

Forecasting U.S. recessions with probit models

Thomaz Bastos Fraga

Engenharia de Produção

Pontifícia Universidade Católica do Rio de Janeiro

Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - 22451-900

Thomaz.bf@hotmail.com

Cristiano Augusto Coelho Fernandes

Departamento de Engenharia Elétrica

Pontifícia Universidade Católica do Rio de Janeiro

Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - 22451-900

cris@ele.puc-rio.br

Davi Michel Valladão

Departamento de Engenharia de Produção

Pontifícia Universidade Católica do Rio de Janeiro

Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - 22451-900

davimv@puc-rio.br

ABSTRACT

In this important and highly researched topic, recession forecasting in the United States, we aim to review previous benchmark works on the econometric literature to draw inspiration and try to improve prediction using different methods. The most widely used variable, and also the one by far with the highest prediction power, is the yield curve. We try to add prediction power to this model with two approaches: using other financial variables and trying different kind of models. We also try to improve the forecasts of previous works with approaches using the Lasso variable selection method and with General Autoregressive Score (GAS) models. These models were able to achieve meaningful prediction power in both short and long time forecast horizons.

KEYWORDS. Recession forecasting; Time series; Lasso; GAS models.

Paper topics: EST&MP – Statistics and Probabilistic Models

1. Introduction

The purpose of this work is to review the set of variables commonly used in recession prediction and the models which can be constructed from those variables.

Recession forecasting is an old and important challenge for both economists and market practitioners. These two agents have different end goals associated with this activity. While economists often use it for academic analysis of business and financial cycles which usually has an end goal of assisting in policy making, market practitioners are interested in protecting their portfolio or profit from market crashes and recession periods.

In the activity of predicting recessions, it is critical to watch the actions taken by the Federal Reserve (FED). Since one of the main duties of the Federal Reserve (FED) is to conduct a nation's monetary policy in pursuit of maximum employment and stable prices, the American Central Bank plays a critical role in this dynamic. This dual mandate has a important role in smoothing the business cycles of the economy.

As the FED uses its monetary tools to achieve these goals, it is fundamental that they use recession probabilities and scenarios as input. Since the FED pursues these goals by acting in the economy with its tools and setting forward guidance, observing these significantly helps the forecast.

The most widely used financial variable for forecasting recessions is the yield curve which is commonly estimated by subtracting the 10-year Treasury Bond Rate from the 3-month Treasury Bill Rate. The yield curve inversion is closely watched by market participants as a powerful indicator of recessions in the near future.

Monetary policy has a great deal of influence on the curve and, consequently, on expectations of real activity. The rise in the short rate is warranted when the economy shows signs of overheating and is associated with slowing real growth in the near future. This dynamic explains the inversion of the yield curve before recessions, as the FED rises the rate to cool down the economy in an effort to curb inflation and the long-term rate does not necessarily rise proportionately. Other factors like the decline in expectations about future economic real activity and change in risk assessments can also provoke shifts in the yield curve.

There are other variables that have predictive power. Credit spreads can point to excesses when junk rate bond start having low implicit default rates in late cycle economies, stock market returns have embedded economic expectations in them as when these are dire, they tend to reflect on prices and financial cycle measures are a couple of examples that will be further analyzed.

The main objective of this paper is to illustrate the ability of the yield curve indicator to forecast recessions. Even though there are many variables that have predictive power in recessions, the job of improving the yield curve recession model proves to be a particularly challenging one.

2. Literature Review

Much of our work was inspired from [Nyberg, 2009], where the author uses financial variables such as the yield curve and the stock market returns for predicting recession. They find that the yield curves proves itself to be a powerful predictor for future economic growth and recession periods. It is also found that the use of foreign yield curves can improve the forecasts. The binary probit models used in [Nyberg, 2009] serve as inspiration for the ones pursued in this work.

The seminal Estrella and [Mishkin, 1998] paper have also proved to be of great assistance to our work. The authors were the firsts to propose that stock market returns can add predictive power to the yield curve model. [Nyberg, 2009] cites their work as the first to incorporate stock market returns as predictors in dynamic probit models.

[Rudebusch and Williams, 2007] compare in their work the forecasts of the yield curve model and the forecasts surveyed in the Survey of Professional Forecasters (SPF). The results show a possible undermining of the predictive power of the yield curve as it consistently outperforms the

SPF. The authors highlight as possible causes of this downweighing both the underestimation of macroeconomic repercussions of changes in monetary policy stance and the belief by the SPF that past performance of this indicator in predicting recessions may not hold in the future.

