April 22, 2015

CF-Indicator: Technical note

1. Introduction

Business cycles are not directly observable. Moreover, there is not a widely accepted definition of what a business cycle is. Burns and Mitchell (1946), among the first to develop a comprehensive analysis of business cycles, argue that:

*Business cycles are a type of fluctuation found in the **aggregate economic activity** of nations that organize their work mainly in business enterprises: a cycle consists of expansions occurring **at about the same time** in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle; in duration, business cycles vary from more than one year to ten or twelve years; they are not divisible into shorter cycles of similar characteristics with amplitudes approximating their own*

Two key features of this definition seem particularly important. First, business cycles are viewed as fluctuations that affect "aggregate economic activity" and "many economic activities." Consequently, we consider a variety of economic indicators rather than just one. Second, the business cycle is apparent across economic activities that move synchronously. Thus, potential business cycle indicators should tend to move in step with one another.

Although the business cycle is not directly observable, one can infer its evolution from the common fluctuations in coincident economic indicators. Among these indicators, we include both broad economic aggregates and sectorial indicators with similar common coincident fluctuations. In addition, monitoring economic activity in real-time requires indicators available with short publication lags.

We incorporate these features to compute the CF-Indicator using the following set of economic indicators. At the aggregate level we include Gross Domestic Product (GDP) or Producto Interior Bruto (PIB in Spanish). In addition, we include sectorial economic indicators, such as the Industrial Production Index (IPI), the Services Sector Activity Index (SSAI) and the Social Security Registered rolls (SSR). Finally, we also include indicators based on surveys.

Surveys have the shortest publication delays and are useful to detect turning points quickly. Among these, we include the Purchasing Managers Index (PMI), the Economic Sentiment Indicator (ESI) and the Consumer Confidence Indicator (CCI).

Next, we infer aggregate economic activity fluctuations from the the common evolution of all these series. The econometric methods used to do this are characterized by: (1) being easily replicated and (ii) converting the process into an algorithm so that it can be routinely used to infer the evolution of the Spanish economy in real time.

The model we use is based on extensions of the seminal dynamic factor model proposed by Stock and Watson (1991). The main assumption of the model is that the evolution of each indicator depends on two elements. The first element is a factor common to all indicators and captures the evolution of the aggregate economic activity. This common factor can be interpreted as the business cycle. The second element captures idiosyncratic movements in the individual indicators.

However, the Stock-Watson approach cannot be directly applied in our context since the individual indicators we use have missing data for three reasons: (1) some indicators are available in recent times; (2) some indicators have long publication lags and (3) some indicators are available quarterly rather than monthly. Therefore we use maximum likelihood estimation with the Kalman filter as in Mariano and Murasawa (2003).

2. Econometric model

This is the most technical part of this note. Readers with less interest on the technical details in the construction of the indicator can skip this section. Readers with interested in the technical details are referred to the extensions of our model that appear in the literature review of Camacho, Pérez-Quirós and Poncela (2013).

2.1. Dynamic factor model

Let X_t be a vector of N economic indicators that may include N_1 quarterly indicators, $X^q_{t,1}$, and N_2 monthly indicators, $X^m_{t,2}$. Although we leave this assumption soon, let us start by assuming that all indicators are observed each month of the sample, including $X^m_{t,1}$, the monthly indicators that are behind the evolution of the indicators that are observed at quarterly frequency $X^q_{t,1}$.

We begin by showing how the quarterly indicators enters into the model. Assume that the log of $X^q_{t,1}$ have a unit root and call the first difference $Y^q_{t,1}$. Among others, Mariano and Murasawa (2003) show that if the sample average of the data within a given quarter can be approximated by the geometric mean, the quarterly growth rates, $Y^q_{t,1}$, can be expressed as a weighted average of past monthly growth rates, $Y^m_{t,1}$,

$$Y^q_{t,1} = \frac{1}{3}Y^m_{t,1} + \frac{2}{3}Y^m_{t-1,1} + Y^m_{t-2,1} + \frac{2}{3}Y^m_{t-3,1} + \frac{1}{3}Y^m_{t-4,1}. \quad (1)$$

Next, consider the monthly indicators. Since IPI and SSR have high seasonal components, they enter the model in annual growth rates. Survey-based indicators enter in levels. Therefore, we include in the vector $Y^m_{t,2}$ both annual growth rates of hard economic indicators and the levels of survey-based indicators.

