

Predicting Economic Recessions

Oliver Titus

December 19, 2019

Abstract

In this paper, we seek to estimate the probability of an economic recession within certain timeframes in the United States. We use a Probit model to predict the probability of recession within as short as three months to as long as two years. We run a total of sixteen models in this study and use a total of five independent variables. We find that the most important predictor is the spread between the ten year and three month Treasury Yield. The federal funds rate was also an important predictor, but only at the three and six month forecast horizons. We also find that the Consumer Price Index as well as the Industrial Production Index are important predictors only at the two year horizon. We also tried to incorporate a variable that denotes whether or not its an election year, but we found that this was not important. We test the performance of our models using the area under the Receiver Operating Characteristic Curves, which is known as the AUC. We also use a stepwise Probit model to determine which model best fits our data at each forecast horizon. Our overall findings suggest that a future recession is more likely to occur within the shorter timeframes as opposed to the longer timeframes from September 2019.

1 Introduction

The objective of this paper is to predict a United States economic recession three months, six months, one and two years ahead using a Probit model. According to the National Bureau of Economic Research (NBER), a recession is defined as a significant decline in economic activity spread across the economy. A recession usually last more than a few months and are visible in the Gross Domestic Product, real income, employment level, industrial production, and wholesale-retail sales. The NBER does not define a recession as a decline of real GDP for two consecutive quarters [1]. Rather, they identify months and quarters of turning points on the business cycle without assigning it a specific date within the period that turning points occurred [11]. Forecasting economic recessions has been a challenge for many U.S. economists, who use many models to predict the path of economic activity and the likelihood of a recession [4]. These models can often be complex and sometimes may not be accurate.

Economic research has found that the leading indicator that predicts recessions is the Treasury Yield Curve, which is the spread between long-and short-term interest rates [5, 15]. The yield curve has a generally positive slope which implies that short-term interest rates are lower than long-term interest rates. When the yield curve inverts, the slope becomes negative and this implies that short-term interests rates are higher than long-term interest rates. Much of the literature has argued that when this inversion happens, this is a strong signal that a recession is imminent [5].

To predict a future recession, we employ the Probit model as in Estrella [5]. We chose the Probit model because it designed for binary dependent variables and it also allows us to predict the probability of a recession at specific forecast horizons. Data was collected from the Federal Reserve Bank of St. Louis that defines NBER-based U.S. recession periods [11]. Our interest rate data is from the Board of Governors of the Federal Reserve System [8]. This data set was used to calculate the term spread and we use it in our model to make a prediction. The Federal Funds Rate from this data set was also used in some of our models. We also incorporated some additional variables such as the Consumer Price Index as well as an Industrial Production Index in our models [9, 10]. We also tested out a dummy variable that denotes if the current

month falls on an election year.

In section 2, we will discuss how other researchers have predicted economic recessions. In section 3, we discuss the data that was collected in more detail. Section 4 will describe the Probit model in detail. We will discuss the results from estimating the model in section 5. In section 6, we will discuss our conclusions as well as acknowledge the limitations of this study.

2 Background

Economic researchers have used many models to forecast recessions. Filardo [6] tests the accuracy of five different recession prediction models. They use simple rules of thumb using the Conference Board's composite index of leading indicators (CLI), Neftçi's probability model of imminent recession using the CLI, the Probit model, a GDP forecasting model, and the Stock and Watson model. They find that all of these models can provide reliable information about future recessions, while some of the models did not detect some past recessions and some had more false positives than others and some were more robust to real time data than others. When it comes to their findings regarding the Probit model, they find that the accuracy of the model varies based on the forecast horizon. They included other variables in the model such as CLI, Corporate Spread (CS), the Term Spread, and others. They find that as the forecast horizon gets larger, many of the independent variables become insignificant except for the Term Spread variable which was significant in all of their models.

Recent economic research has employed even more sophisticated models used to predict recessions. Ng [7] and Döpke [2] use boosting methods in their models. Boosting is a way of combining models that do not perform well individually to improve the accuracy. Ng uses a collection of logit models on a set of 1500 predictors in their study to find that the most important predictors change depending on the forecast horizon. Döpke uses a machine-learning approach known as Boosted Regression Trees (BRT). They find that the most relevant predictors are the short-term interest rate and the term spread. For our study, we are using the Probit model for its simplicity.

There is much evidence in the literature that the Term Spread is an important

predictor for economic recessions. Estrella [5] states two reasons as to why the inversion of the yield curve is highly related to recessions. First is monetary policy. They state that a tightening of monetary policy implies a rise in short-term interest rates. This slows down the economy and flattens (and sometimes inverts) the yield curve. The second reason is changes in investor expectations about the future. The rise in short-term interest rates influenced by monetary policy causes investors to fear of slowdowns in real economic activity and demand for credit. This causes future interest rates to go down which also flattens (and possibly inverts) the yield curve. Previous studies by Estrella [3, 4] yield similar conclusions regarding why the yield curve is an important predictor for recessions.

3 Data

To estimate the probability of recession in the United States, we use several data sources. We obtained our recession indicator data from the Federal Reserve Bank of St. Louis (FRED). The data values take on a 0 or a 1 value, with a 1 indicating an NBER-defined recession period for that month. This occurs when the U.S. business cycle goes from peak to trough. This dummy variable uses an arbitrary convention that the turning point of the business cycle occurred within the period the turning point occurred since the NBER only identifies months or quarters of turning points [11].

We obtained our interest rates data from the Board of Governors of the Federal Reserve System [8]. The data set includes the Federal Funds Effective Rate as well as yields for commercial papers and treasury bills at different levels of maturity. The yield is also given for U.S. treasury securities with short or long-term maturities. We used the 3-month Treasury bill secondary market rate along with the 10-year Treasury securities market rate to calculate the yield spread. We use this time series data starting from April 1953 up to September 2019 since the 10-year rate started in April 1953. A plot of the yield spread between 10-year Treasury Bill and 3-month Treasury securities is given in Figure 2. It is important to note that where the spread is negative (when the plot crosses below the $y = 0$ line) this is a signal of a recession.

An industrial production index was also obtained from FRED [10]. The index measures real output for all facilities in the United States including manufacturing, mining, electric, and gas utilities. The industrial production index is available starting in January 1919 and is calculated monthly. We also obtained consumer price index (CPI) data from FRED [9]. This is a measure of the average monthly charge in the price for goods and services paid by urban consumers. The CPI data is available starting in January 1947 and is calculated monthly. Table 1 provides a summary of all the data used.

