, the probabilities support the view that the multiple change-point model performs a segmentation of the time span into non-overlapping business cycle episodes that is consistent with the SBCDC-referenced cycles.

5 Conclusions

This paper provides an automatic procedure to date the reference cycle turning points in Spain. Based on the novel methodology proposed by Camacho et al. (2021), each pair of specific peaks and troughs from the individual indicators is viewed as a realization of a mixture of an unspecified number of separate variation Gaussian distributions whose different means are the reference turning points. This approach is equivalent to considering a multiple change-point model where the reference cycle is a collection of increasing change-points (peak-trough dates) that segment the time span into K non-overlapping episodes. The computation is carried out with Bayesian techniques using a Markov Chain Monte Carlo (MCMC) algorithm which is implemented using the Gibbs sampler.

The application comprises several steps. Firstly, we collect a broad set of specific indicators that includes aggregate and sectorial variables, both hard and soft. Secondly, we compute the turning points of each individual indicator by using the quarterly Bry-Boschan method and analyze carefully their behavior with respect to the business cycle. A search algorithm using the concordance index allows us to find the optimal set of coincident indicators and build the database of peak-trough

pairs. Thirdly, we select the number of clusters through several measures based on the likelihood function and Bayesian criteria. Finally, we estimate the turning points as the mean of the draw distribution for each cluster and then compute their confidence intervals.

The method identifies seven recessions in the period from 1970.2 to 2020.2. Three of these recessions are dated in the seventies and early eighties. It also dates the recession of the nineties, the double-dip of the global financial crisis and the sovereign debt crisis in the first two decades of the 21st century and, finally, the recent hit of the Covid-19 pandemic. These results support the strong performance of our approach to identifying the Spanish reference cycle, showing little difference from the timing of the turning point dates established by the Spanish Business Cycle Dating Committee (SBCDC). In fact, since the nineties there has been an almost perfect coincidence of the timings. The greatest discrepancies occurred in the seventies and early eighties, a turbulent period in Spain characterized by the juxtaposition of the successive oil crises and the political transition of the 1970s together with the scarcity of statistical sources.

Summing up, the method proposed by Camacho et al. (2021) has successfully overcome the challenge of producing a credible chronology of the reference cycle of the Spanish economy using automatic rules based on a set of specific indicators. It is therefore a useful instrument that complements the work carried out by the Spanish Business Cycle Dating Committee and that can be used by both policy-makers and academics interested in analyzing the economic cycle in Spain.

References

Albert, J., and S. Chib (1993). "Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts", *Journal of Business and Economic Statistics*, 11, pp. 1-15.

Burns, A., and W. Mitchell (1946). *Measuring Business Cycles*, New York, National Bureau of Economic Research.

Bry, G., and Ch. Boschan (1971). *Cyclical Analysis of Time Series: Procedures and Computer Programs*, New York, National Bureau of Economic Research.

Camacho, M., M. D. Gadea, and A. Gómez-Loscos (2021). "A new approach to dating the reference cycle", *Journal of Business and Economics Statistics*, <https://doi.org/10.1080/07350015.2020.1773834> (forthcoming).

Chauvet, M., and J. Piger (2008). "A comparison of the real-time performance of business cycle dating methods", *Journal of Business and Economic Statistics*, 26, pp. 42-49.

Chib, S. (1998). "Estimation and comparison of multiple change-point models", *Journal of Econometrics*, 86, pp. 221-241.

Fraley, Ch., and A. E. Raftery (2002). "Model-Based Clustering, Discriminant Analysis, and Density Estimation", *Journal of the American Economic Association*, 97(458), pp. 611-631.

Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*, Springer Series in Statistics, New York, Springer.

Hamilton, J. (1989). "A new approach to the economic analysis of nonstationary time series and the business cycles", *Econometrica*, 57, pp. 357-384.

Harding, D., and A. Pagan (2002). "Dissecting the cycle: a methodological investigation", *Journal of Monetary Economics*, 49, pp. 365-381.

Harding, D., and A. Pagan (2006). "Synchronization of cycles", *Journal of Econometrics*, 132, pp. 59-79.

Harding, D., and A. Pagan (2016). *The econometric analysis of recurrent events in macroeconomics and finance*, New Jersey, Princeton University Press.

Kass, R., and A. Raftery (1995). "Bayes factors", *Journal of the American Statistical Association*, 90, pp. 773-795.

Stock, J., and M. Watson (2010). "Indicators for dating business cycles: cross-history selection and comparisons", *American Economic Review: Papers and Proceedings*, 100, pp. 16-19.

Stock, J., and M. Watson (2014). "Estimating turning points using large data sets", *Journal of Econometrics*, 178, pp. 368-381.

Table 1: Results of selection of K

K	LogLik	AIC	BIC	Entropy	BIC-Entropy	Bayes factor (k=i/k=i+1)
1	-2229.78	4469.57	4486.30	-	-	-
2	-696.46	1414.91	1451.73	0.03	1451.80	3034.57
3	-400.14	834.27	891.18	0.00	891.18	560.55
4	-141.25	328.51	405.49	0.00	405.50	485.69
5	-25.28	108.57	205.63	0.00	205.63	199.86
6	13.64	42.72	159.87	0.00	159.87	45.76
7	42.61	-3.22	134.01	0.00	134.01	25.86
8	-	-	-	-	-	-

