

Green Shoots? Where, when and how?*

Maximo Camacho
Universidad de Murcia
mcamacho@um.es

Gabriel Perez-Quiros[†]
Banco de España and CEPR
gabriel.perez@bde.es

Pilar Poncela
Universidad Autónoma de Madrid
pilar.poncela@uam.es

Abstract

What is the meaning of *green shoots*? In this paper we provide a statistical definition of this term which allows us to analyze where, when and how the recovery started. With the same methodology, we confirm that the symptoms of recovery are clear in the US, the Euro area and Spain with some differences in timing. In addition, we find some leading behavior from the quotes in the press to the actual confirmation of the data, even when the data include variables with clear expectations contents.

Keywords: Business Cycles, Output Growth, Time Series.

JEL Classification: E32, C22, E27

*The authors thank Antoni Espasa for very helpful comments and suggestions. Maximo Camacho thanks *Fundacion Ramon Areces* for financial support. The views in this paper are those of the authors and do not represent the views of Bank of Spain or the EuroSystem.

[†]Corresponding Author: Gabriel Pérez Quirós, Banco de España, Research Division, Monetary and Financial Studies Department, DG Economics, Statistics, and Research. Alcalá 48, 28014 Madrid, Spain. E-mail: gabriel.perez@bde.es

1 Introduction

In the middle of the current recession, analysts, policy makers, and journalists, have used the term *green shoots* to refer to signals of the end of the recession period. Although this expression was first used in this sense by Norman Lamont, the Chancellor of the Exchequer of the United Kingdom, during the 1991 recession, it was popularized in US by Ben Bernanke, chairman of the Federal Reserve Board, who states on March 15th 2009 that he detected green shoots of economic recovery. From this quote, the use of the expression has been massive, with more than 189 million entries in Google since then.

Needless is to say that the term green shoots has not always been used with scientific criteria mainly for two reasons. First, the term is very imprecise so it leaves the users of the term to identify where, when and how the recovery comes depending basically on their own beliefs. Of course, the signals of recoveries do not appear in all the economic indicators with the same intensity at the same time in different countries. Hence, the skeptical users will be inclined to accentuate the negative signals of some indicators while the optimistic users will be tempted to stress the positive signals of some others. Perhaps, it is the impreciseness of the definition of green shoots what also diminishes the meaning of international comparisons of the existence of these green shoots. Second, in the search of green shoots, the recent advances in information technologies make the number of variables with information about the economy exponentially growing and with an unprecedented updating frequency. The cost of checking in real time the publication calendar of the variables, the latest release and the revisions makes very difficult the task of the analyst of keeping updated to check every day if the shoots are actually green.

The purpose of this paper is to provide economic agents with a statistical definition of the term green shoots which is very easy to interpret for the general public. In particular, we will define green shoots as a low probability of being in a recession in period t with the information available up to that period. This definition overcomes the two problems previously stated associated with the increasingly use of the term green shoots. First, the probability of recession is no longer an imprecise term. The inferences about the state of the cycle are computed from a statistical model applied to data which is then

transparent and objective. In addition, since recession probabilities are free of units of measurement, international comparisons are easily allowed. Second, if the probability of recession is computed based on a set of variables that the agents consider as representative of economic activity, because the chosen variables are good proxies of the general economic activity, this probability of recession should be a “sufficient statistic” for the analysts with the subsequent saving of time and cost for them. We pretend, using a computationally simple algorithm, to compute an easy to interpret number that the users can update when needed.

To compute the probabilities, we propose a dynamic factor model from a set of economic indicators that captures expansion and recession phases as unobserved regime shifts in the mean of the common factor. The unobserved state variable controlling the regime shifts is modeled as following a Markov process as in Hamilton (1989) but differs from other methods employed in the literature in a number of important ways. First, most of the empirical applications fit the Markov switching process to GDP series, assuming that this variable captures all the relevant information about the state of the economy. However, the NBER dating committee defines a recession as a significant decline in economic activity spread across the economy normally visible in production, employment, real income, and other indicators. In addition, although these models usually find good in-sample accuracy to identify business cycles, the time delay in the publication of GDP data makes them useless when addressing the current situation of the economy in real time. For example, on the date in which this paper is written (October 4th 2009) the latest available US GDP release is for the second quarter of 2009. When making a statement about the probability of being in a recession today, one needs to calculate the probability from the latest observation available (2009.Q2) and with this information, to forecast the probability of being in a recession in the current quarter (2009.Q4). This implies missing extremely valuable information from June to October.

Second, following the lines of Diebold and Rudebusch (1996), some recent proposals (Kim and Yoo, 1995, Chauvet, 1998, and Kim and Nelson, 1998) estimate different versions of Markov-switching dynamic factor models which capture the notions of both co-movements across business cycle monthly indicators and regime switching. The indicators

used by these proposals build on the tradition of the linear single-index dynamic factor model of Stock and Watson (1991): following the logic of national accounting that GDP is approximated from the income side, the supply side and the demand side, they choose industrial production index (supply side), total sales (demand side), real personal income (income side) and they add an employment variable to capture the idea that productivity does not change dramatically from one period to the other. Although the inference about the business cycle states from these proposals has been very accurate, there were still important gaps. The main restriction is the need of complete data sets which implies that the models are not able to deal with some typical problems associated to real time forecasting such as mixing frequencies and ragged ends.

We think that incorporating both features in these non linear multivariate dynamic models provides a powerful econometric technique to technically handle the definition of green shoots, and the timing, the intensity and the form of the recovery.

Overall, our results suggest that the Markov-switching dynamic factor model is a potentially very useful filter to transform the information of a broad set of economic indicators into recession probabilities. We provide some formal evidence regarding the speed with which the real-time monitoring can identify the latest turning point in US, the Euro area¹ and Spain. As we expected, our model confirms the presence of green shoots in the US, the Euro area and Spain with different timing in different economies. In addition, we confirm some leading time from the number of quotes in the press to the actual confirmation of the data, even when the data include variables with clear expectations contents.

The paper is structured as follows. Section 2 describes the results of standard univariate analysis using GDP. Section 3 discusses the multivariate extension in a small scale model with no expectation variables. Section 4 analyzes the results of a medium scale model with expectation variables. Section 5 looks at the real time analysis of the forecasts. Section 6 concludes.

¹The details about the application of the model to the Euro area can be found in Camacho et al. (2010b).

2 Univariate analysis

Hamilton (1989) originally proposed an excellent framework to deal with business cycles in time series. Let us assume that the expected value of a time series, y_t , switches between two separate business cycle states which are usually known as expansions and recessions. In mathematical notation, it is assumed that $E(y_t) = \mu_0$ if the economy is in an expansion, and that $E(y_t) = \mu_1$ if the economy is in a recession. In his seminal proposal, Hamilton (1989) assumes that the time series is GDP growth and that apart from the transitions between expansions and recessions the series exhibits autoregressive dynamics. His econometric specification becomes

$$y_t = \mu_{s_t} + u_t, \quad (1)$$

where u_t follows an AR(4) process, and s_t is an unobservable state variable that takes on the value of 0 in expansions and the value of 1 in recessions. The dynamics of the state variable is supposed to follow a Markov chain of order one, which implies that

$$p(s_t = j | s_{t-1} = i, s_{t-2} = h, \dots, I_{t-1}) = p(s_t = j | s_{t-1} = i) = p_{ij} \quad (2)$$

where $i, j = 0, 1$, and I_t is the information set up to period t .

Camacho and Perez Quiros (2007) have recently shown that when appropriately accounting for the different switches between regime dependent means, then the serial correlation that characterizes the regime switches is substituting for the serial correlation that would normally be modeled via autoregressive structures. Accordingly, models that accurately capture the sequence of recessions and expansions are dynamically complete and no additional autoregressive parameters are required to capture the dynamics of the series. In these cases, the model proposed will be

$$y_t = \mu_{s_t} + \varepsilon_t, \quad (3)$$

where ε_t is an uncorrelated sequence of Gaussian errors with zero mean and variance σ^2 .

Figure 1 plots the real GDP growth rate of each of the three economies for the longest available sample (downloaded on October 4th 2009) along with the shaded areas that refer to the NBER-designated recessions. In the US, the data covers from 1953.1 to 2009.2

and for the Euro area from 1991.1 to 2009.2. In the case of Spain, using the latest data set published by the National Statistical Institute, we have data from 1974.1 to 2009.2. As assumed by the simple univariate Markov switching model the GDP series exhibits negative growth rates during most of the NBER recessions. Figure 2 shows the in sample filtered probabilities of being in recessions as estimated in the US, the Euro area and the Spanish economies from the model in (3). As depicted in the US figure, using just GDP data alone without any reference to what NBER may have said, we come up with a very similar recessions dating to the one that the NBER has traditionally relied on. In addition, with the latest available information, the probability that these economies are in recessions is still very high.

The maximum likelihood coefficient estimates of model (3), which are reported in Table 1, reveal some interesting results. First, regarding the sample period and the economic area considered, including the recent recession (left hand panel) leads to positive means in state $s_t = 0$ and negative means in the regime represented by $s_t = 1$. Accordingly, we can associate the first regime with expansions and the second regime with recessions. Second, overall expansions are more persistent than recessions since the estimates of p_{00} are higher than those of p_{11} . The international comparison reveals that although recessions seem to be more persistent in the case of the Euro area and Spain, it seems that this result largely depends on the different samples used to obtain parameter estimates since the persistence of recessions in US becomes similar to that of the European cases when using comparable samples for US. Third, conditional on being in state i , one can derive the expected number of months that the business cycle phases prevail as $(1 - p_{ii})^{-1}$, and the expected amplitude of this state as $\mu_i (1 - p_{ii})^{-1}$. In US, the expected duration and amplitude of a typical expansion are 16.67 quarters and 17%, while those figures fall in the case of recessions to 3.84 quarters and 1.61%. These estimates accord with the well-known fact that recessions are shorter and milder than expansions on average. To examine the extent to which the current recession is different from the previous ones, the right hand panel of Table 1 shows the estimates of the Markov switching parameters obtained from a sample which ends in late 2007. In this case, recessions are expected to last 3.44 quarters and are expected to imply a loss of 1.17% which are close to the previous estimates. So far, this recession has

lasted 7 quarters and has implied a loss of 3.14% so it is actually being longer and harder than expected. Fourth, concerning international comparisons it is worth pointing out that in the Euro area recessions are expected to be longer (5.88 quarters) and deeper (losses of 4.52%) but this results is largely due to the short length in the European GDP. In the case of Spain, the expected recessions are the longest (8.33 quarters) but milder (loss of 0.66%).