We assume that the monthly evolution of individual indicators can be expressed as the sum of two components. The first component refers to the common factor, which captures the fact that the economic fluctuations occur about at the same time in a synchronous manner. Since we include different sectorial indicators, it makes sense to include a second component that captures idiosyncratic fluctuations.

Therefore, each individual indicator is the sum of the common component, f_t, and the idiosyncratic component, U_t:

$$Y_{t,1}^{m} = \beta_1 f_t + U_{t,1}, \tag{2}$$

in the case of quarterly indicators and

$$Y_{t,2}^{m} = \beta_2 \sum_{j=0}^{11} f_{t-j} + U_{t,2}. \tag{3}$$

in the case of monthly indicators.

In this expression, it is worth emphasizing two features. First, monthly indicators depend on a moving average of the common factor since they enter into the model in annual growth rates. Survey-based indicators depend on this moving average as well since the surveys are typically designed to maximize the correlation between the index and the reference series (European Commission, 2006). Second, individual indicators also depend on a moving average of the idiosyncratic component. However, it would increase the model's dimension to the extent that the computational cost would be much higher than the benefits in the quality of the estimation (Camacho and Pérez-Quirós, 2008).

To complete the dynamic specification of the model, assume that $U_{t,i}$, *i*=1,2, and f_t follow autoregressive processes of orders *p*1, *p*2 and *p*3, respectively:

$$U_{t,1} = A_{1,1}U_{t-1,1} + \ldots + A_{p1,1}U_{t-p1,1} + \varepsilon_{t,1}, \tag{4}$$

$$U_{t,2} = A_{1,2}U_{t-1,2} + \ldots + A_{p2,2}U_{t-p2,2} + \varepsilon_{t,2}, \tag{5}$$

$$f_t = b_1 f_{t-1} + \ldots + b_{p3} f_{t-p3} + \varepsilon_{t,f}, \tag{6}$$

where $A_{j,i}$ are diagonal matrices; $\varepsilon_{t,1}$ and $\varepsilon_{t,2}$ are Gaussian univariate random walks with zero mean and diagonal covariance matrices Ω_1 y Ω_2; $\varepsilon_{t,f}$ is a Gaussian random walk with zero mean and variance σ_f^2. All errors are serially uncorrelated in time series and cross section.

2.2. State-space representation

It is convenient to write the model in state-space form. Let $0_{a\times b}$ be a $(a\times b)$ matrix of zeroes and $1_{a\times b}$ be a $(a\times b)$ matrix of ones. Let I_a be the $(a\times a)$ identity matrix a matrix, v be the vector $\left(\frac{1}{3},\frac{2}{3},1,\frac{2}{3},\frac{1}{3}\right)'$ and $\otimes$ be the Kronecker product. Without loss of generality, we assume that all the autoregressive lag orders are of order one; that is *p*1=*p*2=*p*3=1.

The measurement equation, $Y_t = H\beta_t + E_t$, can be obtained from $Y_t = \left(Y_{t,1}^{q}{}', Y_{t,2}^{m}{}'\right)'$, $E_t \sim i.i.d.N(0,R)$, $R=0$,[1] $\beta_t = \left(f_t, f_{t-1},\ldots, f_{t-11}, U_{t,1}', U_{t-1,1}',\ldots,U_{t-4,1}', U_{t,2}'\right)'$ and

$$H = \begin{pmatrix} \frac{\beta_1}{3} & \frac{2\beta_1}{3} & \beta_1 & \frac{2\beta_1}{3} & \frac{\beta_1}{3} & 0_{N_1\times 7} & v'\otimes I_{N_1} & 0_{N_1\times N_2} \\ \beta_2 & \beta_2 & \beta_2 & \beta_2 & \beta_2 & \beta_2\otimes 1_{1\times 7} & 0_{N_2\times 5N_1} & I_{N_2} \end{pmatrix}, \tag{7}$$

where β_1 y β_2 are vectors contain the loading factors of the economic indicators.

