Driving into the Unknown: Mapping and Predicting the S&P 500 in 2023

MGT 6203 - Spring 2023 Group Project - Team 24 Final Report

Alexander Widemeyer awidmeyer3@gatech.edu
Daniel King dking84@gatech.edu
Isabel Toledano itoledano3@gatech.edu
Pawel Madon pmadon@gatech.edu
Qiusi Wang qiusi.wang@gatech.edu

Abstract

This project analyzed the various factors influencing the S&P 500 index, a crucial indicator of the United States (US) equity markets. By studying the historical data, we discovered that the interest rate, inverted yield curve, and Federal Reserve's policy actions are effective leading indicators to forecast an upcoming bear market. Our model predicts that the S&P 500 index could suffer a significant loss in 2023. In response, we explored alternative investment options and offered advice on constructing a portfolio in the current economic environment.

I. Introduction/Background/Motivation

Project Introduction

In 2022, the United States stock market experienced its worst downturn since the 2008 financial crisis. The Dow Jones Index registered a decrease of approximately 8.8%, while the S&P 500 and the technology-laden Nasdaq plummeted by 19.4% and 33.1%, respectively. Conversely, interest rates are rising, and certain banks are currently offering a 5% annual percentage yield on 1-year Certificate of Deposit. This raises a pressing question for individual investors in the US: in 2023, should they continue to bet.

By investigating the fundamental drivers of the S&P 500 index using historical data, this project aims to use a regression analysis to forecast whether it will generate a positive return in 2023 and offer our targeted audience, individual investors in the US, investment advice on whether and how to invest their money in the US equity markets in 2023 and options to hedge against inflation and mitigate market risk in a time of great uncertainty and interest rate risk. Therefore, we have the following initial hypotheses:

- United States Gross Domestic Product (GDP), Inflation (represented by Consumer Price Index (CPI)) and interest rate level are the key drivers of the S&P 500 index return;
- S&P 500 index will not have a positive return in 2023; and
- Gold or silver are not as effective as US Treasuries securities or TIPS ETF in hedging the current high inflation.

Existing Methodology Introduction

When modeling the US stock market index, it is common practice to begin with the most influential domestic indicators. However, C. Liu's (2016) previous research [1] uncovered that global equity markets and foreign exchange market factors can also have a significant impact on the US equity market. These factors include indices such as the Financial Times Stock Exchange 100 Index (FTSE 100), Nikkei Stock Average 225 Index (NIKKEI 225), Shanghai Stock Exchange Composite Index (SSE), as well as exchange rate pairs like USD/CNY, USD/JPY, and USD/GBP.

In addition to these indicators, the yield curve is another important factor in predicting economic activity. The yield curve is a graphical representation of the relationship between the yield on government bonds of different maturities. A. Estrella and F. S. Mishkin (1996) [2] argued that strong empirical evidence suggests the yield curve is a leading predictor of economic activity in the United States and Europe, with an inverted yield curve (where short-term interest rates are higher than long-term rates) typically signaling an impending recession.

Ligita Gasparėnienė et al. (2021) [3] examined the application of machine learning techniques for predicting the S&P 500 index price based on U.S. economic indicators. The study finds that machine learning models outperform the traditional time series analysis in terms of predictive accuracy. The results show that Support Vector Regression and Random Forest models are particularly effective in predicting S&P 500 index prices.

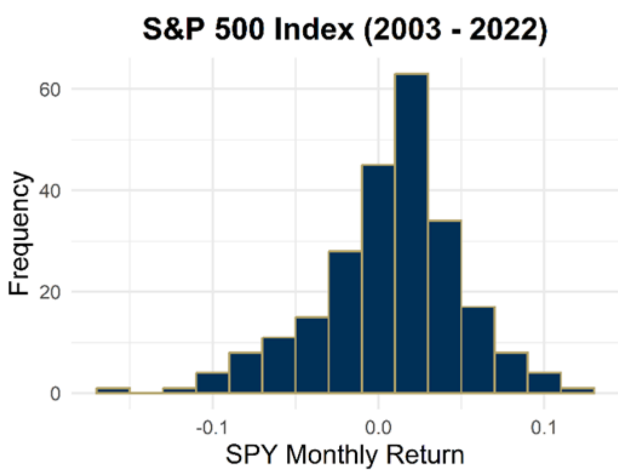
II. Overview of Data

Data Source/Data Cleaning/Exploratory Analysis

The data used in this project includes S&P 500 index history and a number of economic factors such as US GDP, Interest Rate, Inflation. They were obtained from two widely used financial data sources: Yahoo Finance https://finance.yahoo.com/ and Federal Reserve Economic Data (FRED) database https://fred.stlouisfed.org. We access Yahoo Finance data and FRED data using the API provided via R library "tidyquant" and "fredr" respectively.

