Economic Indicators as Predictors of Recessions

A Major Qualifying Project

Submitted to the Faculty of

Worcester Polytechnic Institute

in partial fulfillment of the requirements for the

Degree of Bachelor of Science

By

Alexander Fair

Timothy Day

Date: 4/30/2020

Project Advisors:

Professor Barry Posterro, Advisor

Professor Jon Abraham, Advisor

Table of Contents

Introduction

Even though it may not seem like it, a walk into an average convenience store reveals the amalgamation of an entire global supply chain. The products on the shelves and their ingredients have passed through dozens of hands and across countless borders. No longer is the food at the local grocer the product of the local butchers and farmers. Rather, it is the result of a complex network of international, globalized trade.[1] With this trade comes the complex financial institutions that make it all possible: stock markets, investment banking, insurance companies, and so forth. Clearly, the modern world is governed by complex economic forces. These global markets influence the prices of all of the goods and services people use on a daily basis.

Yet, while the machinery of the modern economy works to bring people the necessities of modern life, this machinery is imperfect. One could point to the myriad of social issues that stem from the modern mode of production. However, one of the most prominent flaws of the modern economy - and one of the most studied - is the cycle of booms and busts. In other words, the economy periodically experiences recessions: the stock markets crash, production slows, and a lot of people lose their jobs.

A recession is commonly defined as two quarters of decline in the GDP. In the U.S., it is the National Bureau of Economic Research which officially declares a recession.[2] Specifically, the NBER's Business Cycle Dating Committee determines the peak and trough of the U.S. business cycle and releases this information publicly.[3] The business cycle is the rise and fall in economic productivity and output of goods and services.[4] The period of decline between the peak

[1] https://www.piie.com/microsites/globalization/what-is-globalization
[2] https://www.investopedia.com/terms/r/recession.asp
[3] https://faculty.fuqua.duke.edu/~charvey/Research/Thesis/Thesis.htm, https://www.nber.org/cycles.html
[4] https://www.investopedia.com/terms/b/businesscycle.asp

of the business cycle and the trough is referred to as a **recession** or economic contraction. The NBER's determination for recessions is more complex than the common definition given above, however.

The Business Cycle Dating Committee considers a number of different indicators beyond simply Gross Domestic Product (GDP) when determining the peaks and troughs of business cycles. It also considers Gross Domestic Income and employment rates to paint a clearer picture of the economy.[5] It is important to remember that the recession start and end dates are not calculated. Rather, they are determined by committee, which introduces a degree of subjectivity.[6] According to the NBER, there have been thirty-three recessions since 1854.[7] On average, these recessions have lasted 17.5 months each.[8] While it is obviously useful to determine recessions after they have occurred, it is more useful to be able to predict when they will happen in the future.

There have been many efforts to crack the code of the business cycle in order to try and predict when the next recession will be. One of the most commonly cited indicators of recessions is known as the "Yield Curve Inversion," which is an event that has preceded each modern recession to this day[9]. The yield curve inversion occurs when the value of short-term debt instruments, commonly 2-year bonds, is greater than that of a long-term instrument, like a 10-year bond. When the 10-year-2-year spread goes negative, it can generally be viewed as a signal that a recession will occur at some time in the near future, but it does not necessarily encapsulate more predictive information beyond that. Other, less "news-worthy" indicators have

[5] https://www.nber.org/cycles/recessions_faq.html
[6] https://www.nber.org/cycles/recessions.html
[7] https://www.nber.org/cycles.html
[8] https://www.nber.org/cycles.html
[9] https://www.investopedia.com/terms/i/invertedyieldcurve.asp

also proven to have some predictive power, such as the 10-year-3-year spread and the **Near-Term Forward Spread**. The latter of these was introduced by Eric C. Engstrom and Stephen A. Sharpe in the Financial Analysts Journal in July 2019, in their article"The Near-Term Forward Yield Spread as a Leading Indicator: A Less Distorted Mirror[10]." Engstrom and Sharpe's work served as the main motivation to build our own predictive model that can be updated regularly in order to determine the probability of a recession six months in the future.

The goal of this project was to glean some understanding of the forces that can push a nation into recession. Through the process of gathering, cleaning, and visualizing this economic data, we get a high-level idea of the relationships between recessions and our indicators. In order to mathematically validate what can be seen on a graph, we created a predictive model for recessions with publicly available data, using several economic indicators. The model is accessible and can be updated continuously. We also introduce a few different definitions for recessions that may offer a way to tell if the nation is in a recession sooner than the National Bureau of Economic Research would retrospectively declare.

[10] https://www.federalreserve.gov/econres/feds/files/2018055r1pap.pdf

Background

It is important to first examine the mathematical and financial underpinnings of the work discussed in this paper. This overview will provide the necessary context required to understand the later sections. It is prudent to start with a dive into the mathematics underpinning the pricing of bonds and other similar securities.

Spot Rates

The spot rate is the annual rate of return that can be earned on capital for a specific duration. This terminology is often used to refer to the yearly rate of return of investments (usually bonds) with different durations at different points in time. The standard notation for the spot rate is $S_n(h)$, where n indicates the time at which the rate is quoted and h indicates the duration of the investment starting from time n. The notation $S_n(h) = i$ means that an investment quoted at time n that lasts h years past time n will accrue i% interest every year.

Spot Rates Over Time

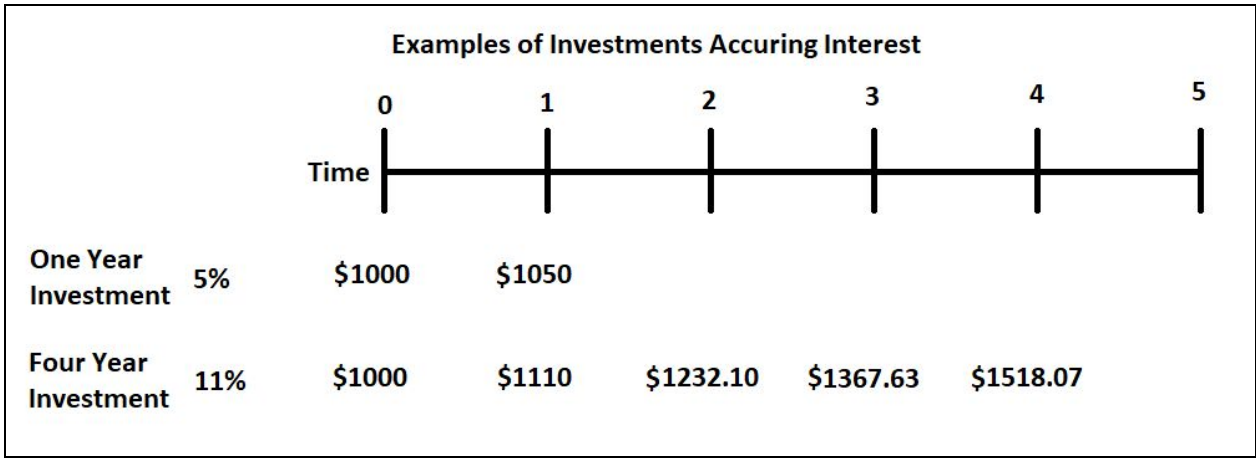
Term (yrs)	1	2	3	4	5
Spot Rate	5%	6%	7%	11%	15%

figure 1-1

Consider Figure 1-1; it displays five different spot rates for investments lasting 1, 2, 3, 4, and 5 years. In the standard notation, this would be represented as $S_0(1) = 0.05$, $S_0(2) = 0.06$, $S_0(3) = 0.07$, $S_0(4) = 0.11$, and $S_0(5) = 0.15$. A one-year investment of 1000 dollars purchased at time 0 would mature at time 1 with 5% interest, making its value $\$1000 \times (1.05) = \$1050$. A four-year investment of 1000 dollars purchased at time 0 would mature at time 4 and be worth $\$1000 \times (1.11)^4 = \$1518.07$. It may be noticed that all of the spot rates discussed thus far are

quoted at time 0. This is because spot rates quoted for future years have a different name: forward rates.