[Liu and Moench, 2016] from the Federal Reserve Bank of New York also wrote a well-regarded paper on this subject. The authors incorporate the receiver operating characteristic (ROC) curve as an evaluation metric for the probit model forecasts.

3. Methodology and goodness-of-fit measures

As a measure of recessions, this paper uses the NBER's (National Bureau of Economic Research) dates of turning points in the United States economy as it is the most widely used metric in academia and between market practitioners for classifying recessions. It defines economic expansionary states with a value of 1 and recessionary states with 0.

“A recession is a significant decline in activity spread across the economy, lasting more than a few months, visible in industrial production, employment, real income, and wholesale-retail sales. A recession begins just after the economy reaches a peak of activity and ends as the economy reaches its trough. Between trough and peak, the economy is in an expansion. Expansion is the normal state of the economy; most recessions are brief and they have been rare in recent decades.” NBER (2003)

The NBER determines these turning points in the economy by relying on a range of macroeconomic indicators. The objective is to have a clear sight that the economy has entered a recession and the analysis of a range of indicators which spread across different sectors of the economy has the purpose of eliminating the bias that any one might have on specific sectors. In addition, it is important to accompany these factors for more than just a month as a single data point can lead to wrongfully labeling a short one-off event as a full recession.

As per the work of the previous cited authors, it is possible to form a set of variables that provide value as recession indicators. We consider the United States Yield Curve (subtraction of the 10y interest rate and the 3-month interest rate), the German Yield Curve (the subtraction of the 10y interest rate and the 3-month interest rate), the S&P 500 index (log-differences from the S&P 500 index), the Baa-Aaa Credit Spread (the subtraction of Baa and Aaa credit ratings for US firms) and the Financial Cycle Index (measure synthetically created by the Bank of International Settlements for measuring financial cycles).

A correlation table was formed to see which lags of the independent variables are most significant in forecasting recessions. The table below shows these correlations between the independent variables and the binary recession indicator.

Lags	US	GE	S&P	BaaAaa	FC
1	-0.228	-0.280	-0.060	0.372	-0.082
2	-0.291	-0.304	-0.080	0.332	-0.080
3	-0.368	-0.326	-0.141	0.278	-0.078
4	-0.444	-0.344	-0.160	0.204	-0.075
5	-0.470	-0.365	-0.227	0.138	-0.074
6	-0.477	-0.392	-0.182	0.101	-0.070
7	-0.504	-0.395	-0.193	0.090	-0.064
8	-0.526	-0.396	-0.182	0.084	-0.060
9	-0.538	-0.394	-0.142	0.078	-0.058
10	-0.548	-0.387	-0.106	0.068	-0.055
11	-0.537	-0.387	-0.069	0.068	-0.050
12	-0.532	-0.382	-0.061	0.070	-0.040
13	-0.529	-0.373	-0.051	0.059	-0.029
14	-0.521	-0.362	-0.012	0.056	-0.020
15	-0.506	-0.354	-0.027	0.075	-0.005
16	-0.478	-0.325	0.046	0.106	0.008

Fig 1: Values of Pearson correlation between the binary recession indicator and the lags of the variables studied where “US” denotes the north-American yield curve, “GE”, the German yield curve, “S&P” the log-differences of the S&P 500 and “FC” the financial cycle index synthetically created by the Bank of International Settlements

With these values we set to find the combination of the most significant lags of the variables and started testing their interaction within the models. We find results similar to [Nyberg, 2009], a combination of the American and German yield curves, together with the S&P lags, results in the best model.

The chosen model for the binary time series forecast is the Probit model where the implementation was done through the “glmnet” package. The conditional probability that the outcome variable y_t takes the value 1 when a recession occurs (and 0 otherwise) is written as:

$$p_t = E_{t-1}(y_t) = P_{t-1}(y_t = 1) = \Phi(\pi_t) \quad (1)$$

Where π_t is a linear function of variables included in the information set Ω_{t-1} and $\Phi(\pi_t)$ is the standard normal cumulative distribution function.

The first two models (equations (2) and (3)) are small variants of the time series models used by [Nyberg, 2009]. For different approaches and in order to improve the forecasts, we try both a Lasso regression on the set of variables on Table (2) and General Autoregressive Score Models using the variables chosen by the Lasso approach as a starting point.

