A Modern Approach to Recession Forecasting

Team #36

Edmund Dale
Kerstin Fontus
Lukas Sledz
Loraine Vizconde

https://github.gatech.edu/MGT-6203-Fall-2023-Canvas/Team-36

ABSTRACT

This study aims to explore alternative predictors for recession, moving beyond the traditionally relied upon yield curve. The predictive models built used different combinations of significant predictors such as Index returns, M-score, yield curve, producer price index (PPI) and weighted average beta. The goal of this study is to shed light on the improvement of recession prediction methodologies, for better preparation and mitigation of economic troubles for both individuals and businesses. We found that a Vector-Autoregressive Model can predict some recessions accurately, but not most of them. A better model to use would be the Vector Error-Correction Model, due to the presence of cointegrated variables.

INTRODUCTION

When is the next recession? This is certainly the question asked after recent aggressive interest rate hikes and persistent inflation. Economists have seen this pattern for years and have come to expect it. But what is a recession? A recession in the United States is defined as a significant decline in economic activity spread across the market lasting more than a few months, normally visible in real Gross Domestic Product (GDP), real income, employment, industrial production, and wholesale retail sales (per the San Francisco Federal Reserve, citing the National Bureau of Economic Research) [6].

Recessions have always been painful to all sectors of the economy. Some view it as a necessary reset, but it still negatively affects individuals and companies. To anticipate these challenges and better prepare for the future, it would be valuable if recession could be predicted.

STATEMENT OF THE PROBLEM AND HYPOTHESIS

The National Association for Business Economics (NABE) is indicating that there are possibly better predictors for a recession than the yield curve [4]. Also, as reported by CNBC, some have a theory that the inverted yield curve is no longer the best measure of an oncoming recession, thus, we ought to find other good or better factors for recession prediction. These predictors are M-Score, stock index returns, PPI and weighted average beta.

Our hypothesis is that adding M-Score, stock index returns, and weighted average beta improves the accuracy of recession forecast models. On the other hand, for our null hypothesis, the combination of these predictors does not enhance the accuracy of the recession prediction models.

KEY VARIABLES AND SOURCES

Real Gross Domestic Product (GDP)

GDP is defined as the total market value of the goods and services produced by a country's economy. This is the response variable of our models. We define the beginning of a recession as when GDP has decreased for 2 consecutive quarters, and the end of a recession when GDP has increased for 2 consecutive quarters; so accurately forecasting GDP is how we will predict an oncoming recession.

M-Score

The Beneish M-Score is associated with the probability that a company is engaged in financial misreporting [1]. M-Score catches fraud in corporate earnings reports using eight financial ratios.

Specifically, if a company's M-Score is between –2.22 and -1.78 then they are classified to be likely manipulators of their reports, and if it rises above –1.78 then they are classified as manipulators. A recent paper finds that an aggregate M-Score predicts recessions five to eight quarters ahead of time because it predicts lower real investment one to four quarters ahead. We believe this variable will be an important addition to our model. The 12 variables used to calculate the required ratios are Current Assets; Total Assets; Cost of Goods Sold; Total Long-Term Debt; Depreciation; Cash Flow from Operations; Income from Operations; Current Liabilities; Property Plant Equipment; Net Receivables; Net Sales; and Sales, General, and Administration Expenses. This data came from Compustat, a comprehensive market and corporate financial database published by Standard and Poor's via Wharton Research Data Services (WRDS).

Yield Curve

This variable is defined as the 10-Year Treasury Constant Maturity Minus 2-Year Treasury Constant Maturity. This is a classic indicator associated with recession.

Producer Price Index (PPI)

This variable measures the average change in domestic selling prices over time. PPI is a leading indicator of inflation and has been found to be correlated with recessions [7].

Market Index Returns

Stock market returns are an indicator of economic performance and investor confidence. The market returns were calculated to use it in determining the correlation between GDP and the indexes. We derived stock market return data from two of the biggest indexes of the US stock market: S&P 500 and Dow Jones. Both data sets were queried from Bloomberg databases. We used the historical data of adjusted closing prices to calculate the market returns.

