Using the Entirety of the Term Structure to Forecast Recessions

Alfredo A. Romero[1,*]

[1]Department of Economics, North Carolina A&T State University, USA

*Corresponding author: Department of Economics, North Carolina A&T State University, USA. E-mail: aaromero@ncat.edu

Abstract

Using a logit model and quarterly data from 1962 to 2021, we test the forecasting power of the yield spread, a popular leading indicator, and show that forecasting models that include the entirety of the term structure of interest rates provide more accurate estimates of future economic downturns. We also show that models that only include the yield spread are implicitly imposing restrictions in the coefficients of the model resulting in lower predictive power and omitted variable bias issues.

Keywords: yield curve, term structure, forecasting, principal components analysis, interest rates, logit models, recession

1. Introduction

Whether the recession of 2020 was caused by changing macroeconomic conditions or as an unintended consequence of deliberate government action, a popular leading indicator of downturns in economic activity signaled a recession in the U.S. economy within its range of influence of 18 months. A yield curve inversion occurred in April 2019, roughly 11 months before the official start of the Covid-19 recession, increasing the record of true positives for this indicator to 90 percent over the last 50 years.

If the tract record of the yield curve can be assessed statistically, it would be of great value to policymakers and market participants alike to hedge their policy or investing positions for at least two reasons: for one, a predictive model using the yield curve does not involve a structural model requiring a sometimes untestable set of assumptions regarding the number of structural equations, macroeconomic variables to include, the identities that link these variables together, the statistical properties of (some of) these variables, and forecasts of these variables that have to be sourced by the user before a model is actually estimated. The second reason, and perhaps the more appealing to practitioners, is the fact that even if these structural models can be statistically and substantively specified, the ever-changing nature of the economy will render them obsolete precisely at the time when their predictive power is needed the most. Forecasting models based on the yield curve, on the other hand, require a more succinct econometric specification, where no a-priori structural form must be derived and calibrated, but also seem to remain impervious to structural changes in the economy. This is of special importance during times of increased uncertainty in both the short-run and long-run path of fundamental economic variables. These uncertain times of structural stability occur during the period immediately preceding the official declaration of a recession and the period immediately preceding the start of the economic recovery, when macroeconomic variables adjust to their, perhaps, new long-run equilibrium.

Several authors have attempted to provide a statistical framework for forecasting models using the yield curve, but to this day, no unified front has been put forth in the applied econometrics literature to assess the reliability of competing models. Disagreements between researchers range from what exactly it is meant by the term yield curve to how to evaluate the forecasts produced by it. In this paper, we draw from the most salient features of competing yield curve models to produce forecasts for real GDP while delineating a framework for the assessment of the reliability of the proposed model. We accomplish this by estimating a statistically adequate logistic regression that uses not only the spread of the long-run versus the short-run interest rates but rather the entirety of the term structure of interest rates, commonly known as the yield curve.

Since most 'baseline' forecasting models of economic activity include only the spread of the yield curve, our first working hypothesis is that [Hypothesis 1] forecasts produced by models that include the entirety of the yield curve will provide more accurate predictions of economic activity as measured by standard forecasting accuracy metrics, such as the MAE and the RMSE. Since we can think of models that include the yield curve as unrestricted models and, analogously, we can think of models that include only the spread as restricted models, our

second hypothesis is that [Hypothesis 2] spread models are imposing implicit and untested restrictions to the yield curve model that could be detrimental to their forecasting ability. In other words, our second hypothesis will be tested by assessing statistically the validity of the restrictions imposed. Furthermore, since most yield curve forecasting models only include a handful of spreads between short-term and long-term yields, thus unintentionally imposing restrictions on the estimated coefficients, this second hypothesis implies that including the entirety of the yield curve without imposing restrictions on the coefficients will increase the predictive power of the model.