The ability of univariate Markov switching models to compute inference of business cycles in real time deserves a final remark. The high commonality in switching times of probabilities and the US business cycle phases identified by NBER observed in Figure 2 gives the impression that the simple univariate Markov-switching models applied to GDP fits the business cycle extremely well. But the good in-sample results of this figure are somewhat tricky in that it plots the filtered probabilities of being in recession for given quarter by using GDP growth rates up to that quarter which are obviously not available when computing inferences in real time. Since the GDP publication lag is about 45 days after the end of the respective quarter, the latest quarter for which inference can be computed in this way is that of the second quarter of 2009. To infer the probability of recession for the current fourth quarter of 2009 one needs to compute two-period ahead forecasts of the probabilities. To analyze the effect of the large publication delay of GDP in business cycle inferences, Figure 3 plots the two-period ahead forecasts of the probabilities.² This figure allows us to put a question mark on the ability of univariate Markov-switching models to infer recession probabilities in real time: for all the recessions the signals to monitor the current business cycle developments are mild and come too late. Accordingly, the natural way to proceed seems to be adding economic monthly indicators which incorporate more timely information about the state of the business cycle.

²Note that this exercise does not account for the effect of data revisions which would amplify the deterioration of the in-sample identification of business cycles.

3 Multivariate analysis

Enhancing the Markov-switching model of GDP to incorporate economic indicators can be helpful for two reasons. First, because they are published with shorter delay so they can incorporate more timely information. Second, because if they are synchronized with GDP, they could help to increase the signal of turning points. For this purpose, Kim and Yoo (1995), Chauvet (1998) and Kim and Nelson (1998) combined the dynamic-factor and Markov-switching frameworks to incorporate the two main characteristics of the business cycle indicators: comovements and asymmetries. Recent applications of the model can be found in Chauvet and Hamilton (2006) and Chauvet and Piger (2008). However, these empirical proposals do not incorporate data measured at different frequencies or unbalanced panel data sets. Both features imply dealing with missing data. Camacho, Perez-Quiros and Poncela (2010a) justify that it is possible to extend to the Markov switching context, the proposal of Mariano and Murasawa (2003) to deal with mixing frequencies in the linear framework. In this paper, we apply it.

3.1 Theoretical framework

Let us start with a single-index dynamic factor model whose common factor follows a Markov switching process. Let $\mathbf{x}_t = (x_{1,t}, \dots, x_{N,t})'$ be the vector of N observed time series which is generated by a non-observed common factor, f_t , and N specific or idiosyncratic components

$$\begin{array}{ccccccc} \mathbf{x}_t & = & \mathbf{\Lambda}(B) & f_t & + & \mathbf{u}_t, & \\ N \times 1 & & N \times 1 & 1 \times 1 & & N \times 1 & \end{array} \quad (4)$$

where $\mathbf{\Lambda}(B) = (\Lambda_1(B), \Lambda_2(B), \dots, \Lambda_N(B))'$ is the factor loading matrix, with $\Lambda_i(B) = \beta_0^i + \beta_1^i B + \dots + \beta_{q_i}^i B^{q_i}$, being B the backshift operator. The common factor follows a Markov switching autoregressive process with changing mean:

$$f_t = \mu_{s_t} + \frac{a_t}{\phi(B)}, \quad (5)$$

where $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$. We assume that s_t evolves according to an irreducible 2-state Markov chain whose transition probabilities are defined by (2). We also consider

that the idiosyncratic components have the dynamic structure

$$\begin{array}{ccc} \mathbf{F}(B) & \mathbf{u}_t & = \quad \boldsymbol{\epsilon}_t, \\ N \times N & N \times 1 & \quad N \times 1 \end{array} \quad (6)$$

where $\mathbf{F}(B) = \text{diag}(F_i(B))$ is a diagonal matrix that collects the specific dynamics of each idiosyncratic shock, with $F_i(B) = 1 - \phi_{i1}B - \dots - \phi_{ip_i}B^{p_i}$, $i = 1, \dots, N$, and $\boldsymbol{\epsilon}_t$ is multivariate zero mean white noise with diagonal covariance matrix $\boldsymbol{\Sigma}_\epsilon$.³

3.2 Variables selection

According to the feasibility restrictions of nonlinear models, the number of variables that can be analyzed must be small and the selection of variables to be included in the analysis needs to be carefully done. However, note that the variable selection problem does not only affect small scale models. The standard linear large scale models never use all the time series available in real time at all levels of disaggregation for all the countries and regions used in the analysis. In addition, the level of complexity that large scale models incorporate to real time analysis is not always justified. In the context of forecasting, Boivin and Ng (2006) have recently suggested that, given the small number of categories that we have in macroeconomic data, the forecast accuracy does not necessarily increase with the number of series included in the model because these series might only add cross correlation in the idiosyncratic noise. Finally, Banbura and Rünstler (2007) find that most of the predictive content of their large scale model is contained in a small set of variables.

Accordingly, we start from a simple model, following the suggestion of Stock and Watson (1991). Their idea follows the logic of national accounting that robust estimates of GDP are obtained by computing GDP from the income side, the supply side and the demand side. Therefore, to obtain robust estimates of activity they choose industrial production index (supply side), total sales (demand side), real personal income less transfer payments (income side) and they add an employment variable to capture the idea that productivity does not change dramatically from one period to the other. In addition, due

³Note that the model allows for common shocks to the economy (a_t) as well as for specific shocks to each economic indicator ($\boldsymbol{\epsilon}_t$).

to its importance in determining the state of the business cycle, we enlarge this model by including the GDP series.

Table 2 shows the description of the series used for each economy, the sample period available and the data source. As we can observe in the table, the data available have all the characteristics mentioned before, ragged ends and mixing frequencies in a multivariate framework. For the case of the Euro area, there are no income variables available so we use wages and salaries released by Eurostat. In addition, the employment series is not available monthly but quarterly. Due to data availability, in Spain we use Large Firm Sales (from the Spanish Revenue Service) instead of retail sales, Social Security Contributors instead of employment, and Salaries Paid (from the Spanish Revenue Service) instead of income.

Some of the selected variables are available monthly while other ones are available quarterly. To use all of them inside the model, we are going to convert all quarterly variables into their monthly counterparts, as in Mariano and Murasawa (2003). This leads to some of the lag polynomials that appear on the factor loading matrix $\Lambda(B)$ in (4).

One final remark deserves some comments. In Spain some series are published monthly but refer to annual growth rates. To obtain comparable results, we transform all the monthly indicators into annual growth rates. Accordingly, the annual growth rates, x_t , can be expressed as the sum of lagged monthly underlying variables:

$$x_t = \sum_{j=0}^{11} z_{t-j}. \quad (7)$$

This also leads to lag polynomials in the factor loading matrix $\Lambda(B)$ in (4).

3.3 Specification and estimation of the model

If we assume that all the variables are observed at monthly frequency, the model admits a simple state space representation. Let us assume that the idiosyncratic part in all the quarterly series in the model is AR(2), which implies that the monthly underlying series for quarterly observations are AR(6). Let f_t^x be the 12×1 vector whose components are the

common factor and its first eleven lags and u_{xt} the vector that contains the idiosyncratic components and their lags for all the variables in the model; finally, let us define the state vector of unobserved components (common or idiosyncratic) h_t as

$$h_t = (f_t^{x'}, u_{xt}')'$$

Hence, one can state the measurement equation (linking the unobserved common factors and idiosyncratic components to the observed variables) as

$$\begin{array}{ccccccc} x_t & = & H_t & h_t & + & v_t & \\ N \times 1 & & N \times k & k \times 1 & & N \times 1 & \end{array}, \quad (8)$$

where v_t is multivariate noise $(0, R)$ with R diagonal. The H_t matrix contains as unknown parameters the factor loadings λ_{is} , $i = 1, \dots, N$ that reflect how much the common factor loads into each observed series. The transition equation (that collects the dynamics in the model) is given by

$$\begin{array}{ccccccc} h_t & = & m_{s_t} & + & F & h_{t-1} & + & V_t, \\ k \times 1 & & k \times 1 & & k \times k & k \times 1 & & k \times 1 \end{array} \quad (9)$$

where V_t is multivariate noise $(0, Q)$ with Q diagonal. All the technical details about the model with the precise definition of all the variables and matrices can be found in Camacho et al (2010a,b). Missing observations (due to mixing frequencies or ragged ends) are handled by adapting the procedure in Mariano and Murasawa (2003) to this nonlinear framework as in Camacho et al. (2010a,b). The model is estimated by approximate maximum likelihood as in Kim and Yoo (1995) and Chauvet (1998). We compute estimates of the filtered probabilities of recession $P(s_t = 1|I_t)$ to assess the state of the business cycle.

3.4 Empirical results

Table 3 shows the maximum likelihood estimates of the more important coefficients for the three economies considered. As expected, GDP and to less extent IPI exhibit the higher factor loadings. Although the magnitude of the other factor loadings depends on the particular country, that of wages in the euro area is not significant probably because wages are a bad proxy of income variable. In addition, the mean of the common factor

is positive in the first state and negative in the second state so we tentatively call them expansions and recessions.

Table 3 is complemented with Figures 4, 5, and 6 where we plot the common factor estimated for each economy and the filtered probabilities of recession periods. In spite of the potential problems associated to the short length of the Euro area aggregates, the evolution of the factors are in clear concordance with the business cycle developments in US and Spain and contains relevant visual information on the intensity of their expansions and recessions.⁴ In addition, these figures plot the evolution of the inferred probabilities of recessions whose peaks are in agreement with the official recessions in US and with the generally accepted downturns in Spain. With respect to the latest probabilities, the models infer probabilities of recession in September 2009 of 0.3 in the USA and of 0.7 in Spain. Both represent significant improvements with respect to probabilities of recessions above 0.9 that both economies exhibited in April and May.

