

Do Leading Indicators Forecast U.S. Recessions? A Nonlinear Re-Evaluation Using Historical Data

Vasilios Plakandaras^{*+}, Juncal Cunado^{**}, Rangan Gupta^{***} and Mark E. Wohar^{****}

^{*} Department of Economics, Democritus University of Thrace, Greece
Email: vplakand@econ.duth.gr

^{***} Department of Economics, University of Navarra, Spain.
Email: jcunado@unav.es

^{***} Department of Economics, University of Pretoria, South Africa.
Email: rangan.gupta@up.ac.za

^{****} College of Business Administration, University of Nebraska at Omaha USA, and
School of Business and Economics, Loughborough University, UK.
Email: mwohar@unomaha.edu

Abstract

This paper analyzes to what extent a selection of leading indicators are able to forecast U.S. recessions by means of both dynamic probit models and Support Vector Machines (SVM) models, using monthly data from January 1871 to June 2016. The results suggest that the probit models foresee U.S. recession periods more closely than SVM models for up to 6 months ahead, while the SVM models are more accurate at longer horizons. Furthermore, SVM models appear to discriminate between recessions and tranquil periods better than probit models do. Finally, the most accurate forecasting models include oil, stock returns and the term spread as leading indicators.

Keywords: Dynamic Probit Models, Support Vector Machines, U.S. Recessions.

JEL Codes: C53, E32, E37.

⁺Corresponding author.

1. Introduction

In September 2010, the NBER's Business Cycle Dating Committee announced that the last expansion in the U.S. economy began in June 2009. The same Committee determined in December 2008 that a peak in economic activity occurred in December 2007, delimiting, thus, that the last U.S. recession took place between December 2007 and June 2009. Economic recessions are accompanied by a higher probability of unemployment, lower wage growth, lower stock returns and decreases in lifetime earnings, which justifies the interest of households, businesses and policymakers to infer the current and future states of the economy. For example, and according to the U.S. Bureau of Labor Statistics, the U.S. unemployment rate increased from 5% (December 2007) to 10% (October 2009) during the recession, while the number of job openings decreased 44 percent. In the wake of this Great Recession, which was considered as the worst global recession since World War II by the International Monetary Fund (IMF, World Economic Outlook, April 2009), a lot has been said about the failure of economic models to forecast recessions (Ng and Wright, 2013; Gadea and Perez-Quiros, 2015), raising the interest of accurately forecasting future recessions.

Despite the recent increasing concern about this topic, a large amount of literature has tried to find leading indicators of future U.S. economic activity since the late eighties (Harvey, 1988, 1989; Stock and Watson, 1989; Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1997, 1998; Hamilton and Kim, 2002; Estrella et al., 2003; Giacomini and Rossi, 2006; Berge and Jordá, 2011; Levanon et al., 2015; Berge, 2015; Liu and Moench, 2016). However, regardless of the great volume of papers on this topic, accurately predicting business cycle turning points is still a pertinent research topic, whose interest has increased in the wake of the largely unpredicted more recent recession. According to the literature, the specific accuracy of the predictions depends on the choice of the leading indicators, on the employed methodology and, on the time period analysed in the empirical studies.

First, and as far as the leading indicators are concerned, the academic literature has proposed a wide variety of variables to predict the U.S. recessions. The slope of the yield curve, that is, the difference between long-term and short-term interest rates

has been found to be one of the most informative leading indicators to predict U.S. recessions (Harvey, 1988; Bernanke, 1990; Estrella and Hardouvelis, 1991; Bernanke and Blinder, 1992; Dueker, 1997; Hamilton and Kim, 2002; Stock and Watson, 2003; Ang et al., 2006; Rudebusch and Williams, 2009; Liu and Moench, 2016). According to these papers, a flat curve indicates weak growth, and conversely, a steep curve will be followed by stronger growth. Other variables that have also been considered as informative leading indicators are stock prices (Estrella and Mishkin, 1998; Hamilton, 2011; Killian and Vigfusson, 2013), the index of leading economic indicators (Stock and Watson, 1989; Berge and Jordá, 2011), the credit market activity (Levanon et al., 2015) or financial intermediary leverage indicators (Liu and Moench, 2016), among others. For example, and according to Hamilton (2005), nine out of ten of the U.S. recessions since World War II were preceded by a spike up in oil prices, which explains the view of oil prices as one of the leading indicators to predict U.S. recessions. Furthermore, Hamilton (2011), Engemann et al. (2010) and Killian and Vigfusson (2013) also find that oil prices have considerable predictive power for U.S. recessions. On the other hand, Estrella and Mishkin (1997, 1998) find that stock returns are an informative leading indicator, mainly in the short run. In order to account for the monetary policy, the literature has also included the short-term interest rate (Estrella and Hardouvelis, 1991) or different monetary aggregates (Hamilton and Kim, 2002) as additional explanatory or leading variables. In this paper, and based on the above economic literature, we will analyse the information content of several leading indicators, such as the yield curve (which we will decompose into an expected short-term interest rate and a term premium component) and nominal and real stock and oil prices.

Second, regarding the methodology, the specific accuracy of the predictions will also depend on the prediction variable. While many papers have focused on predicting continuous variables, such as GNP, GDP, industrial production or investment growth rates (Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1997), most papers use the business cycle chronology proposed by the National Bureau of Economic Research (NBER), a binary variable, to define recessions. The NBER defines a recession as “a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales”. Given the nature of

the variable, much of the empirical literature uses nonlinear probit models to obtain recession forecasts (Estrella and Mishkin, 1998; Dueker, 2005; Kauppi and Saikkonen, 2008; Berge, 2015; Liu and Moench, 2016). This paper will use the U.S. Business Cycle Expansions and Contractions data provided by the NBER as the predicting variable, and non-linear dynamic probit models and non-linear Support Vector Machines (SVM) models to forecast U.S. recessions. While dynamic probit models have been widely used in the literature, SVM models have hardly been used to predict recessions. Among the few papers using this methodology, Gogas et al. (2015) applied Support Vector Machines (SVM) methodology to analyze the ability of the yield curve to forecast US output fluctuations around its long-run trend, using quarterly data for the period 1976:Q3-2011:Q4. Their results show that the SVM methodology outperformed classic econometric models (probit models) on overall forecast accuracy. In this paper, we use both probit and SVM models to analyze the ability of different leading indicators to forecast U.S. recessions and evaluate the prediction accuracy of each of the methods.

Finally, the sample period may influence the results as well. Although many economic variables (i.e., yield curve) have provided useful information about future states of the economy, the relationships between these indicators and the state of the economy might have changed over time (Ng and Wright, 2013). Thus, a number of studies have shown that the predictive power of some leading indicators such as the yield curve have declined since the 1980s (Gertler and Lown, 1999; Estrella et al., 2003; Mody and Taylor, 2003; Rossi and Sekhposyan, 2011). For example, Stock and Watson (2003) and Estrella et al. (2003) document evidence of instability in the relationship between economic activity and leading indicators over time. Moreover, Stock and Watson (2003) and Chauvet and Potter (2002, 2005) found evidence of structural breaks in the relationship between the yield curve and economic activity, and conclude that the credibility of the monetary policy is behind this general result that the predictive power of the yield curve varies over time. Independently of the reasons of these changes, the prediction ability of each of the models will depend on the sample period. In this paper we propose to use a very long time period of data, January 1871-June 2016, which include very distinct monetary policy regimes, and covers nearly the entire history of available data on U.S. recessions.

In this context, the objective of this paper is to determine to what extent a selection of leading indicators are able to forecast U.S. recessions by means of both dynamic probit models and classification methods, such as the linear and non-linear Support Vector Machines (SVM) models, using monthly data from January 1871 to June 2016. The main contributions of the paper are the following. First, the paper uses an ample selection of leading indicators, such as the yield spread, oil price shocks, stock returns and the term premium, and analyze the forecasting ability of each of them. Second, the paper uses both the probit and the SVM models to predict U.S. recessions and compare the accuracy of the predictions obtained with each of the methodologies. Third, and in order to evaluate the accuracy of the predictions, the paper analyzes both in-sample and out-of-sample Quadratic Probability Score (QPS, Diebold and Rudebusch, 1989) for each of the models. Finally, the paper analyzes a long period of data, 1871:01- 2016:06, which include very distinct episodes in the U.S. economy.

The remainder of the paper is structured as follows. Section 2 describes the data and discusses the methodology used in the paper. Section 3 shows the empirical analysis. Section 4 summarizes the main findings.