The first static model denoted as (2) uses the American and German yield curves. The German yield curve is added as a proxy for the European economy.

$$\pi_t = \beta_0 + \beta_1 US_{t-6} + \beta_2 US_{t-12} + \beta_3 GE_{t-12} \quad (2)$$

The second static model denoted as (3) uses the same lags of the yield curves and adds the lags of the log difference of the S&P500 (Standard & Poors 500).

$$\pi_t = \beta_0 + \beta_1 US_{t-6} + \beta_2 US_{t-12} + \beta_3 GE_{t-12} + \beta_4 S\&P_{t-2} + \beta_5 S\&P_{t-4} + \beta_6 S\&P_{t-6} \quad (3)$$

The Lasso model was first introduced by [Tibshirani, 1996] as a new method for variable selection in linear models. Since when choosing our regressors, we are faced with a multitude of variables and lags to choose from, this method proves itself to be a perfect fit to the task. Thus, to improve the choice of predictors and to avoid being trapped in a local maximum kind of phenomenon (where the optimum solution can not be reached by meddling with the nearest lags and variables), we use the Lasso approach.

Variable	Source	Lags
US Yield Curve	Federal Reserve	2, 4, 6, 8, 10, 12
German Yield Curve	Federal Reserve	2, 4, 6, 8, 10, 12
S&P 500 Index	Standard & Poor's	2, 4, 6, 8, 10, 12
Dow Jones Index	Federal Reserve	2, 4, 6, 8, 10, 12
Baa Aaa credit spread	Federal Reserve	2, 4
Financial Cycle Index	BIS	2, 4, 6, 8, 10, 12

Fig 2: Set of variables used as candidates for selection in the Lasso method. The variables chosen by the Lasso method from this table are shown in Table 3 under the “Lasso” column.

The motivation behind the Lasso is to reduce insignificant coefficients to the value of zero, which is done by adding a penalty (λ) to the normal OLS (Ordinary Least Squares) minimization problem. Therefore, penalizing all the estimated parameters and excluding those which are irrelevant. As the tuning parameter λ increases, the penalty for irrelevant variables consequently increases. When λ decreases, the tendency to overfitting also increases since the penalty for the irrelevant variables also decreases. For this work, the λ value is chosen through cross-validation.

When analyzing a significantly large subset of variables, the Lasso approach can prove itself to be very useful since it helps both in reducing over-fitting and with regressor selection. To obtain the Lasso estimators, we must solve the equation below for minimization:

$$L(\lambda, \beta) = \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (4)$$

Besides the Lasso model, we also try to approach the data with Generalized Autoregressive Score (GAS) models. The Bernoulli distribution approach with the logit model used by [Lit and Koopman, 2020] served as inspiration as well as the author's software at timeserieslab.com to fit and forecast the data.

As per [Lit and Koopman, 2020], we switch to using logit model and assume that the recession indicator y_t follows a Bernoulli distribution where π_t is a function of the regression effect $X_t\beta$ where the independent variables used mimic the ones in equation (3) and the dynamic process α_t .

$$\pi_t = \beta'Z_t + \alpha_t \quad (5)$$

$$\alpha_t = \omega + \alpha_{t-1} + \kappa s_t \quad (6)$$

Where κ is the updating parameter and where s_t is the driving force behind updating equations (5) and (6) as is given by:

$$\begin{aligned} s_t &= I_{t|t-1}^{-\frac{1}{2}} * \nabla_t, & \nabla_t &= \frac{y_t - \pi_t}{\pi_t(1-\pi_t)}, & I_{t|t-1} &= \frac{1}{\pi_t(1-\pi_t)} \\ s_t &= I_{t|t-1}^{-\frac{1}{2}} * \nabla_t = \frac{y_t - \pi_t}{\sqrt{\pi_t(1-\pi_t)}} \end{aligned} \quad (7)$$

Equations (5), (6) and (7) form a score-driven model in which the mechanism to update the time-varying parameter over time is the scaled score of the likelihood function.

The in-sample interval used is from January 1974 to December 1994 and the out-of-sample interval is from January 1995 to March-2020. For months ahead forecasts where the independent variables are not available, the explanatory variables are tailored so that for the forecast horizon $t + h$, only information known at time t is used.

