

Evaluating early warning and coincident indicators of business cycles using trends. Do forecasts combinations help?

Marcos Bujosa¹, Antonio García-Ferrer*², Aránzazu de Juan² and Antonio Martín-Arroyo²

¹Universidad Complutense de Madrid, Madrid, Spain. Instituto Complutense del Análisis Económico *ICAE*, Madrid, Spain

²Universidad Autónoma de Madrid, Madrid, Spain

First Draft, December 2017

Abstract

We present two composite coincident and leading indicators designed to capture the state of the Spanish economy and to provide reliable statistical forecasting power, respectively. Our approach, based on trends, guarantees that the resulting indicators are reasonably smooth and issue stable signals, reducing the uncertainty.

The coincident indicator has been checked by comparing it with the one recently proposed by the Spanish Economic Association index. Both indexes show similar behavior and ours captures very well the beginning and end of the official recessions and expansion periods (whereas the leading indicator systematically predicts the peaks and troughs). Our coincident indicator also tracks very well alternative mass media indicators typically used in the political science literature.

The leading indicator provides significant aid in forecasting annual GDP growth rates, and its one-step-ahead forecasts shows improvements over other alternatives.

1 Introduction

Sound leading and coincident indicators of business cycles are essential components for firms, investors and policy makers. Coincident Indicators are designed to capture the present state

*Corresponding author

of the economy or of its global business cycle, while Leading Indicators should be able to show reliable statistical forecasting power. Not surprisingly, economists have devoted a large amount of research in the quest for such indicators following the early works of [Burns and Mitchell \(1946\)](#) for the US economy. This intensive research has produced a vast amount of findings with both, theoretical and empirical implications as well as additional requirements for the indexes to fulfill. Leading indicators, for instance, should systematically provide a precise indication of the future course of the economy (consistency) and the signals need to arrive early enough so prospective policy decisions have time to be effective timeliness. On the other hand, coincident indicators should be able to reproduce the present state of the economy without producing false turning points signals too frequently. The stability of signals is also an important additional requirement that the literature to date has largely overlooked ([Drehmann and Juselius, 2014](#)). Indicators that issue stable or persistent signals reduce the uncertainty regarding trends and avoid confusion for economic agents in interpreting future directions of change.

At present, there is a large amount of literature on how to design coincident and leading indicators. Methods range from *ad-hoc* weighted averages of the time series of observed data to model-based methodologies. In the first approach, the optimality properties of the index are unknown so its usefulness is very limited. Within the model-based approach, however, the methods of diffusion index forecast ([Stock and Watson, 2002](#)) and the other variants of Dynamic Factor Models (DFM) have been able to incorporate information from a large number of predictors into the forecast in a simple and parsimonious way. A practical question in this approach, however, is how much data are really needed? In other words, how to find the best way to extract a subset of variables from a larger data set and, how to use it for real-time forecasting?

Justification for using a very large number of variables has been solely based on statistical properties of final estimates. For instance, if the DFM is estimated using principal components, the number of variables included in the model needs to be large to achieve consistency ([Bai and Ng, 2002](#)). Therefore, a large number of papers in the literature tend to include hundreds of variables without offering a systematic good record of forecasting performance ([Stock and Watson, 2003](#)). However, [Poncela and Ruiz \(2015\)](#) showed that when the DFM is directly estimated from the Kalman Filter equations, no more than a small number of variables is needed to achieve consistency. They also showed that, when model parameters have to be estimated, the parameter and total uncertainties could increase when the number of indicators increases. The related question is then: do we really need consistency when our main goal is forecasting? In this regard, [García-Ferrer and Poncela \(2002\)](#), [Boivin and Ng \(2006\)](#) and [Poncela and García-Ferrer \(2014\)](#) among others, found that expanding the sample size adding data that bear little information about the factor components does not necessarily improve forecasts. Similar results are found in [Álvarez, Camacho, and Pérez-Quirós \(2016\)](#) regarding the use of aggregated versus disaggregated data.

This last issue (reduction search) is central when trying to deal with the permanent

disease that afflicts all real forecasting exercises with non-experimental data: too many predictors without enough data. [Leamer \(2012\)](#) offers an interesting proposal by acknowledging that old and new methods to deal with the overparametrization problem can work well sometimes but not always. Understanding the economic circumstances where each approach is successful becomes a crucial starting point. Then it is the context that determines which procedure to use.¹

Independently of the method used, either spectral methods ([Altissimo, Cristadoro, Forni, Lippi, and Veronese, 2010](#)) or principal components ([Stock and Watson, 2002](#)), estimation of DFM is based on two features: the assumption of stationarity and the use of seasonally adjusted data. Both of them may have potential problems. Because economic data are nonstationary, authors prefilter all series to make each one plausibly stationary by taking first or second differences. But getting rid of nonstationarity by differencing individual series, when the series are cointegrated, throws away vast amount of information and may distort inference ([Corona, Poncela, and Ruiz, 2017](#); [Sims, 2012](#)). The issue of using seasonally adjusted data is also open to controversy [[Ghysels, Osborn, and Rodrigues \(2006\)](#), [Matas-Mir, Osborn, and Lombardi \(2008\)](#)]. In recent years several researchers have found the presence of residual seasonality of the US real gross domestic product (GDP) and other US macroeconomic variables [[Bujosa and García-Ferrer \(2014\)](#)]. In an attempt to reduce the residual seasonality, the US Bureau of Economic Analysis (BEA) revised GDP during the period 2013–2015, seasonally adjusting more of the input data in the aggregated series. In spite of these adjustments, [Phillips and Boldin \(2017\)](#) find that the first quarter data were still on average 0.6% too weak².

For the above reasons; in this paper we will be using original (or log transforms) monthly seasonally unadjusted data to build and analyze coincident and leading indicators for the Spanish business cycles from 1982 to 2014. Our goals in this paper are threefold. Firstly, we will obtain a Composite Coincident Indicator (CCI) using monthly targeted predictors and dynamic factor models with the aim of reproducing the official dating of Spanish business cycles and its relation with mass media indexes. Secondly, and using the same methodology, we will obtain a Composite Leading Indicator (CLI) by updating the results from [Bujosa, García-Ferrer, and de Juan \(2013\)](#) who successfully anticipate the onset of Spanish recessions. Finally, we will evaluate our CLI in comparison with alternative independent forecasts of Spanish GDP for the period 2001–2016, both in one to one basis as well as with alternative forecast combinations.

¹When analysing US business cycles a few years ago, [Leamer \(2009\)](#) found that for macroeconomic variables the borderline between features that repeat and features that do not repeat is constantly changing and, how the contribution to GDP growth of certain economic indicators was radically different during the expansions and during recessions. This empirical finding allowed him to use the so-called *cycle drivers* that systematically anticipated a large percentage of the US recessions. Interestingly, this finding was solely based on detailed examination and monitoring of disaggregated macroeconomic data using very simple methods.

²Prior to the revision in mid-2016 directly seasonally adjusting the official seasonally adjusted GDP would revise the first quarter growth since 2013 upward by an average of about 1.5%.

The paper is organized as follows. Section 2 describes the methodology. Section 2.1 describes the Linear Dynamic Harmonic Regression (LDHR) model to estimate the trends of the individual predictors while Section 2.2 outlines how they can be stacked in the DFM for non-stationary time series. Section 3 is devoted to the construction of our coincident indicator and compares it with the official CF index recently proposed by Spanish Economic Association (AEE) and with other mass media indicators. Section 4 is devoted to the leading indicator construction and its historical leadership regarding business cycle activity. Section 5 assess forecast performance of alternative forecasters including optimal forecast combinations. Finally, Section 6 concludes.

2 Methodology

2.1 Linear Dynamic Harmonic Regression

The Dynamic Harmonic Regression (DHR) model, developed by Young, Pedregal, and Tych (1999), is based on a spectral approach, under the hypothesis that the observed time series \mathbf{y} is periodic or quasi-periodic³ and can be decomposed into several components whose variances are concentrated around certain frequencies: e.g. at a fundamental frequency and its associated sub-harmonics. This model is often appropriate for observed time series with well defined spectral peaks, which implies that variance is concentrated around narrow frequency bands. The DHR model is the sum of several Unobserved Components:

$$\mathbf{y} = \sum_{j=0}^R \mathbf{s}^j + \mathbf{e} \quad (1)$$

where the irregular component, \mathbf{e} is a stationary sequence $\mathbf{e} \equiv \{e_t\}_{t \in \mathbb{Z}}$ of random variables with zero mean and variance σ_e^2 ; and each DHR component $\mathbf{s}^j \equiv \{s_t^j\}_{t \in \mathbb{Z}}$ is an oscillatory processes with form

$$s_t^j = a_t^j \cos(\omega_j t) + b_t^j \sin(\omega_j t). \quad (2)$$

Hence, the frequency ω_j is associated to the j -th component. Usually $j = 0$ corresponds to the trend (or zero frequency term) and the other components ($j = 1, \dots, R$) correspond to the seasonal frequency and its harmonics. Oscillations of each DHR component, \mathbf{s}^j , are modulated by two stochastic processes. Both stochastic processes, $\mathbf{a}^j \equiv \{a_t^j\}_{t \in \mathbb{Z}}$ and $\mathbf{b}^j \equiv \{b_t^j\}_{t \in \mathbb{Z}}$, follow the same AR(1) or AR(2) model with at least one root on the unit circle. The complete DHR model is

$$y_t = \sum_{j=0}^R \{a_t^j \cos(\omega_j t) + b_t^j \sin(\omega_j t)\} + e_t \quad (3)$$

³By “quasi-periodic” we mean that the amplitude and the phase of the periodicity may vary over time.

This model can be considered a straightforward extension of the classical harmonic regression model, in which the gain and phase of the harmonic components can vary randomly due to the stochastic processes \mathbf{a}^j and \mathbf{b}^j . Bujosa, García-Ferrer, and Young (2007) showed that DHR components \mathbf{s}^j have an equivalent representation as ARIMA processes with some AR roots on the unit circle.

$$\phi^j * \mathbf{s}^j = \theta^j * \mathbf{w}^j, \quad j = 0, \dots, R; \quad (4)$$

where, $\mathbf{w}^j \equiv \{w^j\}_{t \in \mathbb{Z}}$ is a white noise process, and where ϕ^j and θ^j are, respectively, the AR and MA polynomials. We say that the trend component ($j = 0$) follows a IRW model when ϕ^0 is a two degree polynomial with two unit roots.