Forward Rates



Consider the forward rate $S_2(3)$. This represents the yearly return of an investment made two years from now with a maturity of three years. It could alternatively be represented by the notation $f_{[2,\,5]}$. It may be tempting to assume that this rate does not change with time i.e. $S_0(3) = S_2(3) = 0.07$ as per Figure 1.1. However, this may not be true. Differences in spot rates may cause differences in the returns offered by these different forward rates investments.

Take the following example. In case A, a 2-year 800 dollar investment is made at time 0. At time 2, the investment matures and yields 1000 dollars, a 25% gain. In case B, a 4-year 600 dollar investment is made at time 0. This investment is worth 1000 dollars at time 4. At time 2 (i.e. 2 years remain), this investment would be the same as getting a 2-year investment that yields 1000 at time 4. Assuming the spot rates remained the same, this would be worth 800 dollars. Hence, the 4-year investment could be sold for 800, a 33.33% gain. Clearly, it is more profitable to make a 4-year investment and sell it at time 2 (case B) than it is to simply make a 2-year investment (case A). Hence, $S_0(2) > S_2(2)$.

Market pressures, like arbitrage, can cause rates to change. In other words, the differences in interest rates for different securities could offer the chance to make money with little to no risk. People acting on this money making opportunity would affect future interest rates by changing the demand for different investments. It is important to calculate what the forward rates will be in the years to come, once market participants act on these arbitrage opportunities. These forward rates are governed by the Law of One Price. This theory states that at any point in time, two identical cash flows have to be worth the same irrespective of their source. Hence, a four-year investment should yield the same as a one-year investment followed by a three-year investment; a five-year investment should yield the same as a three-year investment followed by a two-year investment followed by a one-year investment.

To calculate the forward rates, the following formula may be used:

$$(1+S_0(a))^a(1+f_{[a,\,b]})^{b-a} = (1+S_0(b))^b.$$

This formula states that an investment lasting *b* years must have the same yield as an investment lasting *a* years followed by the investment lasting *b* - *a* years, as per the Law of One Price. Hence, we could find the spot rate of a *b* - *a* year investment at time *a*. Consider the example of $f_{[2,\,5]}$.

$$(1.06)^2(1+f_{[2,\,5]})^3 = (1.15)^5 \text{ yields } f_{[2,\,5]} = 0.214$$

This would be the three-year spot rate at time 2. This could be extended to predict the spot rates one year in the future. With this basic understanding of the mathematics governing bond pricing, it is time to examine the application of these tools. Yet first, a dive into the mechanics of recessions.

Recessions

A recession is commonly defined as two quarters of decline in the GDP. In the U.S., it is the National Bureau of Economic Research which officially declares a recession.[11] Specifically, the NBER's Business Cycle Dating Committee determines the peak and trough of the U.S. business cycle and releases this information publicly.[12] The business cycle is the rise and fall in economic productivity and output of goods and services.[13] The period of decline between the peak of the business cycle and the trough is referred to as a recession or economic contraction. The NBER's determination for recessions is more complex than the common definition given above, however.

The Business Cycle Dating Committee considers a number of different indicators beyond simply Gross Domestic Product (GDP) when determining the peaks and troughs of business cycles. It also considers Gross Domestic Income and employment rates to paint a clearer picture of the economy.[14] It is important to remember that the recession start and end dates are not calculated. Rather, they are determined by committee, which introduces a degree of subjectivity. [15] According to the NBER, there have been thirty-three recessions since 1854.[16] On average, these recessions have lasted 17.5 months each.[17] While it is obviously useful to determine recessions after they have occurred, it is more useful to be able to predict when they will happen in the future.

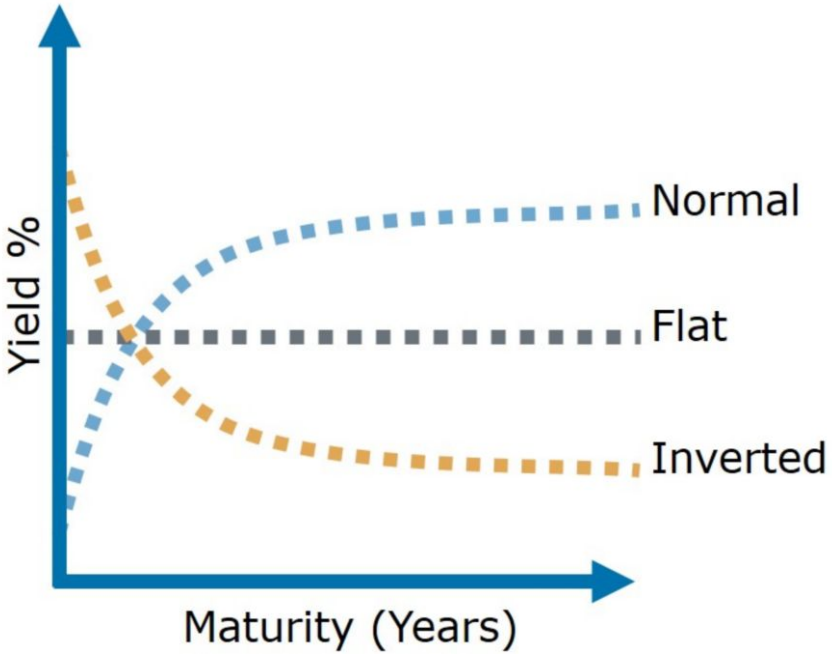
---

[11] https://www.investopedia.com/terms/r/recession.asp
[12] https://www.nber.org/cycles.html, https://www.nber.org/cycles/recessions_faq.html
[13] https://www.investopedia.com/terms/b/businesscycle.asp
[14] https://www.nber.org/cycles/recessions_faq.html
[15] https://www.nber.org/cycles/recessions.html
[16] https://www.nber.org/cycles.html
[17] https://www.nber.org/cycles.html

The Yield Curve

One tool to gauge market expectations is the yield curve. A yield curve is a collection of spot rates, commonly used to display returns for different maturities.[18] The rates of return typically differ between short-term bonds and long-term bonds, giving the chart its characteristic curve.[19] The often cited yield curve is that of U.S. the Treasuries. It is often used to measure market participants' expectations about risks in the U.S. economy.[20] Hence, there exist indicators to gauge risks in the economy based on the yield curve.