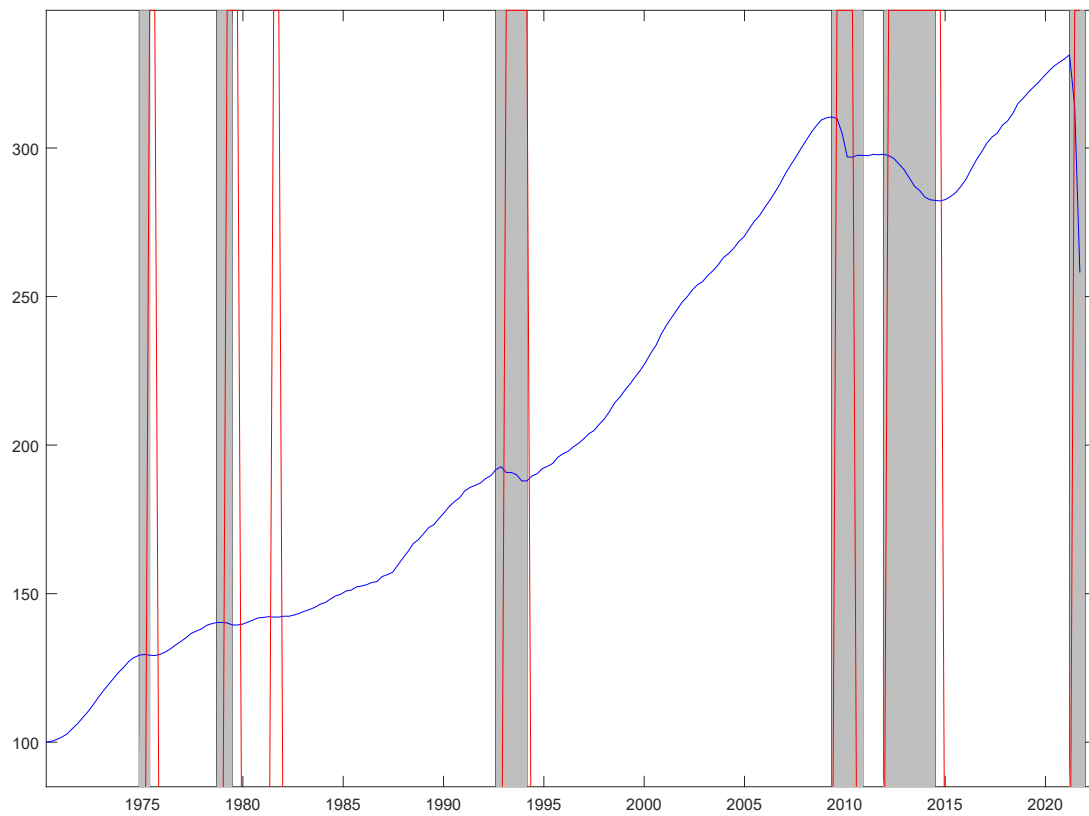
Notes: The first column refers to the marginal log likelihoods. The second and third columns refer to the Bayesian AIC and BIC selection criteria. The third column shows the entropy. The fourth column shows the BIC corrected for misclassification. The last column reports twice the log of the Bayes factor for $K - 1$ versus K clusters, with $K = 2, \dots, 8$.

Table 2: Results of empirical illustration (K=7)

SBCDC		MCPM		Deviation (in quarters)	
Peaks	Troughs	Peaks	Troughs	Peaks	Troughs
1974.4	1975.2	1974.4	1975.3	0	1
		(1974.4,1975.1)	(1975.3,1975.4)		
1978.3	1979.2	1978.3	1979.3	0	1
		(1978.3,1978.4)	(1979.2,1979.3)		
-	-	1980.4	1981.3	-	-
		(1980.4,1981.1)	(1981.3,1981.4)		
1992.1	1993.3	1992.2	1993.4	1	1
		(1992.1,1992.2)	(1993.3,1994.1)		
2008.2	2009.4	2007.4	2009.4	-2	0
		(2007.4,2007.4)	(2009.3,2009.4)		
2010.4	2013.2	2010.3	2013.3	-1	1
		(2010.3,2010.4)	(2013.3,2013.4)		
2019.40	-	2019.4	-	0.00	-
		(2019.4,2019.4)	(-, -)		

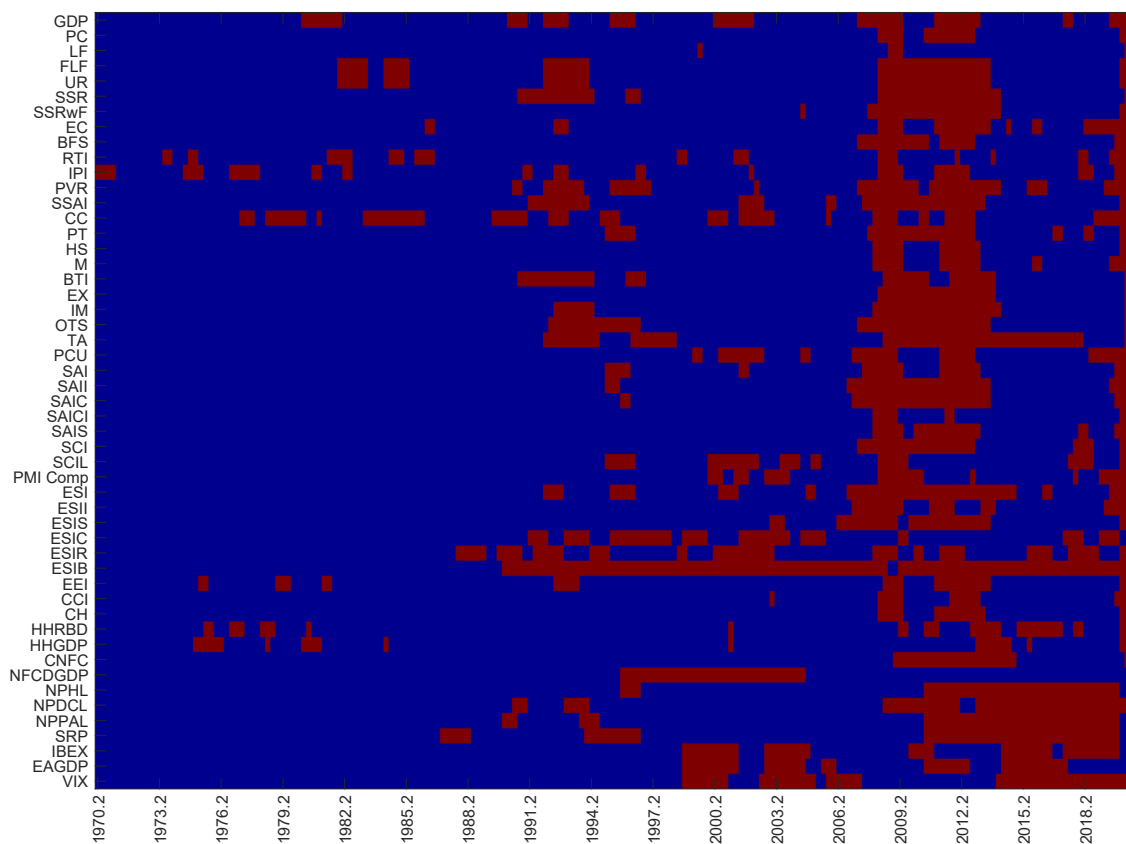
Notes: Columns 1 and 2 give the SBCDC-established turning point dates. Columns 3 and 4 report the turning point dates established with the multiple change-point model, along with the range of values of the posterior probability distribution that includes 95% of the probability in brackets. Deviations are measured as the difference in quarters between the SBCDC turning points and those estimated by the model.