Although it is not easy to determine the threshold of the filtered recession probabilities that marks the end of the recession, let us consider that an economy is in recession if the inferred probabilities of occurrence of this regime go above 0.5.⁵ According to the previous results which use information up to October 4th 2009, the US economy is already out of the recession while the Spanish economy seems to be on its way to the end of the recession period. Interestingly, the timing of the signs of recovery is very distant from the timing about the popularization of the term *green shoots*. The maximum number of searches in Google of this term is in the week of May 10th, and it is still extremely high until the week of the 24th of May. However, the set of hard indicators that we use to infer the recession probabilities did not show up any real change in economic activity in these days.⁶ In the search of explanations of this different timing, in the next section we will enlarge the models with some indicators that include expectations of the agents about the future of the economy.

⁴The problems associated with short lengths of Euro area data will be addressed in the next section.

⁵Hamilton (1989) used this threshold in its seminal proposal.

⁶In May 2009, Marcelle Chauvet still pointed out a probability of US recession of 0.77. See <http://sites.google.com/site/marcellechauvet/probabilities-of-recession>.

4 Enlarging the original specification

In the related literature of small scale dynamic factor models, there are two linear proposals which incorporate indicators of expectations.⁷ The former is the so called Euro-Sting model by Camacho, and Perez Quiros (2010) and analyzes the Euro area economic developments. The latter is known as the Spain-Sting model and has been estimated for the Spanish economy.⁸ Although both models have been designed in linear frameworks, we enlarge them to account for non-linear dynamics.

The Euro area model represents successive enlargements of the original model presented in the previous section. Starting from Stock and Watson (1991), we extend the model to capture the set of expectations indicators that are more promptly available in the Euro area.⁹ In that sense, we add Euro-zone Economic Sentiment Indicator (ESI), the German business climate index (IFO), the Belgian overall business indicator (BNB), and the Euro area Purchasing Managers confidence Indexes (PMI) in the services (PMIS) and manufacturing sectors (PMIM). The main characteristics of these soft indicators are that they represent market expectations, and that they are promptly available so they can be observed on a timely basis within the reference month.

Once the model is enlarged with these indicators, we propose a method to decide whether new indicators should be added to this core. The method, which is based on the assumption that the primary focus of the model is to provide forecasts of GDP growth, consists of adding a variable whenever it increases the percentage of variance of GDP growth explained by the common factor. With this method, we end up adding extra-Euro area exports and the Industrial New Orders index (INO, total manufacturing working on

⁷For feasibility restrictions, recall that we are precluded from using large scale models since we want to propose Markov-switching extensions.

⁸Aruoba, Diebold and Scotti (2009) proposed an interesting small scale factor model for the US which uses weekly variables. We still have not developed the non linear extension of this model, but it is left for further research.

⁹We do not include Personal Disposable Income because we do not have this series for the euro area. As we showed in the previous section Salaries is a poor proxy variable for Income.

orders) to the set of core variables.¹⁰

The model for the Spanish economy shares the same philosophy. It starts from the model estimated in the previous section, Industrial Production (excluding construction), total sales of large firms of Agencia Tributaria (Spanish Internal Revenue Service), social security contributors and total wages paid by the large firms (Spanish Internal Revenue Service) as the set of core variables. The soft indicators used for the Spanish economy are the Product Manufacturing Index, and the Confidence Index produced by the European Commission. To avoid overlap in the information from the supply side, we choose the Industrial Confidence Indicator as an Index of the production sector, and the PMI services (PMISE) for the production of the services sector. From the demand side, we choose the Retail Trade Index.

The resulting model fits the GDP data very precisely with a variance of GDP explained by the common factor of 79%. However, given the importance of the construction sector in explaining the recent boom and the crisis in Spain, we try to separate the supply side in Industry, Services and Construction by choosing the most reliable series of each of these sectors. For Industry, we choose the Industrial Production Series, for Services we use the Overnight Stays (Tourism represents more than 11% of Spanish GDP) and for Construction, we select Consumption of Cement. The demand side was enhanced with Exports (as complementary of the internal demand captured by sales of large firms) and Imports (as indicator of demand that can not be satisfied with internal firms). Finally, we add Total Credit in order to include a variable that captures the transmission of the financial crisis to the real economy. With these enlargements, the variance of GDP explained by the factor increases to 80%.¹¹

Following the lines suggested in Section 3, we extend these two models to account for non-linear Markov-switching dynamics. The results of the estimation are displayed in Table 4 and the filtered probabilities of being in a recession for each date of the sample are

¹⁰To account for data revisions in the GDP, we include flash and first estimates as additional variables in the model.

¹¹This is not a minor increase. Additional variables are usually correlated with the idiosyncratic part of some of the initial variables which implies that the estimation of the factor is biased toward this subgroup. In this case, the variance of GDP explained by the factor usually decreases.

plotted in Figures 7 and 8. According to the results, the expected duration of a recession in the Euro area and Spain is 15 and 13 months respectively. In the Euro area, the probability of recession grew from 0.11 in March 2008 to 0.98 in June 2008 and fell from 0.96 in February 2009 to 0.07 in April 2009. Hence, the model dates the peak in April 2008 and the trough in March 2009 so the recession lasted about one year which was even shorter than expected. In Spain, the probability of recession grew from 0.22 in April 2008 to 0.99 in June so the model tentatively dates the peak in May 2008.¹² Although the model shows signals of recoveries, even when we consider indicators of expectations it is still soon to consider that the recession is over since the probability of recession is in September around 0.3. However, even if the recession ended early, its duration would be longer than one year so it will be longer than expected.

With respect to the amplitude of the last recession, it is worth mentioning that translating the negative growth of the factor associated to the recession period (in Table 4 the estimates are -2.03 and -2.31 for the Euro area and Spain respectively) requires some algebra. First, we obtain the standardized expected quarterly growth rates (-6.09 for the Euro area and -6.93 for Spain). However in order to transform them into a expected quarterly growth rate in a recession period, we have to take into account that they are multiplied by 0.28 and 0.23, respectively, and they have to be transformed with the standard deviation and the mean of the observed variables (0.68 and 0.57 for Spain). After these transformations, the expected growth rate in recessions is -0.79 for the Euro area and -0.22 for Spain. Therefore, the expected amplitude of the recessions would be around 3% in the Euro area, and 1% in the case of Spain, much smaller, in the latter case, than the more than 4% already lost in the current recession.

5 Real time analysis

We based the previous investigation of the business cycle on the analysis of filtered probabilities which are inferences about the state of the economy using currently available

¹²There is a sooner jump in the probability from 0.01 in February to 0.50 in March.

information.¹³ However, the inferred probabilities, which are plotted in Figures 2 to 8, are not computed from the exact amount of information that would be available at the date of the forecasts but from end-of-sample vintages that incorporate data revisions. These data availability restrictions may lead to unrealistically good results, especially concerning the ability of the model to anticipate turning points. To examine the true real-time ability of the model in anticipating turning points, we constructed current-vintage data sets and compute real time probabilities of the forecast which are updated daily in the last two years.

Figure 9 plots the daily probabilities of recession in the Euro area which are computed by using the exact information that would be available each day of the forecasting period. According to this figure, in mid-July 2008 the probability of recession increased up to values that are very close to one. It is worth noting that this prompt signal of bad news about the state of the Euro area economy represents an improvement in the timing of turning points identifications with respect to other standard dating methods. In July, the latest GDP available the figure of 2008.2 and it was still a positive and very high number (0.78). Since the GDP figures for the second and third quarters of 2008 were negative, if one considered that two consecutive falls of GDP growth mark the start of a recession, the recession would not be formally identified before the publication day of the third quarter GDP, November 15th 2008.

In addition, Figure 9 would help us to examine to what extent *green shoots* are real in the Euro area according to the low probabilities of recession definition. About mid-April 2009, the probability of recession dramatically dropped from values of about 0.8 to values close to zero. As in the case of the peak, we find evidence of a trough that marks the end of the recession before other standard dating methods since the latest available figure of GDP growth was still very negative (-2.45% for the first quarter of 2009). Finally, let us examine the mechanics behind these good signals that mark the changes in probabilities. When the probabilities of recession were still high at the beginning

¹³Alternatively, one can obtain smoothed probabilities, which are computed from full-sample information. However, filtered probabilities provide more reliable pictures of the models' accuracy to infer states probabilities since they use information that is not available when computing inferences.

of April, the values of some soft indicators such as ESI and the PMIM were 64.6 and 33.9, respectively. However, the following realizations since that date were 82.8 and 49.3 which imply significant improvements. In addition, the good news were confirmed by hard indicators when they become available: IPI, Retail Sales, INO and Exports experienced a sharp increase from -2.38,-.78, -3.36 or -10.7 to -0.28, -0.21, 2.63 and 4.06, respectively.

Let us now move to the Spanish economy. The inferred daily probabilities of recession, which are displayed in Figure 10, exhibit a similar but unsynchronized pattern. First, as in the case of the Euro area we could call a recession already by mid-February of 2008, while still the latest available figure of GDP was for 2007.4 and still announced an acceleration of the economy from 0.7 in the third quarter of 2007 to 0.8 in the unrevised data for the fourth quarter of 2007. As in the Euro area case, the first indicators that showed early signals of deterioration were the soft indicators, Industrial Confidence Index, PMISE, Retail Trade index that went down in two months. In particular, the Industrial Confidence Index fell from -4.2 to -9.3, the Retail Trade index dropped from -13.1 to -26.3, and PMIS fell from 51 to 46.¹⁴ Again, the signals of recession were confirmed by the next figures of the hard indicators.

In the Spanish case, the probabilities of recession remain close to one until mid-September 2009. The drop in probabilities observed since then was early signaled by the soft indicators. In particular, PMI services grew from 40.8 in August to 45.3 in September, ICI improves upon from its historical low record of -40 in March to -28 in September, and Retail Trade Index grew from -29.4 to -21.8. These latest available releases are, in most cases, similar to the figures observed in the beginning of 2008 and then, compatible with positive growth rates of GDP. However, it is maybe too early to consider that the reduction in recession probabilities up to values of about 0.3 can be considered either clear signals of green shoots or yellow weeds only. We should wait until good news are confirmed or denied by the releases of hard indicators which are not available yet.