This paper does not attempt to corroborate any theoretical models of the relationship between the structure of the yield curve and economic activity. Instead, we propose a framework for assessing the reliability of an empirical regularity: the statistical information contained in the yield curve is a leading indicator of future economic activity. The rest of this paper is organized as follows: Section 2 describes the relevant empirical literature of the relationship between the yield curve and economic activity, particularly recessions. Section 3 presents the data used in the models, the creation of the conditioning set, and the estimation strategy. Section 4 discusses the findings while section 5 concludes the work.

2. Literature Review

The literature on the term structure of interest rates and the economy is vast. This review will focus on the most salient features of the relevant research. It is known that changes to fiscal and monetary policy have ramifications both in the short-run and in the long-run through several macroeconomic channels. In the short-run, they affect consumption decisions via inflation expectations and short-term interest rates. In the long run, they change investors' expectations about investment costs and productivity growth. In competitive markets, these expectations materialize in the different values assigned to interest rates of different maturities. Market efficiency will then drive short-term interest rates to reflect that in the short-run expectations of variables such as inflation whereas long-term interest rates will reflect expectations about the economy's potential GDP growth. Because the relationship between the short-run and long-run expectations about the economy changes throughout the business cycles, and because these expectations are related to the interest rates (Estrella and Mishkin, 1996), monitoring the relationship between the short-run and the long-run interest rates provides important information regarding the current phase of the business cycle and the likelihood of switching to a different phase.

The idea that the term structure of interest rates contains information about economic growth goes back to the beginning of the twentieth century (Fisher, 1907) but was not formalized until the 1970s (see Rubinstein, 1976; Lucas, 1978; Breeden, 1986). Since then, two main areas of researched branched out: forecasting the value of economic aggregates (see Gogas, Papadimitriou, Matthaiou, and Chrysanthidou, 2015, for an extensive literature review on the subject) and forecasting the likelihood of downturns in the economic aggregates, such as the likelihood of a recession. Since recession prediction is the subject matter of this paper, we will limit our review of the literature to that branch.

As early as 1989, researchers have been documenting the predictive power of the slope of the yield curve (Laurent, 1989; Mishkin, 1990; Estrella and Hardouvelis, 1991; Hu 1993). Most of these studies calculate the probability of a recession using probit models using a recession dummy dependent variable and the spread between a long-run and a short-run interest rate as an explanatory variable. Since then, several modifications have been attempted to this basic framework, including adding temporal dynamics to the explanatory variables (Nyberg, 2010), accounting for monetary regime changes (Estrella, 2005), expanding the conditioning set (Bordo and Haubrich, 2008; Ang and Piazzesi, 2003), and expanding the range of countries under investigation (Christiansen, 2012; Moneta, 2005). Most of these papers, however, limit the information contained in the interest rates to simply the spread between long-run and short-run rates.

To our knowledge, there has not been any model that attempts to quantify the likelihood of recession using the entirety of the term structure of the interest rates. Hillebrand, Huang, Lee, and Li (2018) put together a model that uses the entirety of the term structure but to forecast the value of macroeconomic aggregates (personal income growth and inflation), not recessions. Yang (2020) also uses the entire yield curve to forecast GDP growth, but not recessions, and only using two interest rates at a time. Gogas et al. (2015) use a machine learning framework using the entirety of the yield curve, but their definition of a recession is not the official one provided by the NBER but rather their own calculation of whether actual GDP is above of below an estimated long-run trend. Their model did not perform better than a probit of a logit model in forecasting recessions.

3. Data and Methodology

3.1 Data

We constructed a dataset that runs from the first quarter of 1962 to the third quarter of 2021 of both short-term and long-term U.S. federal government interest rates obtained from the Federal Reserve Bank of Saint Louis. The short-term interest rates are from the Treasury-Bills with maturities of 3 and 1 year. The long-term interest rates are from the U.S. government bonds with maturities of 3, 5, and 10 years. We concentrated on these rates and not included other available maturities to allow for a longer estimation period since many of these other rates start at a later time or stopped being traded throughout the 1962Q01-2021Q03 period. We use the quarterly average of the daily closing value of the interest rates to build the conditioning set obtaining a total of 240 observations. Table 1 shows descriptive statistics for the term structure of the interest rate. Notice that, in line with the expectations theory, the annualized yields of the different rates increase with the length of the maturity period. This is evident using both measures of central tendency. Notice also that the volatility of the short-term rates as measured by the standard deviation is higher for the short-term rates than for the long-term rates.