Figure 1: Business cycle SBCDC



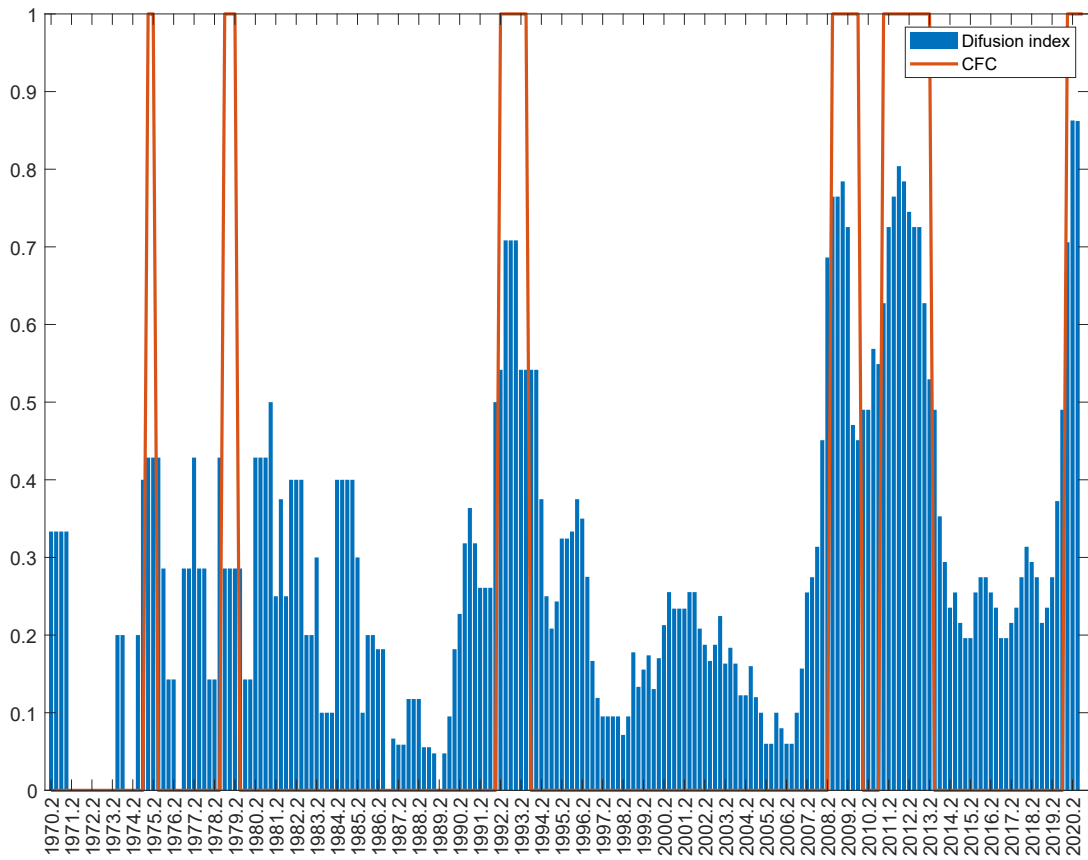
Notes: This Figure displays with grey shadow bars the recessions established by the Spanish Business Cycle Dating Committee (SBCDC). The vertical red lines represent the peaks and troughs detected by the BBQ applied to the GDP. Finally, the blue line shows the evolution of GDP.

Figure 2: Heat map of specific indicators



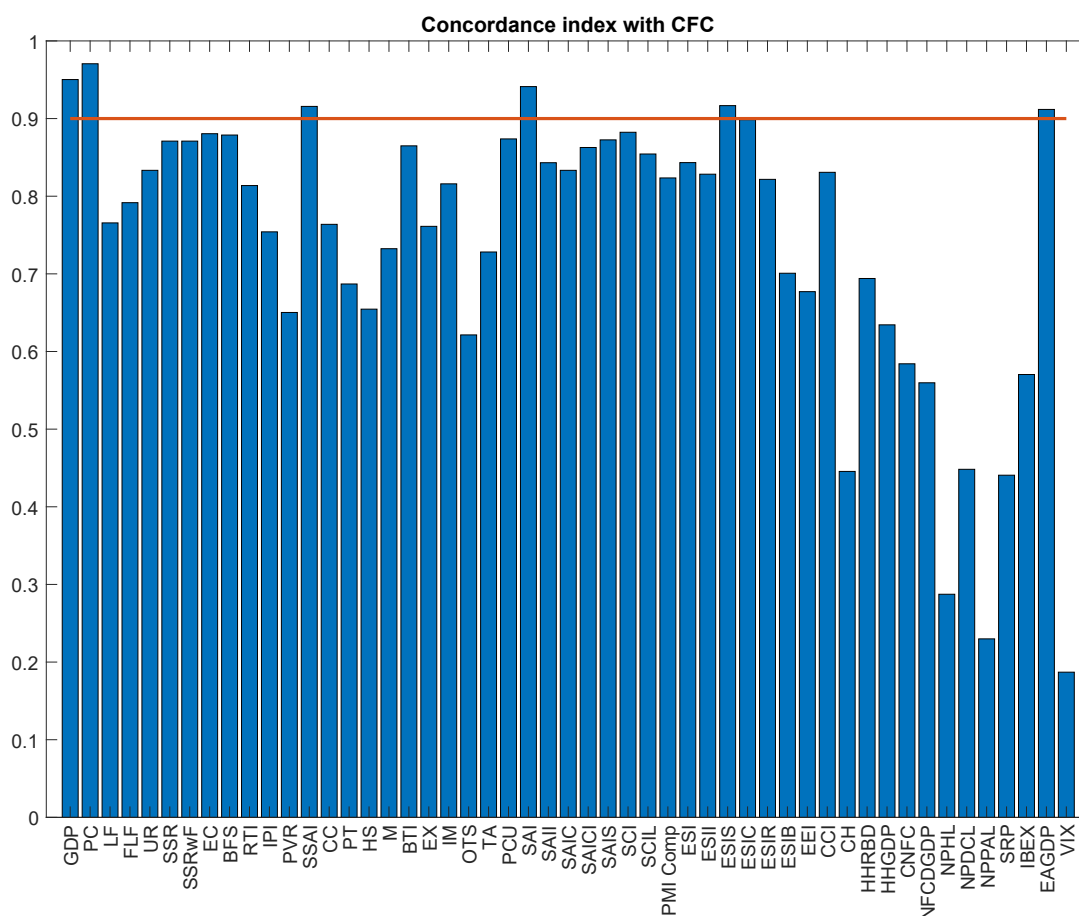
Notes: This Figure represents the periods of expansion (blue) and recession (garnet) for each specific indicator. The switch of the states has been determined from the turning points computed with the Bry-Boschan algorithm.