The Akaike information criterion (AIC) was first proposed by [Akaike, 1974]. In this work, the author reviews the history of the development of statistical hypothesis testing and proposes that hypothesis testing is not adequately defined as the procedure for statistical model identification. He then proposes a measure (AIC) that estimates the relative amount of information lost by a given model.

$$AIC = 2k - 2 \ln(L(\theta)) \quad (8)$$

Where k refers to the number of independently adjusted parameters within the model and $L(\theta)$ is the maximum value of the likelihood function for the model.

We use the AIC as a selection method since it gives a way to measure the goodness-of-fit of the model while penalizing for the over-fitting the data.

The second selection measure utilized is the area under the receiver operating characteristic (ROC) curve. This measure was first tried in recession prediction models by [Liu and Moench, 2016].

To obtain the ROC curve, the first step is obtaining the percentage of true and false positives for a given period while varying the cutoff value (the minimum probability value that classifies a positive signal of a recession). Then, a graph is constructed using the true positive rates (TPR) in the Y-axis, the false positive rates (FPR) in the X-axis for each of the cutoff values, producing the ROC curve

Once the graph is complete, our goodness-of-fit measure is the area under the ROC curve (AUROC). A model that perfectly predicts recessions would only have true positives and no false positives and an AUROC equal to one. The random guess model should, in contrast, have an equal number of true positives and false positives, rendering an AUROC equal to 0.5. An example of a ROC curve graph is shown below.

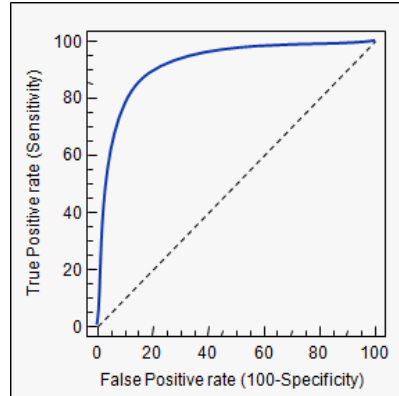


Fig 3: Example of and ROC Curve graph

4. In-sample results

As seen on Table 3 below, the American yield curve is the primary explanatory variable, but the inclusion of the German yield curve does add a statistically significant predictor. Even though the S&P 500 lags do improve the evaluation metrics used, it is difficult to further add significant prediction power beyond the yield curve.

In Table (3), we show the results for the in-sample models. Model 1 refers to equation (2), Model 2 refers to equation (3), Nyberg refers to the Dynamic Model used in [Nyberg, 2009] and GAS refers to our approach to the GAS Model using the variables selected by the Lasso approach.

In-Sample	Model 1	Model 2	Nyberg	Lasso	GAS
Intercept	-0.08 (0.16)	0.02 (0.18)	-0.32 (0.17)		
US_4				-0.09	
US_6	-0.42 (0.10)	-0.39 (0.10)	-0.45 (0.09)	-0.34	-0.72 (0.23)
US_8				-0.18	
US_10				-0.09	-0.34 (0.29)
US_12	-0.37 (0.11)	-0.43 (0.12)		-0.34	-0.37 (0.30)
GE_6			-0.44 (0.09)	-0.35	-0.28 (0.20)
GE_12	-0.42 (0.09)	-0.37 (0.10)		-0.35	-0.58 (0.23)
SPX_2		-14.46 (7.17)	-11.46 (6.51)	-12.68	-11.53 (6.18)
SPX_4		-21.42 (7.03)	-23.42 (6.54)	-10.42	-17.08 (6.62)
SPX_6		-15.35 (7.46)	-19.75 (7.05)	-24.88	-13.93 (6.43)
SPX_8				-15.95	-12.14 (6.76)
SPX_10				-4.48	
BaaAaa_2				13.11	2.81 (1.19)
BaaAaa_4				7.25	-3.00 (1.19)
DJ_4				-17.50	
AIC	112.04	104.73	122.44	105.86	115.70
AUC	0.958	0.960	0.945	0.980	0.966

Fig 4: In-sample results for models 1, 2, the Dynamic Model used by [Nyberg, 2009] and the GAS approach. US denotes the American yield curve, GE the German yield curve, SPX the S&P 500 log-difference, BaaAaa the credit spreads and DJ the Dow Jones Index. The number following the underscore represents the variable's lag. AIC refers to [Akaike, 1974] information criteria and AUC represents the area under the ROC curve. The values under parentheses are the standard deviations.

