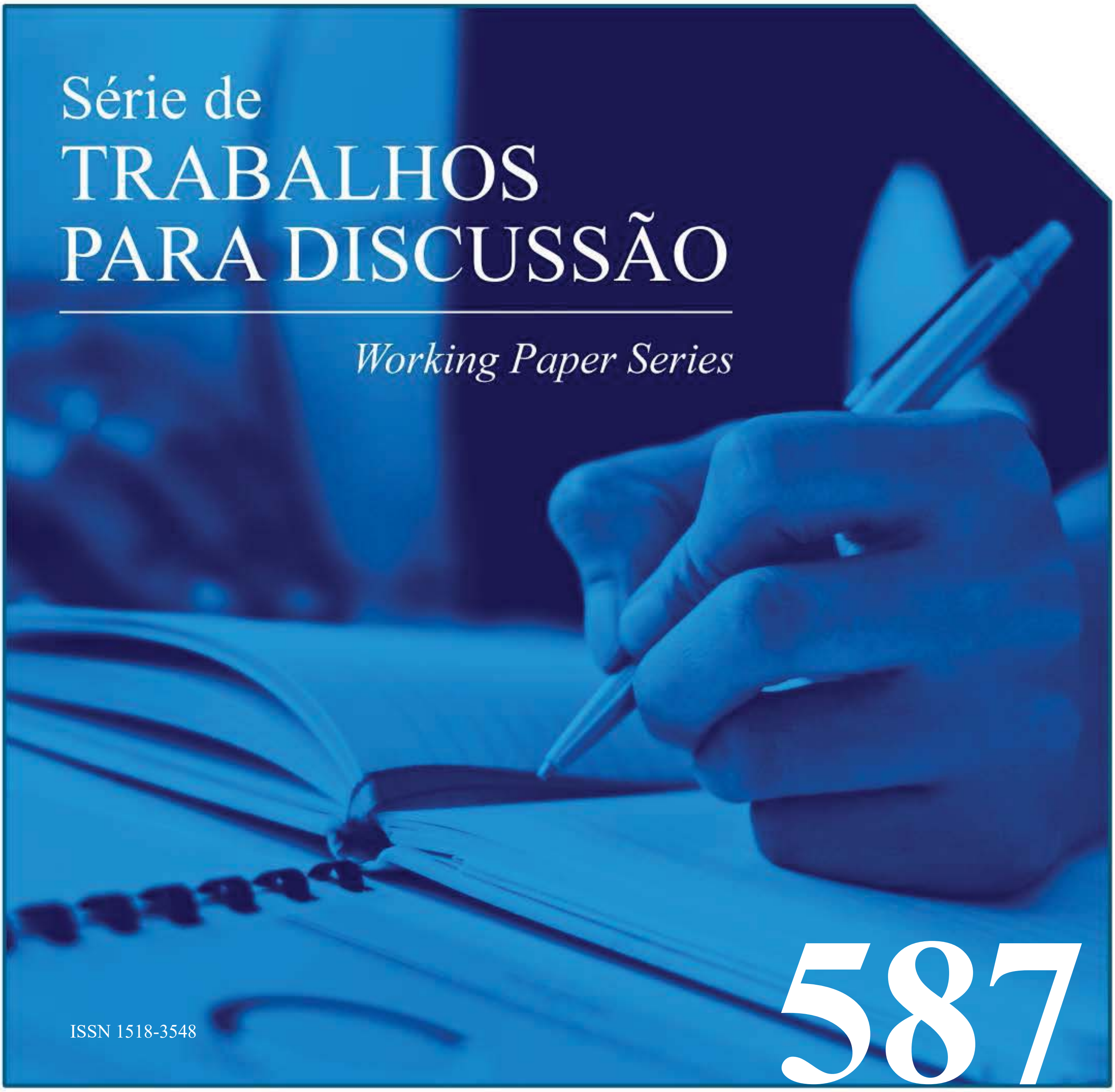


Novembro 2023

Predicting Recessions in (almost) Real Time in a Big-data Setting

Alexandre Bonnet R. Costa, Pedro Cavalcanti G. Ferreira, Wagner Piazza Gaglianone, Osmani Teixeira C. Guillén, João Victor Issler, Artur Brasil Fialho Rodrigues

BANCO CENTRAL DO BRASIL

ISSN 1518-3548
CGC 00.038.166/0001-05

Working Paper Series	Brasília	no. 587	Novembro	2023	p. 3-51

Working Paper Series

Edited by the Research Department (Depep) – E-mail: workingpaper@bcb.gov.br

Editor: Rodrigo Barbone Gonzalez

Co-editor: Eurilton Alves Araujo Jr

Head of the Research Department: André Minella

Deputy Governor for Economic Policy: Diogo Abry Guillen

The Banco Central do Brasil Working Papers are evaluated in double-blind referee process.

Citizen Service Division

Banco Central do Brasil

Deati/Diate

SBS – Quadra 3 – Bloco B – Edifício-Sede – 2º subsolo

70074-900 Brasília – DF – Brazil

Toll Free: 0800 9792345

Fax: +55 (61) 3414-2553

Internet: http//www.bcb.gov.br/?CONTACTUS

Non-technical Summary

In this paper, we explore the specialized literature on economic recession dating models. Our objective is to conduct a comprehensive review of this literature to identify the most effective methods for developing new statistical tools capable of tracking the beginning of economic recessions in real time, or slightly a posteriori.

The basic method proposed by Stock and Watson (1989, 2002, and 2014, among others) for modeling and dating business cycles and recessions assumes the existence of an unobservable factors model that determines the cyclical properties of economic stationary series. Such factors can be identified using principal component analysis or state space models estimated by Kalman filter.

In turn, the canonical correlation approach proposed by Issler and Vahid (2006) follows the *common features* literature and models the probability of occurring a recession using a structural Probit model, which assumes as dependent variable the dates of U.S. recessions, according to the National Bureau of Economic Research (NBER), and as explanatory variables the common cycles of U.S. economic activity.

In this work, we develop a hybrid approach, incorporating the good elements of each technique into a single setup. In other words, we propose a methodology for dating recessions in real time based on the canonical correlation approach of Issler and Vahid (2006), but also using *big data* as defended by Stock and Watson (2014). This novel methodology involves solving the problems of missing data and high dimensionality of the databases, besides establishing a decision rule to choose in real time the best prediction model.

Our empirical results show it is possible to track the state of the economy using the estimated models, as long as appropriate techniques to reduce the dimensionality of the databases are implemented. Depending on the cutoffs chosen, the models predict recessions in real time with an accuracy of 98% and 80%, respectively, for the U.S. and the Euro Area.

Sumário Não Técnico

Neste artigo, exploramos a literatura especializada sobre modelos de datação de recessões na economia. O objetivo é fazer uma revisão crítica dessa literatura para identificar os melhores caminhos a seguir na construção de novas ferramentas estatísticas que possam indicar o início de uma recessão da economia em tempo real, ou levemente *a posteriori*.

O método básico proposto por Stock e Watson (1989, 2002, 2014, dentre outros) para modelagem e datação de ciclos de negócios e recessões assume a existência de um modelo de fatores não observáveis que determina as propriedades cíclicas das séries econômicas estacionárias. Tais fatores podem ser identificados a partir do uso da análise de componentes principais ou de modelos em estado de espaços estimados por filtro de Kalman.

Por sua vez, a abordagem de correlações canônicas proposta por Issler e Vahid (2006) segue a literatura de *common features* e modela a probabilidade de ocorrer uma recessão na economia por meio de um modelo Probit estrutural, que tem como variável dependente as datações de recessões nos EUA, segundo o *National Bureau of Economic Research* (NBER), e como variáveis explicativas os ciclos comuns da atividade econômica norte-americana.

Neste trabalho, desenvolvemos uma abordagem híbrida, incorporando os bons elementos de cada técnica num único arcabouço. Em outras palavras, propomos uma metodologia para datar recessões em tempo real utilizando a abordagem de correlações canônicas, seguindo Issler e Vahid (2006), junto com a ideia de *big data* defendida por Stock e Watson (2014). Esta nova metodologia envolve resolver os problemas de dados faltantes e a alta dimensionalidade das bases de dados utilizadas, além de estabelecer uma regra de decisão para escolher o melhor modelo de previsão em tempo real.

Nossos resultados empíricos mostram ser possível acompanhar o estado da economia usando os modelos estimados, desde que sejam implementadas técnicas apropriadas de redução de dimensionalidade das bases de dados. Dependendo dos limiares de corte escolhidos, é possível acompanhar o estado da economia prevendo recessões em tempo real para os EUA e a Europa, respectivamente, com 98% e 80% de acurácia.

Predicting Recessions in (almost) Real Time in a Big-data Setting*

Alexandre Bonnet R. Costa† Pedro Cavalcanti G. Ferreira‡
Wagner Piazza Gaglianone§ Osmani Teixeira C. Guillén¶
João Victor Issler‖ Artur Brasil Fialho Rodrigues**

Abstract

The objective of this paper is to propose an approach for dating recessions in real time (or slightly *a posteriori*) that is suitable to a *big data* environment. Our proposal is to mix the canonical correlation approach of Issler and Vahid (2006) with the *big data* approach defended by Stock and Watson (2014). We incorporate the good elements of each approach into one. This involves solving both the problem of missing data and high dimensionality in big databases, besides defining a decision rule on how to choose the best forecasting model in real time. Our empirical results show it is possible to track the state of the U.S. and European economies using the models developed here, as long as appropriate techniques to reduce the dimensionality of the databases are implemented – canonical correlations coupled with principal component analysis. Depending on the cutoffs chosen, the models predict recessions in real time with an accuracy of 98% and 80%, respectively, for the U.S. and the Euro Area.

Keywords: Forecasting; Recessions; Big Data; Machine Learning.

JEL Classification: C14; C15; C22; C53; C55; E17; E32.

*We are grateful to Rodrigo Gonzalez, Diogo Guillen, André Minella, João Maurício Moreira, and participants at various workshops and seminars for excellent comments and suggestions. We also gratefully acknowledge the support from CNPq and INCT through different grants. The views expressed in the paper are those of the authors and do not necessarily reflect those of the Banco Central do Brasil, Getulio Vargas Foundation or Petrobras.

†Corporate Risks, Petrobras. E-mail: abrcosta@petrobras.com.br

‡Graduate School of Economics, Getulio Vargas Foundation. E-mail: Pedro.Ferreira@fgv.br

§Research Department, Banco Central do Brasil. E-mail: wagner.gaglianone@bcb.gov.br

¶Open Market Operations Department, Banco Central do Brasil, and Ibmec/RJ. E-mail: osmani.guillen@bcb.gov.br

‖Corresponding Author. Graduate School of Economics, Getulio Vargas Foundation, Praia de Botafogo 190, s.1104, Rio de Janeiro, RJ 22.253-900, Brazil. E-mail: Joao.Issler@fgv.br

**Graduate School of Economics, Getulio Vargas Foundation. E-mail: artur.bfrodrigues@gmail.com

1 Introduction

The attempt to predict the probability of recessions is far from new in the literature. A probabilistic setup was used by Stock and Watson (1989)[1] to build a coincident and leading index of economic activity, as well as an index of recessions. The individual performance of financial variables (interest rates, spreads, stock prices and monetary aggregates) was used by Estrella and Mishkin (1998) to predict the probability of a recession. They found stock prices are good predictors of recessions over the one- to three-quarter horizon, while the slope of the yield curve is a better predictor beyond one quarter.[2]

A dynamic Probit model[3] was used by Nyberg (2010), who found that, in addition to the term spread,[4] lagged stock return values and external spreads are important predictors of a recession. Nonlinear models were used by Anderson and Vahid (1998) to predict the probability of recession in the U.S. using the interest rate spread and the growth of the money stock (M2). Several Probit models were estimated by Wright (2006), who found that adding the Fed Funds rate to the term spread outperforms the Estrella and Mishkin (1998) model. Christiansen, Eriksen and Miller (2013) found sentiment variables have superior predictive power than financial variables.

According to Stock and Watson (2014), there are two main approaches in the business cycle dating literature. The first approach, initiated with Burns and Mitchell (1946), consists of identifying individual inflection points in a large number of time series, and then looking for a common date (aggregate inflection point). Stock and Watson call this approach *date then average*. The second, more recent approach (*average then date*) looks for inflection points in just a few, or just one aggregate time series (e.g., GDP).

In turn, Hamilton (2011) presents a survey of the existing literature and seeks to answer whether there is any effective technique for dating recessions in *real time*.[5] Great importance is given to the approach of Chauvet and Hamilton (2006), which jointly uses the U.S. GDP

[1] Stock and Watson are pioneers in the modern field of econometrics for modeling and dating business cycles and recessions; see Stock and Watson (1988a,b, 1989, 1991, 1993a,b, 2002, 2010, 2014). Its basic method assumes the existence of a model of unobservable factors that determines the cyclical properties of stationary economic series. These can be identified using principal component analysis (PCA) or using state-space models estimated by the Kalman filter. In dating recessions, one can couple this factor model to a Probit model to predict the probability of a recession; see Stock and Watson (1988b).

[2] The predictive power of the term structure is also documented in Rudebusch and Williams (2009), who emphasize the fact that professional analysts do not adequately incorporate the information contained in the yield spread.

[3] Hao and Ng (2011) found dynamic Probit models improve the static Probit, especially when predicting the duration of recessions. This result is expected for the short term in particular and around inflection points. However, the dynamic characteristic of these models makes them unsuitable for a real-time forecasting exercise. A static Probit that uses financial forecasters more frequently does not suffer from this deficiency.

[4] Regarding the use of *term spread*, the multi-period asset pricing literature establishes the relationship between (asset return) spreads of different maturities and the growth of future consumption (and, consequently, of economic activity). See Campbell (2003), and also Issler and Pimentel (2019), who fully derive this relationship.

[5] For real-time turning point prediction, there is a difference between predicting the probability of a recession and predicting the beginning and end of a recession. The last exercise is a little more difficult as it requires decision rules in addition to a probabilistic model. For example, considering a quarterly frequency, the first release of U.S. GDP is available with a lag of one quarter, while the "final" figure is released with a lag of approximately one year. The NBER turning points are reported with at least four lags. These aspects can be ignored if we are only interested in evaluating the in-sample performance of Probit models based on historical data. However, for real-time analysis it is necessary to build a proper strategy to deal with information delays.

and the NBER dates of U.S. recessions. Part of the literature cited there uses the techniques proposed by Hamilton (1989), in a seminal article on models of two (or more) regimes, with a latent state variable representing the regimes of the economy, driven by a Markov Chain.[6] Part of this literature uses common unobservable factors, which can be identified using principal component analysis, e.g., Stock and Watson (1989, 1991, 2002).

On the other hand, the approach of Issler and Vahid (2006) for dating recessions follows the *common features* literature, initiated by Engle and Kozicki (1993), Vahid and Engle (1993) and Engle and Issler (1995). The basic idea is that several economic series have common components that can be removed by linear combination. These common components can be trends, cycles, seasonality, volatility, etc., representing different aspects of the economic series. The best known examples are those of common trends (cointegration) and common cycles (common serial correlation). Issler and Vahid use these techniques to isolate the business cycles of the U.S. economy and model the probabilities of recession using a *Probit* model, which has the dates of recessions according to the NBER as the explained variable and the U.S. economic cycles as the explanatory variables.

More recently, the prediction models used by Kotchoni and Stevanovic (2018) take the form of an AR(1) model that is increased with a recession probability or Inverse Mills Ratio. The authors propose a model capable of producing inflection forecasts of business cycle at various horizons, average forecasts of economic activity and conditional forecasts that depend on whether the horizon of interest belongs to a recession episode or not.[7]

The objective of this paper is to propose a canonical correlation approach that is suitable to *big data* for dating recessions in real-time or slightly *a posteriori*. Our proposal is to mix the approach of Issler and Vahid (2006) with the idea of *big data* defended by Stock and Watson (2014). The results in Issler and Vahid (2006), and their comparisons with previous methods by Stock and Watson, certify this method to be the baseline methodology to be used in a context of few series used for dating recessions.

However, since in the last 10-15 years the *big data* approach has gained strength, as demonstrated by Stock and Watson (2014), we decided to incorporate the good elements of each approach into a single setup. This involves solving the three practical problems: (i) how to complete the missing data in big databases; (ii) how to properly reduce the dimensionality of such data in order to be able to estimate the Probit structural model proposed by Issler and Vahid, since we will have more series than observations in time; and (iii) how to choose the best forecasting models in *real time*.

[6] The Markov Switching (MS) model (Hamilton, 1989) is an alternative way to predict economic activity conditional on the state of the economy.

[7] The objective is not exactly to predict recessions in isolation, but to use the fact that the predictions of certain economic variables depend on the state of the economy, i.e., whether or not the economy is in a recession in a given forecast horizon. Thus, forecasting the state of the economy in real time is one of the intermediate goals of this paper. The empirical results suggest that a static Probit model that uses only the term spread as a regressor already provides a good fit.

Our empirical results show it is possible to track the state of the economy in the U.S. and Europe using the models developed here. Overall, models show a good in-sample fit and the out-of-sample recession predictions are quite good (and, thus, compatible with our goals of dating recessions in real time, or slightly *a posteriori*). For the U.S., depending on the cutoffs chosen, it is possible to predict recessions (expansions) in real time with 98.39% (100%) accuracy. For Europe, in this same context, we can reach 79.17% and 96.36% accuracy for recessions and expansions, respectively.

The outline of the paper is as follows. In Section 2, we present our methodology to predict recessions in real time using big data. Section 3 shows the empirical results of recession/expansion predictions for the U.S. and the Euro Area. Section 4 concludes.

2 Methodology

2.1 Canonical correlation

The canonical correlation approach of Issler and Vahid (2006) follows the common features literature, initiated by Engle and Kozicki (1993), Vahid and Engle (1993), Engle and Issler (1995), Vahid and Engle (1997), and Issler and Vahid (2001). The main idea is that several economic series share common components that can be removed by linear combination. These common components can be trends, cycles, seasonality, volatility, etc., representing different aspects of the economic series. The best known examples are those of common trends (cointegration), Engle and Granger (1987), and common cycles (common serial correlation); see Vahid and Engle (1993), and Engle and Issler (1995).

Issler and Vahid begin their article with the following example: suppose we are asked to construct an index of a patient's health status. Furthermore, suppose we know the best indicator of a patient's health is the results of a blood test. However, blood samples cannot be collected very often and test results are only available with a delay, sometimes too long to be useful in treatment. Our index, therefore, must be a function of variables such as blood pressure, heart rate, and body temperature, which are readily available at regular frequencies. To best combine these variables into an index, should we: (i) only use historical data on blood pressure, heart rate, and body temperature, or (ii) use historical blood test results as well? The answer is, of course, alternative (ii).

Since Burns and Mitchell (1946), there has been a great interest in making inferences about the "state of the economy" from sets of monthly variables, believed to be coincident with or antecedent to the business cycles of the economy (so-called "coincident" and "leading" indicators, respectively). Although the business cycle status of the economy is not directly observable, our best-informed estimate of its turning (inflection) points is embedded in the dummy variable announced by several Business Cycle Committees, the best known of which is that of the NBER – National Bureau of Economic Research, USA. Nowadays, there are several other similar committees. For example, in Europe, recessions are dated by the Euro

Area Business Cycle Dating Committee (EABCDC), made up of five researchers from CEPR – Center for Economic Policy Research; in Brazil, recessions are dated by CODACE – Economic Cycles Dating Committee; and several developing countries are studying the implementation or are implementing this technology, such as Mexico and India, for example.

In particular, the NBER announcements[8] are based on the consensus of a panel of experts and are made some time after (usually six months to a year) the moment a turning point occurs in the business cycle: if a recession starts in month t, the NBER committee dates the start of the recession to month $t + h$.

The article by Issler and Vahid (2006) aims at identifying coincident and leading indicators of economic activity based on the decisions on the state of the economy made by the NBER, which dates the periods of recession and expansion of the U.S. economy. Nonetheless, the methodology can be applied to other countries or regions, as long as they have a dating committee and consistent series and antecedents of economic activity. In the U.S., coincident and leading series are maintained by *The Conference Board* (see The Conference Board, 1997).

The authors impose the restriction that the coincident index is a linear combination of the cyclic components of the coincident variables, all observable. This means the "business cycle" is a linear combination of the cycles of the four coincident series (output, income, employment and sales), which may have more than one cyclical factor, but are, above all, linear functions of the four coincident series. This contrasts with the coincident (single) latent dynamic index view (e.g., Stock and Watson, 1989, and Chauvet, 1998), which constrains the "business cycle" to be a *single* common cyclical factor shared by the coincident variables. To identify the common cycle, the single latent dynamic factor approach must allow the coincident variables to have other idiosyncratic cyclic factors, and this does not provide any control over the intensity of these idiosyncratic cycles relative to the common cycle.

Issler and Vahid define as "cyclic" any variable that can be predicted linearly from the set of past information. This set of leading information includes lags from both sets of coincident and leading variables. The inclusion of lags of leading variables, in addition to lags of coincident variables in the information set, serves two purposes. First, it combines the estimation of coincident indices and leading indicators. Second, it allows for the possibility of asymmetric cycles in coincident series by including lags of variables such as interest rates and the spread between interest rates, which are known as nonlinear processes (Anderson, 1984, Balke and Fomby, 1997), and also as exogenous predictors. There are infinite linear combinations of coincident variables that are predictable in the past, that is, that are cyclic. The authors use canonical correlation analysis to find a basis for the space of these cycles.

Canonical correlation analysis, introduced by Hotelling (1935, 1936), has long been used in multivariate statistics. It was first used in the analysis of multivariate time series by Akaike

[8] The NBER summarizes its deliberations as follows: *"The NBER does not define a recession in terms of two consecutive quarters of decline in real GDP. Rather, a recession is a recurring period of decline in output, income, employment and sales, typically lasting from six months to a year, and marked by widespread contractions in many sectors of the economy."* For further details, see http://www.nber.org/cycles.html

(1976). Akaike properly referred to canonical variables as "the information interface channels between past and present" and referred to canonical correlations as the "strength" of these channels.

Denote the set of coincident variables (income, output, employment and sales) by the vector $x_t = (x_{1t}, x_{2t}, x_{3t}, x_{4t})'$ and the set of m ($m \geq 4$) "predictors" by vector z_t (this includes lags of x_t and the lags of the leading series). Canonical correlation analysis transforms x_t in four independent linear combinations $A(x_t) = (\alpha_1' x_t, \alpha_2' x_t, \alpha_3' x_t, \alpha_4' x_t)$, with the property that $\alpha_1' x_t$ is the linear combination of x_t which is more (linearly) predictable using z_t, $\alpha_2' x_t$ is the second most predictable linear combination of x_t which is more (linearly) predictable using z_t, after controlling for $\alpha_1' x_t$, and so on.

The fact that canonical correlation analysis studies linear dependence channels between x_t and z_t does not necessarily imply it will only be useful for linear multivariate analysis. By including variables non-linearly modeled in z_t (e.g., Fourier series, Tchebyschev polynomials), you can use canonical correlation analysis for nonlinear multivariate modeling; see Anderson and Vahid (1998) for an example and additional references.