2. Data and Methodology

2.1 The Data

Hamilton and Kim (2002) suggest that the spread between the 10 year and 3 month interest rates is a leading indicator of future U.S. economic activity. Moreover, Hamilton (2011) and Killian and Vigfusson (2013) find that both real and nominal stock and oil price returns possess predictive power over the future state of the economy. Overall, we compile a dataset of monthly observations spanning the period January 1871 to June 2016 that covers almost the entire history of available information on U.S. recessions.¹ The dataset consists of the S&P500 index, zero-coupon Treasury bills with maturity of 3 months and 10 years, and West Texas Intermediate (WTI) oil prices. The data on the stock price and long-term interest rates

¹ Data on U.S. recessions available at: <http://www.nber.org/cycles.html> starts from 1854.

are obtained from the website of Professor Robert J. Shiller: <http://www.econ.yale.edu/~shiller/data>. The CPI data used to inflate both the nominal stock and oil prices to create their respective real counterparts, is also obtained from the same website. The WTI oil price is obtained from the Global Financial Database, while the short-term interest rate is obtained from the website of Professor Amit Goyal at: <http://www.hec.unil.ch/agoyal/> till 2015:12, and then updated from the FRED database of the Federal Reserve Bank of St. Louis. Moreover, we decompose the yield curve in an expected short-term interest rate and a term premium component, following Hamilton and Kim (2000):

$$i_t^n = \frac{1}{n} \sum_{j=0}^{n-1} E_t i_{t+j}^1 + TP_t \quad (1)$$

and equivalently

$$i_t^n - i_t^1 = \left(\frac{1}{n} \sum_{j=0}^{n-1} E_t i_{t+j}^1 - i_t^1 \right) + TP_t \quad (2)$$

$$\text{where } TP_t = i_t^n - \frac{1}{n} \sum_{j=0}^{n-1} E_t i_{t+j}^1 - i_t^1$$

where i_t^n denotes the long-term interest rate, i_t^1 the one period short-term interest rate, n is the maturity of the long-term interest rate, $\left(\frac{1}{n} \sum_{j=0}^{n-1} E_t i_{t+j}^1 - i_t^1 \right)$ the future expected short-term interest rate at $t + j$ periods ahead and TP_t the term premium. Equation (2) can be estimated using instrumental variables regression with i_t^n and i_t^1 as instruments. In Figure 1 we depict the term spread and the decomposed expected short-term interest rate and the term premium component.

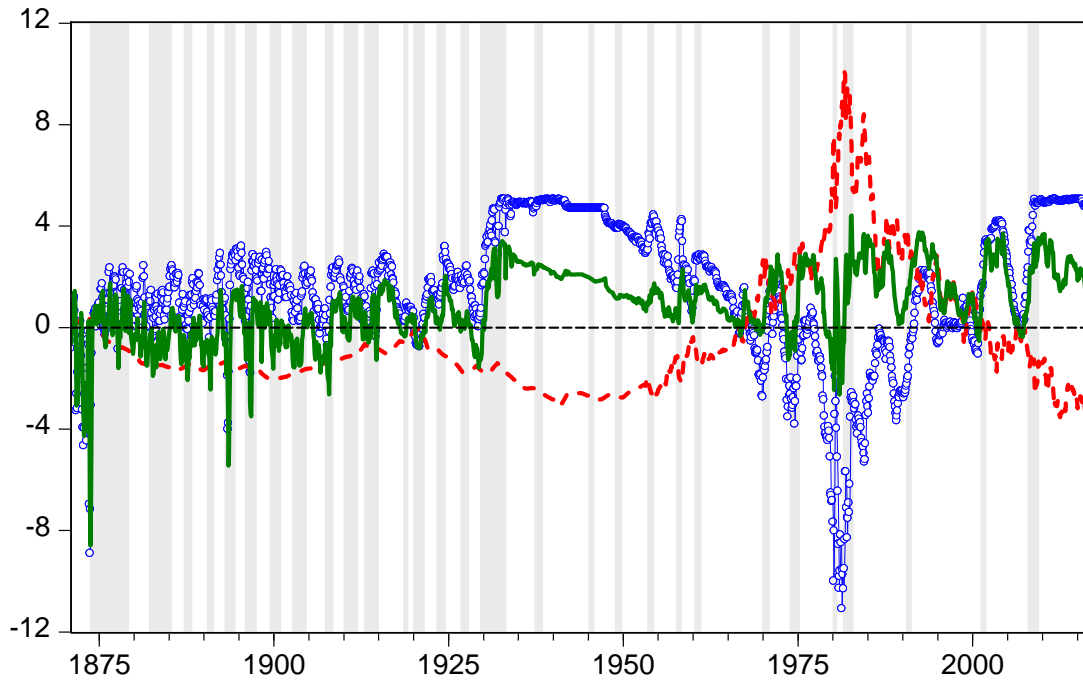


Figure 1: The term spread is depicted with the (green) continuous line, the dotted (blue) line with the circle markers depicts the expected short term interest rate and the (red) dashed line without markers depicts the term premium. Grey areas denote NBER recessions. Colors refer only to the online version of the paper.

As we observe from Figure 1, the term spread exhibits a declining trend during the period 1925 -1975. After 1975, the fluctuation of the term spread is higher, with the term premium reaching a significant positive percentage as a result of the high inflation of the period. In other words, due to the high inflation rates of that period, investors demanded a high compensation in order to hold Treasury Bills of long maturity. Interestingly, in the post Volcker administration period the inflation targeted policies of the Federal Reserve push the term premium towards negative, while the short term expected interest rate is stable after the 2008 financial crisis at around 4%. The reversal between the two term spread components could be attributed to a “flight-in-quality” phenomenon that is common during periods of recession; investors prefer to trust their capital in government bonds that are unlikely to default, than use it in the open market. Thus they push the short-term interest rate up, while they are indifferent to a positive term premium.

In Table 1 we report descriptive statistics and unit root tests for all variables. As we observe, we reject the null hypothesis of normality (Jarque – Bera test) for all variables. Thus, stock prices indices and oil prices are transformed into their natural logarithms. Moreover, we reject the null hypothesis of non-stationarity at the 5% level

of significance for the term spread and real oil prices, but we cannot reject the null of unit root for stock prices (real or nominal) and term premium. The unit root test results are inconclusive regarding the inclusion or not of a trend term in the unit root test for the nominal oil prices and the expected interest rates, so we also consider them as non-stationary. In all cases where the variables are found to follow a unit root process we use first differences; nominal oil prices, real and nominal stock prices, term premium, short and long-term interest rates and the expected short-term interest rate².

Table 1: Descriptive Statistics and Unit Root tests of monthly data

	Mean	Standard Deviation	Skewness	Kurtosis	Jarque–Bera test	Augmented Dickey-Fuller test		Phillips–Perron test	
						Constant	Constant and trend	Constant	Constant and trend
S&P500	3.58	1.93	0.69	2.14	193.06**	0.98	-1.87	-0.81	-1.95
Real SP500	2.78	0.58	0.61	2.41	131.04**	-0.70	-2.60	-0.82	-2.62
WTI	1.47	1.37	0.75	2.48	185.13**	-0.87	-3.77*	-0.74	-3.58*
Real WTI	0.67	0.45	0.50	2.61	84.86**	-4.29**	-4.59**	-4.02**	-4.31**
Term Spread	0.92	1.39	-0.64	5.25	488.73**	-5.69**	-7.27**	-6.56**	-8.11**
Term premium	-0.54	2.27	1.84	6.79	2032.54**	-1.65	-1.60	-1.75	-1.71

Note: ** and * denote rejection of the null hypothesis at 1% and 5% levels of significance.

2.2 Methodology

2.2.1 Support Vector Machines

The Support Vector Machines is a supervised machine learning method used for data classification. Roughly, the basic concept of an SVM is to select a small number of data points from a dataset, called Support Vectors (SV) that can define a linear boundary separating the data points into two classes. When the problem is not linearly-separable, then SVM is coupled with a nonlinear kernel function, projecting the data points to a higher dimensional space, called feature space, where a linear separation is feasible. In the following we describe briefly the mathematical derivations of the SVM theory.

We consider a dataset (vectors) $\mathbf{x}_i \in R^2$ ($i = 1, 2, \dots, n$) belonging to 2 classes (output vectors or targets³) $y_i \in \{-1, +1\}$. If the two classes are linearly separable, then we define a separator as:

$$f(\mathbf{x}_i) = \mathbf{w}^T \mathbf{x}_i - b = 0, \quad y_i f(\mathbf{x}_i) > 0 \forall i \quad (3)$$

² The lag order of the unit root tests are determined according to the minimum SIC criterion and are available upon request from the authors.