The in-sample results shown on the graphs below illustrate these relations. The first model with the American and German yield curves as the predictors shows promising results. In the inflationary crisis of the 1970's and 1980's, the federal funds rate was abruptly increased, and

the 10-year rate did not follow. The result was a negative yield curve in these three recessions which value was higher than the fourth recession in the in-sample period. Thus, the probability implied using only the yield curve can significantly undermine the probability of a recession.

The inclusion of the S&P lags in the second model increases the accuracy of the yield curve model. The probability of the fourth recession of the in-sample period has significantly improved, which given the weak yield curve signal was underestimated by the yield curve model. The erratic nature of the stock market is also translated in Figure 2 and can be observed when compared to the first Figure.

The change of the simple probit model from models 1 and 2 to the Lasso and the GAS Models approach have small and almost imperceptible effects to the plotted graphs. Although the Lasso model graph does appear to be less erratic and to have a slighter better fit.

The AUROC scores in table 4 indicates towards the inclusion of the S&P lags to the first model not adding much value to the predictions, while the Lasso approach shows promising results.

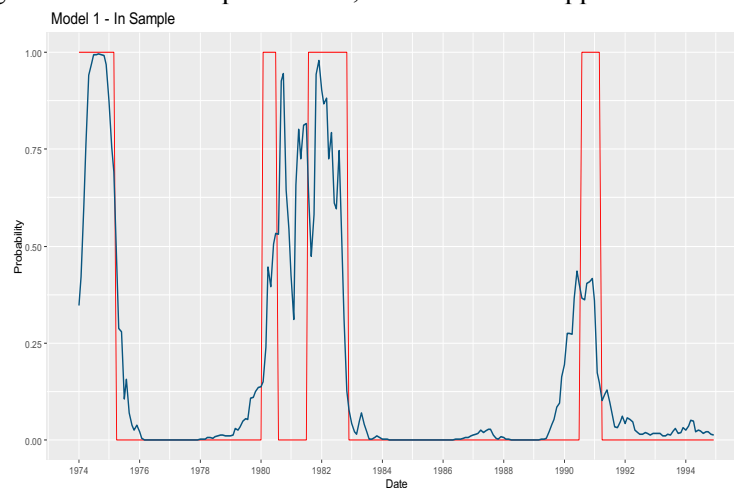


Fig 5: In-sample recession probabilities implied by the model 1 with the yield curves as the only predictors.

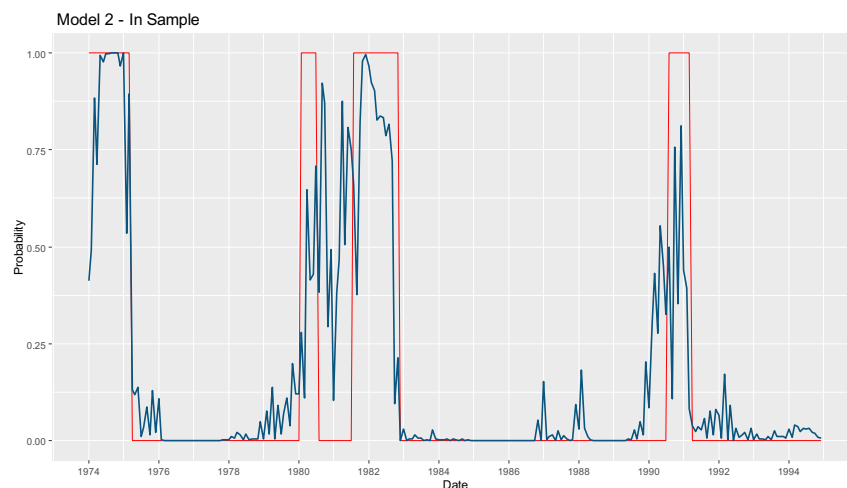


Fig 6: In-sample recession probabilities implied by the model 2, with the addition of the log-difference of S&P lags to the yield curve model (model 1).