The transition equation is $\beta_t = F\beta_{t-1} + V_t$. If $\bar{b}_1 = (b_1, 0_{1\times 10})$ and

$$B = \begin{pmatrix} \bar{b}_1 & 0 \\ I_{11} & 0_{11\times 1} \end{pmatrix}, \tag{8}$$

The matrix *F* becomes

[1] Note that E_t is fixed when there are no missing data. However, defining E_t in this way facilitates the treatment of the model when handling missing data.

$$F=\begin{pmatrix} B & 0_{12\times N_1} & 0_{12\times N_1} & 0_{12\times N_1} & 0_{12\times N_1} & 0_{12\times N_1} & 0_{12\times N_2} \\ 0_{N_1\times 12} & A_{1,1} & 0_{N_1\times N_1} & 0_{N_1\times N_1} & 0_{N_1\times N_1} & 0_{N_1\times N_1} & 0_{N_1\times N_2} \\ 0_{N_1\times 12} & I_{N_1} & 0_{N_1\times N_1} & 0_{N_1\times N_1} & 0_{N_1\times N_1} & 0_{N_1\times N_1} & 0_{N_1\times N_2} \\ 0_{N_1\times 12} & 0_{N_1\times N_1} & I_{N_1} & 0_{N_1\times N_1} & 0_{N_1\times N_1} & 0_{N_1\times N_1} & 0_{N_1\times N_2} \\ 0_{N_1\times 12} & 0_{N_1\times N_1} & 0_{N_1\times N_1} & I_{N_1} & 0_{N_1\times N_1} & 0_{N_1\times N_1} & 0_{N_1\times N_2} \\ 0_{N_1\times 12} & 0_{N_1\times N_1} & 0_{N_1\times N_1} & 0_{N_1\times N_1} & I_{N_1} & 0_{N_1\times N_1} & 0_{N_1\times N_2} \\ 0_{N_2\times 12} & 0_{N_2\times N_1} & 0_{N_2\times N_1} & 0_{N_2\times N_1} & 0_{N_2\times N_1} & 0_{N_2\times N_1} & A_{1,2} \end{pmatrix}. \tag{9}$$

Finally, V_t are serially uncorrelated Gausian errors with zero mean and covariance matrix $Q=diag\left(\left(\sigma_f^2, 0_{1\times 11}, \overline{\Omega}_1', 0_{1\times 4N_1}, \overline{\Omega}_2'\right)'\right)$, where $\overline{\Omega}_i$ is the vector that contains the main diagonal elements of Ω_i, *i*=1,2.

2.3. Estimation and signal extraction

The Kalman filter can be used to estimate the model's parameters by maximum likelihood and to compute inferences on the unobserved components. Starting the filter with $\beta_{0|0}$ and its covariance matrix $P_{0|0}$, the prediction equations become

$$\beta_{t+1|t}=F\beta_{t|t} \tag{10}$$

$$P_{t+1|t}=FP_{t|t}F'+Q\,. \tag{11}$$

These equations can be used to compute the prediction error its covariance matrix

$$\eta_{t+1|t}=Y_t-F\beta_{t+1|t} \tag{12}$$

$$\Delta_{t+1|t}=FP_{t|t}F', \tag{13}$$

which can be used to evaluate the log likelihood

$$l_t=-\frac{1}{2}\left[\ln\left(2\pi\left|\Delta_{t|t}\right|\right)+\eta_{+1|t}'\left(\Delta_{t|t}\right)^{-1}\eta_{+1|t}\right]. \tag{14}$$

Finally, the vector of unobserved components and its covariance matrix are updated conveniently from the transition equations

$$\beta_{t+1|t+1} = \beta_{t+1|t} + P_{t+1|t} H' \left(\Delta_{t+1|t} \right)^{-1} \eta_{t+1|t} \tag{15}$$

$$P_{t+1|t+1} = P_{t+1|t} - P_{t+1|t} H' \left(\Delta_{t+1|t} \right)^{-1} H P_{t+1|t} \tag{16}$$

2.4. Dealing with missing data

So far, we have worked under the unrealistic assumption that all the indicators start from the same date, that the indicators exhibit synchronous publication calendars and that the quarterly indicators are available each month. Although these assumptions are too restrictive, the model can easily be extended to account for these data irregularities.

The solution to this data problem is simple. Missing data are replaced by realizations from a random variable $r_t \sim N\left(0, \sigma_r^2\right)$ that cannot depend on the model parameters. The measurement equation can be expressed in such a way that the unobserbed data are skipped from the Kalman recursions. In particular, let Y_{it} be the i-th element of the vector Y_t and let R_{ii} be its covariance matrix. Let H_{it} be the i-th row of H_t, which has z columns. The measurement equation is transformed as follows

$$Y_{it}^{+} = \begin{cases} Y_{it} & \text{if we observe } Y_{it} \\ r_t & \text{otherwise} \end{cases}, \tag{17}$$

$$H_{it}^{+} = \begin{cases} H_i & \text{if we observe } Y_{it} \\ 0_{1\times z} & \text{otherwise} \end{cases}, \tag{18}$$

$$E_{it}^{+} = \begin{cases} E_{it} & \text{if we observe } Y_{it} \\ r_t & \text{otherwise} \end{cases}, \tag{19}$$

$$R_{iit}^{+} = \begin{cases} R_{ii} & \text{if we observe } Y_{it} \\ \sigma_r^2 & \text{otherwise} \end{cases} . \qquad (20)$$