Figure 3: Diffusion index of specific indicators



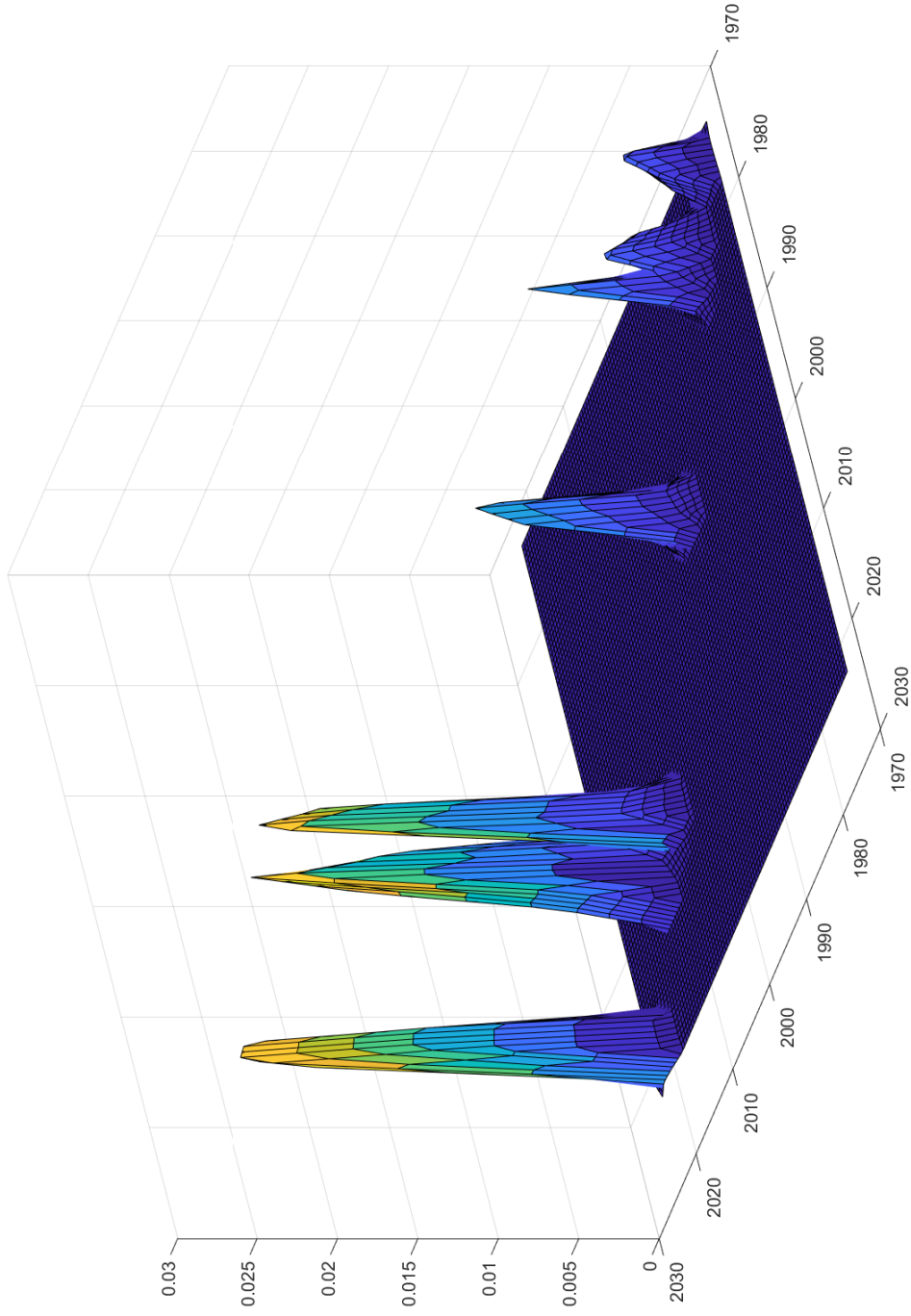
Notes: The blue bars show the percentage of series that are in recession in each period. The garnet boxes represent the recessions established by the Spanish Business Cycle Dating Committee (SBCDC).

Figure 4: Concordance index of specific indicators



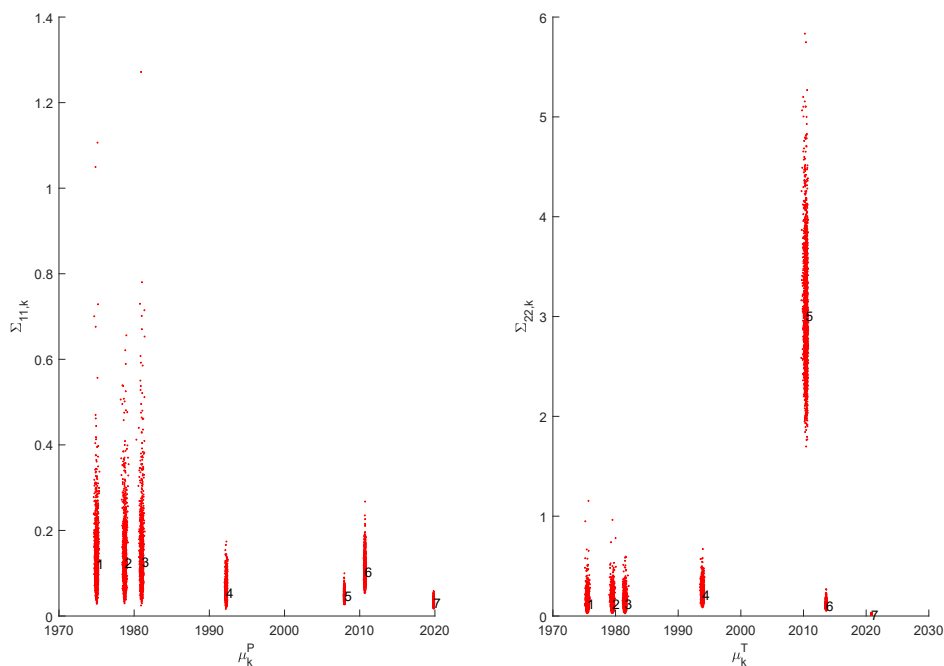
Notes: This Figure displays the concordance index of each indicator and the business cycle established by the Spanish Business Cycle Dating Committee (SBCDC). The garnet line corresponds with a value of 0.9.