Hence, DHR model is a sum of non-stationary stochastic signals plus noise. This is a standard approach in the literature of unobserved component models and seasonal adjustment of economic time series. The non-stationary signal extraction problem in this framework (including its spectral approach) has been studied since the sixties (see Bell, 1984; Hannan, 1967; Pierce, 1979; Sobel, 1967; Tiao and Hillmer, 1978). Nevertheless, and despite the widespread use of the term *pseudo-spectrum* in this literature, there was no proper definition of *pseudo-spectra* in this context⁴ until recently. As a matter of fact, Bujosa, Bujosa, and García-Ferrer (2015) provide the first formal justification of a corner stone in this literature: “the pseudo-spectrum of a sum of signals is the sum of the pseudo-spectra of the signals”. Hence, it is easy to see that the *pseudo-spectrum* of the DHR model is the sum of the *pseudo-spectra* of its components:

$$f_{dhr}(\omega, \boldsymbol{\sigma}^2) = \sum_{j=1}^R \sigma_j^2 \frac{\theta^j(e^{-i\omega})\theta^j(e^{i\omega})}{\phi^j(e^{-i\omega})\phi^j(e^{i\omega})} + \sigma_e^2; \quad (5)$$

where the entries in the variance vector $\boldsymbol{\sigma}^2 = [\sigma_0^2, \dots, \sigma_R^2, \sigma_e^2]$ are the unknown hyper-parameters of the model. Young et al. (1999) propose to fit this model in the frequency domain by the vector $\tilde{\boldsymbol{\sigma}}^2$ that minimizes the euclidean distance⁵

$$\min_{\tilde{\boldsymbol{\sigma}}^2 \in \mathbb{R}^{R+2}} \left\| \widehat{f_z(\omega)} - f_{dhr}(\omega, \tilde{\boldsymbol{\sigma}}^2) \right\|, \quad (6)$$

where $\widehat{f_z(\omega)}$ is the estimated AR-spectrum of the observed time series. This strategy has an intuitive appeal but, since DHR components follow non-stationary ARMA processes, the corresponding *pseudo-spectra* have poles; and therefore the norm is not defined. So they take logs in (6) and truncate the log-functions in the neighborhood of the poles. Then, Young et al. (1999) minimize the alternative objective function using non-linear least squares procedure.

⁴where polynomials ϕ^j and θ^j have constant parameters, so this is not the proper context for evolutionary spectra.

⁵In their original work, Young et al. (1999) seek the vector $\mathbf{NVR} = [1, NVR_0, \dots, NVR_R]$, where $NVR_j = \sigma_j^2 / \hat{\sigma}^2$, using the residual variance $\hat{\sigma}^2$ from a fitted AR model of the observed series \mathbf{z} .

Here we will be using the Linear Dynamic Harmonic Regression (LDHR) alternative developed by Bujosa et al. (2007). The LDHR is a linear alternative procedure that also provides an automatic identification of the complete DHR model. It exploits the algebraic structure of the *pseudo-spectrum* functions (see Bujosa et al. (2015)) to avoid the poles associated to the unit modulus AR roots of $f_{dhr}(\omega, \boldsymbol{\sigma}^2)$. Equation (6) is multiplied by the function $\Psi(\omega) = \varphi(e^{-i\omega})\varphi(e^{i\omega})$, where φ is the minimum order polynomial with all unit modulus AR roots of the complete DHR model, then we can solve by OLS the alternative minimization problem:

$$\min_{\boldsymbol{\sigma}^2 \in \mathbb{R}^{R+2}} \left\| \Psi(\omega) \cdot \widehat{f_y}(\omega) - \Psi(\omega) \cdot f_{dhr}(\omega, \boldsymbol{\sigma}^2) \right\|, \quad (7)$$

where $\widehat{f_y}(\omega)$ is the estimated periodogram $\widehat{I}(\omega)$ of \mathbf{y} . The estimation of $\widehat{I}(\omega)$ is computed after a pre-whitening of \mathbf{y} using an AR polynomial $\widehat{\phi_y}$ fitted to the series. The size, shape and location of the spectral peaks (the roots of $\widehat{\phi_y}$) are used to identify the model of each of the DHR components s^{pj} ; therefore all unit modulus AR roots in $\Psi(\omega)$ are the unit roots in $\widehat{\phi_y}(B)$. Hence, because in (7) all the unit modulus AR roots cancel out, the minimization problem (6) can be solved by OLS (see Bujosa et al., 2015, 2007).

The noise variance ratios ($NVR_j = \sigma_j^2/\sigma_e^2$) work as smoothing parameters; the smaller the NVR_0 , the closer to a linear deterministic trend the estimated trend is. In the limit, when the $NVR_0 = 0$ the estimated trend is linear. In the case of seasonal components, the smaller the NVR_j , the smoother the changes in the amplitude of the oscillations of the j -th DHR component s_t^{pj} (see Young et al., 1999). Our preference for the DHR is due to the advantages of Integrated Random Walk (IRW) and Smooth Random Walk (SRW) trend models over other procedures when dealing with monthly series. Although other alternatives may track the long term behaviour in any time series equally well, when we look at their associated first difference transformations the picture changes dramatically. In some cases, estimated trends actually contain some higher frequency components related to the annual cycle, which are then amplified by the difference operator.

Since the full DHR model is re-estimated when new data is available, small changes in some roots close to the unit circle could lead to different volatility on the final filtered components from one month to the next. In order to track more stable indicators we force the roots close to the unit circle to be on the unit circle so, in all cases, trends follow an IRW model and processes that modulate oscillations of seasonal DHR components follow a RW. This fixed DHR model is justified since it is the model that our LDHR algorithm identifies in most cases along the sample for the targeted variables.

2.2 Dynamic Factor Model

Let $\underline{\mathbf{y}}_t$ be a m -dimensional vector of the m time series (the estimated DHR trends in our case). Following Peña and Poncela (2006) and Lam and Yao (2012), we assume that this

vector can be written as a linear combination of $r < m$ common factors plus noise

$$\underline{\mathbf{y}}_t = \mathbf{P} \underline{\mathbf{f}}_t + \underline{\mathbf{e}}_t \quad (8)$$

where $\underline{\mathbf{f}}_t$ is the r -dimensional vector of common unobserved factors, \mathbf{P} is a factor loading matrix. We assume all the common dynamic structure in the trends comes through the common factors $\underline{\mathbf{f}}_t$. We suppose that the vector of common factors follows a vector autoregressive moving average, VARMA (p, q) model

$$\Phi(B) \underline{\mathbf{f}}_t = \Theta(B) \underline{\mathbf{a}}_t,$$

where $\Phi(B)$ and $\Theta(B)$ are polynomials of $r \times r$ matrices, the roots of the equation $|\Phi(B)| = 0$ can be on or outside the unit circle, and $\underline{\mathbf{a}}_t$ is normally distributed, with zero mean and serially uncorrelated with a full rank variance-covariance matrix Σ_a . The components of the vector of common factors, $\underline{\mathbf{f}}_t$, are non-stationary, and we assume the usual conditions for the invertibility of the VARMA model (see Lütkepohl, 1993, p. 222).

Assume $\underline{\mathbf{y}}_t$ is $I(d)$, and let $\mathbf{C}_y(k)$ be the generalised sample covariance matrices, that is,

$$\mathbf{C}_y(k) = \frac{1}{T^{2d+d'}} \sum_{t=k+1}^T (\underline{\mathbf{y}}_{t-k} - \underline{\bar{\mathbf{y}}})(\underline{\mathbf{y}}_{t-k} - \underline{\bar{\mathbf{y}}})',$$

where $\underline{\bar{\mathbf{y}}} = \frac{1}{T} \sum_{t=1}^T \underline{\mathbf{y}}$ and d' can be either 0 or 1. Peña and Poncela (2006) show that for non-stationary factor identification these matrices play the same role as the sample covariance matrices in the stationary case. They show [Theorem 1] that $\mathbf{C}_y(k)$ converges to a random matrix Γ_y that has r_1 eigenvalues greater than zero almost sure. This r_1 eigenvalues are a basis of the column space of the loading sub-matrix $\mathbf{P}\mathbf{1}$ of the r_1 non-stationary factors.

Unfortunately we can not assume that the sequence of vectors $\underline{\mathbf{e}}_t$ in (8) are normally distributed, have zero mean and full diagonal covariance matrix Σ_e , therefore we can not apply a formal statistical test for the number of common factors as in Peña and Poncela (2006)⁶. We can only inspect the eigenvalues and eigenvectors, in order to find evidence of some large eigenvalues with stable eigenvectors for different values of k , and the remaining eigenvalues relatively small with non-stable associated eigenvectors.

2.3 Cycle characterization

We use the characterization proposed in García-Ferrer and Bujosa-Brun (2000), which is directly linked to the first difference of IRW trend estimations outlined in Section 2.1.

⁶In fact Poncela (2012) claims that the precise estimation of the number of common factors is still a corner stone in both exact or strict as well as large and approximate dynamic factor models. In a less restrictive context, as it is our case, the number of common factors is an open issue.

We define the *anticipation of a recession* at time t as the point where the first difference of trend reaches its local maximum numerical value, and the *confirmation of a recession* as the point where the first difference becomes negative if it remains so for at least six months. Analogously, we define the *anticipation of an expansion* at the local first difference's minimum; and the *confirmation of an expansion* when the first difference becomes positive if it remains so for at least nine months. The empirical time differences for an expansion and a recession are somehow heuristic and based on empirical observations over a large set of IRW trends of monthly economic series. We will follow the same characterization for the cycle when dealing with our composite indicators.

2.4 Building composite indicators

The key issue in building the composite indicators is the selection of the targeted variables and their weighting scheme for their combination into a composite index.