Weighted Average Beta and Market Capitalization

We also investigated the relationships between GDP and aggregated beta values (weighted by market capitalization). We used time series beta values for thousands of companies which were aggregated by date and multiplied by weights (based on market capitalization), to create weighted average data features. This data also comes from WRDS, specifically “Better Market Betas” (data sets by individual contributors) and the Center Research Security Prices (CRSP) annual update stock prices.

Macroeconomic Indicators

Federal Reserve Economic Data (FRED) is an online database with economic time series data from national, international, public, and private sources. It is maintained by the Research Department at the Federal Reserve Bank of St. Louis. We used it as a source for Historical GDP, PPI, and yield curve data.

Data Cleaning and Feature Engineering

Data Exploration and Cleaning

After loading our data sets as .csv files, we imported them in Jupyter Notebooks as Pandas data frames, and then performed cleaning and feature engineering. There are currently 15 of these notebooks available in the "Code" directory for our GitHub project.

Transforming Raw Data

We extracted monthly and quarterly financial information for thousands of companies from different WRDS databases. Since public companies enter and exit the market at different times, and their financials are noisy and difficult to predict; we decided to use the average beta (based on market capitalization) at each point in time, without regard for which companies are in the mix at each date. This allowed us to sidestep the messy problems of data imputation, and inconsistent timelines for each company for this data. We also excluded the few companies with an extremely large market cap to avoid skewing the weighted average.

M-Score values are not readily available, so they were calculated quarterly from publicly available, self-reported financial firm data. Cleaning the data before calculating the M-Scores included excluding companies with mostly incomplete reports and imputing values when appropriate. Calculating the M-Scores resulted in some very large and very small values, so they were bounded or imputed depending on their situation. Creating the quarterly weighted M-Score required matching the companies we had M-Scores for to their market cap. This proved challenging since Compustat uses its own unique keys to identify companies and lacked identifying information like tickers and names for many data points. Any companies that could not be identified had their market cap imputed as that quarter's average market cap. Finally, we computed the quarterly arithmetically averaged M-Score and the weighted average (by market cap).

Creating Trend, Velocity and Acceleration features

For many of our original features, we wanted to examine the first and second derivative (in time) to see if they "lag" GDP. For example, does the change in S&P 500 returns (velocity) or rate of their changes (acceleration) predict GDP change? To create these velocity and acceleration features, we first created a "trend" feature by decomposing each original feature assuming quarterly seasonality (period = 90 for daily data, period = 3 for monthly data, or period = 1 for already quarterly data) and extracting the trend. These decompositions were done with the tsa.seasonal_decompose function from the Python statsmodels.tsa.api library:

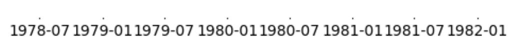


***Fig 1**: Finding the smoothed "trend" component of yield curve data assuming quarterly seasonality.*

Next, we took the gradient of each trend to generate velocity features, then took the gradient again to create acceleration features.

Converting to Quarterly Timescales

Since our real GDP data comes in financial quarter units, we decided to only look at data points occurring at the end of each quarter. We used the pandas.PeriodIndex function to find the date for each quarter, then aggregated by quarter and only selected points with the date closest to the end of each quarter. This method leads to some information loss, but it allows each time series feature to be included in vector autoregressive (VAR) models, and we believe the important long-term trends in each feature are preserved.

MODEL BUILDING

Why We Used a VAR Model

In a VAR model, time series variables that lag a target variable are identified and used to forecast values of the target. Our research suggested that VAR models are commonly employed when forecasting GDP, recessions, and other macroeconomic variables [6]. In addition, there is clear evidence that time series variables such as yield curve are leading indicators of recession (they "lag" recession), so a VAR model seemed like a natural choice.