Table 1. Descriptive Statistics

	3-month	**1-year**	**3-year**	**5-year**	**10-year**
Mean	4.59	4.94	5.35	5.61	5.96
Median	4.72	5.01	5.39	5.56	5.76
Maximum	15.87	16.31	15.78	15.42	14.84
Minimum	0.01	0.06	0.16	0.27	0.65
Std. Dev.	3.36	3.44	3.33	3.19	2.98
Skewness	0.70	0.60	0.52	0.55	0.64
Kurtosis	3.68	3.29	3.08	3.10	3.17
Jarque-Bera	24.3223	15.0047	10.8292	12.3613	16.5623
Probability	0.0000	0.0006	0.0045	0.0021	0.0003
Observations	240	240	240	240	240

We derived two additional variables from these interest rates: the 'spread', or difference between the 10-year and the 3-month interest rate (a measure sometimes colloquially referred to as the 'yield curve), to capture whether the term structure is steep, flat, or inverted; and an indicator variable or signal that takes the value of one if the spread takes a zero or negative value at any point during the quarter and zero otherwise. This inversion signal is traditional used as a back-of-the-envelope predictor of a recession in the near future.

Our dependent variable is an indicator variable that takes the value 1 during the quarter(s) the NBER determined to include the beginning and the end of a recession and zero otherwise. Overall, there were a total of 8 recessions. During the period, there were a total of 41 positive signals of curve inversion, of which only 4 were not followed by a recession within the next 18 months. For instance, the curve inverted during the second and third quarters of 2019, months before the start of the Covid-19 pandemic and almost a full year before the health-induced economic shutdown of 2020.

It is of particular importance to assess the degree of comovement between our recession indicator variable and our conditioning set. Table 2 shows the contemporaneous correlation matrix between the term structure of the interest rate and our recession indicator. All the correlations are statistically different from zero at standard levels of significance. Notice that whereas the correlation between the recession indicator and the 3-month yield is positive and statistically different from zero, the same degree and significance is observed in the rest of the maturities. Non-surprisingly, there is an incredibly high degree of correlation between both the short and long-run maturities but there does not seem to be anything special about the relationship between the 3-month and the 10-year rate. In fact, the correlation between the different yields seem to monotonically decrease the larger their temporal distance.

Table 2. Correlation Matrix of the Yield Curve

	Recession	3-month	1-year	3-year	5-year	10-year
Recession	1.0000					

3-month	0.2313*	1.0000				
	0.000	-----				
1-year	0.2285*	0.9944*	1.0000			
	0.000	0.000	-----			
3-year	0.2201*	0.9738*	0.989*	1.0000		
	0.001	0.000	0.000	-----		
5-year	0.2223*	0.9571*	0.9760*	0.9965*	1.0000	
	0.001	0.000	0.000	0.000	-----	
10-year	0.2220*	0.9288*	0.9511*	0.9819*	0.9936*	1.0000
	0.001	0.000	0.000	0.000	0.000	-----