With this transformation, the standard state-space model is transformed into a time-varying state-space model for which the estimation method described in the previous version applies. Therefore, this transformed model can be estimated in the standard way to compute inferences on the unobserved values of Y_{it}^{+}, H_{it}^{+}, E_{it}^{+} and R_{iit}^{+}.

3. Empirical results

The data used to write this note were downloaded on March, 1st 2015 and the effective sample starts in January 1983 and ends in February 2015. The set of business cycle indicators, their transformations, the source and the publication lags are described in Table 1. GDP enters into the model in quarterly growth rates, the monthly hard indicators enter into the model in annual growth rates and the survey-based indicators enter into the model in levels.[2]

The selection of indicators follow the lines described above. The indicators must follow a common synchronous pattern along with their sectorial idiosyncratic dynamics. Since some indicators should refer to economic aggregates, we start the analysis with Gross Domestic Product, probably the main economic aggregate in industrialized economies. In particular, we use quarterly GDP at market prices, basis 2010, seasonally and calendar adjusted. According to the Quarterly National Accounts of the Spanish National Statistical Institute (INE), the sample starts in the first quarter of 1995. However, we extended back the sample to the first quarter of 1983 by using the quarterly growth rates provided by INE.[3]

Among the monthly hard indicators available for the Spanish economy, we use Industrial Production Index because it his highly correlated with the business cycle. In

[2] Before estimating the model, the variables are standardized to have a zero mean and a variance equal to one.
[3] These data are available at
http://www.ine.es/jaxiBD/tabla.do?per=03&type=db&divi=CNTR&idtab=9

particular, we use the non-seasonally adjusted version of the series because it was available since January 1975.[4] To avoid seasonality problems, IPI enters into the model in annual growth rates. In addition, we include in the model Social Security Registered rolls (SSR) due to the enormous economic role played by employment since in the recent years. The time series is available since January 1982 and enters into the model in annual growth rates.

The indicators based on surveys have the important advantage that they have a very short publication delay. According to the European Commission (2006), these indicators are designed to maximize their correlation with the annual growth rates of the reference series. Therefore, the soft indicators enter into the model in levels although they depend on a moving average of the common factor as if they entered into the model in annual growth rates.

The first soft indicator is the Markit Purchasing Managers Index (PMI). This indicator shows the composite index of manufactures and services and is available since August 1999. Markit uses purchasing managers to produce data on business conditions and computes the index in a way such that values above 50 indicate a change for the better.

The second soft indicator is the INE Services Sector Activity Index (IASS). This index is computed to capture the short-run evolution of firms that produce services in Spain, which is the most important sector (50% of GDP and 43% of employment).[5] In particular, we use the version of the index that is free of seasonality and calendar effects.

The third soft indicator is the European Commission Economic Sentiment Indicator (ESI). This is a composite index based on manufacturing (40%), services (30%), consumers (20%), retail trade (5%) and construction (5%). The index is available since April 1987.

The last soft indicator is the European Commission Consumer Confidence Indicator (CCI), which is available since June 1986. This European Commission indicator is composed of consumers' opinions on their own financial situation, the economic situation in general,

[4] The seasonally and calendar adjusted version of IPI is available only since January 1993.
[5] The sectors included in SSAI represent almost 70% of the value added by total services.

willingness to save and employment in the twelve months to come. The series are seasonally adjusted.

Figure 1 shows the evolution of the individual indicators used to compute the CF-Indicator, which are transformed as they enter into the model. To facilitate graphing, we include in the charts shaded areas that refer to the recessions determined by the Spanish Business Cycle Dating Committee at quarterly frequency. In line with the assumptions of the model, all the indicators share a common dynamic factor in addition to idiosyncratic dynamics. Notably, the common dynamic factor grows in expansions and falls in recessions.

The maximum likelihood estimates of the loading factors, which measure the correlation between the economic indicators and the common factor, appear in Table 2 along with their standard deviations. They show that they are positive and statistically significant. That is, the indicators are procyclical.