For exploring different modeling and variable options: we gathered monthly data from 2003 to 2022 to cover the two economic cycles that include the 2008 financial crisis (December 2007 - June 2009) and the COVID-19 recession (from February 2020 to April 2020). Per our review, there are no missing data points in the data we obtained from the above data sources.

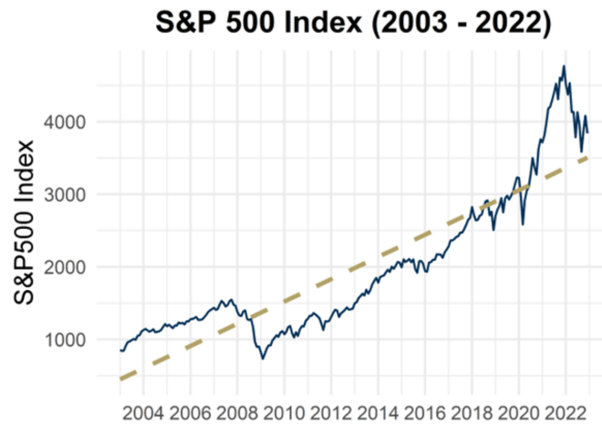
According to the histogram chart below, the monthly returns of the S&P 500 index exhibit a generally normal distribution. However, there is a leftward skew, suggesting that occasional extremely large losses may occur, which is the problem that this project aims to address: whether there is a possibility of significant losses occurring in 2023 and whether investors should take this into account when making investment decisions.



Our final model utilizes a Boolean response variable, the "Bear-Market-Flag," which takes a value of 1 to indicate a bear market and a value of 0 to indicate a non-bear market environment for the S&P 500 index. We derived this variable through a drawdown analysis that identified top drawdowns of the index exceeding 20%. Based on this criterion, we designated the following periods as bear market periods:

- October 2007 to March 2009
- January 2020 to March 2020, and
- June 2022 to December 2022.

These designations align with the movement of the S&P 500 index, as illustrated in the accompanying chart below.



The predictors of our final model include a set of interest variables and variables derived from them as well as Federal Reserve's actions on raising or lowering its target policy interest rate. The detail of each variable is included in **"Section III Overview of the Final Model".**

III. Overview of the Final Model

Final Model Description

The final model we selected for this project is a logistic regression model that can be used to predict whether a bear market is coming. The regression formula is as below

Logit (Odds of Bear Market) =
  $\beta1$* (Fed Fund Rate level with a 15-month lag)
+ $\beta2$* (Boolean Indicator of Negative Spread between 1-year Treasury Yield and Fed Fund Rate)
+ $\beta3$ * (Boolean Indicator whether Federal Reserve is lowing its target policy rate)
+ Intercept

Model Selection

In order to address the question of the S&P 500 index performance, we aimed to predict whether the US equity market, as represented by the index, will enter a bear market. We opted for logistic regression to model the bear market indicator, since a boolean variable is the response variable. Linear regression is not an appropriate option for this type of variable.

Feature engineering

The exploration of using interest rate variables to predict the S&P 500 index performance was inspired by a publication by Federal Reserve of New York economists Arturo Estrella and Frederic S. Mishkin titled “The Yield Curve as a Predictor of the U.S. Recessions.” [2]. Their research suggests that there is a relationship between the slope of the US Treasury yield curve and whether there will be a recession within the next few quarters.