These linear combinations are not correlated with each other and are constrained to have unity variance. The by-products of this analysis are four linear combinations of z_t, $\Gamma(z_t) = (\gamma_1' z_t, \gamma_2' z_t, \gamma_3' z_t, \gamma_4' z_t)$, with the property that $\gamma_i' z_t$ is the linear combination of z_t which has the highest quadratic correlation with $\alpha_i' x_t$, for $i = 1, 2, 3, 4$. Again, the elements of $\Gamma(z_t)$ are not correlated with each other, and are uniquely identified, up to a sign change, with the additional constraint that all four have unity variances. R^2s between $\alpha_i' x_t$ and $\gamma_i' z_t$ for $i = 1, 2, 3, 4$, denoted by $\left(\lambda_1^2, \lambda_2^2, \lambda_3^2, \lambda_4^2\right)$, are the squares of the canonical correlations between x_t and z_t.

One might think that, in order to estimate the weights associated with each basic cycle, it is sufficient to estimate a simple *Probit* model with the NBER indicator as the binary (dummy) dependent variable (1 for U.S. recessions and 0 for non-recessive periods) and the cycles associated with non-zero canonical correlations as explanatory variables. As the basic cycles are linear combinations of the four coincident series, we will end up explaining the NBER indicator by a linear combination of the coincident series.

However, it is important to note that the coincident index we are looking for is a linear combination of the coincident series that has cyclical characteristics similar to the unobserved state of the economy. Using the technical terms introduced in Engle and Kozicki (1993), we are assuming that the (unobserved) business cycle state of the economy and the coincident variables share a serial correlation common feature, and we want to estimate the associated cofeature vector (co-characteristic) to this common feature.

The NBER recession indicator is important because it contains information about the unobserved state of the economy's business cycle. As will become clear below, the linear combination of the coincident series that has a serial correlation pattern similar to that of the unobserved state of the economy is neither the conditional expectation of the NBER recession indicator, given the previous information set, nor the conditional expectation of the

NBER indicator given the coincident series.

Hypothesis 1: *There is a linear index (of the cyclical parts) of the coincident series that has exactly the same pattern of correlation with previous information as the unobserved state of the economy.*

Note that we put "the cyclic parts of" in parentheses because it's redundant. Although the index that has the same pattern of correlation with the past will only involve the basic cycles (i.e., it will not involve combinations of white noise from the coincident series), these basic cycles are themselves linear combinations of the coincident series. Therefore, the index will end up being a linear combination of coincident series.

Let y_t^* be a binary variable representing the unobserved state of the economy and $\{c_{1t}, c_{2t}, c_{3t}\}$ the significant basic cycles of the coincident series in t, advancing the empirical result that only three cycles from the four coincident series were in fact detected. Hypothesis 1 clearly implies there must be a linear combination of y_t^* and $\{c_{1t}, c_{2t}, c_{3t}\}$ that is unpredictable from information ahead of time t, i.e., up to $t-1$. This way,

$$\mathbb{E}(y_t^* - \beta_0 - \beta_1 c_{1t} - \beta_2 c_{2t} - \beta_3 c_{3t} \mid I_{t-1}) = 0, \tag{1}$$

where I_{t-1} is the information available in time $t-1$. If y_t^* were observed, we could estimate β_1, β_2 and β_3 directly by GMM or by maximum likelihood with limited information.

However, the NBER indicator is equal to 1 when, with all the information accumulated by the NBER Committee at the time $t+h$, the economy was in recession at the time t. That is, when the "smoothed" estimate of the unobserved state of the economy based on information up to $t+h$ is below a critical value:

$$\text{NBER}_t = \begin{cases} 1 & \text{if } \mathbb{E}(y_t^* \mid I_{t+h}) < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

This critical value cannot be identified separately from the constant term in equation (1). Therefore, without loss of generality, we assume this critical value is zero: in other words, we let the threshold value be absorbed into the constant term β_0 in (1).

Using equation (1), it follows that:

$$\begin{aligned} \mathbb{E}\left(y_t^* \mid I_{t-1}\right) &= \beta_0 + \beta_1 \mathbb{E}\left(c_{1t} \mid I_{t-1}\right) + \beta_2 \mathbb{E}\left(c_{2t} \mid I_{t-1}\right) + \beta_3 \mathbb{E}\left(c_{3t} \mid I_{t-1}\right) \\ &= \beta_0 + \beta_1 c_{1t} + \beta_2 c_{2t} + \beta_3 c_{3t} + \omega_t, \text{ where } \mathbb{E}\left(\omega_t \mid I_{t-1}\right) = 0, \end{aligned} \tag{3}$$

and obviously ω_t is correlated with c_{it}, $i = 1, 2, 3$. One can always write:

$$\mathbb{E}\left(y_t^* \mid I_{t+h}\right) = \mathbb{E}\left(y_t^* \mid I_{t-1}\right) + \xi_t + \xi_{t+1} \cdots + \xi_{t+h}, \tag{4}$$

where ξ_{t+i} is the "surprise" associated with the new information that arrives in period $t+i$.

Thus, we can show that:

$$\begin{aligned} \mathbb{E}\left(y_t^* \mid I_{t+h}\right) &= \beta_0 + \beta_1 c_{1t} + \beta_2 c_{2t} + \beta_3 c_{3t} + u_t, \\ u_t &= \omega_t + \xi_t + \xi_{t+1} \cdots + \xi_{t+h}, \end{aligned} \tag{5}$$

where u_t is unpredictable given information in period $t-1$, that is, $\mathbb{E}\left(u_t \mid I_{t-1}\right) = 0$, but it has a $MA(h)$ structure, being correlated with c_{it}, $i = 1, 2, 3$, mainly due to the term ω_t.

To consistently estimate β_1, β_2 and β_3 we must use a method with a single structural equation and a bounded dependent variable. All these methods use instrumental variables. In the present case, the obvious instrumental variables would be the variables z_t, that is, lags of the coincident and leading variables. Note that canonical correlation analysis yields estimates of $\gamma_1' z_t$, $\gamma_2' z_t$, $\gamma_3' z_t$ and $\gamma_4' z_t$, which are the best linear predictors for each of the basic cycles, respectively.

Several alternative estimators have been proposed for the consistent estimation of parameters from a single equation with a bounded dependent variable in a simultaneous equations model. These estimators differ in their ease of calculation versus their degree of efficiency. We use the two-stage conditional maximum likelihood estimator (2SCML) proposed by Rivers and Vuong (1988) because of its relative simplicity.

Using the empirical results presented in next section, we assume the four coincident series can be explained by three significant base cycles $\{c_{1t}, c_{2t}, c_{3t}\}$. Denoting the NBER indicator by NBER_t, the first stage of the 2SCML estimation procedure involves regressing $\{c_{1t}, c_{2t}, c_{3t}\}$ onto instruments z_t and save the residuals, which we denote by $\{\hat{v}_{1t}, \hat{v}_{2t}, \hat{v}_{3t}\}$. In the second stage, both basic cycles $\{c_{1t}, c_{2t}, c_{3t}\}$ and the residuals of the first stage $\{\hat{v}_{1t}, \hat{v}_{2t}, \hat{v}_{3t}\}$ are included in the *Probit* model:

$$\Pr\left(\text{NBER}_t = 1\right) = \Phi\left(-\left(\beta_0 + \beta_1 c_{1t} + \beta_2 c_{2t} + \beta_3 c_{3t} + \beta_4 \hat{v}_{1t} + \beta_5 \hat{v}_{2t} + \beta_6 \hat{v}_{3t}\right)\right), \tag{6}$$

where $\Phi(\cdot)$ is the standard Normal cumulative distribution function (cdf). The estimates of β_1, β_2, and β_3 of the second stage of the *Probit* model will be the estimates of 2SCML. The standard errors of the estimated parameters must be modified according to the procedure in Rivers and Vuong, p. 354. Furthermore, as we ignore the dynamic structure of u_t in constructing the likelihood function (i.e., the model is "dynamically incomplete" in the sense of Wooldridge, 1994), autocorrelation-robust standard errors must be used.

The Issler and Vahid (2006) coincident index, denoted by "instrumental variable coincident index" (IVCI), is then given by:

$$\begin{aligned} \Delta IVCI_t &= \widehat{\beta_1} c_{1t} + \widehat{\beta_2} c_{2t} + \widehat{\beta_3} c_{3t} \\ &= \widehat{\beta_1} \alpha_1' x_t + \widehat{\beta_2} \alpha_2' x_t + \widehat{\beta_3} \alpha_3' x_t \\ &= \left(\widehat{\beta_1} \alpha_1' + \widehat{\beta_2} \alpha_2' + \widehat{\beta_3} \alpha_3'\right) x_t, \end{aligned} \tag{7}$$

which shows it is simply a linear combination of the coincident series x_t. Likewise, if we replace c_{1t}, c_{2t}, c_{3t} by their linear optimal predictors $\lambda_1\gamma_1' z_t$, $\lambda_2\gamma_2' z_t$, $\lambda_3\gamma_3' z_t$ in the formula above, we get our "instrumental variable leading index" (IVLI) as a linear combination of the main series z_t, that is:

$$\Delta IVLI_t = \mathbb{E}_{t-1}\left(\widehat{\beta_1}c_{1t} + \widehat{\beta_2}c_{2t} + \widehat{\beta_3}c_{3t}\right) = \left(\widehat{\beta_1}\lambda_1\gamma_1' + \widehat{\beta_2}\lambda_2\gamma_2' + \widehat{\beta_3}\lambda_3\gamma_3'\right) z_t. \tag{8}$$

In summary, our complete statistical model is as follows:

$$\begin{aligned} \text{NBER}_t &= \begin{cases} 1 & \text{if } \mathbb{E}(y_t^* \mid I_{t+h}) < 0 \\ 0 & \text{otherwise.} \end{cases} \\ \mathbb{E}(y_t^* &\mid I_{t+h}) = \psi_0 + \psi' \underset{4\times 1}{x_t} + u_t \\ x_t &= \underset{4\times m}{\Pi} \underset{m\times 1}{z_t} + \varepsilon_t, \end{aligned} \tag{9}$$

where u_t can be correlated with, ε_t; u_t and ε_t have bivariate-Normal distribution; and Π has rank equal to 3.

Next, we present the empirical results obtained by Issler and Vahid (2006). Table 1 below shows that the four coincident series, Industrial Production, Y_t, Income, I_t, Employment, N_t, and Sales, S_t, have three common cycles, since the last canonical correlation is statistically zero, although the first three are not.

Table 1 – Squared canonical correlations and canonical-correlation test

Sq. canonical correlations λ_j^2	degrees of freedom	$H_0 : \lambda_j^2$ and all smaller $\lambda_j^2 = 0$ P-Values (d.f. corrected test)
0.5365	208	0.0000
0.3370	153	0.0000
0.2484	100	0.0000
0.1360	49	0.1768

Source: Issler and Vahid (2006).

The three cycles identified via canonical correlation analysis are as follows:

$$\begin{bmatrix} c_{1t} \\ c_{2t} \\ c_{3t} \end{bmatrix} = \begin{bmatrix} 0.45 & -0.05 & 20.90 & -0.52 \\ 1.43 & -0.69 & 6.72 & -4.78 \\ -0.87 & -7.82 & 14.56 & 2.13 \end{bmatrix} \times \begin{bmatrix} \Delta \ln I_t \\ \Delta \ln Y_t \\ \Delta \ln N_t \\ \Delta \ln S_t \end{bmatrix}. \tag{10}$$

If we normalize these three linear combinations using the respective coefficients of $\Delta \ln N_t$, we will get (approximately) the following: c_{1t} is practically the same as $\Delta \ln N_t$; c_{2t} is approximately equal to $\Delta \ln N_t - \Delta \ln S_t = \Delta \ln \left(\frac{N_t}{S_t}\right)$, i.e., is equal to the growth rate of the number of employees per product sold; lastly, c_{3t} is approximately equal to $\Delta \ln \left(\frac{N_t}{1/2Y_t} \right)$, which is difficult to interpret, as it is far from being equal to the growth rate of employment per unit of output. What is important to note here is the near-ubiquity of employment in these three linear combinations.

Next, Table 2 shows the results of the estimation of the *Probit* model by Conditional Maximum Likelihood in Two Stages (2SCML). The first two common cycles of the four series, c_{1t}, c_{2t}, proved to be significant, although the third, c_{3t}, is only marginally significant.

Table 2 – Two-stage conditional maximum likelihood (2SCML) estimates

Regressor	Est. coeff.	Std. err.
c_{1t}	65.21	(8.86)
c_{2t}	28.05	(6.41)
c_{3t}	13.00	(6.75)
Constant	-0.33	(0.19)
p-value of overall significance		<0.01
McFadden Pseudo-R^2		0.71
% overall correct prediction		94.89%

Source: Issler and Vahid (2006).

With these estimators and the results in hand, we can calculate the weight of each coincident series in the coincident indicator using (7). This is done as follows. First, we rewrite the basic cycles as linear combinations of the coincident series, (7), and, then, the weights are normalized so that they add up to unity, obtaining the following index (HAC robust standard errors are in parentheses):

$$\Delta IVCI_t = \underset{(0.01)}{0.00} \times \Delta \ln I_t + \underset{(0.06)}{0.10} \times \Delta \ln Y_t + \underset{(0.06)}{0.84} \times \Delta \ln N_t + \underset{(0.02)}{0.06} \times \Delta \ln S_t. \qquad (11)$$

Equation (11) shows most of the weight is given to employment, almost no weight is given to income, and employment and industrial production together receive 94% of the weight. This is not surprising, since these two series have a more pronounced consistency with the NBER recession indicator. It also agrees with a memorandum from the Business Cycle Dating Committee (Hall et al. 2002, p. 9), which states that *"employment is probably the most reliable indicator [of recessions]"*.

2.2 Adapting canonical correlation to a *big data* environment

2.2.1 Imputation of Missing Values

The first practical issue to tackle before implementing the methods discussed in previous section is the *missing data*. This occurs in databases for several reasons, but it is up to the final users of the data to treat the missing values using the information available. In the context of canonical correlations, we have two databases in a *big data* format: (i) the coincident series; and (ii) the leading series. The latter database was constructed by Costa et al. (2021), avoiding missing data at the beginning and middle of the sample. But, like any database with many series, there is a mismatch in the updates of the various series, which generates the so-called *ragged edge* problem. The ragged edge problem results in an incomplete database for the final periods of the dataset, causing a jagged edge there; see Wallis (1986). In turn, using the coincident series proposed by Stock and Watson (2014) poses additional problems, as there are more recent series with gaps at the beginning of the sample as well as old series that were discontinued at the end of the sample.

A classic and efficient way to solve the missing data problem is to use the Expectation Maximization (EM) algorithm. There are several forms and versions of this algorithm available in the literature. In this paper, we use a modern form of the EM algorithm proposed in Schneider (2001), which also provides an easy and complete code in MATLAB. Schneider argues the EM algorithm with Gaussian specification for the data is an iterative method both for the estimation of mean values and the variance-covariance matrix of a set of incomplete data, being considered the starting point for the development of a regularized EM algorithm. In contrast to the conventional EM algorithm, the regularized algorithm is applicable to datasets in which the number of variables normally exceeds the sample size, which is usually the case in a big data setup.

The regularized EM algorithm is based on iterated linear regressions of variables with missing values into variables with available values. The regression coefficients are estimated by *Ridge Regressions*, a classic regression method in which a continuous regularization parameter controls the filtering of noise in the data. The regularization parameter is determined by generalized cross-validation, in order to minimize the expected mean squared error of the imputed values. For the imputation of missing values, the regularized EM algorithm can estimate synchronous and diachronic covariance matrices, which may contain information about spatial covariability, stationary temporal covariability or cyclostationary temporal covariability. In this sense, the data to be completed must contain only weakly stationary series, which requires transformations in some original series of the database that are nonstationary, which is very common in the study of coincident and/or leading indicators of economic activity, such as the case of the current study.

2.2.2 Dimensionality reduction

With balanced panels of coincident and leading series, the next step is to implement the analysis of canonical correlations between these two groups of series. It is worth highlighting this is done without adding lags of the coincident series into the leading database, since the past of the coincident series might have good predictive power for the current coincident series.

In several practical cases, the number of series is greater than the number of observations, which is an impediment to the standard implementation of the canonical correlation analysis. A natural solution to deal with this problem is to extract from each database (coincident and leading) its common components, with much less series than the number of series in the original databases. A similar problem was solved by Bai and Ng (2021) using block and subblock factors. The classic way of extracting these factors is through Principal Components Analysis (PCA) – which are simply (orthogonal) linear combinations of the original series, with maximum variance sequentially computed.

In order to give an idea of the dimensionality reduction that we can achieve with this procedure, we note that the coincident database has originally 270 contemporary time series. For example, in the empirical exercise for the USA, to estimate the *Probit* model proposed in Issler and Vahid (2006), we ended up using a maximum of 25 principal components (and a minimum of 5). These 25 principal components explain 75% of the variation in the data, which represents a large reduction in dimensionality, without much loss of information.

In the leading database, this reduction is much greater: there are 329 contemporary series. However, by adding lags of the leading series, besides the coincident series, we can easily exceed $1,000$ series, depending on the lag specification adopted (for two lags we would have $1,198$ series, for example). Again, we managed to use a maximum of 35 principal components (also explaining 75% of the variance in the data) and a minimum of 8 components for the leading series.

Having solved the problems of missing data and dimensionality reduction in the databases, the next step of the analysis by Issler and Vahid is the calculation of the canonical correlations, as these solve the problem of signal extraction, guaranteeing that in the *Probit* model only the cyclic parts of the coincident data explain the state of the economy – expansion *versus* recession.

2.2.3 Estimating the *Probit* structural model

The next step of the empirical analysis is to adjust the *Probit* models proposed in Issler and Vahid (2006). We assume the coincident series can be explained by N basic cycles $\{c_{1t}, c_{2t}, \cdots, c_{Nt}\}$. Let NBER_t denote the NBER indicator. The first stage of Rivers and Vuong's 2SCML estimation procedure involves regressing $\{c_{1t}, c_{2t}, \cdots, c_{Nt}\}$ into k instruments collected in column vector z_t – which are the main components of the leading series – and save the residuals, which we denote by $\{\hat{v}_{1t}, \hat{v}_{2t}, \cdots, \hat{v}_{Nt}\}$. In the second stage, both basic cycles $\{c_{1t}, c_{2t}, \cdots, c_{Nt}\}$ and the first-stage residuals $\{\hat{v}_{1t}, \hat{v}_{2t}, \cdots, \hat{v}_{Nt}\}$ are included as regressors in the *Probit* model:

$$\Pr\left(\text{NBER}_t = 1\right) = \Phi\left(-\left(\beta_0 + \beta_1 c_{1t} + \cdots + \beta_N c_{Nt} + \beta_{N+1}\hat{v}_{1t} + \beta_{N+2}\hat{v}_{2t} + \beta_{2N}\hat{v}_{Nt}\right)\right), \quad (12)$$

where $\Phi(\cdot)$ is the standard Gaussian cumulative distribution function (CDF). The parameters $\beta_0, \beta_1, \beta_2, \cdots, \beta_N$ used in the second-stage of the *Probit* model will be the 2SCML estimates, denoted by $\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2, \cdots, \widehat{\beta}_N$. Based on these estimates, we can predict in *real time* the presence of recessions in period t using:

$$\Pr\left(\widehat{\text{NBER}_t} = 1\right) = \Phi\left(-\left(\widehat{\beta}_0 + \widehat{\beta}_1 c_{1t} + \widehat{\beta}_2 c_{2t} + \cdots + \widehat{\beta}_N c_{Nt}\right)\right), \quad (13)$$

provided that we can update the database up to the current period, period t, using the EM algorithm to fill the missing values in the database as discussed above.

Note that predictions $\Pr\left(\widehat{\text{NBER}_t} = 1\right)$ do not require the knowledge of NBER_t, since we estimate the model using available information up to period $t-1$ to generate $\Pr\left(\widehat{\text{NBER}_t} = 1\right)$. In fact, to date recessions in *real time*, we need a relevant group of series available in t, in the same period – which is hardly the case. However, if we have a relevant group of series available at $t-1$, we can date recessions with a period of delay. On a monthly basis, this represents a month delay – which is a significant gain over the literature, which generally date recessions with a three- to six-month delay. With respect to the NBER, this delay can reach a year or more.

2.3 Discussion on competing approaches

International experience in forecasting rare and extreme events, such as recessions, shows us that we should not be too optimistic when using them in real time. In fact, it is often observed that the fit of the different models is very good within the sample, but they often fail in their behavior outside the sample. This idea is endorsed by Hamilton (2011), in a slightly pessimistic tone regarding real-time recession prediction models, which is the timing in which they have maximum utility and greater relevance. As already noted, the NBER usually dates recessions with a delay ranging from six months to a year. In fact, Hamilton suggests that if one could predict recessions, they probably wouldn't happen:

> *"If people could predict recessions, they probably would not happen. Firms would not be stuck with inventories, labor, and capital they turn out not to need, and the Federal Reserve would probably ease its policy stance earlier. Economists are used to viewing magnitudes such as stock prices as difficult or impossible to predict if the market is functioning properly, and it may be that economic recessions, by their very nature, imply similar fundamental limitations for forecasting."*

Obviously, arguments like these have never been an obstacle for many economists working in the financial market to try to predict asset prices or returns. Likewise, they are not obstacles for economists (academics or not) to try to anticipate recessions. The question to be posed is about what expectations can indeed be achieved. In that sense, taking advantage of the NBER committee was apparently Hamilton's goal in defining his real-time timing. But this does not prevent us from seeking to overcome Hamilton's objectives, as those who seek to predict asset prices in financial markets do on a daily basis.

Stock and Watson (2014) compares two business cycle modeling techniques: *date then average* and *average then date*. The first is the preferred approach of these authors, the second being the preferred technique of Hamilton (1989, 2011) and some followers such as Chauvet and Hamilton (2006). Apparently, the main message of Stock and Watson (2014) is about the usefulness of using a large database of coincident variables in the dating of business cycles. In fact, they used a coincident database containing 270 series representing four categories of real monthly economic activity for the U.S. from 1959:M1 to 2010:M9: employment, industrial production, income, and sales. Apparently, the use of this large database by Stock and Watson corrects the procedure used by them in the 1990-91 recession, as reported by Hamilton (2011):

> *"What went wrong [in the 1990-91 recession]? One of the intriguing new leading indicators that Stock and Watson discovered was the spread between the yield on commercial paper and Treasury bills ... this spread had a dramatic spike prior to each of the recessions in their original sample, but did very little out of the ordinary in the 1990-91 recession, for which their model was on real-time display."*

One way to avoid having a specific variable dictate the behavior of recession dates, therefore, is to use a large database, extracting their respective common components. In the area of multivariate statistics, there are two classic ways of doing this: (i) principal component analysis, proposed by Stock and Watson in several articles; and (ii) canonical correlation analysis, proposed by Issler and Vahid (2006). The first extracts from the database (coincident series or leading series, separately) the factors that are most responsible for the joint variation of the database, ordered in descending order of importance.

As explained above, the second way generalizes the idea behind a least squares regression (or maximum likelihood, under Gaussian assumption). In this context, from two sets of data, coincident series and leading series, linear combinations of the first set that have maximum correlation with linear combinations of the second set are repeatedly sought, until all possibilities are exhausted.

The objective of these multivariate procedures is exactly to extract what is common from the database. This acts as insurance against using idiosyncratic variations of certain isolated series, as did the spread between *Commercial Paper* and the *six-month T-bill* in the 1990-91 recession.

Comparing the two ways to extract common components – principal component analysis and canonical correlation analysis – we see advantages for the use of canonical correlations, as these, in addition to separating *noise* and *signal* from the series in each database (coincident and leading), still makes a bridge between them. In fact, Issler and Vahid emphasize the separation of noise and signal when proposing the use of canonical correlations, when they use only the cyclical properties of the coincident series explaining the state of the economy, recession versus expansion; see equation (6).