States of Yield Curve



[21]

[18] https://www.morningstar.com/articles/921187/what-does-inverted-yield-curve-mean
[19] https://www.morningstar.com/articles/921187/what-does-inverted-yield-curve-mean
[20] https://www.fidelity.com/learning-center/investment-products/fixed-income-bonds/bond-yield-curve
[21] https://www.colotrust.com/the-shape-of-the-u-s-treasury-yield-curve/

The 10 yr - 2 yr Spread

A common financial indicator based on the yield curve is the 10-year - 2-year spread. This indicator is based upon the 10-year Treasury Note and the 2-year Treasury Note. A Treasury Note is a form of US government debt security with a fixed interest rate and maturity between one and ten years.[22] The indicator is calculated by subtracting the 2-year Treasury rate from the 10-year Treasury rate i.e

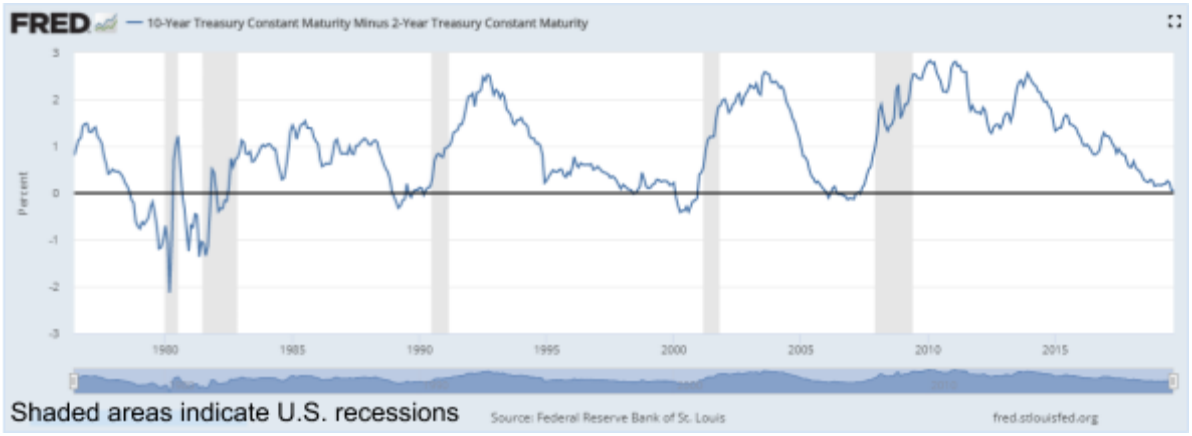
$$10yr - 2yr\ Spread = (10yr\ Treasury\ Rate - 2yr\ Treasury\ Rate).$$[23]

This indicator is used to attempt to predict the occurrence of a recession.

A positive 10-year - 2-yr spread implies that investors have a positive outlook on the economy; the higher 10-year yield implies that investors believe that there is more risk over the next decade than in the near term.[24] A negative 10-year - 2-year spread implies that investors have a negative outlook on the economy; investors believe there is a lot of risk in the near term, signaling an impending recession.[25] The 10-year - 2-year spread has gained popularity due to its accuracy in predicting impending recessions, as first shown in economist Campbell Harvey's 1986 dissertation.[26]

[22] https://www.investopedia.com/terms/t/treasurynote.asp
[23] https://seekingalpha.com/article/4055368-10-year-2-year-spread-reliable-recession-predictor
[24] https://seekingalpha.com/article/4055368-10-year-2-year-spread-reliable-recession-predictor
[25] https://seekingalpha.com/article/4055368-10-year-2-year-spread-reliable-recession-predictor
[26] https://faculty.fuqua.duke.edu/~charvey/Research/Thesis/Thesis.htm

10-year - 2-year Spread Since 1977



[27]

Near-Term Forward Spread

The Near-Term Forward Spread is another financial indicator based on the yield curve. It was introduced recently in a July 2019 paper by Eric C. Engstrom and Steven A. Sharpe.[28] While this indicator is also the difference between two Treasury bill rates, it differs from the 10-year - 2-year spread in that it uses rates from securities with much shorter maturities. The indicator is calculated by finding the difference between the forward rate that is expected on a three month Treasury bill six quarters in the future and the current interest rate on a three-month Treasury bill:

$$Near-Term\ Forward\ Spread = f_{[18month,\ 21month]} - f_{[0month,\ 3month]}$$

The Near-Term Forward Spread could also be used to give insight into the market's outlook on the economy and the likelihood of a recession, much as with the 10-year - 2-year spread.

Engstrom and Sharpe claim that the Near-Term Forward Spread is a better indicator of impending recession than the 10-year - 2-year spread.[29] The reason for their claim is that the

[27] https://fred.stlouisfed.org/
[28] https://www.cfainstitute.org/en/research/financial-analysts-journal/2019/0015198X-2019-1625617
[29] https://www.cfainstitute.org/en/research/financial-analysts-journal/2019/0015198X-2019-1625617

forward rate spread captures the market expectations for conventional monetary policy in the near-term. Conventional monetary policy refers to the changes to the interest rate at which the Federal Reserve lends money as a response to various economic drivers such as inflation, employment, and overall consumer prices. A negative near-term forward spread is an indicator that the market expects the Federal Reserve to lower its interest rates in response to a heightened risk of recession. Since the yield curve is simply an average of forward rates over the time to maturity, it can "tend to blur the signal embedded in forward rates."[30] Engstrom and Sharpe found that the probability of a recession increases by 47 percent when the near-term forward spread is decreased by one standard deviation, "an economically large effect that is statistically significant with a *p*-value of 1%."[31] They further extend their methodology to use the near-term spread as an indicator for GDP growth and excess equity returns.

Probit Model for Recession Probabilities

A probit (a portmanteau of probability and unit) model is a form of regression in which the dependent variable can only have two values.[32] This type of model is typically used to estimate the probability that a particular observation will land in one of the binary categories. We use the probit model instead of the more common ordinary least squares (OLS) regression, as the predictions for OLS regression are not bound between 0 and 1. In the case of predicting recessions, the binary variable is *Y=0* if the country is not in a recession, and *Y=1* if the country is experiencing a recession. Probit models typically take the form:

$$Pr(Y = 1|X) = \Phi(X^T\beta),$$

[30] https://www.cfainstitute.org/en/research/financial-analysts-journal/2019/0015198X-2019-1625617
[31] https://www.cfainstitute.org/en/research/financial-analysts-journal/2019/0015198X-2019-1625617
[32] https://en.wikipedia.org/wiki/Probit_model

in which Y represents the binary variable (mentioned above) and X represents a vector of independent variables assumed to affect the outcome. Hence, this model says that the probability that the event occurs (and $Y = 1$) is calculated by finding the value of the cumulative distribution function of the standard normal distribution using the value of the different independent variables. The parameters (betas) are usually estimated using maximum likelihood estimation.

Our independent variables are the 10-Year - 2-Year spreads, the 10-Year - 3-Month spread, and the Near-Term Forward Spread. The output of the regression gives us coefficients, of which we can interpret the sign, but not the magnitude. The coefficients tell us whether an increase in an independent variable increases or decreases the likelihood of Y=1.

Data and Methodology

The data for the 10-year - 2-year rates, 10-year - 3-month rates, the 3-month Treasury Bill rates, and the instantaneous forward rates is provided by the Economic Research Division of the Federal Reserve Bank of St. Louis.[33] The quarterly data for Gross Domestic Product is also provided by the same division of the Federal Reserve Bank.[34] The daily S&P 500 data was sourced from Yahoo finance.[35] And finally, the recession start and end dates are sourced from the National Bureau of Economic Research.[36]

There are several metrics that were derived from this data. The daily S&P 500 data was averaged to obtain the change in the index over a full quarter. From this, the 'S&P 500 Recession' was defined as any quarter with a decline in the index worse than the 7.5th percentile (in our data, this is a -0.061672% change in the S&P 500 over a quarter). Two other definitions of recession were used. The 'NBER Recession' was defined as any quarter that fell within the peak and trough in the business cycle as declared by the NBER. The 'GDP Recession' was defined as any quarter with a decline in the GDP worse than the 10th percentile (in our data, this is a 0.513519% change in the GDP over a quarter). In addition to these metrics, several bond spreads were calculated.