2.4.1 The selection of the targeted variables

In building both, leading and coincident, composite indicators we initially rely on a starting data base that includes 46 monthly seasonal Spanish economic variables covering all economic sectors included in the *National Accounts* (see Bujosa et al., 2013, Table I, for details). For each composite indicator a small number of targeted variables is selected. Prior to the selection process all series are filtered in the same way (using their LDHR models) to extract their DHR components. Then, their cycles are obtained according to the definitions developed in Section 2.3. The variable selection depends on the (leading or coincident) nature of each composite indicator. In this regard, we use the mean and the range periods for the anticipation and confirmation of the recessions along the sample range (see Table I in Bujosa et al. (2013)), statistical criteria based on dynamic correlation analysis, and also economic information taking expert knowledge into account.

As a result, for the CCI case, our targeted variables include air traffic passengers, affiliates to the social security system, fuel consumption, IPI manufacturing and electricity consumption. For the CLI case, our targeted variables set includes cement consumption, car registrations, housing starts and commercial vehicle registrations (see Section 4). All individual data are available at least since January 1982.

2.4.2 Weighting the trends

Once a small set of targeted variables is selected, using the corresponding trend estimates, we proceed to estimate the DFM Equation (8) and the corresponding generalized sample covariance matrix $[\mathbf{C}_y(k)]$ at the five lags, $k = 1, \dots, 5$. For each small set described above, results corroborate the existence of a single nonstationary common factor. Hence, the nonstationary common factor for trends corresponding to the small set of targeted variables that

drive the cycle will be our CLI. The non-stationary common factor corresponding to the set of *coincident* targeted variables will be our CCI. In each case the composite indicator is finally computed as a weighted sum of the corresponding targeted variable trends. In each case, the vector of weights is the normalized eigenvector corresponding to the largest eigenvalue of $\mathbf{C}_y(5)$.

The remaining of this section is devoted to the building details for the CCI.⁷

2.4.3 Building the CCI

Table 1 includes estimates of the corresponding Noise Variance Ratio (NVR) coefficients for the trend as well as for the main seasonal component and its harmonics. The low values of the NVR estimates imply considerable smoothness. Differences among the coincident indicators are minor with the exception of *Workers in SS System*. Also the seasonal NVR estimates are quite similar for all series, but the seasonal harmonic component with period 2 is not present in three monthly indicators.

Series	Trend	S12	S6	S4	S3	S2.4	S2
Air traffic passengers	0.0022	0.0154	0.0078	0.0052	0.0028	0.0019	
Fuel consumption	0.0005	0.0019	0.0008	0.0056	0.0017	0.0044	
Workers in SS System	0.0304	0.0071	0.0006	0.0010	0.0013	0.0021	
IPI manufacturing	0.0012	0.0015	0.0024	0.0045	0.0063	0.0106	0.0004
Electricity consumption	0.0005	0.0069	0.0053	0.0023	0.0025	0.0037	0.0013

Table 1: Estimated NVRs of individual monthly variables on the CCI index.

In Figure 1, we have plotted the targeted variables in logs, and their estimated IRW trends; where both non-stationarity in the mean and strong seasonality are evident. Using the trend estimates, we have obtained and plotted the estimated individual cycles (IRW trend derivatives) for the five monthly indicators (right hand side column of Figure 1).

In Figure 2 we show the five eigenvalues of the generalized sample covariance matrix of the five trends using the whole sample $[\mathbf{C}_y(k)]$ at the first five lags, $k = 1, \dots, 5$. As we can see, although the five eigenvalues are quite stable at different lags, the first one λ_1 is much larger than the remaining ones, indicative of the existence of only a single nonstationary factor. Figure 3 shows that the corresponding eigenvector is also quite stable at different lags (k). In fact, it is more stable than the other eigenvectors, that are real at some lags,

⁷For details for the CLI see Bujosa et al. (2013). Information about the individual NVR estimates, estimated individual cycles, eigenvalues of $\mathbf{C}_y(k)$ for different lags and the normalized eigenvector associated with the larger eigenvalue of $\mathbf{C}_y(k)$, as well as the evolution of the weights used to build de CLI are available from the authors upon request.

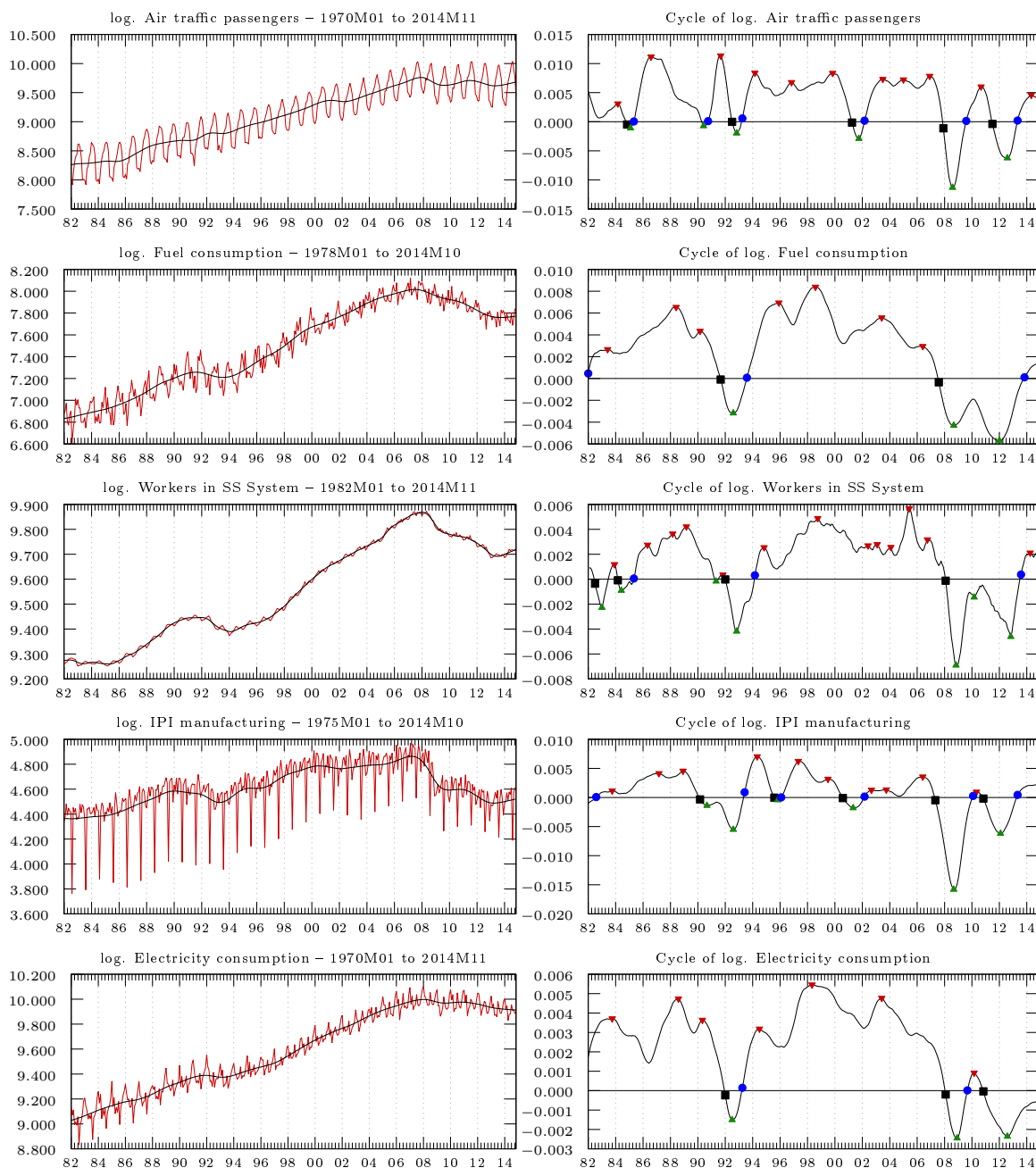


Figure 1: Left hand side column: the coincident indicators in logs and their smoothed IRW trends. Right hand side column: the estimated individual cycles (IRW first differences). Red triangles mark *anticipations of a recession*, black squares mark *confirmations of a recession*, green triangles mark *anticipation of an expansion* and blue dots mark *confirmations of an expansion* (see Section 2.3).

but two of them (not always the same pair) become complex conjugates for different values of k . This also corroborates the existence of a single nonstationary common factor.

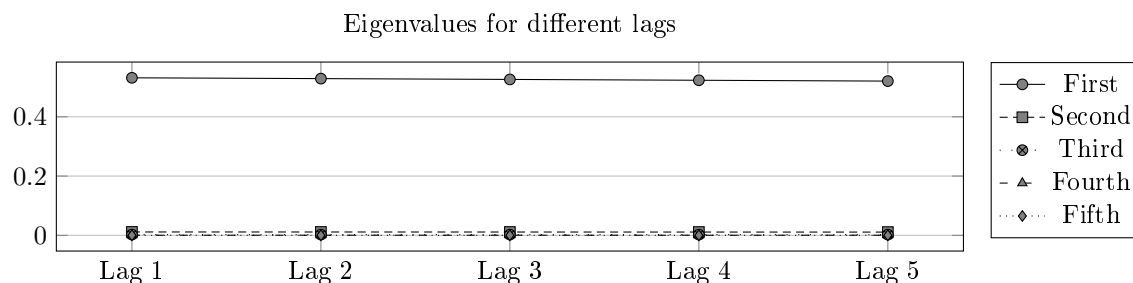


Figure 2: Eigenvalues of $\mathbf{C}_y(k)$ for different lags k .