Figure 5: Bivariate distribution of specific turning point dates

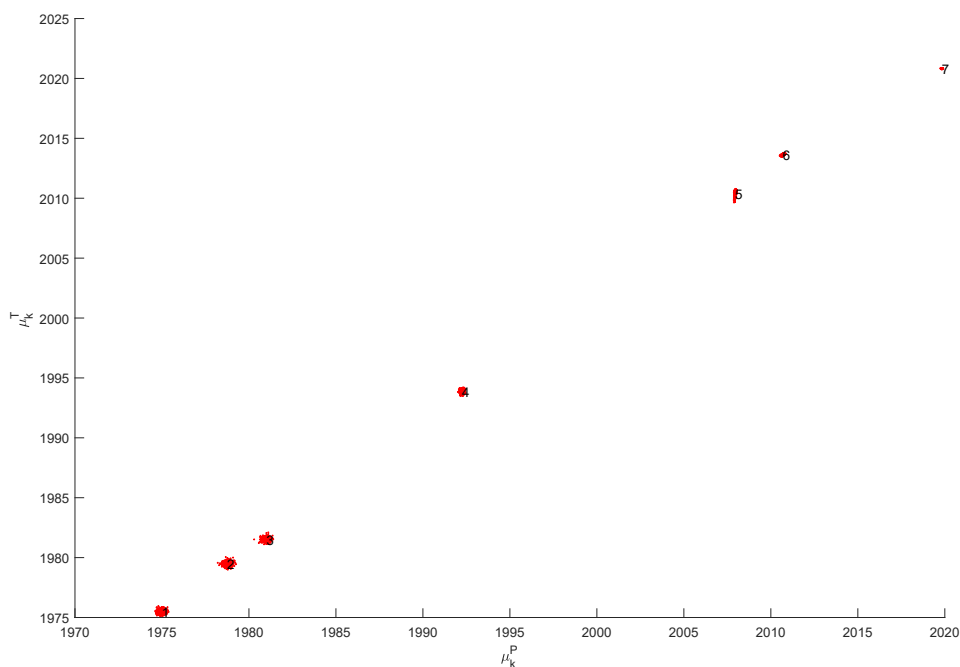


Notes: The figure displays the two-dimensional scatter plots of the MCMC draws for (μ_k^P, μ_k^T) , $(\mu_k^P, \Sigma_{11,k}^P)$ and $(\mu_k^T, \Sigma_{22,k}^T)$ for each of the $K = 8$ clusters.

Figure 6: Scatter plots of means and variances



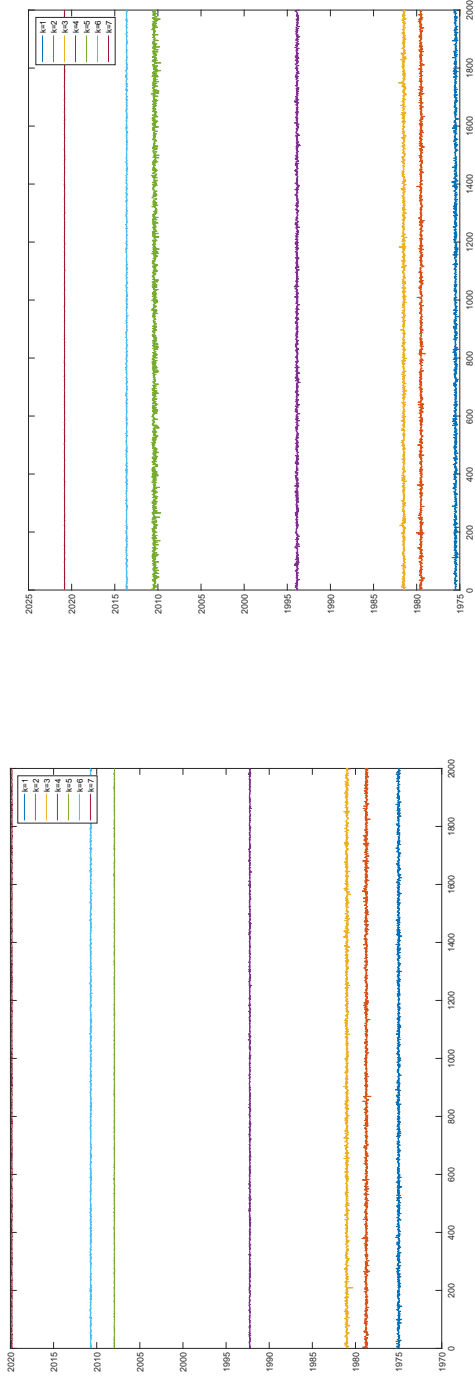
Panel A. Draws of $(\mu_k^P, \Sigma_{11,k})$ and $(\mu_k^T, \Sigma_{22,k})$



Panel B. Draws of (μ_k^P, μ_k^T)

Notes: The figure displays the two-dimensional scatter plots of the MCMC draws for (μ_k^P, μ_k^T) , $(\mu_k^P, \Sigma_{11,k})$ and $(\mu_k^T, \Sigma_{22,k})$ for each of the $K = 7$ clusters.