³In the SVM jargon.

where \mathbf{w} is the weight vector and b is the bias.

The optimal hyper plane is selected as the decision boundary that classifies each data vector to the correct class and has the maximum distance from each class. This distance is often called “margin”. In Figure 2, the SV’s are represented with the pronounced contour, the margin lines (defining the distance of the hyperplane with each class) are represented with the continuous lines and the hyper plane is represented with the dotted line.

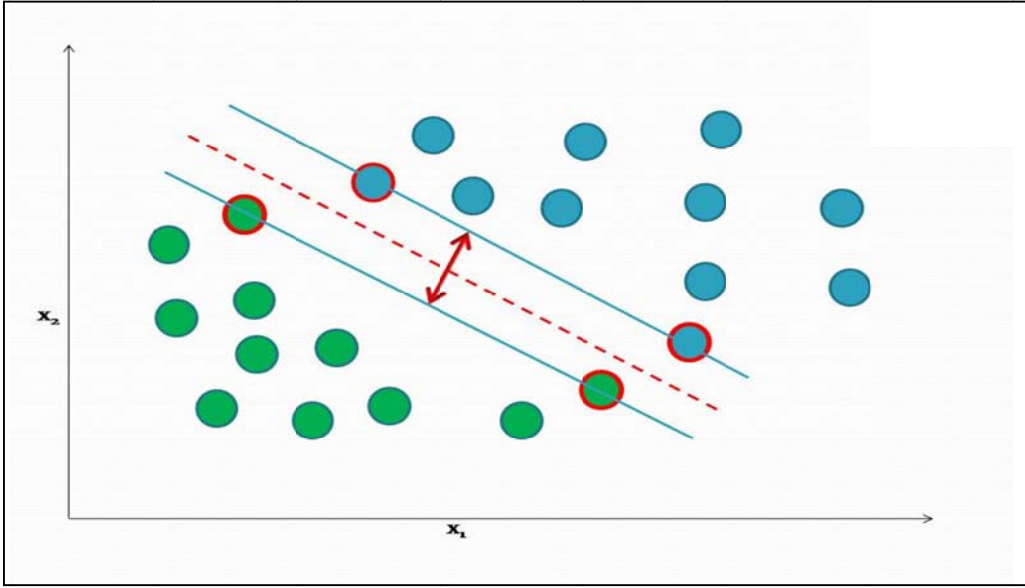


Figure 2: Hyper plane selection and support vectors. The SV’s are represented with the pronounced red contour, the margin lines are represented with the continuous lines and the hyper plane is represented with the dotted line.

The solution to the problem of finding the hyper plane can be dealt through the Lagrange relaxation procedure on the following equation:

$$\min_{\mathbf{w}, b} \max_{\mathbf{a}} \left\{ \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^N a_i [y_i (\mathbf{w}^T \mathbf{x}_i - b) - 1] \right\} \quad (4)$$

where $\mathbf{a} = a_1, \dots, a_n$ are the non negative Lagrange multipliers. Equation (4) is never used to estimated the solution. Instead we always solve the dual problem, defined as:

$$\max_{\mathbf{a}} \left\{ \sum_{i=1}^N a_i - \sum_{j=1}^N \sum_{k=1}^N a_j a_k y_j y_k \mathbf{x}_j^T \mathbf{x}_k \right\}, \quad i = 1, \dots, a_i y_i = 0 \text{ and } 0 \leq a_i \forall i \quad (5)$$

The solution of (3) gives the location of the hyper plane defined by:

$$\hat{\mathbf{w}} = \sum_{i=1}^N a_i y_i \mathbf{x}_i \quad (6)$$

$$\hat{b} = \hat{\mathbf{w}}^T \mathbf{x}_i - y_i, i \in V \quad (7)$$

where $V = \{i: 0 < y_i\}$ is the set of the support vector indices.

In order to allow for a predefined level of error tolerance in the training procedure Cortes and Vapnik (1995) introduced non-negative slack variables $\xi_i \geq 0, \forall i$ and a parameter C describing the desired tolerance to classification errors. Equation (4) is now defined as:

$$\min_{\mathbf{w}, b, \xi} \max_{\mathbf{a}, \mu} \left\{ \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i - \sum_{j=1}^N a_j [y_j (\mathbf{w}^T \mathbf{x}_j - b) - 1 + \xi_j] - \sum_{k=1}^N \mu_k \xi_k \right\} \quad (8)$$

where ξ_i measures the distance of vector \mathbf{x}_i from the hyper plane when classified erroneously.

The hyper plane is then defined as:

$$\hat{\mathbf{w}} = \sum_{i=1}^N a_i y_i \mathbf{x}_i \quad (9)$$

$$\hat{b} = \hat{\mathbf{w}}^T \mathbf{x}_i - y_i, i \in V \quad (10)$$

where $V = \{i: 0 < y_i < C\}$ is the set of the support vector indices.

When the two class dataset cannot be separated by a linear separator then the SVM classification is paired with kernel methods. The concept is quite simple: the dataset is projected through a kernel function into a richer space of higher dimensionality (called feature space) where the dataset is linearly separable. In Figure 3 we depict a dataset of two classes that are not linearly separable at the initial dimensional space (left graph). With the projection in a higher dimensional space (right graph) the linear separation is feasible.

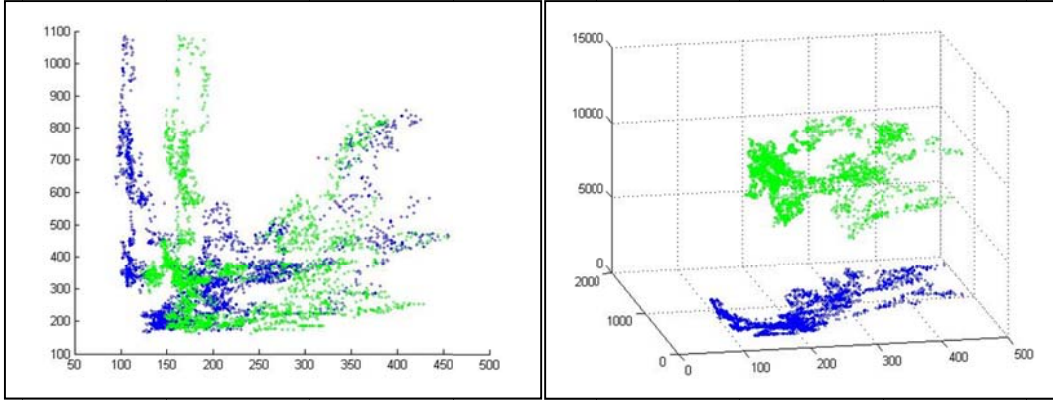


Figure 3: The Data Space. The non-separable two class scenario (left) and the separable case in the feature space after the projection (right)

The solution to the dual problem with projection of eq. (6) now transforms to:

$$\max_{\mathbf{a}} = \sum_{i=1}^N a_i - \frac{1}{2} \sum_{j=1}^N \sum_{k=1}^N a_j a_k y_j y_k K(\mathbf{x}_j, \mathbf{x}_k) \quad (11)$$

Under the constraints $\sum_{i=1}^N a_i y_i = 0$ and $0 \leq a_i \leq C, \forall i$ where $K(\mathbf{x}_j, \mathbf{x}_k)$ is the kernel function.

In our models we examine two kernels: the linear and the radial basis function (RBF)⁴. The linear kernel detects the separating hyperplane in the actual dimensional space of the dataset, while the RBF projects the initial dataset in a higher dimensional space. The mathematical representation of each kernel is:

$$\text{Linear} \quad K_1(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_1^T \mathbf{x}_2 \quad (12)$$

$$\text{RBF} \quad K_2(\mathbf{x}_1, \mathbf{x}_2) = e^{-\gamma \|\mathbf{x}_1 - \mathbf{x}_2\|^2} \quad (13)$$

Platt (2000) proposes a parametric method in mapping the binary output of equation (3) to class probabilities. Fitting a sigmoid model on the posterior of $P(y = 1|f)$ he computes the posterior probability as:

$$P(y = 1|f) = \frac{1}{1 + \exp(Af + B)} \quad (14)$$

where A,B are parameters to be computed through a minimization procedure. Bearing in mind that the targets follow $y_i \in \{-1, +1\}$, we assume that target probabilities follow $k_i = y_i + 1/2$. Thus finding parameters A, B that fit the sigmoid to the output

of equation (4) is the equivalent of minimizing the negative log likelihood of

⁴ Our implementation of SVR models is based on LIBSVM (Chang and Lin, 2011). Software available at: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.

$$\min(-\sum_i k_i \log(p_i) + (1 - k_i)\log(1 - p_i)) \quad (15)$$

where

$$p_i = \frac{1}{1 + \exp(Af + B)}$$

2.3 Dynamic Probit models

The majority of empirical studies in the field exploit Markov switching models. Although the specific category of models has the advantage of providing state-dependent inferences, the main drawback is that it is based on an unobservable Markov switching process and an unobservable state variable. In our study we consider a binary response model that predicts recessions as a binary time series response that is directly observable. We denote the binary state variable as s_t :

$$s_t = \begin{cases} 1, & \text{the economy is in recession at time } t \\ 0, & \text{the economy in is expansion at time } t \end{cases} \quad (16)$$

for $t=1,2,\dots,n$ the range of the monthly observations.