Our modeling process takes into account the widely used benchmark interest rates in the United States, such as the effective Fed Fund rate and US Treasury yields with varying maturities. We also consider the impact of the Federal Reserve's policy actions. To do so, we have developed two Boolean variables:

- A Boolean variable “falling rate” indicating the Federal Reserve takes actions to lower its target policy rate. In general, when the Federal Reserve lowers interest rates, it makes borrowing cheaper, which can stimulate economic growth and increase the attractiveness of stocks as an investment. This can lead to a rise in stock prices. The source of this information is from the Federal Reserve’s press release.
- A Boolean variable “rising rate” indicating the Federal Reserve takes actions to raise its target policy rate. For the periods that the Federal Reserve does not increase or decrease its policy rate, both “rising rate” and “falling rate” variables will be zero.

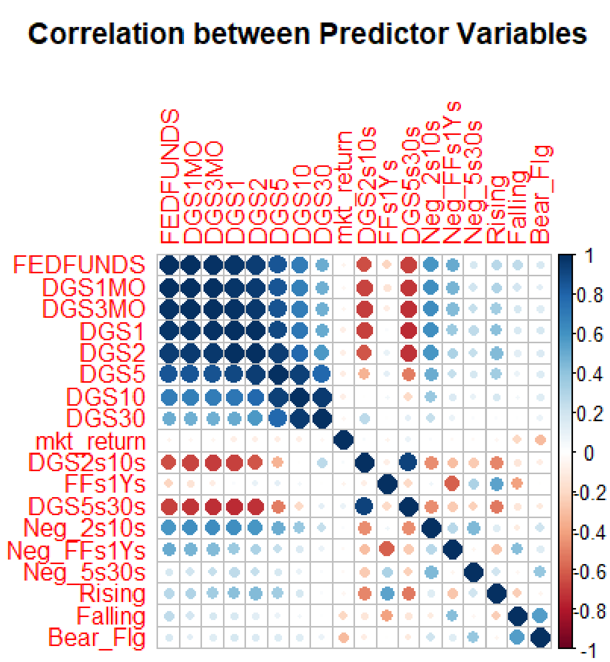
In order to explore the relationship between these yield curves as suggested by Arturo Estrella and Frederic S. Mishkin’s paper [2], we created an additional two predictors:

- A Boolean variable “negative spread between 10-year and 2-year US Treasury yield”. If this variable has a value of one, it indicates an unusual situation where a 2-year Treasury bond has a higher yield than a 10-year Treasury bond, as longer-term bonds typically have higher yields to compensate investors for the greater risks. This often suggests that investors have a pessimistic outlook for future economic growth and inflation.
- A Boolean variable “negative spread between 1-year US Treasury yield and effective fed fund rate”. Similar to the above, this variable indicates an inverted yield curve where the shorter-term rate - effective fed fund rate (overnight rate) - is higher than the longer-term 1-year US Treasury yield.

In sum, we have total 12 predictors to select from

Eight Interest variables	• Effective Fed Fund rate • 1-month treasury yield • 3-month treasury yield • 1-year treasury yield • 2-year treasury yield • 5-year treasury yield • 10-year treasury yield • 30-year treasury yield
Four Derived variables	• Negative spread between 10-year and 2-year treasury yield • Negative spread between 1-year treasury yield and effective fed fund rate • Federal Reserve action to raise its target policy rate • Federal Reserve action to lower its target policy rate

We generated a correlation plot for the variables mentioned above, which revealed some noteworthy findings. Firstly, the interest rate variables (fed fund rate and treasury yields) were found to be highly correlated with each other, highlighting the need for careful consideration when selecting variables to avoid overfitting. Additionally, the bear market indicator was observed to have a strong correlation with the "falling" policy rate indicator, indicating a potential relationship between them.



Different lag structures

It is worth noting that interest rates are typically expected to have a delayed impact on stock market performance. This presents an opportunity for us to investigate whether any interest rates can serve as effective leading indicators for an impending bear market. We created six lag structures for all the variables and derived variables above: 3 months, 6 months, 9 months, 12 months, 15 months and 18 months. For example, a 3-month lag structure means we are attempting to find the relationship between predictor values at month 0 and bear market flag at month 3.