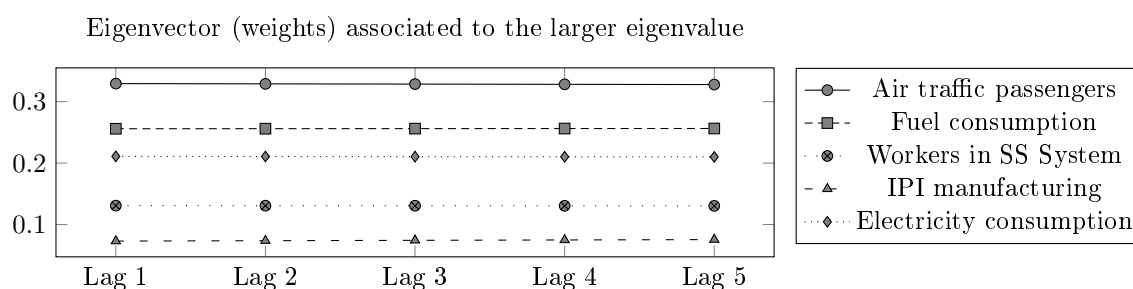


Figure 3: Normalized eigenvector associated to the larger eigenvalue of $\mathbf{C}_y(k)$ for different lags k .

Therefore, we will consider only one non-stationary common factor (since $\lambda_1/\sum \lambda_i \approx 97\%$ of the variance). That common factor will be our Composite Coincident Indicator (CCI); a weighted sum of the five monthly indicators. The vector of weights is the eigenvector corresponding to the largest eigenvalue. In order to carry on the out-of-sample exercise we have computed the generalized sample covariance matrix $\mathbf{C}_y(5)$ using the iteratively estimated trends. Figure 4 shows the evolution of the components of the first eigenvector (the weights used to build the CLI) when the sample is updated.

3 Composite Coincident Indicator

3.1 The CF and the CCI indexes

In April 2015, the AEE launched the first “official” dating of the Spanish Business Cycle through its Dating Committee. This achievement meant an important milestone for researchers

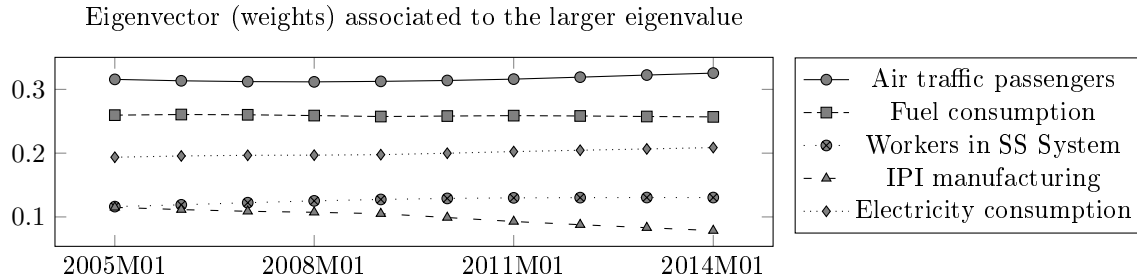


Figure 4: Evolution of the weights used to build the Composite Leading Indicator (using only the information available at each point in time).

and policy makers trying to analyze and compare alternative Spanish business cycle characterizations⁸ As it happens with the NBER for the US data or the CEPR for the euro area, we now have a yardstick to compare historical and future results.

The AEE index (CF) is built by using time series data on different frequencies (quarterly and monthly) following the initial dynamic factor model proposed by [Stock and Watson \(1991\)](#) and further extensions by [Mariano and Murasawa \(2003\)](#) and [Aruoba, Diebold, and Scotti \(2009\)](#) (see [Camacho, Pérez-Quirós, and Poncela \(2013\)](#) for a recent survey). The model is estimated by maximum likelihood using the Kalman Filter. The main assumption behind this approach is that the dynamics of each economic variable depends on two factors. The first one (a common factor that governs the simultaneous movements of all variables) provides a business cycle interpretation while the second factor would represent the idiosyncratic movements of each individual variable.

The set of variables chosen to build the CF index includes quarterly GDP, three monthly indicators of economic activity; namely, industrial production index, social security affiliates and the indicator of service sector activity, plus three indicators based on surveys: purchasing manager index, economic sentiment indicator and the consumer confidence index.⁹

When comparing our restricted data set with the one used in the CF index a major difference appears. In our CCI, there are not expectation variables included. Two main reasons for their absence. First, both the purchasing manager index (1999M08) and the economic sentiment indicator (1987M04) series start their publication too late for our interest. Second, we have doubts regarding the value of the survey responses to questions about the present

⁸Its web page: <http://asesec.org/CFCweb/> gives access to the data and methodology used, the precise turning points and an explanation for each of them. In addition to this, the Dating Committee maintains an update the index of economic activity, which combines monthly and quarterly information from several data sources to extract the latent level of economic activity in real time.

⁹Data enter the model under different transformations. Survey data enter in levels, GDP in seasonally and calendar effects adjusted while the monthly economic variables are seasonally adjusted except the industrial production index that is not.

and future state of the economy that is not already embodied in other economic variables.¹⁰

In Figure 5 we have plotted the official CF and the CCI monthly indexes from 1983M07 to 2014M12. Both indexes show a very similar behavior along the business cycle (with correlation 0.86) and our CCI seems to capture very well the beginning of the official AEE recessions (shaded areas) and expansion periods during this period. Also, as Table 2 shows, our CCI seems to show a major lead at the beginning of 1994 expansion and a minor one at the beginning of the 2008 recession.

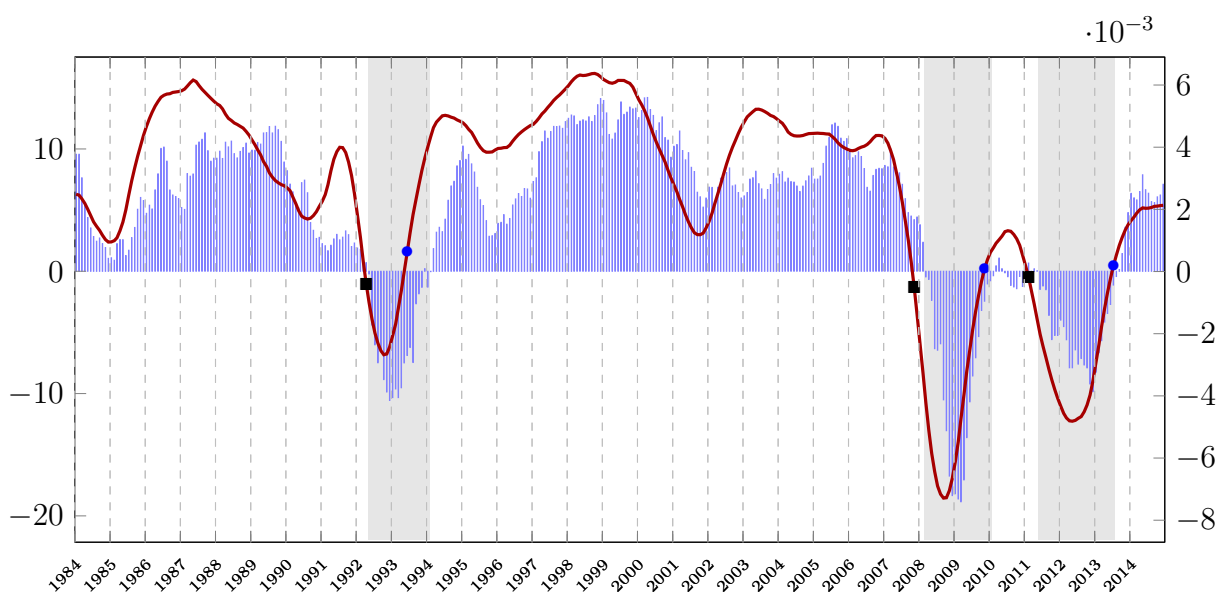


Figure 5: The “official” CF (bars) and CCI (line) indexes: 1984M01–2014M12. In all figures, shaded areas correspond to the “official” AEE recessions (see Section 3.1)

3.2 The Economy, the News media and the CCI index

In the political science literature there is a growing body of work exploring the relationship between the economy, economic news and public perception in various countries. For the US, [Doms and Morin \(2004\)](#) show how the news media affects consumers’ perceptions of the economy through different channels. In particular, they show that the greater the volume of news about the economy, the greater the likelihood that consumers will update their expectations about the economy. More recently, [Soroka, Stecula, and Wlezien \(2015\)](#) using

¹⁰[\(Leamer, 2009, Chap. 13\)](#) provides a good test of the usefulness of these consumer sentiment variables. After the 9/11 terrorist attack, the Conference Board index of consumer expectations plummeted. The plummet in the index after 9/11 should have been evident in retail sales. However, while retail sales tumbled in September 2001, those sales were completely recaptured in October and then back to the normal trend.

Recession begins		Expansion begins	
CF index	CCI	CF index	CCI
1992M05	1992M04	1994M02	1993M06
2008M03	2007M11	2010M02	2009M11
2011M06	2011M02	2013M08	2013M07

Table 2: Recessions dating of coincident indexes. Dates on the CCI columns correspond to *confirmation* of recessions and expansions of our CCI following the definitions given in Section 2.3

31,180 news stories over 20 years in the US find how media measures just do a very good job capturing the economy itself, better even than particular economic indicators. Similarly, Lischka (2015) uses VAR/Granger methodologies to check the dynamic relationships among real-world indicators, public expectations, and aggregated news on the general economy for the four most popular news outlets in Germany from 2002 to 2011.

For the Spanish case, media variables are based on a comprehensive monthly data set of frontpage economic news stories from three national newspapers: “ABC”, “El Mundo” (EM) and “El País” (EP). The final data set includes 5,095 stories for the three newspapers obtained by Soto (2017) from May 2004 until December 2015. A human-based content analysis has been performed to convert the array stories into time series of economic news coverage¹¹. The resulting data has been arranged into two indicators:

1. Volume of Coverage (Volume), a straightforward monthly economy-related article count
2. Tone of Content (Tone), net monthly balance of all the economy-related stories (positive minus negative stories).

The volume and tone for the three newspapers and its weighted average of coverage are illustrated in Figure 6¹²

Note that tone and volume series are negatively correlated. This reflects the fact that coverage increases when the news are bad, a well known aspect of modern mass media research (Soroka et al., 2015). Also, as the first column of Table 3 indicates, tone is positively related to the business cycle while volumen is negatively related.