Denoting the conditional expectation $E_{t-1}(s_t|\Omega_{t-1})$ in the information set Ω_{t-1} at time $t-1$, the conditional probability at time t that the market is in a recession is:

$$p_t = E_{t-1}(s_t|\Omega_{t-1}) = P_{t-1}(s_t = 1) = \Phi(\pi_t) \quad (17)$$

where π_t is a linear combination of variables and $\Phi(\cdot)$ is the normal cumulative distribution function. Naturally, the conditional probability of a recession is the complement of the probability that the economy is not in recession $P_{t-1}(s_t = 0) = 1 - p_t$. In order to predict the linear function π_t we study static and dynamic models. We use as benchmark the univariate probit model (Chen, 2009):

$$\pi_t = \omega + \mathbf{x}'_{t-h}\boldsymbol{\beta} \quad (18)$$

where ω is a constant, $\boldsymbol{\beta}$ is the coefficients vector and \mathbf{x}_{t-h} a matrix of predictive regressors. The index h denotes the forecasting horizon. The popular static model can be extended by adding lags of the state variable s_t resulting in the dynamic autoregressive model

$$\pi_t = \omega + \alpha(s_{t-1}) + \mathbf{x}'_{t-h}\boldsymbol{\beta} \quad (19)$$

or by adding lags of the dependent variable π_t , leading to the autoregressive dynamic model (Kauppi and Saikkonen, 2008):

$$\pi_t = \omega + \delta(\pi_{t-1}) + \mathbf{x}'_{t-h}\boldsymbol{\beta} \quad (20)$$

By recursive substitution, equation (18) can be seen as an infinite order static equation (17) where the whole history of the values of the predictive variables has an effect χ_{t-h} on the conditional probability. Thus, if the longer history of explanatory variables included in χ_{t-h} are useful to predict the future market status, the autoregressive equation (19) may offer a parsimonious way to specify the predictive model. A natural extension would be the dynamic autoregressive model:

$$\pi_t = \omega + \alpha(s_{t-1}) + \delta(\pi_{t-1}) + \mathbf{x}'_{t-h}\boldsymbol{\beta} \quad (21)$$

Forecasts on the future state of the economy based on a probit model follow the basic principles discussed in Kauppi and Saikkonen (2008). Of course one could consider higher order lags of variables s_t and π_t , but as argued in Kauppi and Saikkonen (2008) a first order lag structure usually suffices in forecasting. In general a forecast of the state variable s_t at time $t-h$ is the conditional expectation $E_{t-h}(s_t)$. According to the law of iterated conditional expectations and equation (17):

$$E_{t-h}(s_t) = E_{t-h}(\Phi(\pi_t)) \quad (22)$$

In order to compute the linear model of equation (18) or the autoregressive model of equation (20) we just have to plug the linear model into the conditional expectations expression (22). For instance for $h=2$ the autoregressive model (20) is given by recursive substitution from equation:

$$\pi_t = \omega + \delta_1(\pi_{t-1}) + \mathbf{x}'_{t-2}\boldsymbol{\beta} = (1 + \alpha_1)\omega_1 + \alpha_1^2\pi_{t-2} + \alpha_1\mathbf{x}'_{t-3}\boldsymbol{\beta} + \mathbf{x}'_{t-2}\boldsymbol{\beta} \quad (23)$$

which shows that the dependent π_t depends only on past information. Nevertheless, when the lagged values of the state variable s_t are considered, the case is more complicated since we depend upon the past values and the possible paths of the state variable that is unknown at the time of the forecast. To illustrate that, let's consider again two periods ahead forecasts as:

$$E_{t-2}(s_t) = E_{t-2}(\Phi(\omega + \alpha_1 s_{t-1} + \delta_1 \pi_{t-1} + \mathbf{x}'_{t-2}\boldsymbol{\beta})) \quad (24)$$

$$\text{given } E_{t-2}(s_t) = \begin{cases} \tilde{\Phi}(0), & \text{if } s_{t-1} = 0 \\ \tilde{\Phi}(1), & \text{if } s_{t-1} = 1 \end{cases}$$

where $\tilde{\Phi}(0)$ and $\tilde{\Phi}(1)$ are two possible outcomes depending on the value of the state variable at $t-1$. With a little manipulations, equation (24) can be written as

$$E_{t-2}(s_t) = (1 - p_{t-1})\tilde{\Phi}(0) + p_{t-1}\tilde{\Phi}(0) \quad (25)$$

As we observe, the forecasted probability is conditioned upon the probability of the possible outcomes of the previous periods. When the forecasting horizon extends to more than two periods ahead, the conditional probability becomes very complicated since we have to calculate the probability of all possible outcomes. For instance for six periods ahead forecasting we have to compute the probabilities for all $2^7=128$ possible combinations. In order to overcome this obstacle, we use several one-period ahead models, trained iteratively in order to obtain forecasts of the value of the s_{t-1} and then we return to the initial forecasting model with $h>1$ and obtain direct forecasts.

3. Empirical Results

We build our forecasting models starting from a simple model of only a constant term and add the real (or nominal) oil price and stock returns, the term spread, the future expected term spread and the term premium. Thus, following the notation of the probit models, the regressors' matrix and the coefficients' vector would be $\mathbf{x}_{t-h} = \{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS), TP\}$ and $\boldsymbol{\beta} = \{1, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5\}$ included recursively to the model. For the oil and stock prices we considered either real or nominal prices.

In order to measure the statistical significance of each model we perform a recursive Likelihood Ratio (LR) test, comparing the log-likelihood between the constrained "poorer" model with less regressors against a "richer" unconstrained model where we consider more explanatory variables. The null hypothesis is that the restricted model predicts better than the unrestricted one.

We forecast recessions for $h=1, 3, 6, 12, 24$ and 36 months ahead. All models are trained in the period January 1871 - December 1945 (901 observations), while the period January 1946 - June 2016 (845 observations) is kept aside during training for out-of-sample forecasting. We evaluate the forecasting accuracy of each model according to the quadratic probability score (QPS, Diebold and Rudebusch, 1989):

$$QPS = \frac{1}{m} \sum_{i=1}^m 2(s_t - E_{t-h}(s_t))^2 \quad (26)$$

where $E_{t-h}(s_t) = E_{t-h}(\Phi(\pi_t))$. QPS can be seen as the equivalent to Mean Square Error for classification models. It ranges from 0 to 2 with the smallest prices denoting the smallest forecasting error. In Table 2 we report the in-sample QPS statistic for the static and the dynamic probit models that are statistically significant at 5% level of significance.

Table 2: In- sample QPS of probit models								
Variables (χ_{t-h})	Model	h=1	h=3	h=6	h=12	h=18	h=24	h=36
Real prices								
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500))\}$	Static					0.494		0.495
	AR Static		0.212					
	Dynamic							0.495
	AR Dynamic	0.077	0.212					
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	Static	0.471	0.471	0.477	0.488	0.493	0.484	0.495
	AR Static	0.077	0.212	0.356	0.472	0.478	0.466	0.479
	Dynamic	0.397	0.414	0.441	0.483	0.442	0.444	0.495
	AR Dynamic	0.077	0.212	0.355	0.405			0.430
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS)\}$	Static							
	AR Static							
	Dynamic							
	AR Dynamic	0.076						
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS), TP\}$	Static	0.076	0.429	0.430				
	AR Static	0.076	0.204	0.338		0.441	0.425	
	Dynamic	0.353	0.368	0.396		0.416		0.425
	AR Dynamic	0.076	0.204	0.341				0.351
Nominal Prices								
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500))\}$	Static					0.494		0.494
	AR Static		0.212					
	Dynamic							
	AR Dynamic	0.077	0.213	0.356				
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	Static	0.460	0.462	0.472	0.487	0.493	0.486	0.494
	AR Static	0.077	0.212	0.354	0.461			
	Dynamic	0.362	0.393	0.435	0.480	0.416		
	AR Dynamic	0.077	0.212	0.354	0.393			0.403
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS)\}$	Static							
	AR Static							
	Dynamic							
	AR Dynamic	0.076						
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS), TP\}$	Static	0.428	0.421	0.424				
	AR Static	0.076	0.204	0.335	0.423			
	Dynamic	0.326	0.352	0.390				0.421
	AR Dynamic	0.076	0.204	0.333				0.340

Note: Static model corresponds to the model of equation (18), AR Static to the model of equation (19), Dynamic to model (20) and AR Dynamic to model (21). We report only the statistically significant results at 5% level of significance according to the LR test of model specification. We compare the log-likelihood between the constrained “poorer” model with less regressors against a “richer” unconstrained model where we consider more explanatory variables following the augmentation of the variable matrix in the order discussed in the Table. The null hypothesis is that the restricted model predicts better than the unrestricted one. We do not use tests based in the Mean Square Error (MSE) as Clark and West (2007) and McCracken (2007) since SVM models produce only directional forecasts where a MSE is not applicable.