Figure 7: Behaviour of the Gibbs sampler



Panel A. Draws of μ_k^P

Panel B. Draws of μ_k^T

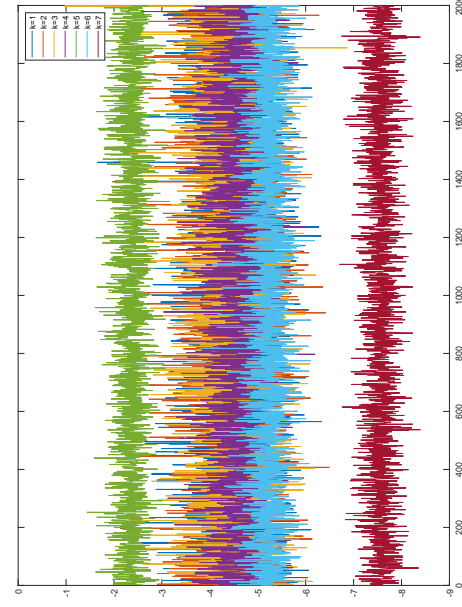


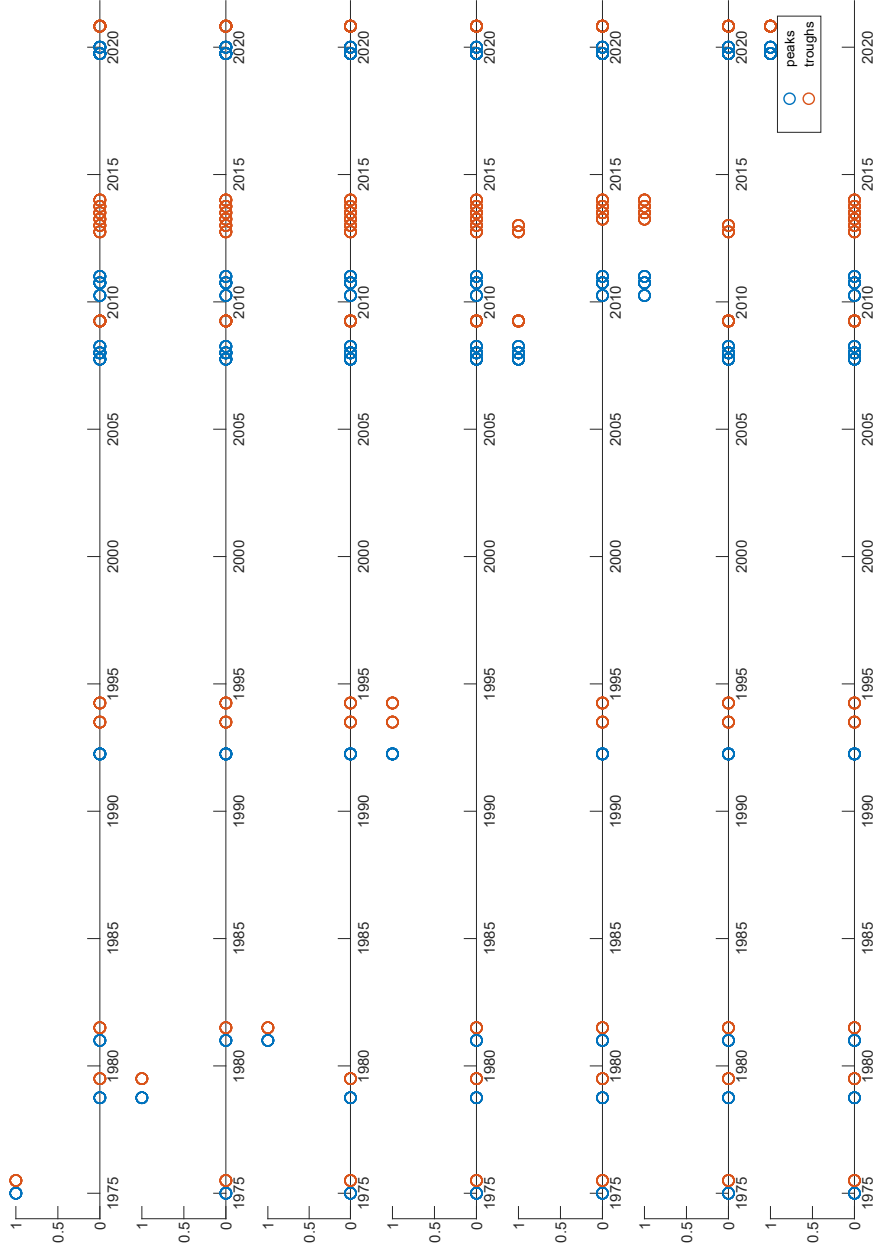
Figure 3: Behaviour of Gibbs sampler. Draws of transition matrix

Panel C. Draws of $\log(\Sigma_k)$

Panel D. Draws of $p_{i,j}$

Notes: MCMC draws for μ_k^P , μ_k^T , $\log(\Sigma_k)$ and $p_{i,j}$ for each of the $K = 7$ clusters.

Figure 8: Classification probabilities



Notes. For each $k = 1, \dots, 8$ and $i = 1, \dots, N$, the figure plots the estimated posterior classification probabilities $\Pr(s_i = k|\theta)$ on a timeline that spans from January 1959 to August 2010. The probabilities represent the average number of times that observation i is allocated to cluster k across the M MCMC replications.

6 Appendix

Table A1: Variable definitions

Variable	Acronym	Source	Sample
Gross Domestic Product	GDP	National Statistics Institute (INE)	1970:I-2020:II
Private consumption	PC	INE	1995:I-2020:II
Labor force	LF	INE	1972:III-2020:II
Female labor force	FLF	INE	1972:III-2020:II
Unemployment rate	UR	INE	1989:01-2020:07
Social Security registrations	SSR	Ministerio de Inclusión, Seguridad Social y Migraciones	1982:01-2020:09
Social Security registrations without workers on furlough	SSRwF	Ministerio de Inclusión, Seguridad Social y Migraciones	1982:01-2020:09
Electricity consumption	EC	Red eléctrica de España	1981:01-2020:08
Big firms sales	BFS	Agencia Tributaria	1996:01-2020:07
Retail trade index	RTI	INE	1995:01-2020:06
Industrial production index	IPI	INE	1975:01-2020:08
Private vehicles registration	PVR	Asociación Española de Fabricantes de Automóviles y Camiones (ANFAC)	1975:01-2020:07
Services sector activity index	SSAI	INE	2000:01-2020:07
Cement consumption	CC	Oficemen	1989:02-2020:05
New construction permits	PT	Ministerio de Transportes, movilidad y agenda urbana	1992:01-2020:07
Home sales	HS	INE	2007:01-2020:08
Mortgages	M	INE	2003:01-2020:07
Business turnover index	BTI	INE	2002:01-2020:05
Exports	EX	Departamento de Aduanas y Ministerio de Asuntos Económicos y Transformación Digital	1970:06-2020:05
Imports	IM	Departamento de Aduanas y Ministerio de Asuntos Económicos y Transformación Digital	1970:06-2020:05
Overnight tourist stays	OTS	INE	1995:01-2020:09
Tourist arrivals	TA	INE	1995:01-2020:09
Productive capacity utilization	PCU	Ministerio de Asuntos Económicos y Transformación Digital	1995:I-2020:III
Synthetic activity indicator	SAI	EDE Business	1995:01-2020:06
Synthetic activity indicator. Industry	SAII	EDE Business	1995:01-2020:06
Synthetic activity indicator. Construction	SAIC	EDE Business	1995:01-2020:06
Synthetic activity indicator. Construction investment	SAICI	EDE Business	1995:01-2020:06
Synthetic activity indicator. Services	SAIS	EDE Business	1995:01-2020:06
Synthetic consumption indicator	SCI	Ministerio de Asuntos Económicos y Transformación Digital	1995:01-2020:06
Synthetic consumption indicator. Large chain stores	SCIL	Ministerio de Asuntos Económicos y Transformación Digital	1995:01-2020:06
Composite produce manager index	PMI Comp	IHS Markit	1999:08-2020:06
Economic sentiment indicator	ESI	European Commission	1987:04-2020:10
Economic sentiment indicator. Industry	ESII	European Commission	1987:04-2020:08
Economic sentiment indicator. Services	ESIS	European Commission	1996:01-2020:08
Economic sentiment indicator. Consumption	ESIC	European Commission	1986:06-2020:08
Economic sentiment indicator. Retail	ESIR	European Commission	1988:01-2020:08
Economic sentiment indicator. Building	ESIB	European Commission	1989:01-2020:08
Employment expectations indicator	EEI	Eurostat	1996:01-2020:08
Consumer confidence index	CCI	Centro de investigaciones sociológicas (CIS)	2004:09-2020:10
Credit to households (% GDP)	CH	Banco de España	1995:IV-2020:IV
Credit to Non Financial Corporate (% GDP)	HHRBD	Banco de España	1995:IV-2020:IV
Ratio of households debt over disposable income	HHGDP	Banco de España	1987:I-2020:II
Ratio of households debt over GDP	CNFC	Banco de España	1987:I-2020:II
Ratio of non financial corporate debt over GDP	NFCDGDP	Banco de España	1987:I-2020:II
Ratio of non performing housing loans	NPHL	Banco de España	1998:IV-2020:II
Ratio of non performing durable consumption loans	NPDC	Banco de España	1998:IV-2020:II
Ratio of non performing production activities loans	NPPAL	Banco de España	1998:IV-2020:II
Sovereign risk premia ES-DE	SRP	Datastream	1991:07-2020:09
IBEX 35 index	IBEX	Datastream	1987:01-2020:09
EA Gross Domestic Product	EAGDP	Eurostat	1995:I-2020:II
VIX index	VIX	Datastream	1990:01-2020:09