Stepwise Regression

To determine the optimal lag structure with the greatest fit and explanatory power, we conducted stepwise regression for each structure. For the regression formula with the lowest AIC within each structure, we evaluated the P-values of the predictors. The resulting outcomes are presented below:

Lag Structure	Best AIC	Statistical Significance of Predictors
No lag	Poor goodness of fit	
3-month Lag	Poor goodness of fit	
6-month Lag	Poor goodness of fit	
9-month Lag	79	7 out of 10 variables are significant at 95% confidence level. Intercept is not significant.
12-month Lag	54	6 out of 7 variables are significant at 95% confidence level. Intercept is significant.
15-month Lag	61	8 out of 8 variables are significant at 95% confidence level. Intercept is significant
18-month Lag	70	5 out of 6 variables are significant at 95% confidence level. Intercept is not significant.

The 15-month lag structure produces the best outcome after considering both AIC and P-values. The stepwise regression suggests the following seven predictors for regressing a bear market indicator under a 15-month lag structure:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	8.590	4.230	2.031	0.04227 *
Fed Fund	4.936	2.498	1.976	0.04811 *
3M TSY	-6.794	2.802	-2.425	0.01531 *
10Y TSY	19.410	5.936	3.270	0.00108 **
30Y TSY	-19.097	6.000	-3.183	0.00146 **
Negative Spread (1Y TSY - FF)	4.147	1.280	3.239	0.00120 **
Rising Policy Rate	-4.119	1.631	-2.525	0.01157 *
Falling Policy Rate	-4.160	1.438	-2.894	0.00381 **

Null deviance: 156.974 on 224 degrees of freedom
Residual deviance: 45.352 on 217 degrees of freedom
AIC: 61.352

Further feature selection

Among the seven variables selected by the stepwise regression analysis above, we further eliminated four of them for the following reasons:

- Rising Policy Rate was eliminated because its coefficient sign was contrary to business intuition. Although it was expected that a rising policy rate would increase the likelihood of a bear market, the coefficient indicated the opposite effect.
- We observed that the effective fed fund rate and three treasury yields were statistically significant at a 95% confidence level in the best regression model picked by stepwise regression. However, as previously noted, these variables are highly correlated, and including all of them can lead to overfitting. Therefore, we needed to include only one of them in the regression formula. We performed further regression using each of the four interest rate variables along with negative spread and falling policy rate, and found that the effective fed fund rate produced the best fitting. Consequently, we eliminated the three treasury yield variables.

Choice of level interest rate variable	AIC	Statistical Significance of Predictors
Effective fed fund rate	79	At 95% confidence level, 3 out of 3 variables are significant. Intercept is significant.
3-Month Treasury Yield	79	At 95% confidence level, 2 out of 3 variables are significant as falling policy rate is not significant. Intercept is significant.
10-Year Treasury Yield	85	At 95% confidence level, 2 out of 3 variables are significant as falling policy rate is not significant. Intercept is

		significant.
30-Year Treasury Yield	89	At 95% confidence level, 1 out of 3 variables is significant as falling policy rate and 30-year Treasury yields are not significant. Intercept is significant.

Final Model Outcome

Therefore, our final model has three predictors: effective fed fund rate, negative spread between 1-year treasury yield and effective fed fund rate, and the falling Federal Reserve target policy rate. The logistic regression using these three predictors under 15-month lagging structure has the following results:

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.1764	-0.1798	-0.13	-0.1275	3.1058

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-4.8607	0.6938	-7.006	2.45e-12 ***
Fed Fund	0.5724	0.1769	3.237	0.00121 **
Negative Spread (1Y TSY - FF)	4.1313	0.777	5.317	1.05e-07 ***
Falling Policy Rate	-1.7528	0.7829	-2.239	0.02516 *

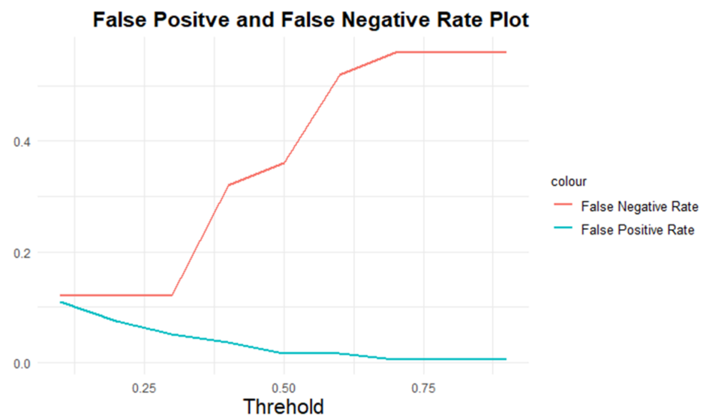
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null deviance: 156.974 on 224 degrees of freedom
Residual deviance: 70.878 on 221 degrees of freedom
AIC: 78.878

