FEDERAL RESERVE BANK OF SAN FRANCISCO

WORKING PAPER SERIES

A Chronology of Turning Points in Economic Activity: Spain 1850-2011

Travis J. Berge
Federal Reserve Bank of Kansas City

Òscar Jordà
Federal Reserve Bank of San Francisco
and University of California Davis

November 2011

Working Paper 2011-28
http://www.frbsf.org/publications/economics/papers/2011/wp11-28bk.pdf

November, 2011

A Chronology of Turning Points in Economic Activity: Spain, 1850-2011*

Abstract

This paper codifies in a systematic and transparent way a historical chronology of business cycle turning points for Spain reaching back to 1850 at annual frequency, and 1939 at monthly frequency. Such an exercise would be incomplete without assessing the new chronology itself and against others –this we do with modern statistical tools of signal detection theory. We also use these tools to determine which of several existing economic activity indexes provide a better signal on the underlying state of the economy. We conclude by evaluating candidate leading indicators and hence construct recession probability forecasts up to 12 months in the future.

Keywords: business cycle, correct classification frontier, area under the curve.

JEL Codes: C14, C82, E32, E65

Travis J. Berge (Federal Reserve Bank of Kansas City)

e-mail: travis.j.berge@kc.frb.org

Òscar Jordà (Federal Reserve Bank of San Francisco and University of California, Davis)

e-mail: oscar.jorda@sf.frb.org

*This paper was prepared for the Plenary Lecture of the XXXVI Simposium of the Spanish Economic Association held in Málaga, December 15-17, 2011. We thank Victor Aguirregabiria and Carlos Hervés Beloso for their kind invitation. Máximo Camacho Alonso and Gabriel Pérez Quirós were gracious enough to share their data with us. We are grateful to Jesús Fernández Villaverde and his collaborators at FEDEA for sharing their data with us. The views expressed herein are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of Kansas City, San Francisco or the Board of Governors of the Federal Reserve System.

1 Introduction

Late in the third quarter of 2007, as the fuse of the Global Financial Recession was being lit across the globe, 20.5 million Spaniards held a job.[1] Four years later, that number stood at 18.2 million –a loss of over 2,350,00 jobs a time when the working age population grew by about 800,000 individuals. Measured by the peak to trough decline in GDP –a 5% loss– one would have to reach back to the Great Depression (excluding the Spanish Civil War) to find a steeper decline in output. Moreover, employment prospects remain dim in the waning hours of 2011 for many that joined the ranks of the unemployed back in 2007. Given this environment, dating turning points in economic activity may thus seem the epitome of the academic exercise. Yet the causes, consequences and solutions to the current predicament cannot find their mooring without an accurate chronology of the Spanish business cycle.

Not surprisingly, the preoccupation with business cycles saw its origin in the study of crises. Whereas early economic historians found the roots of economic crises in "war or the fiscal embarrassments of governments,"[2] by the early twentieth century it became clear that economies experienced contractions in economic activity whose origin could not be easily determined.

As economies became less dependent on agriculture, more industrialized, more globalized and therefore more financialized, the vagaries of the weather were soon to be replaced by the vagaries of the whim. Asset price bubbles and financial crises littered the mid-nineteenth and early twentieth centuries (see Schularick and Taylor, forthcoming). The period from 1870 to 1929 saw no less than four global financial panics, each engulfing a large portion of the industrialized world –and by most accounts upwards of 50% of global GDP at the time (see Jordà, Schularick and Taylor, 2011).

Against this backdrop, 1920 saw the creation of the National Bureau of Economic Research (NBER). The NBER now views as its core mission "the aggregate economy, examining in detail

[1] Source: Encuesta de Población Activa, Ocupados. Instituto Nacional de Estadística.

[2] Wesley C. Mitchell (1927, p. 583).

the business cycle and long-term economic growth."[3] Early exponents of this mission can be found in "Simon Kuznets' pioneering work on national income accounting, Wesley Mitchell's influential study of the business cycle, and Milton Friedman's research on the demand for money and the determinants of consumer spending [...]" In fact, it is the work of Wesley C. Mitchell and Arthur F. Burns (1946) which laid the foundations for the study of the business cycle at the NBER. And since 1978 a standing Business Cycle Dating Committee (BCDC) was formed to become the arbiter of the American business cycle, a chronology that now reaches back to 1854. Slowly, other countries have been creating similar committees, such as the Euro Area Business Cycle Dating Committee of the Center for Economic Policy Research, founded in 2002. But to our knowledge, no such independent arrangement has been created in Spain.

A chronology of the Spanish business cycle is not only a necessity for the modern study of the origins of macroeconomic fluctuations and the design of optimal policy responses, it is a necessity that as of September 7, 2011 would appear to be a matter of constitutional law. The constitutional reform of article 135 passed by parliament that day states that: "The limits of the structural deficit and public debt volume may be exceeded only in case of natural disasters, *economic recession* or extraordinary emergency situations that are either beyond the control of the State or significantly impair the financial situation or the economic or social sustainability of the State, as *appreciated* by an absolute majority of the members of the Congress of Deputies" (emphasis added). It would appear that the whimsy of the business cycle is at the purview of the legislature rather than the economic brain trust. If nothing else, this observation serves to cement the importance that an independent committee, whose job is to determine turning points in economic activity, can play in the economic and political life of a country.

But what is a recession? The BCDC offers a clear yet less than operational definition:[4]

> A recession is a significant decline in economic activity spread across the economy,

[3] From the NBER's website on the History of the NBER available at: http://www.nber.org/info.html.

[4] www.nber.org/cycles/

> lasting more than a few months, normally visible in production, employment, real income, and other indicators.—Determination of the December 2007 Peak in Economic Activity, December 2008. Business Cycle Dating Committee of the National Bureau of Economic Research.

And most institutions in the business of keeping a chronology of economic cyclical activity use a similarly intuitive yet entirely mathematically imprecise definition of what a recession is. How then would one determine whether or not a business cycle dating committee (or a legislature) is appropriately sorting the historical record into periods of expansion and periods of recession? After all, the true state of the economy (expansion or recession) is inherently unobservable –an infinite sample of data can only improve the precision of the estimated probabilities associated to each state, but it does not reveal the states themselves.

Our quest to formalize a chronology of the Spanish business cycle begins with a brief description of the statistical methods that have been used in the literature to achieve a classification of turning points. That journey begins with the early methods that Gerhard Bry and Charlotte Boschan introduced in (1971) at the NBER. The Bry and Boschan (1971) algorithm comes closest to translating the NBER's definition into practice: take the data (there is no need to detrend), remove seasonals, smooth lightly, constrain cycles to have a minimum duration of six months or two quarters and to alternate, make sure that completed cycles (recession+expansion) last at least 15 months, and then spot the local minima and maxima in the series. A local minimum is a *trough* and the following local maximum a *peak* so that the period between trough and peak is an expansion, and from peak to trough a recession. The original Bry and Boschan (1971) algorithm saw its most recent revival in work of Harding and Pagan (2002a, b), which for quarterly data they dub the BBQ algorithm, and the string of papers by Kose, Prasad and Terrones (2003), Kose, Otrok and Whiteman (2008) and Kose, Otrok and Prasad (2008), to cite a few. Arbitrary as the Bry and Boschan (1971) algorithm may seem, it is simple to implement, reproducible, and perhaps more critically, it does not require that the data be detrended.

A more structural view of how fluctuations around trend-growth are determined is to suppose that the data are generated by a mixture process. In the econometrics literature, characterizing the stochastic process of economic fluctuations as a mixture finds its most celebrated reference in the pioneering work of Hamilton (1989). The idea is to conceive of the data as being generated by two distributions (one for each state, expansion or recession) and to characterize the transition between states as a hidden-Markov process. In the statistics literature, the problem of identifying the underlying state of the economy closely resembles pattern recognition problems in computational learning, or more briefly *decoding*.

Decoding is most often referred to in information theory as an algorithm for recovering a sequence of code words or messages from a given sequence of noisy output signals (Geman and Kochanek, 2001). In fact, almost every cell-phone on earth uses a version of the celebrated Viterbi algorithm (Viterbi, 1967), itself based on filtering a hidden-Markov process. More recently, an application of these principles with non-parametric computational techniques was introduced by Fushing, Hwang, Lee, Lang and Horng (2006) in what they call the *hierarchical factor segmentation* (HFS) algorithm. An application of HFS to economic data is found in Fushing, Chen, Berge and Jordà (2010). The basic principle of the HFS algorithm is to use the recurrence time distribution of certain events (say, record each time output grows below a given threshold) to come up with an optimal non-parametric mixture using the maximum entropy principle of Jaynes (1957a, b). Interestingly, the idea of using recurrence times dates back to Poincaré (1890).

Each method can be applied to different series out of which one obtains a multiplicity of chronologies, one for each variable. Or one could combine the data first with a factor model, and then use the factor to date the business cycle. The combine-then-date approach appears to be the most commonly used at present (a good example is Stock and Watson, 2010), probably reflecting the popularity that factor models currently enjoy in other areas of economics. Moreover, a single indicator of economic activity has the advantage of being a succinct tool of communication. From that perspective, our investigation will take us to consider a variety of such indicators that have been proposed to characterize business conditions in Spain. Among these, we will investigate the OECD's composite leading indicator (CLI) index,[5] the index of economic activity constructed by the Spanish think tank FEDEA,[6] and two recent more sophisticated indexes, the MICA-BBVA index[7] of Camacho and Doménech (2011), and Spain-STING[8] by Camacho and Pérez Quirós (2011).

Yet as we shall see, variables do not always fluctuate synchronously –a prime example can be seen by comparing employment and output across the business cycle– an observation that would seem to favor the date-then-combine approach if interest is tilted toward constructing a single series of turning points. Moreover, the variables in our data set are observed over different spans and at different frequencies, features that make the factor approach less attractive. Instead, a simple method of date-combination based on the network connectivity properties of each chronology (see, e.g., Watts and Strogatz, 1998), turns out to provide insight into the determinants of economic fluctuations and a straightforward method to generate a single chronology of turning points.

It is not enough to come up with a chronology of turning points, one must also formally assess the quality of any given chronology. A scientific defense of the quality of such a chronology requires formal statistical assessment and to this end we reach back to 1884 and Charles S. Peirce's "Numerical Measure of the Success of Predictions," the direct precursor to the Youden index (Youden, 1950) for rating medical diagnostic tests, and the receiver operating characteristics (ROC) curve by Peterson and Birdsall (1953) in the field of radar detection theory. Today, the ROC curve is a standard statistical tool in the assessment of medical diagnostic procedures (going back to Lusted, 1960), but it is also used routinely in atmospheric sciences (see Mason, 1982) and

[5] The OECD's CLI index can be downloaded directly from the OECD's website: www.oecd.org/std/cli.

[6] FEDEA stands for Fundación de Estudios de Economía Aplicada, and their website is www.fedea.es

[7] MICA-BBVA stands for factor Model of economic and financial Indicators which is used to monitor Current development of Economic Activity by Banco Bilbao Vizcaya Argentaria (BBVA). We thank Máximo Camacho for making these data readily available to us.

[8] STING stands for Short-Term Indicator of Growth. We thank Máximo Camacho and Gabriel Pérez Quirós for making the data readily available to us.

machine learning (Spackman, 1989). In economics, early uses appear for the problem of credit scoring, but more recently for the evaluation of zero-cost investments, such as the carry trade (see Jordà and Taylor, 2009 and Berge, Jordà and Taylor, 2011). Jordà and Taylor (2011) provide perhaps the most detailed overview of this literature and emphasize the *correct classification frontier*, a relative of the ROC curve, as the more appropriate tool in economics. Applications of these techniques to the classification of economic data into expansions and recessions in the U.S. is done in Berge and Jordà (2011).

Our pursuits end by gazing into the future: What can we say about the problem of predicting future turning points? In another departure from traditional econometric practice, the problem of choosing good predictors for classification purposes does not require that the predictors be accurate in the usual root mean squared error sense. Moreover, we will argue that, unlike conventional time series modelling, it is best to tailor the set of predictors to the forecast horizon under consideration. In our experience, we have found that variables can be good classifiers in the short-run but poor classifiers in the long-run and vice versa. If, as is common practice, one fits a model based on short-run prediction and then iterates forward to longer horizons, the model will tend to put too much weight on the short-run classifiers and generate worse predictions than if a different model is chosen for each horizon –a practice commonly referred to as *direct forecasting*. Seen through this lens, the outlook for the Spanish economy over the next few months remains grim.

2 Dating Turning Points

The BCDC's September 10, 2010 press release pronounced the U.S. trough of economic activity to have occurred June 2009.[9] In that release, the committee made available the data and figures used to make that determination, thus offering a more intimate glimpse at how decisions on turning points are made. The Bry and Boschan (1971) algorithm is perhaps the most direct expression of this process. At its heart, this algorithm attempts to identify peaks and troughs in the level of a

[9] See http://www.nber.org/cycles/sept2010.html

business cycle indicator. We explain the details of how this is done below taking note of the data that the BCDC analyzes in order to replicate a similar analysis with Spanish data. The results of this analysis form the basis of our proposed chronology of the Spanish business cycle.

If instead one focuses on the rates of growth in economic activity, so that the data can be reasonably thought of as being stationary and therefore trend-free, an alternative way to conceive of cyclical phenomena is to speculate that the data are generated by a mixture process whose alternating pattern is driven by a hidden-Markov process. Thinking of the data generation process (DGP) in this manner calls for a filtering method. Hamilton's (1989) filter is the most commonly used in economics, which we briefly describe below. If one prefers to be less specific about the stochastic processes describing the evolution of the data in each regime, there exist a number of non-parametric filtering algorithms within the statistics literature. One that has been applied to the problem of classifying business cycles is the hierarchical factor segmentation (HFS) algorithm, which is also described below. In our application to Spanish data, these two hidden Markov models are estimated on real GDP growth data to serve as a counterpoint to the cyclical turning points we identify with the Bry and Boschan (1971) approach.

However, the application of these methods to Spanish data leave a jumble of dates and discrepancies across series to contend with. This we do using network connectivity measures. The result is a unique candidate chronology of Spanish recessions that at least forms the basis for a more informed conversation about the Spanish business cycle. In the next section we will examine different tools that can be used to evaluate this chronology against other available chronologies (such as those produced by the OECD and the Economic Cycle Research Institute or ECRI). Perhaps not surprisingly, we find strong empirical evidence in support of our chronology.