¹¹As a result of the content analysis, each article was assigned to a positive, negative, or neutral category. A story was coded as positive if it presented the economy as improving. Conversely, stories that depicted the economy as declining were coded negative. Articles with no clear tone or mixed messages were coded as neutral.

¹²Besides the tone and volume of the individual journals, we also use the respective weighted tone and volume averages where the weights are computed as a function of the journals circulation. Consequently weights change from year to year.

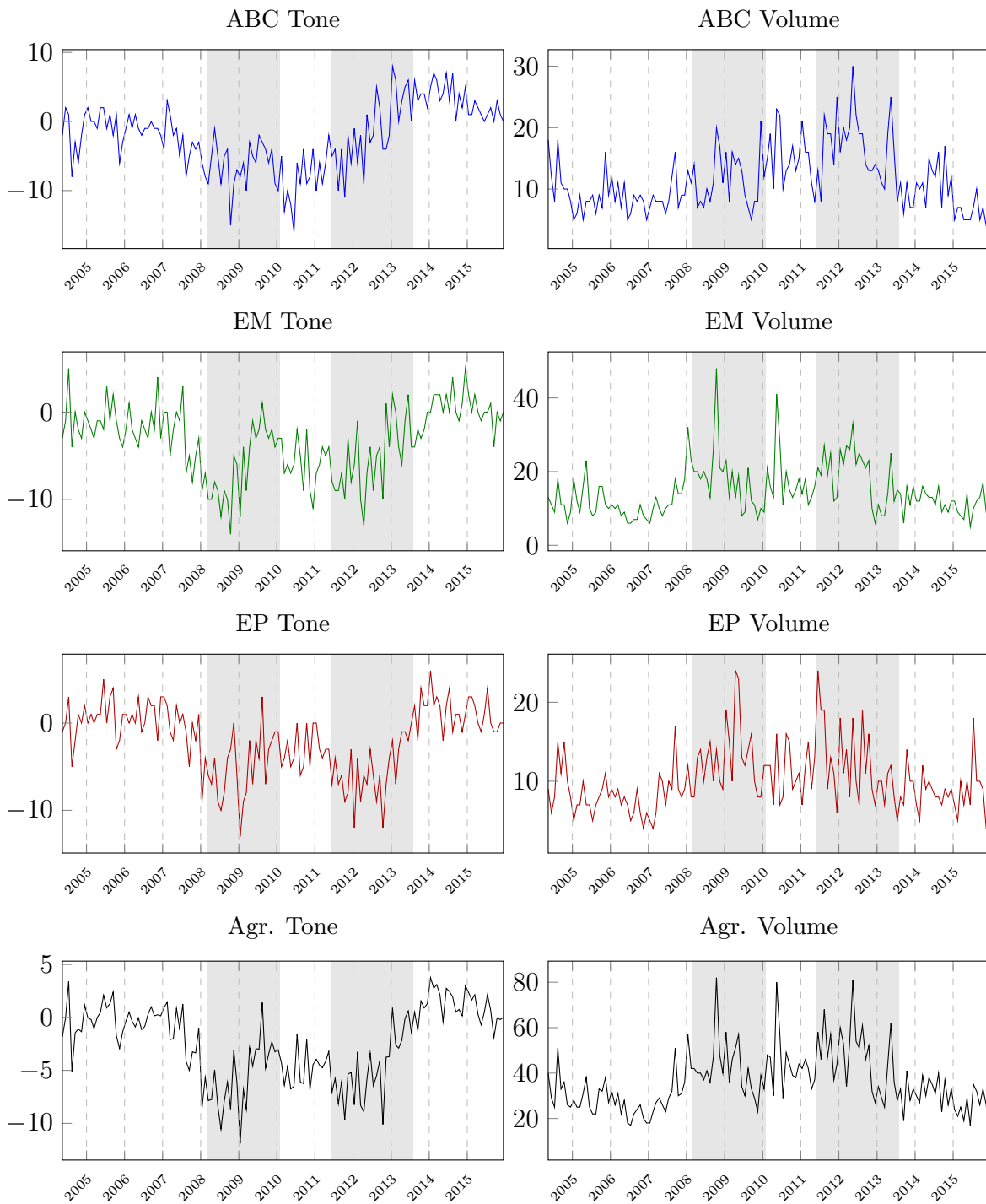


Figure 6: Tones and Volumes of the main Spanish newspapers and their aggregate.

Variables	CCI	ABC(T)	EM(T)	EP(T)	AG(T)	ABC(V)	EM(V)	EP(V)
ABC (Tone)	.336	–						
EM (Tone)	.627	.641	–					
EP (Tone)	.713	.461	.640	–				
AG (Tone)	.710	.717	.884	.903	–			
ABC (Vol.)	-.410	-.356	-.425	-.421	-.474	–		
EM (Vol.)	-.552	-.399	-.597	-.521	-.610	.586	–	
EP (Vol.)	-.451	-.276	-.383	-.538	-.508	.351	.470	–
AG (Vol.)	-.578	-.438	-.578	-.586	-.654	.796	.885	.650

Table 3: Bi-variate Correlations: Economic and Public media measures. *Note:* ABC, EM=El Mundo, EP=El País, AG = Weighted Press Index. *Data:* 2004M05 – 2015M12. Negative correlations are shaded.

In Figure 7 we show the normalized EP tone and the normalized CCI. In general, the media coverage of the economy tracks very well the state of economy. Although its signals are not always consistent (stability condition), exceptions are few ¹³.

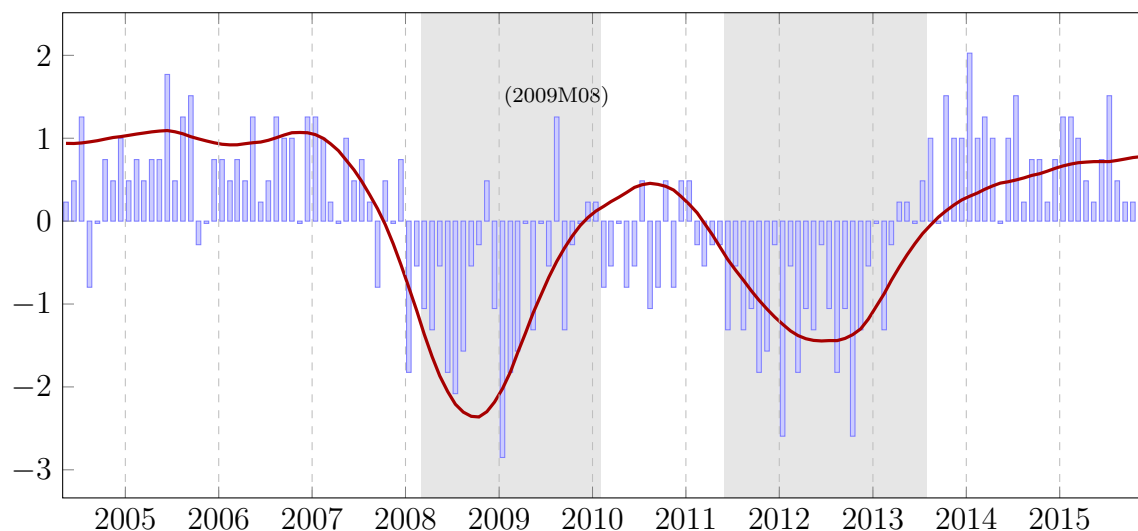


Figure 7: El País Tone (bars and left scale) and the CCI (solid line and right scale).

¹³The most notable one corresponds to 2009M08. At that time, however, Spanish Government was repeatedly announcing the possibility of “green shoots”, indicative of a potential recovery.

4 Composite Leading Indicator

One of the key issues in building leading indicator models is the selection of the leading variables, and of their weighting scheme for their combination into a composite leading index. Years ago, [Emerson and Hendry \(1996\)](#) acknowledged that the variables selected as leading indicators change all too often, suggesting that they didn't lead for very long, and that picking up leading indicators by maximizing in sample correlation was unreliable. This reflects one of the main problems in constructing leading indicator models: certain categories such as *leading*, *coincidental* and *lagging* are not invariant through time; neither are the same ones across different countries. More recently, however, leading indicators have been giving a new lease of life by exploiting the fact that leading indicators (available at monthly frequency) can be stacked in the model. In a univariate context, the MIDAS approach by [Ghysels, Santa-Clara, and Valkanov \(2006\)](#) that directly links low-frequency to high-frequency data has been extensively used in macroeconomic forecasting [e.g. [Clements and Galvão \(2009\)](#)] to obtain larger predictive gains, particularly at short horizons. Also, in a multivariate context, the mixed-frequency VAR suggested by [Ghysels \(2016\)](#) and the U-MIDAS model of [Forni, Marcellino, and Schumacher \(2015\)](#) offers potential forecasting gains.

As it is described in Section 2.4, for our CLI, the selected small number of leading targeted variables includes cement consumption, car registrations, housing starts and commercial vehicle registrations. From the economic point of view, the selected variables are good proxies of the Spanish economy since they belong to the main contributing sectors to GDP, namely private consumption investment and construction. For the period 1980M01 to 2009M12, [Bujosa et al. \(2013\)](#) show how well individual indicators lead the reference business cycle with a lead ranging from 8 to 14 months. Also, when looking at growth cycles, ([Bujosa et al., 2013](#), Table 4) also show that the CLI always precedes the highs and lows of the GDP growth rates. The average leading anticipation is 3.5 quarters in peaks and 3.6 quarters in troughs; confirming the greater length and amplitude symmetry of the growth cycles than their corresponding business cycles [see, e.g. [Niemira and Klein \(1994\)](#) and [García-Ferrer, Queralt, and Blazquez \(2001\)](#)].

In Figure 8 we show the CLI index together with official AEE's CF index. Also in Table 4 we report information regarding the turning points dating (peaks and troughs) of both indexes from March 1983 to December 2014. As can be seen, the CLI leads CF by eighteen months at peaks but only eight months at troughs. This suggest that in terms of timeliness, the CLI is most useful for signalling oncoming recessions. As in other cases [e.g. [Paap, Segers, and van Dijk \(2009\)](#)] there seems to be convincing evidence favoring the presence of a non-synchronous common cycle with asymmetric lead times.