4 Methods

Before calculating the probability of recession, we converted our three-month secondary market rate from a discount basis d to a bond-equivalent basis in order to match the results from Estrella 2006 [5]. We apply their conversion:

$$\text{Bond-Equivalent Yield} = 100 \left(\frac{365d}{360 - 91d} \right)$$

To calculate the probability of recession, we consider the spread between the 10-year Treasury securities rate and the the 3-month Treasury bill rate. We calculate this variable as:

$$SPREAD = R_{10Y} - R_{3M},$$

where R is the interest rate and the subscripts denote the maturity. The model used to make our prediction is a Probit equation given by:

$$P(USREC_{t+h}) = F(\alpha + \beta_1 SPREAD_t), \quad (\text{A})$$

where h is the forecast horizon t months ahead. $USREC$ is the NBER defined recession period and F is the standard normal cumulative density function which can represented as an improper integral:

$$F(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

The value of F gives us the probability of recession. A plot of F is given in Figure 1. We test three other models. Our second model is:

$$P(USREC_{t+h}) = F(\alpha + \beta_1 SPREAD_t + \beta_2 FFR_t) \quad (\text{B})$$

where FFR is the Federal Funds rate. The following model is also tested:

$$P(USREC_{t+h}) = F(\alpha + \beta_1 SPREAD_t + \beta_2 FFR_t + \beta_3 CPI_t + \beta_4 INDPRO_t) \quad (\text{C})$$

where CPI is the Consumer Price Index and $INDPRO$ is the Industrial Production Index. We also test this model with an election dummy E_t that we created. This variable equals 1 if the month t is during the year of an election, 0 otherwise. This model is given by:

$$P(USREC_{t+h}) = F(\alpha + \beta_1 SPREAD_t + \beta_2 FFR_t + \beta_3 CPI_t + \beta_4 INDPRO_t + \beta_5 E_t). \quad (\text{D})$$

To evaluate how these models fit our data, we plot Receiver Operating Characteristic (ROC) curves. These curves shows the true positive percentage (TPP) on the y -axis which is known as the sensitivity, as well as the false positive percentage (FPP) on the x -axis known as the specificity. The sensitivity is calculated as follows:

$$\text{TPP} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100.$$

The specificity is calculated as follows:

$$\text{FPP} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100.$$

Note that specificity can also be calculated as $1 - \text{TPP}$ [13]. The area under the curve known as the AUC measures the performance of the model. A higher AUC implies a better model. We use the `pROC` package in R [12] to plot the ROC curve and compute the AUC. The `pROC` package calculates the area under the curve using trapezoids.

To determine the model that best fits our data at each forecast horizon, we use the

`stepAIC` function in the MASS package of R. We use these resulting models to plot our estimated probabilities of recession at three-month, six-month, one-year, and two-year forecast horizons. The best model is determined by minimizing the Akaike information criterion (AIC) [14]:

$$\text{AIC}_i = 2K_i - 2 \log(L_i)$$

where L_i is the maximum likelihood for the best model i and K_i is a penalty based on the number of variables included in model i .

5 Results

We test the hypothesis that the Term Spread, Federal Funds Rate, the Consumer Price Index, Industrial Production Index, as well as whether or not it is an election year matter in terms of predicting an economic recession. The coefficients from the sixteen Probit models ran are shown in Tables 2 through 5. Each table shows Models A through D at certain forecast horizons.

Looking at results from Model A, the coefficient for the term spread is negative and highly significant across all forecast horizons. This implies that as the spread increases, the probability of recession decreases. It is also negative and highly significant for Model B across all forecast horizons. The coefficient for the Federal Funds Rate is positive and significant for the three-and six-month horizons implying that an increase in federal funds yields an increase in the probability of recession, but this coefficient is insignificant for both the one and two-year horizons. The coefficients for the term spread and the federal funds rate yielded the same conclusion in model C. The coefficient for the consumer price index is negative and significant at the two-year forecast horizon in model C implying that an increase in the CPI yields a decrease in the predicted probability of recession, but is insignificant for the other forecast horizons. The coefficient for the industrial production index is positive and highly significant for the two-year forecast horizon meaning that an increase in this index yields an increase an increase in the probability of recession, but this is insignificant for the other forecast

horizons. Model D showed the same pattern for all the coefficients, except an election dummy was added to this model which came out as insignificant at all forecast horizons. The intercepts for all the models were negative and highly significant.

Plots of the estimated probability of recession for each model are shown in Figures 3 through 6. These plots result from the stepwise Probit models shown in Table 7. The blue plot is the estimated probability of recession at the specified forecast horizon. The grey columns in the plots are the NBER-defined recession periods. Looking at these plots, we see that the spikes generally line up with the NBER recession indicators. We also see that as the forecast horizons increases, the peaks on the spikes get higher, but drop significantly at the two-year horizon. In other words, the probabilities of recession are generally highest at the one year horizon, but lowest at the two-year horizon.

Results of the fit of the these models are given by the ROC curves in Figures 7 through 10. We see that the AUC value increases as the forecast horizon increases, but drops at the two-year horizon. The highest AUC results from model D at the one-year forecast horizon, implying that this is the best model. This is interesting since this model included the insignificant election year dummy variable. The model with the lowest AUC resulted from model A at the three month forecast horizon, implying that this is the worst model. We see that the ROC curve looks best and has a higher AUC at the one year forecast horizons for all models. This implies that a one year ahead prediction of recession seems to the best forecast horizon based on the data used.

Table 6 shows us the probability of recession t months ahead of September 2019. At the time of writing this paper, data for the variables used are available as far as September 2019. Looking at the table, it appears that a recession is more likely to occur within three or six months of September 2019, but not so likely within the next year or two from September of 2019.

Table 7 shows the stepwise Probit coefficients. We see that the best model for the three month horizon is Model C. The best model for the six month forecast horizon is Model B. The one and two year forecast horizons yielded two new models not previously tested. All the coefficients in these models have the same implications as mentioned earlier from Models A through D. Another interesting observation to note is that the

forecast horizon with the highest AIC value is the three month horizon, implying this is the worst forecast horizon to use. The horizon with the lowest AIC value is the one year horizon, implying this is the best horizon to use. This is consistent with the AUC curves in Figure 9 and Figure 11 in that the three month forecast horizon yielded a lower AUC value of 72.15% and the one-year horizon yielded a higher AUC value of 87.12%.

6 Conclusion

In summary, we found that the difference in the ten-year and three-month treasury yield was an important predictor for the probability of recession at any forecast horizon. The federal funds rate is important only at the three and six-month horizons. The Consumer Price Index and Industrial Price Index are significant only at the two-year forecast horizon. The election year dummy was insignificant at all forecast horizons. Based on the models ran, a recession appears to be more likely to occur within the next six months of September 2019. A recession is less likely to occur within a year or two of September 2019 based on the data used in the models.

There are some limitation we want to acknowledge for this study. First is that our models likely are biased due to omitted variables such as Gross Domestic Product (GDP) or housing prices. Because we used data that was collected monthly, we could not use data collected quarterly, like GDP, since there may not be enough variation to produce a signal. We could not incorporate a housing price index variable since data was found to go as far back as 1980, and this may not be enough data for our models to pick up a signal. The AUC values for our models were within fair range, but they could be better. Running a model with more highly significant variables would likely increase the AUC value. Due to the limitations pointed out here as well as by other economic researchers, we will never know what the true probability of recession is due to their unpredictable nature.