2.1 Bry and Boschan (1971)

Understanding the basics of the Bry and Boschan (1971) algorithm is best achieved using a yearly frequency data example. And to that end, figure 1 displays the time series of Spanish real GDP per capita from 1850 to 2008 (assembled by Prados de la Escosura, 2003), along with recession shadings whose construction we will now discuss. Let y_t denote the logarithm of real GDP per capita in 2000 euros, let P_t be a binary indicator that takes the value of 1 if date t is a peak of economic activity, 0 otherwise, and let T_t be a binary indicator that takes the value of 1 if date t is a trough of economic activity, 0 otherwise. Then peak and trough dates can be calculated as follows:

$$\begin{aligned} P_t &= 1 \text{ if } \Delta y_t > 0 \text{ and } \Delta y_{t+1} < 0 \\ T_t &= 1 \text{ if } \Delta y_t < 0 \text{ and } \Delta y_{t+1} > 0. \end{aligned} \tag{1}$$

In other words, the algorithm looks for local maxima and minima in the raw data. We will use the acronym BBY to refer to the application of this algorithm to yearly frequency data, which is done in Figure 1. Recessions arrive more frequently in the early part of the sample, likely reflecting among other things, the preponderance of an agricultural sector that is much more sensitive to fluctuations in weather patterns. A simple calculation of the average growth rate of GDP during the period prior to the start of the civil war puts average annual per capita GDP growth at about 1.2%. The destruction of economic activity during the Civil War is massive, with a loss of per capita output near 35%, and a recovery to trend growth that would take almost the entirety of Franco's dictatorship. Since then, the rate of per capita growth has stabilized around a 2% rate, which is largely comparable to other industrialized economies. Table 1 provides the list of peak and trough dates that we calculate with expression (1).

[Insert Figure 1 here]

[Insert Table 1 here]

To motivate the filtering methods that we discuss below, it is useful to calculate the empirical mixture distribution that results for the annual growth rate in real GDP per capita from the Bry and Boschan (1971) procedure, and this is displayed in figure 2. The kernel density estimates for the recession and expansion distributions overlap roughly over the interval of $\pm 5\%$. This overlap serves to illustrate that the dating of business cycles is not a simple mechanical exercise of recording when output is below or above some threshold (say zero percent). Rather, cyclical activity refers to recurrent patterns of depressed and burgeoning periods of economic activity within which one can countenance some variation.

[**Insert Figure 2 here**]

When the data is quarterly or monthly, several additional adjustments are done to the basic rules in expression (1).[10] First, the data are seasonally adjusted. Next, it is common to smooth the data with a moving average filter to remove small sources of idiosyncratic variation that matter not for spotting cyclical phenomena (although with quarterly data, the smoothing step is omitted due to the coarse frequency of the data). In addition, two important ad-hoc rules are added to an expression like (1): a restriction on the minimum length of a recession –6 months or two quarters, depending on the frequency of the data–; and a restriction on the minimum length of a complete recession+expansion cycle –15 months or four quarters, again depending on the frequency of the data. These rules reflect the spirit of the definition of recession presented in the introduction and the notions on cyclical activity described in Burns and Mitchell (1946).

The application of the Bry and Boschan (1971) algorithm to quarterly data (with the gastronomical acronym BBQ as Harding and Pagan, 2002a fittingly recognized) is presented in figure 3 and table 2. Figure 3 contains two panels, the top panel displays the raw real GDP data available from the Spanish National Accounts, which comes organized into three overlapping windows depending on the base year used to calculate prices. The samples are 1970Q1 to 1998Q4, 1980Q1-2004Q4, and 1995Q1-2011Q2. The first two samples share two recessions in common and the timing is rather similar, usually within 2 quarters of each other. The second panel displays employment data (total employed from the household survey), which starts a little later, 1976Q3

[10] The specific details are best explained in King and Plosser (1994) and Harding and Pagan (2002a).

to 2011Q2. At the start of the sample and up until the trough of 1985Q2, employment is steadily declining so it is difficult to date the beginning of that recession with employment data alone. However, the dates of the last two recessions overlap reasonably well with those identified with GDP, although employment appears to decline earlier than GDP and recover later. This is presented more clearly in table 2. Moreover, the dates presented in table 2 relate well to the dates we identified using the historical yearly data and presented in table 1.

[**Insert Figure 3 here**]

[**Insert Table 2 here**]

Finally, we show the results of the Bry and Boschan (1971) algorithm when used on monthly data (dubbed here BBM). In an effort to replicate the same series used by the BCDC[11] for Spain, we examine linearly interpolated quarterly data on real GDP and employment (used earlier for the BBQ analysis) and we add the number of registered unemployed, the industrial production index and an index of wage income. The sources and transformations for all the data are provided in more detail in the appendix.

In all, we have five series from which to construct a single chronology of peaks and troughs. But before we show how this can be done, table 3 summarizes the BBM chronology. There are a number of adjustments that deserve comment. To this end, figure 4, which displays the registered unemployed series, serves to highlight where these adjustments come from. The most obvious pattern in the figure is the big run-up in the number of registered unemployed at the end of 1975 and all the way to about 1985. This is a striking change and likely reflects a number of institutional changes: Franco dies in November 1975 and the referendum on the Spanish Constitution takes place in 1978 –two of the early salvos in the creation of the modern democratic architecture of the Spanish state– along with the two oil crises of 1973-4 and 1979. Even separating that subsample

[11] The BCDC looks at lots of data but in their website, special emphasis is made on the following variables: linearly interpolated from quarterly real Gross Domestic Product (GDP); linearly interpolated from quarterly real Gross Domestic Income (GDI); Industrial Production Index (IPI); real Personal Income less transfers (PI); payroll employment (PE); household employment (HE); real Manufacturing and Trade Sales (MTS).

from the rest, it is easy to see that the cyclical behavior of the data after 1985 is quite different than it was before 1975. Clearly, it would be very difficult to come up with a model that could describe the entire sample and here is where the Bry and Boschan (1971) algorithm can be quite useful. In addition, notice that there is clearly an adjustment in the series in November 1995 that has nothing to do with the business cycle. We reconciled the dates of peaks and troughs accordingly to avoid detecting a spurious recession. Before we discuss how all this information can be reconciled to generate a unique chronology, we discuss two alternative methods that we use as a cross-check of the results reported here.

[**Insert Figure 4 here**]

[**Insert Table 3 here**]

2.2 Hamilton's Markov Switching Model

With nearly 5,000 citations in scholar.google.com, Hamilton's Markov switching model is one of the most commonly applied methods for identifying business cycles in economic data. A complete description of the model introduced in Hamilton (1989) is beyond the scope of this paper. However, the basic ideas can be expressed succinctly. In its simplest specification, suppose y_t refers to the annualized growth rate of quarterly real GDP and conjecture that the stochastic process describing the data is given by:

$$(y_t - \mu_s) = \rho(y_{t-1} - \mu_s) + \varepsilon_t \qquad \varepsilon_t \sim N(0, \sigma) \tag{2}$$

where $\mu_s \in \{\mu_1, \mu_2\}$, that is, the unconditional mean is assumed to attain one of two values depending on the state $s \in \{1, 2\}$. When $\rho = 0$, equation (2) is the expression of a Gaussian mixture with common variance but different means. There are many dimensions in which the model can be made more complex (such as allowing the dynamics and the variance to be state dependent, considering more than two states, and many other variations that are discussed in the literature).[12]

[12] There are many sources of code available to estimate Markov switching models, including code available from Hamilton's own website at: http://weber.ucsd.edu/~jhamilto/. We used MATLAB code available from Perlin, M.

The transition between states is assumed to be described by a first-order, two-state Markov process with transition probabilities:

$$p_{ij} = P(s_t = i|s_{t-1} = j) = P(s_t = i|s_{t-1} = j; s_{t-2} = k, ...)$$

where $i, j, k = 1, 2$ so that information prior to time $t-2$ is not needed. The true state of the process, s_t, is not directly observable but can be inferred from the sample. One way to estimate the model and make inferences about the unobserved state is to cast the model in state-space form (see e.g. Kim and Nelson, 1999). The model can then be estimated by maximum likelihood and the transition probabilities can be calculated as a by-product of the estimation. Moreover, the specification of the filter permits a convenient way to obtain accurate estimates of these transition probabilities through a backwards smoothing step. The resulting probability estimates are the quantities that we will report in our examples.

As an illustration, consider the annualized growth rate of real GDP at a quarterly frequency provided in the Spanish national accounts since 1970Q1. Figure 5 compares the smoothed transition probabilities for the recession state against the recession regions identified with the BBM algorithm on the linearly interpolated data for GDP only. Table 4 collects the BBQ dates for GDP, those from the Hamilton (1989) filter, and those from the HFS algorithm, to be discussed briefly. Figure 5 and table 4 show that the Hamilton (1989) filter selects fewer recessions: three in the 1970Q1 to 20011Q2 period against the five selected by BBQ, and the eight selected by HFS. However, the dates of those three recessions largely coincide across methods. If anything, the evidence from the five monthly indicators discussed in the previous section would suggest that the Hamilton (1989) dates are perhaps too conservative –see table 3. Nevertheless, it is reassuring that for those recessions detected, the dates largely line up with those from other methods.

[**Insert Figure 5 here**]

[**Insert Table 4 here**]

(2010) MS Regress available at SSRN: http//ssrn.com/abstract=1714016.

2.3 The Hierarchical Factor Segmentation Algorithm: HFS

Introduced by Fushing, Hwang, Lee, Lang and Horng (2006), the hierarchical factor segmentation (HFS) algorithm is a non-parametric, pattern-recognition procedure that exploits the recurrence distribution of separating events, an idea that traces its origins perhaps as far back as Poincaré (1890). The reader is referred to the original source for a more in-depth description. HFS belongs to the larger class of hidden Markov models and in that sense, it can be considered as the non-parametric cousin to Hamilton's (1989) model. Here we provide a succinct summary.

HFS is a procedure whose underlying premise is that the data has been generated by a mixture model –much like the specification of the Markov switching model presented above. However, rather than specifying the complete stochastic process of the data, one proceeds in a series of steps. First, determine a separating event –that is, a feature of the data more likely to belong to one distribution than the other–, which is used to generate a preliminary partition of the data. In our application, this separating event is based on observations in the bottom 30^{th} percentile of the empirical distribution of quarterly real GDP growth. This step may appear ad-hoc, but the success of HFS does not depend on a precise determination of this separating event (see Fushing, Chen, Berge and Jordà, 2010).

Next, the data is further partitioned into clusters, that is, periods where the observed frequency of separating events is high and periods when it is low. Entropy arguments (Jaynes, 1957a, b) suggest that the duration between events can be best characterized by a Geometric mixture (see Fushing, Chen and Hwang, 2010a, b) and the final partition into expansions and recessions is the result of maximizing the empirical likelihood of this mixture.

As a way to illustrate the procedure in practice, we used the same real GDP growth data that we used to estimate the Hamilton (1989) model described in the previous section. The dates of peaks and troughs are described in table 4, which we discussed previously. Relative to BBQ and Hamilton (1989), HFS tends to identify more recessions: eight versus five and three respectively.

However, as the monthly analysis suggests, some of these additional recessions appear to find a counterpart in the monthly variables that we analyzed in table 3.

2.4 Summary, Network Connectivity and a Chronology

The previous sections have generated a multiplicity of business cycle chronologies, each derived from a particular method and using different underlying data. Along the way we have learned several lessons worth summarizing. A chronology of peaks and troughs facilitates the cataloguing of basic empirical facts and for this reason, we think the Bry and Boschan (1971) algorithm codifies that which is more likely to be of interest to researchers. Moreover the Bry and Boschan (1971) method is robust: it does not require detrending the data, the dates will not change as a result of expanding the sample over time, and it is easy to communicate. On the downside, the algorithm feels ad-hoc and it requires a few observations past the turning point to make a sound determination on its precise date (undoubtedly, one of the reasons the NBER takes anywhere between 12 to 18 months, thus eliciting the jeers of those that would prefer a more timely release schedule). On the other hand, methods based on the hidden Markov approach, such as Hamilton's (1989) and HFS, have more solid statistical justification and can generate more timely pronouncements (subject to inevitable revisions in the data), but have a less intuitive feel. When we calculate the employment loss in a recession, we think of the employment level at the peak minus the employment level at the trough and those are easy concepts to grasp. It is less clear why that calculation should be done by comparing those periods when the transition probability is, say, above 0.5 and then below it.

We conclude this section by discussing how we reconcile the patchwork of dates that we have uncovered using different economic indicators, to generate a single chronology. At the NBER such a procedure is done by committee. Here we propose procedures based on the theory of networks (see Watts and Strogatz, 1998) and in particular, two popular measures of network connectivity: the incidence rate and the wiring ratio. Suppose that we generate a binary indicator of recession out of each of the five indicators that we considered above using the periods from a peak to the subsequent trough identified in table 3. The incidence rate computes the ratio of the number of indicators flashing recession relative to the total number of indicators at every point in the sample:

$$\rho_t = \frac{r_t}{n}; \qquad r_t = \sum_{i=1}^{n} r_{it} \tag{3}$$

where the binary recession indicator i is $r_{it} \in \{0, 1\}$ for $t = 1, ..., T$; $i = 1, ..., n$ and n is the total number of indicators.

The incidence rate is very intuitive, but attributes the same marginal weight to an additional indicator flashing recession when going from 0 to 1 indicators, than when going from 4 to 5 indicators. If instead one wants the marginal value of an additional signal to be low when few indicators flash recession and high otherwise, the wiring ratio offers an attractive alternative. The wiring ratio is based on the number of pair-wise active connections relative to the total number of possible pair-wise connections and hence can be calculated as:

$$\omega_t = \frac{r_t(r_t - 1)/2}{n(n-1)/2} = \frac{r_t(r_t - 1)}{n(n-1)}. \tag{4}$$

The samples available for each of the five indicators that we consider vary greatly. Prior to 1970 we can only rely on data for the number of registered unemployed. As time goes by, we are able to incorporate information from the other indicators and by 1985 we have information on all of them. This is easily accommodated by our two measures since all that is required is to adjust n over time to reflect the number of indicators available –another score in the simplicity scale.

Both of these network connectivity measures are displayed in figure 6 along with an interpolated measure of real GDP to provide some context. Moreover, the figure displays recessions calculated as those periods when the incidence rate is above 50%. The resulting dates are also listed in table 5. The first column simply summarizes the yearly chronology of peaks and troughs using the historical data of Prados de la Escosura (2003) described earlier, where as the second column contains monthly dates of peaks and troughs based on increasingly more data and the 50% incidence rule.

[**Insert Figure 6 here**]

[Insert Table 5 here]

We present this chronology because its construction is transparent and replicable, but not because we think it is the last word on the Spanish business cycle. There are certainly other variables one may have considered and at all times one must be aware of what economic history tells us to be able to refine the dates that we present. But we think this chronology is a reasonable starting point that we hope will be of service to other researchers.

3 Tools of Evaluation

If the true state of the economy (expansion versus recession) is not directly observable, by what metric would one then judge one chronology as being superior to another? This would seem to be an impossible question to answer but statistical methods dating back to the nineteenth century provide ways to get a handle on this question. Before we get there, we find it useful to begin our discussion taking a chronology of business cycles as given, and then asking how good is a given variable in sorting history into expansions and recessions. Such a problem, it turns out, is not all that different from evaluating a medical diagnostic procedure, determining whether an e-mail is spam or not, or judging a tornado warning system, to mention a few applications. In all cases the object we wish to predict is a binary outcome and how we judge the quality of a variable as a classifier depends to a great extent, on the costs and benefits associated with each possible classifier, outcome pair.