Stationary Time Series Data

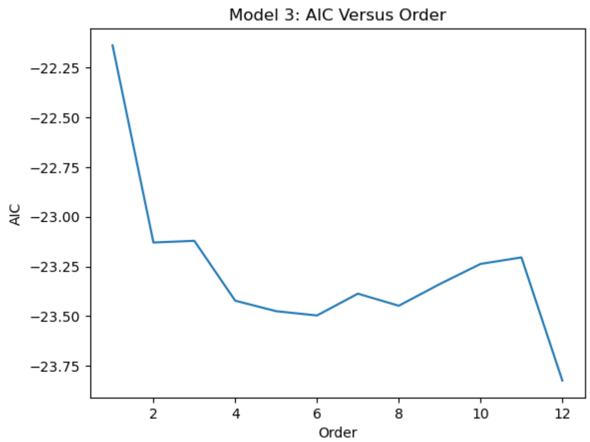
One requirement of using a VAR model is having stationary time series data, meaning that none of the variables' mean or variance change over time. We ran the Augmented Dickey-Fuller Test which determined that many of our datasets are non-stationary. To resolve this issue, we differenced every variable, then ran the ADF test again confirming all transformed variables were stationary, then we used different train/test splits of the differenced data to build our models.

Causality of Variables

The goal of this model is to predict GDP, so we needed to find the independent variables that lag it. We used the Granger Causality Test on the differenced data to study this, with a maximum lag value of twelve quarters, or three years. The variables that resulted in a p-value less than or equal to 0.05 are determined to be lagging GDP. It turns out that our data set that runs from 1976 to 2023 had only three variables that lagged GDP compared to the twenty from the data set running until 2013. We hypothesized that this was due to the non-traditional recession during the 2020 COVID pandemic. There were likely no economic indicators that this recession was coming. Running the Granger Causality test on the same data set limited to data points before 2020 revealed more variables that lagged GDP. As a result, we decided to build most of our models using only the data set ending in 2013, which contained all of our time series variables including M-Score, weighted average beta, and average MCAP.

Choosing Order of the Model

Creating a VAR model requires choosing a lag order. Forecasts are highly dependent on lag order so it must be selected with care. We calculated AIC values for models using orders 1 through 12 and



discovered that a lag order of 5 (5 quarters) is near the first local minimum in AIC (left). A recent paper also claims that M-Score lags the start of a recession by 5 to 8 quarters [1]. Therefore, we decided to stick with a lag order of 5 for each of our models.

Fig 2*: Model Order vs. Model AIC*

Train and Test Split

We chose to start by building a model that predicts one recession, the most recent one before 2013 being the Great Recession which started in December of 2007. This recession has one of the most dramatic drops in GDP of all the different recessions, so it makes a good baseline for our model building. Therefore, we split our test data at the end of the third quarter of 2007. Later on, we used different train test splits to forecast other recessions.

Variable Selection

We obtained the first set of features used to build our model from the Granger Causality Test step. They were average weighted M-Score; average M-Score velocity and acceleration; average weighted M-Score velocity and acceleration; S&P 500 return trend and velocity; Dow Jones (Indu) return trend and velocity; the PPI Curve, its trend, velocity, and acceleration; the yield curve and its velocity; weighted average beta, its trend, velocity, and acceleration; and average market cap velocity.

This first model built (called model1) made terribly inaccurate forecasts, prompting us to explore our variables further. We calculated the variance inflator factor (VIF) of each variable to check for any multicollinearity among them. We also used the Durbin Watson Statistic to check if there was any correlation in the residuals, meaning that there are some patterns in the time series variables that are not explained by the model. For this test, a result closer to 0 indicates positive serial correlation, while a result closer to 4 indicates negative serial correlation, with 2 being the optimal result. We used a combination of these two results to reduce the number of variables in the next version of our model.

The resulting variables in this more accurate model (called model2) were average weighted M-Score; average weighted M-Score velocity and acceleration; Dow Jones return trend and velocity; the PPI Curve's trend, velocity, and acceleration; the yield curve and its velocity; weighted average beta trend, velocity, and acceleration; and average market cap velocity.