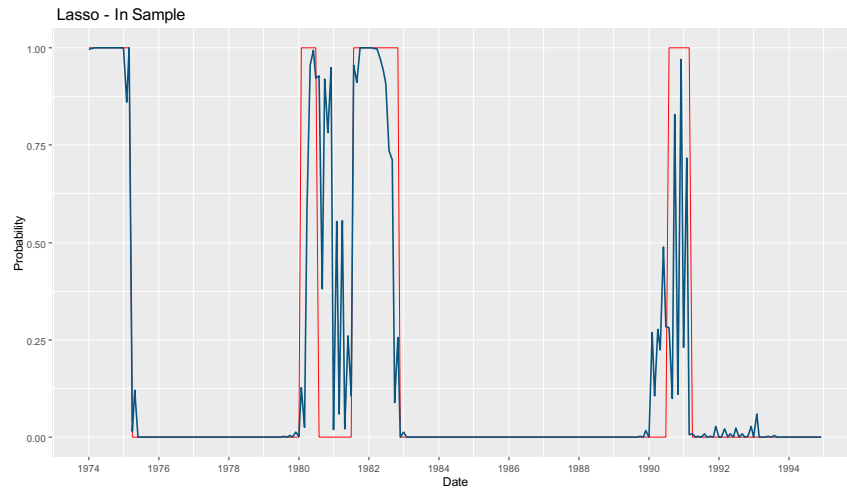


Fig 6: In-sample recession probabilities implied by the Lasso model.

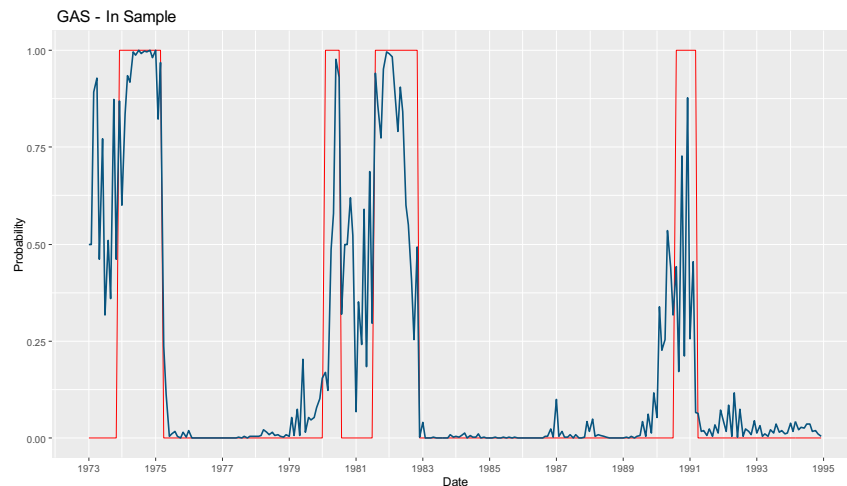


Fig 7: In-sample recession probabilities implied by the GAS model.

AUC of ROC graphs - In Sample	
	AUC
Model 1	0.958
Model 2	0.960
Lasso	0.980
GAS	0.966

Fig 8: Area under the ROC curve for the in sample models.

5. Out-of-sample results

While we can observe a great deal of predictability for recession and expansion periods in the in-sample period, that does not necessary translate to out-of-sample predictability. For the out-of-sample period, the predictions will start in January 1995, but the end date will depend on how many months ahead the forecast will be. In the $t+1$ models, for example, the last prediction will be for October 2020 and in the $t+12$ models the last one will be for September 2021.

For months where the forecast horizon exceeds a variable's lag, we tailor the lag so that only variables with known values are used. For example, if we are forecasting a $t+h$ model which has a set of variables with lags of which we do not know the value at the moment of the forecast,

we reestimate the model without these variables and using only known lagged values at the moment t .

When forecasting the out-of-sample period, we pursue the following methodology. We start estimating the forecasts for January 1996. In the $t+1$ models, we use as sample the period until December 1995 and forecast one step ahead. In the $t+6$ models, we use as sample the period until July 1995 and forecast six steps ahead. Finally, in the $t+12$ models, we use as sample the period until January 1995 and forecast twelve steps ahead. Then, we continue to do the same procedure for the following months until the end date. This same methodology is followed for the forecasts using the Lasso method.

The methodology outlined above is referred as direct forecasting and was scrutinized alongside iterated multistep AR methods for forecasting macroeconomic time series by [Marcellino et al., 2006]. Even though the authors find that the multistep AR methods typically outperform direct forecasting methods in the prediction of macroeconomic time series, the independent variables present in this work are financial variables. Therefore, we opted for the direct forecasting method.

First, using only the out-of-sample plots as an analysis tool, we can observe a great deal of predictability. The yield curve model performs well for the out-of-sample period for all the time horizons, illustrating the efficacy of the term spread as a recession predictor. However, the probabilities implied by this model struggle to surpass the 50% threshold. The reason for this is the fact that the term spread for the recessions since 1995 has not widened as much as the recessions in the in-sample period.