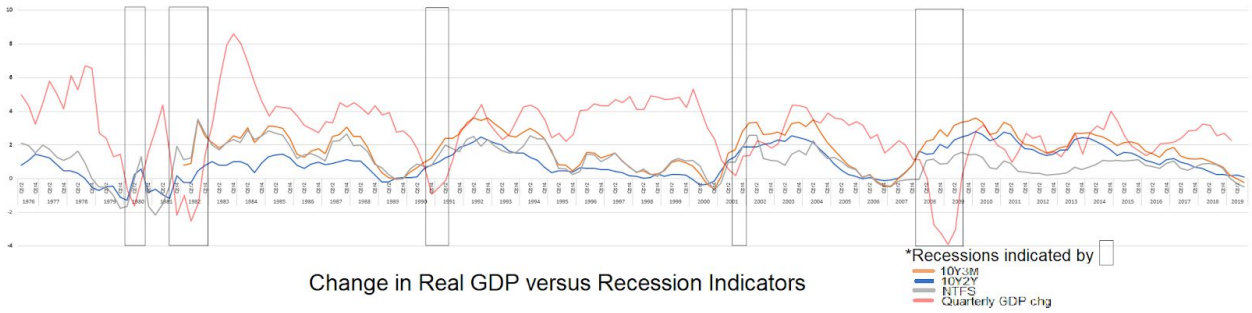
As mentioned earlier, the data used to calculate forward rates is provided by the Federal Reserve Bank of St. Louis. They produce a regularly-updated estimate of the yield curve dating back to 1961, based on the *Journal of Monetary Economics* article written by Gürkaynak et al.

[33] https://fred.stlouisfed.org/
[34] https://fred.stlouisfed.org/
[35] https://finance.yahoo.com/quote/%5EGSPC/history/
[36] https://www.nber.org/cycles.html

titled, “The U.S. Treasury yield curve: 1961 to the present.”[37] These data were also used by Engstrom and Sharpe, whose formula for calculating forward rates is:

$$f_t(n,0) = \beta_0 + \beta_1 exp(-\frac{n}{\tau_1}) + \beta_2(\frac{n}{\tau_1})exp(-\frac{n}{\tau_1}) + \beta_3(\frac{n}{\tau_2})exp(-\frac{n}{\tau_2}),$$

with parameters $[\beta_0, \beta_1, \beta_2, \beta_3, \tau_1, \tau_2]$, which are provided by the Federal Reserve Board.[38] We estimate the 3 month forward rate, six quarters ahead by setting n, the maturity of the forward date, equal to 1.625. Engstrom and Sharpe took the average of instantaneous forward rates over the period n=1.5 to n=1.75, but noted that using the midpoint of this interval in estimating the forward rates results in a difference of more than 1 basis point on only 0.13% of days, meaning 99.87% of days differ from the interval average by less than 1 basis point.[39] The spreads are then calculated by taking the difference between the 3 month Treasury bill rate and the forward rate. We compare these rates to the 10-year - 2-year spread and the 10-year 3-month spread as well as the change in GDP, the data for which is provided by the St. Louis Fed.[40]



Change in Real GDP versus Recession Indicators

This graph shows each of the spreads against change in real GDP with recessions indicated.

[37] https://www.sciencedirect.com/science/article/pii/S0304393207000840?via%3Dihub#sec3
[38] https://www.cfainstitute.org/en/research/financial-analysts-journal/2019/0015198X-2019-1625617, https://www.sciencedirect.com/science/article/pii/S0304393207000840?via%3Dihub#sec3
[39] https://www.cfainstitute.org/en/research/financial-analysts-journal/2019/0015198X-2019-1625617
[40] https://fred.stlouisfed.org/

Probit Model Methodology

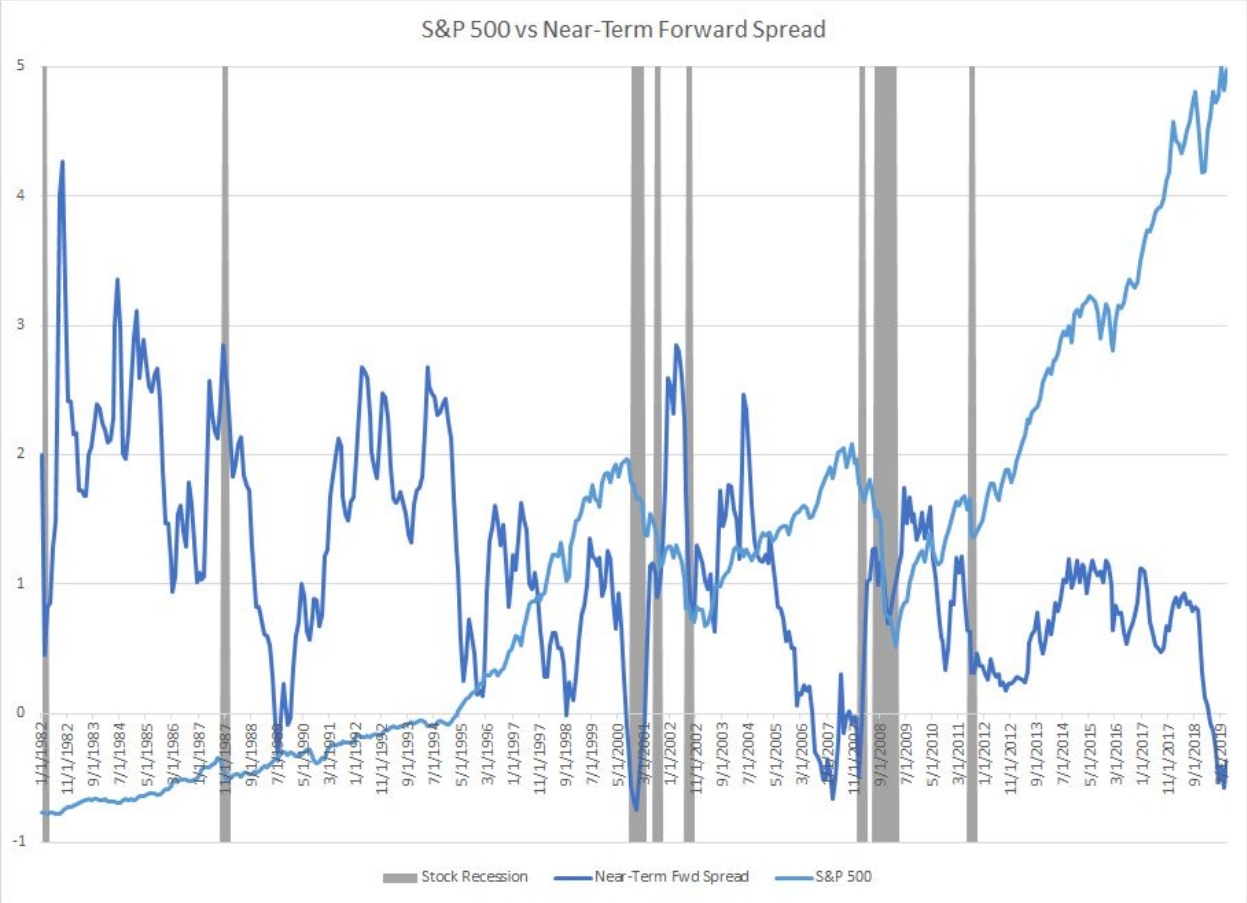
In order to get an understanding of how these indicators impact the probability of a recession, we utilized R, which has many powerful built-in features that cover a variety of statistical analyses. Using the generalized linear model (GLM) function in R, we can set the link function as "probit," which has already been set as the cumulative distribution function of the standard normal distribution. The glm function automatically performs a maximum likelihood estimation to determine the coefficients of each explanatory variable. As previously mentioned, the dependent variable we are concerned about is an indicator with a value of 0 if there is no recession, and 1 if there is a recession. Our independent variables are the 10-year - 2-year term spread, the 10-year - 3-month term spread, and the near-term forward spread. In order to train the model and predict 6 months ahead, we created a column that indicates whether the US experiences a recession at any time in the next 6 months, and used that as the dependent variable in the R model. Additionally, we created a model for each category of recession as defined previously.