Nevertheless, as shown in Figure 1, that while the comovements between the different yields are highly correlated, there is a considerable degree of temporal dynamics of divergence and convergence in the rates. The red line in the graph represents the difference between the 10-year and the 3-month yield. Notice that its value consistently approaches zero when the interest rates are both increasing in value and converging. It makes sense that a model that only includes two of these five series would be missing additional statistical information contained in the data.
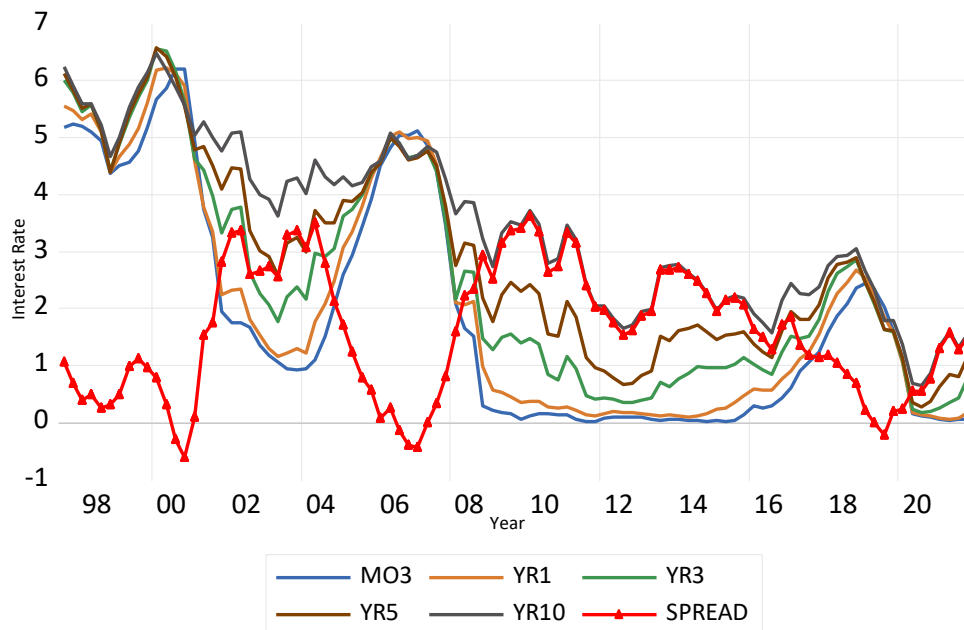

Figure 1. Dynamic Evolution of the Yield Curve

3.2 Estimation Strategy

We estimate the probability of a recession for five different forecasting periods: same quarter and 1, 2, 3, and 4 quarters ahead using logistic regression. We include a same-quarter forecasting period based on the fact that the determination of the beginning of a recession done by the NBER is released several months after the fact. Thus, it could be useful information to calculate the probability that the economy is already experimenting a recession before it is officially confirmed. We used the following logistic framework:

$$P(Y_t = 1 | I_{t-h}) = e^{\boldsymbol{\beta}^T x} / \left(1 + e^{\boldsymbol{\beta}^T x}\right)$$

where $P(Y_t = 1 | I_{t-h})$ is the probability that a recession occurs in quarter t given information at time t-h, $h = 0, 1, 2, 3, 4$. The vector $\boldsymbol{x}$ contains the explanatory variables and $\boldsymbol{\beta}$ is the vector of parameters to be estimated. Three different models were estimated for each forecasting period: a base model where the conditioning set was simply an indicator variable of an inversion of the curve, $\boldsymbol{M_0}$, a 'spread' model that included only the yield spread as defined in Estrella and Hardouvelis (1991), $\boldsymbol{M_1}$, and a model that included the entirety of the term structure of interest rates, $\boldsymbol{M_2}$.

3.3 Forecasts Evaluations

In line with the empirical literature (see Harvey 1991; Rudebusch and Williams, 2007), we evaluate the accuracy of the forecasts using the mean absolute error (MAE) and the root mean squared error (RMSE) over the entirety of the periods. These measures evaluate the closeness, on average, of the probabilities predicted by the models to the observed recession realizations as measured by our dependent variable. The statistics can evaluate each type of model and forecast with:

$$MAE(\boldsymbol{M}_i, h) = \frac{1}{T} \sum_{t=1}^{T} \left| P_t^{M_i} - Y_t \right|$$

$$RMSE(\boldsymbol{M}_i, h) = \sqrt{\frac{1}{T} \sum_{t=1}^{T} \left(P_t^{M_i} - Y_t \right)^2}$$

