

# **Yield curve and Recession Forecasting in a Machine Learning Framework**

**Theophilos Papadimitriou, Periklis Gogas, Maria Matthaiou, Efthymia Chrysanthidou**

Department of Economics  
Democritus University of Thrace  
Komotini 69100, Greece

## **Abstract**

In this paper, we investigate the forecasting ability of the yield curve in terms of the U.S. real GDP cycle. More specifically, within a Machine Learning (ML) framework, we use data from a variety of short (treasury bills) and long term interest rates (bonds) for the period from 1976:Q3 to 2011:Q4 in conjunction with the real GDP for the same period, to create a model that can successfully forecast output fluctuations (inflation and output gaps) around its long-run trend. We focus our attention in correctly forecasting the instances of output gaps referred for the purposes of our analysis here as recessions. In this effort, we applied a Support Vector Machines (SVM) technique for classification. The results show that we can achieve an overall forecasting accuracy of 66,7% and a 100% accuracy in forecasting recessions.

## **1. Introduction**

The yield curve depicts the relationship of the yields of bonds with different maturities issued by the same entity, usually a sovereign government. The stylized facts show that it is usually upward sloping reflecting a positive liquidity premium. An inverted yield curve (negative slope) is considered a sign of an upcoming recession.

After the recent crises in the economy worldwide, the issue of forecasting output gaps is very topical and has significant implications in terms of economic policy. Prompt warning of an upcoming output downturn is of course extremely important to policymakers as they can swiftly adjust fiscal and monetary policy to either steer the economy away from a recession or dampen its impact to the real economy. The yield curve contains information on bonds' returns of various maturities; short term interest rates have a direct correlation with monetary policy implementation while long term bonds' yield-to-maturity reflect the expectations of investors on the future economic activity (Estrella & Mishkin 1996, 1997). For this reason many studies have tried to exploit the information included in the yield curve in forecasting the economic activity of a country or a region.

The methodology used by these studies can be distinguished in two groups in terms of methodology employed to forecast the GDP. In one hand, we have linear models as we can see in studies such as Ang, Piazzesi & Wei (2006), Mishkin (1997), Bernanke,

Boivin & Elias (2005), Diebold & Li (2006) and Moench (2008), that use methods based on Ordinary Least Square (OLS) and Vector Auto regression (VAR). Other studies have used non-linear models; see for example Ang, Piazzesi & Wei, (2006), Chauvet & Potter (2002), Chionis, Gogas & Pragidis (2009), Estrella & Hardouvelis (1991), Estrella & Mishkin (1996) and Wright (2006) use probit models in order to forecast economic output by defining the value of a dummy variable where value "0" indicates an output gap and value "1" an inflationary gap.

Most of the studies use the spread of interest rates and more specifically the spread between a long and a short term rate, in order to forecast the economic activity. There are also a number of studies claiming that the inclusion of macro variables and financial variables can improve the forecasting accuracy of such a model. Bordo & Haubrich (2008) use both slope and level (a short term interest rate such as the Federal Funds Rate) of the curve in their forecasting model. Ang & Piazzesi (2003) estimate models that include macroeconomic factors of inflation and economic growth and also measures of level, curvature and slope of the yield curve called latent factors. The study finds that macro factors explain up to 85% of the forecast variance for long forecast horizons, while 60% is explained for a one-month forecast horizon. In a later work Ang, Piazzesi & Wei (2006) use a dynamic model in order to characterize the expectations of GDP. The model that includes a short one-month term rate, a five years spread and lagged GDP is more efficient in forecasting real output.

Diebold, Rudebusch & Aruoba (2004) in order to find the dynamic interactions between the macroeconomy and the yield curve, they construct a model with the latent factors (level, slope and curvature) of the yield curve and macroeconomic variables of real activity, inflation and the stance of monetary policy. The results show that the macroeconomic variables have a significant effect in future movements of the yield curve. Another study that includes a variety of macro variables is that of Moench (2008) which uses the common components of a large number of macroeconomic variables and the short rate term in order to estimate the model. In this article no latent factors are used and it concludes in a 50% reduction of the root mean squared forecast errors for short and 20% reduction for long maturities. Finally, Chionis, Gogas & Pragidis (2009) include in their model the European Central Bank's euro area government benchmark bonds of various maturities, as well as non-monetary policy variables (the unemployment and a composite European stock price index). The results show that the yield curve augmented with the composite stock index has significant forecasting power in terms of the EU15 real output.