Much of this discussion borrows from Berge and Jordà (2011) and Jordà and Taylor (2011) and finds its origin in the work of Charles S. Peirce (1884) and the theory of signal detection in radars by Peterson and Birdsall (1953). Specifically, let y_t denote the classifier, an object that can be any number of things: an indicator variable (say an index of economic activity), a real-time probability prediction (say from a binary probability model), a single index (say from a simple regression, or a neural network, or some other model), a factor (say from a principal component decomposition), and so on. The distinction is unnecessary for the methods we describe. y_t together with a threshold c define a binary prediction *recession* with $s_t = 1$ whenever $y_t \leq c$, and *expansion* with $s_t = 0$ whenever $y_t > c$. Obviously the sign convention is for convenience. If we used the unemployment rate as our classifier y_t, we could just as easily reformulate the problem in terms of the negative of the unemployment rate.

Associated with these variables, there are four possible classifier, outcome $\{y_t, s_t\}$ probability pairs: the *true positive rate* $TP(c) = P[y_t \leq c|s_t = 1]$, the *false positive rate* $FP(c) = P[y_t \leq c|y_t = 0]$, the *true negative rate* $TN(c) = P[y_t > c|s_t = 0]$ and the *false negative rate* $FN(c) = P[y_t > c|s_t = 1]$. It is straight-forward to see that $TP(c) + FN(c) = TN(c) + FP(c) = 1$ with $c \in (-\infty, \infty)$. Clearly, as $c \to \infty$ $TP(c) \to 1$ but $TN(c) \to 0$ and vice versa when $c \to -\infty$. To an economist, this trade-off is familiar since it has the same ring as the production possibilities frontier: for a given technology and a fixed amount of input, dedicating all the input to the production of one good restricts production of the other good to be zero and vice versa. And the better the technology the more output of either good or a combination can be produced. For this reason Jordà and Taylor (2011) label the curve representing all the pairs $\{TP(c), TN(c)\}$ for $c \in (-\infty, \infty)$ as the *correct classification frontier* (CCF). In biostatistics, the curve representing all the pairs $\{TP(c),\ FP(c)\}$ is called the *receiver operating characteristics curve* or ROC curve, but this is just the mirror image of the CCF and it shares the same statistical properties.

A good classifier is one that has high values of $TP(c)$ and $TN(c)$ regardless of the choice of c and in the ideal case it turns out that $TP(c) = TN(c) = 1$ for any c. In that case, it is easy to see that the CCF is just the unit square in $TP(c) \times TN(c)$ space, as shown in figure 7. At the other extreme, an uninformative classifier is one in which $TP(c) = 1 - TN(c)$ for any c and the CCF is the diagonal bisecting the unit-square in $TP(c) \times TN(c)$ space. Using the colorful language of the pioneering statistician Charles Sanders Peirce (1884), the classifiers corresponding to these two extreme cases would be referred to as the "infallible witness" and the "utterly ignorant person" (Baker and Kramer 2007). In practice, the CCF is a curve that sits between these two extremes as depicted in figure 7.

[Insert Figure 7 here]

Depending on the trade-offs associated with $TP(c)$ and $TN(c)$ (and implicitly $FP(c)$ and $FN(c)$), Peirce (1884) tells us that the "utility of the method" can be maximized by choosing c such that:

$$U(c) = [U_{pP}TP(c)\pi + U_{nN}TN(c)(1-\pi)] + [U_{pN}(1-TN(c))(1-\pi) + U_{nP}(1-TP(c))\pi] \quad (5)$$

where π is the unconditional probability $P(s_t = 1)$. A good rule of thumb is to assume that $U_{pP} = U_{nN} = 1$ and $U_{nP} = U_{pN} = -1$ so that we are equally happy correctly identifying periods of expansion and recession, and equally unhappy when we make a mistake. Yet to a policymaker these trade-offs are unlikely to be symmetric, specially if the costs of intervening are low relative to the costs of misdiagnosing a recession as an expansion. Therefore, figure 7 plots a generic utility function that makes clear, just as in the production possibilities frontier textbook model, that the optimal choice of c is achieved at the point where the CCF and the utility function are tangent (assuming no corner solutions such as when we have a perfect classifier). This is sometimes called the *optimal operating point*.

In the canonical case with equal utility weights for hits and misses and $\pi = 0.5$, the optimal operating point maximizes the distance between the average correct classification rates and 0.5, the average correct classification rate of an uninformative classifier –the utterly ignorant person. This is just another way of expressing the well-known Kolmogorov-Smirnov (KS) statistic (see Kolmogorov, 1933 and Smirnov, 1939):

$$KS = \max_c 2\left|\frac{TP(c)+TN(c)}{2} - \frac{1}{2}\right|.$$

Intuitively, the KS statistic measures the distance between the empirical distribution of y_t when $s_t = 1$, and the empirical distribution of y_t when $s_t = 0$. An example of this situation is displayed in figure 2 presented earlier, which shows the kernel estimates for the empirical distribution of per capita real GDP in expansions and in recessions.

The beginnings of an evaluation strategy begin to materialize. In the situation where the chronology of business cycles is a given and y_t is, say, a linear combination of leading indicators, the more clear the separation between the empirical distribution of y_t when $s_t = 1$ from when $s_t = 0$, the easier it will be to correctly sort the data into expansion and recession when making predictions. But this argument can be inverted to judge the chronology itself. If a candidate chronology, by which we mean the sequence $\{s_t\}_{t=1}^{T}$, is "good" then it should be the case that cyclical candidate variables y_t have empirical distributions in each state that are easily differentiated. Consider again figure 2. If the chronology of recessions and expansions carried no useful information, then the two conditional empirical distributions would lie on top of one another, so that any given observation of real GDP would be as likely to have been drawn in expansion as in recession.

There are several reasons why the KS statistic is somewhat unappealing, among them the fact that we do not know what the utility weights are, and we would want some statistical metric that somehow summarizes the space of all possible trade-offs as a function of the threshold c. Moreover, when looking at expansions and recessions, we know for a fact that π is not 1/2. In fact, in the Spanish business cycle –as we have dared to characterize it– if we reach back to 1939, periods of recession represent about 1/3 of the sample (closer to 1/4 in more recent times). Finally, the KS statistic has a non-standard asymptotic distribution.

Luckily, the CCF presented earlier provides a simple solution to these shortcomings and in particular, the *Area Under the CCF* or AUC (to use the same acronym that is used when the area is calculated with the ROC curve for which a voluminous literature exists. For a summary of that literature see, e.g., Pepe, 2003). In its simplest form, the AUC can be easily calculated non-parametrically since Green and Swets (1966) show that $AUC = P[Z > X]$, where Z denotes the random variable associated with observations z_t drawn from y_t when $s_t = 0$; and similarly, X denotes the random variable associated with observations x_t drawn from y_t when $s_t = 1$. Hence, a simple non-parametric estimate is:

$$\widehat{AUC} = \frac{1}{T_0 T_1} \sum_{i=1}^{T_0} \sum_{j=1}^{T_1} I\left(z_i > x_j\right) \tag{6}$$

where i, j are a convenient way to break down the index t into those observations for which $s_t = 0, 1$ respectively, $I(A)$ is the indicator function that takes the value of 1 when the event A is true, 0 otherwise and $T_0 + T_1 = T$ simply denote the total number of observations for which $s_t = 0, 1$ respectively. There are more sophisticated non-parametric estimates of (6) using kernel weights and there are also parametric models (for a good compilation see Pepe, 2003), but expression (6) has intuitive appeal. Under mild regularity conditions and based on empirical process theory (see Kosorok, 2008), Hsieh and Turnbull (1996) show that

$$\sqrt{T}\left(\widehat{AUC} - P[Z > X]\right) \rightarrow N(0, \sigma^2),$$

although in general (specially when y_t is itself the generated from an estimated model), it is recommended that one use the bootstrap. In what follows, we use the AUC as our preferred tool to evaluate our proposed business cycle chronology in a variety of ways.

3.1 Evaluation of Alternative Chronologies

This section compares our proposed business cycle chronology with a chronology proposed by the OECD,[13] and two chronologies provided by the Economic Cycle Research Institute (ECRI):[14] their *business cycle* chronology and their *growth rate* chronology. The latter may or may not result in recessions as their website explains, but we include it for completeness. Table 6 summarizes several experiments used to assess each chronology. First we consider each individual indicator separately and ask how well does each chronology classify the data into the two empirical distributions expansion/recession of the series considered. As the previous section explains, this is the approach that we use to determine how good each chronology is. Next, we repeat the exercise, but

[13] http://www.oecd.org/document/34/0,3746,en_2649_34349_1891170_1_1_1_1,00.html

[14] http://www.businesscycle.com/business_cycles/international_business_cycle_dates

by allowing up to 12 leads and lags of each series to search for that horizon that would maximize the AUC. We do this because some of the chronologies that we consider may be tailored to a single indicator rather than being a combination of dates as we have proposed. This can be particularly problematic since labor related indicators tend to lead into the recession, but exit much later than production indicators. By searching for the optimal horizon, we handicap our own chronology, but also uncover some interesting timing issues associated with each indicator.

Broadly speaking, we find that ECRI's business cycle chronology and ours deliver very similar results whereas ECRI's growth rate and the OECD's chronologies are clearly far inferior, in many cases, no better than the null of no-classification ability. Our proposed chronology tends to do better with labor related indicators (employment, registered unemployed and the wage income index) whereas ECRI's does better with production indicators (GDP and IPI). Looking at the horizon at which our chronology maximizes the AUC, we note that leads between 3 to 8 months would generate slightly higher AUCs. At the front end, this implies delaying the start and/or end of the recessions slightly. However, one has to be careful because the samples available for each indicator are slightly different and in fact, as we will show, the synchronicity between each indicator and chronology at which the AUC is maximized is much better in recent times.

[**Insert Table 6 here**]

If we compare –indicator to indicator– the AUCs of our chronology for Spain against those of the BCDC for the U.S. to provide a benchmark. The results for the U.S. can be found in table 3 of Berge and Jordà (2011). The AUC for GDP in the U.S. is 0.93 compared with 0.82 in Spain; for personal income in the U.S. it is 0.85 compared with 0.94 for the wage index in Spain; industrial production has an AUC of 0.89 in the U.S. versus 0.84 in Spain; and personal/household employment in the U.S. has an AUC of 0.82/0.78 versus an AUC of 0.96 in Spain. Broadly speaking, both chronologies appear to have similar properties, an observation that is further supported by the evaluation of economic activity indexes in the next section.

[**Insert Tables 7 and 8 here**]

Finally, it may be useful to summarize some of the salient features of the business cycles identified for Spain with each available chronology against the business cycles for U.S. data identified by the NBER. A summary of the raw peak and trough dates for each is provided in table 7. Table 8 summarizes the salient features of the recessions using each method and compares these features to U.S. recessions. If we set aside the ECRI-growth chronology for a moment (which ECRI itself warns is not meant to be a chronology of business cycles properly speaking), it is clear that Spain and the U.S. suffer a similar number of recession periods but recessions in the U.S. last less time. In the U.S., the average recession lasts about one year whereas in Spain recessions last over two years on average. The number of months in recession represents less than 20% of the sample in the U.S. but close to 30% in Spain. And as one looks at more recent samples, these differences seem to stay fairly constant or if anything, to be even somewhat worse.

4 Evaluating Economic Activity Indices

A historical record of turning points in economic activity serves primarily as a reference point for academic studies. Moreover, determining the precise date of a turning point requires some time after the event has passed. Due to data revisions and because it is important not to have to revise the dating, the NBER will usually delay by between 12 to 18 months any public announcement of business cycle turning points. But, it is important to have a means to communicate effectively and in real time what is the current situation of the economy. In the U.S., the Chicago Fed National Activity Index[15] or CFNAI, and the Philadelphia Fed Business Conditions Index[16] or ADS to use the more common acronym representing the last names of the authors (Aruoba, Diebold and Scotti, 2009), are two examples of frequently updated economic activity indexes commonly cited in the press and in policy circles. In Spain, we consider four similar indexes: an index produced

[15] http://www.chicagofed.org/webpages/publications/cfnai/index.cfm

[16] http://www.philadelphiafed.org/research-and-data/real-time-center/business-conditions-index/

by FEDEA,[17] a composite index of leading indicators constructed by the OECD,[18] the MICA-BBVA[19] index of Camacho and Doménech (2011), and the Spain-STING[20] of Camacho and Pérez Quirós (2011). In very broad terms, we can characterize these indexes as factors from a model that combines activity indicator variables, sometimes observed at different frequencies. The most commonly cited precursor for this type of index is Stock and Watson (1991).

Figure 8 presents a time series plot of each of the four indexes for Spain, each chart also displaying the recession shaded regions based on the chronology we introduced in table 5. For each index, we then calculated the optimal threshold that would maximize expression (5) but these optimal values are virtually identical to the mean at which the indexes are centered –zero for FEDEA, MICA-BBVA and STING, and 100 for OECD CLI. In terms of how well the indexes correspond to our recession periods, it is easy to see that FEDEA, MICA-BBVA and STING conform rather well so that observations below the zero threshold indicate mostly periods of recession. The OECD index is somewhat more variable and appears to fluctuate by a larger amount between our preferred periods of recession.

[**Insert Figure 8 here**]

[**Insert Table 9 here**]

The observations in figure 8 are confirmed by a more formal analysis presented in table 9. In order to cover our bases, we consider how well our proposed chronology sorts the empirical distributions of expansion/recession for each of the indexes contemporaneously, as well as up to 12 leads and lags of the index. This will reveal whether the indexes work better as lagging or leading indicators. We also consider the sorting ability of chronologies produced by the OECD

[17] FEDEA stands for fundación de estudios de economía aplicada. The index can be found at: http://www.crisis09.es/indice/

[18] www.oecd.org/std/cli

[19] MICA-BBVA stands for factor Model of economic and financial Indicators which is used to monitor the Current develpment of the economic Activity by Banco Bilbao Vizcaya Argentaria. We thank Máximo Camacho and Rafael Doménech for making the data available to us.

[20] STING stands for short-term INdicator of euro area Growth. We thank Máximo Camacho and Gabriel Pérez Quirós for making the data available to us.

and ECRI. In principle, the former ought to match well with the OECD CLI. The exercise thus serves several purposes: it is another form of evaluating the chronology that we propose; it helps determine the lagging, coincident or leading properties of the indexes; and it serves to compare the performance across the indexes themselves.

Off hand, it becomes readily apparent that the OECD and ECRI-growth chronologies are not very good at sorting the data –something we already suspected from the results in the previous section. Their AUC values are often not meaningfully different from the null of no classification ability: to paraphrase Charles S. Peirce, they are the "utterly ignorant" chronologies. As we knew from the analysis in the previous section, our chronology and ECRI's are both similar. Our's attains the highest AUC values across all indexes both contemporaneously or at the optimal lead/lag, but the differences are minor. Within indexes, the suspicions we raised when discussing figure 8 are confirmed. Focusing on our proposed chronology, the STING index achieves the highest contemporaneous score with an AUC = 0.96, which is very close to the perfect classifier ideal of 1. This is closely followed by MICA-BBVA (AUC = 0.93), followed by FEDEA (AUC = 0.89), and far behind OECD (AUC = 0.69). In terms of the optimal lead/lag, STING comes closest to being a contemporaneous indicator with a one-month lag (the OECD CLI attains its maximum contemporaneously, but the AUC is much lower), followed by FEDEA (which attains its maximum with a one-quarter lag) and finishing with MICA-BBVA (with a 5-month lag at which point its AUC is virtually identical to STING's). Except for the OECD, at their optimal the three remaining indexes all achieve AUCs above 0.9.