The Stock and Watson model, based on principal components, admits past predictability outside the common cycle, which is vetoed in the analysis of canonical correlations by Issler and Vahid, as argued above. This can be seen most clearly from the full Stock and Watson model described earlier.

Consider X_{it} one of four series in $X_t = (Y_t, N_t, S_t, I_t)'$, representing industrial production, employment, sales and income in the period t, $i = 1, 2, 3, 4$; let $y_{it} = \Delta \ln (X_{it})$ be the set of coincident series and z_{t-h} a vector containing the set of leading series, lags of y_{it}, $i = 1, 2, 3, 4$, and other variables with explanatory power for y_{it}. For exposition purposes, we next show Stock and Watson's (1989) stylized factor model discussed earlier:

$$
\begin{aligned}
y_{it} &= \alpha_i + \lambda_i f_t + \gamma_i z_{t-h} + u_{it}, && (14)\\
f_t &= \alpha_f + \phi_1 f_{t-1} + \phi_2 f_{t-2} + \varepsilon_t, && (15)\\
u_{it} &= \rho_{i1} u_{it-1} + \rho_{i2} u_{it-2} + e_{it}, && (16)
\end{aligned}
$$

$$
\begin{bmatrix} \varepsilon_t \\ e_{1t} \\ e_{2t} \\ e_{3t} \\ e_{4t} \end{bmatrix} \sim i.i.d.\ \mathcal{N} \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\varepsilon & 0 & 0 & 0 & 0 \\ 0 & \sigma_1 & 0 & 0 & 0 \\ 0 & 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_3 & 0 \\ 0 & 0 & 0 & 0 & \sigma_4 \end{bmatrix} \right). \qquad (17)
$$

Note that factor f_t is an $AR(2)$ process, thus, it has a cyclical behavior. However, the error term u_{it} also has a cyclical behavior, being also an $AR(2)$. This divides the cyclic part of the series into $X_t = (Y_t, N_t, S_t, I_t)'$ between two components: one common and the other idiosyncratic, which does not control the importance of the common factor vis-à-vis the term idiosyncratic. For example, one can get an idiosyncratic component u_{it} that explains much of the variance of y_{it}, with low explanatory power for f_t. In such a context, the idea of a common component of business cycles, which has been the cornerstone of research in this area since Burns and Mitchell (1946), would fall apart.

Next, we present the empirical results of business cycles expansions/recessions predictions, for the U.S. and the Euro Area, using the canonical correlation approach of Issler and Vahid (2006), adapted to large databases, as advocated by Stock and Watson (2014).

3 Empirical exercise

3.1 United States

For the U.S. economy, in order not to sacrifice too much the number of recessions, we decided to start the sample at 1970:M1 and end it at 2021:M5 (617 observations), which gives us a total of 8 recessions in 51 years, that is, one recession in every 6 and a half years, approximately.

The *coincident* series come from the large database of Stock e Watson (2014); see also McCracken and Ng (2015). From the total of 270 series, we decided not to use the *sales* series in the analysis, as they have a lot of missing data. This way, we use 178 series to form our database for the coincident series. It is worth mentioning that selecting an adequate sample in both time and cross-section dimensions made it possible to apply the EM algorithm with a small number of missing data, which favors a healthy database from the point of view of collinearity between the final series. Regarding the *leading* series, we use in great part the database of Costa et al. (2021), with an amount of 329 contemporaneous variables; see Appendix A for the full list of leading variables for the U.S.

For the coincident series, we used a maximum of 25 principal components and a minimum of five. In the leading database, where the final number of series can reach more than $1,000$ series (depending on the number of lags and the specification adopted), we used a maximum of 35 principal components and a minimum of eight.

In possession of the principal components of the coincident and leading series, following the procedure in Issler and Vahid (2006), we compute the canonical correlations between both groups, in order to have the basic cycles, $c_{1t}, c_{2t}, \cdots, c_{Nt}$, of the coincident series. Using these cycles, we estimate the structural *Probit* model, obtaining the maximum likelihood estimates of $\beta_0, \beta_1, \beta_2, \cdots, \beta_N$, needed to predict US recessions in real time (dated by NBER).

Next, we present (for illustrative purposes only) the estimates of $\beta_0, \beta_1, \beta_2, \cdots, \beta_N$ for the full sample, noting that in real-time forecasting exercises these will be done recursively (month after month) to allow tracking of the state of the U.S. economy. We present the estimates for the U.S. of a model with five principal components for the coincident series, seven principal components for the leading series, and with two lags of the latter. This will be our *benchmark* model for the U.S.

Table 3 – Probit model estimates (USA)
full sample (2SCML)

Regressor	β_i	Robust std. dev. of β_i
Constant	-2.47***	(0.258)
c_{1t}	17.93***	(6.699)
c_{2t}	-113.10***	(18.928)
c_{3t}	8.61	(14.094)
c_{4t}	-8.86	(10.463)
c_{5t}	24.54**	(11.546)

Note: ** and *** denotes 5% and 1% significance levels, respectively.

Once we fit the models for the entire sample, we can recursively use these estimates in forecasting exercises relevant to predicting recessions in real time. To do this, we use the following procedure:

1. From an initial estimation period of the *Probit* models (called *burning period*) the estimated coefficients are used to make out-of-sample predictions using data actually observed up to period t using the equation (13). This is done for the following episodes: 1996:M7-2001:M1, 2001:M2-2004:M3, 2004:M4-2007:M11, 2007:M12-2011:M8, 2011:M9-2015:M10, 2015:M11-2019:M12, and 2020:M1 onwards.

2. To try to reproduce the exercises carried out in real time, for each forecast episode, we lock the estimates of the *Probit* model at an earlier period, where we are sure the dependent variable (0 or 1) was already dated by NBER. Using only the locked estimates of $\beta_0, \beta_1, \beta_2, \cdots, \beta_N$, we can then predict the probability of recession in each episode, varying over time only N significant basic cycles $c_{1t}, c_{2t}, \cdots, c_{Nt}$ that are available in real time.

Note that, in these episodes, there are some recessive periods included, but the important thing here is that the parameters are fixed for the period immediately before each episode, which is exactly what would be done in real time. Indeed, one does not know if there is a recession (or not) in the current period, which will only be known some time later, when the NBER in fact declares that a recession has occurred.

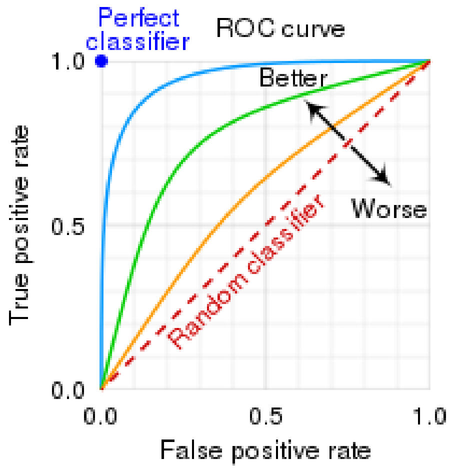
In this procedure, we disregarded the *ragged edge* problem due to the different update delays of the series in the database. In real time, we will have to fill in the missing data to be able to use the equation (13) to predict recessions out-of-sample.

A relevant empirical question, from a practical point of view in real time, is related to the threshold from which one declares that a recession is occurring (or not) in period t using (13), noting that all forecasts will be in the range $[0, 1]$. One way to assess out-of-sample forecasting performance is to use a *receiver operating characteristic (ROC) curve*. This curve is a graph that illustrates the diagnostic capability of a binary classifier system as its discrimination threshold is varied. The method was originally developed for operators of military radar receivers, hence its name.

The ROC curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various classification threshold settings. The true positive ratio is also known as *recall* or *detection probability* in the machine learning literature. The false positive rate is also known as the *false alarm probability* and can be calculated as $(1 - specificity)$.

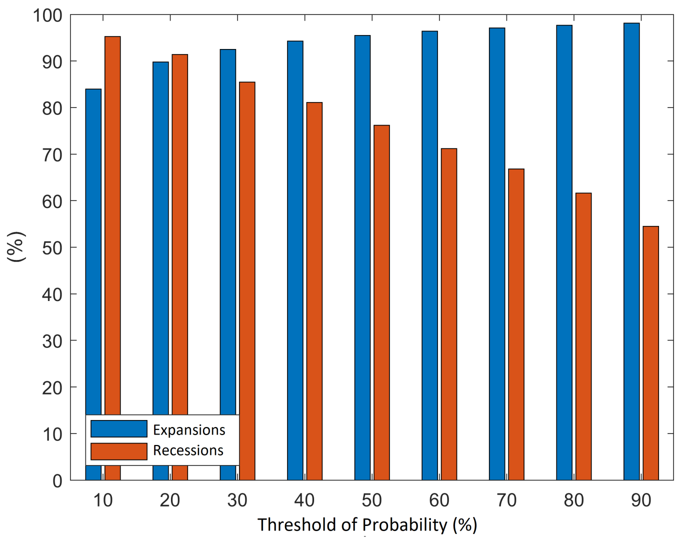
The ROC curve can also be thought of as a graph of power versus Type I error of the decision rule. In short, if used correctly, the ROC curve is a powerful tool as a statistical measure of performance in detection/classification theory and hypothesis testing, as it allows having all relevant quantities on a single graph; as follows:

Figure 1 – Receiver Operating Characteristic (ROC) curve



Instead of presenting an ROC curve, we present here a graph including the trade-off between the correct rate of expansion and the correct rate of recession, for various levels of cut-off, ranging from 10% to 90%. For the U.S. database, considering all possible estimated models, we present the results in Figure 2.

Figure 2 – Frequency of events correctly predicted out-of-sample for the U.S. (all models)

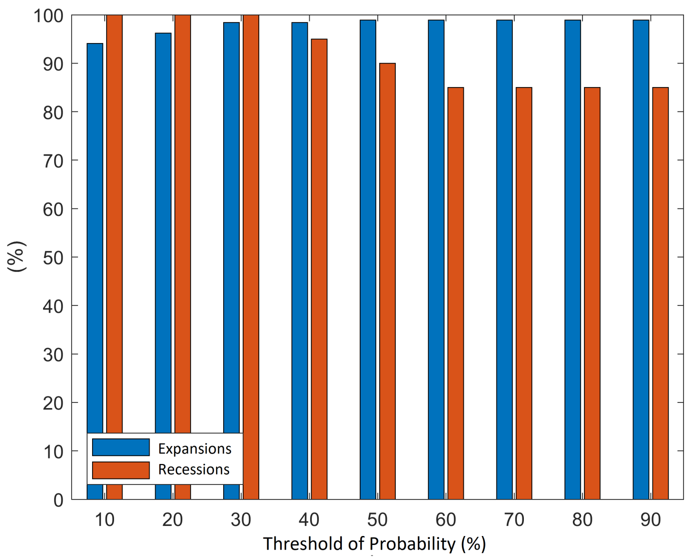


The separation of the expansion and recession forecast events allows us to clearly see the trade-off between the forecast of both cases. For the 10% threshold, the models correctly predict a little less than 100% of recessions, but correctly predict expansions with accuracy around 85%. As the threshold increases to 90%, the correct prediction of recessions worsens (reaching less than 60%), but the correct prediction of expansions improves (reaching almost 100%).

For example, for the 20% threshold, both events are correctly predicted with probabilities almost equal to 90%. This would be a good choice if the econometrician equally weights the correct forecasts of expansion and recession. More generally, it is up to the econometrician to choose the threshold of preference, taking into account the probabilities of false classification of expansions and recessions for the different cut-off thresholds.

Now, if we want to calculate these same statistics for the U.S. *benchmark* model, with 5 principal components for the coincident series, 7 principal components for the leading series, and with two lags, the results are even more promising, as shown in Figure 3:

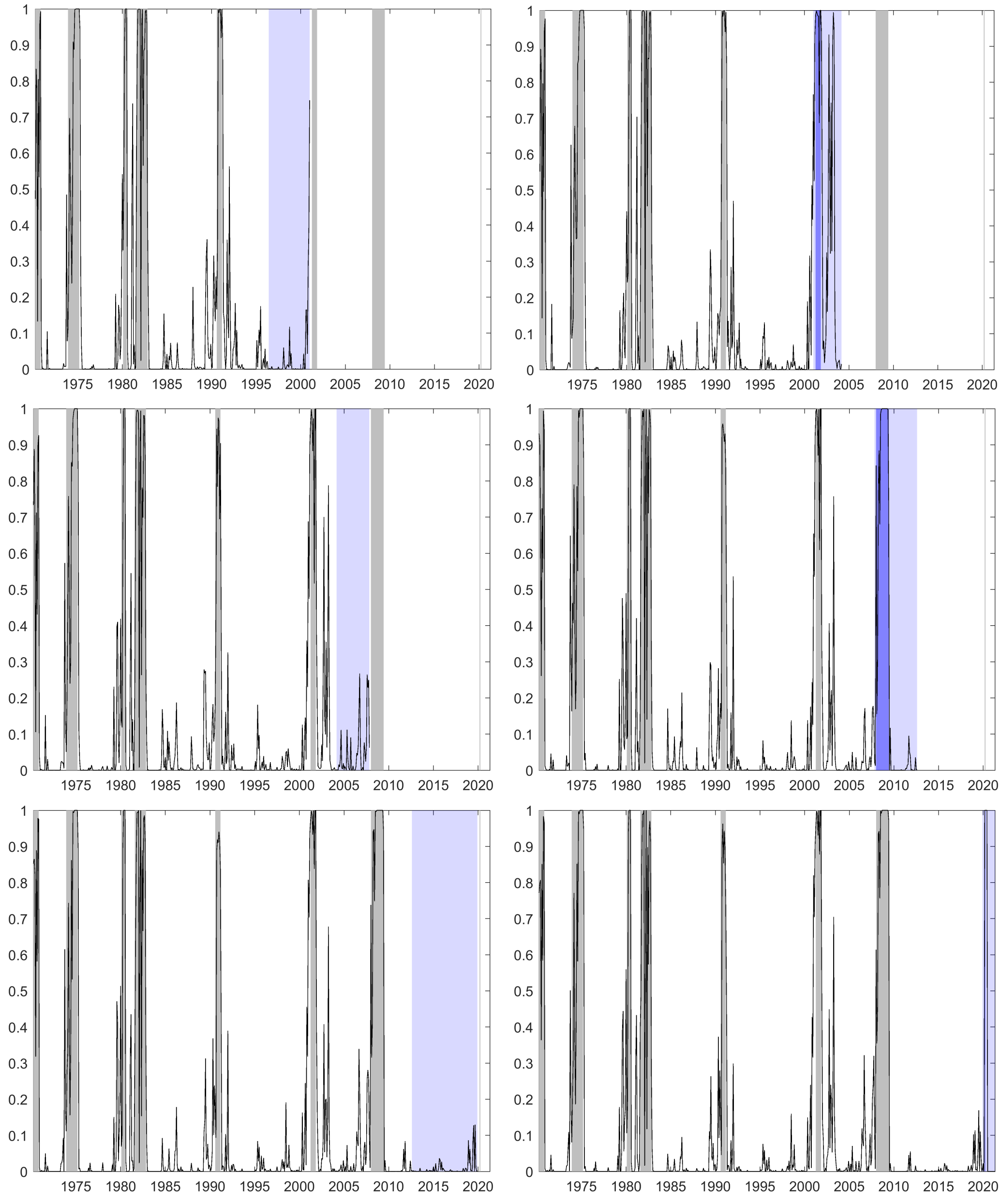
Figure 3 – Frequency of events correctly predicted out-of-sample for the U.S. (*benchmark* model)



Note the results are much more robust for the benchmark model. Regarding the 30% threshold, the model correctly predicts almost 100% of expansions and recessions (98.39% and 100%, respectively). Any other probability thresholds do not involve two-way improvements. For example, increasing the threshold to 70% generates an improvement in the probability of expansion and a worsening in the probability of recession estimates (98.92% and 85%, respectively). On the other hand, decreasing the threshold to 10% generates a worsening in the estimates of probability of expansion and maintenance of the probability of recession (94.09% and 100%, respectively).

In our first exploratory forecasting exercise, we show below the sequence of *real-time* forecasts in period t using (13) for all episodes (1996:M7-2001:M1, 2001:M2-2004:M3, 2004:M4-2007:M11, 2007:M12-2011:M8, 2011:M9-2015:M10, 2015:M11-2019:M12, and 2020:M1 onwards) for the model using six coincident principal components and eight leading principal components. The number of lags in the leading series is three (months), but in this case, using two lags generates very similar results and using four lags generates worse results. Each real-time forecast episode is colored blue and may or may not include recessions dated by the NBER. This allows investigating the behavior of the model in the two contexts of activity: expansion and contraction. The charts below are shown in chronological order of episodes in Figure 4.

Figure 4 – Probability of recessions in real time for the U.S.



Notes: Gray bars indicate the NBER recession periods, and blue bars denote real-time forecast episodes.

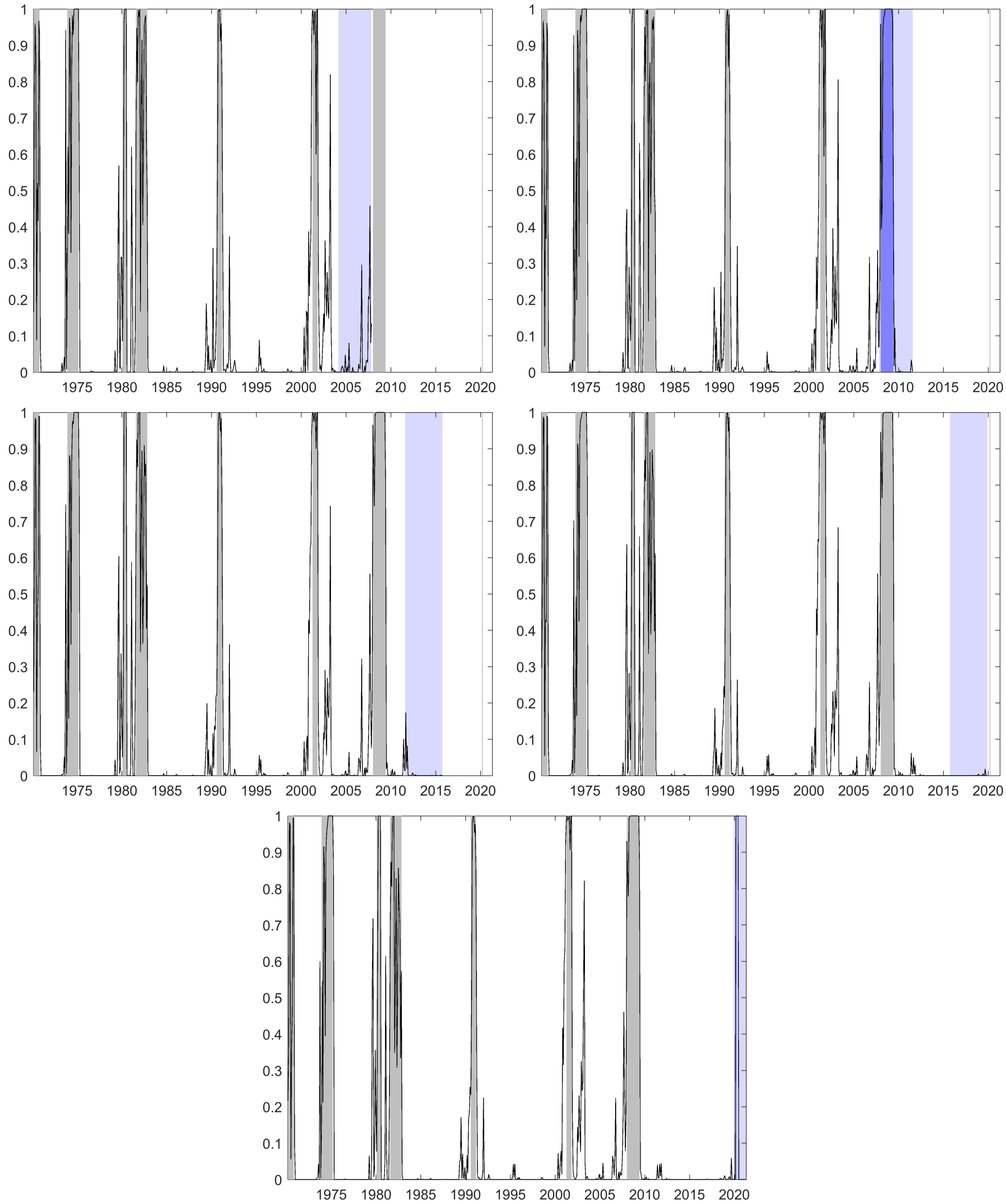
Empirically, the good forecast in real time is remarkable in the various episodes, perhaps with the exception of the second – 2001:M2-2004:M3 – in which we observe potentially false forecasts of recession in periods of actual expansion according to the NBER.

It is worth noting that, in the different exercises for the U.S., we estimated a total of 48 models, varying the number of coincident and leading principal components and the number of lags. Some of these exercises were in-sample and others were out-of-sample. It was not by chance that we showed the out-of-sample forecasting exercise for the model using six coincident principal components and eight leading principal components with three lags. This and some other models with few principal components also performed well. This guided us to apply the ever-present *principle of parsimony* to out-of-sample forecasts. The preferred number of lags was between two and three, leaning more towards the first, depending on the specification.

A second point to note is that, as we have more observations (and therefore more recessive episodes), the real-time forecasts improve both for recessions as well as for expansions. This is a natural outcome, since there are few recessive episodes at the beginning of the sample to estimate more precisely the parameters of the model. In our second exploratory exercise, we will take into account these results and focus only on the most recent episodes.

In our final exploratory exercise, using the benchmark model, based on five coincident principal components and seven leading principal components with two lags, the results of the out-of-sample forecasts are shown in Figure 5 for the last five episodes (2004:M4-2007:M11, 2007:M12-2011:M8, 2011:M9-2015:M10, 2015:M11-2019:M12, and 2020:M1 onwards). Note that, in general, the results improve compared to those from our preliminary exercise.

Figure 5 – Probability of recessions in real time for the U.S. (*benchmark* model)



Notes: Gray bars indicate the NBER recession periods, and blue bars denote real-time forecast episodes.

3.2 Euro Area

The CEPR-EABCN (Euro Area Cycle Dating Committee) sets the chronology of recessions and expansions for the eleven original Euro area member countries plus Greece for 1970-1998, and for the entire Euro area from 1999 onwards. The Committee also publishes, in spring and autumn, the current state of aggregate economic activity in the Euro area and launches research initiatives aimed at monitoring and better understanding aggregate economic developments in the Euro area.

Due to the limitation of aggregate economic activity data in pre-Euro area periods, we decided to start the sample in 1970:M1 and end it in 2021:M5, which gives us a total of 5 recessions in 51 years, i.e. roughly one every 10 years. Again, this allowed applying the EM algorithm with a small number of missing data, which favors a healthy database from the point of view of collinearity between the final series.