The output of the GLM function is a maximum likelihood estimate of the coefficients for each of the independent variables, along with their standard errors. For each independent variable we created a univariate probit model, meaning that we are only considering one variable at a time. Additionally, we created a multivariate model using all of the variables in our consideration. For the multivariate model, we calculated the average marginal effects for each of the variables in order to understand how the individual indicators change the probability of a recession per unit change. We then used the R predict function to test the accuracy of each of our

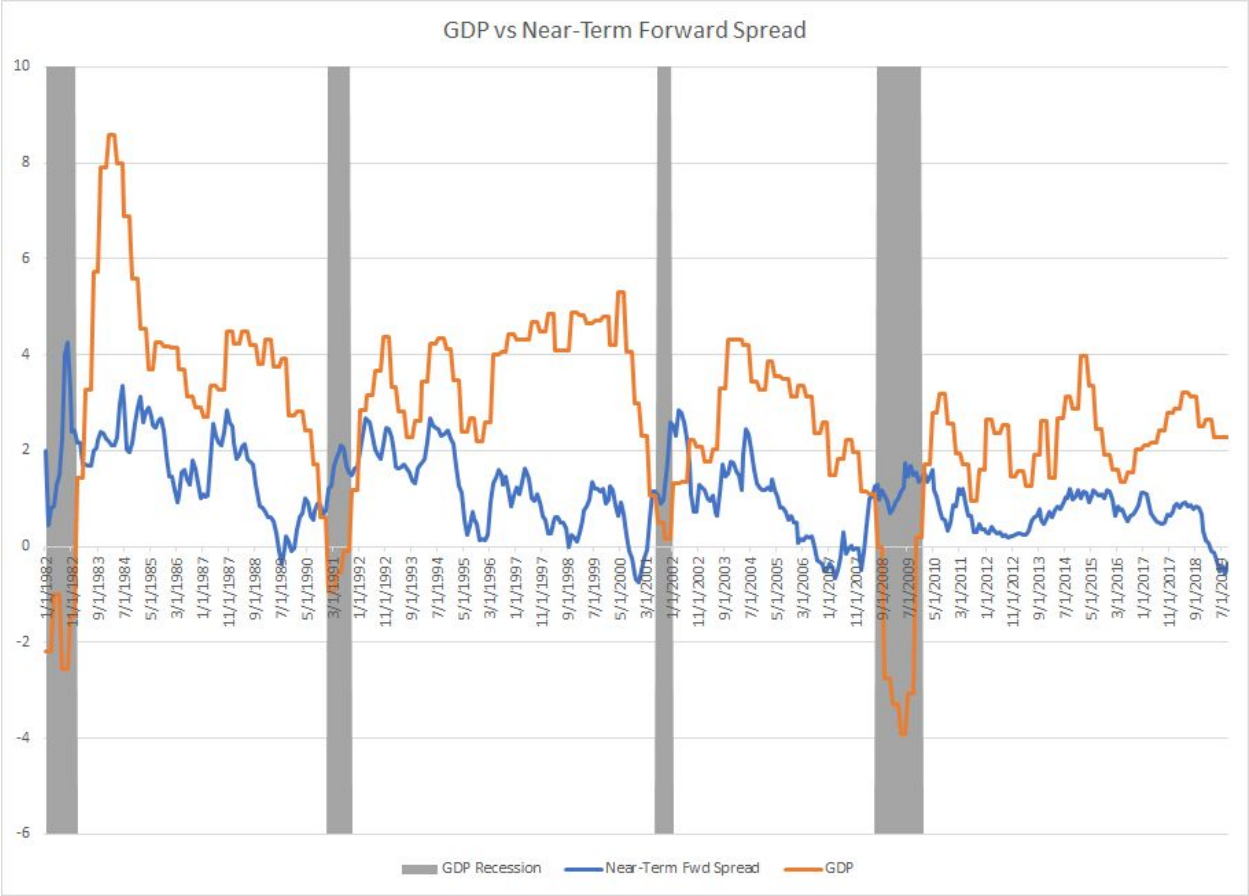
models, based on the percent of recession days that were accurately predicted. From these models, we are able to derive a probability of recession given a value of an indicator.

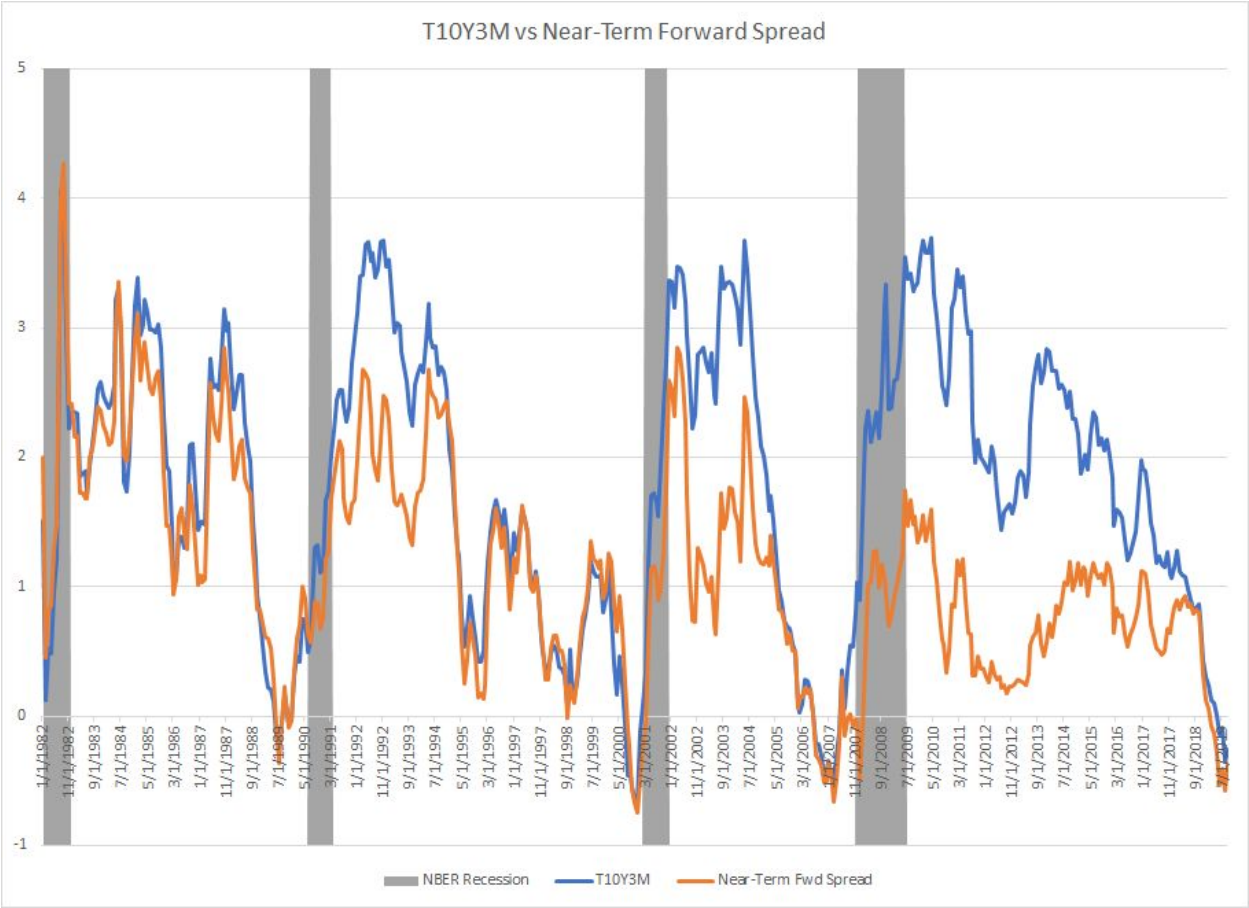
Results and Conclusion

The following graphic illustrates a direct comparison between the different recession definitions outlined earlier. The red bars represent quarters which the definition on the left designate as a time of recession and the green bars, of course, represent no recession. The recessions primarily revolve around a few time frames: 1982, 1990-91, 2001, 2008-09.