As we observe from Table 2, when we use as regressors the oil and stock returns we do not get models that outperform a model with only a constant with statistical significance. In contrast, the addition of the term spread (TS) leads to models that are statistically significant more accurate than only with the oils and stock

returns. The same applies for up to 6 months with the inclusion of the term premium (TP) and not with the expected term spread ($E_{t-h}(TS)$). Regarding the model structure, we observe that the dynamic AR model expressions of equations (19) and (22) (inclusion of the first lag of the state variable and both the state variable and the probability) outperform the static (18) and the dynamic AR (20) models (inclusion only of the lagged probability). Nevertheless, the true forecasting ability of a model is measured in out-of-sample forecasting (Table 3).

Table 3: Out- of - sample QPS of probit models								
Variables (χ_{t-h})	Model	h=1	h=3	h=6	h=12	h=18	h=24	h=36
Real prices								
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500))\}$	Static					0.435		0.439
	AR Static		0.065					
	Dynamic							0.426
	AR Dynamic	0.050	0.064					
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	Static	0.419	0.410	0.410	0.426	0.434	0.443	0.439
	AR Static	0.050	0.064	0.129	0.309	0.321	0.320	0.324
	Dynamic	0.438	0.431	0.439	0.448	0.505	0.503	0.428
	AR Dynamic	0.050	0.063	0.127	0.200			0.215
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS)\}$	Static							
	AR Static							
	Dynamic							
	AR Dynamic	0.048						
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS), TP\}$	Static		0.541	0.497				
	AR Static	0.048	0.057	0.128		0.476	0.463	
	Dynamic	0.409	0.401	0.424		0.510		0.511
	AR Dynamic	0.048	0.057	0.127				0.290
Nominal Prices								
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500))\}$	Static					0.437		0.440
	AR Static		0.065					
	Dynamic							
	AR Dynamic	0.050	0.064	0.125				
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	Static	0.402	0.395	0.401	0.422	0.437	0.448	0.440
	AR Static	0.050	0.064	0.127	0.276			
	Dynamic	0.416	0.409	0.424	0.448	0.513		
	AR Dynamic	0.050	0.063	0.125	0.185			0.199
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS)\}$	Static							
	AR Static							
	Dynamic							
	AR Dynamic	0.048						
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS), TP\}$	Static	0.493	0.434	0.434				
	AR Static	0.048	0.059	0.131	0.301			
	Dynamic	0.367	0.358	0.390				0.502
	AR Dynamic	0.048	0.059	0.124				0.263

Note: Static model corresponds to the model of equation (18), AR Static to the model of equation (19), Dynamic to model (20) and AR Dynamic to model (21). We report only the statistically significant results at 5% level of significance according to the LR test of model specification. We compare the log-likelihood between the constrained “poorer” model with less regressors against a “richer” unconstrained model where we consider more explanatory variables following the augmentation of the variable matrix in the order discussed in the Table. The null hypothesis is that the restricted model predicts better than the unrestricted one. We do not use tests based in the Mean Square Error (MSE) as Clark and West (2007) and McCracken (2007) since SVM models produce only directional forecasts where a MSE is not applicable.

The reported QPS statistic is higher in longer forecasting horizons denoting higher forecasting error, as expected due to the higher uncertainty of long –term forecasts over shorter ones. Again we observe that the most accurate forecasting

models are the ones that include the first lag of the state variable s_{t-1} as regressor. We find no significant difference in the forecasting error whether we include nominal or real prices of oil and stock returns. In contrast, a significant increase in the forecasting ability comes with the inclusion of the term spread.

In Tables 4 and 5 we depict the QPS statistics of an SVM-linear model in in-sample and out-of-sample forecasting, respectively. Motivated by the high accuracy of including first lag of the state variable s_{t-1} we train models only with the variables as regressors (coded Static) and with the inclusion of the state variable (coded AR). We also use the iterative approach in order to forecast the values of the first lag of the state variable when we forecast further than one month ahead. The reported results are all statistically significant at 5% level of significance according to the LR tests.

Table 4: In-sample QPS of SVM-linear models								
Variables (χ_{t-h})	Model	h=1	h=3	h=6	h=12	h=18	h=24	h=36
Real prices								
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500))\}$	Static	0.478	0.472	0.482	0.494	0.495	0.496	0.497
	AR	0.077	0.147	0.148	0.149	0.150	0.151	0.149
$\left\{ \begin{array}{c} \omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), \\ TS \end{array} \right\}$	Static	0.473	0.472	0.482	0.494	0.495	0.496	0.496
	AR	0.077	0.147	0.148	0.149	0.150	0.151	0.149
$\left\{ \begin{array}{c} \omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), \\ TS, E_{t-h}(TS) \end{array} \right\}$	Static	0.445	0.434	0.434	0.464	0.495	0.494	0.448
	AR	0.077	0.147	0.148	0.149	0.150	0.151	0.149
$\left\{ \begin{array}{c} \omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), \\ TS, E_{t-h}(TS), TP \end{array} \right\}$	Static	0.439	0.429	0.429	0.458	0.481	0.461	0.437
	AR	0.077	0.147	0.148	0.149	0.150	0.151	0.149
Nominal Prices								
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500))\}$	Static	0.464	0.464	0.473	0.494	0.495	0.496	0.499
	AR	0.077	0.147	0.148	0.149	0.150	0.151	0.149
$\left\{ \begin{array}{c} \omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), \\ TS \end{array} \right\}$	Static	0.460	0.462	0.473	0.494	0.495	0.496	0.500
	AR	0.077	0.147	0.148	0.149	0.150	0.151	0.149
$\left\{ \begin{array}{c} \omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), \\ TS, E_{t-h}(TS), TP \end{array} \right\}$	Static	0.430	0.423	0.428	0.463	0.495	0.496	0.445
	AR	0.077	0.147	0.148	0.149	0.150	0.151	0.149
$\left\{ \begin{array}{c} \omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), \\ TS, E_{t-h}(TS), TP \end{array} \right\}$	Static	0.427	0.419	0.423	0.456	0.485	0.467	0.439
	AR	0.077	0.147	0.148	0.149	0.150	0.151	0.149

Note: Static denotes models that include the various variables as regressors, while AR denotes the additional inclusion of the first lag of the state variable s_{t-1} . The reported are statistically significant at 5% level of significance according to an LR test of model specification. We compare the log-likelihood between the constrained “poorer” model with less regressors against a “richer” unconstrained model where we consider more explanatory variables following the augmentation of these variable matrix in the order discussed in the Table. The null hypothesis is that the restricted model predicts better than the unrestricted one. We do not use tests based in the Mean Square Error (MSE) as Clark and West (2007) and McCracken (2007) since SVM models produce only directional forecasts where a MSE is not applicable.