All the coefficients align with business intuition. The effective fed fund rate and negative spread between 1-year treasury yield and effective fed fund rate both have a positive coefficient, which is as anticipated since a rising interest rate and an inverted yield curve are commonly seen as signals of a potential bear market. Additionally, we expect the falling policy rate variable to have a negative coefficient, as a decreasing policy rate indicates that the Federal Reserve is taking measures to bolster the economy when an economic stress is developing.

Confusion Matrix

To select the threshold for logistic regression, we consider our level of risk aversion. As our investment suggestion is geared towards the average investors, who tend to be more risk averse than institutional investors, we prioritize a sufficiently low false negative rate in our choice of cut-off point. This ensures that we minimize the risk of failing to predict a coming bear market, which is especially important for our target audience. As indicated by the chart below, which is produced using the final model, a threshold of 0.3 is the most suitable choice as it has a low false negative rate of 0.12 and a good false positive rate of 0.05.

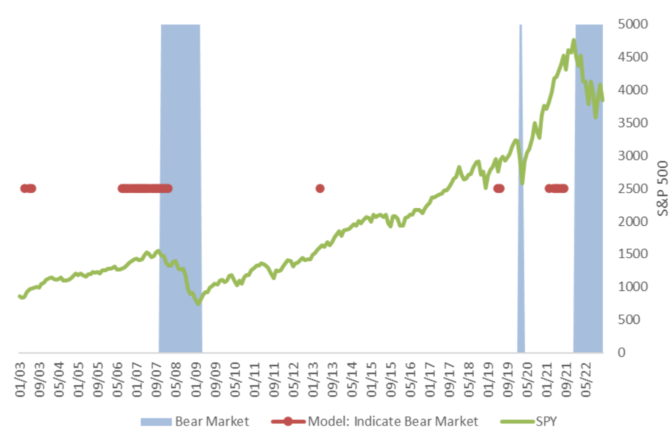


The confusion matrix using a threshold of 0.3 is as below:

Confusion Matrix	Actual (no bear market)	Actual (bear market)
Model (no bear market)	190	3
Model (bear market)	10	22

The chart below shows the model's performance compared to actual bear markets:

Model Comparison to Actual Bear Markets (2003-2022)



- The blue shade indicates the actual bear market period
- The red line represents the bear market period predicted by the model.

Ideally, the red dots should overlap or precede the blue shading. In our time horizon, the model correctly predicts all three bear markets, but it does suffer from false positives. However, we consider this to be a more favorable outcome than failing to predict a bear market, which could result in significant losses for investors. Future research could expand the time horizon and explore whether multiple, consecutive months of yield-curve inversion could lead to a more accurate model with fewer false positives.

Prediction using the final model

As of April 2023, our final model predicts a 33% likelihood of a bear market in 2023, which is above our threshold of 30%. This is consistent with our initial hypothesis that the S&P 500 index will not have a positive return in 2023. Our analysis has also revealed a probable correlation between interest rates and the S&P 500 return as suggested by our initial hypothesis. We recommend adopting a cautious investment approach. However, if the interest rates start to decline, this recommendation may not hold true.