How does this performance compare with the performance of CFNAI and ADS for the U.S.? Berge and Jordà's (2011) table 5 reports the AUC for CFNAI to be 0.93 and for ADS to be 0.96 using the NBER's business cycle dates, which are essentially the values we have found for MICA-BBVA and STING using our chronology for Spain. This is another dimension one can use to assess our chronology and by and large the results are not materially different from what one finds in the U.S.

5 Turning Point Prediction

A historical chronology of business cycle fluctuations between periods of expansion and recession is an important tool for research. We provide such a chronology in this paper but more helpfully, we present simple methods by which one can generate such a record in a replicable manner, and how one can evaluate whether the proposed chronology is any "good." Determining turning points demands some patience to sort out data revisions and other delays –in real time, indexes on economic activity such as FEDEA, MICA-BBVA and STING offer a reliable indication about the current state of the economy. What about future turning points? This section investigates a collection of potential indicators of future economic activity and constructs turning point prediction tools. The predictions we obtain indicate that economic activity is likely to remain subdued at least until summer of next year (our forecast horizon ends in August 2012). Here is how we go about making this determination.

We begin by exploring a number of candidate variables listed in table 10 and described in more detail in the appendix. The choice of variables does not follow an exhaustive search and we expect that others will come up with additional variables with useful predictive properties. But the variables listed in table 10 will probably resonate with most, and offer a reasonable benchmark. Variables such as cement and steel production; new vehicle registrations; and air passenger and cargo transportation among others, are meant to provide leading indicators on economic output. Financial variables such as Madrid's stock market index and the spread between the three-month and one-year interbank rates have often been found to be good predictors of future economic activity in the U.S. –the S&P 500 index and the spread between the federal funds rate and the 10 year T-bond rate are two of the variables in the index of leading economic indicators produced by the Conference Board. Finally, more recently available survey data, such as consumer confidence, outlook on household finances and economic outlook expectations find its mirror in the consumer confidence survey maintained, among others, by the University of Michigan, which is also a leading indicator used by the Conference Board in the U.S.

In previous work (see Berge and Jordà, 2011), we found that different variables have predictive power at different horizons. This observation suggests that for the purposes of generating forecasts at a variety of distant horizons, it is generally a bad idea to use a one-period ahead model, and then iterate forward to the desired horizon. The reason is that the loadings on the different predictors should probably differ depending on the forecast horizon and iterating the one-period ahead model is likely to put to much weight on good short-term predictors. Moreover, because the important metric here is classification ability rather than model fit, issues of parameter estimation uncertainty play a more secondary role than in traditional forecasting environments, where the root mean square error metric and the usual trade-offs between bias and variance often favor more parsimonious approaches.

Table 10 provides a summary of each variable's classification ability using the AUC and also reports the lead horizon over which the AUC is maximized. For example, the stock market index data has a maximum AUC = 0.65 seven months in the future meaning that this variable should probably receive a relatively high weight when predicting turning points around the half-year mark. The survey data tend to have very high AUCs (all three surveys surpass 0.90), but we should point out that these data go back about 25 years only. By the same token, cement production and Madrid's stock market index have more middling AUCs but go back over 50 years –a more turbulent period that includes the end of the dictatorship, a new Constitution, and a coup d'état attempt– and for which we have to rely on less information to come up with the chronology of turning points.

With these considerations in mind, we are interested in modeling the posterior probabilities $P[s_{t+h} = s|x_t]$ for $h = 1, ..., 12$ and where $s = 0, 1$ with 0 for expansion, 1 for recession and where x_t is a $k \times 1$ vector of indicator variables. We then assume that the log-odds ratio of the expansion/recession probabilities at time h is an affine function of x_t so that

$$\log \frac{P[s_{t+h}=0|x_t]}{P[s_{t+h}=1|x_t]} = \beta_{h0} + \boldsymbol{\beta}_h' x_t; \qquad h = 1, .., 12. \tag{7}$$

This is a popular model for classification in biostatistics and with a long tradition in economics: it is the logit model. In principle, one could use other classification models, for example linear discriminant analysis (LDA), a standard classification algorithm that combines a model such as (7) with a marginal model for x_t. However, Hastie, Tibshirani and Friedman (2009) have argued that it is often preferable to stick to a model such as (7) rather than rely on LDA in practice. For these reasons, we feel the model in expression (7) is a reasonable choice.

Figure 9 displays the results of fitting a model such as (7) to three samples of data. We estimate a *long-range* sample beginning in January 1961 that only includes data on cement production, new car registrations and Madrid's stock market index. For brevity, we have omitted a graph of these predictions although figure 10 displays the AUC of the in-sample classification ability of such a modeling approach. We see that these long-range predictors carry a modest degree of predictive ability, again with the caveat that this sample covers a series of turbulent periods. The *medium-range* sample begins in January 1976 and adds data on imports and exports, new registered firms, steel production and new truck registrations. The top panel of figure 9 displays the one-step ahead probability predictions of recession against our recession dates. Because the data ends in August 2011, after that date we produce out of sample predictions on the odds of recession up to 12 months into the future. The middle panel of figure 10 displays the AUC of the in-sample classification ability for each horizon when using this set of predictors. Finally, the *short-range* sample begins in June 1986 and adds to the set of predictors consumer confidence survey data. The top panel presents the one-period ahead probability of recession forecasts and the out of sample forecasts starting August 2011 and ending August 2012. The right-hand panel of figure 10 displays the AUC of the in-sample classification ability of the model 1 to 12 periods into the future.

Estimates based on all the data (but for the short-range sample) have very good in-sample classification ability (the sample is too short for any serious out-of-sample evaluation) and even at the 12-month ahead mark, the AUC remains above 0.90. This is easy to see in the top panel of figure 9 as well, with well delineated probabilities that coincide well with our proposed chronology.

Few will be surprised by the predictions that we report: Regardless of the sample chosen, the outlook of the Spanish economy going to August 2012 is dim. The medium-range model uses less data but contains more observations. As we can see in figure 10, the forecast is somewhat noisier. At forecast horizons of one month, the model produces an in-sample AUC near 0.90, which then tapers toward 0.75 at the 12-month mark. The top panel of figure 9 therefore displays a noisier predicted probability series but the forecasts beginning in August 2011 are all above 0.5 (they are in fact increasing over time). Focusing on the model that uses all indicators, we see in figure 10 that this model seems to produce a more accurate signal of the risks of recession at all horizons. Again, in figure 9 we see that this model also portends troubled times for the Spanish economy, as the out-of-sample forecasts from this model continue to hover near 100%.

6 Conclusion

A major area of macroeconomic research investigates the alternating periods of expansion and contraction experienced by economies as they grow. Business cycle theory, seeks to understand the causes, consequences and policy alternatives available to tame these economic fluctuations. One of the empirical foundations on which this research edifice rests is an historical record of when the economy drifts from one state to another. This paper shows how to construct and assess such a record and applies the proposed methods to Spanish economic activity.

The most venerable of business cycle chronologies is surely to be found in the U.S., with the NBER as its custodian. One of the objectives of this paper was to systematize the manner in which the NBER's BCDC determines turning points to generate a similar historical record for Spain. The overriding principles we sought was to strive for simplicity, transparency, reproducibility and formal assessment. We hope on that score to have provided the beginnings of a formal reconsideration of the Spanish business cycle chronology.

A historical record of expansion and recession periods has significant academic value. However a policymaker's actions are guided by the current and future state of the economy. We find that existing indexes of economic activity provide a clear picture in real time about that state, much like similar indexes available for the U.S. speak about the American economic cycle.

Preemptive policymaking requires an accurate reading of the future. As Charles S. Peirce recognized back in 1884, the actions taken as a result of a forecast require that we rethink how probability forecasts are constructed and evaluated. The usual bias-variance trade-offs neatly encapsulated in the traditional mean-square error loss need to make way for methods that reorient some of the focus toward assessing classification ability. Using this point of view, we construct predictive models on the odds of recession that have good classification skill for predictions 1- to 12-months into the future.

The last word on the past, present and future of the Spanish business cycle has not yet been written. We hope instead that our modest contribution serves to organize the conversation on how our chronology of turning points could be improved.

7 Data Appendix

7.1 Yearly Frequency

- **Real GDP per capita** (Producto Interior Bruto per capita, precios constantes de 2000, en euros). **Sample:** 1850-2008. **Source:** Leandro Prados de la Escosura (2003), see: `http://e-archivo.uc3m.es/bitstream/10016/4518/1/wh0904.pdf`

7.2 Quarterly Frequency

- **Real GDP** (Producto Interior Bruto a precios constantes, 1986 pta, 1995 euro, 2000 euro). **Samples:** 1970Q1- 1998Q4; 1980Q1-2004Q4; 1995Q1-2011Q2. **Source:** Contabilidad Nacional Trimestral de España. Instituto Nacional de Estadística: `www.ine.es`

- **Real GDP yearly growth rate (**Tasa de variación anual del Producto Interior Bruto, base 2000 en euros). **Sample:** 1970Q1-2011Q2. **Source:** Contabilidad Nacional Trimestral de España. Instituto Nacional de Estadística: `www.ine.es`

- **Employment** (Ocupados, Encuesta de Población Activa). **Sample:** 1976Q3-2011Q2. **Source:** Encuesta de Población Activa, Instituto Nacional de Estadística: `www.ine.es`

7.3 Monthly Frequency

7.3.1 Series Used for Turning Point Chronology

- **Real GDP linearly interpolated from quarterly to monthly.**

- **Employment linearly interpolated from quarterly to monthly.**

- **Industrial Production Index (**Índice de Producción Industrial, base 2005). **Sample:** January 1975 to August 2011. **Source:** Instituto Nacional de Estadística: `www.ine.es`

- **Registered Unemployed (**Paro registrado, personas). **Sample:** September 1939 to September 2011. **Source:** Instituto de Empleo Servicio Público de Empleo Estatal (INEM), Instituto Nacional de Estadística: `www.ine.es`

- **Real Wage Income Index** (Indicador de Renta Salarial). **Sample:** January 1977-September 2011. **Source:** Ministerio de Economía y Hacienda: `www.meh.es`

7.3.2 Indexes of Economic Activity

- **OECD Composite Leading Indicators. Sample:** September 1963-August 2011. **Source:** `www.oecd.org/std/cli`. **Component series:** Production: future tendency manufacturing, % balance; Order books/demand: future tendency in manufacturing, % balance; Finished goods stocks: level manufacturing, % balance, inverted; **Source:** Ministerio de Industria, Comercio y Turismo. Nights in hotels (number). **Source:** Instituto Nacional de Estadística: `www.ine.es`. Yield over 2-year government bonds (% per annum) inverted. **Source:** Banco de España.

- **FEDEA. Sample:** January 1984-October 2011. **Source:** `www.crisis09.es/indice/calendario.html`. **Component series:** beginning 1982, real GDP (PIB, **Source:** Contabilidad Nacional Trimestral de España. Instituto Nacional de Estadística: www.ine.es), electricity consumption (**Source:** Red Eléctrica de España), social security afiliations (**Source:** Ministerio de Trabajo). Beginning 1987, add survey of consumer sentiment (**Source:** European Comission). Beginning 1989, add new car registrations (**Source:** Asociación Española de Fabricantes de Automoviles). Beginning 1993 add industrial production index (Instituto Nacional de Estadística). Beginning 1995 add retail sales (**Source:** Instituto Nacional de Estadística).

- **MICA-BBVA. Sample:** January 1981-October 2011. **Source:** see Camacho and Doménech (2011) for details.

- **Spain-STING. Sample:** January 1984-October 2011. **Source:** see Camacho and Pérez Quirós (2011) for details.

7.3.3 Leading Indicators

Rather than listing individual sources we note that these data can be downloaded from the Boletín Estadístico del Banco de España at: `http://www.bde.es/webbde/es/estadis/infoest/bolest.html`

- **Electricity Production in Kw/hr (millions). Sample:** January 1977 to July 2011.

- **Cement Production in metric tons. Sample:** January 1955 to September 2011.

- **Steel Production in metric tons. Sample:** January 1968 to July 2011.

- **New Truck and Bus Registrations. Sample:** January 1964 to September 2011.

- **New Car Registrations. Sample:** January 1960 to September 2011.

- **Number of Hotel Nights. Sample:** April 1965 to August 2011.

- **Number of Air Passengers and Metric Tons of Air Cargo. Sample:** January 1965 to July 2011.

- **Consumer Confidence, Household Outlook and Economic Outlook Surveys. Sample:** June 1986 to October 2011.

- **Exports and Imports. Sample:** September 1971 to August 2011.

- **Madrid Stock Exchange. Sample:** January 1950 to October 2011.

- **Interbank Rates. Sample:** September 1979 to September 2011.

- **New Registered Firms. Sample:** January 1967 to August 2011.

References

Aruoba, Borağan, Francis X. Diebold and Chiara Scotti. 2009. Real-Time Measurement of Business Conditions. *Journal of Business and Economic Statistics,* 27(4): 417-427.

Baker, Stuart G., and Barnett S. Kramer. 2007. Pierce, Youden, and Receiver Operating Characteristics Curves. *The American Statistician*, 61(4): 343-346.

Berge, Travis J. and Òscar Jordà. 2011. Evaluating the Classificatin of Economic Activity into Recessions and Expansions. *American Economic Journal: Macroeconomics,* 3(2): 246-277.

Berge, Travis J., Òscar Jordà, and Alan M. Taylor. 2011. Currency Carry Trades. In International Seminar on Macroeconomics, 2010, Richard H. Clarida, Jeffrey A. Frankel, and Francesco Giavazzi (eds.), Chicago: The University of Chicago Press for the National Bureau of Economic Research.

Bry, Gerhard and Charlotte Boschan. 1971. Cyclical Analysis of Time Series: Selected Procedures and Computer Programs. New York: National Bureau of Economic Research.

Burns, Arthur F. and Wesley C. Mitchell. 1946. Measuring Business Cycles. New York: National Bureau of Economic Research.

Camacho, Máximo and Rafael Doménech. 2011. MICA-BBVA: A Factor Model of Economic and Financial Indicators for Short-Term GDP Forecasting. *SERIEs: Journal of the Spanish Economic Association,* DOI 10.1007/s13209-011-0078-z.

Camacho, Máximo and Gabriel Pérez Quirós. 2011. Spain-STING: Spain short term Indicator of Growth. *The Manchester School,* 79: 594-616.

Fushing Hsieh, Shu-Chun Chen and Chii-Ruey Hwang. 2010a. Non-parametric Decoding on Discrete Time Series and Its Applications in Bioinformatics. *Statistics in Biosciences*, forthcoming.

Fushing Hsieh, Shu-Chun Chen and Chii-Ruey Hwang. 2010b. Discovering Stock Dynamics through Multidimensional Volatility-Phases. *Quantitative Finance.* doi:10.1080/14697681003743040.

Fushing Hsieh, Chii-Ruey Hwang, T-K Leee, Yen-Chiu Lan, Shwu-Bin Horng. 2006. Exploring and Reassembling Patterns in Female Been Weevil's Cognitive Processing Networks. *Journal of theoretical Biology,* 238(4): 805-816.