The double dip behavior and the long lead of CLI preceding the 1992 recession were caused, in a different way, by two special events that took place in Spain in 1992: the Olympic Games of Barcelona and the Universal Exposition in Seville (Expo 92). Whereas

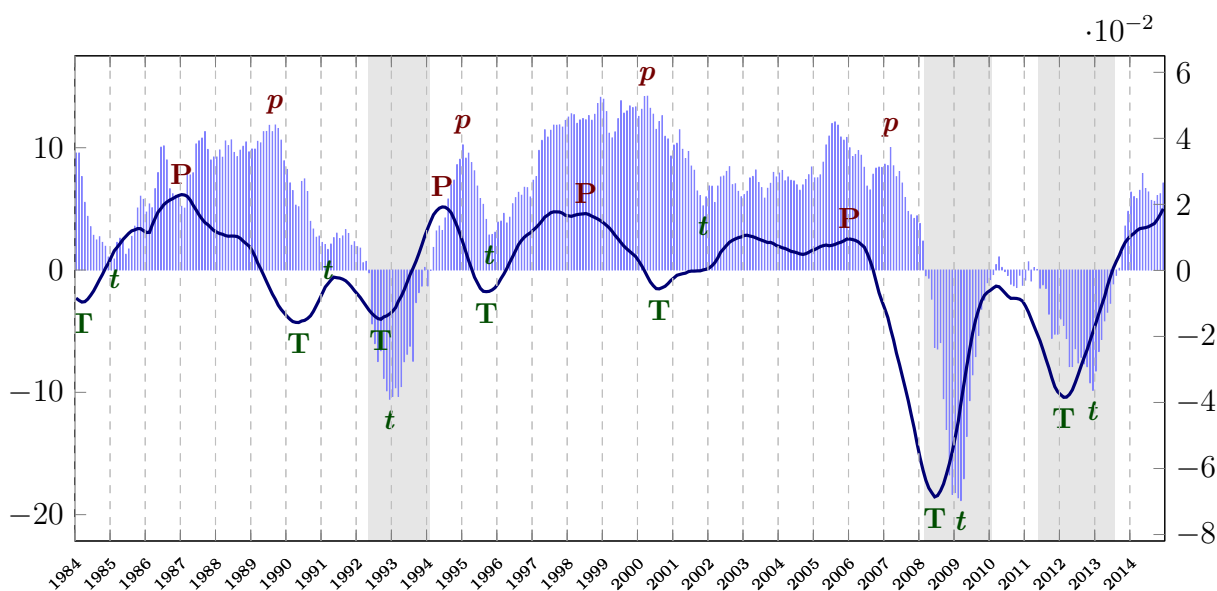


Figure 8: CF (bars and left scale) and CLI (solid line and right scale) indexes 1984M03–2014M12. Peaks and troughs marked with letters (Upper case for CLI and italic lower case for CF).

Peaks			Troughs		
CLI	CF	# months	CLI	CF	# months
			84M03	85M02	11
87M01	89M09	32	90M05	91M03	10
			92M09	92M12	3
94M06	95M01	6	95M09	95M10	1
98M07	00M04	21	00M08	01M11	15
06M01	07M03	13	08M06	09M03	9
			12M02	12M12	10
Average		18	Average		8.4

Table 4: Turning Points Dating of CLI and CF indexes

most developed economies were in recession by mid 1991, Spain was heavily investing on these events and the Spanish economy was not in a contraction phase. But when the rest of the world was taking off from the recession, the Spanish economy came into a contraction. Thus the 1992 special events in Spain mask the recession (see [Bujosa et al. \(2013\)](#)). In this regard, the CLI anticipates the 2008 recession as early as 2006M01, 10 months earlier than the CCI and 13 months earlier than the official CF index.

5 Real-time business cycle forecasting

The full-sample estimation results discussed in the previous section demonstrate two facts. The first one is that using the CCI within the LDHR model delivers an accurate description of the Spanish business cycle dynamics; and secondly, that the CLI is a good leading indicator of both CCI and CF indexes. However, both facts are ex-post features since the trends used to construct the CCI and the CLI indexes are smoothed estimates that, at each data point, use information pertaining to the whole sample. But the practical usefulness of leading indicators can only be assessed on their ability to signal changes in the business cycle ex-ante. Therefore, a realistic test of this issue requires the use of information that were actually available at the beginning of each forecast period.

Our forecasting exercise will be based on the one-step ahead forecast errors of annual Spanish GDP growth from 2001 to 2016. The forecast period include 16 years under three well defined paths: a normal growth period (2001–2007), a double-dip recession period (2008–2012) and a posterior recovery (2013–2016). As such, it is a good yardstick to check the ability of alternative models to produce, not only accurate forecasts but, also the right signaling of the corresponding turning points. The CLI forecasts will be compared with those of other 19 individual public and private institutions published by Consensus Forecasts and FUNCAS¹⁴.

The empirical literature on the performance of professional macroeconomic forecasts is abundant. Using different data sets and methodologies it has analyzed several issues such as accuracy, unbiasedness, and efficiency extensively. At first, it focused on large advanced countries (see, e.g. [Ager, Kappler, and Osterloh, 2009](#); [Dovern and Weisser, 2011](#)). More recently, however, attention has been spreaded to emerging countries like Brazil ([Carvalho and Minella, 2012](#)), Mexico ([Capistrán and López-Moctezuma, 2014](#)), China ([Deschamps and Bianchi, 2012](#)) and individual or groups of the Asian-Pacific countries ([Chen, Costantini, and Deschamps, 2016](#)). As far as we know, however, no such comparative analysis has been carried on in the case of Spain.

To be sure that this is a genuine out-of-sample forecasting exercise all models should

¹⁴Every month, Consensus Economic surveys over 250 prominent financial and economic forecasters for their estimates of a range of variables including future growth, inflation, interest rates and exchange rates. More than 20 countries are covered and the reference data, together with analysis and polls on topical issues is rushed to subscribers by express mail and e-mail.

share the same time information when the forecasts are made. In the case of the institutions included in the Consensus Forecasts and FUNCAS, the annual GDP growth forecasts for 2001 are those published in the January 2001 report that use information available up to December 2000. Similarly, GDP forecasts for 2002 are those published in the 2002 January report using information up to December 2001... and so on. Therefore, for each institution we collect 16 one-step-ahead forecasts and their corresponding 16 one-step-ahead forecast errors.

In the case of the monthly CLI, annual growth rates are computed directly as the difference between January and December for every year in the sample. As with its competitors, the newest information used to produce the $(t + 1)$ forecast is December of year t . The initial estimation period runs from 1980 to 2000 where the GDP growth rate is regressed on three lags of CLI growth rates plus a constant. So, the 2001 GDP growth rate forecast is obtained. Since the CLI is updated along 2001, the model is re-estimated again to obtain the 2002 forecast... and so on. As its contenders, we will obtain a vector of 16 one-step-ahead forecasts and errors.¹⁵

Absolute and relative Root Mean Square Error (RMSE) results for individual institutions are shown in Table 5, where the consensus represents the simple average of all institutions excluding the CLI¹⁶ and is used as the reference to calculate the relative performance. For the whole forecasting period (2001–2016), the CLI shows considerable improvement over its competitors (including the consensus) which tend to perform similarly, indicating a potential degree of herding behavior (e.g. Rülke, Silgoner, and Würz, 2016). For the normal growth period (2001–2007), however, all institutions perform similarly but the Spanish Government seems to outperform the other alternatives. For this particular period, the CLI results are in line with those obtained by the consensus. Again, for the final forecasting period (2008–2016), CLI provides the most accurate forecasts and the gain relative to the other alternatives is statistically significant; see Table 6.

We test whether the differences in RMSE found in Table 5 are statistically significant using the Diebold and Mariano test of equal predictive accuracy of the CLI relative to the other models in the panel. Statistical results in Table 6 confirm the evidence given in Table 5. In particular, the CLI model provides significantly more accurate forecasts than the other 19 institutions for the whole and final forecasting periods. For the initial period (2001–2007), however, the differences are not statistically significant.

¹⁵Detailed information about empirical results are available from the authors upon request. Also data for the updated CLI vintages are available at <http://uam-ucm-economic-indicators.es/>.

¹⁶Other forecasts accuracy measures results like Mean Absolute Error (MAE) or Mean Absolute Scaled Error (MASE) are available from the authors upon request. However, their messages are identical to those shown by the RMSE. For instance, the correlation coefficients between the RMSE and MASE for the three forecasting periods in Table 5 are 0.91, 0.96 and 0.97, respectively.

Institution	Root Mean Square Error (RMSE)			Relative RMSE over Consensus		
	2001–2016	2001–2007	2008–2016	2001–2016	2001–2007	2008–2016
AFI	1.582	0.720	2.011	0.931	0.970	0.877
BANKIA	1.652	0.782	2.091	0.972	1.053	0.912
BBVA	1.828	0.798	2.333	1.076	1.075	1.017
Cataluña Caixa	1.787	0.783	2.335	1.052	1.055	1.018
CEPREDE	1.788	0.763	2.288	1.053	1.028	0.998
Econ Intel Unit	1.760	1.130	2.002	1.036	1.522	0.873
Euro Comission	1.870	0.688	2.418	1.101	0.927	1.054
FMI	1.911	0.756	2.459	1.125	1.019	1.072
FUNCAS	1.610	0.867	2.006	0.948	1.168	0.875
IFL	1.656	0.765	2.103	0.975	1.030	0.917
ICAE	1.655	0.729	2.112	0.975	0.982	0.921
IEE	1.664	0.760	2.116	0.980	1.023	0.922
Intermoney	1.737	0.788	2.210	1.023	1.062	0.964
La Caixa	1.718	0.780	2.185	1.012	1.050	0.953
OCDE	1.820	0.720	2.205	1.072	0.969	0.962
Santander	1.715	0.895	2.032	1.010	1.205	0.886
Spanish Gov.	2.246	0.560	2.954	1.323	0.755	1.288
The Economist	2.014	0.978	2.403	1.186	1.318	1.048
UBS	1.602	1.014	1.827	0.943	1.366	0.796
Consensus	1.698	0.742	2.293	1.000	1.000	1.000
CLI	0.820	0.793	0.892	0.483	1.068	0.389