IV. Other Modeling Approaches Explored

In addition to the final model presented above, we also explored the impact of some of the most relevant indexes against the S&P 500 such as the index fund return for Technology Sector and the Health Care Sector, the returns on commodities such as Gold and Silver, and the 10-year Treasury yield. We added some macroeconomic measures such as the Real GDP, CPI, Population Growth, Inflation, Consumer Sentiment, and a measure of the U.S. money stock (M2). Given the current rise in house prices we also included the deltas in Median House Prices. Subsequently, we sought to determine which predictors correlate to the response variable to use these predictors as part of our linear regression. The predictors highly correlated with our S&P 500 were: Technology sector (XLK), with a 0.917 correlation, and the Healthcare sector (XLV), with a 0.813 correlation. Other predictors would be silver, with a 0.261 correlation, followed by the population growth rate, with a negative correlation of 0.259, and inflation, with a negative correlation of 0.120. We ran several models, including a combination of the predictors with the highest correlation with our response variable (S&P 500), and overall, we determined that the Technology sector (XLK), Healthcare sector (XLV) with a 0.813 correlation, and population growth with a negative are the most relevant because they are the ones that are consistently statistically significant throughout our different models. As a result, we reduced our dataset to contain the statistically significant predictors using a variable selection method. It was noted that Technology Sector Fund, Healthcare Sector Fund, and Population Growth rate were statistically significant, thereby eliminating GDP, CPI, and interest rates (represented by 10-year treasury yield) through variable selection. Although those which were eliminated included the key variables which formed our initial hypothesis, we re-ran the model with the three statistically significant variables and the R^2 dropped to 0.928 but adjusted R^2 remained 0.928 and all three variables remained significant at 99% confidence level, indicating a high proportion of variance in the dependent variable that can be explained by the independent variables. In addition, we validated if there was a strong correlation among these three variables using plots, regressions, and the Variance Inflation Factors. We did not see any significant indication of multicollinearity.

Next, we analyzed a second set of data as proposed by Liu, Wang, Xiao, and Liang [1], which includes the major indexes such as Nasdaq, FTSE 100, NIKKEI 225 (stock market index for the Tokyo Stock Exchange), SSE Composite, as well as the currencies against the US dollar such as Chinese Yuan, Japanese Yen, Great Britain Pound, Euro, and finally the crude oil. We reduce our dataset to contain statistically significant predictors. Next, we ran the Nasdaq Index against the other two independent variables. The VIF turned out to be even smaller, so we do not suspect any major problem using these variables. Then we run the plots to determine whether this is a good linear model. We can see no pattern to the points in the Residuals vs. Fitted plot. This would mean that the error terms are uncorrelated and there is homoscedasticity (constant variance). Additionally, for the Normal Q-Q plot, even though it is not a perfect line, most of the points fall within the expected line.

Once we have these analyses, we create a third model with the statistically significant variables of the two previous models and run a new linear regression model using the Technology sector (XLK), Healthcare sector (XLV) and population growth, Nasdaq return, FTSE 100 return, SSE Composite, and Chinese Yuan and Great Britain Pound against the US dollar. After running selection variables to this model, we drop the SSE Composite. Again, we tested for correlation, and this time there was a very high correlation between Nasdaq and the Technology sector, with a VIF more significant than 10, so we decided to work only with the Tech and Health index, population growth, FTSE 100 and the three currencies listed above. Although our initial research rejected our hypothesis that GDP, CPI and interest rates would drive the S&P 500, our team wanted to further explore the relationship between interest rates and overall market performance, rooted in the S&P 500, to make an assessment on our second hypothesis: “The S&P 500 index will not have a positive return in excess of risk-free rate in 2023.”

V. Other Investment Options than the S&P 500 Index

Finally, we tested our initial hypothesis of “Gold or silver are not as effective as US Treasuries securities or TIPS ETF in hedging the current high inflation”. Similar to previous approaches, the first step was to gather data using the tidyquant function in R. Variables obtained, spanning a timeframe of January 2007 to December 2022, include the returns of the S&P 500; SPDR Gold Shares, which is an ETF that attempts to mimic the performance of gold bullion and a way for investors to access the gold bullion market; iShares Silver Trust, which reflects the price of silver owned by the fund; and iShares TIPS Bond ETF, which seeks to replicate the Barclays Capital U.S. Treasury Inflation Notes Index. In addition, similar to our second hypothesis, the market yield on US Treasury Securities: 1 month, 3 month, 1 year, 2 year, 5 year, 10 year, and 30 year were gathered using the get_fred function. To assess the best strategy for hedging the current high inflation and upcoming forecasted bear market, individual beta values were obtained for this time period to determine the asset class which has moved at the least volatile rate over the aforementioned period. The following results were obtained:

- For all US Treasury Securities, betas were noticeably negative. Logically, this makes sense as it proves the inverse relationship between the general market and treasury securities. The four with statistically significant beta values (over 90%) were 1, 2, 5, and 10 year Treasury securities with beta’s of -3.51, -3.53, -3.32, and -2.84, respectively.
- Gold, silver, and TIPS ETFs all had positive beta values of 0.08, 0.53, and 0.12, respectively. Most notably, only silver and TIPS were statistically significant (over 99%).