The purpose of this study is to create a forecasting model using information from the yield curve that will be able to forecast future economic activity (real GDP above or below the potential GDP) by means of a Support Vector Machine (SVM) classifier. The Support Vector Machines are binary classifiers that to the best of our knowledge are employed for the first time to forecast the GDP cycle using the information provided by the yield curve. For this reason the dependent variable series data are

classified in two classes: a) Class "-1" denoting an output gap (real output below long-run trend) and b) Class "1" that denotes an inflationary gap (real output above the long run trend). The results show that both the monetary policy signals and the expectations of investors about future economic activity play an important role in forecasting future deviations of real GDP from the long-run trend and our model is especially successful in forecasting future recessions.

The rest of the article is structured as follows: Section II outlines the data set used and describes the methodology of the SVM. The empirical results from our model selection process are presented in Section III and finally the last Section concludes.

## **II. Data and Methodology**

Both short-term and long-term U.S federal government interest rates were used in our study. The short-term interest rates are from Treasury-Bills with maturities of 3 and 6 months. The long-term interest rates are from the U.S government bonds with maturities of 2, 3, 5, 7 and 10 years. Real seasonally adjusted U.S GDP is used as the basis for our binary dependent variables. The data span the period from 1967:Q3 to 2011:Q4 and they were available from the database of the Federal Reserve Bank of Saint Louis, Federal Reserve Economic Data (FRED). The real GDP figures were transformed into natural logarithms. We define recession in our study as deviations of GDP under the long-term trend (output gap). To decompose the GDP series and obtain the cyclical component and the trend series we use the Hodrick-Prescott filter by setting the  $\lambda$  parameter equal to 1600 (Hodrick, Prescott 1997). The cyclical component series is then transformed into a binary (dummy) variable that has the value of one whenever the cyclical component is greater than zero and minus one elsewhere. We focused on forecasting the cyclical component of real GDP in 1, 2 and 3 quarters ahead forecasting windows using the SVM methodology.

The Support Vector Machines represent a supervised machine learning method used for two-class data classification. Essentially, SVM try to locate a small number of data points from our dataset, called Support Vectors (SV) that can define a hyperplane separating the two classes' data points perfectly or as perfect as possible<sup>1</sup> (see Figure 1). The method has two basic steps: the training step and the testing step. During the training step, the largest part of the dataset is used for the estimation of the separating hyperplane through a minimization process; in the testing step, the generalization ability of the model is evaluated by investigating the model's performance in the small subset that was left aside in the first step.

---

<sup>1</sup> imposing a weight parameter  $C$  for each erroneously classified data point in the minimization procedure