Table 5: Absolute and relative RMSE for individual institutions at different forecasting periods

Institutions	2001–2016	2001–2007	2008–2016
AFI	1.803* (0.931)	−0.124 (0.333)	3.303* (1.190)
BANKIA	2.023** (0.905)	−0.030 (0.391)	3.620** (1.120)
BBVA	2.638* (1.390)	−0.001 (0.379)	4.692** (1.882)
Cataluña Caixa	2.450** (1.085)	−0.021 (0.058)	4.612** (1.340)
CEPREDE	2.502* (1.320)	−0.064 (0.340)	4.498** (1.764)
EIU	2.428*** (0.712)	0.943* (0.360)	3.171*** (0.850)
Eu Commision	2.790* (1.311)	−0.163 (0.396)	5.087** (1.606)
FMI	2.943* (1.409)	−0.069 (0.387)	5.285** (1.687)
FUNCAS	1.894** (0.868)	0.106 (0.291)	3.285** (1.239)
Spanish Gov.	4.343* (2.336)	−0.326 (0.363)	7.975 (3.086)
IFL	2.043** (0.928)	−0.060 (0.353)	3.679** (1.166)
ICAE	2.035** (0.925)	−0.110 (0.244)	3.704** (1.173)
IEE	2.065** (0.976)	−0.066 (0.448)	3.728** (1.218)
Intermoney	2.330** (1.087)	−0.017 (0.355)	4.156** (1.415)
La Caixa	2.533** (1.045)	−0.032 (0.371)	4.032** (1.379)
OCDE	2.669** (1.134)	0.066 (0.356)	4.115** (1.332)
Santander	1.927* (0.910)	0.050 (0.392)	3.387** (1.239)
The Economist	2.833** (1.207)	0.034 (0.531)	5.010*** (1.290)
UBS	4.002*** (1.244)	4.817* (2.620)	3.368*** (0.908)

Table 6: Diebold and Mariano forecast test for individual institutions

(HAC standard errors are in parenthesis. The superscripts ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively, of the [Diebold and Mariano \(1995\)](#) statistic for testing the null hypothesis that the difference in Mean Square Error (MSE) of the Individual institutions and the CLI is equal to 0).

5.1 Do forecast combinations help?

When several forecasts of the same event are available, forecast combination seeks to reduce the information in a vector of forecasts to a single combined forecast using weights which are

chosen to minimize expected loss. A variety of methods have been proposed in the literature, including performance-based combinations, principal components, projection on the mean, optimal weighting and Bayesian shrinkage (see [Timmermann, 2006](#), chap. 4 for details). Moreover, empirical studies and extensive simulations ([Genre, Kenny, Meyler, and Timmermann, 2013](#)) show that the estimated optimal forecast combination does not perform better than the single arithmetic average, a phenomenon dubbed the “forecast combination puzzle” —see [Claeskens, Magnus, Vasnev, and Wang \(2016\)](#) for a simple theoretical explanation—. In this paper we will also assess two forecast combination procedures recently developed by [Arroyo and de Juan \(2017\)](#). The main idea is to use the metric vector space of non-negative weights adding up to one (the so-called simplex) as the natural sample space for vectors of weights assigned to J forecasters. First, we apply usual multivariate statistics to log ratios of weights to form combinations. Then, we select those forecasters with weights larger than the benchmark average $1/J$.

Again, absolute and relative RMSE results for alternative forecast combinations are shown in [Table 7](#) for different forecasting periods. Please, note that the whole period 2001–2016 loses one observation since the computation of the two new combinations needs the 2001 forecast error to start their iterations. Also, the 2008–2016 forecast period has been splitted into two sub-periods to differentiate the crisis (2007–2012) and the posterior recovery (2013–2016). As in the previous section, the CLI shows the best forecasting results for all periods but the initial 2002–2006 one where results are very similar among the different combinations and the CLI. Interestingly, however, for this particular data set and time period, the new two combinations seem to improve the consensus, particularly so in the case of the last expansion period (2013–2016). When the CLI is included in the panel of forecasts both combinations assign it a larger weight in their combinations than the fixed $1/J$ value that it is used in the Consensus.

	RMSE				Relative RMSE over consensus			
	2002–2016	2002–2006	2007–2012	2013–2016	2002–2016	2002–2006	2007–2012	2013–2016
CLI	0.755	0.815	1.051	0.189	0.432	1.068	0.417	0.176
Consen.	1.747	0.763	2.519	1.071	1.000	1.000	1.000	1.000
Comb.	1.551	0.825	2.276	0.634	0.888	1.081	0.903	0.593
Selec.	1.421	0.719	2.105	0.528	0.814	0.942	0.836	0.493

Table 7: Absolute and relative RMSE of the CLI and alternative forecast combinations (consensus, combination with CLI, and section with CLI).

Statistical significance of the previous results are shown in [Table 8](#) using the [Diebold and Mariano \(1995\)](#) test of equal predictive accuracy of the CLI relative to the alternative combinations. The CLI model provides more accurate forecasts than the other alternative combinations in three out of four forecasting periods. As before, however, for the initial

period (2002–2006) the differences are not statistically significant.

Combinations	2002 – 2016	2002 – 2006	2007 – 2012	2013 – 2016
Consensus	2.451** (1.028)	0.251 (0.302)	5.205** (1.596)	3.837** (0.279)
Combination	2.490** (0.849)	0.325 (0.195)	4.040*** (0.854)	2.872 (2.277)
Selection	2.120** (0.826)	0.161 (0.241)	3.294*** (0.772)	2.808 (2.309)

Table 8: Testing equal predictive accuracy for the combinations

Heteroskedasticity and Autocorrelation Consistent (HAC) standard errors are in parenthesis. The superscripts ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively, of the Diebold-Mariano (1985) statistic for testing the null hypothesis that the difference in MSE of the combination and the CLI is equal to 0.

In summary, our CLI model produces more accurate forecasts than the other individual models and their combinations by considerable margin. Also, as Figure 9 shows, our model produces sharper and more accurate turning points estimates, both, for business cycles peaks and troughs and captures very well the path of the double dip recession.

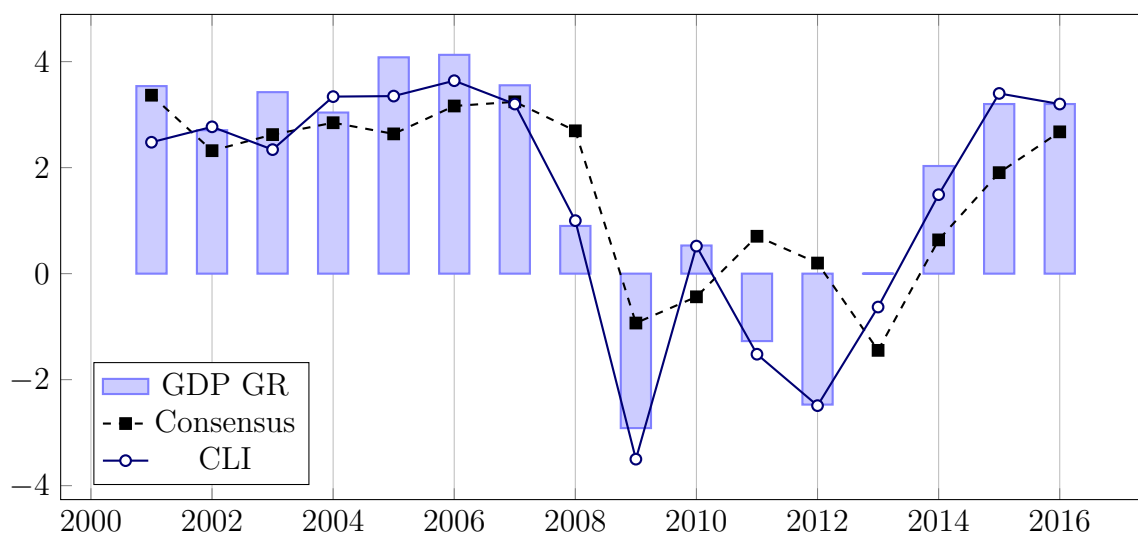


Figure 9: Observed and forecast annual GDP growth rates 2001–2016.

6 Conclusions

In this article we have presented two new composite coincident and leading indicators designed to capture the present state of the Spanish economy and to provide reliable statistical forecasting power, respectively. One novelty of our approach is that both indicators are based on IRW trend components of non-stationary monthly economic variables rather than on the usual stationary transformations of seasonally adjusted series. This approach guarantees that the resulting indicators are reasonably smooth and issue stable or persistent signals, reducing the uncertainty shown in many coincident and leading indicators in the literature. Such a volatility may be confusing for economic agents in interpreting future directions of change. For the long period of time considered in this paper, the signaling quality of both indicators remains intact over different samples and is not overly sensitive to the specific crisis dating employed.

As regards the usefulness of our coincident indicator, this has been checked by comparing it with the one recently proposed by the Spanish Economic Association, CF index. Both indexes show a very similar behavior along the business cycle and our coincident indicator seems to capture very well the beginning and end of the official AEE recessions and expansion periods. Additionally, our coincident indicator tracks very well alternative mass media indicators typically used in the political science literature. On the other hand, the suitability of the leading indicator has been assessed through both historical (in-sample) behavior and out-of-sample forecasting performance. When using the whole sample our leading indicator systematically predicts the peaks and troughs of the CF index with leads that change among different historical periods. It also provides significant aid in forecasting annual GDP growth rates for the 2001–2016 period. Using only real data available at the beginning of each forecast period, our leading indicator one-step-ahead forecasts shows sizable improvements over other alternatives, including panels of individual professional forecasters and different forecast combinations.