Notes: The variables include either monthly (arabic numbers) or quarterly data (roman numbers).

BANCO DE ESPAÑA PUBLICATIONS

WORKING PAPERS

- 2020 VICTORIA IVASHINA, LUC LAEVEN and ENRIQUE MORAL-BENITO: Loan types and the bank lending channel.
- 2021 SERGIO MAYORDOMO, NICOLA PAVANINI and EMANUELE TARANTINO: The impact of alternative forms of bank consolidation on credit supply and financial stability.
- 2022 ALEX ARMAND, PEDRO CARNEIRO, FEDERICO TAGLIATI and YIMING XIA: Can subsidized employment tackle long-term unemployment? Experimental evidence from North Macedonia.
- 2023 JACOPO TIMINI and FRANCESCA VIANI: A highway across the Atlantic? Trade and welfare effects of the EU-Mercosur agreement.
- 2024 CORINNA GHIRELLI, JAVIER J. PÉREZ and ALBERTO URTASUN: Economic policy uncertainty in Latin America: measurement using Spanish newspapers and economic spillovers.
- 2025 MAR DELGADO-TÉLLEZ, ESTHER GORDO, IVÁN KATARYNIUK and JAVIER J. PÉREZ: The decline in public investment: "social dominance" or too-rigid fiscal rules?
- 2026 ELVIRA PRADES-ILLANES and PATROCINIO TELLO-CASAS: Spanish regions in Global Value Chains: How important? How different?
- 2027 PABLO AGUILAR, CORINNA GHIRELLI, MATÍAS PACCE and ALBERTO URTASUN: Can news help measure economic sentiment? An application in COVID-19 times.
- 2028 EDUARDO GUTIÉRREZ, ENRIQUE MORAL-BENITO, DANIEL OTO-PERALÍAS and ROBERTO RAMOS: The spatial distribution of population in Spain: an anomaly in European perspective.
- 2029 PABLO BURRIEL, CRISTINA CHECHERITA-WESTPHAL, PASCAL JACQUINOT, MATTHIAS SCHÖN and NIKOLAI STÄHLER: Economic consequences of high public debt: evidence from three large scale DSGE models.
- 2030 BEATRIZ GONZÁLEZ: Macroeconomics, Firm Dynamics and IPOs.
- 2031 BRINDUSA ANGHEL, NÚRIA RODRÍGUEZ-PLANAS and ANNA SANZ-DE-GALDEANO: Gender Equality and the Math Gender Gap.
- 2032 ANDRÉS ALONSO and JOSÉ MANUEL CARBÓ: Machine learning in credit risk: measuring the dilemma between prediction and supervisory cost.
- 2033 PILAR GARCÍA-PEREA, AITOR LACUESTA and PAU ROLDAN-BLANCO: Raising Markups to Survive: Small Spanish Firms during the Great Recession.
- 2034 MÁXIMO CAMACHO, MATÍAS PACCE and GABRIEL PÉREZ-QUIRÓS: Spillover Effects in International Business Cycles.
- 2035 ÁNGEL IVÁN MORENO and TERESA CAMINERO: Application of text mining to the analysis of climate-related disclosures.
- 2036 EFFROSINI ADAMOPOULOU and ERNESTO VILLANUEVA: Wage determination and the bite of collective contracts in Italy and Spain: evidence from the metal working industry.
- 2037 MIKEL BEDAYO, GABRIEL JIMÉNEZ, JOSÉ-LUIS PEYDRÓ and RAQUEL VEGAS: Screening and Loan Origination Time: Lending Standards, Loan Defaults and Bank Failures.
- 2038 BRINDUSA ANGHEL, PILAR CUADRADO and FEDERICO TAGLIATI: Why cognitive test scores of Spanish adults are so low? The role of schooling and socioeconomic background
- 2039 CHRISTOPH ALBERT, ANDREA CAGGESE and BEATRIZ GONZÁLEZ: The Short- and Long-run Employment Impact of COVID-19 through the Effects of Real and Financial Shocks on New Firms.
- 2040 GABRIEL JIMÉNEZ, DAVID MARTÍNEZ-MIERA and JOSÉ-LUIS PEYDRÓ: Who Truly Bears (Bank) Taxes? Evidence from Only Shifting Statutory Incidence.
- 2041 FELIX HOLUB, LAURA HOSPIDO and ULRICH J. WAGNER: Urban air pollution and sick leaves: evidence from social security data.
- 2042 NÉLIDA DÍAZ SOBRINO, CORINNA GHIRELLI, SAMUEL HURTADO, JAVIER J. PÉREZ and ALBERTO URTASUN: The narrative about the economy as a shadow forecast: an analysis using Banco de España quarterly reports.
- 2043 NEZIH GUNER, JAVIER LÓPEZ-SEGOVIA and ROBERTO RAMOS: Reforming the individual income tax in Spain.
- 2101 DARÍO SERRANO-PUENTE: Optimal progressivity of personal income tax: a general equilibrium evaluation for Spain.
- 2102 SANDRA GARCÍA-URIBE, HANNES MUELLER and CARLOS SANZ: Economic uncertainty and divisive politics: evidence from the *Dos Españas*.
- 2103 IVÁN KATARYNIUK, VÍCTOR MORA-BAJÉN and JAVIER J. PÉREZ: EMU deepening and sovereign debt spreads: using political space to achieve policy space.
- 2104 DARÍO SERRANO-PUENTE: Are we moving towards an energy-efficient low-carbon economy? An input-output LMDI decomposition of CO₂ emissions for Spain and the EU28.