Let us point out a final remark with respect to the potential lack of synchronicity between the aim of searching for green shoots and the actual confirmation from macroeco-

¹⁴For some of these indicators, the drops observed in these figures were the most severe in the history of the indicators.

conomic data that the recovery is a fact. For this purpose, we plot in Figure 11 the evolution of the number of searches of the term green shoots in the world and in the US during the year 2009.¹⁵ The signals of recoveries that we found in the Euro area in mid-April 2009 are roughly synchronized with the maximum number of search of green shoots. However, the signals of recoveries exhibit clear lags with respect to the global anxiety in looking for green shoots.

6 Conclusions

Over the year 2009 the term *green shoots* has been highly popularized as a term that represents the beginnings of economic growth after a recession. But the term is very imprecise and has not been defined in economically meaningful ways. In this paper, we define green shoots as low probabilities of recessions. In this sense, we provide the term with economic sense so it allows us to examine where, when and how the recovery comes. To infer the probabilities of recession, we propose a Markov-switching extension of the single-index dynamic factor model proposed by Camacho and Perez Quiros (2010). The model is able to handle indicators which are available at different frequencies, and to account for the gaps that characterize the ragged edges behind the asynchronous data publication.

Using data up to October 4th 2009, we find symptoms of recoveries in US, the Euro area and Spain, with some differences in the timing and clarity of those symptoms. When the analysis is developed with real-time data sets, in the Euro area the probabilities of recession are close to zero since May 2009 which is the month where the number of searches of the term green shoots in Google becomes maximum. In Spain, the signals of recoveries are milder, especially when soft indicators are excluded from the data vintages.

Finally, there is a growing literature which concerns how one should expect that the recession will finish, that is, whether it will be a “V-shaped” or an “L-shaped” recession. The former type of recessions refers to the case in which the economy springs back rapidly from its slump and is viewed as evidence in favor of recessions having only temporary

¹⁵To construct Figure 11, we used Google Trends which is a public web facility of Google that shows how often a particular search term is entered in a given period.

effects. The latter type of recessions are followed by flat recoveries and is viewed as having permanent effects on the level of production. In this context, Camacho, Perez Quiros and Rodriguez (2009) present evidence about the loss of the high-growth phase of the cycle typically observed at the end of the US recessions and show that the loss of the “plucking effect” can explain part of the Great Moderation.¹⁶ They postulate that these two phenomena may be due to changes in inventory management brought about by improvements in information and communications technologies.

¹⁶Camacho, Perez Quiros and Saiz (2008) show that the absence of the third-phase in the last recoveries holds for all major industrialized economies.

References

- [1] Aruoba, B., Diebold, F.X. and Scotti, Ch. 2009. Real-Time Measurement of Business Conditions *Journal of Business and Economic Statistics* 27: 417-427.
- [2] Boivin, J., and Ng, S. 2006. Are more data always better for factor analysis? *Journal of Econometrics*, 132: 169-194.
- [3] Banbura, M., and Rünstler, G. 2007. A look into the factor model black box: publication lags and the role of hard and soft data in forecasting GDP. European Central Bank working paper 751.
- [4] Camacho, M., and Perez Quiros, G. 2007. Jump-and-rest effect of U.S. business cycles. *Studies in Nonlinear Dynamics and Econometrics* 11(4): article 3.
- [5] Camacho, M., and Perez Quiros, G. 2010. Introducing the Euro-STING: Short Term INDicator of Euro Area Growth. *Journal of Applied Econometrics*, forthcoming.
- [6] Camacho, M., Perez Quiros, G., and Poncela, P. 2010a. Markov-switching dynamic factor models in real time. Mimeo.
- [7] Camacho, M., Perez Quiros, G., and Poncela, P. 2010b. The Euro-STING in forecasting the state of the business cycle. Mimeo
- [8] Camacho, M., Perez Quiros, G., and Rodríguez, H. 2009. Are the high-growth recovery periods over? Banco de España working paper 1209.
- [9] Camacho, M., Perez Quiros, G., and Saiz, L. 2008. Do European business cycles look like one? *Journal of Economic Dynamics and Control* 32: 2165-2190.
- [10] Chauvet, M. 1998. An econometric characterization of business cycle dynamics with factor structure and regime switches. *International Economic Review* 39: 969-96.
- [11] Chauvet, M., and Hamilton, J. 2006. Dating business cycle turning points in real time. In *Nonlinear Time Series Analysis of Business Cycles*, edited by C. Milas, P. Rothman, and D. Van Dijk. Amsterdam Elsevier Science.

- [12] Chauvet, M., and Piger, J. 2008. A comparison of the real-time performance of business cycle dating methods. *Journal of Business and Economic Statistics* 26: 42-49.
- [13] Diebold, F., and Rudebusch, G. 1996. Measuring business cycles: A modern perspective. *Review of Economics and Statistics* 78, 67-77.
- [14] Hamilton, J. 1989. A new approach to the economic analysis of nonstationary time series and the business cycles. *Econometrica* 57: 357-384.
- [15] Kim, C., and Nelson, C. 1998. Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching. *Review of Economics and Statistics* 80: 188-201.
- [16] Kim, C., and Yoo, J.S. 1995. New index of coincident indicators: A multivariate Markov switching factor model approach. *Journal of Monetary Economics* 36: 607-630.
- [17] Mariano, R., and Murasawa, Y. 2003. A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics* 18: 427-443
- [18] Stock, J., and Watson, M. 1991. A probability model of the coincident economic indicators. In *Leading Economic Indicators: New Approaches and Forecasting Records*, edited by K. Lahiri and G. Moore. Cambridge University Press.

Table 1. Markov-switching estimates

Including the latest recession						Excluding the latest recession					
sample	μ_0	μ_1	σ^2	p_{00}	p_{11}	sample	μ_0	μ_1	σ^2	p_{00}	p_{11}
US											
53.1	1.02	-0.42	0.60	0.94	0.74	53.1	1.02	-0.34	0.60	0.94	0.71
09.2	(0.08)	(0.26)	(0.06)	(0.02)	(0.10)	07.4	(0.08)	(0.31)	(0.06)	(0.03)	(0.12)
74.1	0.92	-0.53	0.50	0.95	0.74	74.1	0.93	-0.43	0.50	0.96	0.69
09.2	(0.07)	(0.27)	(0.06)	(0.02)	(0.13)	07.4	(0.07)	(0.35)	(0.06)	(0.02)	(0.12)
91.3	0.76	-0.92	0.24	0.99	0.94	91.3	0.95	0.46	0.18	0.92	0.88
09.2	(0.07)	(0.26)	(0.06)	(0.02)	(0.10)	07.4	(0.11)	(0.16)	(0.04)	(0.09)	(0.11)
Euro area											
91.3	0.56	-0.77	0.17	0.97	0.83	91.3	0.64	0.01	0.09	0.95	0.80
09.2	(0.05)	(0.15)	(0.03)	(0.02)	(0.14)	07.4	(0.05)	(0.11)	(0.02)	(0.04)	(0.13)
Spain											
74.1	0.91	-0.08	0.21	0.97	0.88	74.1	1.17	0.45	0.17	0.93	0.95
09.2	(0.05)	(0.12)	(0.03)	(0.02)	(0.06)	07.4	(0.13)	(0.09)	(0.02)	(0.04)	(0.03)
91.3	0.84	-0.74	0.12	0.97	0.86	91.3	0.85	-0.51	0.09	0.98	0.78
09.2	(0.05)	(0.12)	(0.03)	(0.02)	(0.06)	07.4	(0.04)	(0.14)	(0.02)	(0.01)	(0.18)

Notes. The estimated model is $y_t = \mu_{s_t} + \varepsilon_t$, where y_t is GDP growth rate, $\varepsilon_t \sim iidN(0, \sigma^2)$, and $p(s_t = i/s_{t-1} = j) = p_{ij}$.

Table 2. Indicators used in (5 series) Markov-switching factor model

Series	Sample	Source	Frequency
US			
Industrial production	60.01-09.08	Datastream	monthly
Retail sales	60.01-09.08	Datastream	monthly
Employees in non farm	60.01-09.09	Bureau of Labor Statistics	monthly
Personal income less transfer payments	60.01-09.08	Datastream	monthly
GDP	60.I-09.II	St. Louis FRED	quarterly
Euro area			
Industrial production	90.01-09.07	Eurostat	monthly
Retail sales	95.01-09.07	Eurostat	monthly
Employment	91.I-09.II	Eurostat	quarterly
Wages	95.01-08.12	Eurostat	monthly
GDP	90.I-09.II	Eurostat	quarterly
Spain			
Industrial production	83.01-09.08	INE	monthly
Large firm sales	95.01-09.08	Agencia Tributaria	monthly
Employment	83.01-09.09	Ministerio de Trabajo	monthly
Wages paid by large firms	95.01-08.12	Agencia Tributaria	monthly
GDP	83.I-09.II	INE	quarterly

Notes. To describe the sample, first two digits refer to the year, and last (two) digit(s) refers to the (month) quarter.