Table 5: Out-of-sample QPS of SVM-linear models								
Variables (χ_{t-h})	Model	h=1	h=3	h=6	h=12	h=18	h=24	h=36
Real prices								
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500))\}$	Static	0.429	0.421	0.422	0.438	0.433	0.448	0.415
	AR	0.051	0.099	0.099	0.099	0.099	0.099	0.099
$\left\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), \right. \\ \left. TS \right\}$	Static	0.425	0.427	0.421	0.436	0.435	0.448	0.440
	AR	0.051	0.099	0.099	0.099	0.099	0.099	0.099
$\left\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), \right. \\ \left. TS, E_{t-h}(TS) \right\}$	Static	0.335	0.295	0.273	0.327	0.438	0.436	0.350
	AR	0.051	0.099	0.099	0.099	0.099	0.099	0.099
$\left\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), \right. \\ \left. TS, E_{t-h}(TS), TP \right\}$	Static	0.535	0.522	0.415	0.527	0.547	0.807	0.690
	AR	0.051	0.099	0.099	0.099	0.099	0.099	0.099
Nominal Prices								
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500))\}$	Static	0.408	0.401	0.407	0.435	0.438	0.451	0.460
	AR	0.051	0.099	0.099	0.099	0.099	0.099	0.099
$\left\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), \right. \\ \left. TS \right\}$	Static	0.402	0.396	0.405	0.435	0.439	0.451	0.394
	AR	0.051	0.099	0.099	0.099	0.099	0.099	0.099
$\left\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), \right. \\ \left. TS, E_{t-h}(TS), TP \right\}$	Static	0.312	0.284	0.266	0.318	0.433	0.440	0.338
	AR	0.051	0.099	0.099	0.099	0.099	0.099	0.099
$\left\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), \right. \\ \left. TS, E_{t-h}(TS), TP \right\}$	Static	0.411	0.396	0.405	0.499	0.505	0.796	0.456
	AR	0.051	0.099	0.099	0.099	0.099	0.099	0.099

Note: Static denotes models that include the various variables as regressors, while AR denotes the additional inclusion of the first lag of the state variable s_{t-1} . The reported are statistically significant at 5% level of significance according to an LR test of model specification. We compare the log-likelihood between the constrained “poorer” model with less regressors against a “richer” unconstrained model where we consider more explanatory variables following the augmentation of these variable matrix in the order discussed in the Table. The null hypothesis is that the restricted model predicts better than the unrestricted one. We do not use tests based in the Mean Square Error (MSE) as Clark and West (2007) and McCracken (2007) since SVM models produce only directional forecasts where a MSE is not applicable.

The results of the SVR-linear model indicate the same pattern as in the dynamic probit models; the inclusion of the lagged state variable exhibits the smallest forecasting error. Since the linear kernel achieves quantitative similar forecasting accuracy with the most accurate probit models, we repeat our forecasting exercise based on the RBF kernel. The in-sample and out-of-sample QPS statistics for the SVM-RBF model are reported in Tables 6 and 7, respectively.

Table 6: In-sample QPS of SVM-RBF models								
Variables (χ_{t-h})	Model	h=1	h=3	h=6	h=12	h=18	h=24	h=36
Real prices								
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500))\}$	Static	0.472	0.472	0.461	0.483	0.491	0.454	0.469
	AR	0.077	0.147	0.148	0.149	0.150	0.151	0.149
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	Static	0.461	0.458	0.409	0.457	0.491	0.447	0.458
	AR	0.075	0.129	0.146	0.147	0.148	0.149	0.149
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS)\}$	Static	0.404	0.409	0.415	0.441	0.291	0.271	0.407
	AR	0.075		0.146	0.147	0.148	0.149	0.149
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS), TP\}$	Static	0.351	0.342	0.358	0.416	0.022	0.037	
	AR	0.076	0.147	0.148	0.148	0.150	0.151	0.148
Nominal Prices								
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500))\}$	Static	0.453	0.455	0.463	0.488	0.487	0.473	0.491
	AR	0.077	0.148	0.148	0.149	0.150	0.151	0.149
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	Static	0.401	0.416	0.450	0.475	0.489	0.467	0.477
	AR	0.075	0.146	0.146	0.147	0.148	0.149	0.149
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS), TP\}$	Static		0.332	0.326	0.407	0.399	0.428	0.446
	AR	0.075	0.146	0.146	0.147	0.148	0.149	0.149
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS), TP\}$	Static	0.397			0.414	0.291	0.220	0.424
	AR	0.076	0.147	0.148	0.149	0.150	0.151	0.149

Note: Static denotes models that include the various variables as regressors, while AR denotes the additional inclusion of the first lag of the state variable s_{t-1} . The reported are statistically significant at 5% level of significance according to an LR test of model specification. We compare the log-likelihood between the constrained “poorer” model with less regressors against a “richer” unconstrained model where we consider more explanatory variables following the augmentation of the variable matrix in the order discussed in the Table. The null hypothesis is that the restricted model predicts better than the unrestricted one. We do not use tests based in the Mean Square Error (MSE) as Clark and West (2007) and McCracken (2007) since SVM models produce only directional forecasts where a MSE is not applicable.

Table 7: Out-of-sample QPS of SVM-RBF models								
Variables (χ_{t-h})	Model	h=1	h=3	h=6	h=12	h=18	h=24	h=36
Real prices								
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500))\}$	Static	0.431	0.490	0.422	0.430	0.438	0.473	0.558
	AR	0.051	0.099	0.099	0.099	0.099	0.099	0.099
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	Static	0.412	0.402	0.422	0.427	0.438	0.448	0.591
	AR	0.051	0.133	0.099	0.099	0.099	0.124	0.100
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS)\}$	Static	0.324	0.299	0.278	0.312	0.480	0.482	0.538
	AR	0.051		0.099	0.099	0.100	0.101	0.100
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS), TP\}$	Static	0.682	0.763	0.657	0.478	0.434	0.450	
	AR	0.097	0.148	0.152	0.137	0.146	0.157	0.258
Nominal Prices								
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500))\}$	Static	0.386	0.383	0.398	0.423	0.433	0.444	0.439
	AR	0.051	0.098	0.099	0.099	0.099	0.099	0.099
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	Static	0.366	0.374	0.403	0.423	0.434	0.437	0.427
	AR	0.051	0.098	0.098	0.099	0.099	0.099	0.099
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS), TP\}$	Static		0.336	0.318	0.338	0.444	0.403	0.339
	AR	0.051	0.098	0.099	0.099	0.099	0.099	0.099
$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS), TP\}$	Static	0.376			0.342	0.424	0.453	0.659
	AR	0.093	0.145	0.148	0.145	0.151	0.164	0.267

Note: Static denotes models that include the various variables as regressors, while AR denotes the additional inclusion of the first lag of the state variable s_{t-1} . The reported are statistically significant at 5% level of significance according to an LR test of model specification. We compare the log-likelihood between the constrained “poorer” model with less regressors against a “richer” unconstrained model where we consider more explanatory variables following the augmentation of the variable matrix in the order discussed in the Table. The null hypothesis is that the restricted model predicts better than the unrestricted one. We do not use tests based in the Mean Square Error (MSE) as Clark and West (2007) and McCracken (2007) since SVM models produce only directional forecasts where a MSE is not applicable.

The forecasting results based on the RBF kernel do not improve the forecasting accuracy over the probit or the SVR-linear models. Again the most

accurate models are the ones that include the first lag value of the state variable. Nominal prices of oil and the stock returns exhibit marginally higher out-of-sample forecasting accuracy than real prices in out-of-sample forecasting. Given the volume of the reported results, in Table 8 we present the models that exhibit the lowest forecasting error per forecasting horizon for the probit and SVR models. In doing so, we select the models with the highest out-of-sample forecasting accuracy. In cases where models exhibit similar forecasting accuracy, we follow an *Occam's razor* approach selecting the models with the highest parsimony (less input variables) that exhibit the smallest forecasting error in in-sample and out-of-sample forecasting.

Table 8: Most accurate models per horizon				
Horizon	Variables (χ_{t-h})	Model	In-sample QPS	Out-of-sample QPS
Probit				
1	$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS)\}$	AR Dynamic	0.076	0.048
3	$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS), TP\}$	AR Static AR Dynamic	0.204	0.057
6	$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS, E_{t-h}(TS), TP\}$	AR Dynamic	0.333	0.124
12	$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	AR Dynamic	0.393	0.124
18	$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	AR Static	0.478	0.321
24	$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	AR Static	0.466	0.320
36	$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	AR Dynamic	0.403	0.199
SVM				
1	$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	AR RBF	0.075	0.051
3	$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	AR RBF	0.146	0.098
6	$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	AR RBF	0.146	0.098
12	$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	AR RBF	0.147	0.099
18	$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	AR RBF	0.148	0.099
24	$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	AR RBF	0.141	0.099
36	$\{\omega, \Delta(\ln(WTI)), \Delta(\ln(SP500)), TS\}$	AR RBF AR Linear	0.149	0.099

Note: AR Static to the model of equation (19), AR Dynamic to model (21) and AR denotes the additional inclusion of the first lag of the state variable s_{t-1} in the SVM models. When models exhibit similar forecasting accuracy, we follow an Occam's razor approach selecting the models with the highest parsimony (less variables) and the best in-sample and out-of sample forecasting accuracy.