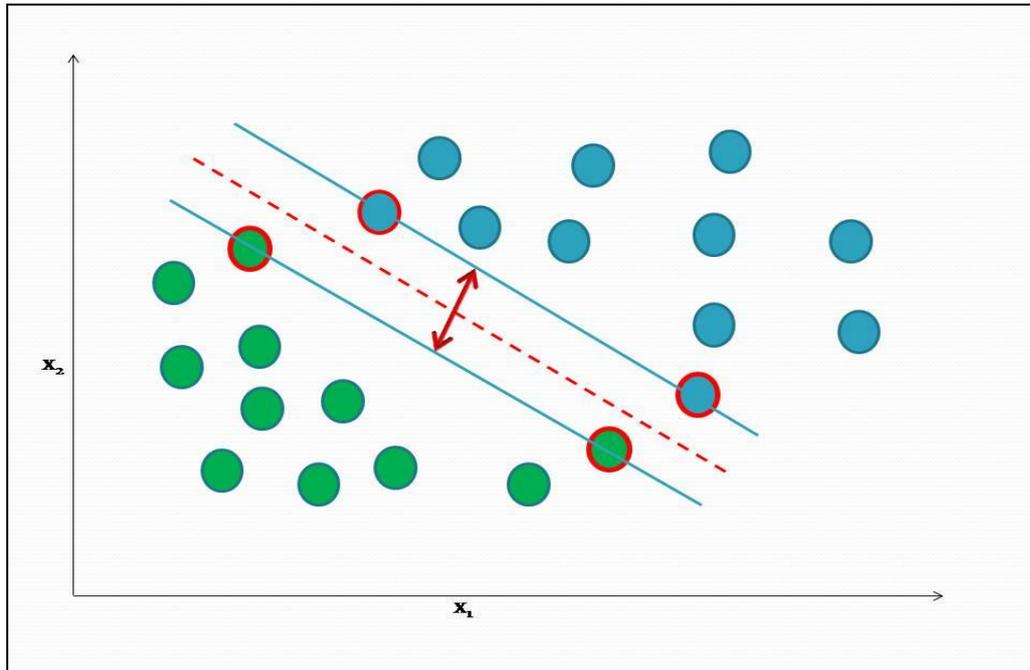


Figure 1 Hyper plane selection and support vectors. The SV's are represented with the pronounced red contour, the margin lines are represented with the continuous lines and the hyper plane is represented with the dotted line.

In cases that the dataset is not linearly-separable, then SVM is coupled with a non-linear Kernel mapping procedure, projecting the data points to a higher dimensional space, called feature space, where the classes are linearly separable (see Figure 2).

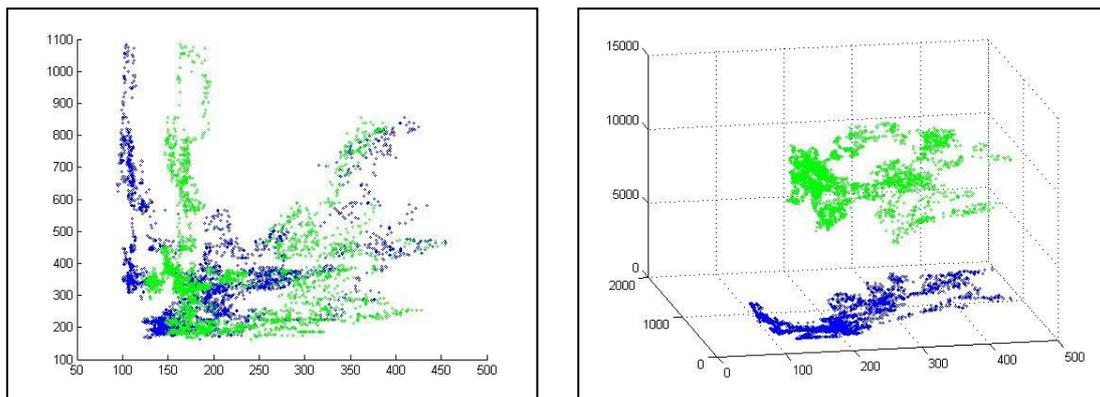


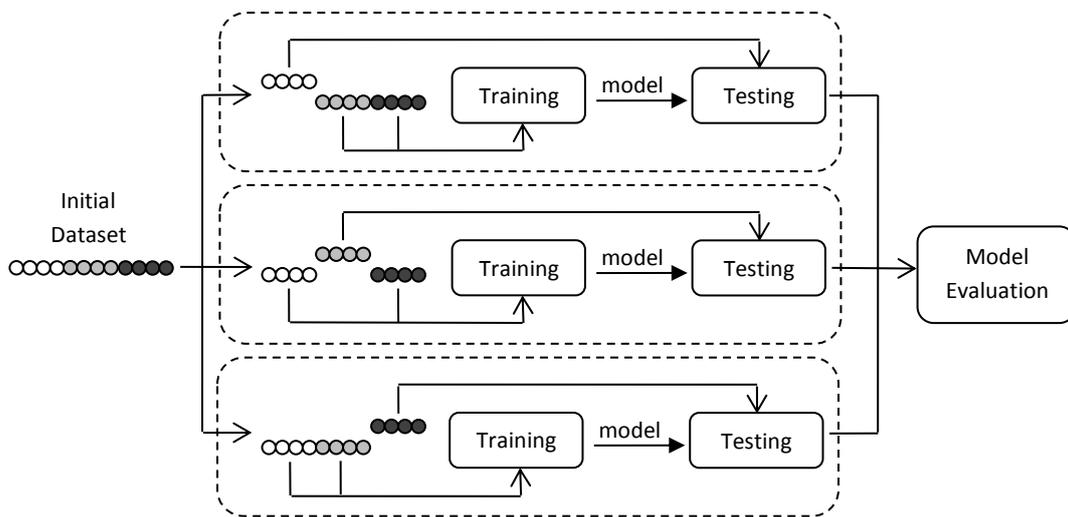
Figure 2: (left) the data space. The green and blue classes cannot be separated linearly. (right) the feature space. The data points are projected into a richer space and the two classes are linearly separable.