Table 3. (Main) Parameter estimates from (5 series) Markov-switching factor model

Factor loadings (λ_i)					Markov-switching parameters			
GDP	Income	Sales	IP	Employ	μ_0	μ_1	p_{00}	p_{11}
US								
0.36 (0.02)	0.07 (0.01)	0.08 (0.01)	0.16 (0.01)	0.05 (0.01)	0.22 (0.06)	-1.11 (0.13)	0.98 (0.01)	0.93 (0.03)
Euro area								
0.49 (0.04)	0.006 (0.03)	0.08 (0.03)	0.18 (0.01)	0.03 (0.01)	0.08 (0.08)	-2.03 (0.47)	0.99 (0.01)	0.91 (0.13)
Spain								
0.26 (0.04)	0.08 (0.02)	0.08 (0.01)	0.11 (0.01)	0.06 (0.01)	0.14 (0.07)	-2.06 (0.28)	0.99 (0.01)	0.94 (0.08)

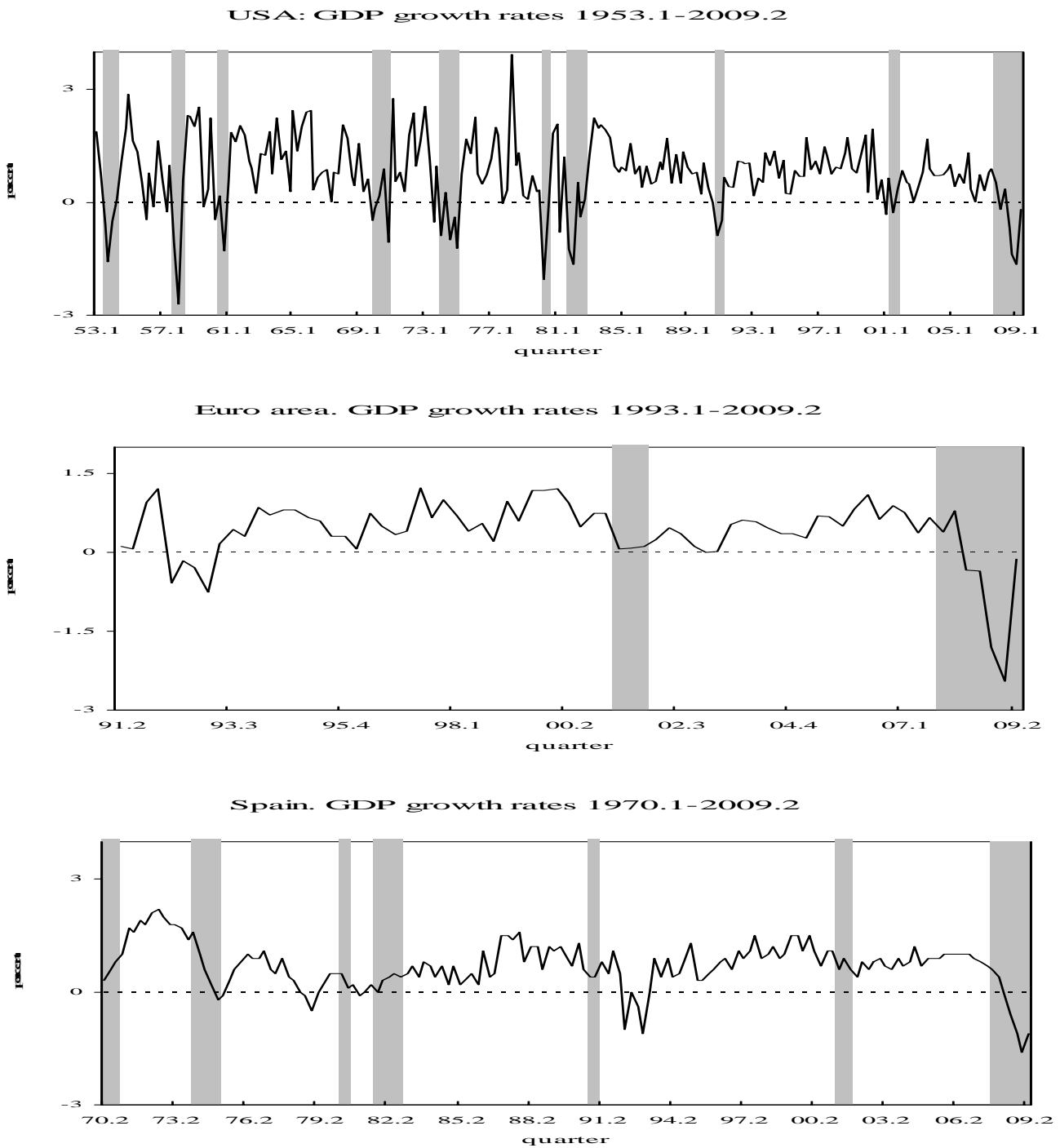
Notes. The factor loadings (standard errors are in parentheses) measure the correlation between the common factor and each of the indicators appearing in columns. See Table 2 for a description of the indicators.

Table 4. (Main) Parameter estimates from Markov-switching extensions of Euro-STING and Spain-STING

Euro area			Spain		
Indicator	Estimates	Standard deviations	Indicator	Estimates	Standard deviations
GDP	0.286	(0.034)	GDP	0.236	(0.044)
IPI	0.361	(0.045)	Wages	0.073	(0.018)
Sales	0.101	(0.033)	Sales	0.083	(0.014)
INO	0.331	(0.047)	IPI	0.095	(0.009)
Export	0.202	(0.052)	Employment	0.062	(0.003)
ESI	0.078	(0.010)	Export	0.069	(0.016)
BNB	0.100	(0.023)	Imports	0.091	(0.012)
IFO	0.084	(0.012)	Over-stays	0.055	(0.020)
PMIM	0.113	(0.014)	Cement	0.076	(0.012)
PMIS	0.101	(0.018)	Credit	0.018	(0.007)
Employ	0.125	(0.037)	ICI	0.061	(0.009)
			Ret. Trade Index	0.026	(0.019)
			PMIS	0.047	(0.019)
Markov-switching parameters					
μ_0	0.37	(0.11)	μ_0	0.22	(0.09)
μ_1	-2.03	(0.38)	μ_1	-2.31	(0.31)
p_{00}	0.97	(0.02)	p_{00}	0.99	(0.01)
p_{11}	0.93	(0.06)	p_{11}	0.92	(0.09)

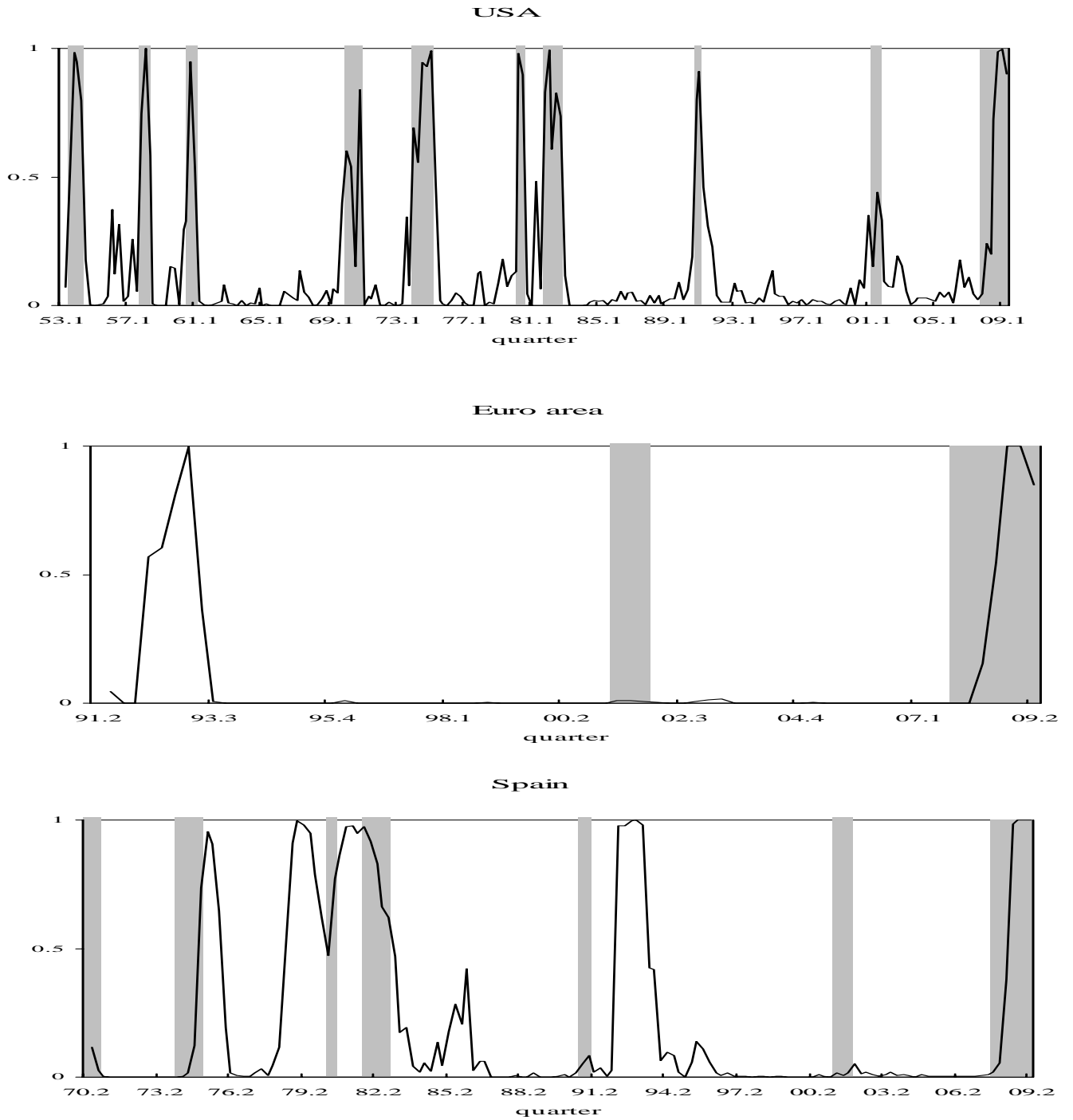
Notes. First block refers to factor loading parameters. Second block refers to within expansion means (μ_0), within recession means (μ_1), and probabilities of staying in expansions (p_{00}) and recessions (p_{11}).