As we observe from Table 8, the probit models foresee recession periods more closely than SVM models for up to 6 months ahead, while the SVM models are more accurate in longer horizons. The most accurate forecasting model is the one that includes oil and stock returns and the term spread, while the expected term spread and the term premium add to the forecasting ability of the best models only in short-term forecasting. Bearing in mind that there is a significant delay between the economy falling in recession and the imprint in the economic indicators of the monetary authority, short term forecasting less than 6 months is of limited practical interest. Thus the stability of the AR SVM-RBF models in forecasting seems ideal in using such a methodology in an early warning mechanism of future recessions. In figures 4 and 5 we depict the forecasted probabilities of recession for the dynamic probit and SVM – RBF models of Table 8.

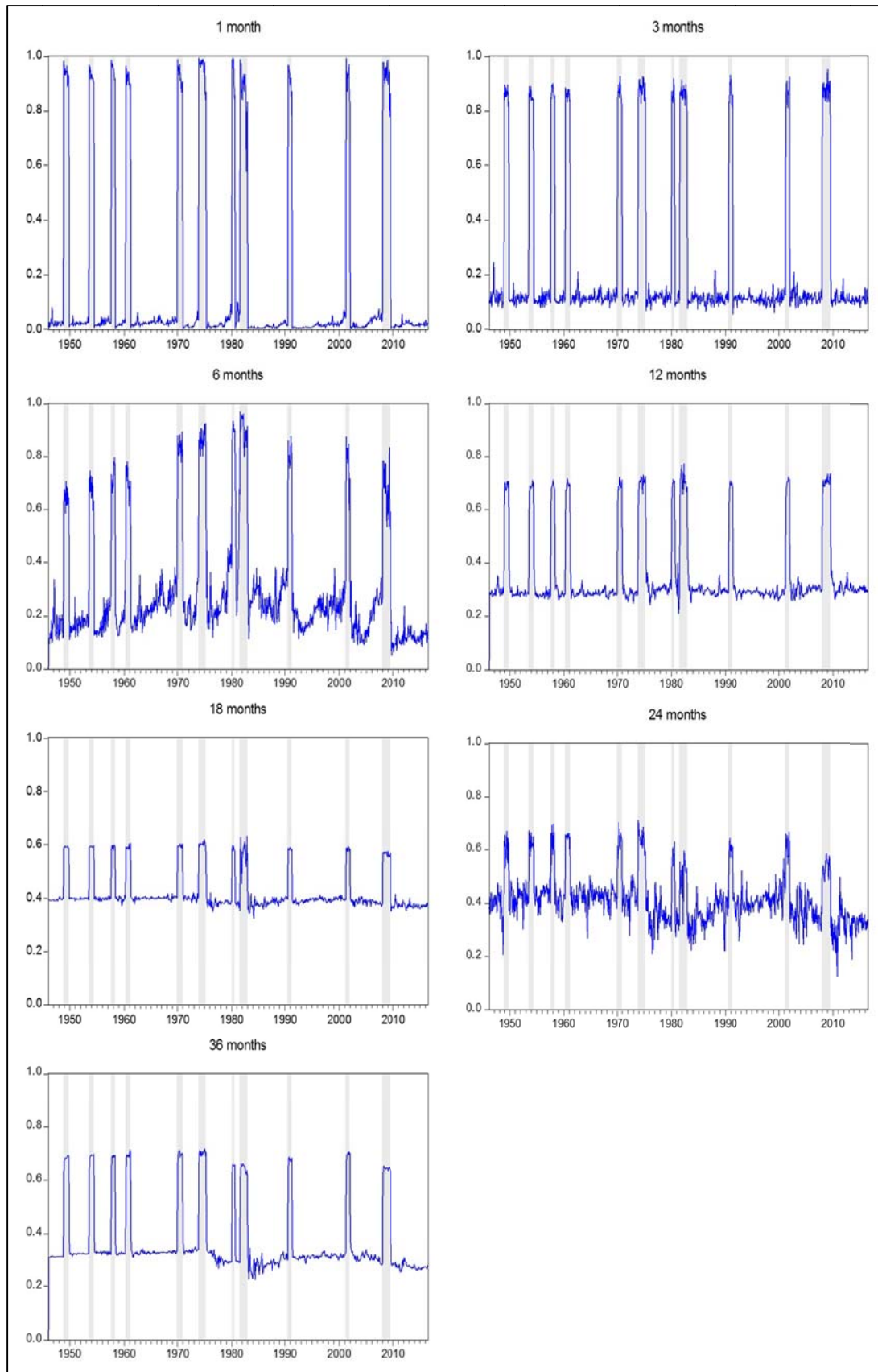


Figure 4: Out-of sample forecasted probabilities of the dynamic probit models. Grey areas denote NBER recessions.

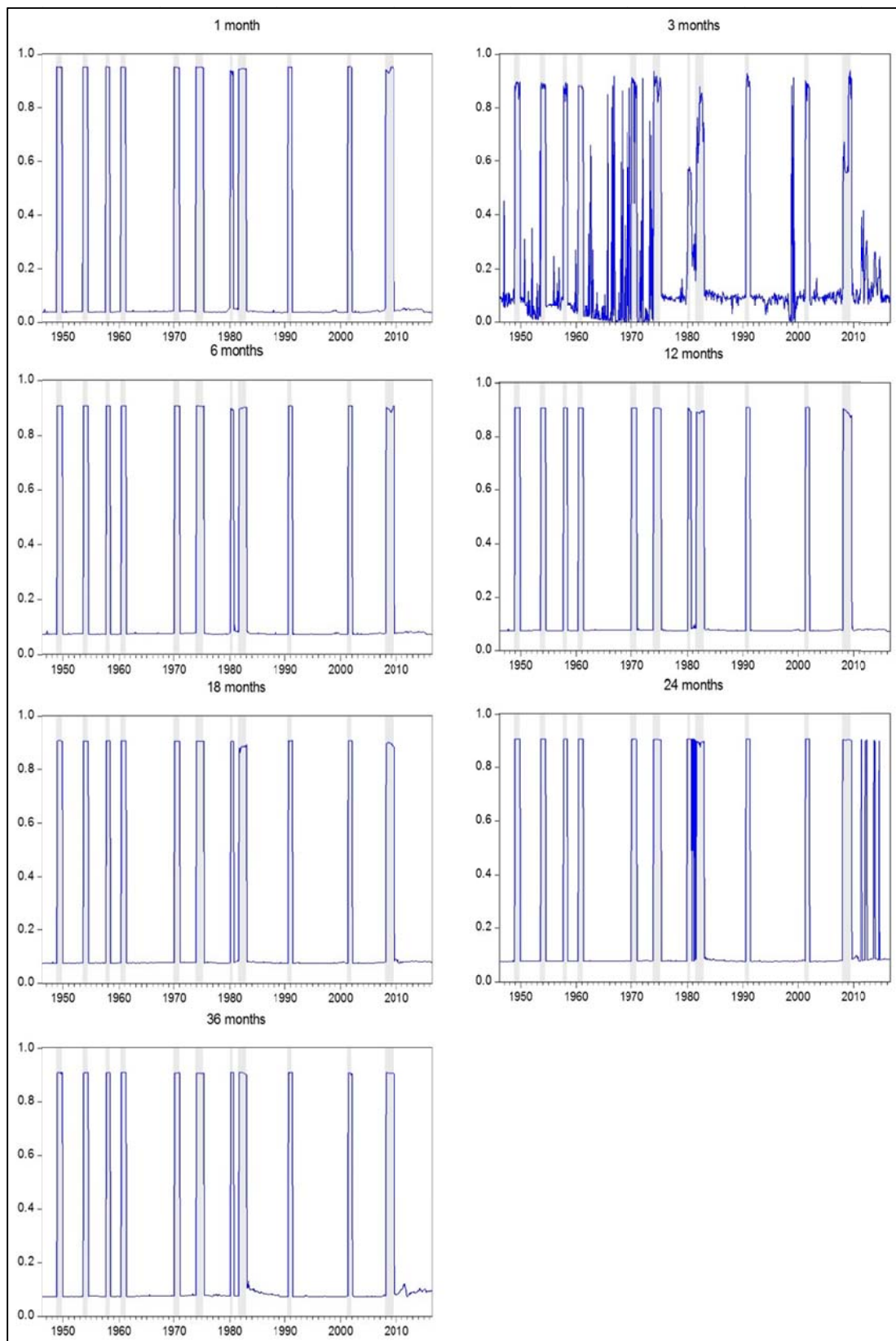


Figure 5: Out-of sample forecasted probabilities of the AR SVM-RBF models. Grey areas denote NBER recessions.