Fushing Hsieh, Shu-Chun Chen, Travis J. Berge and Òscar Jordà. 2010. A Chronology of International Business Cycles Through Non-Parametric Decoding. U.C. Davis, mimeo.

Geman, Stuart and Kevin Kochanek. 2001. Dynamic Programming and the Graphical Representation of Error-Correcting Codes. *IEEE Transactions on Information Theory,* 47(2): 549-568.

Green, David M., and John A. Swets. 1966. Signal Detection Theory and Psychophysics. Los Altos, CA.: Peninsula Publishing.

Hamilton, James D. 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica,* 57(2): 357-384.

Harding, Don and Adrian Pagan. 2002a. Dissecting the Cycle: A Methodological Investigation. *Journal of Monetary Economics*, 49: 365-381.

Harding, Don and Adrian Pagan. 2002b. A Comparison of Two Business Cycle Dating Methods. *Journal of Economic Dynamics and Control,* 27: 1681-1690.

Hastie, Trevor, Robert Tibshirani and Jerome Friedman. 2009. The Elements of Statistical Learning: Data Mining, Inference and Prediction. New York: Springer

Hsieh, Fushing and Bruce W. Turnbull. 1996. Nonparametric and Semiparametric Estimation of the Receiver Operating Characteristics Curve. *Annals of Statistics*, 24: 25-40.

Jaynes, Edwin T. 1957a. Information Theory and Statistical Mechanics I. *Physical Review,* 106: 620.

Jaynes, Edwin T. 1957b. Information Theory and Statistical Mechanics II, *Physical Review,* 108: 171.

Jordà, Òscar, Mortiz Schularick and Alan M. Taylor. 2011. Financial Crises, Credit Booms and External Imbalances: 140 Years of Lessons. *IMF Economic Review*, 59: 340-378.

Jordà, Òscar and Alan M. Taylor. 2009. The Carry Trade and Fundamentals: Nothing to Fear but FEER itself. NBER Working Papers no. 15518.

Jordà, Òscar and Alan M. Taylor. 2011. Performance Evaluation of Zero Net-Investment Strategies. NBER Working Papers no. 17150.

King, Robert G. and Charles I. Plosser. 1994. Real Business Cycles and a Test of the Adelmans. *Journal of Monetary Economics*, 33(2): 405-438.

Kim, Chang-Jin and Charles R. Nelson. 1999. State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications. Cambridge, MA: MIT Press.

Kolmogorov, Andrey N. 1933. Sulla Determinazione Empirica di una Legge di Distribuzione. *Giornale dell'Istituto Italiano degli Attuari,* 4:83-91.

Kose, M. Ayhan, Eswar S. Prasad and Marco E. Terrones. 2003. How Does Globalization Affect the Synchronization of Business Cycles? *American Economic Review, Papers and Proceedings,* 93(2): 57-62.

Kose, M. Ayhan, Christopher Otrok and Charles H. Whiteman. 2008. Understanding the Evolution of World Business Cycles. *Journal of International Economics,* 75(1): 110-130.

Kose, M. Ayhan, Christopher Otrok and Eswar S. Prasad. 2008. Global Business Cycles: Convergence or Decoupling? NBER Working Papers no. 14292.

Kosorok, Michael R. 2008. Introduction to Empirical Processes and Semiparametric Inference. New York: Springer.

Lusted, Lee B. 1960. Logical Analysis in Roentgen Diagnosis: Memorial Fund Lecture. *Radiology,* 74(2): 178-193.

Mason, I. 1982. A Model for Assessment of Weather Forecasts. *Australian Metereological Magazine,* 30(4): 291-303.

Mitchell, Wesley C. 1927. Business Cycles: The Problem and Its Setting. New York: National Bureau of Economic Research.

Peirce, Charles S. 1884. The Numerical Measure of the Success of Predictions. *Science* 4: 453-454.

Pepe, Margaret S. 2003. The Statistical Evaluation of Medical Tests for Classification and Prediction. Oxford: Oxford University Press.

Peterson, W. Wesley, and Theodore G. Birdsall. 1953. The Theory of Signal Detectability: Part I. The General Theory. Electronic Defense Group, Technical Report 13, June 1953. Available from EECS Systems Office, University of Michigan.

Poincaré, J. Henri. 1890. Sur les Equations aux Dérivées Partielles de la Physique Mathématique. *American Journal of Mathematics*, 12(3): 211-294.

Prados de la Escosura, Leandro. 2003. El Progreso Económico de España (1850-2000). Bilbao, Spain: Fundación BBVA.

Schularick, Moritz and Alan M. Taylor. Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008. NBER Working Papers no. 15512. Forthcoming *American Economic Review.*

Smirnov, Nikolai V. 1939. Estimate of Deviation between Empirical Distribution Functions in Two Independent Samples. *Bulletin Moscow University*, 2: 3-16.

Spackman, Kent A. 1989. Signal Detection Theory: Valuable Tools for Evaluating Inductive Learning. *Proceedings of the Sixth International Workshop on Machine Learning.* Ithaca, June.

Stock, James H. and Mark W. Watson. 1991. A Probability Model of the Coincident Economic Indicators. In: Lahiri K., Moore G. (eds.) Leading Economic Indicators, New Approaches and Forecasting Records. Cambridge, U.K.: Cambridge University Press.

Stock, James H. and Mark W. Watson. 2010. Estimating Turning Points Using Large Data Sets. NBER working paper 16532.

Viterbi, Andrea J. 1967. Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm. *IEEE Transaction on Information Theory,* 13(2): 260-269.

Watts, Duncan J. and Stephen H. Strogatz. 1998. Collective Dynamics of 'Small-World' Networks. *Nature*, 393(4): 440-442.

Youden, W. J. 1950. Index for Rating Diagnostic Tests. *Cancer* 3, 32-35.

Table 1. Dates of Peaks and Troughs of Economic Activity Based on Yearly Real GDP per Capita: 1850-2008. Based on the Bry and Boschan (1971) Algorithm

1850-1900		1901-1939		1940-2010	
Peak	**Trough**	**Peak**	**Trough**	**Peak**	**Trough**
1852	1853	1901	1902	1940	1941
1855	1856	1903	1905	1944	1945
1863	1865	1909	1910	1952	1953
1866	1868	1911	1912	1958	1959
1873	1874	1913	1914	1974	1975
1877	1879	1916	1918	1978	1979
1883	1887	1925	1926	1980	1981
1892	1893	1927	1928	1992	1993
1894	1896	1929	1931	2007	
		1932	1933		
		1935	1938		

Notes: Source of the data Prados de La Escosura (2003). See text for a description of the Bry and Boschan (1971) algorithm used to generate these dates.

Table 2. Recession Dates Based on Quarterly Data on Real GDP and Employment Using BBQ

RGDP (1986b, SA, Pta) 1970Q1-1998Q4		RGDP(1995b, SA, Euro) 1980Q1-2004Q4		RGDP(2000, SA, Euro) 1995Q1-2011Q2		Employment, SA 1976Q3-2011Q2	
Peak	Trough	Peak	Trough	Peak	Trough	Peak	Trough
1974Q4	1975Q2	-	-	-	-	-	-
1978Q2	1979Q1	-	-	-	-	-	-
1980Q3	1981Q1	1980Q4	1981Q2	-	-	-	-
						-	1985Q2
1992Q2	1993Q3	1992Q1	1993Q2	-	-	1991Q1	1994Q1
-	-	-	-	2008Q2	2009Q3	2008Q1	-

Notes: See appendix for data sources. The differences in dating for the 1992-1993 recession are due to differing base-years for price deflators in the underlying data.

Table 3. Recession Dates Based on Monthly Data on Interpolated Real GDP and Interpolated Employment, Registered Unemployed, Industrial Production Index, and Wage Income Index Using Bry-Boschan Algorithm

RGDP-Q		RGDP-M		EMP-Q		EMP-M		Reg. Un.		IPI		Wage Inc	
P	T	P	T	P	T	P	T	P	T	P	T	P	T
-	-	-	-	-	-	-	-	Dec-1939	Jan-1941	-	-	-	-
-	-	-	-	-	-	-	-	Apr-1945	May-1946	-	-	-	-
-	-	-	-	-	-	-	-	Mar-1948	Jan-1951	-	-	-	-
-	-	-	-	-	-	-	-	Apr-1953	Oct-1954	-	-	-	-
-	-	-	-	-	-	-	-	Mar-1959	Nov-1960	-	-	-	-
-	-	-	-	-	-	-	-	Nov-1962	Jan-1965	-	-	-	-
-	-	-	-	-	-	-	-	Mar-1970	Jan-1972	-	-	-	-
Q4-1974	Q2-1975	Feb-1974	May-1975	-	-	-	-	Feb-1974	May-1975	-	-	-	-
Q2-1978	Q1-1979	May-1978	Feb-1979	-	-	-	-	Oct-1977	Mar-1978	-	-	-	-
Q2-1980	Q1-1981	Jun-1980	Mar-1981	-	-	-	-	Feb-1980	Jan-1983	Oct-1979	Feb-1982	Aug-1978	-
-	-	-	-	-	Q2-1985	-	Jun-1985	Nov-1983	Sep-1984	Feb-1984	Aug-1986	-	May-1985
Q1/Q2-1992	Q2/Q3-1993	Feb-1992	Apr-1993	Q1-1991	Q1-1994	Mar-1991	Feb-1994	Jun-1992	Jan-1994	Oct-1989	Jul-1993	Jun-1992	Feb-1994
-	-	-	-	-	-	-	-	-	-	Jun-1995	Sep-1998	-	-
-	-	-	-	-	-	-	-	Mar-2001	Jan-2004	Jun-2000	Aug-2001	-	-
-	-	-	-	-	-	-	-			Jun-2004	Aug-2006	-	-
Q2-2008	Q3-2009	Jun-2008	Sep-2009	Q1-2008	-	Feb-2008	-	Mar-2007	-	Jun-2007	Sep-2009	Jul-2007	-
-	-	-	-	-	-	-	-	-	-	Jun-2010	-	-	-

Notes: quarterly data are linearly interpolated to monthly. This is the same procedure that the BCDC uses at the NBER and is described in: www.nber.org/cycles/. All data are seasonally adjusted first. For industrial production, we also corrected for seasonal means due to strong August effects. The monthly data are smoothed with a double sided symmetric moving average of length 6 in either direction (6 lags, 6 leads and the current observation). If a recession is shorter than 6 months, it is disregarded. There are a handful of small adjustments relative to the automatic procedures coded in STATA that replicate this table due to beginning- or end-of-sample effects but these are very few.

Table 4. A Comparison of BBQ, Hamilton (1989) and HFS on Real GDP Growth: 1971Q1 – 2011Q2.

BBQ		Hamilton		HFS	
Peak	**Trough**	**Peak**	**Trough**	**Peak**	**Trough**
1974Q4	1975Q2	-	-	1973Q1	1975Q2
1978Q2	1979Q1	1979Q1	1979Q3	1977Q1	1979Q1
1980Q2	1981Q1	-	-	1980Q3	1981Q3
-	-	-	-	1984Q2	1985Q4
-	-	-	-	1987Q4	1991Q3
1992Q1/2	1993Q2/3	1992Q3	1993Q3	1992Q1	1992Q4
-	-	-	-	2000Q1	2002Q2
2008Q2	2009Q3	2008Q3	2010Q2	2007Q1	2009Q1

Notes: The first column records the same dates reported in tables 2 and 3. The second column record the peak and trough dates using Hamilton's (1989) smoothed recession probabilities and a cutoff recession probability of 0.5. The third column refers to the dates uncovered with the Hierarchical Factor Segmentation (HFS) algorithm and replicate those reported in Fushing, Chen, Berge and Jordà (2010).

Table 5. A Chronology of the Business Cycle since 1850

Based on Yearly GDP 1850-1939		Based on Monthly Indicators Jan-1939 to Oct-2011	
Peak	**Trough**	**Peak**	**Trough**
1852	1853	Dec-1939	Jan-1941
1855	1856	Apr-1945	May-1946
1863	1865	Mar-1948	Jan-1951
1866	1868	Apr-1953	Oct-1954
1873	1874	Mar-1959	Nov-1960
1877	1879	Nov-1962	Jan-1965
1883	1887	Mar-1970	Jan-1972
1892	1893	Feb-1974	May-1975
1894	1896	Aug-1978	Feb-1979
1901	1902	Feb-1980	Feb-1982
1903	1905	Feb-1984	Sep-1984
1909	1910	Feb-1992	Jan-1994
1911	1912	Jul-2007	-
1913	1914		
1916	1918		
1925	1926		
1927	1928		
1929	1931		
1932	1933		
1935	1938		

Notes: Dates from 1850 to 1939 based on applying the Bry and Boschan (1971) algorithm to historical real GDP per capita data in 2000 Euros constructed by Prados de la Escosura (2003). Dates from January 1939 to October 2011 constructed using five monthly indicators: registered unemployed, linearly interpolated real GDP, linearly interpolated employment (household survey), industrial production index, and wage income index. We apply the Bry and Boschan (1971) to each series in the manner described in the text. Then we generate the reported chronology using the dates that correspond to an incidence rate above 50%. We do not date the last trough although some of the data would indicate that it occurred sometime in late 2009 early 2010.

Table 6. Evaluating Four Business Cycle Chronologies against Indicators of Economic Activity

			Berge-Jordà		ECRI		ECRI-growth		OECD	
		N	h=0	max	h=0	max	h=0	max	h=0	max
Employment	AUC	406	**0.963**	**0.974**	0.921	0.941	0.500	0.645	0.563	0.617
	s.e.		(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.03)	(0.03)
	Horizon			3		6		12		8
GDP	AUC	484	0.823	0.864	**0.918**	**0.921**	0.424	0.697	0.579	0.667
	s.e.		(0.02)	(0.02)	(0.01)	(0.01)	(0.03)	(0.02)	(0.03)	(0.02)
	Horizon			8		2		12		8
IPI	AUC	435	0.841	0.841	**0.865**	**0.880**	0.520	0.816	0.696	0.776
	s.e.		(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)
	Horizon			0		-3		10		6
Reg. Unem.	AUC	501	**0.774**	**0.864**	0.742	0.743	0.409	0.583	0.405	0.661
	s.e.		(0.02)	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)
	Horizon			6		-1		12		12
Wages	AUC	405	**0.935**	**0.958**	0.909	0.937	0.408	0.589	0.550	0.607
	s.e.		(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.03)	(0.03)
	Horizon			6		7		12		10

Notes: For each chronology we calculate the AUC contemporaneously and then by allowing the dependent variable up to 12 leads and lags. We then report the horizon that maximizes the classification ability of the chronology. For example, the employment entry for the Berge-Jordà dates achieves its maximum AUC of 0.974 at h = 3. This means that a 3-month lead on the employment data is the best classifier for our chronology contemporaneously. We mark in bold those entries that are highest across chronologies.

Table 7. Comparing Chronologies: The List of Dates for Spain and the U.S.