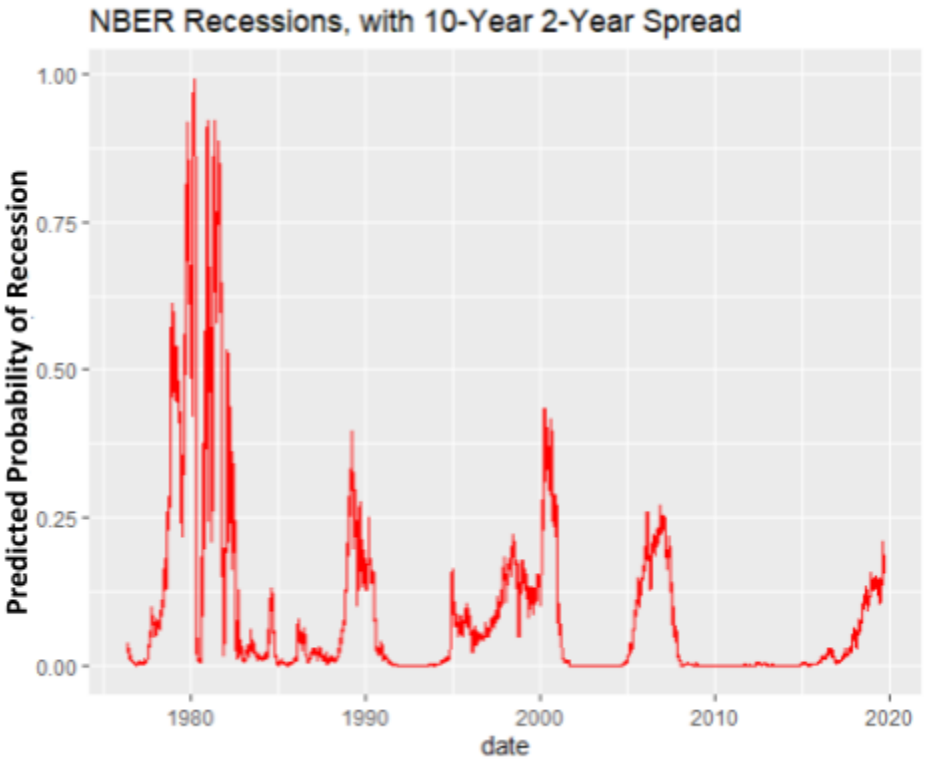
In the following charts, the differing recession definitions are compared to relevant indicators. The first of these is the S&P 500. Most notably, each 'stock recession' is accompanied by a drop in S&P 500 index and preceded by an inversion of the Near-Term Forward spread (neglecting the 1982 recession, which lacks adequate data). The next chart compares the standard NBER definition with two bond indices. As expected, both the Near-Term Forward spread and the 10YR - 3M spread invert preceding the recession; this is a visual representation of the predictive potential of these bond indices. The final chart is a comparison of the 'GDP recession' definition with the GDP and the Near-Term Forward spread. This definition, once again, is accompanied by drop in the Near-Term Forward spread preceding each recession. With this overview of the available data, it is important to analyse the effectiveness of the probit model.



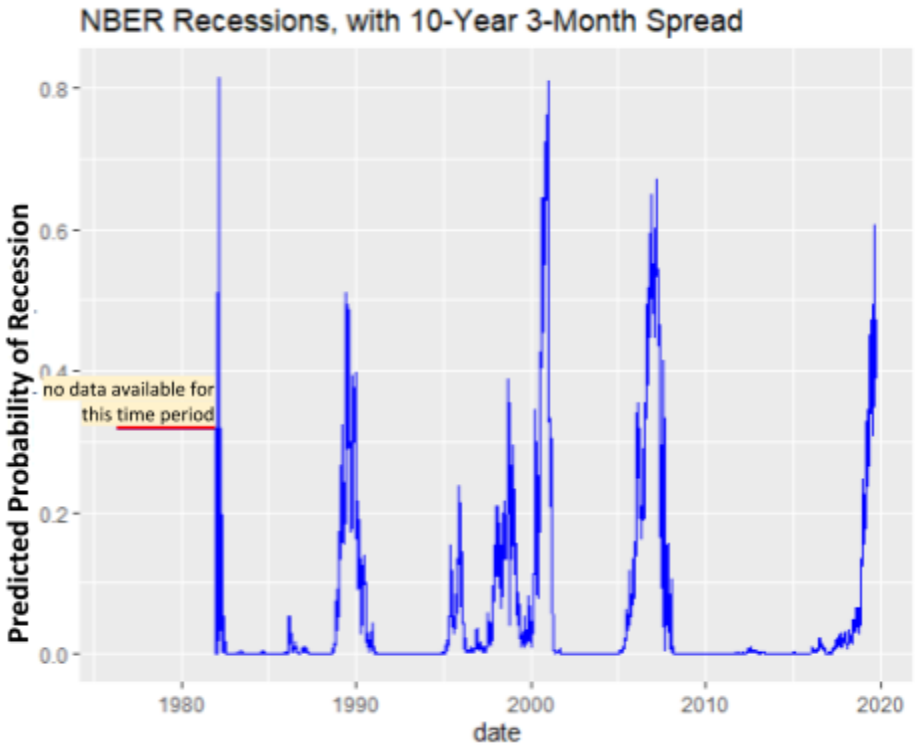


Probit Model Results

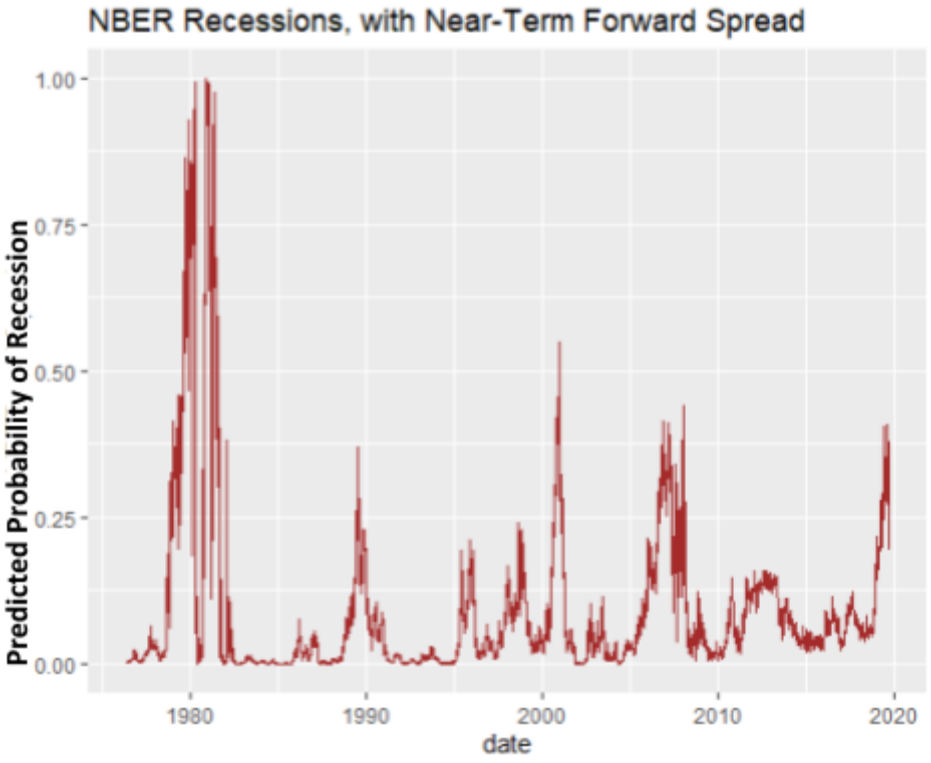
For each version of the model, we graphed the predicted probabilities of recession implied by each indicator, or combinations of indicators. Graphs 1-3 depict the predicted probabilities of the univariate models with each of the indicators. Graph 4 shows the multivariate probabilities with all of the indicators.