In future work, it would be useful to convert the data used in this study to quarterly data so that we can use other data that is only collected quarterly, like GDP. It would also be interesting to see how using quarterly data would affect the performance of the

models ran in this study. It would also be interesting to try a logistic regression and see how its performance compares to the Probit regression. It also may be useful to try bootstrapping methods to test data only available for a limited amount of time, such as the housing price index.

7 Appendix

7.1 Figures

Figure 1: The Standard Normal Cumulative Density Function

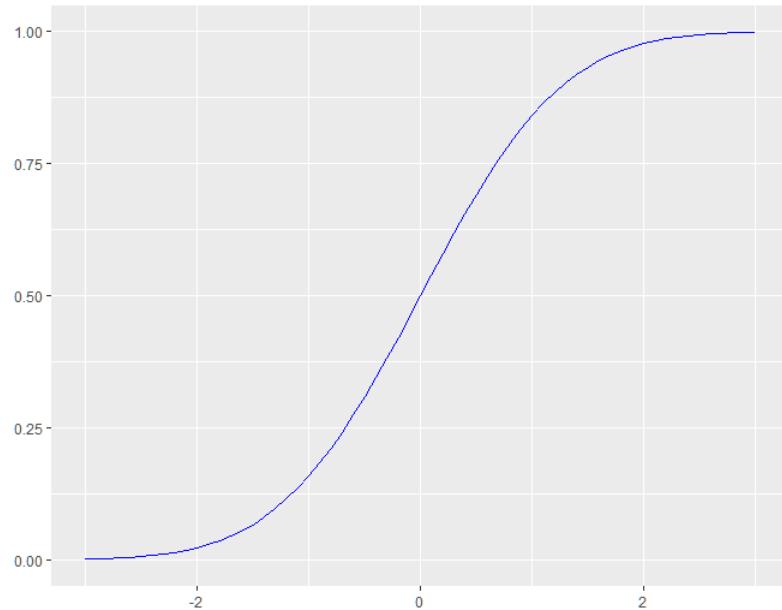


Figure 2: Yield Spread Between 3-month Treasury Bill and 10-year Treasury Securities