- 2105 ANDRÉS ALONSO and JOSÉ MANUEL CARBÓ: Understanding the performance of machine learning models to predict credit default: a novel approach for supervisory evaluation.
- 2106 JAVIER ANDRÉS, ÓSCAR ARCE and PABLO BURRIEL: Market polarization and the Phillips curve.
- 2107 JUAN de LUCIO and JUAN S. MORA-SANGUINETTI: New dimensions of regulatory complexity and their economic cost. An analysis using text mining.
- 2108 DANILO LEIVA-LEON and LUIS UZEDA: Endogenous time variation in vector autoregressions.
- 2109 CORINNA GHIRELLI, ANDREA GONZÁLEZ, JOSÉ LUIS HERRERA and SAMUEL HURTADO: Weather, mobility and the evolution of the Covid-19 pandemic.
- 2110 KLODIANA ISTREFI, FLORENS ODENDAHL and GIULIA SESTIERI: Fed communication on financial stability concerns and monetary policy decisions: revelations from speeches.
- 2111 YUNUS AKSOY, HENRIQUE S. BASSO and CAROLYN ST AUBYN: Time Variation in Lifecycle Consumption and Income.
- 2112 JENNIFER PEÑA and ELVIRA PRADES: Price setting in Chile: micro evidence from consumer on-line prices during the social outbreak and Covid-19.
- 2113 NEZIH GUNER, YULIYA A. KULIKOVA and ARNAU VALLADARES-ESTEBAN: Does the added worker effect matter?
- 2114 RODOLFO G. CAMPOS and JACOPO TIMINI: Unequal trade, unequal gains: the heterogeneous impact of MERCOSUR.
- 2115 JAVIER QUINTANA: Import competition, regional divergence, and the rise of the skilled city.
- 2116 PATRICK MACNAMARA, MYROSLAV PIDKUYKO and RAFFAELE ROSSI: Marginal Tax Changes with Risky Investment.
- 2117 RODOLFO G. CAMPOS, JACOPO TIMINI and ELENA VIDAL: Structural gravity and trade agreements: does the measurement of domestic trade matter?
- 2118 ALESSANDRO MELCARNE, JUAN S. MORA-SANGUINETTI and ROK SPRUK: Democracy, technocracy and economic growth: evidence from 20 century Spain.
- 2119 ÁNGEL ESTRADA and DANIEL SANTABÁRBARA: Recycling carbon tax revenues in Spain. Environmental and economic assessment of selected green reforms.
- 2120 ALEJANDRO FERNÁNDEZ-CEREZO, BEATRIZ GONZÁLEZ, MARIO IZQUIERDO and ENRIQUE MORAL-BENITO: Firm-level heterogeneity in the impact of the COVID-19 pandemic.
- 2121 EDUARDO GUTIÉRREZ and CÉSAR MARTÍN MACHUCA: The effect of tariffs on Spanish goods exports.
- 2122 JACOPO TIMINI: Revisiting the 'Cobden-Chevalier network' trade and welfare effects.
- 2123 ALEJANDRO BUESA, JAVIER J. PÉREZ and DANIEL SANTABÁRBARA: Awareness of pandemics and the impact of COVID-19.
- 2124 ALICIA AGUILAR and DIEGO TORRES: The impact of COVID-19 on analysts' sentiment about the banking sector.
- 2125 SILVIA ALBRIZIO, IVÁN KATARYNIUK, LUIS MOLINA and JAN SCHÄFER: ECB euro liquidity lines.
- 2126 ANTHONY BALD, ERIC CHYN, JUSTINE HASTINGS and MARGARITA MACHELETT: The causal impact of removing children from abusive and neglectful homes.
- 2127 IRMA ALONSO, PEDRO SERRANO and ANTONI VAELO-SEBASTIÀ: The impact of heterogeneous unconventional monetary policies on the expectations of market crashes.
- 2128 MARÍA T. GONZÁLEZ-PÉREZ: Lessons from estimating the average option-implied volatility term structure for the Spanish banking sector.
- 2129 SIMÓN A. RELLA, YULIYA A. KULIKOVA, EMMANOUIL T. DERMITZAKIS and FYODOR A. KONDRASHOV: Rates of SARS-COV-2 transmission and vaccination impact the fate of vaccine-resistant strains.
- 2130 MATÍAS LAMAS and DAVID MARTÍNEZ-MIERA: Sectorial holdings and stock prices: the household-bank nexus.
- 2131 ALBERT BANAL-ESTAÑOL, ENRIQUE BENITO, DMITRY KHAMETSHIN and JIANXING WEI: Asset encumbrance and bank risk: theory and first evidence from public disclosures in Europe.
- 2132 ISABEL ARGIMÓN and MARÍA RODRÍGUEZ-MORENO: Business complexity and geographic expansion in banking.
- 2133 LUIS GUIROLA: Does political polarization affect economic expectations?: Evidence from three decades of cabinet shifts in Europe.
- 2134 CHRISTIANE BAUMEISTER, DANILO LEIVA-LEÓN and ERIC SIMS: Tracking weekly state-level economic conditions.
- 2135 SERGI BASCO, DAVID LÓPEZ-RODRÍGUEZ and ENRIQUE MORAL-BENITO: House prices and misallocation: The impact of the collateral channel on productivity.
- 2136 MANUEL ARELLANO, STÉPHANE BONHOMME, MICOLE DE VERA, LAURA HOSPIDO and SIQI WEI: Income risk inequality: Evidence from Spanish administrative records.
- 2137 ANGELA ABBATE and DOMINIK THALER: Optimal monetary policy with the risk-taking channel.
- 2138 MARTA BAÑBURA, DANILO LEIVA-LEÓN and JAN-OLIVER MENZ: Do inflation expectations improve model-based inflation forecasts?
- 2139 MÁXIMO CAMACHO, MARÍA DOLORES GADEA and ANA GÓMEZ LOSCOS: An automatic algorithm to date the reference cycle of the Spanish economy.