With the addition of the S&P lags, the $t+1$ and $t+6$ horizons show a better fit to the recession periods. Although, given the erratic nature of the stock market, the plots have more false positives during expansion periods. In the $t+12$ horizon, the second model resembles the first model almost exactly, as the S&P lags do not have great predictability for longer term forecasts.

The Lasso models have an impressive fit to the recession periods, greatly diminishing the number of false positives from the previous models.

The AUROC scores shown in table 5 deliver interesting results. The first model using only the yield curves as explanatory variables has larger scores for all time horizons when compared to model 2. These results point to the inclusion of the S&P lags as not being significant towards improving the forecasts. Meanwhile, the results for the Lasso model point to it being more accurate at shorter time horizons although it does not add much prediction value for the $t+12$ forecasts when compared to the first model.

AUC of ROC graphs - Out of Sample			
	t+1	t+6	t+12
Model 1	0.922	0.888	0.933
Model 2	0.904	0.887	0.928
Lasso	0.983	0.957	0.934
GAS	0.967	0.962	0.953

Fig 9: Area under the ROC curve for the out of sample models.

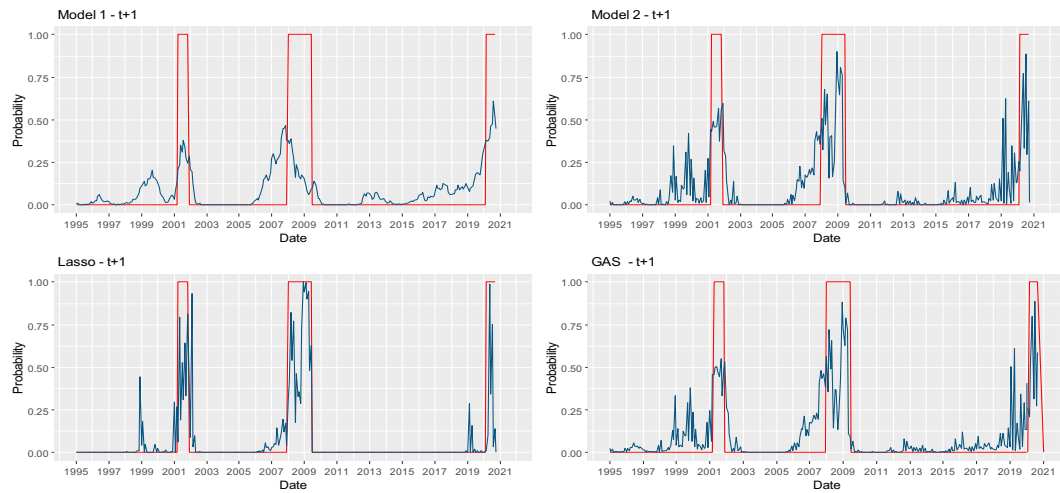


Fig 10: Out-of-sample forecasts for $t+1$ steps ahead.

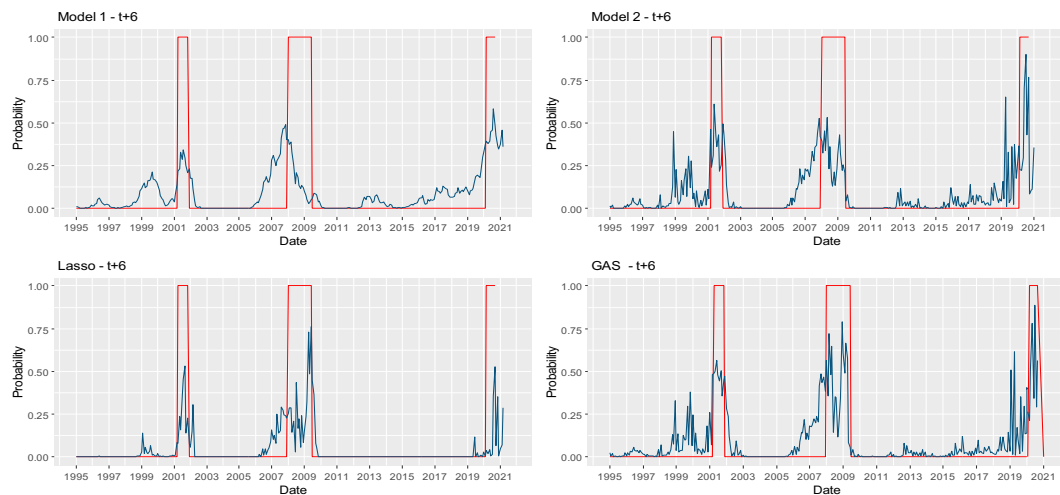


Fig 11: Out-of-sample forecasts for $t+6$ steps ahead.