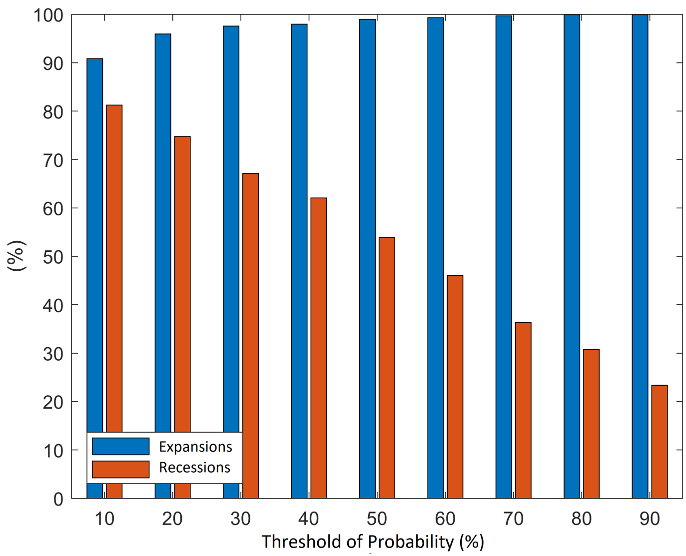
The *leading* database is formed by 220 series selected from the data of Costa et al. (2021). The *coincident* database is formed by 126 series from the ECB, Fred-OECD, and IMF; see Appendix A for the full list of coincident and leading variables for the Euro Area.

For the coincident series, we used a maximum of eleven principal components and a minimum of four. In the leading database, where the number of series again can reach more than $1,000$ depending on the specification used, we used a maximum of 22 principal components and a minimum of seven.

With the principal components of the coincident and leading series, following the procedure in Issler and Vahid (2006), we estimated the basic cycles by canonical correlations and later the *Probit* model, obtaining the maximum likelihood estimates of $\beta_0, \beta_1, \beta_2, \cdots, \beta_N$, necessary to predict in real time the European recessions dated by the CEPR-EABCN.

Next, we show the general results for the choice of cut-off thresholds, based on the ROC analysis, similar to the one made for the U.S. above. It is clear that the European results show a worsening in relation to the results for the U.S., since the best estimate of the probability of recession is approximately 80%, with the threshold of 10%. This outcome is possibly due to the shorter sample size of the European data, compared to the U.S. data, and fewer number of recessions/expansions used as input in the *Probit* model.

Figure 6 – Frequency of events correctly predicted out-of-sample for Europe (all models)



The search for a European *benchmark* model led us to the model with 4 coincident principal components (53.1% of the coincidental variation) and 5 leading PCs (45.0% of the leading variation) with two lags. Table 4 presents the estimates of our chosen benchmark *Probit* model for Europe.

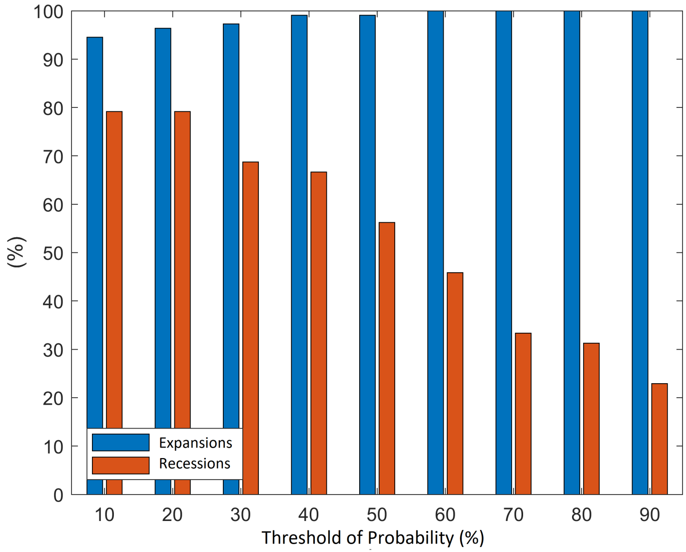
Table 4 – Probit model estimates (Euro area)
full sample (2SCML)

Regressor	β_i	Robust std. dev. of β_i
Constant	-3.19***	(0.362)
c_{1t}	21.50***	(2.609)
c_{2t}	-27.03***	(3.538)
c_{3t}	15.70***	(3.721)
c_{4t}	9.21	(5.711)

Note: ** and *** denote 5% and 1% significance levels, respectively.

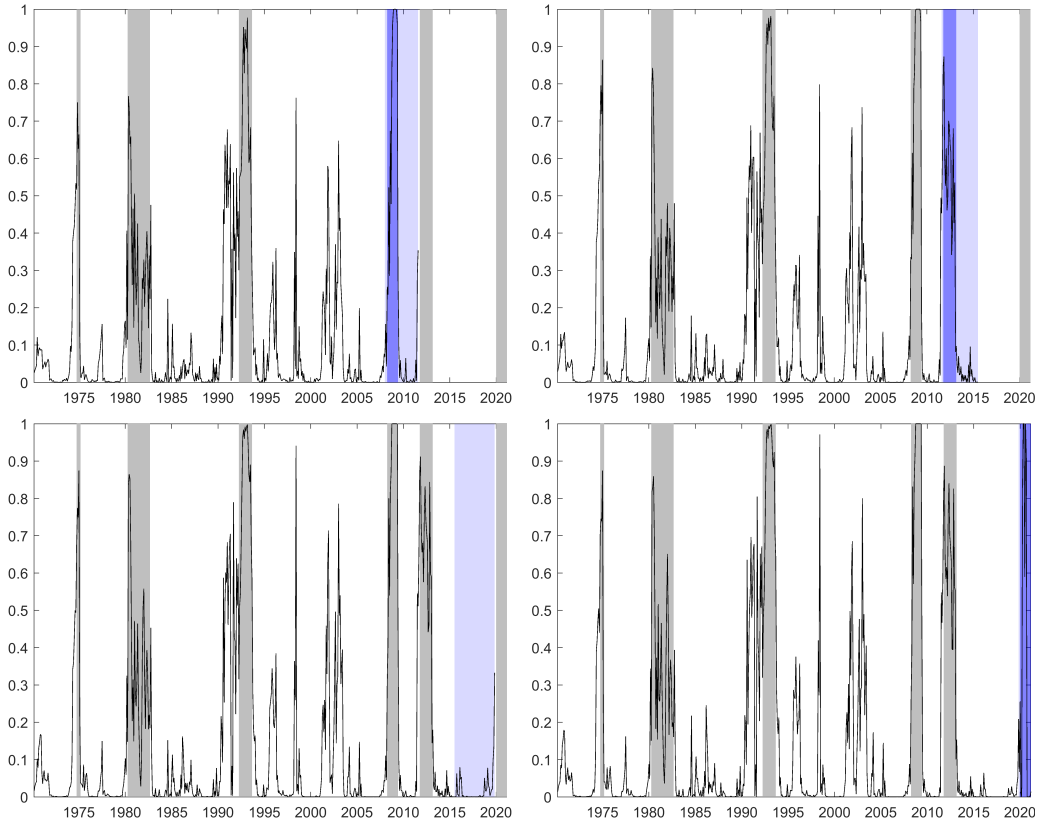
For the benchmark model, the results for the choice of cut-off thresholds are slightly more encouraging compared to the previous graph, especially for the threshold of 20% – 96.36% and 79.17%, respectively for the probability of expansion and recession.

Figure 7 – Frequency of events correctly predicted
out-of-sample for Europe (*benchmark* model)



In our last exercise for Europe, which uses the benchmark model, with four coincident principal components and five leading principal components with two lags, we show the results of the out-of-sample forecasts, for the following episodes, 2007:M12-2011:M8, 2011 :M9-2015:M7, 2015:M8-2019:M11, and 2019:M12 onwards.

Figure 8 – Probability of recessions in real time for Europe (*benchmark* model)



Notes: Gray bars indicate the CEPR-EABCN recession periods, and blue bars denote real-time forecast episodes.

Empirically, the good forecast in real time is remarkable in the various episodes, perhaps with the exception of the first one – 2007:M12-2011:M8 – in which we observed potentially false forecasts of recession in periods of expansion according to the CEPR-EABCN, noting they were slightly worse than those for the U.S.

3.3 Brazil and China

Despite Brazil having a business cycle dating committee – CODACE – the results using the methodology proposed here were disappointing, in the sense that the model is barely able to distinguish between recessive and non-recessive periods. In our assessment, the problem is not the methodology itself, as it worked well for the U.S. and Europe, but it has to do with the lack of longer data on economic activity for a broad set of coincident series and even the lack of this same set for the leading series.

The Chinese economy has shown unusual resilience since the 1990s, when GDP only once experienced negative quarterly growth. For this reason, we understand that it makes little sense to investigate the probability of a recession based on current historical data. Thus, we leave the investigation of these two countries for future research.

3.4 Practical recommendations

As a result of the critical analysis of the previous literature, and of the empirical exercises of dating of recessions presented here, we provide practical recommendations for the dating of recessions in real time using *big data*, summarized as follows:

(i) **Scope**: The techniques discussed in this paper can be immediately applied to countries or regions that have official business cycle dates, the most relevant today being the U.S. and the Euro Area;

(ii) **Leading series**: It seems very difficult to have a model for dating recessions in *real time* that does not contain a good group of leading variables, which must be large enough to benefit from the use of *big data* techniques and generate accurate predictions of the state of the economy;

(iii) **Coincident series**: An important question is about the format of the coincident series. The set of series must be large, as in Stock and Watson (2014), who used a set of 270 series representing four main groups of variables: industrial production, income, sales and employment. Ideally, one would use the same database that Stock and Watson used. However, the missing data present in most of the sample bring a practical empirical issue to be solved; and

(iv) **Principal components *versus* canonical correlations**: Stock and Watson (2014) use principal component analysis and state-space models to identify and estimate common factors. In this approach, there is no direct connection between coincident and leading series. In Issler and Vahid (2006), who use canonical correlation analysis, only the predictable parts of the coincident series using the leading series are used in the structural *Probit* model that predicts recession probabilities.[9]

[9] This eliminates the noise present in the coincident series. Despite this, one must use principal components to reduce the size of the coincident and leading databases, since, in some cases, there are more series than temporal observations, which makes the direct application of regression models not feasible.

4 Conclusions

First, this article focused on a brief review of the specialized literature on recession dating models, whether *real-time* or slightly *a posteriori*. Both are of practical interest, as, with such models, firms can anticipate sudden drops (or increases) in asset prices in order to adjust their input needs, which is fundamental for cost-minimizing firms. On the other hand, such models are useful for policymakers to better track the state of the economy in a timely manner, providing valuable insights for decision-making.

In this sense, this paper presents several relevant articles from this literature, including Chauvet and Hamilton (2006), Issler and Vahid (2006), Stock and Watson (1989, 1991, 1993a, 1993b, 2002, 2010, 2014), Hamilton (2011), and Kotchoni and Stevanovic (2018).

Second, in the light of the results of our literature survey, we consider which are the best tools that can indicate the beginning of a recession in real time, or slightly *a posteriori*. At this point, we decided to use a modified version of the Issler and Vahid methodology, adapted for a *big data* environment based on the advantages presented by Stock and Watson (2014).

Our results show it is possible to follow the state of the economy of some key countries or agglomerations of countries (Europe) using these models, provided that appropriate techniques for reducing the dimensionality of the databases are implemented – canonical correlations and principal component analysis. Depending on the cut-off thresholds, for the U.S. benchmark model, it is possible to track the state of the economy by predicting out-of-sample recessions with 98.39% and almost 100% accuracy for recessions and expansions, respectively. For the Euro Area, in this same context, we can reach 79.17% and 96.36% accuracy for recessions and expansions, respectively.

References

[1] Akaike, H., 1976. Canonical Correlation Analysis of Time Series and the Use of an Information Criterion. In R.K. Mehra and D.G. Lainiotis (Eds) *System Identification: Advances and Case Studies,* New York: Academic Press, 27-96.

[2] Anderson, T.W., 1984. *An Introduction to Multivariate Statistical Analysis*. Second edition. NewYork: John Wiley and Sons.

[3] Anderson, H.M., Vahid, F., 1998. Testing Multiple Equation Systems for Common Nonlinear Factors. *Journal of Econometrics* 84, 1-37.

[4] Bai, J., Ng, S., 2021. Matrix Completion, Counterfactuals, and Factor Analysis of Missing Data. *Journal of the American Statistical Association* 116 (536), 1746-1763.

[5] Balke, N.S., Fomby, T.B., 1997. Threshold Cointegration. *International Economic Review* 38, 627-647.

[6] Burns, A.F., Mitchell, W.C., 1946. *Measuring Business Cycles*. New York: National Bureau of Economic Research.

[7] Campbell, J.Y., 2003. Consumption-based asset pricing. In *Handbook of the Economics of Finance*, vol.1(B), 803-887.

[8] Chauvet, M., 1998. An Econometric Characterization of Business Cycle Dynamics with Factor Structure and Regime Switching. *International Economic Review* 39, 969-996.

[9] Chauvet, M., Hamilton, J.D., 2006. Dating business cycle turning points. In C. Milas, P. Rothman, & D. van Dijk (Eds.), *Nonlinear time series analysis of business cycles*, 1-54. Amsterdam: Elsevier.

[10] Christiansen, C., Eriksen, J.N., Miller, S.V., 2013. Forecasting US recessions: The role of sentiments. Technical report, CREATES Research Paper 2013-14.

[11] Costa, A.B.R., Ferreira, P.C., Gaglianone, W.P., Guillén, O.T.C., Issler, J.V., Lin, Y., 2021. Machine Learning and Oil Price Point and Density Forecasting. *Enegy Economics* 102, 105494.

[12] Croushore, D., Stark, T., 2003. A real-time data set for macroeconomists: does the data vintage matter? *Review of Economics and Statistics* 85, 605-617.

[13] Engle, R.F., Granger, C.W.J., 1987. Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica* 55(2), 251-276.

[14] Engle, R.F., Kozicki, S., 1993. Testing for Common Features. *Journal of Business and Economic Statistics* 11, 369-395.

[15] Engle, R.F., Issler, J.V., 1995. Estimating Sectoral Cycles Using Cointegration and Common Features. *Journal of Monetary Economics* 35, 83-113.

[16] Estrella A., Mishkin, F.S., 1998. Predicting U.S. Recessions: Financial Variables as Leading Indicators. *Review of Economics and Statistics* 80, 45-61.

[17] Hall, R., Feldstein, M., Bernanke, B., Frankel, J., Gordon, R. and Zarnowitz, V., 2002. The NBER's Business-Cycle Dating Procedure. Downloadable from http://www.nber.org/cycles/november2001/recessions.pdf.

[18] Hamilton, J.D., 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica* 57, 357-384.

[19] Hamilton, J.D., 2011. Calling recessions in real time. *International Journal of Forecasting* 27, 1006-1026.

[20] Hao, L., Ng, E.C., 2011. Predicting Canadian recessions using dynamic probit modelling approaches. *Canadian Journal of Economics* 44(4), 1297-1330.

[21] Hotelling, H., 1935. The Most Predictable Criterion. *Journal of Educational Psychology* 26, 139-142.

[22] Hotelling, H., 1936. Relations Between Two Sets of Variates. *Biometrika* 28, 321-377.

[23] Issler, J.V., Vahid, F., 2001. Common Cycles and the Importance of Transitory Shocks to Macroeconomic Aggregates. *Journal of Monetary Economics* 47, 449-475.

[24] Issler, J.V., Vahid, F., 2006. The missing link: Using the NBER recession indicator to construct coincident and leading indices of economic activity. *Journal of Econometrics* 132, 281-303.

[25] Issler, J.V., Pimentel, L.M.M., 2019. Uma Medida de PIB Mensal Para o Brasil Usando o Term Spread. *Revista Brasileira de Economia* 73(1), 53-75.

[26] Kotchoni, R., Stevanovic, D., 2018. Forecasting U.S. Recessions and Economic Activity. Working paper, Université du Québec à Montréal.

[27] McCracken, M.W., Ng, S., 2015. FRED-MD: A Monthly Database for Macroeconomic Research. Working Paper 2015-012B, Federal Reserve Bank of St. Louis.

[28] Nyberg, H., 2010. Dynamic probit models and financial variables in recession forecasting. *Journal of Forecasting*, Special Issue: Advances in Business Cycle Analysis and Forecasting, 29, 215-230.

[29] Rivers, D., Vuong, Q.H., 1988. Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models. *Journal of Econometrics* 39, 347-366.

[30] Rudebusch, G., Williams, J., 2009. Forecasting recessions: The puzzle of the enduring power of the yield curve. *Journal of Business and Economic Statistics* 27, 492-503.

[31] Schneider, T., 2001. Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *Journal of Climate* 14, 853-871.

[32] Stock J.H., Watson M.W., 1988a. A New Approach to Leading Economic Indicators. Mimeo, Harvard University, Kennedy School of Government.

[33] Stock J.H., Watson M.W., 1988b. A Probability Model of the Coincident Economic Indicator. NBER Working Paper # 2772.

[34] Stock J.H., Watson M.W., 1989. New Indexes of Leading and Coincident Economic Indicators. *NBER Macroeconomics Annual*, 351-95.

[35] Stock J.H., Watson M.W., 1991. A Probability Model of the Coincident Economic Indicators. In *"Leading Economic Indicators: New Approaches and Forecasting Records,"* K. Lahiri and G. Moore, Eds. New York: Cambridge University Press.

[36] Stock J.H., Watson M.W., 1993a. A Procedure for Predicting Recessions with Leading Indicators: Econometric Issues and Recent Experiences. In "*New Research on Business Cycles, Indicators and Forecasting,*" J.H. Stock and M.W. Watson, Eds., Chicago: University of Chicago Press, for NBER.

[37] Stock J.H., Watson M.W., 1993b. *New Research on Business Cycles, Indicators and Forecasting.* Chicago: University of Chicago Press, for NBER.

[38] Stock J.H., Watson M.W., 2002. Forecasting Using Principal Components From a Large Number of Predictors. *Journal of the American Statistical Association* 97, 1167-1179.

[39] Stock, J.H., Watson, M.W., 2010. New Indexes of Monthly GDP. Available at http://www.princeton.edu/~mwatson/mgdp_gdi.html

[40] Stock, J.H., Watson, M.W., 2014. Estimating turning points using large data sets. *Journal of Econometrics* 178, 368-381.

[41] The Conference Board. 1997. Business Cycle Indicators. Mimeo, The Conference Board. Downloadable from http://www.tcb-indicators.org/bcioverview/bci4.pdf.

[42] Vahid, F., Engle, R.F., 1993. Common Trends and Common Cycles. *Journal of Applied Econometrics* 8, 341-360.

[43] Vahid, F., Engle, R.F., 1997. Codependent Cycles. *Journal of Econometrics* 80(2), 199-221.

[44] Wallis, K.F., 1986. Forecasting with an econometric model: The 'ragged edge' problem. *Journal of Forecasting* 5(1), 1-13.

[45] Wooldridge, J.M., 1994. Estimation and Inference for Dependent Processes. In "*Handbook of Econometrics IV,*" R.F. Engle and D. McFadden, Editors. Amsterdam: Elsevier Press.

[46] Wright, J.H., 2006. The yield curve and predicting recessions. Working paper, Federal Reserve Board.

Appendix A. Database

Table A1 – Leading series of economic activity for the United States

	Category	Name	Source	Nickname	tcode
1	Output and Income	Real Personal Income	FRED-MD	RPI	5
2	Output and Income	Real personal income ex transfer receipts	FRED-MD	W875RX1	5
3	Output and Income	IP Index	FRED-MD	INDPRO	5
4	Output and Income	IP: Final Products and Nonindustrial Supplies	FRED-MD	IPFPNSS	5
5	Output and Income	IP: Final Products (Market Group)	FRED-MD	IPFINAL	5
6	Output and Income	IP: Consumer Goods	FRED-MD	IPCONGD	5
7	Output and Income	IP: Durable Consumer Goods	FRED-MD	IPDCONGD	5
8	Output and Income	IP: Nondurable Consumer Goods	FRED-MD	IPNCONGD	5
9	Output and Income	IP: Business Equipment	FRED-MD	IPBUSEQ	5
10	Output and Income	IP: Materials	FRED-MD	IPMAT	5
11	Output and Income	IP: Durable Materials	FRED-MD	IPDMAT	5
12	Output and Income	IP: Nondurable Materials	FRED-MD	IPNMAT	5
13	Output and Income	IP: Manufacturing (SIC)	FRED-MD	IPMANSICS	5
14	Output and Income	IP: Residential Utilities	FRED-MD	IPB51222s	5
15	Output and Income	IP: Fuels	FRED-MD	IPFUELS	5
16	Output and Income	Capacity Utilization: Manufacturing	FRED-MD	CUMFNS	2
17	Labor market	Help-Wanted Index for United States	FRED-MD	HWI	2
18	Labor market	Ratio of Help Wanted/No. Unemployed	FRED-MD	HWIURATIO	2
19	Labor market	Civilian Labor Force	FRED-MD	CLF16OV	5
20	Labor market	Civilian Employment	FRED-MD	CE16OV	5
21	Labor market	Civilian Unemployment Rate	FRED-MD	UNRATE	2
22	Labor market	Average Duration of Unemployment (Weeks)	FRED-MD	UEMPMEAN	2
23	Labor market	Civilians Unemployed - Less Than 5 Weeks	FRED-MD	UEMPLT5	5
24	Labor market	Civilians Unemployed for 5-14 Weeks	FRED-MD	UEMP5TO14	5
25	Labor market	Civilians Unemployed - 15 Weeks & Over	FRED-MD	UEMP15OV	5
26	Labor market	Civilians Unemployed for 15-26 Weeks	FRED-MD	UEMP15T26	5
27	Labor market	Civilians Unemployed for 27 Weeks and Over	FRED-MD	UEMP27OV	5
28	Labor market	Initial Claims	FRED-MD	CLAIMSx	5
29	Labor market	All Employees: Total nonfarm	FRED-MD	PAYEMS	5
30	Labor market	All Employees: Goods-Producing Industries	FRED-MD	USGOOD	5
31	Labor market	All Employees: Mining and Logging: Mining	FRED-MD	CES1021000001	5
32	Labor market	All Employees: Construction	FRED-MD	USCONS	5
33	Labor market	All Employees: Manufacturing	FRED-MD	MANEMP	5
34	Labor market	All Employees: Durable goods	FRED-MD	DMANEMP	5
35	Labor market	All Employees: Nondurable goods	FRED-MD	NDMANEMP	5
36	Labor market	All Employees: Service-Providing Industries	FRED-MD	SRVPRD	5
37	Labor market	All Employees: Trade, Transportation & Utilities	FRED-MD	USTPU	5
38	Labor market	All Employees: Wholesale Trade	FRED-MD	USWTRADE	5
39	Labor market	All Employees: Retail Trade	FRED-MD	USTRADE	5
40	Labor market	All Employees: Financial Activities	FRED-MD	USFIRE	5
41	Labor market	All Employees: Government	FRED-MD	USGOVT	5
42	Labor market	Avg Weekly Hours : Goods-Producing	FRED-MD	CES0600000007	1
43	Labor market	Avg Weekly Overtime Hours : Manufacturing	FRED-MD	AWOTMAN	2
44	Labor market	Avg Weekly Hours : Manufacturing	FRED-MD	AWHMAN	1
45	Labor market	Avg Hourly Earnings : Goods-Producing	FRED-MD	CES0600000008	5
46	Labor market	Avg Hourly Earnings : Construction	FRED-MD	CES2000000008	5
47	Labor market	Avg Hourly Earnings : Manufacturing	FRED-MD	CES3000000008	5
48	Housing	Housing Starts: Total New Privately Owned	FRED-MD	HOUST	4
49	Housing	Housing Starts, Northeast	FRED-MD	HOUSTNE	4
50	Housing	Housing Starts, Midwest	FRED-MD	HOUSTMW	4
51	Housing	Housing Starts, South	FRED-MD	HOUSTS	4
52	Housing	Housing Starts, West	FRED-MD	HOUSTW	4
53	Housing	New Private Housing Permits (SAAR)	FRED-MD	PERMIT	4
54	Housing	New Private Housing Permits, Northeast (SAAR)	FRED-MD	PERMITNE	4
55	Housing	New Private Housing Permits, Midwest (SAAR)	FRED-MD	PERMITMW	4
56	Housing	New Private Housing Permits, South (SAAR)	FRED-MD	PERMITS	4
57	Housing	New Private Housing Permits, West (SAAR)	FRED-MD	PERMITW	4
58	Consumption, orders, and inventories	Real personal consumption expenditures	FRED-MD	DPCERA3M086SBEA	5
59	Consumption, orders, and inventories	Real Manu. and Trade Industries Sales	FRED-MD	CMRMTSPLx	5
60	Consumption, orders, and inventories	Retail and Food Services Sales	FRED-MD	RETAILx	5
61	Consumption, orders, and inventories	New Orders for Consumer Goods	FRED-MD	ACOGNO	5
62	Consumption, orders, and inventories	New Orders for Durable Goods	FRED-MD	AMDMNOx	5
63	Consumption, orders, and inventories	New Orders for Nondefense Capital Goods	FRED-MD	ANDENOx	5
64	Consumption, orders, and inventories	Unfilled Orders for Durable Goods	FRED-MD	AMDMUOx	5
65	Consumption, orders, and inventories	Total Business Inventories	FRED-MD	BUSINVx	5
66	Consumption, orders, and inventories	Total Business: Inventories to Sales Ratio	FRED-MD	ISRATIOx	2
67	Consumption, orders, and inventories	Consumer Sentiment Index	FRED-MD	UMCSENTx	2

Note: The column "tcode" denotes the following series transformations:
(1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $\ln(x_t)$; (5) $\Delta\ln(x_t)$; (6) $\Delta^2\ln(x_t)$.