U.S.		SPAIN							
NBER		Berge-Jordà		OECD		ECRI		ECRI-Growth	
Peak	**Trough**	**Peak**	**Trough**	**Peak**	**Trough**	**Peak**	**Trough**	**Peak**	**Trough**
-	-	Dec-39	Jan-41	-	-	-	-	-	-
Feb-45	Oct-45	Apr-45	May-46	-	-	-	-	-	-
Nov-48	Oct-49	Mar-48	Jan-51	-	-	-	-	-	-
Jul-53	May-54	Apr-53	Oct-54	-	-	-	-	-	-
Aug-57	Apr-58	-	-	-	-	-	-	-	-
Apr-60	Feb-61	Nov-62	Jan-65	Dec-61	Jun-63	-	-	-	-
-	-	-	-	Apr-66	Apr-68	-	-	-	-
Dec-69	Nov-70	Mar-70	Jan-72	May-69	Apr-71	-	-	-	-
Nov-73	Mar-75	Feb-74	May-75	Feb-74	Feb-76	-	-	Jan-73	Jan-75
-	-	Aug-78	Feb-79					Jul-76	Mar-79
Jan-80	Jul-80	Feb-80	-	Aug-79	-	Mar-80	-	Mar-80	Sep-81
Jul-81	Nov-82	-	Feb-82	-	Aug-82	-	-	-	-
-	-	Feb-84	Sep-84	Dec-83	May-85	-	May-84	May-83	May-85
Jul-90	Mar-91	-	-	Jun-89	Mar-91	-	-	Aug-89	-
-	-	Feb-92	Jan-94	Dec-91	Apr-93	Nov-91	Dec-93	-	Feb-93
-	-	-	-	Feb-95	Aug-96	-	-	-	-
-	-	-	-	Apr-98	-	-	-	Sep-97	Feb-99
Mar-01	Nov-01	-	-	-	Apr-02	-	-	Feb-00	Jun-02
-	-	-	-	-	-	-	-	Oct-03	Aug-04
Dec-07	Nov-09	Jul-07	?	Jan-08	Apr-09	Feb-08	?	Jan-07	Mar-09
-	-	-	-	-	-	-	-	Apr-10	?

Notes: Columns denote peak and trough dates from NBER (first column), our proposed chronology (second column), the OECD's chronology (third column), ECRI's business cycle chronology (fourth column), and ECRI's growth chronology (fifth column). See text for details.

Table 8. Comparing the Cyclical Properties of Chronologies for Spain against the NBER's Chronology for the U.S.

	Berge-Jorda	OECD	ECRI	ECRI-Growth	U.S. - NBER
Sample	September 1939 to September 2011				
Number of recessions	13	--	--	--	12
Total months in recession	284	--	--	--	147
Percentage of months in recession	32.8	--	--	--	17.0
Average length of recession (months)	21.8	--	--	--	12.3
Sample	January 1960 to September 2009				
Number of recessions	9	11	--	--	8
Total months in recession	192	270	--	--	106
Percentage of months in recession	30.9	43.4	--	--	17.1
Average length of recession (months)	21.3	24.5	--	--	13.3
Sample	January 1970 – September 2011				
Number of recessions	7	9	3	12	7
Total months in recession	154	218	121	271	94
Percentage of months in recession	30.7	43.5	24.2	54.1	18.8
Average length of recession (months)	22.0	24.2	40.3	22.6	13.4

Notes: The latest release for ECRI recession dates is February 2011. The latest release for ECRI-growth dates is October 2011. The table assumes that both series continue up to September 2011 with no known trough date. The first Berge-Jordà recession in the January 1960 sample begins in March 1959. The first OECD recession in the January 1970 sample begins in May 1969. The first NBER recession in the January 1970 begins in December 1969.

Table 9. Indexes of Economic Activity and Business Cycle Chronologies

			Berge-Jordà		OECD		ECRI		ECRI-growth	
		N	**h = 0**	**max**	**h = 0**	**max**	**h = 0**	**max**	**h = 0**	**max**
FEDEA	**AUC**	358	0.888	0.904	0.506	0.787	0.872	0.893	0.421	0.728
	se		(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)
	Horizon		0	3	0	-12	0	-7	0	12
MICA BBVA	AUC	371	0.932	0.958	0.508	0.663	0.927	0.950	0.480	0.637
	se		(0.01)	(0.01)	(0.03)	(0.03)	(0.01)	(0.01)	(0.03)	(0.03)
	Horizon		0	5	0	-12	0	5	0	-12
OECD CLI	AUC	576	0.687	0.687	0.772	0.855	0.678	0.685	0.484	0.696
	se		(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)
	Horizon		0	0	0	6	0	-4	0	12
STING	AUC	334	0.959	0.961	0.563	0.764	0.939	0.954	0.532	0.720
	se		(0.01)	(0.01)	(0.03)	(0.03)	(0.01)	(0.01)	(0.03)	(0.03)
	Horizon		0	1	0	-12	0	-5	0	11

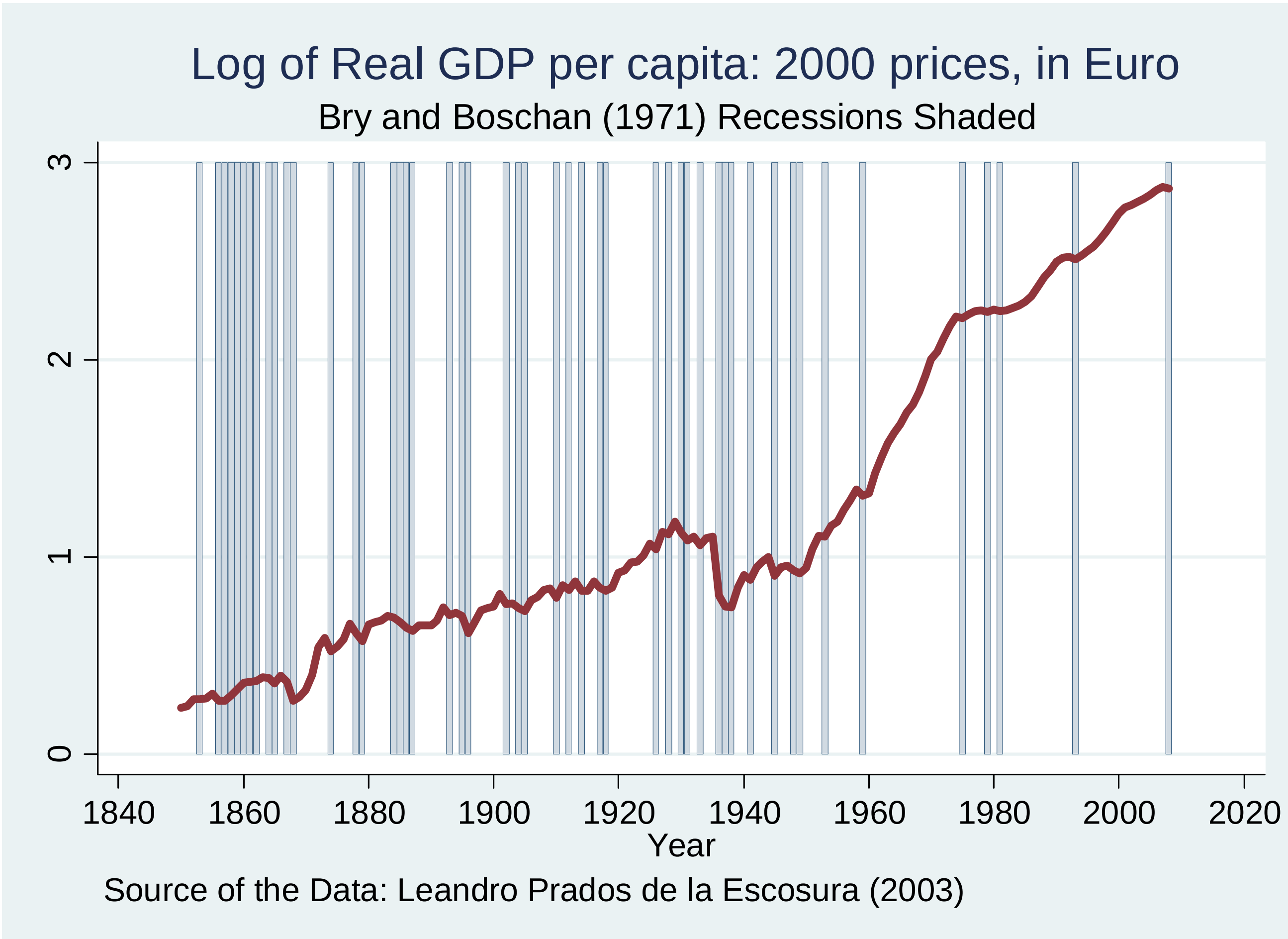
Notes: We consider four chronologies of business cycles beginning with the one proposed by us in table 5 above, the OECD's, ECRI Business Cycle Peak and Trough Dates, and ECRI Growth Rate Cycle Peak and Trough Dates. For each chronology we calculate the AUC for each index of economic activity contemporaneously and then calculate the lead/lag (+/-) at which the AUC is maximized. For example for the FEDEA index and the Berge-Jordà chronology, the maximum AUC of 0.904 is achieved when the 3-month lead of FEDEA is used to classify our chronology contemporaneously.

Table 10. Classification Ability of Different Candidate Leading Indicators

Indicator	N	h = 0	h-max	Indicator	N	h = 0	h-max
Air Cargo	547	0.55	0.53	Hotel Nights	533	0.54	0.55
se		0.03	0.03	se		0.03	0.03
h			12	h			10
Air Passengers	547	0.62	0.64	Spread	373	0.64	0.64
se		0.03	0.03	se		0.03	0.03
h			4	h			0
New Bus Registrations	549	0.59	0.61	Madrid's Stock Exchange	718	0.64	0.65
se		0.03	0.03	se		0.02	0.02
h			3	h			7
Consumer Confidence	305	0.97	0.97	Imports	456	0.55	0.55
se		0.01	0.01	se		0.03	0.03
h			0	h			0
Car Registrations	597	0.66	0.67	New Firm Registrations	512	0.61	0.61
se		0.03	0.03	se		0.03	0.03
h			3	h			0
Cement Production	657	0.72	0.72	Steel Production	499	0.55	0.58
se		0.02	0.02	se		0.03	0.03
h			0	h			12
Economic Outlook	293	0.96	0.96	New Registered Truck	549	0.78	0.78
se		0.01	0.01	se		0.02	0.02
h			0	h			0
Electricity Production	391	0.71	0.71	Exports	456	0.56	0.56
se		0.03	0.03	Se		0.03	0.03
h			0	h			0
Household Outlook	293	0.98	0.98				
se		0.01	0.01				
h			0				

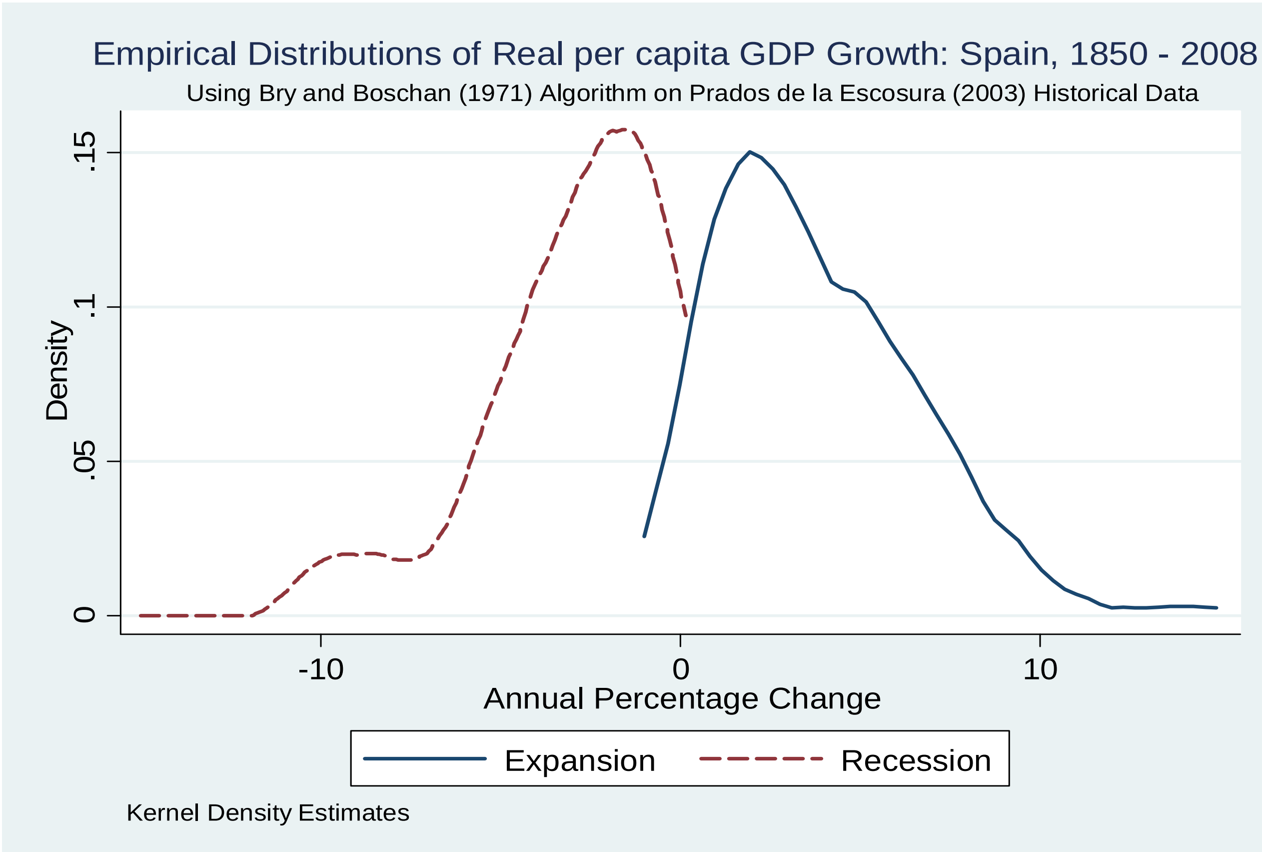
Notes: Each indicator is transformed by taking the year-on-year log difference to obtain a yearly growth rate except for "Spread" which is the spread between the 1-year and the 3-months interbank rates. We refer to the contemporaneous classification ability as h = 0 whereas h-max refers to that horizon in the future for which the current observation of the indicator attains the highest AUC. N refers to the number of observations.

Figure 1. Real GDP per Capita: 1850-2008. Bry and Boschan (1971) Recessions



Notes: The shaded regions are the recession periods reported in table 1. See text for details.