Figure 2 plots the common factor and the shaded areas that represent the CF-referenced recessions. To facilitate its economic interpretation, the CF-Indicator has been normalized with the threshold that marks the business cycle phase changes, following the procedure described in Berge and Jordà (2011). The evolution of the indicator agrees well with dating of expansions and recessions. During periods classified as expansions the CF-Indicator is positive, and it is negative in recessions. The indicator signals peaks by sudden falls from positive to negative readings and it shows troughs as rapid changes from negative to positive values. Accordingly, we think that the CF-Indicator is a useful tool to monitor the evolution of Spanish aggregate economic activity in real time.

4. Conclusions

This note describes the monthly CF-Indicator of the Spanish economic activity. The note has two main features. First, the note is self-contained so that all interested readers can replicate the results. Second, all the assumptions used in the computation of the index are supported by the data.

The CF-Indicator will be updated (almost) every month. We hope it to be useful as a way to monitor economic activity in Spain in real-time.

Acknowledgements

The Spanish Business Cycle Dating Committee is grateful to the Spanish Economic Association for its support.

References

Berje, T., y Jorda, O. 2011. Evaluating the classification of economic activity into recessions and expansions. American Economic Journal. Macroeconomics 3: 246-277

Camacho M, y Perez Quiros G. 2010. Introducing the Euro-STING: Short Term INdicator of euro area Growth. *Journal of Applied Econometrics* 25: 663-694.

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

Comisión Europea, 2006. The Joint Harmonised EU Programme of Business and Consumer Surveys. Directorate General for Economic and Financial Affairs Special Report 5/2006.

Mariano, R., y Murasawa, Y. 2003. A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics* 18: 427-443.

Table 1: List of business cycle indicators

Indicator	Sample	Source	Delay	Transform.
Gross Domestic Product (GDP, 2010 basis. Seasonally and calendar adjustment)	1983T1 2014T4	INE	2	TCT
Industrial Production Index (IPI, 2010 basis)	1983M01 2014M12	INE	1	TCA
Social Security Registrations (SSR, End-of-month all registrations. Seasonally adjusted)	1983M01 2015M01	MTIN MEH	0	TCA
Services Sector Activity Index (SSAI, total index Seasonally and calendar adjustment)	2001M01 2014M12	INE	2	Nivel
Purchasing Managers Index (PMI, composed: manufactures and services)	1999M08 2015M01	BEA	0	Nivel
Economic Sentiment Indicators (ESI, total)	1987M04 2015M02	European Commission	0	Nivel
Consumer Confidence Indicator (CCI, seasonally adjusted)	1962.M01 2015.M02	European Commission	0	Nivel