Table A2 – Leading series of economic activity for the United States (cont.)

	Category	Name	Source	Nickname	tcode
68	Money and credit	M1 Money Stock	FRED-MD	M1SL	5
69	Money and credit	M2 Money Stock	FRED-MD	M2SL	5
70	Money and credit	Real M2 Money Stock	FRED-MD	M2REAL	5
71	Money and credit	Monetary Base	FRED-MD	BOGMBASE	5
72	Money and credit	Total Reserves of Depository Institutions	FRED-MD	TOTRESNS	5
73	Money and credit	Reserves Of Depository Institutions	FRED-MD	NONBORRES	7
74	Money and credit	Commercial and Industrial Loans	FRED-MD	BUSLOANS	5
75	Money and credit	Real Estate Loans at All Commercial Banks	FRED-MD	REALLN	5
76	Money and credit	Total Nonrevolving Credit	FRED-MD	NONREVSL	5
77	Money and credit	Nonrevolving consumer credit to Personal Income	FRED-MD	CONSPI	2
78	Money and credit	MZM Money Stock	FRED-MD	MZMSL	5
79	Money and credit	Consumer Motor Vehicle Loans Outstanding	FRED-MD	DTCOLNVHFNM	5
80	Money and credit	Total Consumer Loans and Leases Outstanding	FRED-MD	DTCTHFNM	5
81	Money and credit	Securities in Bank Credit at All Commercial Banks	FRED-MD	INVEST	5
82	Interest and exchange rates	Effective Federal Funds Rate	FRED-MD	FEDFUNDS	2
83	Interest and exchange rates	3-Month AA Financial Commercial Paper Rate	FRED-MD	CP3Mx	2
84	Interest and exchange rates	3-Month Treasury Bill	FRED-MD	TB3MS	2
85	Interest and exchange rates	6-Month Treasury Bill	FRED-MD	TB6MS	2
86	Interest and exchange rates	1-Year Treasury Rate	FRED-MD	GS1	2
87	Interest and exchange rates	5-Year Treasury Rate	FRED-MD	GS5	2
88	Interest and exchange rates	10-Year Treasury Rate	FRED-MD	GS10	2
89	Interest and exchange rates	Moody's Seasoned Aaa Corporate Bond Yield	FRED-MD	AAA	2
90	Interest and exchange rates	Moody's Seasoned Baa Corporate Bond Yield	FRED-MD	BAA	2
91	Interest and exchange rates	3-Month Commercial Paper Minus FEDFUNDS	FRED-MD	COMPAPFFx	1
92	Interest and exchange rates	3-Month Treasury C Minus FEDFUNDS	FRED-MD	TB3SMFFM	1
93	Interest and exchange rates	6-Month Treasury C Minus FEDFUNDS	FRED-MD	TB6SMFFM	1
94	Interest and exchange rates	1-Year Treasury C Minus FEDFUNDS	FRED-MD	T1YFFM	1
95	Interest and exchange rates	5-Year Treasury C Minus FEDFUNDS	FRED-MD	T5YFFM	1
96	Interest and exchange rates	10-Year Treasury C Minus FEDFUNDS	FRED-MD	T10YFFM	1
97	Interest and exchange rates	Moody's Aaa Corporate Bond Minus FEDFUNDS	FRED-MD	AAAFFM	1
98	Interest and exchange rates	Moody's Baa Corporate Bond Minus FEDFUNDS	FRED-MD	BAAFFM	1
99	Interest and exchange rates	Trade Weighted U.S. Dollar Index	FRED-MD	TWEXAFEGSMTHx	5
100	Interest and exchange rates	Switzerland / U.S. Foreign Exchange Rate	FRED-MD	EXSZUSx	5
101	Interest and exchange rates	Japan / U.S. Foreign Exchange Rate	FRED-MD	EXJPUSx	5
102	Interest and exchange rates	U.S. / U.K. Foreign Exchange Rate	FRED-MD	EXUSUKx	5
103	Interest and exchange rates	Canada / U.S. Foreign Exchange Rate	FRED-MD	EXCAUSx	5
104	Prices	PPI: Finished Goods	FRED-MD	WPSFD49207	5
105	Prices	PPI: Finished Consumer Goods	FRED-MD	WPSFD49502	5
106	Prices	PPI: Intermediate Materials	FRED-MD	WPSID61	5
107	Prices	PPI: Crude Materials	FRED-MD	WPSID62	5
108	Prices	Crude Oil, spliced WTI and Cushing	FRED-MD	OILPRICEx	5
109	Prices	PPI: Metals and metal products	FRED-MD	PPICMM	5
110	Prices	CPI : All Items	FRED-MD	CPIAUCSL	5
111	Prices	CPI : Apparel	FRED-MD	CPIAPPSL	5
112	Prices	CPI : Transportation	FRED-MD	CPITRNSL	5
113	Prices	CPI : Medical Care	FRED-MD	CPIMEDSL	5
114	Prices	CPI : Commodities	FRED-MD	CUSR0000SAC	5
115	Prices	CPI : Durables	FRED-MD	CUSR0000SAD	5
116	Prices	CPI : Services	FRED-MD	CUSR0000SAS	5
117	Prices	CPI : All Items Less Food	FRED-MD	CPIULFSL	5
118	Prices	CPI : All items less shelter	FRED-MD	CUSR0000SA0L2	5
119	Prices	CPI : All items less medical care	FRED-MD	CUSR0000SA0L5	5
120	Prices	Personal Cons. Expend.: Chain Index	FRED-MD	PCEPI	5
121	Prices	Personal Cons. Exp: Durable goods	FRED-MD	DDURRG3M086SBEA	5
122	Prices	Personal Cons. Exp: Nondurable goods	FRED-MD	DNDGRG3M086SBEA	5
123	Prices	Personal Cons. Exp: Services	FRED-MD	DSERRG3M086SBEA	5
124	Stock market	S&P's Common Stock Price Index: Composite	FRED-MD	S&P500	5
125	Stock market	S&P's Common Stock Price Index: Industrials	FRED-MD	S&P_indust	5
126	Stock market	S&P's Composite Common Stock: Dividend Yield	FRED-MD	S&P_div_yield	2
127	Stock market	S&P's Composite Common Stock: Price-Earnings Ratio	FRED-MD	S&P_PE_ratio	5
128	Stock market	CBOE S&P 100 Volatility Index: VXO	FRED-MD	VXOCLSx	1
129	Industrial Production	Production of Total Industry in Austria	FRED	AUTPROINDMISMEI	5
130	Industrial Production	Production of Total Industry in Belgium	FRED	BELPROINDMISMEI	5
131	Industrial Production	Production of Total Industry in Brazil	FRED	BRAPROINDMISMEI	5
132	Industrial Production	Production of Total Industry in Canada	FRED	CANPROINDMISMEI	5
133	Industrial Production	Production of Total Industry in Chile	FRED	CHLPROINDMISMEI	5
134	Industrial Production	Production of Total Industry in Czech Republic	FRED	CZEPROINDMISMEI	5

Note: The column "tcode" denotes the following series transformations:
(1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $\ln(x_t)$; (5) $\Delta \ln(x_t)$; (6) $\Delta^2 \ln(x_t)$.

Table A3 – Leading series of economic activity for the United States (cont.)

	Category	Name	Source	Nickname	tcode
135	Industrial Production	Production of Total Industry in Denmark	FRED	DNKPROINDMISMEI	5
136	Industrial Production	Production of Total Industry in Finland	FRED	FINPROINDMISMEI	5
137	Industrial Production	Production of Total Industry in France	FRED	FRAPROINDMISMEI	5
138	Industrial Production	Production of Total Industry in Germany	FRED	DEUPROINDMISMEI	5
139	Industrial Production	Production of Total Industry in Greece	FRED	GRCPROINDMISMEI	5
140	Industrial Production	Production of Total Industry in Hungary	FRED	HUNPROINDMISMEI	5
141	Industrial Production	Production of Total Industry in Ireland	FRED	IRLPROINDMISMEI	5
142	Industrial Production	Production of Total Industry in Israel	FRED	ISRPROINDMISMEI	5
143	Industrial Production	Production of Total Industry in Italy	FRED	ITAPROINDMISMEI	5
144	Industrial Production	Production of Total Industry in Japan	FRED	JPNPROINDMISMEI	5
145	Industrial Production	Production of Total Industry in Korea	FRED	KORPROINDMISMEI	5
146	Industrial Production	Production of Total Industry in Netherlands	FRED	NLDPROINDMISMEI	5
147	Industrial Production	Production of Total Industry in Norway	FRED	NORPROINDMISMEI	5
148	Industrial Production	Production of Total Industry in Poland	FRED	POLPROINDMISMEI	5
149	Industrial Production	Production of Total Industry in Portugal	FRED	PRTPROINDMISMEI	5
150	Industrial Production	Production of Total Industry in Russian Federation	FRED	RUSPROINDMISMEI	5
151	Industrial Production	Production of Total Industry in Spain	FRED	ESPPROINDMISMEI	5
152	Industrial Production	Production of Total Industry in Sweden	FRED	SWEPROINDMISMEI	5
153	Industrial Production	Production of Total Industry in Turkey	FRED	TURPROINDMISMEI	5
154	Industrial Production	Production of Total Industry in United Kingdom	FRED	GBRPROINDMISMEI	5
155	Industrial Production in the U.S.	Industrial Production: Durable Manufacturing	FRED	IPDMAN	5
156	Industrial Production in the U.S.	Industrial Production: Durable Goods: Iron and Steel Products	FRED	IPG3311A2S	5
157	Industrial Production in the U.S.	Industrial Production: Durable Goods: Alumina and Aluminum Production and Processing	FRED	IPG3313S	5
158	Industrial Production in the U.S.	Industrial Production: Durable Goods: Raw Steel	FRED	IPN3311A2RS	5
159	Industrial Production in the U.S.	Industrial Production: Durable Goods: Automotive Products	FRED	IPB51110S	5
160	Industrial Production in the U.S.	Industrial Production: Durable Goods: Cement and Concrete Product	FRED	IPG3273S	5
161	Industrial Production in the U.S.	Industrial Production: Durable Goods: Primary Metal	FRED	IPG331S	5
162	Industrial Production in the U.S.	Industrial Production: Durable Goods: Machinery	FRED	IPG333S	5
163	Industrial Production in the U.S.	Industrial Production: Durable Goods: Aerospace and Misc. Transportation Equipment	FRED	IPG3364T9S	5
164	Industrial Production in the U.S.	Industrial Production: Non-Durable Goods: Food	FRED	IPG311S	5
165	Industrial Production in the U.S.	Industrial Production: Non-Durable Goods: Petroleum and Coal Products	FRED	IPG324S	5
166	Industrial Production in the U.S.	Industrial Production: Non-Durable Goods: Chemical	FRED	IPG325S	5
167	Industrial Production in the U.S.	Industrial Production: Non-Durable Goods: Plastics and Rubber Products	FRED	IPG326S	5
168	Industrial Production in the U.S.	Industrial Production: Non-Durable Goods: Petroleum Refineries	FRED	IPG32411S	5
169	Industrial Production in the U.S.	Industrial Production: Non-Durable Goods: Pharmaceutical and Medicine	FRED	IPG3254S	5
170	Industrial Production in the U.S.	Industrial Production: Non-Durable Goods: Plastics Material and Resin	FRED	IPN325211S	5
171	Industrial Production in the U.S.	Industrial Production: Construction Supplies	FRED	IPB54100S	5
172	Industrial Production in the U.S.	Industrial Production: Non-Energy, Total	FRED	IPX5001ES	5
173	Industrial Production in the U.S.	Industrial Production: Energy, Total	FRED	IPB50089S	5
174	Industrial Production in the U.S.	Industrial Production: Electric and Gas Utilities	FRED	IPUTIL	5
175	Industrial Production in the U.S.	Industrial Production: Electric Power Generation, Transmission, and Distribution	FRED	IPG2211S	5
176	Industrial Production in the U.S.	Industrial Production: Mining: Crude Oil	FRED	IPG211111CS	5
177	Industrial Production in the U.S.	Industrial Production: Mining: Crude Petroleum and Natural Gas Extraction	FRED	IPG211111S	5
178	Industrial Production in the U.S.	Industrial Production: Mining: Oil and Gas Extraction	FRED	IPG211S	5
179	Industrial Production in the U.S.	Industrial Production: Mining: Copper, Nickel, Lead, and Zinc Mining	FRED	IPG21223S	5
180	Industrial Production in the U.S.	Industrial Production: Mining: Natural Gas	FRED	IPN211111GS	5
181	Industrial Production in the U.S.	Industrial Production: Mining: Coal Mining	FRED	IPN2121S	5
182	Industrial Production in the U.S.	Industrial Production: Mining: Iron Ore Mining	FRED	IPN21221S	5
183	Industrial Production in the U.S.	Industrial Production: Mining: Drilling Oil and Gas Wells	FRED	IPN213111S	5
184	Economic uncertainty	Policy-related economic uncertainty index for Australia	Economic Policy Uncertainty	EPU_Australia	1
185	Economic uncertainty	Policy-related economic uncertainty index for Brazil	Economic Policy Uncertainty	EPU_Brazil	1
186	Economic uncertainty	Policy-related economic uncertainty index for Canada	Economic Policy Uncertainty	EPU_Canada	1
187	Economic uncertainty	Policy-related economic uncertainty index for Chile	Economic Policy Uncertainty	EPU_Chile	1
188	Economic uncertainty	Policy-related economic uncertainty index for China	Economic Policy Uncertainty	EPU_China	1
189	Economic uncertainty	Policy-related economic uncertainty index for Colombia	Economic Policy Uncertainty	EPU_Colombia	1
190	Economic uncertainty	Policy-related economic uncertainty index for France	Economic Policy Uncertainty	EPU_France	1
191	Economic uncertainty	Policy-related economic uncertainty index for Germany	Economic Policy Uncertainty	EPU_Germany	1
192	Economic uncertainty	Policy-related economic uncertainty index for Greece	Economic Policy Uncertainty	EPU_Greece	1
193	Economic uncertainty	Policy-related economic uncertainty index for India	Economic Policy Uncertainty	EPU_India	1
194	Economic uncertainty	Policy-related economic uncertainty index for Ireland	Economic Policy Uncertainty	EPU_Ireland	1
195	Economic uncertainty	Policy-related economic uncertainty index for Italy	Economic Policy Uncertainty	EPU_Italy	1
196	Economic uncertainty	Policy-related economic uncertainty index for Japan	Economic Policy Uncertainty	EPU_Japan	1
197	Economic uncertainty	Policy-related economic uncertainty index for Korea	Economic Policy Uncertainty	EPU_Korea	1
198	Economic uncertainty	Policy-related economic uncertainty index for Netherlands	Economic Policy Uncertainty	EPU_Netherlands	1
199	Economic uncertainty	Policy-related economic uncertainty index for Russia	Economic Policy Uncertainty	EPU_Russia	1
200	Economic uncertainty	Policy-related economic uncertainty index for Spain	Economic Policy Uncertainty	EPU_Spain	1
201	Economic uncertainty	Policy-related economic uncertainty index for Singapore	Economic Policy Uncertainty	EPU_Singapore	1
202	Economic uncertainty	Policy-related economic uncertainty index for UK	Economic Policy Uncertainty	EPU_UK	1
203	Economic uncertainty	Policy-related economic uncertainty index for US	Economic Policy Uncertainty	EPU_US	1

Note: The column "tcode" denotes the following series transformations:
(1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $\ln(x_t)$; (5) $\Delta\ln(x_t)$; (6) $\Delta^2\ln(x_t)$.

Table A4 – Leading series of economic activity for the United States (cont.)

	Category	Name	Source	Nickname	tcode
204	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_ARGENTINA	1
205	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_BRAZIL	1
206	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_CHINA	1
207	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_COLOMBIA	1
208	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_HONG_KONG	1
209	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_INDIA	1
210	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_INDONESIA	1
211	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_ISRAEL	1
212	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_KOREA	1
213	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_MALAYSIA	1
214	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_MEXICO	1
215	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_PHILIPPINES	1
216	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_RUSSIA	1
217	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_SAUDI_ARABIA	1
218	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_SOUTH_AFRICA	1
219	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_THAILAND	1
220	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_TURKEY	1
221	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_UKRAINE	1
222	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_VENEZUELA	1
223	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR	1
224	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_THREAT	1
225	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_ACT	1
226	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_BROAD	1
227	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_NARROW	1
228	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPRH	1
229	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPRHT	1
230	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPRHA	1
231	Leading Indicator	OECD Composite Leading Indicator (CLI) for Australia	OECD	CLI_Australia	2
232	Leading Indicator	OECD Composite Leading Indicator (CLI) for Austria	OECD	CLI_Austria	2
233	Leading Indicator	OECD Composite Leading Indicator (CLI) for Belgium	OECD	CLI_Belgium	2
234	Leading Indicator	OECD Composite Leading Indicator (CLI) for Brazil	OECD	CLI_Brazil	2
235	Leading Indicator	OECD Composite Leading Indicator (CLI) for Canada	OECD	CLI_Canada	2
236	Leading Indicator	OECD Composite Leading Indicator (CLI) for Chile	OECD	CLI_Chile	2
237	Leading Indicator	OECD Composite Leading Indicator (CLI) for China	OECD	CLI_China	2
238	Leading Indicator	OECD Composite Leading Indicator (CLI) for Colombia	OECD	CLI_Colombia	2
239	Leading Indicator	OECD Composite Leading Indicator (CLI) for Czech Republic	OECD	CLI_Czech Republic	2
240	Leading Indicator	OECD Composite Leading Indicator (CLI) for Denmark	OECD	CLI_Denmark	2
241	Leading Indicator	OECD Composite Leading Indicator (CLI) for Estonia	OECD	CLI_Estonia	2
242	Leading Indicator	OECD Composite Leading Indicator (CLI) for Finland	OECD	CLI_Finland	2
243	Leading Indicator	OECD Composite Leading Indicator (CLI) for France	OECD	CLI_France	2
244	Leading Indicator	OECD Composite Leading Indicator (CLI) for Germany	OECD	CLI_Germany	2
245	Leading Indicator	OECD Composite Leading Indicator (CLI) for Greece	OECD	CLI_Greece	2
246	Leading Indicator	OECD Composite Leading Indicator (CLI) for Hungary	OECD	CLI_Hungary	2
247	Leading Indicator	OECD Composite Leading Indicator (CLI) for Iceland	OECD	CLI_Iceland	2
248	Leading Indicator	OECD Composite Leading Indicator (CLI) for India	OECD	CLI_India	2
249	Leading Indicator	OECD Composite Leading Indicator (CLI) for Indonesia	OECD	CLI_Indonesia	2
250	Leading Indicator	OECD Composite Leading Indicator (CLI) for Ireland	OECD	CLI_Ireland	2
251	Leading Indicator	OECD Composite Leading Indicator (CLI) for Israel	OECD	CLI_Israel	2
252	Leading Indicator	OECD Composite Leading Indicator (CLI) for Italy	OECD	CLI_Italy	2
253	Leading Indicator	OECD Composite Leading Indicator (CLI) for Japan	OECD	CLI_Japan	2
254	Leading Indicator	OECD Composite Leading Indicator (CLI) for Korea	OECD	CLI_Korea	2
255	Leading Indicator	OECD Composite Leading Indicator (CLI) for Mexico	OECD	CLI_Mexico	2
256	Leading Indicator	OECD Composite Leading Indicator (CLI) for Netherlands	OECD	CLI_Netherlands	2
257	Leading Indicator	OECD Composite Leading Indicator (CLI) for Norway	OECD	CLI_Norway	2
258	Leading Indicator	OECD Composite Leading Indicator (CLI) for Poland	OECD	CLI_Poland	2
259	Leading Indicator	OECD Composite Leading Indicator (CLI) for Portugal	OECD	CLI_Portugal	2
260	Leading Indicator	OECD Composite Leading Indicator (CLI) for Russia	OECD	CLI_Russia	2
261	Leading Indicator	OECD Composite Leading Indicator (CLI) for Slovak Republic	OECD	CLI_Slovak Republic	2
262	Leading Indicator	OECD Composite Leading Indicator (CLI) for Slovenia	OECD	CLI_Slovenia	2
263	Leading Indicator	OECD Composite Leading Indicator (CLI) for South_Africa	OECD	CLI_South_Africa	2
264	Leading Indicator	OECD Composite Leading Indicator (CLI) for Spain	OECD	CLI_Spain	2
265	Leading Indicator	OECD Composite Leading Indicator (CLI) for Sweden	OECD	CLI_Sweden	2
266	Leading Indicator	OECD Composite Leading Indicator (CLI) for Switzerland	OECD	CLI_Switzerland	2
267	Leading Indicator	OECD Composite Leading Indicator (CLI) for Turkey	OECD	CLI_Turkey	2
268	Leading Indicator	OECD Composite Leading Indicator (CLI) for United Kingdom	OECD	CLI_UK	2
269	Leading Indicator	OECD Composite Leading Indicator (CLI) for United States of America	OECD	CLI_USA	2
270	Leading Indicator	OECD Composite Leading Indicator (CLI) for Euro area (19 countries)	OECD	CLI_Euro_area	2
271	Leading Indicator	OECD Composite Leading Indicator (CLI) for Big four European	OECD	CLI_Big4_European	2

Note: The column "tcode" denotes the following series transformations:
(1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $\ln(x_t)$; (5) $\Delta \ln(x_t)$; (6) $\Delta^2 \ln(x_t)$.

Table A5 – Leading series of economic activity for the United States (cont.)