Figure 1. Growth rates of GDP



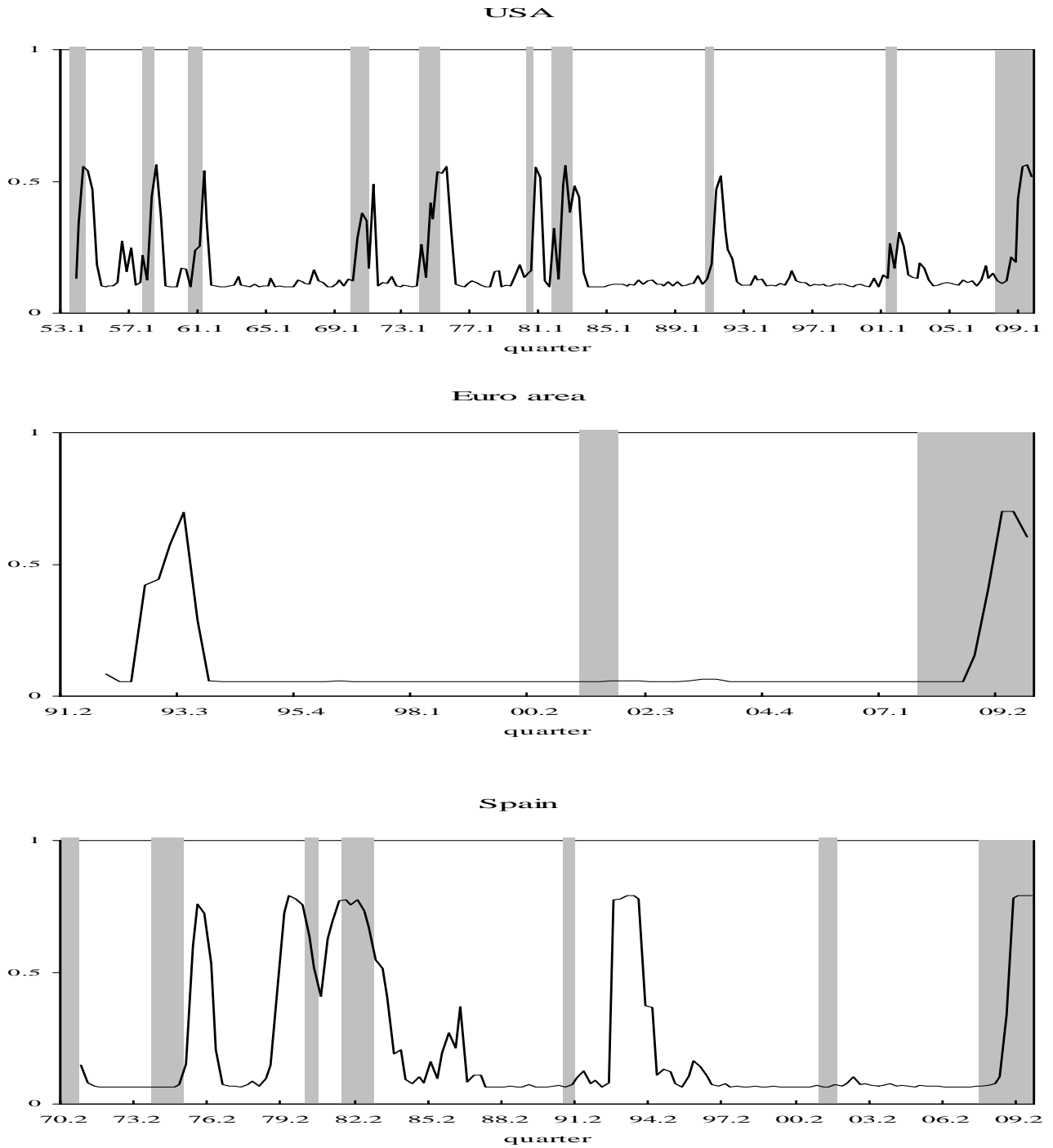
Note: Shaded areas correspond to recessions as documented by the NBER.

Figure2. Filtered probabilities of recessions from GDP growth rates



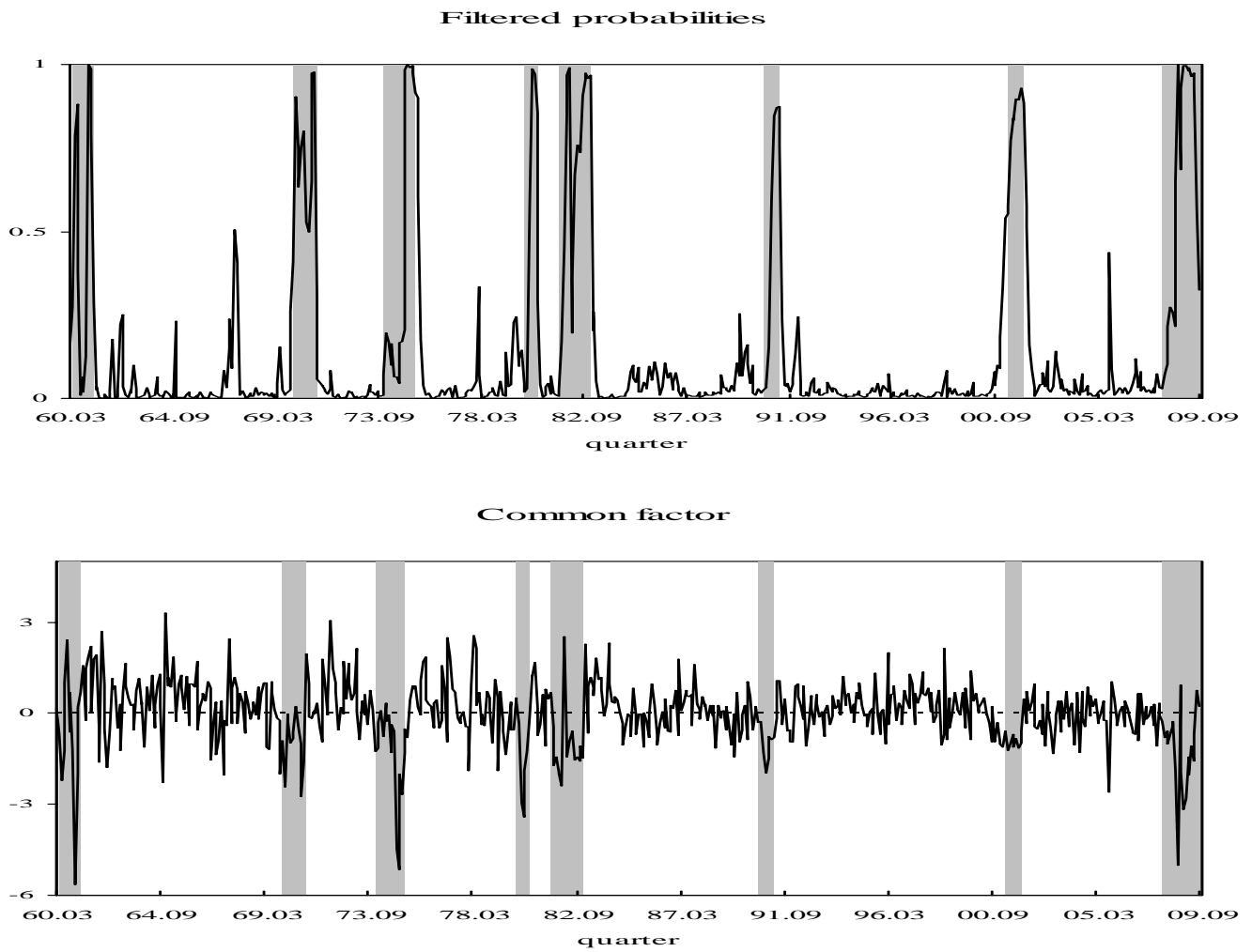
Note: Shaded areas correspond to recessions as documented by the NBER.

Figure 3: Two periods ahead filtered probabilities of recession from GDP growth rates



Note: Shaded areas correspond to recessions as documented by the NBER.

Figure 4: US filtered probabilities of recessions and common factor from Markov-switching multivariate model (5 variables)



Note: Shaded areas correspond to recessions as documented by the NBER.

Figure 5: Euro area filtered probabilities of recessions and common factor from Markov-switching multivariate model (5 variables)