Both dynamic probit and SVM models forecast recession periods accurately, since they appear to have a probability of recession higher than 50% in each period. An exception is the 3-month ahead SVM model that forecasts more than actual periods as recessions. Interestingly, SVM models appear to discriminate between recessions and tranquil periods better than probit ones, since the forecasted probabilities between the two periods have a sharp difference.

4. Conclusion

The objective of this paper is to determine to what extent a selection of leading indicators are able to forecast U.S. recessions by means of both dynamic probit models and classification methods, such as the linear and non-linear Support Vector Machines (SVM) models, using monthly data from January 1871 to June 2016. As leading indicators, we use an ample selection of leading indicators, such as the yield spread, oil price shocks, stock returns and the term premium, and analyze the forecasting ability of each of them. In order to define recessions, we use the business cycle chronology proposed by the National Bureau of Economic Research (NBER). As far as the methodology is concerned, the paper uses both the probit and the SVM models to predict U.S. recessions and compare the accuracy of the predictions obtained with each of the methodologies. Furthermore, and in order to evaluate the accuracy of the predictions, the paper analyzes both in-sample and out-of-sample Quadratic Probability Score (QPS, Diebold and Rudebusch, 1989) for each of the models.

The main results suggest the following. First, concerning the methodology, the results suggest that the probit models foresee U.S. recession periods more closely than SVM models for up to 6 months ahead, while the SVM models are more accurate in longer horizons. Furthermore, the most accurate forecasting models include oil, stock returns and the term spread as leading indicators. Taking into account that there is a significant delay between the economy entering into a recession and the imprint in the economic indicators of the monetary authority, short term forecasting could be of limited practical interest. Furthermore, SVM models appear to discriminate between recessions and tranquil periods better than probit ones. Therefore, and according to our results, SVM models seem more appropriate to predict economic recessions than the usually employed probit models.

Finally, and in line with most of the literature, the results also suggest that the most accurate forecasting models include the oil prices, stock returns and the term spread as leading indicators.

References

- Ang, A., Piazzesi, M., and Wei, M., 2006. What does the yield curve tell us about GDP? *Journal of Economics* 131, 1–2.
- Berge, T. J., & Jordà, Ò., 2011. Evaluating the classification of economic activity into recessions and expansions. *American Economic Journal: Macroeconomics* 3, 246–277.
- Berge, T.J., 2015. Predicting Recessions with Leading Indicators: Model Averaging and Selection over the Business Cycle, *Journal of Forecasting* 34, 455-471.
- Bernanke, B.S., 1990. On the predictive power of interest rates and interest rate spread. NBER Working Paper 3486.
- Bernanke, B.S. and Blinder, A.S., 1989. The Federal funds rate and the channels of monetary transmission. *American Economic Review* 82, 901-927.
- Chen, S.S., 2009. Predicting the bear market: macroeconomic variables as leading indicators. *Journal of Banking and Finance* 33, 211–223.
- Chauvet, M. and Potter, S., 2002. Predicting a recession: Evidence from the yield curve in the presence of structural breaks. *Economic Letters* 77, 245-253.
- Chauvet, M. and Potter, S., 2005. Forecasting recessions using the yield curve. *Journal of Forecasting* 24, 77-103.
- Clark T. and West K. (2007) Approximately Normal Tests for Equal Predictive Accuracy in Nested Models, *Journal of Econometrics*, vol. 138, pp. 291-311.
- Diebold, F.X., Rudebusch, G.D., 1989. Scoring the leading indicators. *Journal of Business* 62, 369–391.
- Dueker, M.J. 1997. Strengthening the case for the yield curve as a predictor of U.S. recessions. *St. Louis Federal Review*, 41-51.
- Dueker, M.J. 2005. Dynamic forecasts of qualitative variables: A Qual VAR model of U.S. recessions. *Journal of Business and Economic Statistics* 23, 96-104.
- Engemann, K. M., Kliesen, K. L., and Owyang, M. T., 2010. Do Oil Shocks Drive Business Cycles? - Some U.S. and International Evidence. Federal Reserve Bank of St. Louis Working Paper Series 2010 007A.

- Estrella, A. and Hardouvelis, G. A., 1991. The term structure as a predictor of real economic activity. *Journal of Finance* 46, 555-576.
- Estrella, A., and F. S. Mishkin, 1997. The Predictive Power of the Term Structure of Interest Rates in Europe and the United States: Implications for the European Central Bank. *European Economic Review* 41, 1375–401.
- Estrella, A., and Mishkin, F. S., 1998. Predicting US recessions: Financial variables as leading indicators. *The Review of Economics and Statistics* 80, 45–61.
- Estrella, A., Rodrigues, A. P. and Schich, S., 2003. How stable is the predictive power of the yield curve? Evidence from Germany and the United States. *Review of Economics and Statistics* 85, 629-644.
- Gadea, M.D. and Perez-Quiros, G., 2015. The failure to predict the Great Recession. A view through the role of credit. *Journal of the European Economic Association* 13, 534-559.
- Gertler, M. and Lown, C. S., 1999. The information in the high-yield bond spread for the business cycle: evidence and some implications. *Oxford Review of Economic Policy* 15, 132-150.
- Giacomini, R. and Rossi, B., 2006. How stable is the forecasting performance of the yield curve for output growth? *Oxford Bulletin of Economics and Statistics* 68, 783-795.
- Gogas, P., Papadimitriou, T., and Chrysanthidou, E., 2015. Yield Curve Point Triplets in Recession Forecasting. *International Finance* 18, 207-226.
- Hamilton, J. D. and Kim, D. H., 2002. A re-examination of the predictability of the yield spread for real economic activity. *Journal of Money, Credit, and Banking* 34, 340-360.
- Hamilton, J.D. 2005. Oil and the macroeconomy. UCSD Economics Working Paper 0508.
- Hamilton J.D., 2011. Calling recessions in real time. *International Journal of Forecasting* 27, 1006-1026.
- Harvey, C. R., 1988. The real term structure and consumption growth. *Journal of Financial Economics* 22, 305-333.

- Harvey, C. R., 1989. Forecasts of Economic Growth from the Bond and Stock Markets. *Financial Analysts Journal* 45, 38–45.
- International Monetary Fund, 2009. *World Economic Outlook*, April 2009.
- Kauppi, H. and Saikkonen, P. 2008. Predicting U.S. recessions with dynamic binary response models. *The Review of Economics and Statistics* 90, 777-791.
- Killian L. and Vigfusson R., 2013. Do oil prices help forecast U.S. real GDP? The role of nonlinearities and asymmetries. *Journal of Business and Economic Statistics* 31, 78-93.
- Levanon, G., Manini, J.C., Ozyildirim, A., Schaitkin, B., and Tanchua, J., 2015. Using a leading credit index to predict turning points in the US business cycle. *International Journal of Forecasting* 31, 392–398.
- Liu, W. and Moench, E., 2016. What predicts US recessions? *International Journal of Forecasting* 32, 1138-1150.
- McCracken W. (2007) Asymptotics for out of sample tests of Granger causality, *Journal of Econometrics*, vol. 140, pp. 719-752.
- Mody, A. and Taylor, M. P., 2003. The high-yield spread as a predictor of real economic activity: evidence of a financial accelerator for the United States. *IMF Staff Papers* 50, 373-402.
- Ng, S. and Wright, J. H., 2013. Facts and challenges from the Great Recession for forecasting and macroeconomic modeling. *Journal of Economic Literature* 51, 1120-1154.
- Platt J., 2000. Probabilistic Output for Support Vector Machines and comparison to Regularized Log Likelihood Methods. In: A. Smola, P. Bartlett, B. Schölkopf, and D. Schuurmans (eds.): *Advances in Large Margin Classifiers*. Cambridge, MA.
- Rossi, B. and Sekhposyan, T., 2011. Have economic models' forecasting performance for US output growth and inflation changed over time, and when? *International Journal of Forecasting* 26, 808-835.
- Rudebusch, G.D. and Williams, J.C., 2009. Forecasting recessions: the puzzle of the enduring power of the yield curve. *Journal of Business and Economic Statistics* 27, 492-503.

Stock, J. H., and Watson, M. W., 1989. New indexes of coincident and leading economic indicators. In NBER Macroeconomics Annual, 4, 351-409. National Bureau of Economic Research, Inc., NBER Chapters.

Stock, J. H., and Watson, M. W., 2003. Forecasting output and inflation: the role of asset prices. *Journal of Economic Literature* 41, 788-829.