	Category	Name	Source	Nickname	tcode
272	Leading Indicator	OECD Composite Leading Indicator (CLI) for G7	OECD	CLI_G7	2
273	Leading Indicator	OECD Composite Leading Indicator (CLI) for NAFTA	OECD	CLI_NAFTA	2
274	Leading Indicator	OECD Composite Leading Indicator (CLI) for Major five Asia	OECD	CLI_Major5_Asia	2
275	Leading Indicator	OECD Composite Leading Indicator (CLI) for OECD Europe	OECD	CLI_OECD_Europe	2
276	Leading Indicator	OECD Composite Leading Indicator (CLI) for OECD Total	OECD	CLI_OECD_Total	2
277	Leading Indicator	OECD Composite Leading Indicator (CLI) for OECD Major six NME	OECD	CLI_OECD_Major6_NME	2
278	Real business conditions in the U.S.	Aruoba-Diebold-Scotti Business Conditions Index	Federal Reserve Bank of Philadelphia	ADS_index	1
279	Quantitative Easing	Total Assets (US$ trillions), Federal Reserve	Federal Reserve Bank of St. Louis	QE_FED	5
280	Quantitative Easing	Total Assets (US$ trillions), Federal Reserve + European Central Bank + Bank of Japan	Federal Reserve Bank of St. Louis	QE_FED_ECB_BOJ	5
281	Energy Outlook	Liquid Fuels Consumption, World (million barrels per day)	Short-Term Energy Outlook, U.S. EIA	STEO.PATC_WORLD.M	5
282	Energy Outlook	Liquid Fuels Consumption, OECD (million barrels per day)	Short-Term Energy Outlook, U.S. EIA	STEO.PATC_OECD.M	5
283	Energy Outlook	Liquid Fuels Consumption, non-OECD (million barrels per day)	Short-Term Energy Outlook, U.S. EIA	STEO.PATC_NON_OECD.M	5
284	Energy Outlook	Crude Oil Production Capacity, OPEC (million barrels per day)	Short-Term Energy Outlook, U.S. EIA	STEO.COPC_OPEC.M	5
285	Energy Outlook	Petroleum Product Supply, Total (million barrels per day)	Short-Term Energy Outlook, U.S. EIA	STEO.PASUPPLY.M	5
286	Energy Outlook	Crude Oil Production, U.S. (million barrels per day)	Short-Term Energy Outlook, U.S. EIA	STEO.COPRPUS.M	5
287	Energy Outlook	Crude Oil and Other Liquids Inventory, U.S. (million barrels)	Short-Term Energy Outlook, U.S. EIA	STEO.PASC_US.M	5
288	Energy Outlook	Petroleum Net Imports, U.S. (million barrels per day)	Short-Term Energy Outlook, U.S. EIA	STEO.PAIMPORT.M	5
289	Energy Outlook	Net Inventory Withdrawals, Crude Oil and Other Liquids, U.S. (million barrels per day)	Short-Term Energy Outlook, U.S. EIA	STEO.T3_STCHANGE_US.M	5
290	Energy Outlook	Natural Gas Henry Hub Spot Price, U.S. (dollars per thousand cubic feet)	Short-Term Energy Outlook, U.S. EIA	STEO.NGHHMCF.M	5
291	Energy Outlook	Cost of Coal Delivered to Electric Generating Plants, U.S. (dollars per million Btu)	Short-Term Energy Outlook, U.S. EIA	STEO.CLEUDUS.M	5
292	Energy Outlook	Coal Production, U.S. (million short tons)	Short-Term Energy Outlook, U.S. EIA	STEO.CLPRPUS_TON.M	5
293	Energy Outlook	Coal Consumption, U.S. (million short tons)	Short-Term Energy Outlook, U.S. EIA	STEO.CLTCPUS_TON.M	5
294	Energy Outlook	Consumption of Electricity, U.S. (billion kilowatthours)	Short-Term Energy Outlook, U.S. EIA	STEO.ELCOTWH.M	5
295	Energy Outlook	Raw Steel Production, U.S. (million short tons per day)	Short-Term Energy Outlook, U.S. EIA	STEO.RSPRPUS.M	5
296	Energy Outlook	Aircraft Utilization, U.S. (revenue ton-miles/day thousands)	Short-Term Energy Outlook, U.S. EIA	STEO.RMZZPUS.M	5
297	Energy Outlook	Vehicle Miles Traveled, U.S. (million miles/day)	Short-Term Energy Outlook, U.S. EIA	STEO.MVVMPUS.M	5
298	Financial markets	Baltic Exchange Dry Index (BDI)	Thomson Reuters	BALTIC_DRY	1
299	Financial markets	CBOE SPX VOLATILITY VIX	Thomson Reuters	VIX	1
300	Financial markets	US Dollar index DXY	Thomson Reuters	US_DOLLAR_INDEX	5
301	Financial markets	MSCI Emerging Markets U$	Thomson Reuters	MSCI_EM	5
302	Financial markets	MSCI World U$	Thomson Reuters	MSCI_WORLD	5
303	Financial markets	EURO STOXX 50	Thomson Reuters	EURO_STOXX50	5
304	Financial markets	S&P500 ES ENERGY	Thomson Reuters	SP500_ENERGY	5
305	Financial markets	S&P GSCI Energy Total Return - RETURN IND. (OFCL)	Thomson Reuters	SP_GSCI_ENERGY	5
306	Financial markets	CRB BLS Spot Index (1967=100)	Thomson Reuters	CRB	5
307	Financial markets	CRB BLS Spot Index Raw Industrials	Thomson Reuters	CRB_RAW_IND	5
308	Financial markets	CRB BLS Spot Index Metals	Thomson Reuters	CRB_METALS	5
309	Financial markets	CRB BLS Spot Index Foodstuffs	Thomson Reuters	CRB_FOOD	5
310	Financial markets	CRB BLS Spot Index Fats & Oils	Thomson Reuters	CRB_FATS	5
311	Financial markets	CRB BLS Spot Index Livestock	Thomson Reuters	CRB_LIVESTOCK	5
312	Financial markets	CRB BLS Spot Index Textiles	Thomson Reuters	CRB_TEXTI	5
313	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 1 month	Thomson Reuters	FUTURE_BRENT_M1	5
314	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 2 months	Thomson Reuters	FUTURE_BRENT_M2	5
315	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 3 months	Thomson Reuters	FUTURE_BRENT_M3	5
316	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 4 months	Thomson Reuters	FUTURE_BRENT_M4	5
317	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 5 months	Thomson Reuters	FUTURE_BRENT_M5	5
318	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 6 months	Thomson Reuters	FUTURE_BRENT_M6	5
319	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 7 months	Thomson Reuters	FUTURE_BRENT_M7	5
320	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 8 months	Thomson Reuters	FUTURE_BRENT_M8	5
321	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 9 months	Thomson Reuters	FUTURE_BRENT_M9	5
322	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 10 months	Thomson Reuters	FUTURE_BRENT_M10	5
323	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 11 months	Thomson Reuters	FUTURE_BRENT_M11	5
324	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 12 months	Thomson Reuters	FUTURE_BRENT_M12	5
325	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 24 months	Thomson Reuters	FUTURE_BRENT_M24	5
326	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 36 months	Thomson Reuters	FUTURE_BRENT_M36	5
327	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 48 months	Thomson Reuters	FUTURE_BRENT_M48	5
328	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 60 months	Thomson Reuters	FUTURE_BRENT_M60	5
329	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 72 months	Thomson Reuters	FUTURE_BRENT_M72	5

Note: The column "tcode" denotes the following series transformations:
(1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $\ln(x_t)$; (5) $\Delta\ln(x_t)$; (6) $\Delta^2\ln(x_t)$.

Table A6 – Leading series of economic activity for the Euro Area

	Category	Name	Source	Nickname	tcode
1	Output and Income	Real Personal Income	FRED-MD	RPI	5
2	Output and Income	IP Index	FRED-MD	INDPRO	5
3	Output and Income	IP: Consumer Goods	FRED-MD	IPCONGD	5
4	Output and Income	IP: Durable Consumer Goods	FRED-MD	IPDCONGD	5
5	Output and Income	IP: Nondurable Consumer Goods	FRED-MD	IPNCONGD	5
6	Output and Income	IP: Materials	FRED-MD	IPMAT	5
7	Output and Income	Capacity Utilization: Manufacturing	FRED-MD	CUMFNS	2
8	Labor market	Help-Wanted Index for United States	FRED-MD	HWI	2
9	Labor market	Civilian Labor Force	FRED-MD	CLF16OV	5
10	Labor market	Civilian Employment	FRED-MD	CE16OV	5
11	Labor market	Civilian Unemployment Rate	FRED-MD	UNRATE	2
12	Labor market	Average Duration of Unemployment (Weeks)	FRED-MD	UEMPMEAN	2
13	Labor market	Initial Claims	FRED-MD	CLAIMSx	5
14	Labor market	All Employees: Total nonfarm	FRED-MD	PAYEMS	5
15	Labor market	Avg Weekly Hours : Goods-Producing	FRED-MD	CES0600000007	1
16	Labor market	Avg Weekly Overtime Hours : Manufacturing	FRED-MD	AWOTMAN	2
17	Labor market	Avg Weekly Hours : Manufacturing	FRED-MD	AWHMAN	1
18	Labor market	Avg Hourly Earnings : Goods-Producing	FRED-MD	CES0600000008	5
19	Labor market	Avg Hourly Earnings : Construction	FRED-MD	CES2000000008	5
20	Labor market	Avg Hourly Earnings : Manufacturing	FRED-MD	CES3000000008	5
21	Housing	Housing Starts: Total New Privately Owned	FRED-MD	HOUST	4
22	Housing	New Private Housing Permits (SAAR)	FRED-MD	PERMIT	4
23	Consumption, orders, and inventories	Real personal consumption expenditures	FRED-MD	DPCERA3M086SBEA	5
24	Consumption, orders, and inventories	Real Manu. and Trade Industries Sales	FRED-MD	CMRMTSPLx	5
25	Consumption, orders, and inventories	Retail and Food Services Sales	FRED-MD	RETAILx	5
26	Consumption, orders, and inventories	New Orders for Consumer Goods	FRED-MD	ACOGNO	5
27	Consumption, orders, and inventories	New Orders for Durable Goods	FRED-MD	AMDMNOx	5
28	Consumption, orders, and inventories	Total Business Inventories	FRED-MD	BUSINVx	5
29	Consumption, orders, and inventories	Total Business: Inventories to Sales Ratio	FRED-MD	ISRATIOx	2
30	Consumption, orders, and inventories	Consumer Sentiment Index	FRED-MD	UMCSENTx	2
31	Money and credit	Commercial and Industrial Loans	FRED-MD	BUSLOANS	5
32	Money and credit	Real Estate Loans at All Commercial Banks	FRED-MD	REALLN	5
33	Money and credit	Total Consumer Loans and Leases Outstanding	FRED-MD	DTCTHFNM	5
34	Money and credit	Securities in Bank Credit at All Commercial Banks	FRED-MD	INVEST	5
35	Interest and exchange rates	Effective Federal Funds Rate	FRED-MD	FEDFUNDS	2
36	Interest and exchange rates	3-Month Treasury Bill	FRED-MD	TB3MS	2
37	Interest and exchange rates	1-Year Treasury Rate	FRED-MD	GS1	2
38	Interest and exchange rates	5-Year Treasury Rate	FRED-MD	GS5	2
39	Interest and exchange rates	10-Year Treasury Rate	FRED-MD	GS10	2
40	Interest and exchange rates	Moody's Seasoned Aaa Corporate Bond Yield	FRED-MD	AAA	2
41	Interest and exchange rates	Moody's Seasoned Baa Corporate Bond Yield	FRED-MD	BAA	2
42	Interest and exchange rates	Trade Weighted U.S. Dollar Index	FRED-MD	TWEXAFEGSMTHx	5
43	Interest and exchange rates	Switzerland / U.S. Foreign Exchange Rate	FRED-MD	EXSZUSx	5
44	Interest and exchange rates	Japan / U.S. Foreign Exchange Rate	FRED-MD	EXJPUSx	5
45	Interest and exchange rates	U.S. / U.K. Foreign Exchange Rate	FRED-MD	EXUSUKx	5
46	Interest and exchange rates	Canada / U.S. Foreign Exchange Rate	FRED-MD	EXCAUSx	5
47	Prices	PPI: Finished Goods	FRED-MD	WPSFD49207	5
48	Prices	PPI: Intermediate Materials	FRED-MD	WPSID61	5
49	Prices	PPI: Crude Materials	FRED-MD	WPSID62	5
50	Prices	Crude Oil, spliced WTI and Cushing	FRED-MD	OILPRICEx	5
51	Prices	PPI: Metals and metal products	FRED-MD	PPICMM	5
52	Prices	CPI : All Items	FRED-MD	CPIAUCSL	5
53	Prices	Personal Cons. Exp: Durable goods	FRED-MD	DDURRG3M086SBEA	5
54	Prices	Personal Cons. Exp: Nondurable goods	FRED-MD	DNDGRG3M086SBEA	5
55	Prices	Personal Cons. Exp: Services	FRED-MD	DSERRG3M086SBEA	5
56	Stock market	S&P's Common Stock Price Index: Composite	FRED-MD	S&P500	5
57	Stock market	S&P's Common Stock Price Index: Industrials	FRED-MD	S&P_indust	5
58	Stock market	S&P's Composite Common Stock: Dividend Yield	FRED-MD	S&P_div_yield	2
59	Stock market	S&P's Composite Common Stock: Price-Earnings Ratio	FRED-MD	S&P_PE_ratio	5
60	Stock market	CBOE S&P 100 Volatility Index: VXO	FRED-MD	VXOCLSx	1

Note: The column "tcode" denotes the following series transformations:
(1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $\ln(x_t)$; (5) $\Delta \ln(x_t)$; (6) $\Delta^2 \ln(x_t)$.

Table A7 – Leading series of economic activity for the Euro Area (cont.)

	Category	Name	Source	Nickname	tcode
61	Industrial Production	Production of Total Industry in Austria	FRED	AUTPROINDMISMEI	5
62	Industrial Production	Production of Total Industry in Belgium	FRED	BELPROINDMISMEI	5
63	Industrial Production	Production of Total Industry in Brazil	FRED	BRAPROINDMISMEI	5
64	Industrial Production	Production of Total Industry in Canada	FRED	CANPROINDMISMEI	5
65	Industrial Production	Production of Total Industry in Chile	FRED	CHLPROINDMISMEI	5
66	Industrial Production	Production of Total Industry in Czech Republic	FRED	CZEPROINDMISMEI	5
67	Industrial Production	Production of Total Industry in Denmark	FRED	DNKPROINDMISMEI	5
68	Industrial Production	Production of Total Industry in Finland	FRED	FINPROINDMISMEI	5
69	Industrial Production	Production of Total Industry in France	FRED	FRAPROINDMISMEI	5
70	Industrial Production	Production of Total Industry in Germany	FRED	DEUPROINDMISMEI	5
71	Industrial Production	Production of Total Industry in Greece	FRED	GRCPROINDMISMEI	5
72	Industrial Production	Production of Total Industry in Hungary	FRED	HUNPROINDMISMEI	5
73	Industrial Production	Production of Total Industry in Ireland	FRED	IRLPROINDMISMEI	5
74	Industrial Production	Production of Total Industry in Israel	FRED	ISRPROINDMISMEI	5
75	Industrial Production	Production of Total Industry in Italy	FRED	ITAPROINDMISMEI	5
76	Industrial Production	Production of Total Industry in Japan	FRED	JPNPROINDMISMEI	5
77	Industrial Production	Production of Total Industry in Korea	FRED	KORPROINDMISMEI	5
78	Industrial Production	Production of Total Industry in Netherlands	FRED	NLDPROINDMISMEI	5
79	Industrial Production	Production of Total Industry in Norway	FRED	NORPROINDMISMEI	5
80	Industrial Production	Production of Total Industry in Poland	FRED	POLPROINDMISMEI	5
81	Industrial Production	Production of Total Industry in Portugal	FRED	PRTPROINDMISMEI	5
82	Industrial Production	Production of Total Industry in Russian Federation	FRED	RUSPROINDMISMEI	5
83	Industrial Production	Production of Total Industry in Spain	FRED	ESPPROINDMISMEI	5
84	Industrial Production	Production of Total Industry in Sweden	FRED	SWEPROINDMISMEI	5
85	Industrial Production	Production of Total Industry in Turkey	FRED	TURPROINDMISMEI	5
86	Industrial Production	Production of Total Industry in United Kingdom	FRED	GBRPROINDMISMEI	5
87	Economic uncertainty	Policy-related economic uncertainty index for Australia	Economic Policy Uncertainty	EPU_Australia	1
88	Economic uncertainty	Policy-related economic uncertainty index for Brazil	Economic Policy Uncertainty	EPU_Brazil	1
89	Economic uncertainty	Policy-related economic uncertainty index for Canada	Economic Policy Uncertainty	EPU_Canada	1
90	Economic uncertainty	Policy-related economic uncertainty index for Chile	Economic Policy Uncertainty	EPU_Chile	1
91	Economic uncertainty	Policy-related economic uncertainty index for China	Economic Policy Uncertainty	EPU_China	1
92	Economic uncertainty	Policy-related economic uncertainty index for Colombia	Economic Policy Uncertainty	EPU_Colombia	1
93	Economic uncertainty	Policy-related economic uncertainty index for France	Economic Policy Uncertainty	EPU_France	1
94	Economic uncertainty	Policy-related economic uncertainty index for Germany	Economic Policy Uncertainty	EPU_Germany	1
95	Economic uncertainty	Policy-related economic uncertainty index for Greece	Economic Policy Uncertainty	EPU_Greece	1
96	Economic uncertainty	Policy-related economic uncertainty index for India	Economic Policy Uncertainty	EPU_India	1
97	Economic uncertainty	Policy-related economic uncertainty index for Ireland	Economic Policy Uncertainty	EPU_Ireland	1
98	Economic uncertainty	Policy-related economic uncertainty index for Italy	Economic Policy Uncertainty	EPU_Italy	1
99	Economic uncertainty	Policy-related economic uncertainty index for Japan	Economic Policy Uncertainty	EPU_Japan	1
100	Economic uncertainty	Policy-related economic uncertainty index for Korea	Economic Policy Uncertainty	EPU_Korea	1
101	Economic uncertainty	Policy-related economic uncertainty index for Netherlands	Economic Policy Uncertainty	EPU_Netherlands	1
102	Economic uncertainty	Policy-related economic uncertainty index for Russia	Economic Policy Uncertainty	EPU_Russia	1
103	Economic uncertainty	Policy-related economic uncertainty index for Spain	Economic Policy Uncertainty	EPU_Spain	1
104	Economic uncertainty	Policy-related economic uncertainty index for Singapore	Economic Policy Uncertainty	EPU_Singapore	1
105	Economic uncertainty	Policy-related economic uncertainty index for UK	Economic Policy Uncertainty	EPU_UK	1
106	Economic uncertainty	Policy-related economic uncertainty index for US	Economic Policy Uncertainty	EPU_US	1
107	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_ARGENTINA	1
108	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_BRAZIL	1
109	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_CHINA	1
110	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_COLOMBIA	1
111	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_HONG_KONG	1
112	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_INDIA	1
113	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_INDONESIA	1
114	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_ISRAEL	1
115	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_KOREA	1
116	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_MALAYSIA	1
117	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_MEXICO	1
118	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_PHILIPPINES	1
119	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_RUSSIA	1
120	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_SAUDI_ARABIA	1

Note: The column "tcode" denotes the following series transformations:
(1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $\ln(x_t)$; (5) $\Delta\ln(x_t)$; (6) $\Delta^2\ln(x_t)$.

Table A8 – Leading series of economic activity for the Euro Area (cont.)

	Category	Name	Source	Nickname	tcode
121	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_SOUTH_AFRICA	1
122	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_THAILAND	1
123	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_TURKEY	1
124	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_UKRAINE	1
125	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_VENEZUELA	1
126	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR	1
127	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_THREAT	1
128	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_ACT	1
129	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_BROAD	1
130	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPR_NARROW	1
131	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPRH	1
132	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPRHT	1
133	Economic uncertainty	Geopolitical Risk Index of Caldara and Iacoviello	Geopolitical Risk	GPRHA	1
134	Leading Indicator	OECD Composite Leading Indicator (CLI) for Australia	OECD	CLI_Australia	2
135	Leading Indicator	OECD Composite Leading Indicator (CLI) for Austria	OECD	CLI_Austria	2
136	Leading Indicator	OECD Composite Leading Indicator (CLI) for Belgium	OECD	CLI_Belgium	2
137	Leading Indicator	OECD Composite Leading Indicator (CLI) for Brazil	OECD	CLI_Brazil	2
138	Leading Indicator	OECD Composite Leading Indicator (CLI) for Canada	OECD	CLI_Canada	2
139	Leading Indicator	OECD Composite Leading Indicator (CLI) for Chile	OECD	CLI_Chile	2
140	Leading Indicator	OECD Composite Leading Indicator (CLI) for China	OECD	CLI_China	2
141	Leading Indicator	OECD Composite Leading Indicator (CLI) for Colombia	OECD	CLI_Colombia	2
142	Leading Indicator	OECD Composite Leading Indicator (CLI) for Czech Republic	OECD	CLI_Czech Republic	2
143	Leading Indicator	OECD Composite Leading Indicator (CLI) for Denmark	OECD	CLI_Denmark	2
144	Leading Indicator	OECD Composite Leading Indicator (CLI) for Estonia	OECD	CLI_Estonia	2
145	Leading Indicator	OECD Composite Leading Indicator (CLI) for Finland	OECD	CLI_Finland	2
146	Leading Indicator	OECD Composite Leading Indicator (CLI) for France	OECD	CLI_France	2
147	Leading Indicator	OECD Composite Leading Indicator (CLI) for Germany	OECD	CLI_Germany	2
148	Leading Indicator	OECD Composite Leading Indicator (CLI) for Greece	OECD	CLI_Greece	2
149	Leading Indicator	OECD Composite Leading Indicator (CLI) for Hungary	OECD	CLI_Hungary	2
150	Leading Indicator	OECD Composite Leading Indicator (CLI) for Iceland	OECD	CLI_Iceland	2
151	Leading Indicator	OECD Composite Leading Indicator (CLI) for India	OECD	CLI_India	2
152	Leading Indicator	OECD Composite Leading Indicator (CLI) for Indonesia	OECD	CLI_Indonesia	2
153	Leading Indicator	OECD Composite Leading Indicator (CLI) for Ireland	OECD	CLI_Ireland	2
154	Leading Indicator	OECD Composite Leading Indicator (CLI) for Israel	OECD	CLI_Israel	2
155	Leading Indicator	OECD Composite Leading Indicator (CLI) for Italy	OECD	CLI_Italy	2
156	Leading Indicator	OECD Composite Leading Indicator (CLI) for Japan	OECD	CLI_Japan	2
157	Leading Indicator	OECD Composite Leading Indicator (CLI) for Korea	OECD	CLI_Korea	2
158	Leading Indicator	OECD Composite Leading Indicator (CLI) for Mexico	OECD	CLI_Mexico	2
159	Leading Indicator	OECD Composite Leading Indicator (CLI) for Netherlands	OECD	CLI_Netherlands	2
160	Leading Indicator	OECD Composite Leading Indicator (CLI) for Norway	OECD	CLI_Norway	2
161	Leading Indicator	OECD Composite Leading Indicator (CLI) for Poland	OECD	CLI_Poland	2
162	Leading Indicator	OECD Composite Leading Indicator (CLI) for Portugal	OECD	CLI_Portugal	2
163	Leading Indicator	OECD Composite Leading Indicator (CLI) for Russia	OECD	CLI_Russia	2
164	Leading Indicator	OECD Composite Leading Indicator (CLI) for Slovak Republic	OECD	CLI_Slovak Republic	2
165	Leading Indicator	OECD Composite Leading Indicator (CLI) for Slovenia	OECD	CLI_Slovenia	2
166	Leading Indicator	OECD Composite Leading Indicator (CLI) for South_Africa	OECD	CLI_South_Africa	2
167	Leading Indicator	OECD Composite Leading Indicator (CLI) for Spain	OECD	CLI_Spain	2
168	Leading Indicator	OECD Composite Leading Indicator (CLI) for Sweden	OECD	CLI_Sweden	2
169	Leading Indicator	OECD Composite Leading Indicator (CLI) for Switzerland	OECD	CLI_Switzerland	2
170	Leading Indicator	OECD Composite Leading Indicator (CLI) for Turkey	OECD	CLI_Turkey	2
171	Leading Indicator	OECD Composite Leading Indicator (CLI) for United Kingdom	OECD	CLI_UK	2
172	Leading Indicator	OECD Composite Leading Indicator (CLI) for United States of America	OECD	CLI_USA	2
173	Leading Indicator	OECD Composite Leading Indicator (CLI) for Euro area (19 countries)	OECD	CLI_Euro_area	2
174	Leading Indicator	OECD Composite Leading Indicator (CLI) for Big four European	OECD	CLI_Big4_European	2
175	Leading Indicator	OECD Composite Leading Indicator (CLI) for G7	OECD	CLI_G7	2
176	Leading Indicator	OECD Composite Leading Indicator (CLI) for NAFTA	OECD	CLI_NAFTA	2
177	Leading Indicator	OECD Composite Leading Indicator (CLI) for Major five Asia	OECD	CLI_Major5_Asia	2
178	Leading Indicator	OECD Composite Leading Indicator (CLI) for OECD Europe	OECD	CLI_OECD_Europe	2
179	Leading Indicator	OECD Composite Leading Indicator (CLI) for OECD Total	OECD	CLI_OECD_Total	2
180	Leading Indicator	OECD Composite Leading Indicator (CLI) for OECD Major six NME	OECD	CLI_OECD_Major6_NME	2

Note: The column "tcode" denotes the following series transformations:
(1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $\ln(x_t)$; (5) $\Delta \ln(x_t)$; (6) $\Delta^2 \ln(x_t)$.