Since different models will produce different values for the MAE and the RMSE, we assess the statistical significance of their difference using the Diebold-Mariano test. This test is based on the mean accuracy differential between the competing forecasts from models i and j.

$$Diff_t(MAE) = \left| P_t^{\boldsymbol{M}_j} - Y_t \right| - \left| P_t^{\boldsymbol{M}_i} - Y_t \right|$$

$$Diff_t(MSE) = \left(P_t^{\boldsymbol{M}_j} - Y_t \right)^2 - \left(P_t^{\boldsymbol{M}_i} - Y_t \right)^2$$

The test is simply a t-test for the hypothesis of a zero population mean differential. This is

computed by regressing the aforementioned differentials on a constant and testing for the statistical significance of the constant at standard levels of significance (Diebold and Mariano, 1995). Note that there is no test for the RMSE, but its conclusion can be inferred from the MSE version of the test (Rudebusch and Williams, 2007).

4. Results

4.1 Probability Forecasts

The results from the proposed models provide important corroborating information regarding the predictive power of the term structure of interest rates. Table 3 shows the calculations of the MAE, the RMSE and a measure of predictive accuracy using McFadden's pseudo- R^2 for the three models and the four forecasting horizons. Notice that for same-quarter forecasts, the term structure model, M_2, provides statistically significant better predictions by lower values of the MAE and RMSE (at standard levels of significance) than the base model. This is signified with the asterisk next to their value. It also yields a higher value for the R^2. The results remain the same for the one-quarter-ahead recession probabilities, with statistically significant lower values for M_2 and substantively higher R^2.

Table 3. Accuracy of Probability Forecasts

Model / Statistic	MAE	RMSE	pseudo-R^2
Same-quarter Recession Prediction			
M_0: Inversion	0.2325	0.3410	0.0216
M_1: Yield Spread	0.2351	0.3428	0.0103
M_2: Term Structure	0.2129*	0.3251*	0.0974
One-quarter-ahead Recession Prediction			
M_0: Inversion	0.2165	0.3290	0.0907
M_1: Yield Spread	0.2160	0.3273	0.0944
M_2: Term Structure	0.1873*	0.3068*	0.1951
Two-quarter-ahead Recession Prediction			
M_0: Inversion	0.2013	0.1007	0.1529
M_1: Yield Spread	0.1843	0.0900	0.2233
M_2: Term Structure	0.1653*	0.0843	0.2934
Three-quarter-ahead Recession Prediction			
M_0: Inversion	0.1653	0.2875	0.2932
M_1: Yield Spread	0.1630	0.2817	0.3165
M_2: Term Structure	0.1500*	0.2723	0.3579
Four-quarter-ahead Recession Prediction			
M_0: Inversion	0.1744	0.2953	0.2605
M_1: Yield Spread	0.1711	0.2911	0.2959
M_2: Term Structure	0.1604*	0.2819	0.3272

For the two, three, and four-quarter-ahead probabilities of recession, the MAE and the RMSE for M_2 continue to be lower than for M_0 and M_1 (and its R^2 higher than the other two) but only statistically so for the MAE measure. Surprisingly, the predictive power of all models seems to peak in the three-quarter-ahead horizon. These results corroborate our expectation for Hypothesis 1.