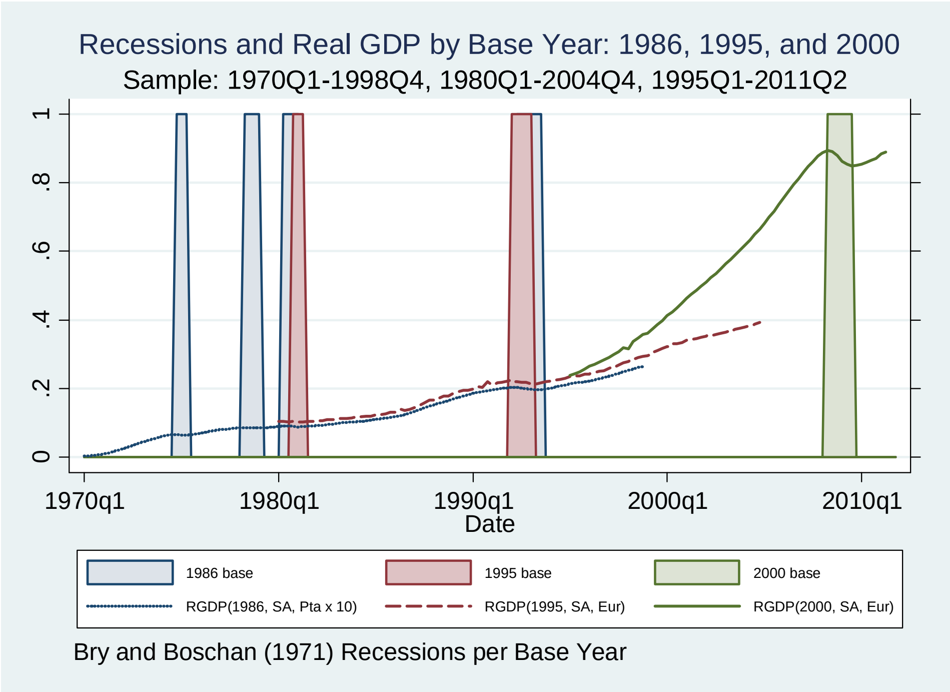
Figure 2. Kernel Estimates of the Empirical Mixture Distribution of Real GDP per Capita, 1850 – 2008.



Notes: Kernel density estimates of the empirical distribution of per capita real GDP growth annualized in expansions (solid line) and in recessions (dashed line). Recession dates are based on the dates in table 1 and displayed in figure 1.

Figure 3. Recession Dating Using Quarterly Real GDP Data and Employment Based on the Bry and Boschan (1971) Algorithm

Panel 1. Real GDP Windows based on Original Source Using 1986, 1995, and 2000 as Base Years.



Panel 2. Total Employment

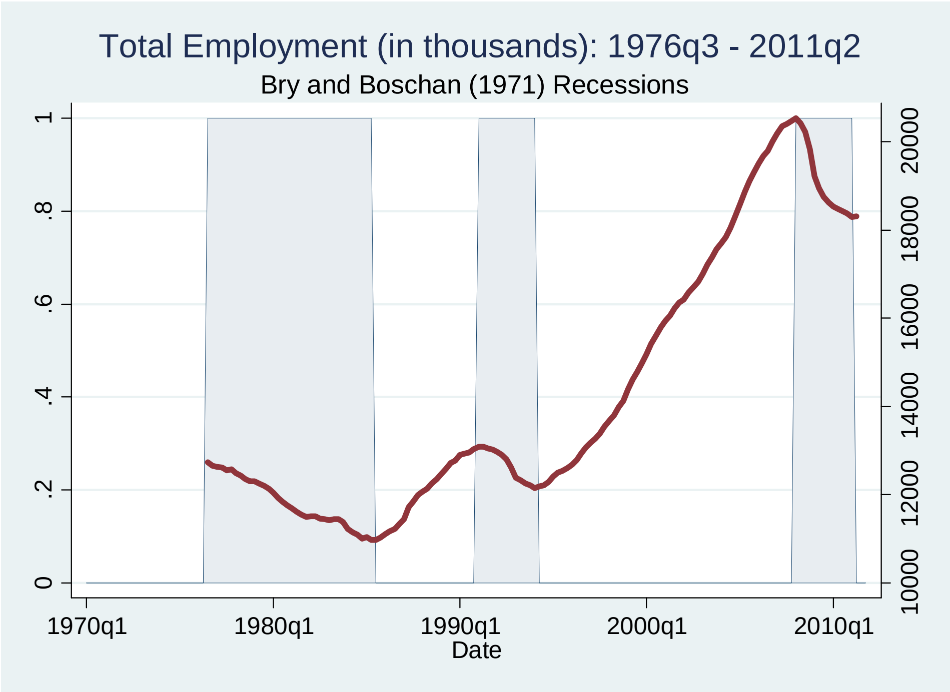
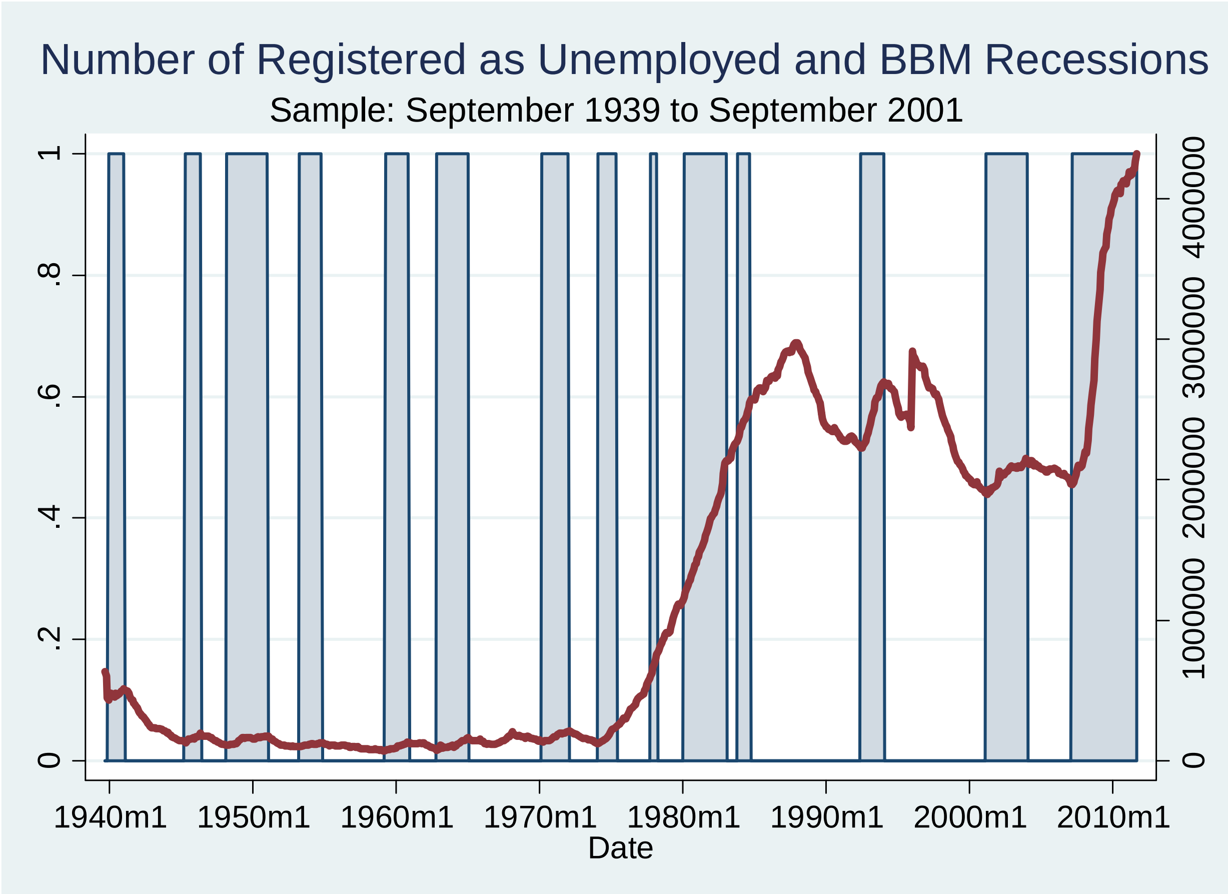


Figure 4. Number of Registered Unemployed and BBM Recessions



Notes: Recessions are computed separately over three regions determined by breaks in 1975 and 1985. For the middle regime, we detrend the data linearly. Also notice that there is a break in the way registered unemployed are accounted for the observation dated November 1995.

Figure 5. The Hamilton (1989) Smoothed Transition Probabilities against the BBM Recessions. Based on Yearly Growth Rate of Real GDP: 1970Q1 – 2011Q2 then Interpolated Linearly to Monthly

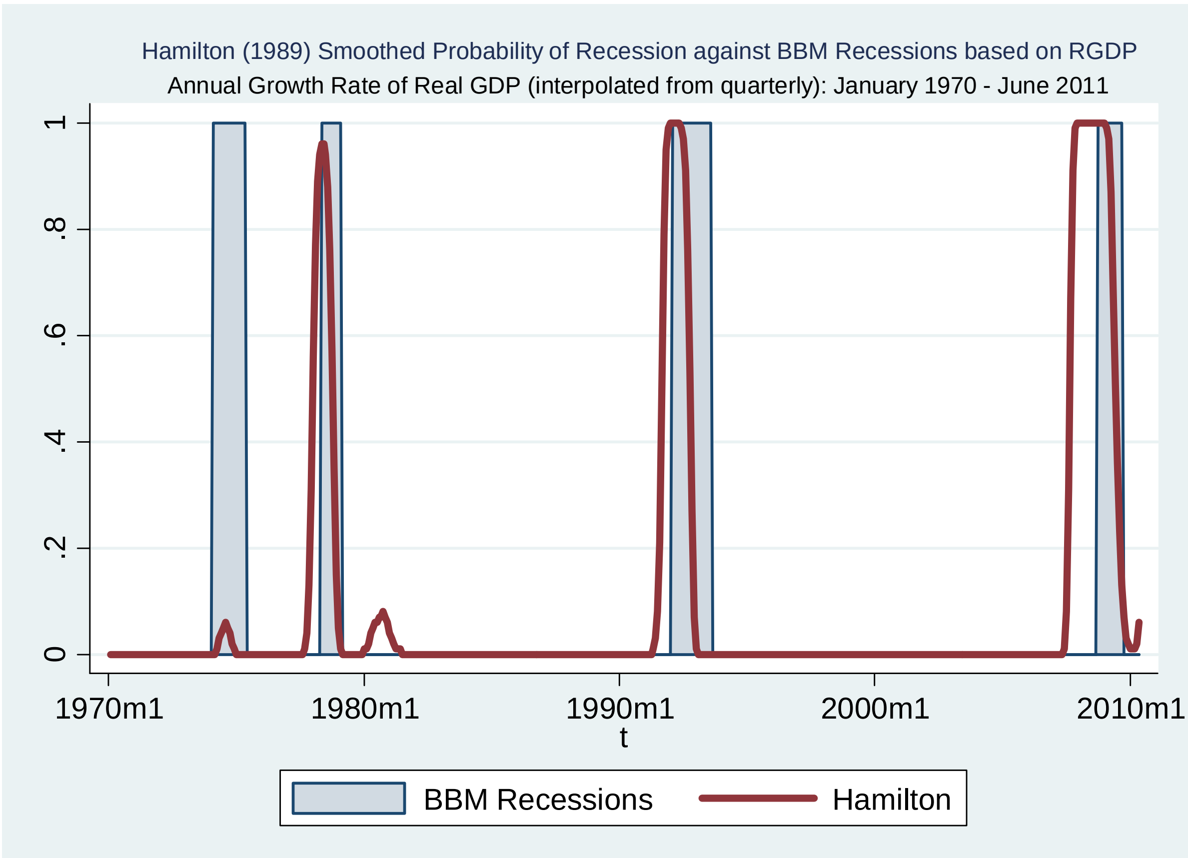
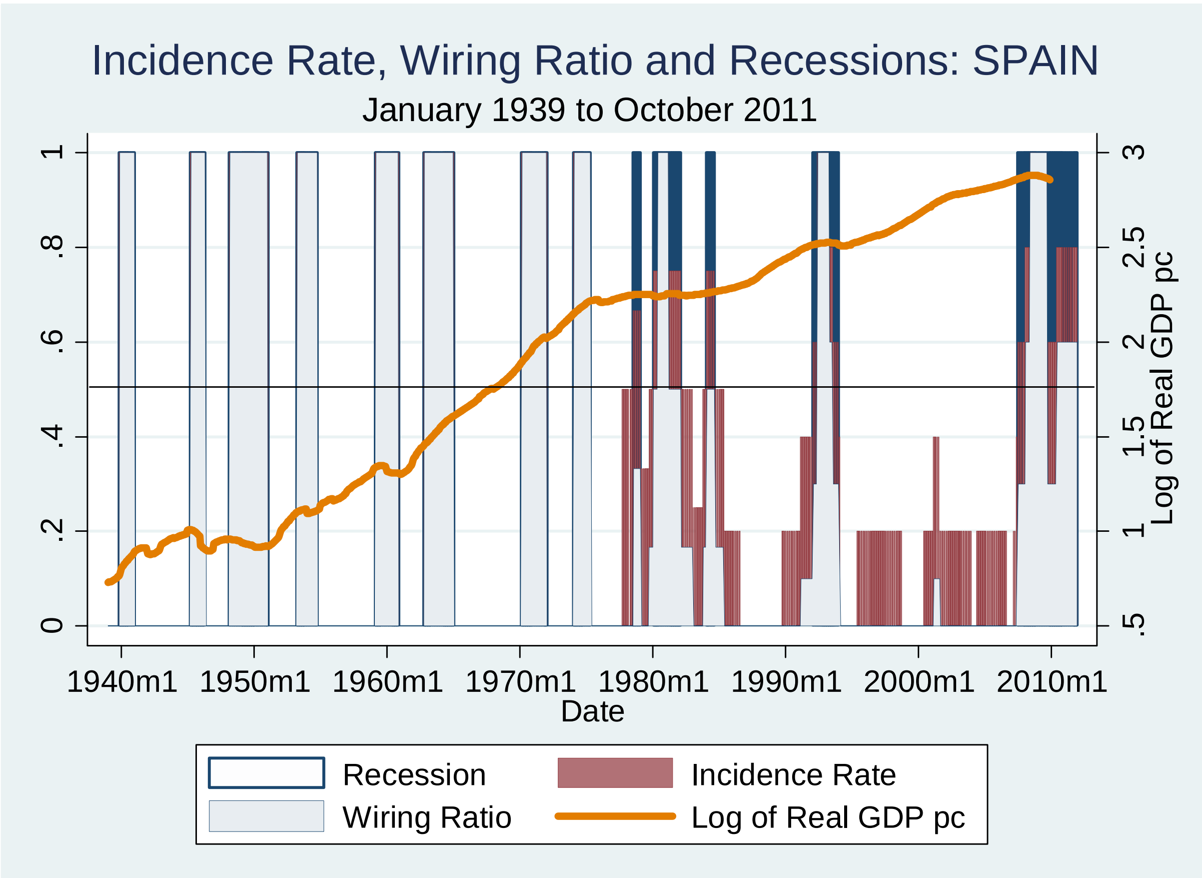
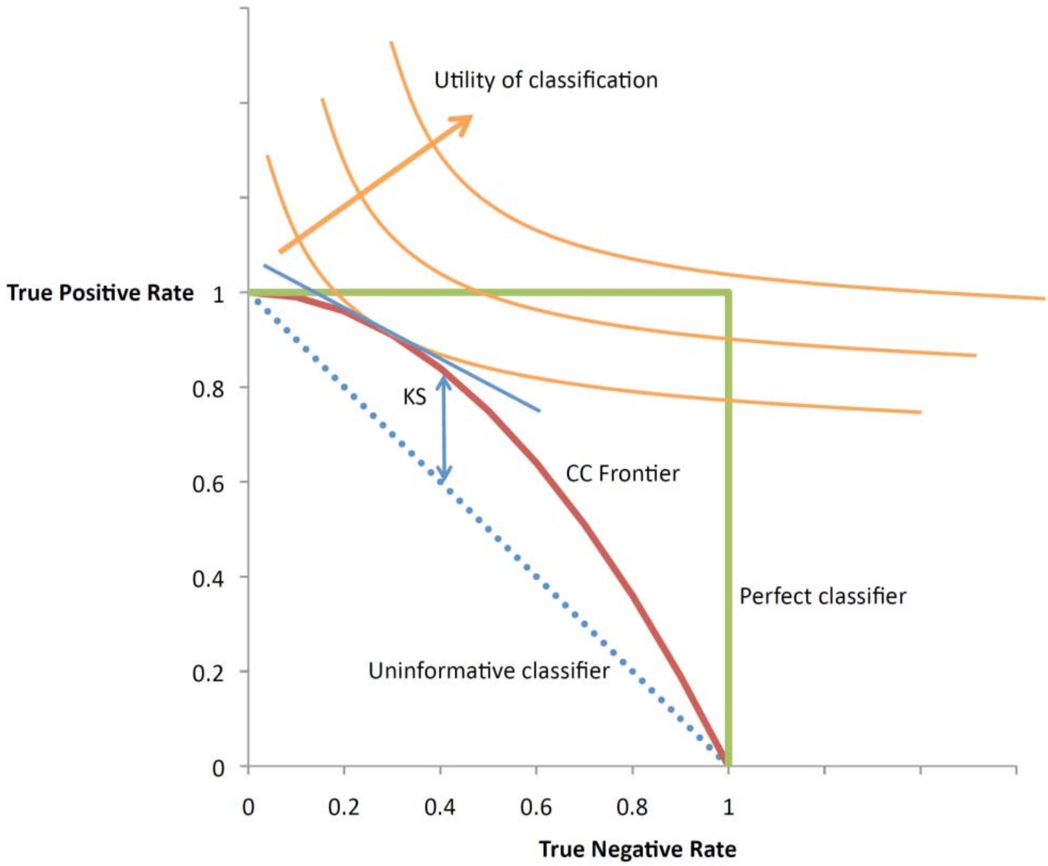