In our tests, we used a) the Linear kernel and b) Radial basis function (RBF). Their mapping function is:

- *Linear*:  $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$
- *Radial basis function- RBF*:  $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2)$ ,  $\gamma > 0$

The error tolerance parameter  $C$  and the RBF parameter  $\gamma$  were investigated through a grid search.

In a machine learning scheme, training results in overfitting when the model produced is significantly affected by possible noise in the sample in hand instead of the true underlying relationship that describes the phenomenon. Usually, overfitting yields a very high performance on the training step and significantly lower accuracy on the testing step.  $k$ -fold cross validation, is adopted to avoid overfitting. The dataset is cut into  $k$  chunks and the training-testing steps are repeated  $k$  times. In each turn a different chunk is used as the test dataset, while the rest  $k-1$  chunks are grouped together to form the training dataset. The model is evaluated by averaging the performance of the model on every fold.



**Figure 3:** Overview of a 3-fold Cross Validation Evaluation System

As it is indicated by its name -*Machine Learning framework*- the training procedure is responsible for the creation of the forecasting model (hyperplane), so we use the largest sample of our dataset in order to construct the best model that is possible using the described Cross Validation scheme. We enhanced this algorithm by adding a further step of out-of-sample evaluation. In this final step the trained model performance is tested against a totally unknown data chunk that was left out of the cross validation training step. As train data we used data from 1976:Q3 to 2006:Q3 while the out-of-sample data where from 2006:Q4 to 2011:Q3.

**Table 1. Descriptive Statistics**

	GDP	3 months interest	6 months interest	2 years interest	3 years interest	5 years interest	7 years interest	10 years interest
Mean	9,083	5,320	5,437	6,224	6,411	6,722	6,961	7,123
Median	9,080	5,035	5,105	5,935	6,190	6,425	6,700	6,855
Minimum	8,546	0,020	0,050	0,280	0,470	1,060	1,620	2,150
Maximum	9,505	15,020	14,740	15,540	15,500	15,410	15,330	15,150
Standard deviation	0,301	3,368	3,342	3,485	3,378	3,179	3,042	2,9110
C.V.	0,033	0,633	0,614	0,559	0,526	0,473	0,4369	0,408
Asymmetry	-0,126	0,592	0,529	0,469	0,487	0,570	0,620	0,686
Cumulant	-1,337	0,302	0,128	-0,079	-0,076	-0,055	-0,083	-0,099

It is important to notice that the period in which we try to forecast the GDP cycle is the period of the global financial crisis. This later crisis was a really unusual event in economic history. We decided to include it in our out-of-sample forecasting as if our model accurately forecasts the performance of the economy during this turbulent period with rare economic events, its value will be higher.

### III. Empirical Results

We conducted 3501 tests using the linear kernel and 293371 tests using the RBF. The best accuracy was achieved by an RBF model with error tolerance parameter  $C = 34.5$  and  $\gamma = 0.6$ . The Cross Validation test accuracy was 74.2%, and the out-of-sample accuracy was 66.7%. The model intrigued 6 false alarms (cases that the model erroneously provides a “Recession” flag). However this is a burden we can tolerate, since in the same time it forecasted correctly all the true “Recession” cases.

**Table 2: The estimated cases and the ground truth**

<u>Forecasted</u>	<u>Actual</u>	
	Growth	Recession
Growth	4	0
Recession	6	8

The same out of sample forecasting score was achieved from a linear model though with lower Cross Validation Accuracy (63,3%).