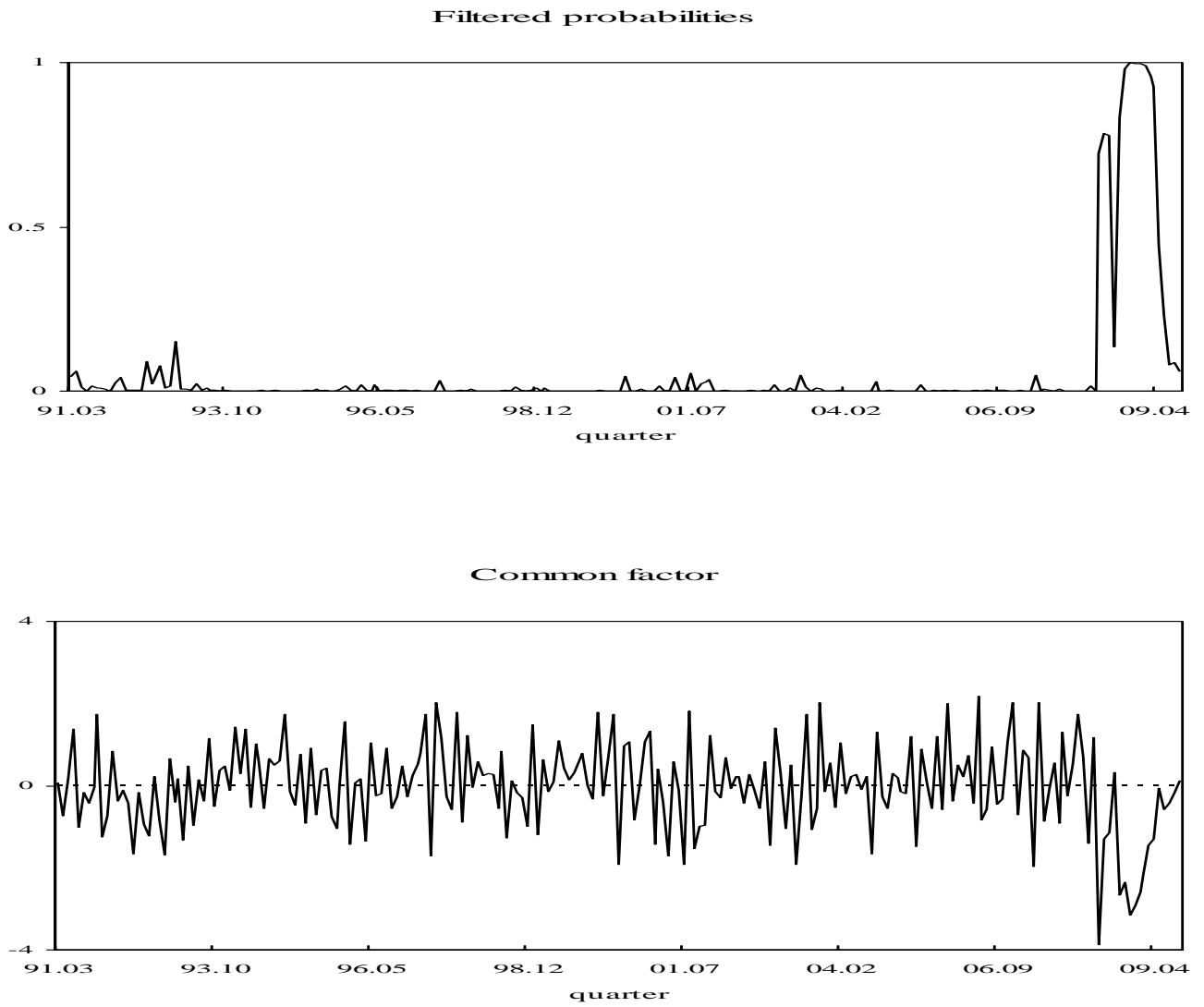
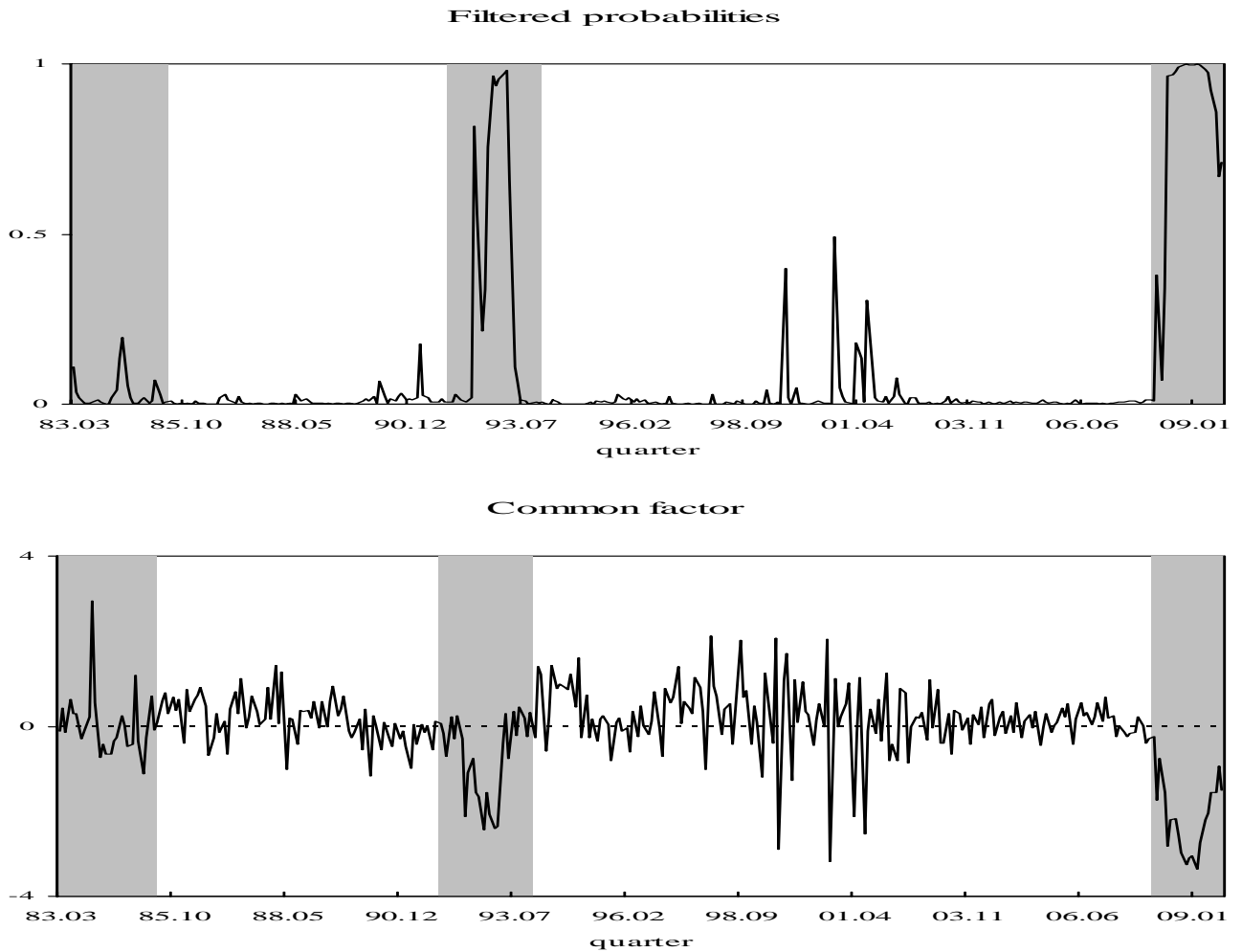


Figure 6: Spanish filtered probabilities of recessions and common factor from Markov-switching multivariate model (5 variables)



Note: Shaded areas correspond to recessions as documented by the ECRI.

Figure 7: Euro area filtered probabilities of recessions and common factor from Markov-switching Euro-STING model

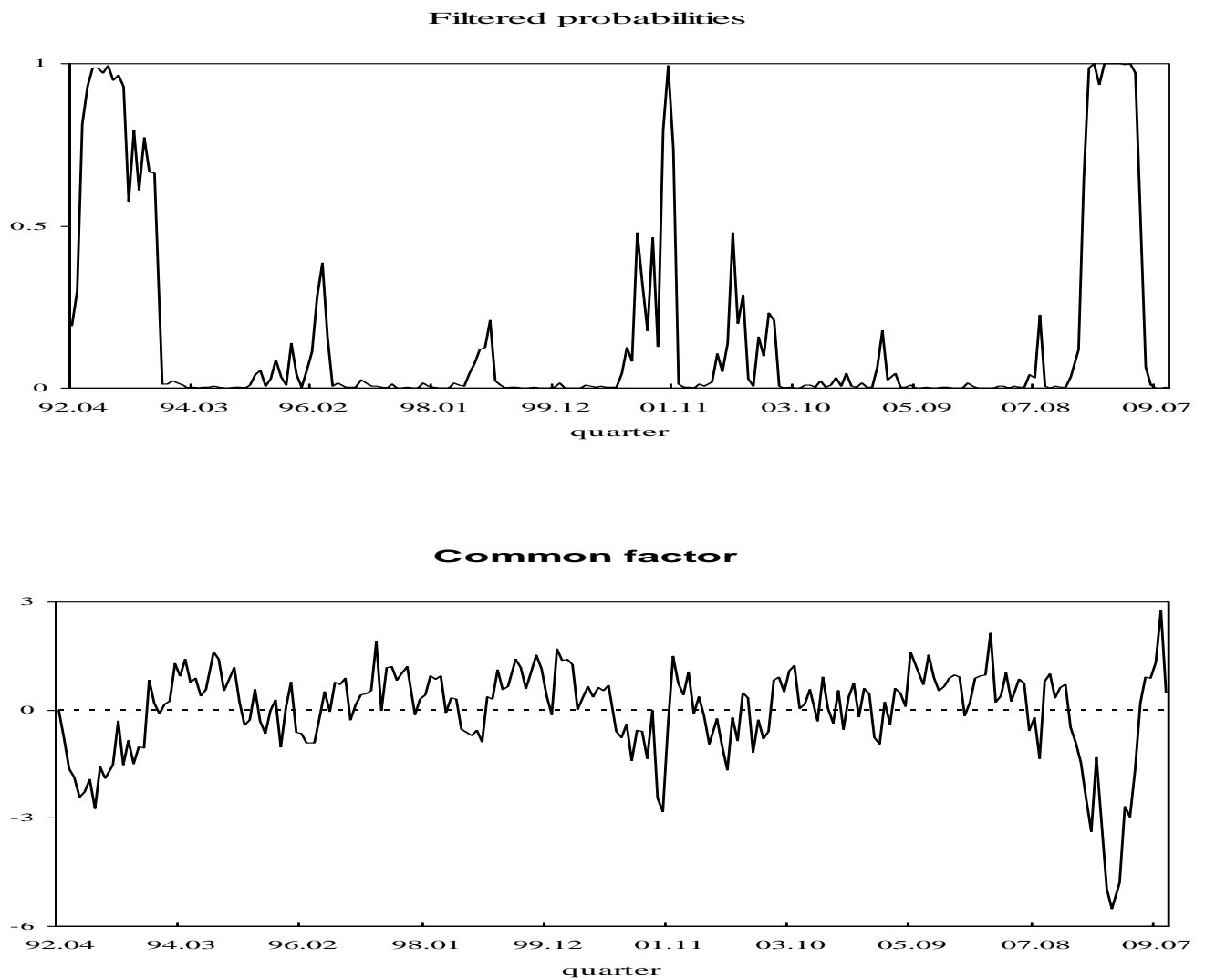
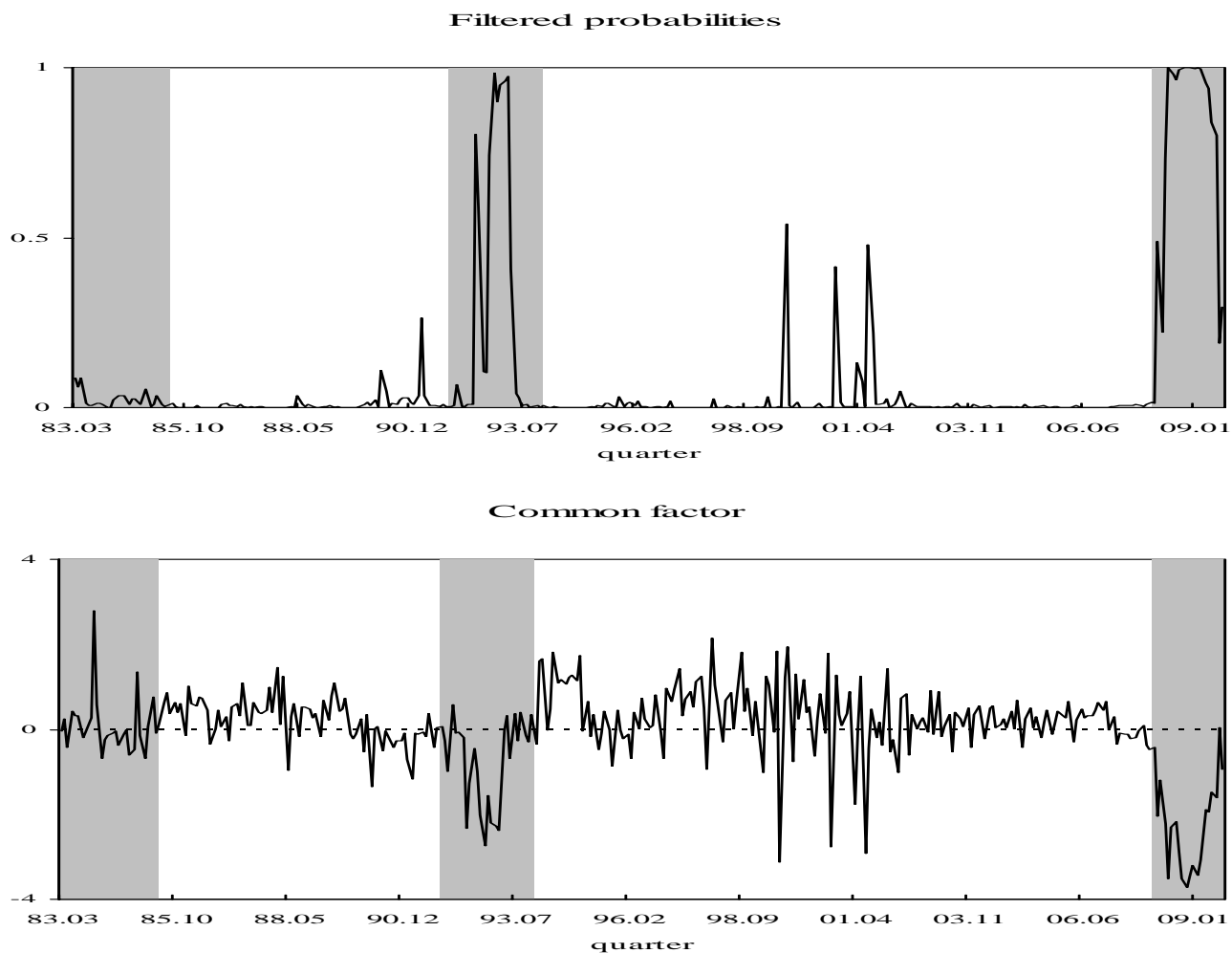