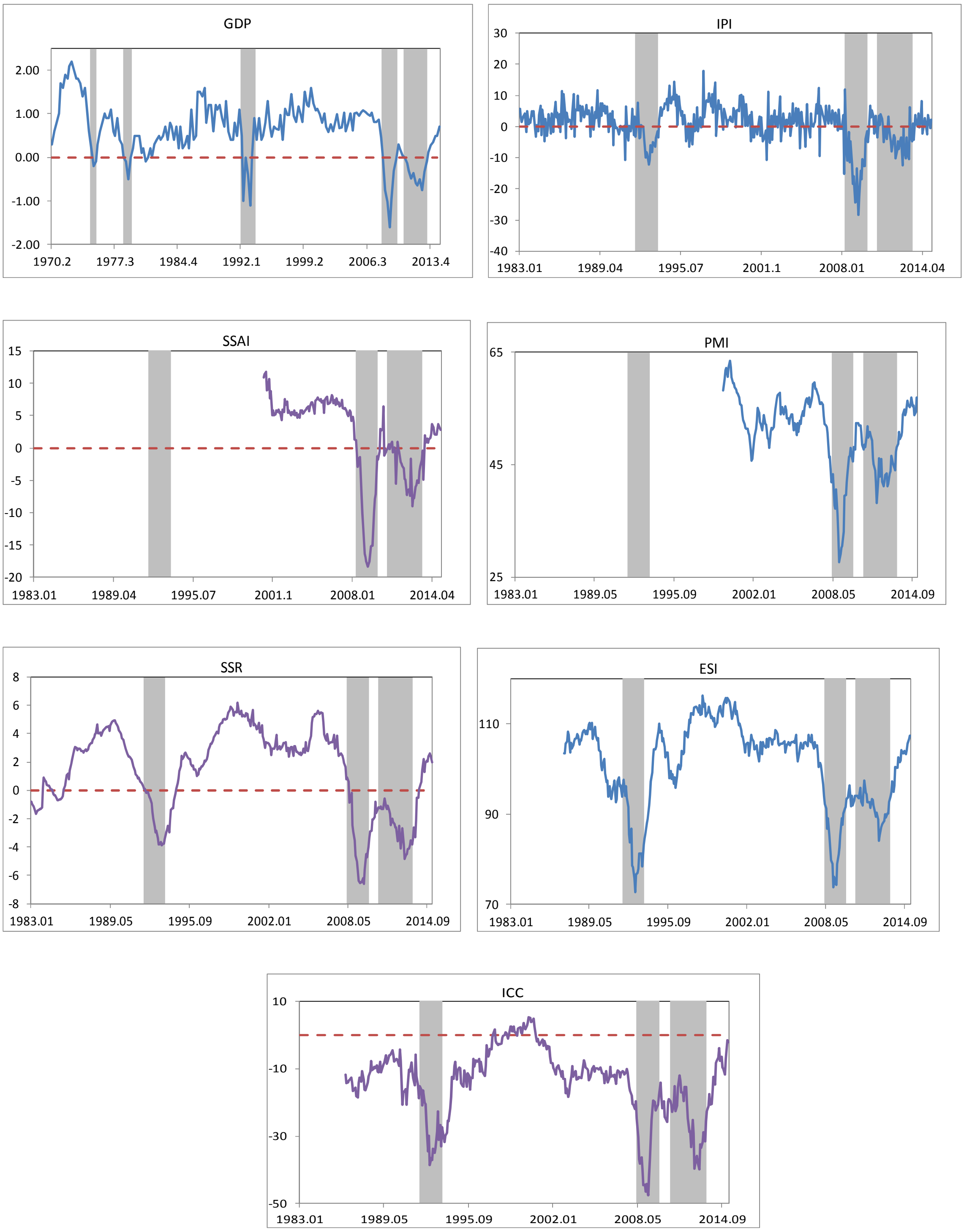
Note. The second column refers to the effective sample (data downloaded on March, 1st 2015). The delay refers to the month between the reference period and the publication.

Table 2: Loading factors

GDP	IPI	SSR	IASS	PMI	ESI	ICC
0.108	0.022	0.031	0.027	0.030	0.030	0.027
(0.019)	(0.004)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)

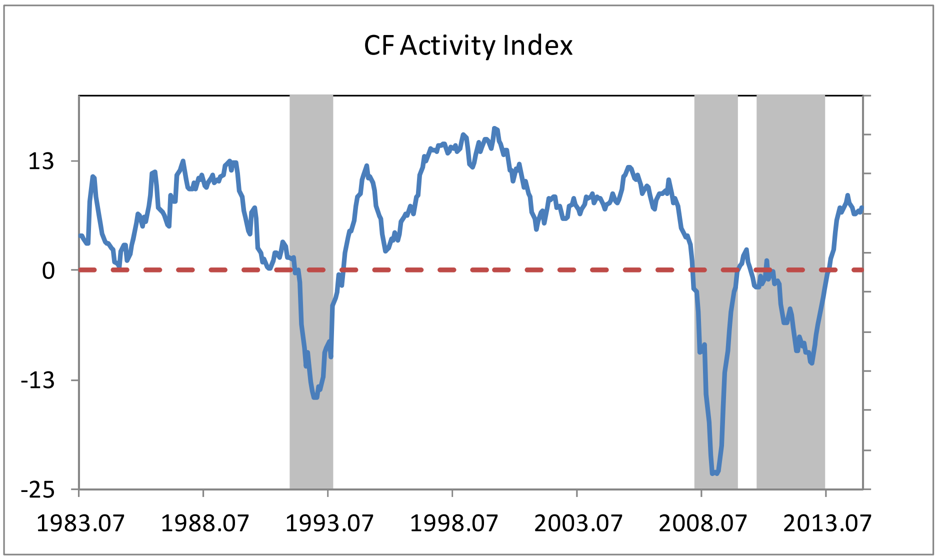
Note. The loading factors (standard deviations in brackets) refer to the degree of comovement with the business cycle.

Figure 1. Indicators and business cycle



Note. The indicators are described in Table 1. Shaded areas refer to the recessions determined by the Dating Committee of the Spanish business cycle at quarterly frequency.

Figure 2. CF Activity Index



Note. The figure plots the common cycle. Shaded areas refer to the recessions determined by the Dating Committee of the Spanish business cycle at quarterly frequency.