Next, we wanted to see if it was possible or necessary to reduce our list of variables even further. If time series variables are cointegrated, then they have a long running statistical relationship with each other,

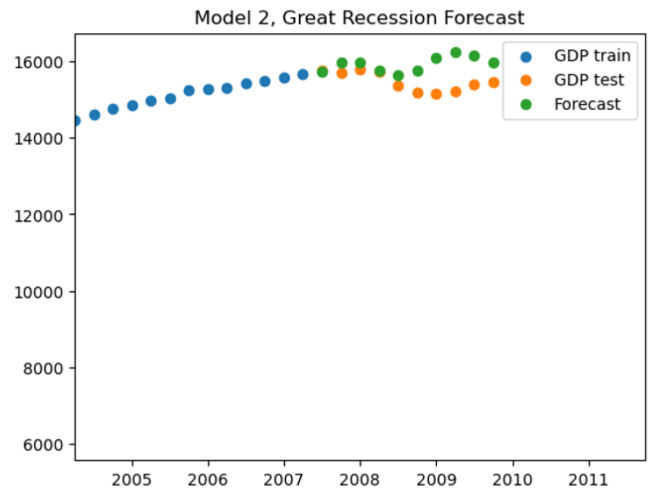


***Fig 3:** Forecast of 2008 Great Recession after 1st round of variable selection (model 2).*

and a vector error correction model (VECM) may be more appropriate to use than a VAR model. We performed the Engle Granger Two Step Cointegration on every combination of variables, which revealed 5 variables that were all cointegrated with each other (average weighted M-Score velocity, Dow Jones return trend and velocity, and PPI velocity and acceleration, and yield curve velocity). After some empirical testing, the most accurate model (called model3) used all of these velocity variables. Studying the variables used to build model3 shows consistent behavior across the different recessions we studied. They all have periods of inclines and declines either before or during these time periods (as shown in figure 4):

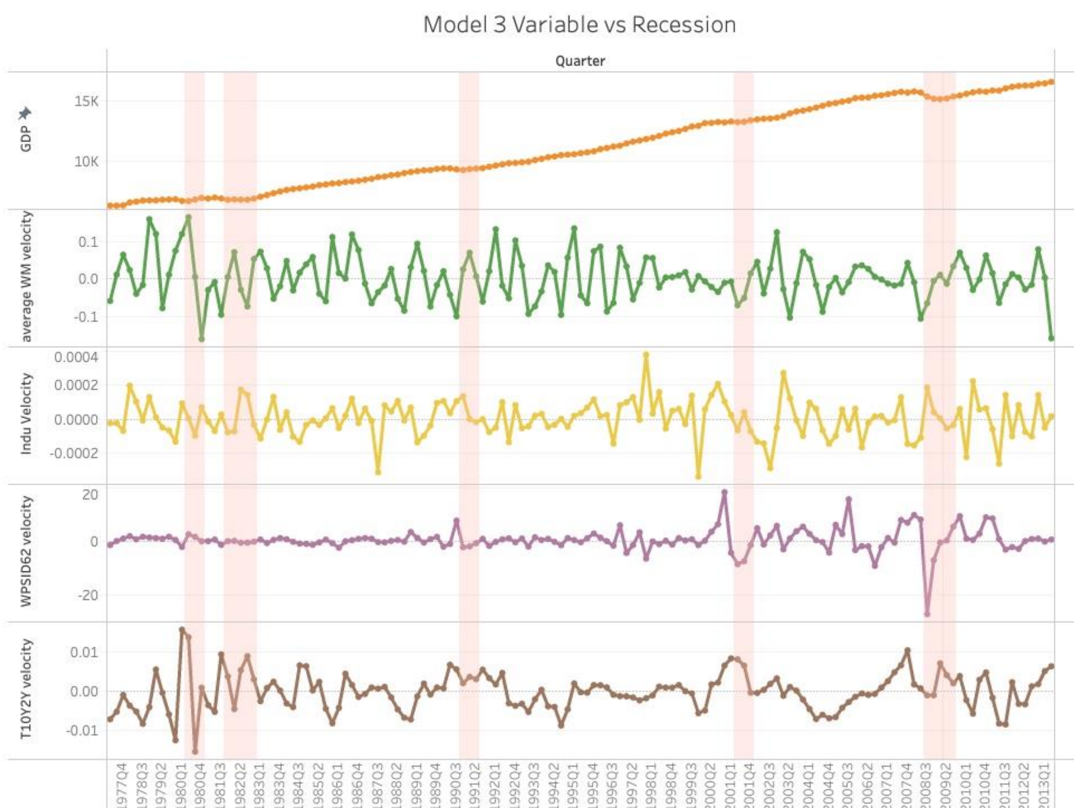


***Fig 4**: Values of Model 3 variables over time (pink bars indicate recession years).*

In addition to model 3 we created one more model (model4) which is based on a different dataset that goes up to 2023. We used model 4 to forecast the 2020 "covid-19 recession" as well as forecast GDP beyond 2023:

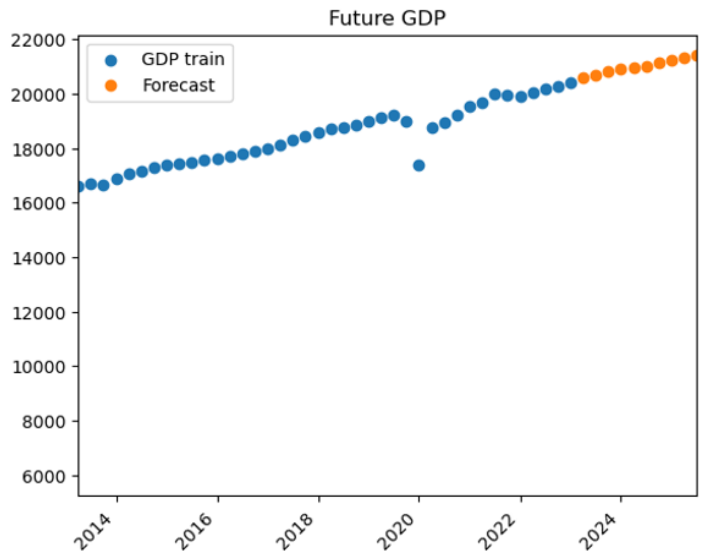


Fig 5: *Forecast of GDP past 2023 using model 3*

Model Evaluation

We evaluated the accuracy of forecasts of recessions from 1982 to 2020 for all 4 of our VAR models (model4 was used to forecast the 2020 recession). Overall, model3 Forecasts had the lowest RMSE and the highest correlation coefficients with actual GDP data (see figure 6). Interestingly, model3 is completely composed of "velocity" variables which may mean that the rate of change in recession predictors is at least as important as the variable values themselves.

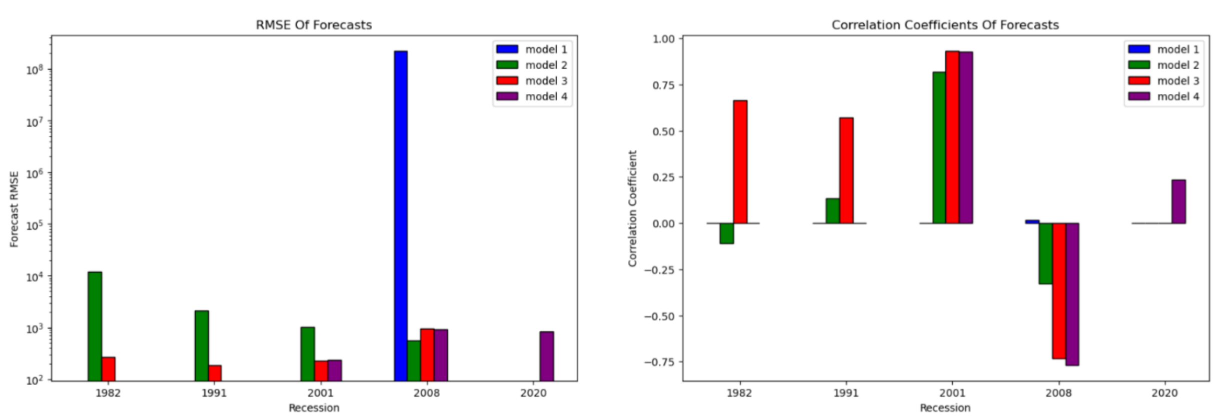


Fig 6: *(left) RMSE for forecasts of recessions from 1982 – 2020. (right) correlation coefficients between forecasts and actual data at each recession.*

We plotted all of the forecasts made with model 3 alongside the actual GDP data with highlights at each recession (pink bars). The plot below shows that our best model was only able to accurately predict the 1982 recession:

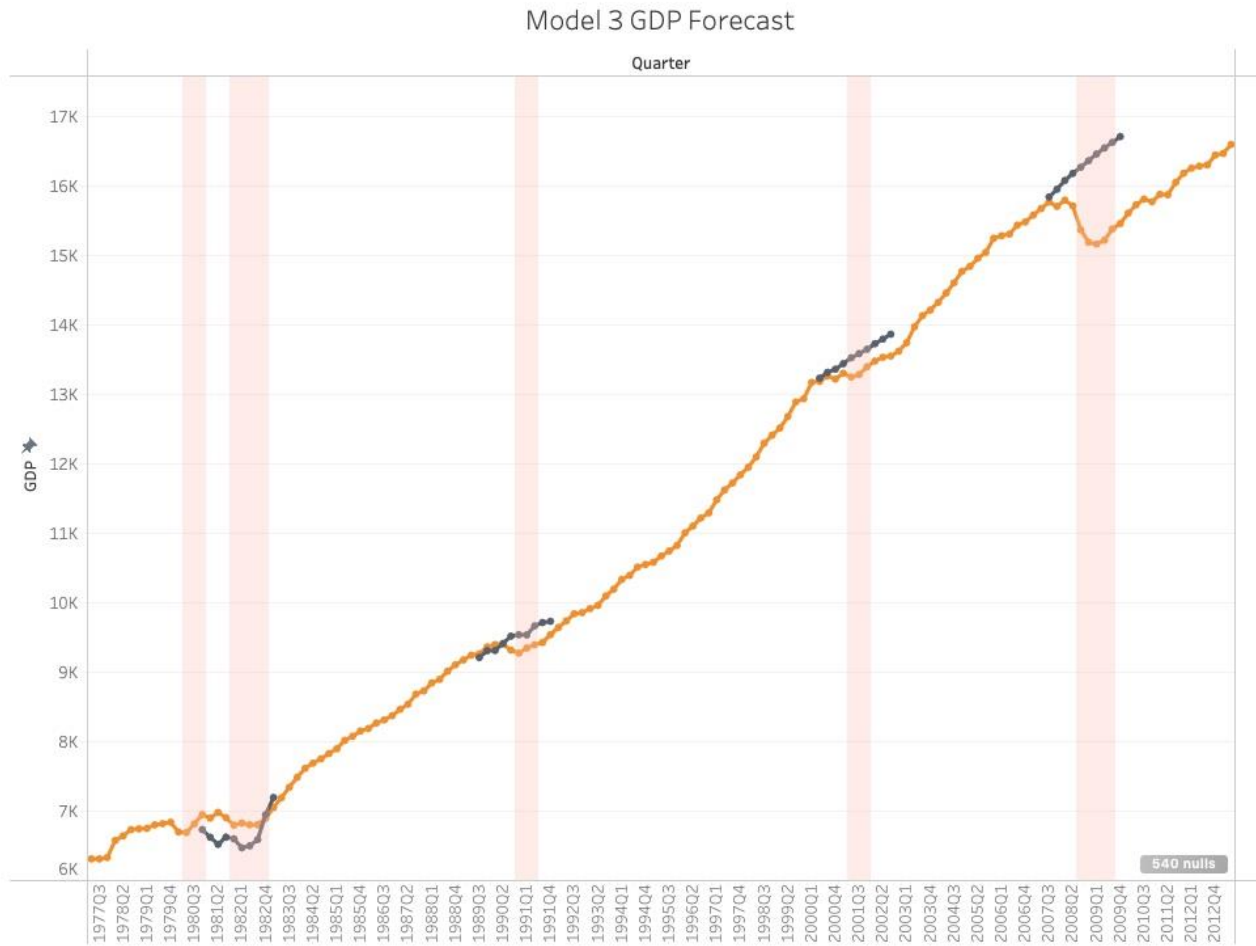


Fig 7: *Real and Forecasted GDP vs. Recessions*

CONCLUSION

The question of when the next recession will occur remains a vital concern. The exploration of alternative predictors aims to foster better preparation and mitigation strategies. The findings of our study indicated that a Vector-Autoregressive model can exhibit some level of accuracy in predicting some recessions, but it fell short in providing consistent predictions. Given more time, we would undertake an in–depth look using Vector Error-Correction model, as the presence of cointegrated variables was apparent in our findings. Our initial hypothesis of using these variables – M-score, Index returns, PPI, and weighted average beta – in increasing the accuracy of forecasting a slowdown in economic activity was not supported by our findings. We were only able to predict this type of activity during the 1981-1982 recession, thus our models yielded insufficient evidence to reject our null hypothesis (that adding modern variables like M-score does not improve accuracy of recession forecasts), indicating that the combination of these predictors does not enhance the accuracy of the recession prediction models.