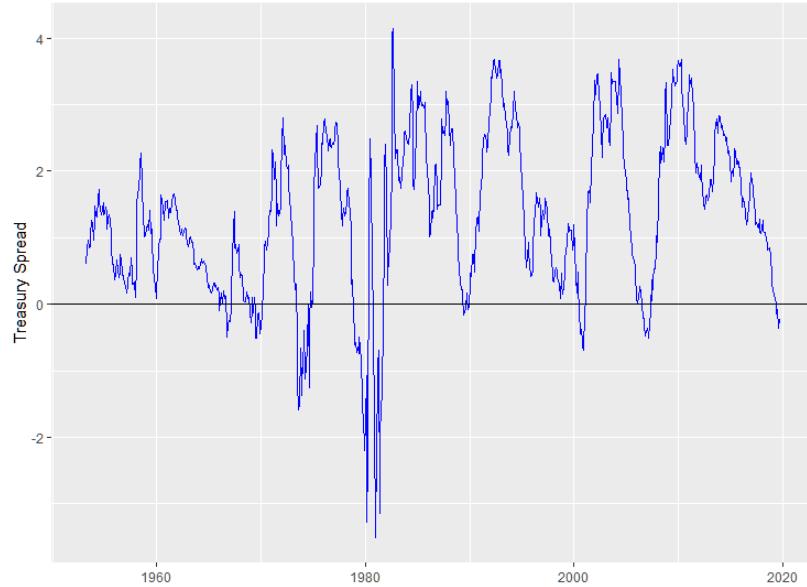


Figure 3: Probability of Recession Three Months Ahead

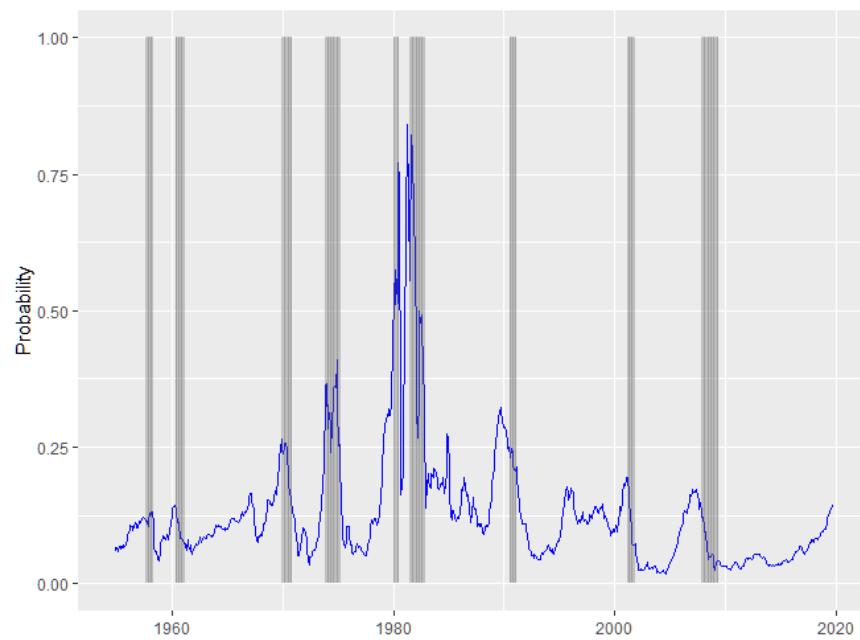


Figure 4: Probability of Recession Six Months Ahead

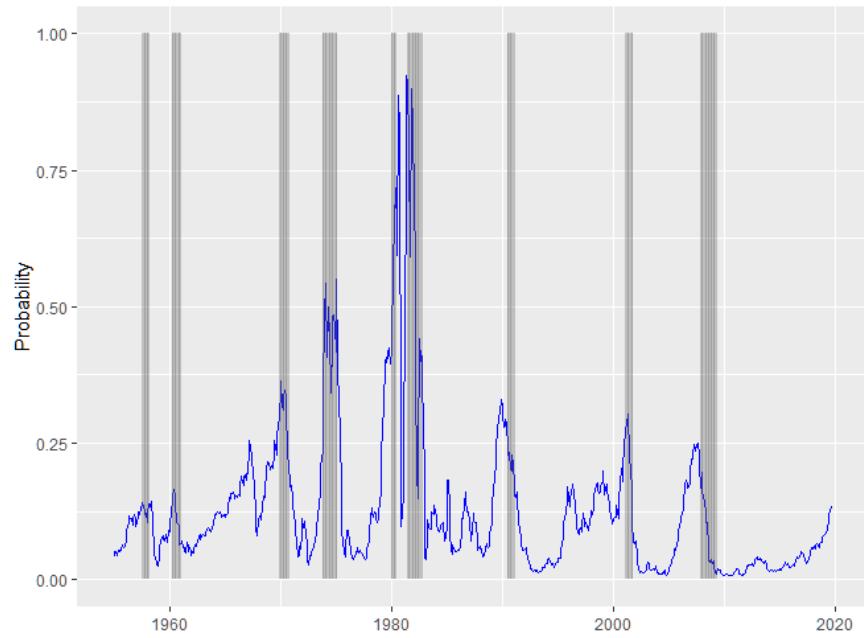


Figure 5: Probability of Recession One Year Ahead

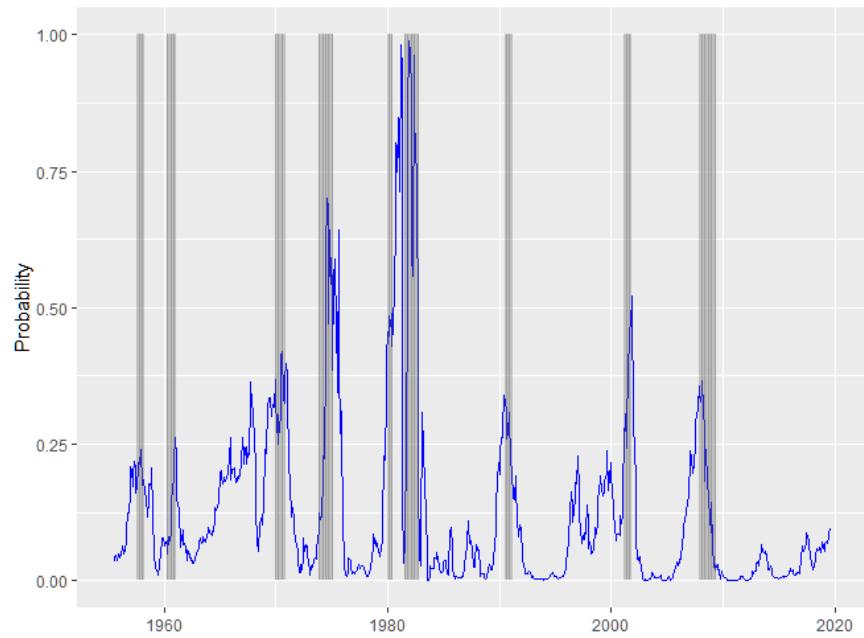


Figure 6: Probability of Recession Two Years Ahead

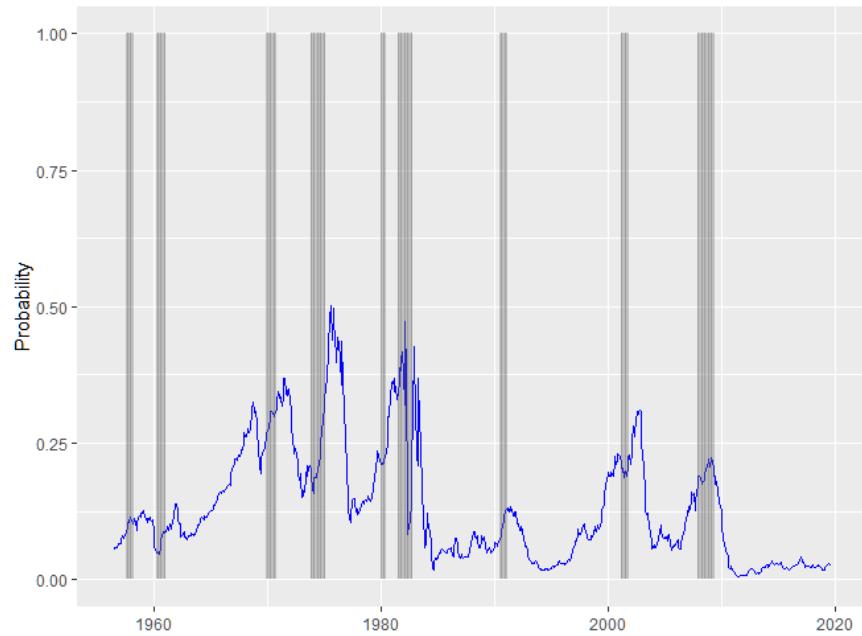
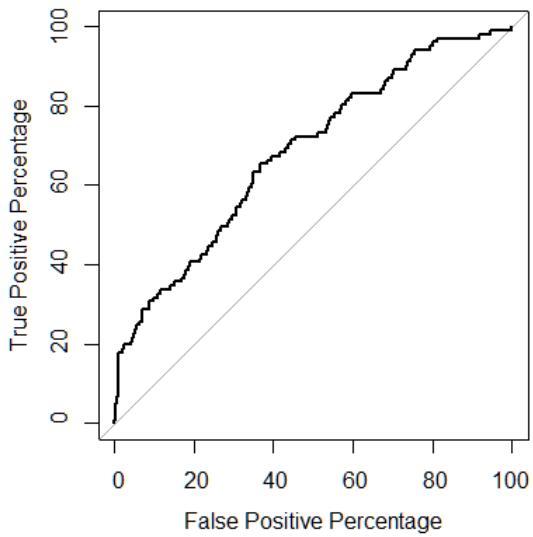


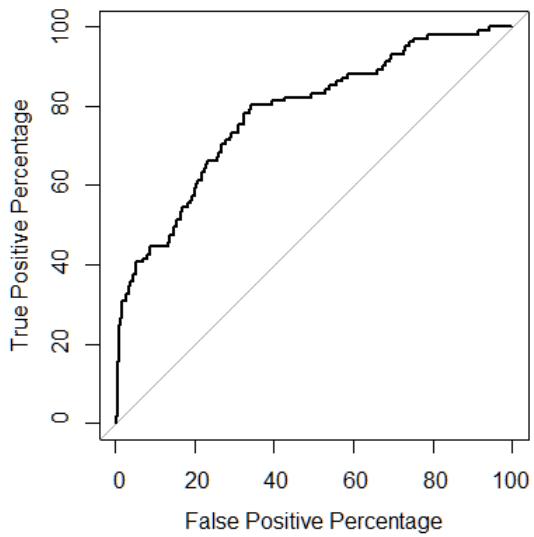
Figure 7: ROC Curves for Model A

Three months:



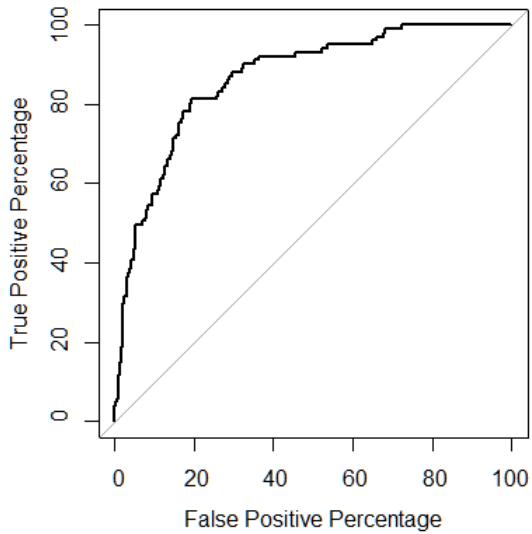
AUC = 68.26%

Six months:



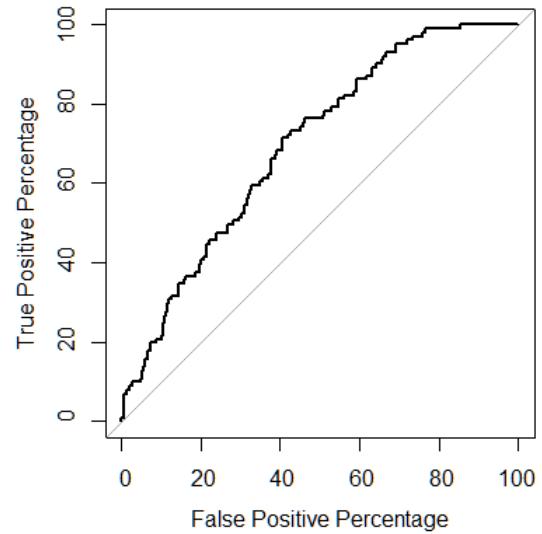
AUC = 77.87%

One year:



AUC = 86.44%

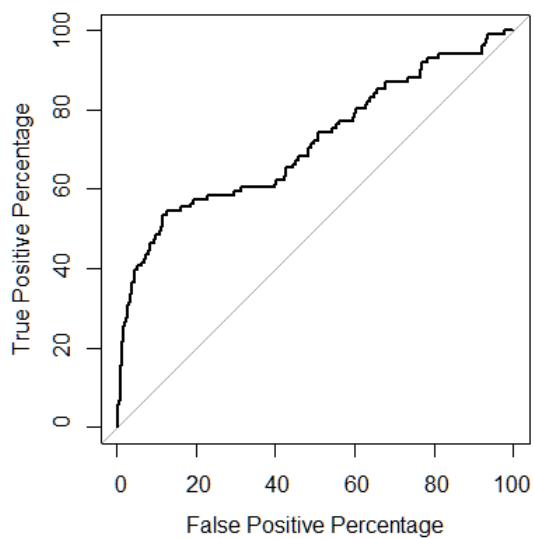
Two years:



AUC = 69.43%

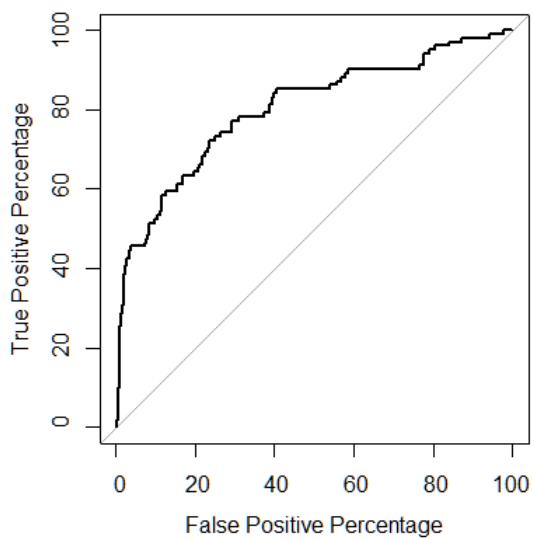
Figure 8: ROC Curves for Model B

Three months:



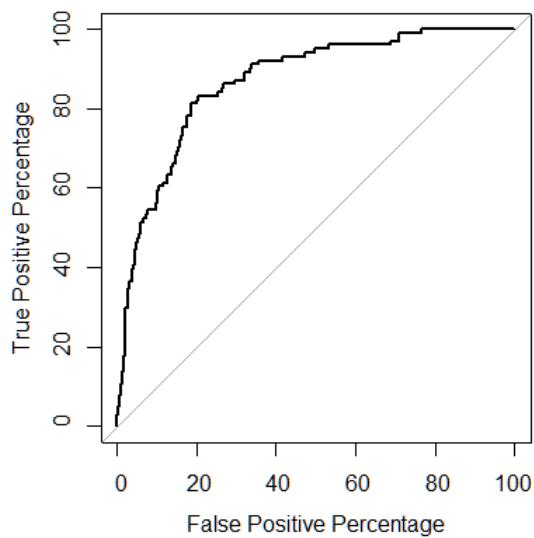
AUC = 71.31%

Six months:



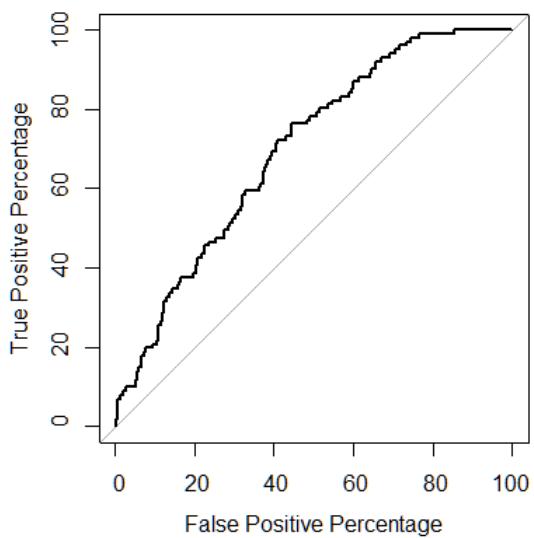
AUC = 79.88%

One year:



AUC = 86.72%

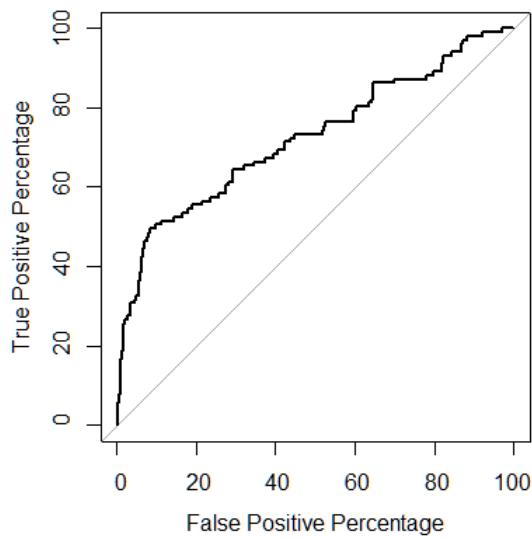
Two years:



AUC = 69.59%

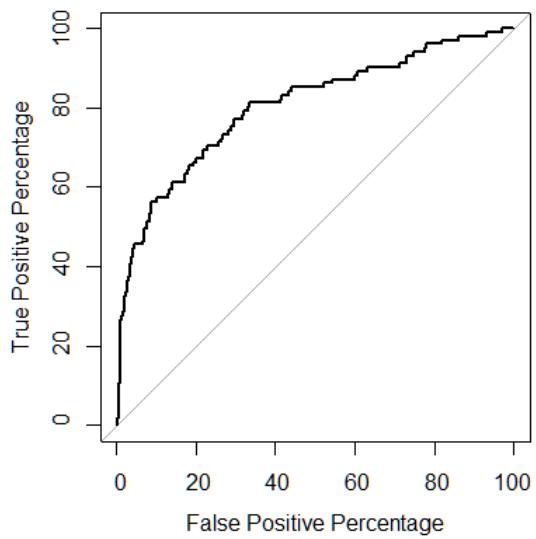
Figure 9: ROC Curves for Model C

Three months:



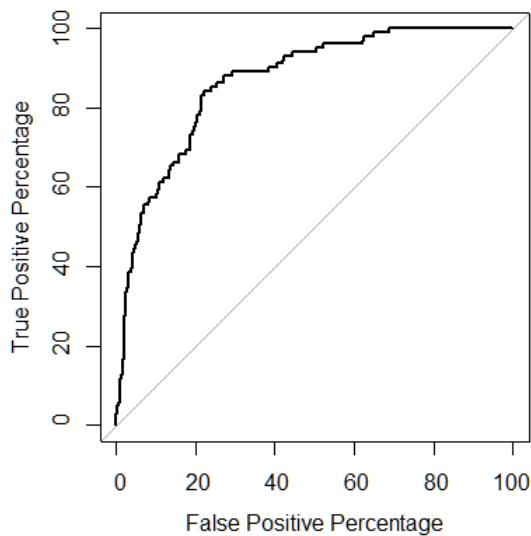
AUC = 72.15%

Six months:



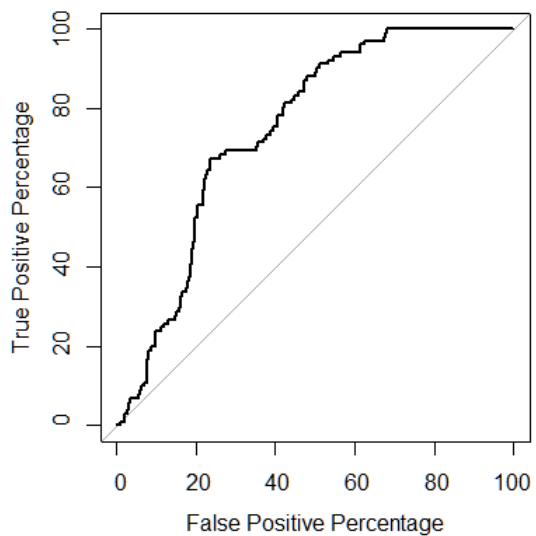
AUC = 80.19%

One year:



AUC = 86.71%

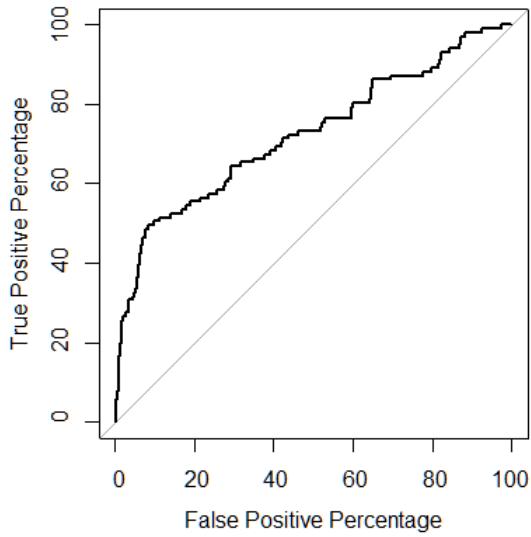
Two years:



AUC = 74.76%

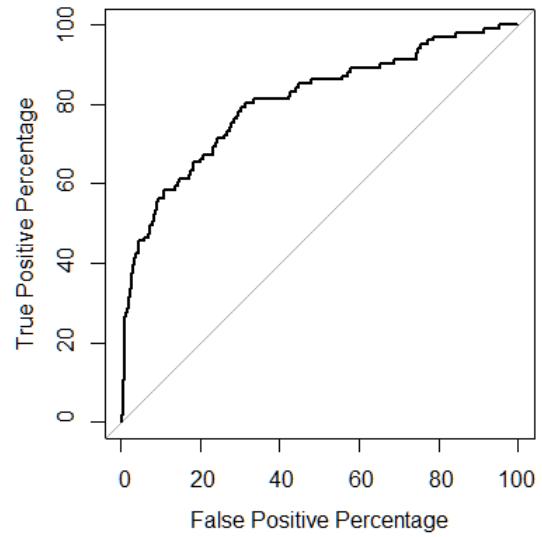
Figure 10: ROC Curves for Model D

Three months:



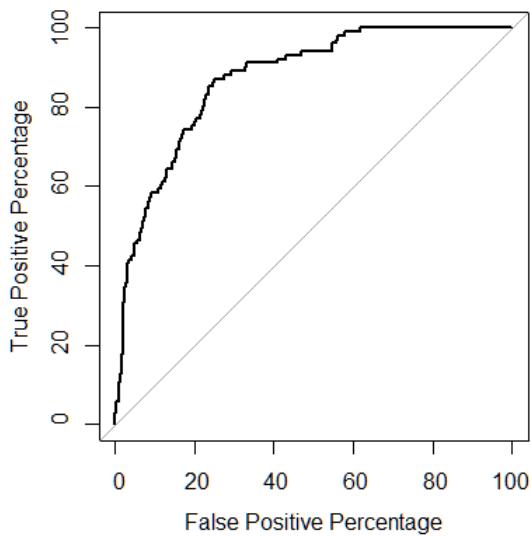
AUC = 72.12%

Six months:



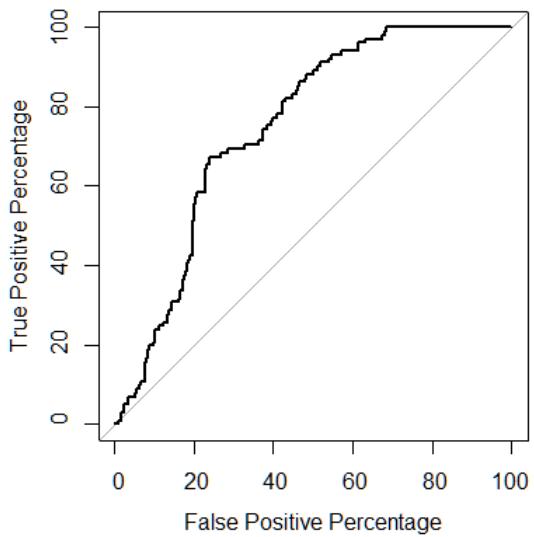
AUC = 80.35%

One year:



AUC = 87.04%

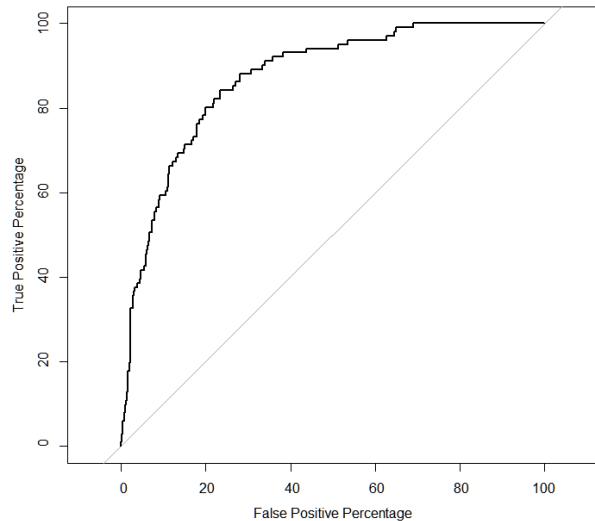
Two years:



AUC = 74.83%

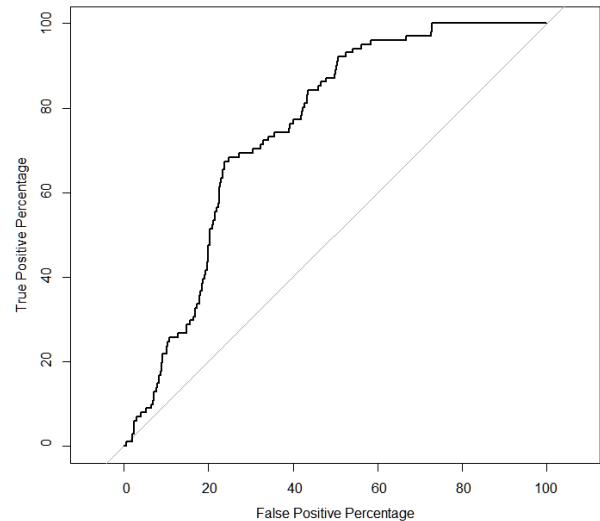
Figure 11: ROC Curves for Unique Stepwise Models

One year:



AUC = 87.12%

Two years:



AUC = 74.88%

7.2 Tables

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
NBER Recession Indicator	0.129	0.335	0	1
Three Month Treasury Yield (%)	4.502	3.226	0.01	17.237
Ten Month Treasury Yield (%)	5.839	2.852	1.5	15.32
Spread (%)	1.337	1.236	-3.505	4.146
Federal Funds Rate	4.785	3.595	0.07	19.1
Consumer Price Index	119.33	75.864	26.71	256.36
Industrial Production Index	63.73	28.505	18.61	110.55
Election Year	0.246	0.431	0	1

n=783

Table 2: Probit Coefficients for Forecasting NBER Recessions Three Months Ahead

Model	A	B	C	D
Intercept	-0.792*** (0.075)	-1.515*** (0.157)	-1.259*** (0.292)	-1.257*** (0.294)
Ten-Year less Three Month Spread	-0.309*** (0.048)	-0.142** (0.055)	-0.174** (0.059)	-0.174** (0.059)
Federal Funds Rate		0.097*** (0.018)	0.102*** (0.018)	0.102*** (0.018)
Consumer Price Index			0.006 (0.004)	0.006 (0.004)
Industrial Production Index			-0.016 (0.011)	-0.016 (0.011)
Election Year				-0.005 (0.141)

****, p ≈ 0, ***, p ≤ 0.001, **, p ≤ 0.01. *, p ≤ 0.05

Standard errors are in parentheses.

Table 3: Probit Coefficients for Forecasting NBER Recessions Six Months Ahead

Model	A	B	C	D
Intercept	-0.662*** (0.075)	-1.269*** (0.163)	-1.292*** (0.305)	-1.312*** (0.308)
Ten-Year less Three Month Spread	-0.481*** (0.055)	-0.326*** (0.062)	-0.347*** (0.067)	-0.347*** (0.067)
Federal Funds Rate		0.082*** (0.020)	0.086*** (0.020)	0.086*** (0.020)
Consumer Price Index			0.003 (0.005)	0.003 (0.005)
Industrial Production Index			-0.005 (0.012)	-0.005 (0.012)
Election Year				0.066 (0.146)

****, p ≈ 0, ***, p ≤ 0.001, **, p ≤ 0.01. *, p ≤ 0.05

Standard errors are in parentheses.

Table 4: Probit Coefficients for Forecasting NBER Recessions One Year Ahead

Model	A	B	C	D
Intercept	-0.543*** (0.078)	-0.789*** (0.179)	-1.178*** (0.320)	-1.239*** (0.323)
Ten-Year less Three Month Spread	-0.697*** (0.068)	-0.624*** (0.080)	-0.605*** (0.086)	-0.613*** (0.086)
Federal Funds Rate		0.034 (0.023)	0.031 (0.023)	0.033 (0.023)
Consumer Price Index			-0.005 (0.005)	-0.005 (0.005)
Industrial Production Index			0.016 (0.012)	0.016 (0.012)
Election Year				0.235 (0.158)

****, p ≈ 0, ***, p ≤ 0.001, **, p ≤ 0.01. *, p ≤ 0.05

Standard errors are in parentheses.

Table 5: Probit Coefficients for Forecasting NBER Recessions Two Years Ahead

Model	A	B	C	D
Intercept	-0.774*** (0.076)	-0.803*** (0.160)	-1.545*** (0.285)	-1.557*** (0.289)
Ten-Year less Three Month Spread	-0.308*** (0.049)	-0.301*** (0.060)	-0.223*** (0.065)	-0.223*** (0.065)
Federal Funds Rate		0.004 (0.019)	-0.019 (0.022)	-0.018 (0.022)
Consumer Price Index			-0.022*** (0.004)	-0.022*** (0.004)
Industrial Production Index			0.051*** (0.011)	0.051*** (0.011)
Election Year				0.038 (0.141)

****, p \approx 0, ***, p \leq 0.001, **, p \leq 0.01. ., p \leq 0.05

Standard errors are in parentheses.

Table 6: Probability of Recession t months from September 2019

Model	A	B	C	D
Three Months	0.23	0.10	0.14	0.14
Six Months	0.24	0.13	0.18	0.17
One Year	0.13	0.11	0.11	0.10
Two Years	0.13	0.13	0.03	0.03

Table 7: Stepwise Probit Coefficients

Forecast Horizon	Three Months (AIC = 536.35)	Six Months (AIC = 492.91)	One Year (AIC = 442.47)	Two Years (AIC = 535.41)
Intercept	-1.259*** (0.292)	-1.269*** (0.163)	-0.856*** (0.184)	-1.646*** (0.258)
Ten-Year less Three Month Spread	-0.174** (0.059)	-0.326*** (0.062)	-0.632*** (0.081)	-0.195*** (0.055)
Federal Funds Rate	0.102*** (0.019)	0.082*** (0.020)	0.036 (0.023)	
Consumer Price Index	0.006 (0.004)			-0.021*** (0.004)
Industrial Production Index	-0.016 (0.011)			0.050*** (0.011)
Election Year			0.239 (0.157)	

****, p ≈ 0, **, p ≤ 0.001, *, p ≤ 0.01. ., p ≤ 0.05

Standard errors are in parentheses.