The importance of these results is obvious. Policy makers, the government and the central bank, can use such a model to forecast the deviations of real output from its long-run trend , the full employment output. By doing so they can efficiently and moreover swiftly without any delays design and implement the necessary monetary and fiscal policy mix that can steer the economy to the full employment GDP. The results show that this scheme was able to identify and forecast one quarter ahead all instances of actual output gaps below the long run trend. Our out-of-sample data included eight quarters of actual output below full employment and our model forecasted correctly all of them. Moreover, we correctly forecasted eight quarters of inflationary gaps. The only downside of this model is that it produced some false alarms in the case of below-trend output: although the model forecasted all eight such instances of an unemployment gap it also provided six false alarms. These are cases where the model forecasted an unemployment gap and the actual output was above trend. Evaluating the overall performance of this model we can state that it has significant power to forecast upcoming recessions so that the government and the central bank can react swiftly and effectively to dampen the effects of the downturn in real output or even avoid it all together. As the model in the effort to forecast all recessionary events produces some false alarms, we can expect that the expansionary

monetary and fiscal policy associated with those false warnings will result in some inflationary pressure to the economy.

#### **IV. Conclusion**

This paper is the first empirical investigation on the relation between the yield curve and an economy's real output, using an SVM classifier. Our goal is to create a forecasting model that can accurately inform us about future output gaps. Such a model can be a very useful tool for policy makers: governments and central banks. It can be employed for the efficient and prompt implementation of the fiscal and monetary policy mix in aiming to minimize deviations of real output from the long-term trend. We apply the SVM methodology, using information from the yield curve; i.e. the interest rates on U.S. federal government treasury bills for 3 and 6 months maturity and long-term government bonds with maturities 2, 3, 5, 7 and 10 years. We used both the linear and the RBF kernels in an out-of-sample one-quarter-ahead forecasting scheme. Our out-of-sample data included eight quarters of below-trend output and ten quarters of inflationary gaps. We achieved to have a 74.2% overall accuracy using the RBF kernel. Nonetheless, focusing on the incidents of an unemployment gap (an actual GDP value below the long-run trend) we can see that our model had a 100% accuracy in forecasting while at the same time it produced six false alarms: cases where the model forecasted an unemployment gap but the actual GDP was above trend. Overall, the model has a significant power to forecast upcoming recessions. This can be effectively exploited by policy makers within the government and the central bank and by implementing the necessary policy mix they can dampen the effects of the recession, minimize its duration, or steer the economy away from it all together. As the model provides some false alarms, we expect that implementing fiscal and monetary policy in this manner may put some inflationary pressure to the economy.

#### **References**

Ang A, Piazzesi M, Wei M, (2006) "What does the yield curve tell us about GDP?", *Journal of Economics* 131, p:1–2

Ben-Hur A, Weston J, "A User's Guide to Support Vector Machines", Technical report, Department of Computer Science Colorado State University, NEC Labs America Princeton

Bernanke BS, Boivin J, Elias P, (2005) "Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach", *Quart J Econ* 120, p:387–422

Chauvet M, Potter S, (2002) "Predicting a recession: evidence from the yield curve in the presence of structural breaks", *Economic Letters* 77(2), p: 245-253

Chung C, Lin C, (2001), "Library for Support Vector Machines", Technical paper, Department of Computer Science, National Taiwan University, Taipei, Taiwan, May

Chionis D, Gogas P, Pragidis I, (2009) "Predicting European Union Recessions in the Euro Era: The Yield Curve as a Forecasting Tool of Economic Activity", *International Advances in Economic Research* 16 (1)

Diebold F, Li C, (2006) "Forecasting The Term Structure of Government Bond Yields", *Journal of Economics*, 130, p: 337-364

Estrella, A., Hardouvelis, G.A., (1991) "The term structure as a predictor of real economic activity", *Journal of Finance* 46, p: 555–576

Estrella A, Mishkin FS, (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 (7), p: 1375–1401

Estrella A, Mishkin FS, (1996) "The Yield curve as a predictor of U.S Recessions", *Federal Reserve Bank of New York: Current Issues in Economic and Finance* 2(7), June

Hodrick R, Prescott EP, (1997) "Postwar business cycles: an empirical investigation", *J Money Credit Bank* 29 (1), p: 1–16

Moench E, (2008) "Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented VAR approach", *Journal of Econometrics* 146(1), p:26-43, September

Peel D, Taylor M, (1998) "The slope of the yield curve and real economic activity: Tracing the transmission mechanism", *Economic Letters* 59(3), p:353-360

Robert G. King and Sergio T. Rebelo, (1998) "Low Frequency Filtering and Real Business Cycles", Rochester Center for Economic Research, Working Paper 205, p: 1-31

Wright J, (2006) "The yield curve and predicting recessions", *Finance and Economics Discussion Series*, Federal Reserve Board 2006–7, February