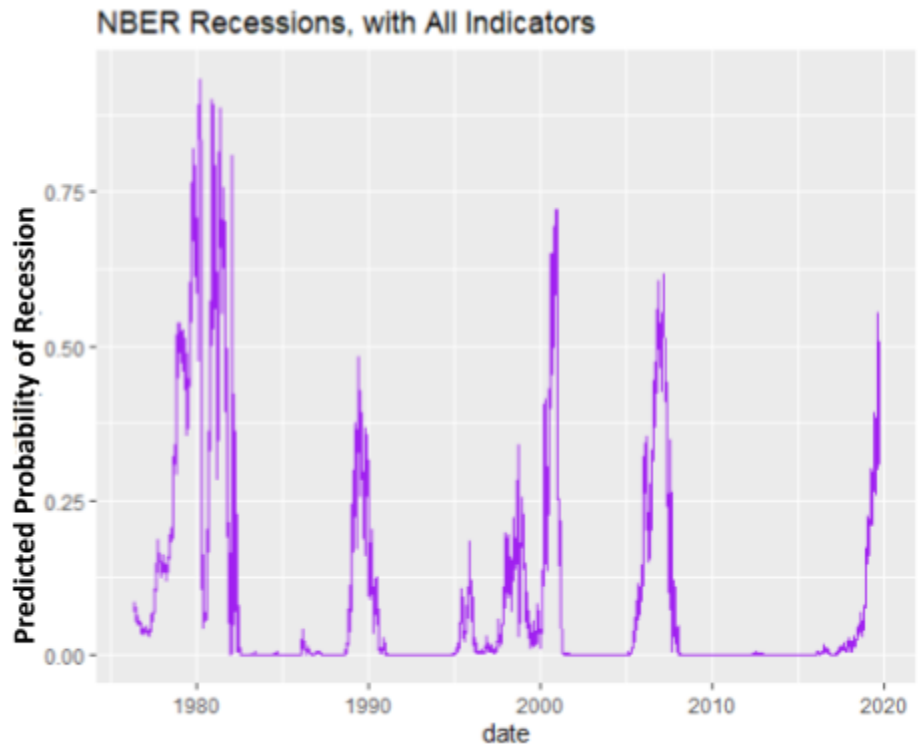
Graph 1



Graph 2

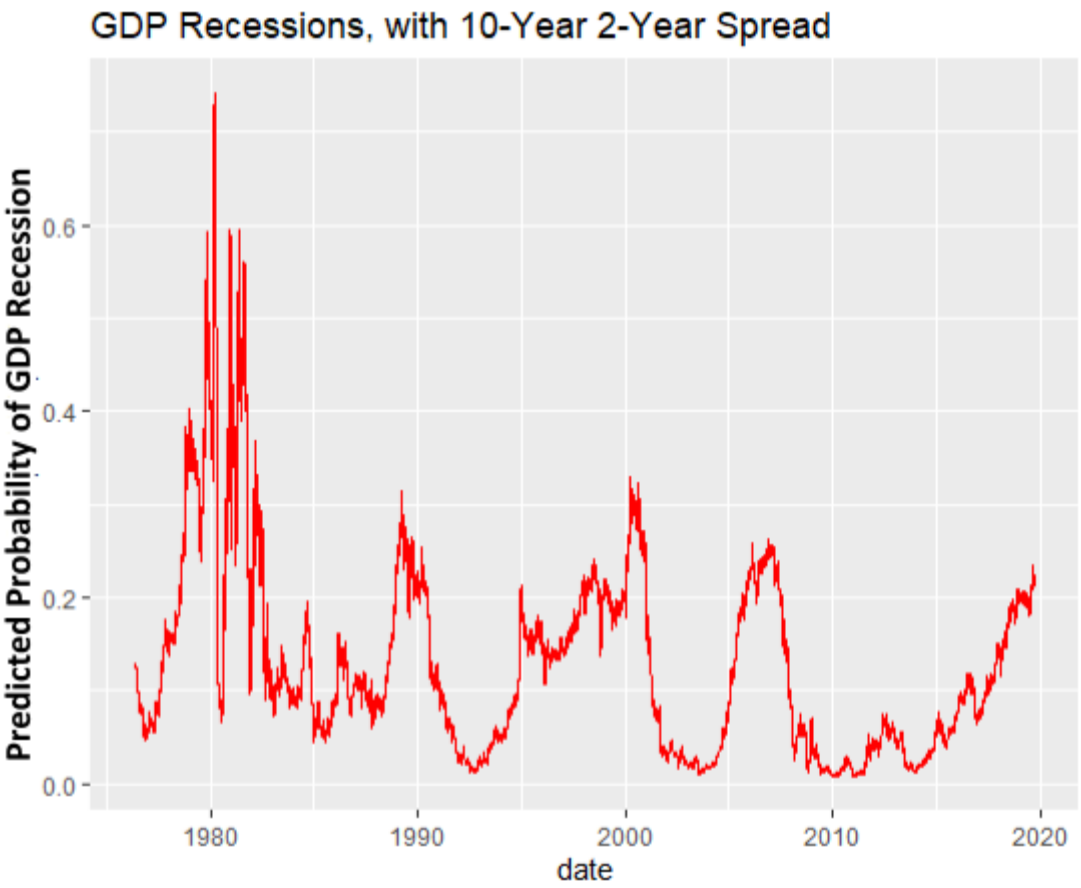


Graph 3

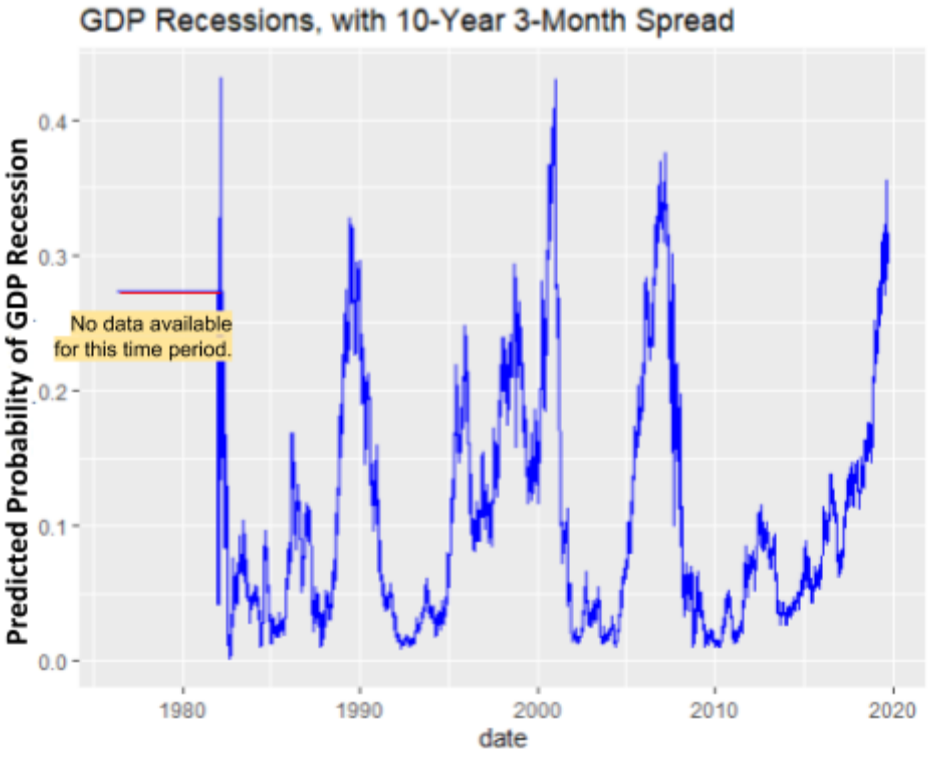


Graph 4

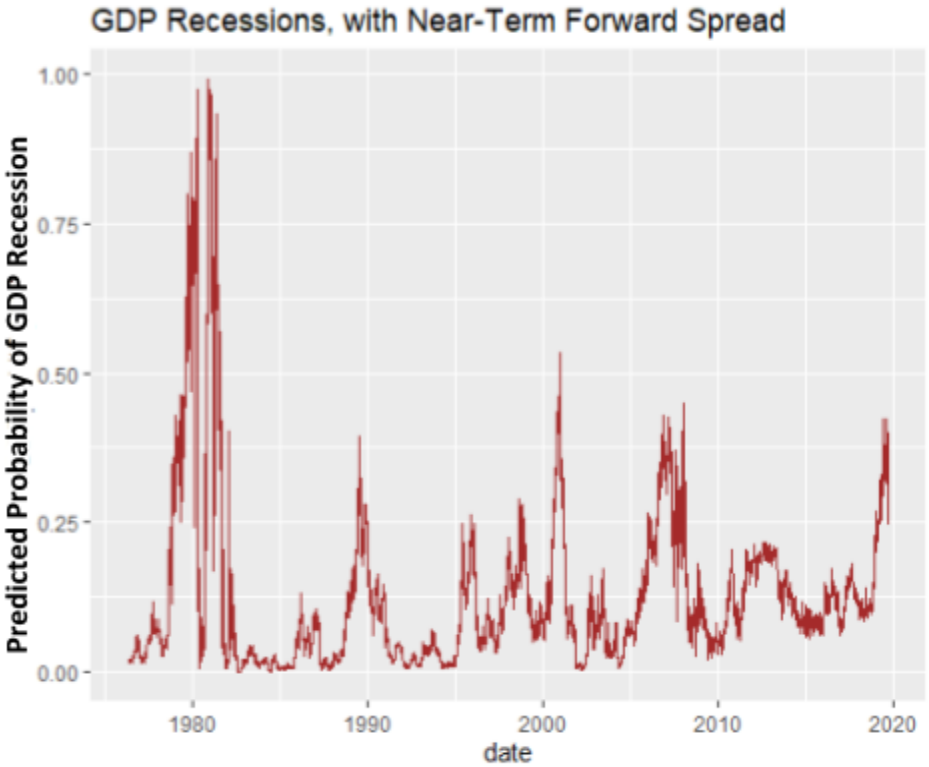
In a similar way, we graphed the probabilities for our other definitions of recession.



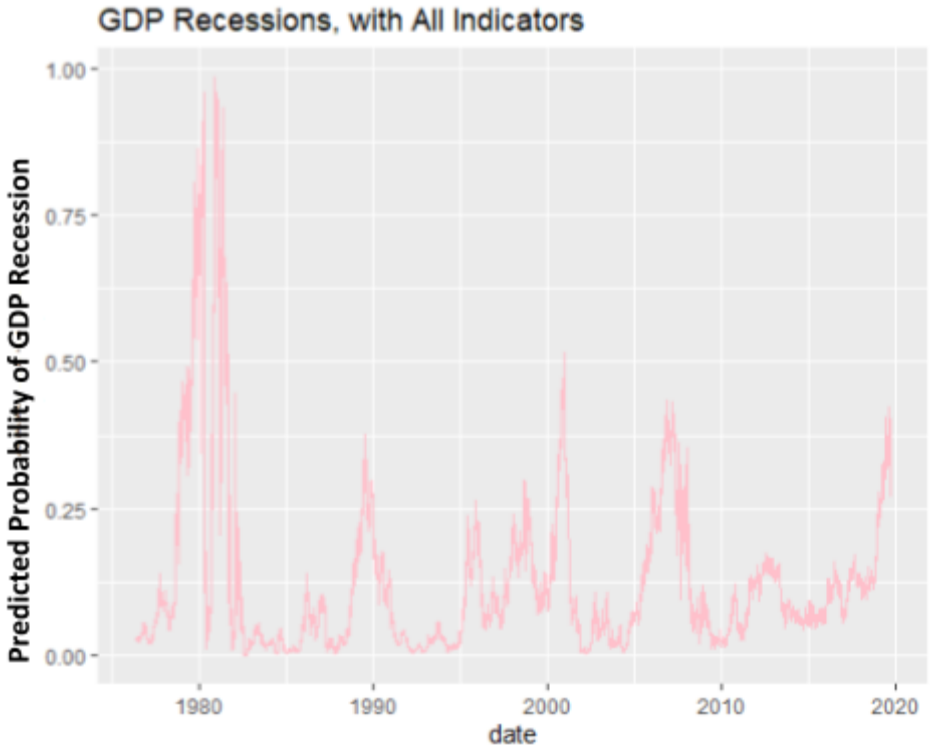
Graph 5



Graph 6



Graph 7



Graph 8

The average marginal effects for each of the indicators in the multivariate models were also calculated. The average marginal effects shows the change in probability that is implied by a one-unit change in the indicator.

Average Marginal Effects			
Indicator	10-Year 2-Year	10-Year 3-Month	Near-Term Forward
NBER Model	.06124	.09067	.01737
GDP Model	.02872	.00409	.09054
Stock Model	.00205	.01114	.06066

Additionally, for the NBER and GDP models, we calculated the percent of recession and non-recession days that were accurately predicted.

Percent Correctly Predicted (All Days)				
Indicator	10-Year 2-Year	10-Year 3-Year	Near-Term Forward	All Indicators
NBER Model	91.56%	90.49%	91.89%	92.06%
GDP Model	87.83%	87.30%	88.82%	89.09%

Discussion of Model Results

The graphs provided above offer a way to visualize the way our model works. We can see that the univariate models tend to be less consistent, and predict lower probabilities, which is to be expected, as the predictions are based on less information compared to the multivariate model. However, we can visually see that the near-term forward spread is peaking generally higher than the other indicators. We can also see that the near-term forward spread model correctly predicts a slightly higher percentage of days. In all the models, we can see a sharp uptick in probability at the end of the timeline. Unfortunately, we were unable to obtain updated data from when we started this project to see if the trajectory of this probability continued upward. However, we can see that some of these models suggest higher probabilities in the most recent quarter than in the quarter preceding the Great Recession of 2008.

Conclusion and Recommendations

At the time of writing this, it is extremely likely that the United States is currently facing a recession as a result of the COVID-19 pandemic. However, it appears as though this global event is not solely responsible for the current state of the economy, but rather served as a catalyst to bring about what these indicators had begun to show.

If real-time data were available for these indicators, we could get an idea of the probability that we are currently in a recession with our model. The model serves as a base structure that would benefit from the addition of more indicators in order to create a more robust prediction. The creation of this model also allows for the discovery of potential new indicators that have not been studied yet.