7.3 R Code

Listing 1: Code used to Generate Election Dummy Variable

```
indicators %>% mutate(election.y = (indicators$raDATE
%>% parse_date(format = "%Y-%m")
%>% year() %>% mod(4) %>% {ifelse(. == 0, 1, 0)}))
```

Listing 2: Code used to Plot the Standard Normal Cumulative Density Function

```
library(tidyverse)
library(ggfortify)
ggdistribution(pnorm, seq(-3, 3, 0.1), mean = 0, sd = 1, colour = 'blue')
```

Listing 3: Code used to Plot the Term Spread

```
ggplot(indicators, aes(x=raDATE, y=SPREAD))
+ geom_line(color="blue") + geom_abline(slope=0, intercept=0)
+ xlab("") + ylab("Treasury_Spread")
```

Listing 4: Code used to Predict the Probability of Recession Three Months Ahead

```
library (MASS)

indicators <- read.csv ("Data/Aggregated_Data/indicatorsfedfunds.csv",
  col_types = cols (USREC = col_integer (),
  raDATE = col_date (format = "%Y-%m")))

spread <- lag (indicators$SPREAD, 3)
spread.t <- tail (spread, 780)
USREC.t <- tail (indicators$USREC, 780)
date.t <- tail (indicators$raDATE, 780)
INDPRO <- lag (indicators$INDPRO, 3)
INDPRO.t <- tail (INDPRO, 780)
CPI <- lag (indicators$CPIAUCSL, 3)
CPI.t <- tail (CPI, 780)
FFR <- lag (indicators$FEDFUNDS, 3)
FFR.t <- tail (FFR, 780)
election.t <- tail (election.y, 780)
model <- glm (USREC.t ~ spread.t + FFR.t + CPI.t + INDPRO.t + election.t,
  family = binomial (link="probit"))

step <- stepAIC (model)
summary (model)
summary (step)

stepModel <- glm (USREC.t ~ spread.t + CPI.t + INDPRO.t,
  family = binomial (link="probit"))

zscores <- predict (stepModel)
probs <- pnorm (zscores)

p1 <- ggplot () + geom_line (aes (x=date.t, y=probs), color="blue")
p1+geom_col (aes (x=date.t, y=USREC.t), alpha=0.3) + xlab ("")
+ ylab ("Probability")
```

Listing 5: Code used to Predict the Probability of Recession Six Months Ahead

```
spread <- lag (indicators$SPREAD, 6)
```

```

spread.t <- tail(spread, 777)
USREC.t <- tail(indicators$USREC, 777)
date.t <- tail(indicators$raDATE, 777)
INDPRO <- lag(indicators$INDPRO, 6)
INDPRO.t <- tail(INDPRO, 777)
CPI <- lag(indicators$CPIAUCSL, 6)
CPI.t <- tail(CPI, 777)
FFR <- lag(indicators$FEDFUNDSD, 6)
FFR.t <- tail(FFR, 777)
election.t <- tail(election.y, 777)

```

Listing 6: Code used to Predict the Probability of Recession One Year Ahead

```

spread <- lag(indicators$SPREAD, 12)
spread.t <- tail(spread, 771)
USREC.t <- tail(indicators$USREC, 771)
date.t <- tail(indicators$raDATE, 771)
INDPRO <- lag(indicators$INDPRO, 12)
INDPRO.t <- tail(INDPRO, 771)
CPI <- lag(indicators$CPIAUCSL, 12)
CPI.t <- tail(CPI, 771)
FFR <- lag(indicators$FEDFUNDSD, 12)
FFR.t <- tail(FFR, 771)
election.t <- tail(election.y, 771)

```

Listing 7: Code used to Predict the Probability of Recession Two Years Ahead

```

spread <- lag(indicators$SPREAD, 24)
spread.t <- tail(spread, 759)
USREC.t <- tail(indicators$USREC, 759)
date.t <- tail(indicators$raDATE, 759)
INDPRO <- lag(indicators$INDPRO, 24)
INDPRO.t <- tail(INDPRO, 759)

```

```
CPI <- lag(indicators$cPIAUCSL, 24)
CPI.t <- tail(CPI, 759)
FFR <- lag(indicators$FEDFUNDS, 24)
FFR.t <- tail(FFR, 759)
election.t <- tail(election.y, 759)
```

Listing 8: Code used to Produce Summary Statistics

```
summary(indicators)
sd(indicators$T10YCM)
sd(indicators$USREC)
sd(indicators$cPIAUCSL)
sd(indicators$SPREAD)
sd(indicators$FEDFUNDS)
sd(indicators$INDPRO)
```

Listing 9: Code used to Produce Table 6

```
probs[780]
probs[777]
probs[771]
probs[759]
```

Listing 10: Code used to Plot ROC Curves

```
library(pROC)
roc(USREC.t, probs, plot=TRUE, legacy.axes=TRUE,
percent=TRUE, xlab="False Positive Percentage",
ylab="True Positive Percentage")
```

References

- [1] Us business cycle expansions and contractions. <https://www.nber.org/cycles.html>. Accessed: 11-06-2019.
- [2] Jörg Döpke, Ulrich Fritzsche, and Christian Pierdzioch. Predicting recessions with boosted regression trees. *International Journal of Forecasting*, 33(4):745–759, Oct 2017.
- [3] Arturo Estrella. The term structure as a predictor of real economic activity. *The journal of finance*, 46(2):555–576, 1991.
- [4] Arturo Estrella. The yield curve as a predictor of recessions in the united states and europe. pages 324–337, 1996.
- [5] Arturo Estrella and Mary R. Trubin. The yield curve as a leading indicator: some practical issues. *Current Issues in Economics and Finance*, (125), 2006.
- [6] Andrew J. Filardo. How reliable are recession prediction models? *Economic review*, 84(2):35–55, 1999.
- [7] Serena Ng. Boosting recessions.
- [8] Board of Governors of the Federal Reserve System. H.15 selected interest rates. <https://www.federalreserve.gov/datadownload/Download.aspx?rel=H15&series=d7e27b7b09a3a7feae9>. Accessed: 11-16-2019.
- [9] Federal Reserve Bank of St. Louis. Consumer price index for all urban consumers: All items in u.s. city average. <https://fred.stlouisfed.org/series/CPIAUCSL>. Accessed: 11-16-2019.
- [10] Federal Reserve Bank of St. Louis. Industrial production index. <https://fred.stlouisfed.org/series/INDPRO>. Accessed: 11-16-2019.
- [11] Federal Reserve Bank of St. Louis. Nber based recession indicators for the united states from the period following the peak through the trough. <https://fred.stlouisfed.org/series/USREC>. Accessed: 11-16-2019.
- [12] Xavier Robin, Natacha Turck, Alexandre Hainard, Natalia Tiberti, Frédérique Lisacek, Jean-Charles Sanchez, and Markus Müller. proc: an open-source package

for r and s+ to analyze and compare roc curves. *BMC Bioinformatics*, 12:77, 2011.

- [13] Thomas G. Tape. Interpreting diagnostic tests. <http://gim.unmc.edu/dxtests/Default.htm>. Accessed: 12-1-2019.
- [14] E. J. Wagenmakers and S. Farrell. Aic model selection using akaike weights. *Psychonomic Bulletin & Review*, 11(1):192–196, 2004.
- [15] Jonathan H. Wright. The yield curve and predicting recessions. Technical report, 2006.