As argued above, it could be the case that M_1 is an implicitly restricted version of Model 2, where only the coefficient for the 3-month yield is equal to the negative value of the 10-year, letting all other coefficients be equal to zero. In other words, applying these restrictions to Model 2 yields Model 1. If these restrictions are valid, then the additional interest rates contained in the term structure provide no additional information about the probability of a recession and Model 1 should be preferred. If, on the other hand, the restrictions are not statistically valid, then the rest of the interest rates contain information regarding the probability of recession and not including them in the model will lead to omitted variable bias in Model 1. Table 4 shows the results of the parameter restrictions tests to M_2 for the five forecasting horizons. For the same quarter prediction and up to the two-quarter-ahead prediction horizons, the restrictions are not valid using the F-test and the χ^2 tests resulting from a Wald coefficients restrictions test. This implies that the forecasting ability of the model is maximized when the entirety of the term structure of the interest rate is used. These results corroborate our expectation regarding Hypothesis 2. Surprisingly, the implicit restrictions that only the difference between the 10-year and the 3-month yield are necessary to forecasting accuracy is valid for the three and the four-quarter-ahead models, even though the unrestricted model provides a higher R^2 and lower MAE and RMSE. This could be indicative that, for the medium-term forecasting horizons, M_2 cannot do any worse than M_1 and can potentially do better. In other words, a researcher trying to compute the probability of a recession in the near future and medium-term horizons should primarily use an unrestricted model of the term structure and then hone down on the number of needed coefficients as long as such model provide higher accuracy in the forecasts.

Table 4. Implicit Restrictions Tests

Statistic / h =	0	1	2	3	4
F-test	0.0049	0.0031	0.0235	0.1356	0.2345
χ^2-Test	0.0041	0.0025	0.0213	0.1316	0.2308
Decision	M_2	M_2	M_2	M_1	M_1

4.2 Robustness Checks

We performed a similar analysis using a more recent sample starting instead from 1987Q1 to be consistent with the idea that the relationship between these macroeconomic aggregates structurally changed after 1988. The results are not shown here but are available upon request. This period is known as the Great Moderation, where the long-run paths and volatility of both real GDP and interest rates decreased (Estrella, Rodrigues, and Schich, 2003). We did not find

a systematic change to our previous conclusions, corroborating the claim that no structural change to the parameters of the models can be detected by changing the period of analysis, as shown in Rudebusch and Williams (2007) and Croushore and Marsten (2016).

As a second robustness test, we performed a principal components regression using the entire term structure, M_3. Some authors have shown that the first two principal components account for 99.7 percent of the variation observed in the quarterly yield curve variation (Ang, Piazzesi, and Weid, 2006). Using exploratory factor analysis and the k-means method, we decided to retain two factors and calculated the scores from the data. The updated conditioning set that contains the two latent variables is given by $x_t = \Lambda f_t$, where Λ are the factor loadings and f are the factors. Model 3 performed better than the spread model, as expected, but not better than the model using the unrestricted term structure, M_2. Overall, we cannot change the conclusion that the unrestricted term structure model provides more accurate estimates for the short-run, and just as good estimates than the yield spread model in the long-run.

5. Conclusion

For decades, researchers have used the spread of the long-run and the short-run interest rates to predict the likelihood of economic downturns. These models, although succinct, ignore additional information contained in the rest of the variables in the yield curve. We demonstrate that including the rest of the rates increases the forecasting ability of the models since the entirety of the curve accounts for the convergence and divergence dynamics present in the business cycle. We also provide evidence that yield spread model implicitly impose structural restrictions in the model coefficients that reduce their forecasting power by inducing omitted variable bias.

References

Ang, A., & Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, *50*(4), 745-787. https://doi.org/10.1016/S0304-3932(03)00032-1

Ang A., Piazzesi, M., & Weid, M. (2006). What does the yield curve tell us about GDP growth? *Journal of Econometrics*, *131*(1-2), 359-403. https://doi.org/10.1016/j.jeconom.2005.01.032

Bordo, M. D., & Haubrich, J. G. (2008). Forecasting with the yield curve; level, slope, and output 1875-1997. *Economic Letters*, *99*(1), 48-50. https://doi.org/10.1016/j.econlet.2007.05.026

Breeden, D. T. (1986). Consumption, Production and Interest Rates: A Synthesis. *Journal of Financial Economics*, *16*, 3-39. https://doi.org/10.1016/0304-405X(86)90041-3