Table A9 – Leading series of economic activity for the Euro Area (cont.)

	Category	Name	Source	Nickname	tcode
181	Real business conditions in the U.S.	Aruoba-Diebold-Scotti Business Conditions Index (U.S.)	Federal Reserve Bank of Philadelphia	ADS_index	1
182	Quantitative Easing	Total Assets (US$ trillions), Federal Reserve	Federal Reserve Bank of St. Louis	QE_FED	5
183	Quantitative Easing	Total Assets (US$ trillions), Federal Reserve + European Central Bank + Bank of Japan	Federal Reserve Bank of St. Louis	QE_FED_ECB_BOJ	5
184	Energy Outlook	Liquid Fuels Consumption, World (million barrels per day)	Short-Term Energy Outlook, U.S. EIA	STEO.PATC_WORLD.M	5
185	Energy Outlook	Liquid Fuels Consumption, OECD (million barrels per day)	Short-Term Energy Outlook, U.S. EIA	STEO.PATC_OECD.M	5
186	Energy Outlook	Liquid Fuels Consumption, non-OECD (million barrels per day)	Short-Term Energy Outlook, U.S. EIA	STEO.PATC_NON_OECD.M	5
187	Energy Outlook	Crude Oil Production Capacity, OPEC (million barrels per day)	Short-Term Energy Outlook, U.S. EIA	STEO.COPC_OPEC.M	5
188	Energy Outlook	Petroleum Product Supply, Total (million barrels per day)	Short-Term Energy Outlook, U.S. EIA	STEO.PASUPPLY.M	5
189	Financial markets	Baltic Exchange Dry Index (BDI)	Thomson Reuters	BALTIC_DRY	1
190	Financial markets	CBOE SPX VOLATILITY VIX	Thomson Reuters	VIX	1
191	Financial markets	US Dollar index DXY	Thomson Reuters	US_DOLLAR_INDEX	5
192	Financial markets	MSCI Emerging Markets U$	Thomson Reuters	MSCI_EM	5
193	Financial markets	MSCI World U$	Thomson Reuters	MSCI_WORLD	5
194	Financial markets	EURO STOXX 50	Thomson Reuters	EURO_STOXX50	5
195	Financial markets	S&P500 ES ENERGY	Thomson Reuters	SP500_ENERGY	5
196	Financial markets	S&P GSCI Energy Total Return - RETURN IND. (OFCL)	Thomson Reuters	SP_GSCI_ENERGY	5
197	Financial markets	CRB BLS Spot Index (1967=100)	Thomson Reuters	CRB	5
198	Financial markets	CRB BLS Spot Index Raw Industrials	Thomson Reuters	CRB_RAW_IND	5
199	Financial markets	CRB BLS Spot Index Metals	Thomson Reuters	CRB_METALS	5
200	Financial markets	CRB BLS Spot Index Foodstuffs	Thomson Reuters	CRB_FOOD	5
201	Financial markets	CRB BLS Spot Index Fats & Oils	Thomson Reuters	CRB_FATS	5
202	Financial markets	CRB BLS Spot Index Livestock	Thomson Reuters	CRB_LIVESTOCK	5
203	Financial markets	CRB BLS Spot Index Textiles	Thomson Reuters	CRB_TEXTI	5
204	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 1 month	Thomson Reuters	FUTURE_BRENT_M1	5
205	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 2 months	Thomson Reuters	FUTURE_BRENT_M2	5
206	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 3 months	Thomson Reuters	FUTURE_BRENT_M3	5
207	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 4 months	Thomson Reuters	FUTURE_BRENT_M4	5
208	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 5 months	Thomson Reuters	FUTURE_BRENT_M5	5
209	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 6 months	Thomson Reuters	FUTURE_BRENT_M6	5
210	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 7 months	Thomson Reuters	FUTURE_BRENT_M7	5
211	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 8 months	Thomson Reuters	FUTURE_BRENT_M8	5
212	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 9 months	Thomson Reuters	FUTURE_BRENT_M9	5
213	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 10 months	Thomson Reuters	FUTURE_BRENT_M10	5
214	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 11 months	Thomson Reuters	FUTURE_BRENT_M11	5
215	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 12 months	Thomson Reuters	FUTURE_BRENT_M12	5
216	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 24 months	Thomson Reuters	FUTURE_BRENT_M24	5
217	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 36 months	Thomson Reuters	FUTURE_BRENT_M36	5
218	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 48 months	Thomson Reuters	FUTURE_BRENT_M48	5
219	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 60 months	Thomson Reuters	FUTURE_BRENT_M60	5
220	Financial markets	Futures Brent crude oil, Intercontinental Exchange (ICE), 72 months	Thomson Reuters	FUTURE_BRENT_M72	5

Note: The column "tcode" denotes the following series transformations:
(1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $\ln(x_t)$; (5) $\Delta\ln(x_t)$; (6) $\Delta^2\ln(x_t)$.

Table A10 – Coincident series of economic activity for the Euro Area

	Category	Name	Source	Nickname	tcode
1	Labor market	Standardised unemployment, Total (all ages), Female, Percentage	ECB	RTD.M.S0.S.L_FETOT.F	2
2	Labor market	Standardised unemployment, Total (all ages), Female, Persons	ECB	RTD.M.S0.S.L_FETOT.M	5
3	Labor market	Standardised unemployment, Age 25 and over, Total (male and female), Percentage	ECB	RTD.M.S0.S.L_FMO25.F	2
4	Labor market	Standardised unemployment, Age 25 and over, Total (male and female), Persons	ECB	RTD.M.S0.S.L_FMO25.M	5
5	Labor market	Standardised unemployment, Age under 25, Total (male and female), Percentage	ECB	RTD.M.S0.S.L_FMU25.F	2
6	Labor market	Standardised unemployment, Age under 25, Total (male and female), Persons	ECB	RTD.M.S0.S.L_FMU25.M	5
7	Labor market	Standardised unemployment, Total (all ages), Male, Percentage	ECB	RTD.M.S0.S.L_MATOT.F	2
8	Labor market	Standardised unemployment, Total (all ages), Male, Persons	ECB	RTD.M.S0.S.L_MATOT.M	5
9	Labor market	Standardised unemployment, Total (all ages), Total (male and female), Percentage	ECB	RTD.M.S0.S.L_UNETO.F	2
10	Labor market	Standardised unemployment, Total (all ages), Total (male and female), Persons	ECB	RTD.M.S0.S.L_UNETO.M	5
11	Trade Balance	Trade - Import, Capital goods - Value, Euro	ECB	RTD.M.S0.S.T_MCAP.E	5
12	Trade Balance	Trade - Import, Capital goods - Unit value, Index	ECB	RTD.M.S0.S.T_MCAP_U.X	5
13	Trade Balance	Trade - Import, Capital goods - Volume, Index	ECB	RTD.M.S0.S.T_MCAP_V.X	5
14	Trade Balance	Trade - Import, Consumption goods (consumer goods and cars and petrol), Euro	ECB	RTD.M.S0.S.T_MCOM.E	5
15	Trade Balance	Trade - Import, Consumption goods (consumer goods and cars and petrol) - Unit value, Index	ECB	RTD.M.S0.S.T_MCOM_U.X	5
16	Trade Balance	Trade - Import, Consumption goods (consumer goods and cars and petrol) - Volume, Index	ECB	RTD.M.S0.S.T_MCOM_V.X	5
17	Trade Balance	Trade - Import, Intermediate goods - Value, Euro	ECB	RTD.M.S0.S.T_MINT.E	5
18	Trade Balance	Trade - Import, Intermediate goods - Unit value, Index	ECB	RTD.M.S0.S.T_MINT_U.X	5
19	Trade Balance	Trade - Import, Intermediate goods - Volume, Index	ECB	RTD.M.S0.S.T_MINT_V.X	5
20	Trade Balance	Trade - Import, Manufactured products - Value, Euro	ECB	RTD.M.S0.S.T_MMAN.E	5
21	Trade Balance	Trade - Import, Manufactured products - Unit value, Index	ECB	RTD.M.S0.S.T_MMAN_U.X	5
22	Trade Balance	Trade - Import, Manufactured products - Volume, Index	ECB	RTD.M.S0.S.T_MMAN_V.X	5
23	Trade Balance	Trade - Import, Petroleum, petroleum products and related materials - Value, Euro	ECB	RTD.M.S0.S.T_MOIL.E	5
24	Trade Balance	Trade - Import, Petroleum, petroleum products and related materials - Unit value, Index	ECB	RTD.M.S0.S.T_MOIL_U.X	5
25	Trade Balance	Trade - Import, Petroleum, petroleum products and related materials - Volume, Index	ECB	RTD.M.S0.S.T_MOIL_V.X	5
26	Trade Balance	Trade - Import, Total - Value, Euro	ECB	RTD.M.S0.S.T_MTTT.E	5
27	Trade Balance	Trade - Import, Total - Unit value, Index	ECB	RTD.M.S0.S.T_MTTT_U.X	5
28	Trade Balance	Trade - Import, Total - Volume, Index	ECB	RTD.M.S0.S.T_MTTT_V.X	5
29	Trade Balance	Trade - Export, Capital goods - Value, Euro	ECB	RTD.M.S0.S.T_XCAP.E	5
30	Trade Balance	Trade - Export, Capital goods - Unit value, Index	ECB	RTD.M.S0.S.T_XCAP_U.X	5
31	Trade Balance	Trade - Export, Capital goods - Volume, Index	ECB	RTD.M.S0.S.T_XCAP_V.X	5
32	Trade Balance	Trade - Export, Consumption goods (consumer goods and cars and petrol) - Value, Euro	ECB	RTD.M.S0.S.T_XCOM.E	5
33	Trade Balance	Trade - Export, Consumption goods (consumer goods and cars and petrol) - Unit value, Index	ECB	RTD.M.S0.S.T_XCOM_U.X	5
34	Trade Balance	Trade - Export, Consumption goods (consumer goods and cars and petrol) - Volume, Index	ECB	RTD.M.S0.S.T_XCOM_V.X	5
35	Trade Balance	Trade - Export, Intermediate goods - Value, Euro	ECB	RTD.M.S0.S.T_XINT.E	5
36	Trade Balance	Trade - Export, Intermediate goods - Unit value, Index	ECB	RTD.M.S0.S.T_XINT_U.X	5
37	Trade Balance	Trade - Export, Intermediate goods - Volume, Index	ECB	RTD.M.S0.S.T_XINT_V.X	5
38	Trade Balance	Trade - Export, Manufactured products - Value, Euro	ECB	RTD.M.S0.S.T_XMAN.E	5
39	Trade Balance	Trade - Export, Manufactured products - Unit value, Index	ECB	RTD.M.S0.S.T_XMAN_U.X	5
40	Trade Balance	Trade - Export, Manufactured products - Volume, Index	ECB	RTD.M.S0.S.T_XMAN_V.X	5
41	Trade Balance	Trade - Export, Total - Value, Euro	ECB	RTD.M.S0.S.T_XTTT.E	5
42	Trade Balance	Trade - Export, Total - Unit value, Index	ECB	RTD.M.S0.S.T_XTTT_U.X	5
43	Trade Balance	Trade - Export, Total - Volume, Index	ECB	RTD.M.S0.S.T_XTTT_V.X	5
44	Survey of Expectations	Construction Survey - Assessment of order books, Percentage	ECB	RTD.M.S0.S.Y_COAOB.F	2
45	Survey of Expectations	Construction Survey - Construction Confidence Indicator, Percentage	ECB	RTD.M.S0.S.Y_COCCI.F	2
46	Survey of Expectations	Construction Survey - Employment expectations for the months ahead, Percentage	ECB	RTD.M.S0.S.Y_COEEX.F	2
47	Survey of Expectations	Consumer Survey - Consumer Confidence Indicator, Percentage	ECB	RTD.M.S0.S.Y_CSCCI.F	2
48	Survey of Expectations	Consumer Survey - Financial situation over next 12 months, Percentage	ECB	RTD.M.S0.S.Y_CSF12.F	2
49	Survey of Expectations	Consumer Survey - General economic situation over next 12 months, Percentage	ECB	RTD.M.S0.S.Y_CSG12.F	2
50	Survey of Expectations	Consumer Survey - Savings over next 12 months, Percentage	ECB	RTD.M.S0.S.Y_CSS12.F	2
51	Survey of Expectations	Consumer Survey - Unemployment expectations over next 12 months	ECB	RTD.M.S0.S.Y_CSU12.F	2
52	Survey of Expectations	Economic Sentiment Indicator	ECB	RTD.M.S0.S.Y_ESIND.F	2
53	Survey of Expectations	Industry Survey - Assessment of order-book levels	ECB	RTD.M.S0.S.Y_ISAOB.F	2
54	Survey of Expectations	Industry Survey - Assessment of stocks of finished products	ECB	RTD.M.S0.S.Y_ISASP.F	2
55	Survey of Expectations	Industry Survey - Industrial Confidence Indicator	ECB	RTD.M.S0.S.Y_ISICI.F	2
56	Survey of Expectations	Industry Survey - Production expectations for the months ahead	ECB	RTD.M.S0.S.Y_ISPEX.F	2
57	Survey of Expectations	Retail Trade Survey - Assessment of stocks	ECB	RTD.M.S0.S.Y_RSAOS.F	2
58	Survey of Expectations	Retail Trade Survey - Expected business situation	ECB	RTD.M.S0.S.Y_RSBEX.F	2
59	Survey of Expectations	Retail Trade Survey - Present business situation	ECB	RTD.M.S0.S.Y_RSPBS.F	2
60	Survey of Expectations	Retail Trade Survey - Retail Confidence Indicator	ECB	RTD.M.S0.S.Y_RSRCI.F	2
61	Survey of Expectations	Service Survey - Assessment of the business climate	ECB	RTD.M.S0.S.Y_SSABC.F	2
62	Survey of Expectations	Service Survey - Evolution of demand expected in the months ahead	ECB	RTD.M.S0.S.Y_SSEDE.F	2
63	Survey of Expectations	Service Survey - Evolution of demand in recent months	ECB	RTD.M.S0.S.Y_SSEDR.F	2
64	Survey of Expectations	Service Survey - Service Confidence Indicator	ECB	RTD.M.S0.S.Y_SSSCI.F	2

Note: The column "tcode" denotes the following series transformations:
(1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $\ln(x_t)$; (5) $\Delta\ln(x_t)$; (6) $\Delta^2\ln(x_t)$.

Table A11 – Coincident series of economic activity for the Euro Area (cont.)

	Category	Name	Source	Nickname	tcode
65	Survey of Expectations	Business Tendency Surveys for Construction: Confidence Indicators	FRED	BCCICP02EZM460S	2
66	Survey of Expectations	Business Tendency Surveys for Construction: Employment	FRED	BCEMFT02EZM460S	2
67	Survey of Expectations	Business Tendency Surveys for Construction: Selling Prices	FRED	BCSPFT02EZM460S	2
68	Survey of Expectations	Business Tendency Surveys for Manufacturing: Confidence Indicators	FRED	BSCICP03EZM665S	2
69	Survey of Expectations	Business Tendency Surveys for Manufacturing: Employment	FRED	BSEMFT02EZM460S	2
70	Survey of Expectations	Business Tendency Surveys for Manufacturing: Finished Goods Stocks	FRED	BSFGLV02EZM460S	2
71	Survey of Expectations	Business Tendency Surveys for Manufacturing: Order Books	FRED	BSOBLV02EZM460S	2
72	Survey of Expectations	Business Tendency Surveys for Manufacturing: Production	FRED	BSPRFT02EZM460S	2
73	Survey of Expectations	Business Tendency Surveys for Manufacturing: Production: Tendency	FRED	BSPRTE02EZM460S	2
74	Survey of Expectations	Business Tendency Surveys for Manufacturing: Selling Prices	FRED	BSSPFT02EZM460S	2
75	Survey of Expectations	Business Tendency Surveys for Manufacturing: Export Order Books or Demand	FRED	BSXRLV02EZM086S	2
76	Survey of Expectations	Business Tendency Surveys for Services: Confidence Indicators	FRED	BVCICP02EZM460S	2
77	Survey of Expectations	Business Tendency Surveys for Services: Demand Evolution	FRED	BVDEFT02EZM460S	2
78	Survey of Expectations	Business Tendency Surveys for Services: Demand Evolution: Tendency	FRED	BVDETE02EZM460S	2
79	Survey of Expectations	Business Tendency Surveys for Services: Employment: Tendency	FRED	BVEMTE02EZM460S	2
80	Survey of Expectations	Business tendency surveys (construction): Business situation - activity	FRED	EA19BCBUTE02STSAM	2
81	Survey of Expectations	Business tendency surveys (retail trade): Confidence indicators	FRED	EA19BRCICP02STSAM	2
82	Survey of Expectations	Business tendency surveys (retail trade): Employment: Future tendency	FRED	EA19BREMFT02STSAM	2
83	Survey of Expectations	Business tendency surveys (retail trade): Order intentions or Demand	FRED	EA19BRODFT02STSAM	2
84	Survey of Expectations	Business tendency surveys (retail trade): Volume of stocks	FRED	EA19BRVSLV02STSAM	2
85	Economic Activity	Leading Indicators OECD: CLI: Amplitude adjusted for the Euro Area	FRED	EA19LOLITOAASTSAM	2
86	Economic Activity	Leading Indicators OECD: CLI: Trend restored for the Euro Area	FRED	EA19LOLITOTRGYSAM	2
87	Economic Activity	Leading Indicators OECD: CLI: Trend restored for the Euro Area (index)	FRED	EA19LOLITOTRSTSAM	2
88	Economic Activity	Leading Indicators OECD: Reference series: Gross Domestic Product	FRED	EA19LORSGPRTSTSAM	2
89	Production	Production: Construction: Total for the Euro Area (growth rate, previous period))	FRED	EA19PRCNTO01GPSAM	1
90	Production	Production: Construction: Total for the Euro Area (growth rate, previous year)	FRED	EA19PRCNTO01GYSAM	2
91	Production	Production: Construction: Total for the Euro Area (index)	FRED	EA19PRCNTO01IXOBSAM	5
92	Production	Production: Manufacturing: Consumer goods: Durable goods for the Euro Area	FRED	EA19PRMNCG02IXOBSAM	5
93	Production	Production: Manufacturing: Consumer goods: Non durable goods for the Euro Area	FRED	EA19PRMNCG03IXOBSAM	5
94	Production	Production: Manufacturing: Intermediate goods: Total for the Euro Area	FRED	EA19PRMNIG01IXOBSAM	5
95	Production	Production: Manufacturing: Investment goods: Total for the Euro Area	FRED	EA19PRMNVG01IXOBSAM	5
96	Sales	Sales: Manufacturing: Consumer goods durable: Value for the Euro Area	FRED	EA19SLMNCD02IXOBSAM	5
97	Sales	Sales: Manufacturing: Consumer goods non durable: Value for the Euro Area	FRED	EA19SLMNCN02IXOBSAM	5
98	Sales	Sales: Manufacturing: Intermediate goods: Value for the Euro Area	FRED	EA19SLMNIG02IXOBSAM	5
99	Sales	Sales: Manufacturing: Total manufacturing: Value for the Euro Area	FRED	EA19SLMNTO02IXOBSAM	5
100	Sales	Sales: Manufacturing: Investment goods: Value for the Euro Area	FRED	EA19SLMNVG02IXOBSAM	5
101	Labor market	Harmonized Unemployment: Aged 15-24: Females for the Euro Area, Persons	FRED	LFHU24FEEZM647S	5
102	Labor market	Harmonized Unemployment: Aged 15-24: Males for the Euro Area, Persons	FRED	LFHU24MAEZM647S	5
103	Labor market	Harmonized Unemployment: Aged 15-24: All Persons for the Euro Area, Persons	FRED	LFHU24TTEZM647S	5
104	Labor market	Harmonized Unemployment: Aged 25 and Over: Females for the Euro Area, Persons	FRED	LFHUADFEEZM647S	5
105	Labor market	Harmonized Unemployment: Aged 25 and Over: All Persons for the Euro, Persons	FRED	LFHUADTTEZM647S	5
106	Labor market	Total Harmonized Unemployment: Females for the Euro Area, Persons	FRED	LFHUTTFEEZM647S	5
107	Labor market	Total Harmonized Unemployment: Males for the Euro Area, Persons	FRED	LFHUTTMAEZM647S	5
108	Labor market	Harmonized Unemployment: Aged 15-24: Females for the Euro Area, Percentage	FRED	LRHU24FEEZM156S	2
109	Labor market	Harmonized Unemployment: Aged 15-24: Males for the Euro Area, Percentage	FRED	LRHU24MAEZM156S	2
110	Labor market	Harmonized Unemployment: Aged 15-24: All Persons for the Euro Area, Percentage	FRED	LRHU24TTEZM156S	2
111	Labor market	Harmonized Unemployment: Aged 25 and Over: Females for the Euro Area, Percentage	FRED	LRHUADFEEZM156S	2
112	Labor market	Harmonized Unemployment: Aged 25 and Over: All Persons for the Euro, Percentage	FRED	LRHUADTTEZM156S	2
113	Labor market	Harmonized Unemployment: Total: Females for the Euro Area, Percentage	FRED	LRHUTTFEEZM156S	2
114	Labor market	Harmonized Unemployment: Total: Males for the Euro Area, Percentage	FRED	LRHUTTMAEZM156S	2
115	Trade balance	Net Trade: Value Goods for the Euro Area (national currency)	FRED	XTNTVA01EZM664S	5
116	Trade balance	Net Trade: Value Goods for the Euro Area (USD)	FRED	XTNTVA01EZM667S	5
117	Monetary and Financial Accounts	Liabilities to Central Bank, Euros	IMF-IFS	IMF_Liabilities	5
118	Monetary and Financial Accounts	Net Claims on Central Government, Claims on Central Government, Euros	IMF-IFS	IMF_Net_Claims1	5
119	Monetary and Financial Accounts	Net Claims on Central Government, Euros	IMF-IFS	IMF_Net_Claims2	5
120	Monetary and Financial Accounts	Net Claims on Central Government, Liabilities to Central Government, Euros	IMF-IFS	IMF_Net_Claims3	5
121	Monetary and Financial Accounts	Net Foreign Assets, Claims on Non-residents, Euros	IMF-IFS	IMF_Net_Foreign_Assets	5
122	Prices	Prices, Consumer Price Index, Harmonized, Index	IMF-IFS	IMF_CPI	5
123	Prices	Prices, Producer Price Index, All Commodities, Index	IMF-IFS	IMF_PPI_All	5
124	International reserves	International Reserves and Liquidity, Reserves, Official Reserve Assets, Official Gold Price, SDRs	IMF-IFS	IMF_International_Reserves1	5
125	International reserves	International Reserves and Liquidity, Reserves, Official Reserve Assets, Official Gold Price, US Dollar	IMF-IFS	IMF_International_Reserves2	5
126	International reserves	International Reserves and Liquidity, Reserves, Official Reserve Assets, Market Value/Price, US Dollar	IMF-IFS	IMF_International_Reserves3	5

Note: The column "tcode" denotes the following series transformations:
(1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $\ln(x_t)$; (5) $\Delta \ln(x_t)$; (6) $\Delta^2 \ln(x_t)$.