References

- Ager, P., M. Kappler, and S. Osterloh (2009). The accuracy and efficiency of the consensus forecasts: A further application and extension of the pooled approach. *International Journal of Forecasting* 25(1), 167–181.
- Altissimo, F., R. Cristadoro, M. Forni, M. Lippi, and G. Veronese (2010). New eurocoin: tracking economic growth in real time. *The review of economics and statistics* 92(4), 1024–1034.
- Álvarez, R., M. Camacho, and G. Pérez-Quirós (2016). Aggregate versus disaggregate in-

- formation in dynamic factor models. *International Journal of Forecasting* 32(3), 680 – 694.
- Arroyo, A. S. M. and A. de Juan (2017, September). Simplex combination and selection of forecasts. Granada. International Work-Conference of Time Series.
- Aruoba, S. B., F. X. Diebold, and C. Scotti (2009). Real-time measurement of business conditions. *Journal of Business & Economic Statistics* 27(4), 417–427.
- Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica* 70(1), 191–221.
- Bell, W. (1984, June). Signal extraction for nonstationary time series. *The Annals of Statistics* 12(2), 646–664.
- Boivin, J. and S. Ng (2006). Are more data always better for factor analysis. *Journal of Econometrics* 132, 169–194.
- Bujosa, M., A. Bujosa, and A. García-Ferrer (2015). Mathematical framework for pseudo-spectra of linear stochastic difference equations. *IEEE Transactions on Signal Processing* 63(24), 6498–6509.
- Bujosa, M. and A. García-Ferrer (2014, December). Outlier detection in unobserved harmonic regression models. ERCIM-CFE, University of Pisa, Italy.
- Bujosa, M., A. García-Ferrer, and A. de Juan (2013). Predicting recessions with factor linear dynamic harmonic regressions. *Journal of Forecasting* 32(6), 481–499.
- Bujosa, M., A. García-Ferrer, and P. C. Young (2007, October). Linear dynamic harmonic regression. *Comput. Stat. Data Anal.* 52(2), 999–1024.
- Burns, A. F. and W. C. Mitchell (1946). *Measuring the Business Cycle*. New York: National Bureau of Economic Research (NBER).
- Camacho, M., G. Pérez-Quirós, and P. Poncela (2013). Short-term forecasting for empirical economists: A survey of the recently proposed algorithms. *Foundations and Trends® in Econometrics* 6(2), 101–161.
- Capistrán, C. and G. López-Moctezuma (2014). Forecast revisions of mexican inflation and gdp growth. *International Journal of Forecasting* 30(2), 177–191.
- Carvalho, F. A. and A. Minella (2012). Survey forecasts in brazil: a prismatic assessment of epidemiology, performance, and determinants. *Journal of International Money and Finance* 31(6), 1371–1391.

- Chen, Q., M. Costantini, and B. Deschamps (2016). How accurate are professional forecasts in asia? evidence from ten countries. *International Journal of Forecasting* 32(1), 154–167.
- Claeskens, G., J. R. Magnus, A. L. Vasnev, and W. Wang (2016). The forecast combination puzzle: A simple theoretical explanation. *International Journal of Forecasting* 32(3), 754–762.
- Clements, M. P. and A. B. Galvão (2009). Forecasting us output growth using leading indicators: An appraisal using midas models. *Journal of Applied Econometrics* 24(7), 1187–1206.
- Corona, F., P. Poncela, and E. Ruiz (2017). Determining the number of factors after stationary univariate transformations. *Empirical Economics* 53(1), 351–372.
- Deschamps, B. and P. Bianchi (2012). An evaluation of chinese macroeconomic forecasts. *Journal of Chinese Economic and Business Studies* 10(3), 229–246.
- Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics* 20(1), 134–144.
- Doms, M. E. and N. J. Morin (2004). Consumer sentiment, the economy, and the news media. Technical Report WP-09, Federal Reserve Bank of San Francisco.
- Dovern, J. and J. Weisser (2011). Accuracy, unbiasedness and efficiency of professional macroeconomic forecasts: An empirical comparison for the g7. *International Journal of Forecasting* 27(2), 452–465.
- Drehmann, M. and M. Juselius (2014). Evaluating early warning indicators of banking crises: Satisfying policy requirements. *International Journal of Forecasting* 30(3), 759–780.
- Emerson, R. A. and D. F. Hendry (1996, July). An evaluation of forecasting using leading indicators. *Journal of Forecasting* 15(4), 271–91.
- Froni, C., M. Marcellino, and C. Schumacher (2015). Unrestricted mixed data sampling (midas): Midas regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 178(1), 57–82.
- García-Ferrer, A. and M. Bujosa-Brun (2000, Jun). Forecasting OECD industrial turning points using unobserved components models with business survey data. *International Journal of Forecasting* 16(2), 207–227.
- García-Ferrer, A. and P. Poncela (2002). Forecasting european gnp data through common factor models and other procedures. *Journal of Forecasting* 21(4), 225–244.

- García-Ferrer, A., R. Queralt, and C. Blazquez (2001). A growth cycle characterisation and forecasting of the Spanish economy: 1970–1998. *International Journal of Forecasting* 17(3), 517–532.
- Genre, V., G. Kenny, A. Meyler, and A. Timmermann (2013). Combining expert forecasts: Can anything beat the simple average? *International Journal of Forecasting* 29(1), 108–121.
- Ghysels, E. (2016). Macroeconomics and the reality of mixed frequency data. *Journal of Econometrics* 193(2), 294–314.
- Ghysels, E., D. Osborn, and P. Rodrigues (2006). Leading indicators. In G. Elliot, C. Granger, and A. Timmerman (Eds.), *Handbook of Forecasting*, pp. 659–711. Amsterdam: Elsevier.
- Ghysels, E., P. Santa-Clara, and R. Valkanov (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics* 131(1), 59–95.
- Hannan, E. J. (1967, Apr). Measurement of a wandering signal amid noise. *Journal of Applied Probability* 4(1), 90–102.
- Koopman, S. J. and N. Shephard (Eds.) (2015). *Unobserved components and time series econometrics*. Oxford University Press.
- Lam, C. and Q. Yao (2012). Factor modeling for high-dimensional time series: inference for the number of factors. *The Annals of Statistics* 40(2), 694–726.
- Leamer, E. E. (2009). *Macroeconomic Patterns and Stories*. Berlin: Springer-Verlag.
- Leamer, E. E. (2012). The context matters: Comment on Jerome H. Friedman, "Fast sparse regression and classification". *International Journal of Forecasting* 28(3), 741–748.
- Lischka, J. A. (2015). What follows what? relations between economic indicators, economic expectations of the public, and news on the general economy and unemployment in germany, 2002-2011. *Journalism & Mass Communication Quarterly* 92(2), 374–398.
- Lütkepohl, H. (1993, August). *Introduction to Multiple Time Series Analysis* (Second ed.). Heidelberg, Germany: Springer.
- Mariano, R. S. and Y. Murasawa (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of applied Econometrics* 18(4), 427–443.
- Matas-Mir, A., D. R. Osborn, and M. J. Lombardi (2008). The effect of seasonal adjustment on the properties of business cycle regimes. *Journal of Applied Econometrics* 23, 257–278.

- Niemira, M. P. and P. A. Klein (1994). *Forecasting Financial and Economic Cycle*. Nueva York: John Wiley & Sons.
- Paap, R., R. Segers, and D. van Dijk (2009). Do leading indicators lead peaks more than troughs? *Journal of Business & Economic Statistics* 27(4), 528–543.
- Peña, D. and P. Poncela (2006, april). Nonstationary dynamic factor analysis. *Journal of Statistical Planning and Inference* 136(4), 1237–1257.
- Phillips, K. and M. Boldin (2017, June). Testing and adjusting for sharp breaks in seasonal patterns: The case of US GDP data. Paper presented at the 37th International Symposium on Forecasting Cairns, Australia.
- Pierce, D. A. (1979, November). Signal extraction error in nonstationary time series. *The Annals of Statistics* 7(6), 1303–1320.
- Poncela, P. (2012). Further research on conditionally heteroskedastic factor models. *International Journal of Forecasting* 28, 94–96.
- Poncela, P. and A. García-Ferrer (2014). The effects of disaggregation on forecasting non-stationary time series. *Journal of Forecasting* 33(4), 300–314.
- Poncela, P. and E. Ruiz (2015). More is not always better: Back to the kalman filter in dynamic factor models. See [Koopman and Shephard \(2015\)](#), Chapter 11, pp. 250–271.
- Rülke, J.-C., M. Silgoner, and J. Wörz (2016). Herding behavior of business cycle forecasters. *International Journal of Forecasting* 32(1), 23 – 33.
- Sims, C. A. (2012, Spring). Comments and discussion on stock, j.h. and watson, m.w. ”disentangling the channels of the 2007–09 recession”. *Brookings Papers on Economic Activity*, 141–148.
- Sobel, E. (1967, Aug). Prediction of noise-distorted, multivariate, non-stationary signal. *Journal of Applied Probability* 4(2), 330–342.
- Soroka, S. N., D. A. Stecula, and C. Wlezien (2015). It’s (change in) the (future) economy, stupid: economic indicators, the media, and public opinion. *American Journal of Political Science* 59(2), 457–474.
- Soto, I. (2017). La información económica ante los cambios de ciclo. sesgos profesionales y editoriales de la prensa nacional durante las legislaturas VIII, IX y X. In *13er Congreso de la Asociación Española de la Ciencia Política y de la Administración*, Santiago de Compostela. AECPA.

- Stock, J. and M. Watson (2003). How leading indicator forecasts perform during the 2001 recession? *Economic Quarterly* 89, 71–90.
- Stock, J. H. and M. W. Watson (1991). A probability model of coincident economic indicators. In K. Lahiri and G. H. Moore (Eds.), *Leading Economic Indicators: New Approaches and Forecasting Records*, pp. 63–90. Cambridge: Cambridge University Press.
- Stock, J. H. and M. W. Watson (2002). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association* 97(460), 1167–1179.
- Tiao, G. C. and S. C. Hillmer (1978, Dec). Some consideration of decomposition of a time series. *Biometrika* 65(3), 497–502.
- Timmermann, A. (2006). Forecast combinations. *Handbook of economic forecasting* 1, 135–196.
- Young, P. C., D. Pedregal, and W. Tych (1999, November). Dynamic Harmonic Regression. *Journal of Forecasting* 18, 369–394.