Figure 8: Spanish filtered probabilities of recessions and common factor from Markov-switching Spain-STING model



Note: Shaded areas correspond to recessions as documented by the ECRI.

Figure 9: Real-time daily Euro area filtered probabilities of recessions and common factor from Markov-switching Euro-STING model

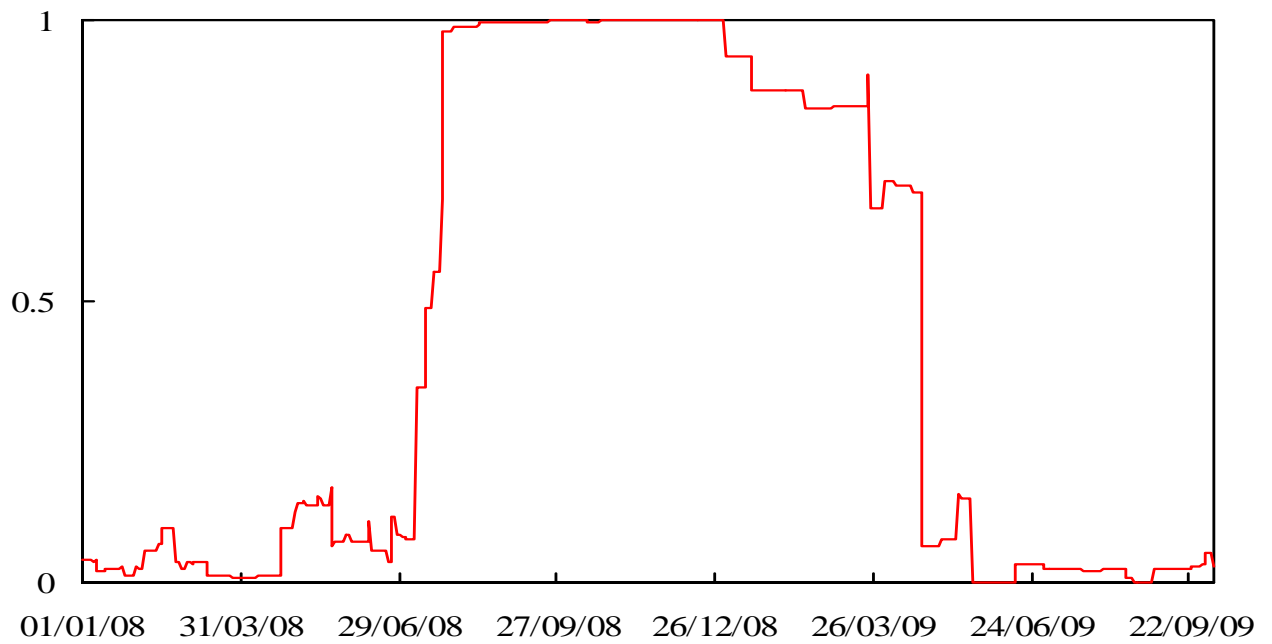


Figure 10: Real-time daily Spanish filtered probabilities of recessions and common factor from Markov-switching Spain-STING model

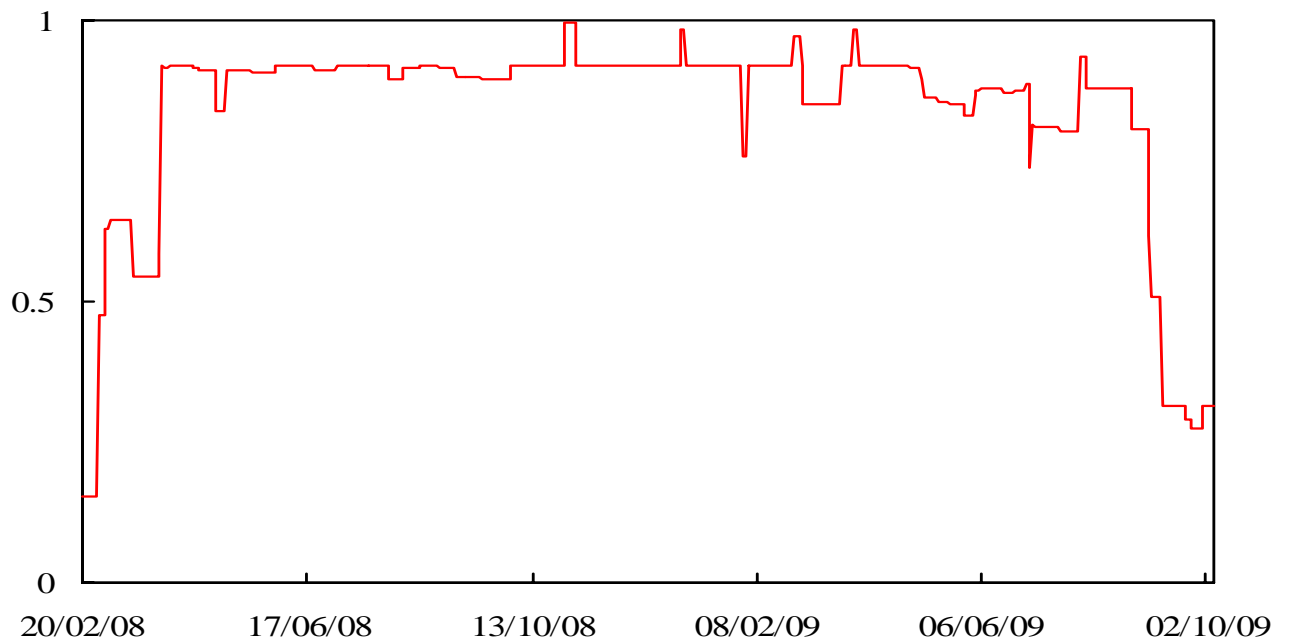
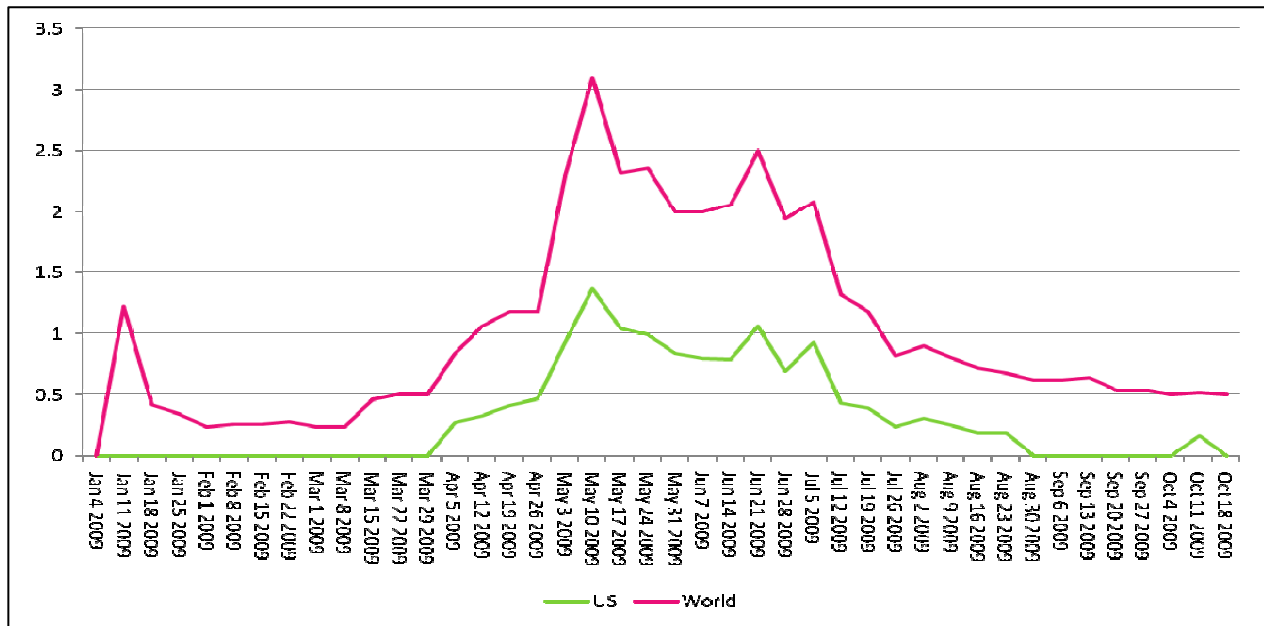


Figure 11: Number of searches of the term *green shoots* in Google



Notes. The number of searches are in millions.