Figure 6. Incidence Rate, Wiring Ratio and Recessions. Spain January 1939 to October 2011



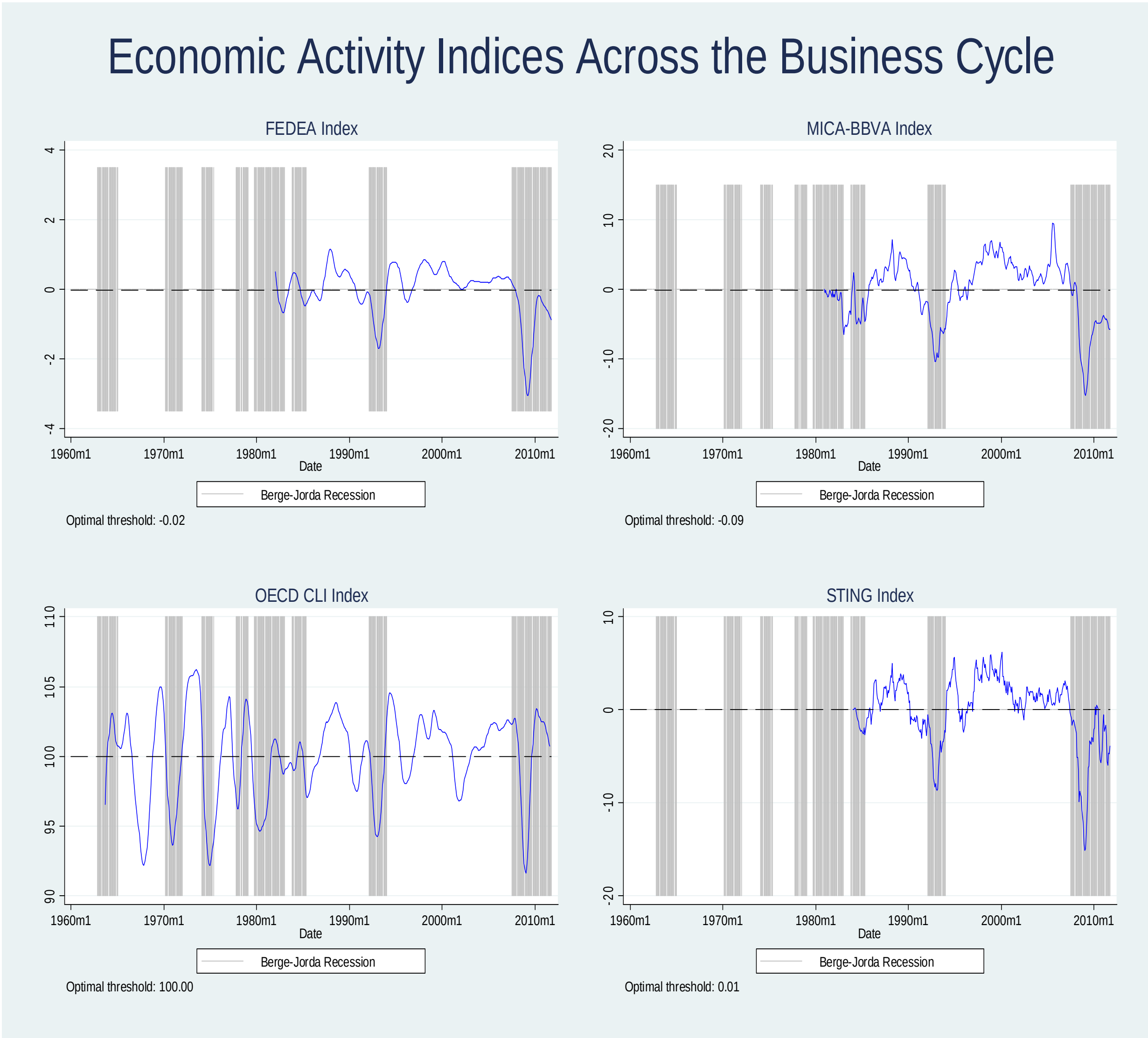
Notes: Recession shading based on a 0.5 threshold for the incidence rate (displayed). Real GDP per capita interpolated from yearly frequency observations (Prados de la Escosura, 2003). Early recessions coincide with the wiring ratio because there is no other data available. More series become available starting in the early 1970's. See text for details.

Figure 7. The Correct Classification Frontier



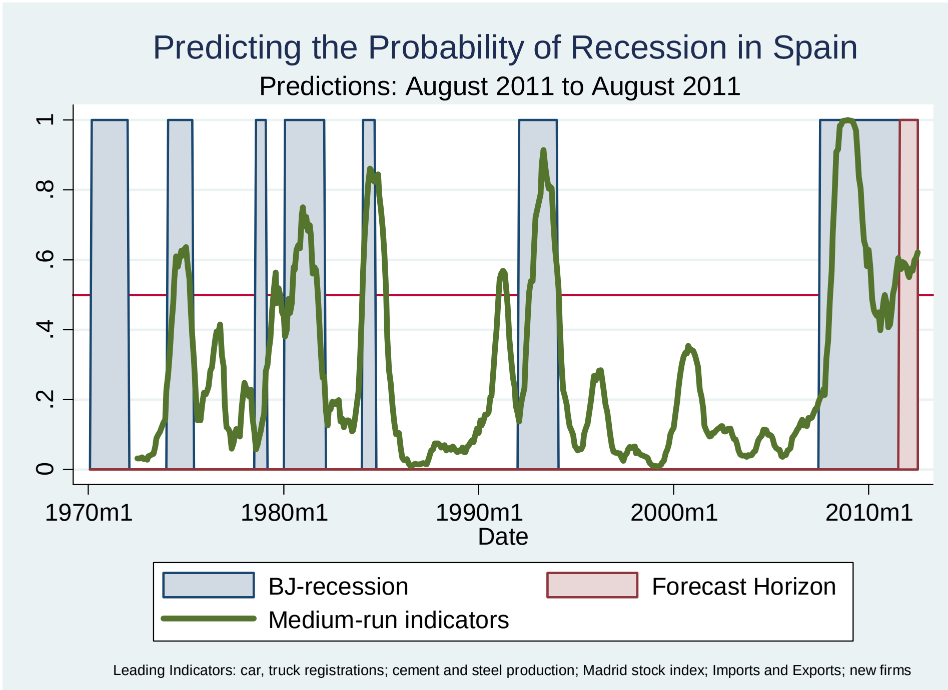
Notes: See text for details.

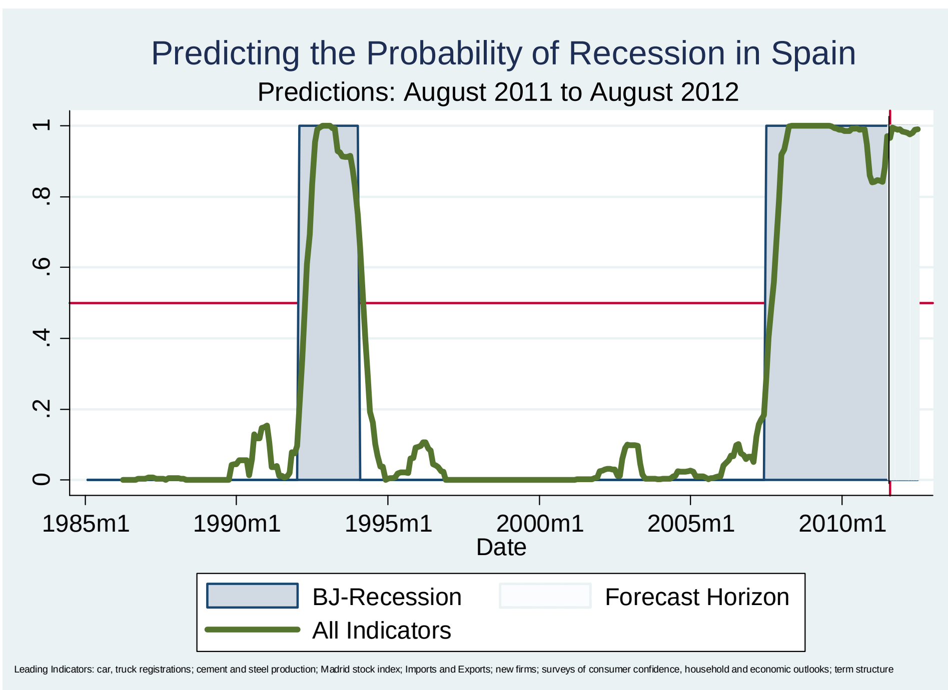
Figure 8. Four Economic Activity Indexes



Notes: Recessions shaded using table 5. We report the optimal thresholds for each index that would allow one to determine whether the economy is in expansion or recession using equal weights. See text for more details.

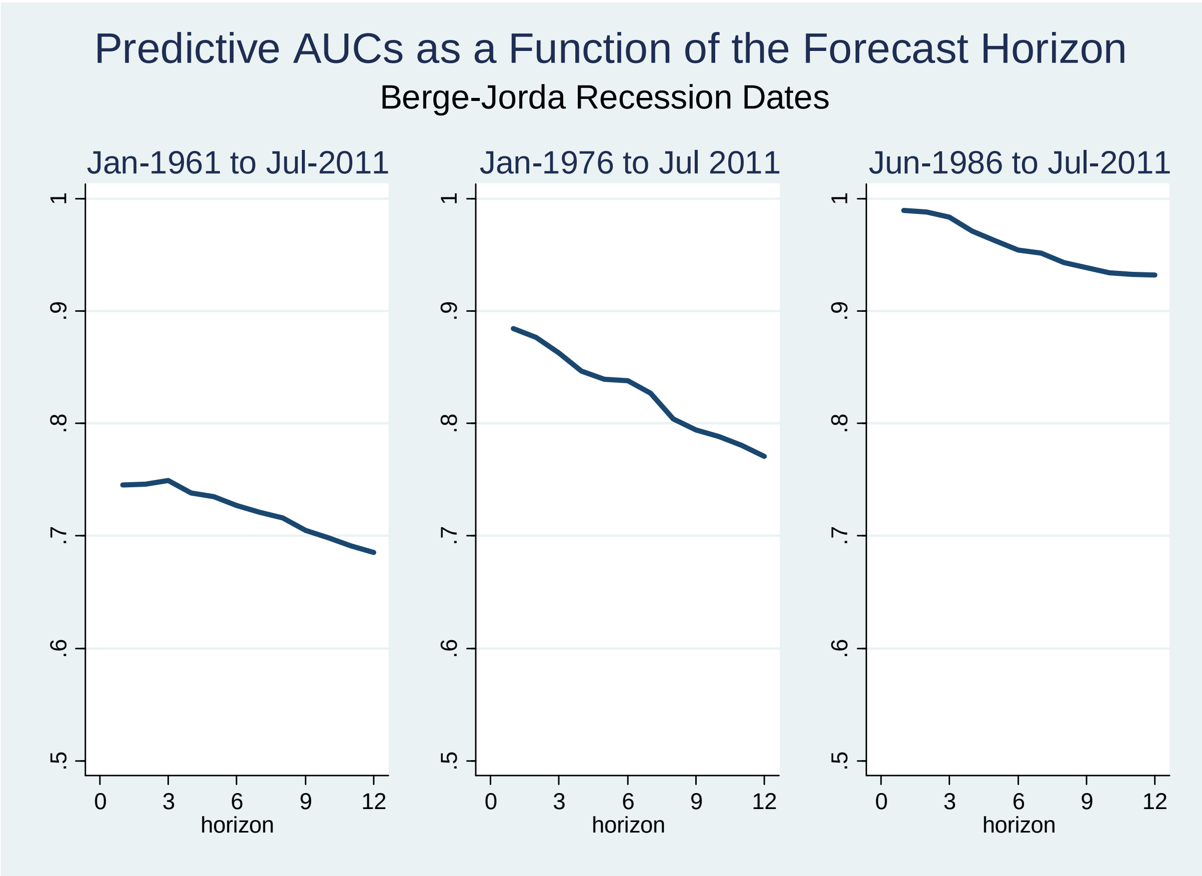
Figure 9. Predicting the Probability of Recession. Berge and Jordà (2011) vs. OECD Recessions





Notes: The figures reports in-sample one period ahead probability forecasts and out-of-sample forecasts up to August 2012. Recessions shaded are those reported in table 5. Top panel uses all but survey and term structure data (for a longer sample); bottom panel uses all indicators (available over a shorter sample).

Figure 10. Classification Ability of the Predictive Models. Berge and Jordà (2011) vs. OECD Recessions



Notes: Areas Under the correct classification Curve (AUC) per forecast horizon (from 1 to 12 months), depending on the length of the sample available to fit the predictive model. For the first sample (January 1961 to July 2011) the indicators considered are: car registrations, cement production, and Madrid's stock market. For the second sample (January 1976 to July 2011), we add to the previous indicators: imports and exports, new registered firms, steel production and truck registration. The last sample (June 1986 to July 2011) includes in addition: consumer confidence survey data, economic outlook survey data, and household outlook survey data.