In conclusion, our research contributes to important understandings of recession prediction methodologies, emphasizing the significance of critically evaluating the combination of alternative predictors.

Works Cited

1. Beneish, Messod D; Farber, David B.; Glendening, Matthew; Shaw, Kenneth W. "Aggregate Financial Misreporting and the Predictability of U.S. Recessions and GDP Growth." *The Accounting Review* 1 September 2023; 98 (5): 129–159. https://doi.org/10.2308/TAR-2021-0160
2. Estrella, Arturo; Mishkin, Frederic S. "The Yield Curve as a Predictor f U.S. Recessions." *Current Issues in Economics & Finance.* Volume 2, Number 7. Federal Reserve Bank of New York. June 1996. https://www.newyorkfed.org/medialibrary/media/research/current_issues/ci2-7.pdf
3. Franz, Jared. Spence, Darrell. "Guide to Recessions: 9 things you need to know." *Capital Group.* 7 Sept. 2023. https://www.capitalgroup.com/advisor/insights/articles/guide-to-recessions.html
4. Konish, Lorie. "This Recession Indicator May Be a 'broken Odometer for the Economy,' ExpertSays. Where Forecasts Stand Now." *CNBC.* CNBC. 5 Sept. 2023. www.cnbc.com/2023/09/05/recession-indicator-may-be-broken-odometer-for-economy-says-expert.html
5. Prabhakaran, Selva. "Vector Autoregression (VAR) - Comprehensive Guide with Examples in Python." Machine Learning Plus. https://www.machinelearningplus.com/time-series/vector-autoregression-examples-python/#google_vignette
6. Unknown. "What is the difference between a recession and a depression?". *Federal Reserve Bank of San Francisco.* Federal Reserve Bank of San Francisco. February 2007. https://www.frbsf.org/education/publications/doctor-econ/2007/february/recession-depression-difference/.
7. Vrontos, Spyridon D.; Galakis, John; Vrontos, Ioannis D. "Modeling and predicting U.S. recessions using machine learning techniques." *International Journal of Forecasting*. Volume 37, Issue 2. 2021. Pages 647-671. ISSN 0169-2070. https://doi.org/10.1016/j.ijforecast.2020.08.005

Data References

- Final Data Set: https://www.dropbox.com/scl/fi/umx2vjq4a1ppstb2kh1yv/merged_data_mscore_wab20231031125447.csv?rlkey=jmlquopgzrbf9ggtgenja0lmh&dl=0
- Center Research Security Prices (CRSP), LLC. (1975 – 2023). CRSP Stock (Quarterly). Wharton Research Data Service. https://wrds-www.wharton.upenn.edu/data-dictionary/crsp_q_stock/.
- Federal Reserve Bank of St. Louis, 10-Year Treasury Constant Maturity Minus 2-Year Treasury Constant Maturity [T10Y2Y] (1976 – 2023), retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/T10Y2Y, November 2, 2023.
- Individual Contributors. (1926-2019). Better Market Betas. Wharton Research Data Service. https://wrds-www.wharton.upenn.edu/data-dictionary/contrib_general/
- S&P Global Market Intelligence. (1968-2013). Compustat Snapshot – As First Reported - Quarterly. Wharton Research Data Service. https://wrds-www.wharton.upenn.edu/data-dictionary/compsamp_snapshot/wrds_csq_unrestated/.

- U.S. Bureau of Economic Analysis, Real Gross Domestic Product [GDPC1] (1975 – 2023), retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/GDPC1, November 2, 2023.
- U.S. Bureau of Labor Statistics, Producer Price Index by Commodity: Intermediate Demand by Commodity Type: Unprocessed Goods for Intermediate Demand [WPSID62] (1975 – 2023), retrieved from FRED, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/WPSID62, November 2, 2023.