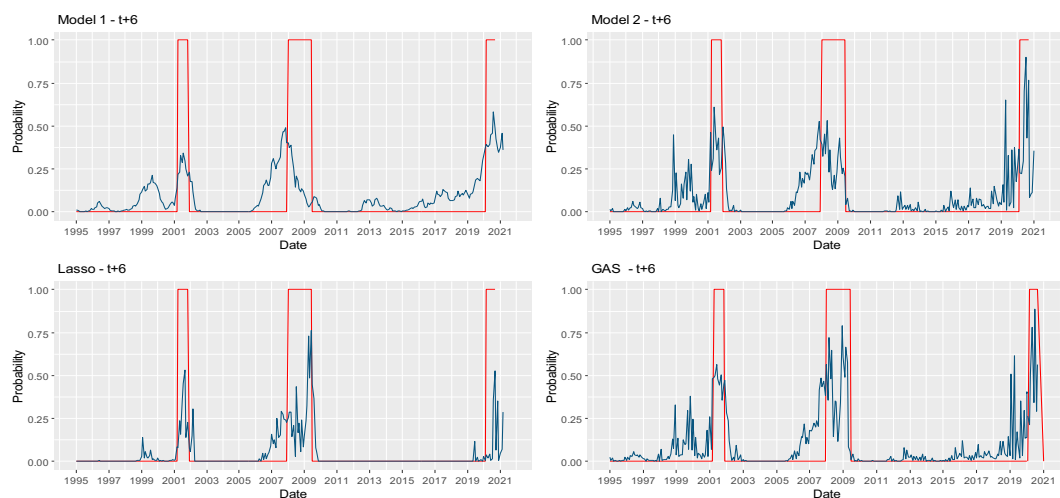


Fig 12: Out-of-sample forecasts for $t+12$ steps ahead.

6. Conclusion

Upon examining the performance of these recession prediction models, we can reach some conclusions. The primary insight from our research is the power of the yield curve as a predictor and how difficult it is to add further explanatory power to this variable. The reason for this is the fact that financial variables, which have market determined values, have embedded financial and economic expectations on its price formation. Thus, the task of improving the forecast of a financial variable by adding other financial variables to the model is a hard one, as there is often little additional information that is not already priced in.

As the difficulty of finding leading indicators of recession whose recession prediction power is not embedded in financial market expectations is highly elevated, a number of different model approaches were tried. Although expected, an important takeaway is the fact that increasing the complexity of the model does not necessarily improve the prediction by a significant amount. The best path to significantly improve the forecast is to find independent variables which add further information to the financial ones used in this work.

References

AKAIKE, H. (1974). A new look at statistical model identification. *IEEE Transactions on Automatic Control*, p. 713-723.

Business Cycle Dating Committee, National Bureau of Economic Research. (2003). The NBER's Recession Dating Procedure. Available at:
<https://www.nber.org/cycles/r/recessions.html#:~:text=A%20recession%20is%20a%20significant,the%20economy%20reaches%20its%20trough.>

ESTRELLA, A.; MISHKIN, F. (1998). Predicting U.S. recessions: financial variables as leading indicators. *Review of Economics and Statistics*, 80: 45-61.

LIU, W.; MOENCH, E. (2016). What predicts US recessions. *International Journal of Forecasting*, vol. 32, issue 4, p. 1138-1150.

LIT, R.; KOOPMAN. (2020). S.J. Forecasting the 2020 edition of the Boat Race. *Time Series Lab - Article Series*, 2.

MARCELLINO, M.; STOCK, J.; WATSON, M. (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics*, 136, p. 499-526.

NYBERG, H. (2009). Dynamic Probit Models and Financial Variables in Recession Forecasting. *Journal of Forecasting*, 29: 215-230.

RUDEBUSCH, G.; WILLIAMS, J. (2009). Forecasting recessions: The puzzle of the enduring power of the yield curve. *Journal of Business & Economic Statistics*, 27(4): 492-503.

TIBSHIRANI, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society*, vol. 58, no. 1, p. 267-288.