Christiansen, C. (2012). Predicting severe simultaneous recessions using yield spreads as leading indicators. *Journal of International Money and Finance*, *32*, 1032-1043.

https://doi.org/10.1016/j.jimonfin.2012.08.005

Croushore, D., & Marsten, K. (2016). Reassessing the Relative Power of the Yield Spread in Forecasting Recessions, *Journal of Applied Econometrics*, *31*(6), 1183-1191. https://doi.org/10.1002/jae.2485

Diebold, F. X., & Mariano, R. S. (1995). Comparing Forecast Accuracy. *Journal of Business and Economics Statistics*, *13*, 253-265. https://doi.org/10.2307/1392185

Estrella, A., & Hardouvelis, G. A. (1991). The Term Structure as a Predictor of Real Economic Activity. *The Journal of Finance*, *46*, 555-57. https://doi.org/10.1111/j.1540-6261.1991.tb02674.x

Estrella, A., & Mishkin, F. S. (1996). The Yield Curve as a Predictor of U.S. Recessions, *Federal Reserve Bank of New York Current Issues in Economics and Finance*, *2*(7). https://doi.org/10.2139/ssrn.1001228

Estrella, A., Rodrigues, A. P., & 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. https://doi.org/10.1162/003465303322369777

Estrella, A. (2005). Why does the yield curve predict output and inflation? *Economic Journal*, *115*, 722-744. https://doi.org/10.1111/j.1468-0297.2005.01017.x

Fisher, I. (1907). *The Rate of Interest*. New York: MacMillan.

Gogas, P., Papadimitriou, T., Matthaiou, M., & Chrysanthidou, E. (2015). Yield Curve and Recession Forecasting in a Machine Learning Framework. *Computational Economics*, *45*(4), 635-645. https://doi.org/10.1007/s10614-014-9432-0

Harvey, C. R. (1991). The Term Structure and World Economic Growth. *Journal of Fixed Income*, *1*(1), 4-17. https://doi.org/10.3905/jfi.1991.692342

Hillebrand, E., Huang, H., Lee, T., & Li, C. (2018). Using the Entire Yield Curve in Forecasting Output and Inflation. *Econometrics*, *6*(3), 40. https://doi.org/10.3390/econometrics6030040

Hu, Z. (1993). The Yield Curve and Real Activity. *IMF Staff Papers*, 622-37. https://doi.org/10.2139/ssrn.883446

Laurent, R. (1989). *Testing the Spread, Economic Perspectives*. Federal Reserve Bank of Chicago.

Lucas, R. E. Jr. (1978). Asset Prices in and Exchange Economy, *Econometrica*, *46*, 1429-1445. https://doi.org/10.2307/1913837

Mishkin, F. S. (1990). What Does the Term Structure Tell Us About Future Inflation? *Journal of Monetary Economics*, *25*(1), 77-95. https://doi.org/10.1016/0304-3932(90)90046-7

Moneta, F. (2005). Does the yield spread predict recessions in the Euro area? *International Finance*, *8*(2), 263-301. https://doi.org/10.1111/j.1468-2362.2005.00159.x

Nyberg, H. (2010). Dynamic probit models and financial variables in recessions forecasting. *Journal of Forecasting*, *29*, 2115-230. https://doi.org/10.1002/for.1161

Rubinstein, M. (1976). The Valuation of Uncertain Income Streams and the Pricing of Options. *Bell Journal of Economics*, *1*, 407-425. https://doi.org/10.2307/3003264

Rudebusch, G. D., & Williams, J. C. (2007). Forecasting recessions: the puzzle of the enduring power of the yield curve. *Working Paper Series 2007-16*, Federal Reserve Bank of San Francisco. https://doi.org/10.2139/ssrn.1007803

Yang, P. R. (2020). Using the yield curve to forecast economic growth. *Journal of Forecasting*, *39*(7), 1057-1080. https://doi.org/10.1002/for.2676