Appendix B. Further discussion on dating recessions

Chauvet and Hamilton (2006) and Hamilton (2011)

Hamilton's (2011) article aims at answering whether there is any effective technique for dating recessions in real time. Part of the literature uses the techniques proposed by the author himself, Hamilton (1989), in his seminal article on models of two (or more) regimes, with a latent state variable representing the regimes of the economy, driven by a Markov Chain. Part of this literature uses common unobservable factors, which can be identified using principal component analysis, e.g., Stock and Watson (1989, 1991, 2002).

Hamilton (2011) asks the following question: given that we have NBER dates for the U.S., what is the advantage of trying to implement an automatic mechanism (algorithm) for forecasting U.S. recessions? The author answers this question as follows:

1. *Timeliness.* The Business Cycle Dating Committee has issued its announcements of the beginning and end of recessions usually much after the event. For example, the NBER dated the recession to 1990-91, starting in August 1990 and ending in March 1991. It did not make the announcement that the recession had started until April 1991 – a month after which the NBER itself later decided to be the end of the recession. The end of the 2001 recession was announced in July 2003 – 28 months after the recession was considered over.

2. *Apolitical Mechanism.* A purely objective algorithm for determining recession dates in real time ensures that the process is completely apolitical. While no one has accused the NBER of altering its announcements based on political considerations, there is undeniably pressure to delay the announcement that a recession has begun or speed up the announcement that a recovery has begun if the goal is to help the incumbent.

3. *Structural Mechanism.* Creating a mechanical way of recognizing the inflection points of business cycles allows us to elucidate what we really mean when we say "the economy is in recession". If the dates assigned by the NBER represent the answer, what is the question? The whole process seems to assume that there are some very different factors operating in the economy at different times, and that these changes have observable implications. Mechanization of the dating procedure can help clarify exactly how and why we assign the dates we do.

The next question asked by Hamilton is *why is it so difficult to implement a recession-prediction algorithm for the U.S.*? Again, the answer is given in three different ways:

1. *Predictability.* If people could predict recessions, they probably wouldn't happen. Firms wouldn't be stuck with inventories, labor and capital they wouldn't need, and the Fed would likely ease its monetary policy stance sooner. Economists know that stock prices are difficult (or impossible) to predict if the market is working properly. It may be that economic recessions, by their very nature, imply similar limitations to their forecasting.

2. *Data revision.* The data available in real time may send different signals than the same data later reviewed. For example, the real GDP growth rates for each quarter of 2001, as reported in the late 2002 *vintage*, show 3 successive quarters of declining real GDP, which sounds unmistakably like a recession. However, data from the same quarters by the January 30, 2002 *vintage* show that the recession was already over. This shows the difficulty of using GDP data to date recessions, as these are subject to many revisions that, in some cases, may change the perception of the state of the economy depending on the *vintage*.

3. *Changes over time.* One factor that makes it difficult to recognize real-time business cycle turning points is the fact that key economic relationships continually change over time and with available information.

With respect to item (3) above, Hamilton illustrates these difficulties by reviewing the real-time history of two important efforts to predict business cycle turning points. The first is the model of Stock and Watson (1989, 1991):

$$\begin{aligned} y_t &= k + \gamma(L) c_t + u_t, \\ D(L) u_t &= \varepsilon_t, \\ \phi(L) c_t &= \delta + \eta_t, \end{aligned} \tag{18}$$

where y_t includes the four coincident series, industrial production, income, sales and employment, c_t represents the common factor of business cycles, $D(L)$ is a polynomial in L composed of diagonal matrices. $\gamma(L)$ and $\phi(L)$ are also polynomials in L and ε_t and η_t are white-noise terms.

The model is estimated by the real-time *Kalman filter*. Stock and Watson defined an unobservable state variable S_t equal to unity if the economy were in recession on the period t. They interpreted a recession as a particular pattern followed by c_t, $S_t = 1$, if $\{c_{t-j}\}_{j=0}^{8} \in B_t$, where B_t is chosen to best represent the historical pattern of U.S. recessions as determined by the NBER. With this framework, one can adjust the historical pattern of recessions and calculate their probabilities (in-sample) from a historical point of view:

$$P\left(S_{t-h} \middle| y_t, y_{t-1}, \cdots, y_1; \widehat{\theta}\right) = 1, \qquad h = 0, 1, 2, \cdots$$

where $\widehat{\theta}$ includes the estimated parameters of model (18).

In order to have a prospective model that could be used to predict recessions in *real time*, Stock and Watson considered a generalization of the system in (18), which would work as a leading indicator:

$$\begin{aligned} y_t &= k + \gamma(L) c_{t-1} + \Gamma(L) y_{t-1} + \varepsilon_t, \\ c_t &= \delta + \alpha(L) c_{t-1} + \beta(L) y_{t-1} + \eta_t, \end{aligned} \tag{19}$$

where y_t now includes both traditional and alternative coincident indicators. With this new model, one can make forecasts of recessions from a prospective (out of the sample) point of view:

$$P\left(S_{t+h} \middle| y_t, y_{t-1}, \cdots, y_1; \widehat{\theta}\right) = 1, \qquad h = 0, 1, 2, \cdots$$

where $\widehat{\theta}$ now includes the estimated parameters of model (19).

There are two important points to note here. First, the *in-sample* prediction results are relatively successful, which could indicate that the *out-of-sample* predictions would be as well. Second, the *out-of-sample* prediction results are poor. The recession does not appear on the radar 6 months ahead, nor 3 months ahead. Even when it had already started, the model could not date it in real time, which only happens a few months *after* it started.[10]

Hamilton asks: what went wrong in this recession? One of the new leading indicators that Stock and Watson used was the spread between the Commercial Paper and the 6-month T-bill. This spread peaked dramatically before each of the recessions in the sample,

[10] Hamilton explains the problem as follows: "The leading indicator proved to be a disappointment: the 1990s recession came and went, with the model predicting no recession. In fact, in November 1990, the contemporary index, $P(t|t)$, signaled that a recession had begun, but the model thought it would be short-lived. In fact, the 3 and 6 months ahead probabilities always remained below 50%, although the ex-post probabilities ended up recognizing that a recession had occurred at some earlier date."

but it did very little out of the ordinary in the 1990-91 recession, in which the Stock and Watson model was tested in real time. Stock and Watson later released probabilities on their original leading index and an alternate index that did not make use of any interest rates or interest rate spreads. Model performance was similar for the 2001 recession and the 1990-1991 recession. The alternative leading indicator estimated a recession probability of 6 months earlier, $P(t+6\,|t)$, which remained below 20%.

A second example of poor predictions of real-time recessions comes from the work of Wright (2006). Wright defined a historic indicator $H_t = 1$, if the NBER later declared that there was an economic recession in any of the following four quarters. $H_t = 0$, otherwise. With these series, the spread of two treasuries – 10 years minus 3 months – and with the FED funds rate, Wright fitted the following *Probit* model predicting the probability of recessions:

$$P\left(H_t = 1\,|i_{10y,t}, i_{3m,t}, i_{f,t}\right) = \Phi\left(-2.17 - 0.76\left(i_{10y,t} - i_{3m,t}\right) + 0.35 i_{f,t}\right),$$

where $\Phi(\cdot)$ represents the standard Normal distribution function, $\mathcal{N}(0,1)$.

The website *Political Calculations* reported weekly updates of this model's predictions between April 2006 and August 2008. The highest this probability reached was 50% on April 4, 2007, after which the probability monotonically declined. It stood at just 0.1% on August 20, 2008, just before of one of the worst periods the U.S. economy has experienced over the past half century. These two examples illustrate that a good fit of an in-sample model is not a guarantee of a good fit in real-time or out-of-sample.

In a slightly less optimistic tone, Hamilton then proposes to aim for something more modest, trying to recognize a turning point soon after it has occurred using robust algorithms against the revised data and structural changes out-of-sample, which seems to have a reasonable track record (at least until now). The algorithm that Hamilton proposes is based on the approach in Chauvet and Hamilton (2006).

Now consider the following question: what is different about the behavior of GDP during the quarters that the NBER classifies as recessions compared to those characterized as expansions? Chauvet and Hamilton (2006) answered this question by collecting the growth rate of U.S. GDP between 1947Q2 and 2004Q2, in which the NBER ended up describing 45 of these quarters as part of an economic recession. This subsample of 45 observations has an average growth rate of -1.2% (expressed as annual growth) and a standard deviation of 3.5. The remaining 180 expansion quarters had a mean of 4.5 and a standard deviation of 3.2. The observed sample of 225 observations can be seen as a mixture of these two distributions, with 20% coming from the distribution of recession periods and the remaining 80% from the distribution of booming periods.

Let $S_t = 1$ if the quarter t is eventually declared by the NBER as part of a recession, and $S_t = 2$ if quarter t is eventually declared part of an expansion. Denote the GDP growth rate by y_t. We can define the joint probability that the quarter t is declared a recession and has a GDP growth rate of y_t as follows:

$$P\left(S_t = 1, y_t\right) = f\left(y_t\,|S_t = 1\right) P\left(S_t = 1\right),$$

where $f\left(y_t\,|S_t = 1\right)$ is the conditional density of y_t given $S_t = 1$, in which $P\left(S_t = 1\right)$ is the probability of occurrence of a recession, previously calculated as 0.2. Analogously:

$$P\left(S_t = 2, y_t\right) = f\left(y_t\,|S_t = 2\right) P\left(S_t = 2\right),$$

where $P\left(S_t = 2\right) = 0.8$. The optimal inference that interests us is to know what is the probability of a recession being dated by the NBER when we have a certain value of GDP growth as a condition:

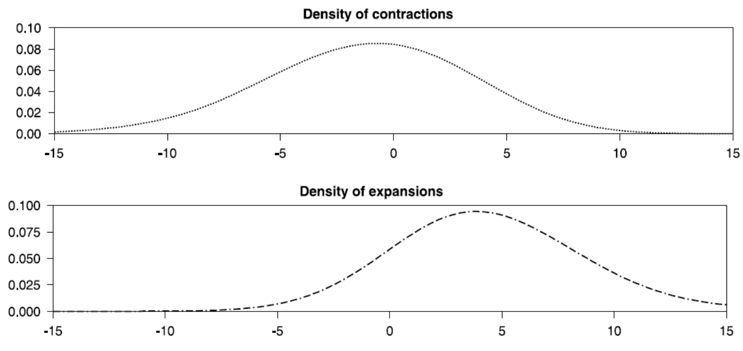
$$P(S_t = 1 | y_t) = \frac{P(S_t = 1, y_t)}{f(y_t)} = \frac{P(S_t = 1, y_t)}{P(S_t = 1, y_t) + P(S_t = 2, y_t)},$$

so we can simplify:

$$\begin{aligned} P(S_t = 1, y_t) &= f(y_t | S_t = 1) \times 0.2, \\ P(S_t = 2, y_t) &= f(y_t | S_t = 2) \times 0.8. \end{aligned}$$

When we separate the observations into contractions and expansion periods, we obtain the following densities:

Figure B1 – Conditional probability density function (pdf)



Source: Hamilton (2011).

This shows these densities have very different means, although their respective variances are not very different. The average of recessive periods is a GDP growth (annualized) of -1.2%, whereas the average of expansion periods is of 4.5%. Therefore, if we observe a GDP growth of 6%, for example, the probability of a recession being dated is quite low. On the other hand, if we observe a GDP growth of -6%, this probability will be quite high.

This rule (or algorithm) seems to satisfy the requirements of being simple and robust. Unfortunately, this is not very useful, since the vast majority of observations will fall into a range where they will not provide clear signals. However, there is a second feature of NBER dates that can be quite useful: the value of S_t is quite likely the same as S_{t-1}. For 95% of the observations for which $S_{t-1} = 2$, S_t was also equal to 2, whereas 78% of the observations for which $S_{t-1} = 1$ were also followed by $S_t = 1$. So, even if y_t alone does not give us a very useful idea of the signal, the value of y_{t-1} can help us refine it. Obviously, there is no reason to stop conditioning in the period $t-1$, when we can condition the model using information up to the beginning of the sample, i.e., calculating:

$$P(S_t = 1 | y_t, y_{t-1}, \cdots, y_1),$$

which is a filtered version of the probability of occurrence of a recession dated by the NBER. We can also consider a smoothed version of it, which uses the entire sample, $y_1, y_2, \cdots, y_T$, as follows:

$$P(S_t = 1 | y_T, y_{T-1}, \cdots, y_1).$$

The model estimation follows the following steps. It is assumed that:

$$\begin{aligned} y_t | S_t = 1 &\sim \mathcal{N}(\mu_1, \sigma^2), \\ y_t | S_t = 2 &\sim \mathcal{N}(\mu_2, \sigma^2). \end{aligned}$$

The density of the first observation is given by:

$$f\left(y_1;\mu_1,\mu_2,\sigma^2,\pi\right)=\pi\phi\left(y_1;\mu_1,\sigma^2\right)+\left(1-\pi\right)\phi\left(y_1;\mu_2,\sigma^2\right),$$

where π is the unconditional probability of occurrence of a recession dated by the NBER and $\phi(\cdot)$ is the probability density function (pdf) of a Normal random variable. This way:

$$P\left(S_t=1\left|y_1\right.\right)=\frac{\pi\phi\left(y_1;\mu_1,\sigma^2\right)}{\pi\phi\left(y_1;\mu_1,\sigma^2\right)+\left(1-\pi\right)\phi\left(y_1;\mu_2,\sigma^2\right)}.$$

For $t=2$, it follows that:

$$P\left(S_t=1\left|y_2,y_1\right.\right)=\frac{\xi_2\phi\left(y_2,\mu_1,\sigma^2\right)}{\xi_2\phi\left(y_2;\mu_1,\sigma^2\right)+\left(1-\xi_2\right)\phi\left(y_2;\mu_2,\sigma^2\right)},$$

where:

$$\xi_2=p_{11}P\left(S_t=1\left|y_1\right.\right)+\left(1-p_{22}\right)P\left(S_t=2\left|y_1\right.\right),$$

such that $p_{11}=P\left(S_t=1\left|S_{t-1}=1\right.\right)$ is the probability that the recession will continue to be dated and $p_{22}=P\left(S_t=2\left|S_{t-1}=2\right.\right)$ is the probability of continuing an expansion, so $\left(1-p_{22}\right)$ is the probability that an expansion will end. From the third observation, we can iterate as follows:

$$\begin{aligned}
P\left(S_t=1\left|y_t,y_{t-1},\cdots,y_1\right.\right) &= \frac{\xi_t\phi\left(y_t,\mu_1,\sigma^2\right)}{\xi_t\phi\left(y_t;\mu_1,\sigma^2\right)+\left(1-\xi_t\right)\phi\left(y_t;\mu_2,\sigma^2\right)},\\
\xi_t &= p_{11}P\left(S_{t-1}=1\left|y_{t-1},y_{t-2},\cdots,y_1\right.\right)\\
&\quad +\left(1-p_{22}\right)P\left(S_{t-1}=2\left|y_{t-1},y_{t-2},\cdots,y_1\right.\right),\\
f\left(y_t\left|y_{t-1},y_{t-2},\cdots,y_1\right.;\theta\right) &= \sum_{j=1}^{2}f\left(y_t\left|S_t=j\right.\right)P\left(S_t=j\left|y_{t-1},y_{t-2},\cdots,y_1\right.\right)\\
&= \xi_t\phi\left(y_t;\mu_1,\sigma^2\right)+\left(1-\xi_t\right)\phi\left(y_t;\mu_2,\sigma^2\right),
\end{aligned}$$

where $\pi=\frac{1-p_{22}}{1-p_{22}+1-p_{11}}$ and θ is a vector of parameters associated with the conditional density $f\left(y_t\left|y_{t-1},y_{t-2},\cdots,y_1\right.;\theta\right)$ as follows, $\theta=\left(\mu_1,\mu_2,\sigma^2,p_{11},p_{22}\right)'$.

The log-likelihood function is then given by:

$$\mathcal{L}\left(\theta;y_T,y_{T-1},\cdots,y_1\right)=\sum_{t=1}^{T}\log\left[f\left(y_t\left|y_{t-1},y_{t-2},\cdots,y_1\right.;\theta\right)\right].$$

Next, the results of the estimation of the model of Chauvet and Hamilton (2006), were compared with those from Hamilton for the GDP data, classified according to the NBER:

Table B1 – Parameter estimates

Table 1
Parameter estimates based on (1) the characteristics of expansions and recessions as classified by the NBER, and (2) the values that maximize the observed sample log likelihood of postwar GDP growth rates over the period 1947:Q2–2004:Q2.

Parameter	Interpretation	Value from NBER classifications	Value from GDP alone
μ_1	Average growth in expansion	4.5	4.62
μ_2	Average growth in recession	−1.2	−0.48
σ	Standard deviation of growth	3.4	3.34
p_{11}	Prob. expansion continues	0.95	0.92
p_{22}	Prob. recession continues	0.78	0.74

Note: Reproduced from Chauvet and Hamilton (2006).

Source: Hamilton (2011).

The results are apparently excellent, with the exception of average GDP growth in NBER recessions. But, we must keep in mind that they were obtained in an *in-sample* analysis. As

Hamilton (in fact, Niels Bohr) reminded us, predictions are very difficult, especially about the future, i.e., *out-of-sample*.

Next, Hamilton shows the results of two experiments. In the first one, the filtered probability is calculated as $P(S_t = 1 | y_t, y_{t-1}, \cdots, y_1)$ using the latest GDP vintage (available in 2004). In the second experiment, to evaluate the usefulness of this algorithm in real time (on date t), the authors use GDP data that would be available in t according to its vintage; see the database maintained by the Federal Reserve Bank of Philadelphia - Croushore and Stark (2003). These data are used both to estimate the parameter vector θ and to make inference about $P(S_t = 1 | \cdot)$. However, the use of such data resulted in a considerable deterioration of results. Therefore, Hamilton recommends waiting an extra quarter to use revised data, which allows for additional precision, before dating recession in t.

Thus, the proposed mechanism is to make inferences about S_t only after the next quarter's GDP growth rate, y_{t+1}, is available, calculating $P\left(S_t = 1 | y_{t+1}, y_t, \cdots, y_1 ; \widehat{\theta}_{t+1}\right)$, where $\widehat{\theta}_{t+1}$ represents the parameters estimated by maximum likelihood using as data $y_1, y_2, \cdots, y_t, y_{t+1}$.

In order to implement the real-time algorithm, based on $P\left(S_t = 1 | y_{t+1}, y_t, \cdots, y_1 ; \widehat{\theta}_{t+1}\right)$, Chauvet and Hamilton (2006) recommended the following decision rule: when $P\left(S_t = 1 | y_{t+1}, y_t, \cdots, y_1 ; \widehat{\theta}_{t+1}\right)$ first exceeds 0.65, a recession is declared to be underway. At this point, the probable start of the recession is attributed to the beginning of the most recent set of observations for which $P\left(S_{t-j} = 1 | y_{t+1}, y_t, \cdots, y_1 ; \widehat{\theta}_{t+1}\right)$ exceeds 0.5. The recession call remains in effect until $P\left(S_t = 1 | y_{t+1}, y_t, \cdots, y_1 ; \widehat{\theta}_{t+1}\right)$ falls below 0.35, time at which the probable end point for the recession is assigned as the beginning of the most recent set of observations for which $P\left(S_{t-j} = 1 | y_{t+1}, y_t, \cdots, y_1 ; \widehat{\theta}_{t+1}\right)$ is smaller than 0.5. The results of this implementation of the real-time algorithm are shown below:

Table B2 – Turning points of business cycle in the U.S.

Table 2
Business cycle turning points and dates on which announcements were issued by NBER and the GDP-based algorithm.

Start of recessions

Peak as determined by NBER	Date NBER made declaration	Recession start as determined by algorithm	Date algorithm made declaration	Algorithm announcement lead (−) or lag (+) in months
1969:Q4	N.A.	1969:Q2	May 1970*	N.A.
1973:Q4	N.A.	1973:Q4	May 1974*	N.A.
1980:Q1	Jun 1980	1979:Q2	Nov 1979*	−7
1981:Q3	Jan 1982	1981:Q2	Feb 1982*	+1
1990:Q3	Apr 1991	1989:Q4	Feb 1991*	−2
2001:Q1	Nov 2001	2001:Q1	Feb 2002*	+3
2007:Q4	Dec 2008	2007:Q4	Jan 2009	+1

Start of expansions

Trough as determined by NBER	Date NBER made declaration	Recession end as determined by algorithm	Date algorithm made declaration	Algorithm announcement lead (−) or lag (+) in months
1970:Q4	N.A.	1970:Q4	Aug 1971*	N.A.
1975:Q1	N.A.	1975:Q1	Feb 1976*	N.A.
1980:Q3	Jul 1981	1980:Q2	May 1981*	−2
1982:Q4	Jul 1983	1982:Q4	Aug 1983*	+1
1991:Q1	Dec 1992	1991:Q4	Feb 1993*	+2
2001:Q4	Jul 2003	2001:Q3	Aug 2002*	−12
2009:Q2	Sep 2010	2009:Q2	Apr 2010	−5

Notes: Starred entries denote simulated real-time declarations, unstarred are actual real-time declarations. N.A. indicates that the information is not available.

Source: Hamilton (2011).

The results in the table above show that the peak and valley dates of the algorithm and the NBER are either identical or very close. As for the timeliness of the algorithm's dates vis-à-vis the NBER, there is no clear advantage. In fact, half the time NBER dates peaks and valleys before the algorithm, and half the time NBER dates peaks and valleys after the algorithm. Anyway, in terms of the algorithm's average gain or loss in relation to the NBER, we have the following result: in the dates in which the algorithm is ahead of the NBER, the average gain is 5.6 quarters and the average loss is 1.6 quarters.