In conclusion, our hypothesis was deemed correct. US Treasury securities display an inverse relationship with the S&P 500. The ones found to be statistically significant indicate their performance is ideal in hedging high inflation during a bear market, a period which we forecast is on the horizon in 2023. Next, between the two statistically significant beta values of silver and TIPS, the TIPS ETF displays a much lower beta value indicating less market volatility in comparison to S&P 500 movements. This would also be the wiser investment in a period which is expected to feature high inflation and a bear market.

VI. Overall Conclusion and Recommendations / Business Impact / Further Analysis

Our team initially hypothesized the following to provide investors a map for predicting the S&P 500 in this tumultuous economic period and investment advice for the year 2023: United States GDP, Inflation (represented by CPI) and interest rate level are the key drivers of the S&P 500 index return; S&P 500 index will not have a positive return in excess of risk-free rate in 2023; and gold or silver are not as effective as US Treasuries securities or TIPS ETF in hedging the current high inflation.

Upon conclusion of our project and discussion as a group, we determined two of our initial three hypotheses were correct. First, from our exploratory data analysis, we determined that GDP, Inflation, and interest rate levels were not the key drivers of the S&P 500 index return, rather, Technology sector (XLK), Healthcare sector (XLV), and population growth were statistically significant when included with the aforementioned variables we initially hypothesized. Further, despite the lack of significance and indication that interest rate levels are key drivers in the S&P 500, we found it prudent to thoroughly research this area given the current interest rate environments and hikes imposed by the Federal Reserve. The results were impressive as our model successfully predicted various bear markets over a period of 20 years. In addition, using the model to predict 2023, it indicates a 33% chance of a bear market in 2023. Using these results, we found it imperative to research best investments, among silver, gold, the TIPS ETFs, and US Treasuries, to hedge the potential bear market and overall high inflation. Our results were intuitive as we discovered inverse beta relationships between the US Treasuries and low positive betas among silver, gold, and TIPS ETFs. Statistically significant betas were 1, 2, 5, and 10 year Treasury securities and silver and TIPS ETF.

In conclusion, prepare for a bear market and invest prudently in various US Treasuries and the TIPS ETF to hedge against market risk and navigate the winding road called 2023.

As a final step in the project, we came up with an investment recommendation based on our project findings. We conducted some research, analyzed a number of assets and we have come up with a sample investment portfolio. Our guiding principle for this portfolio was to have less volatility than S&P 500 (beta less than 1) and well diversified ETFs / businesses across categories that tend to do well during times of recession and bear market.

Our portfolio comprises of following 5 assets:

VDC (20% of portfolio) - Vanguard Consumer Staples Index Fund ETF seeks to track the performance of a benchmark index that measures the investment return of stocks in the consumer staples sector.

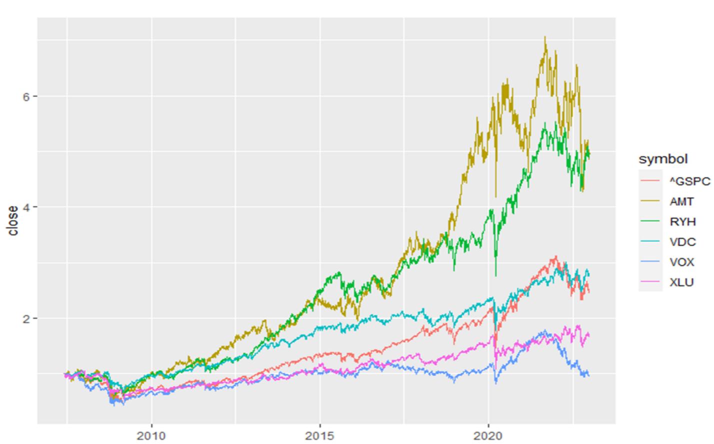
AMT (20% of portfolio) - American Tower Corporation is an American real estate investment trust and an owner and operator of wireless and broadcast communications infrastructure in several countries worldwide

XLU (20% of portfolio) - The Utilities Select Sector SPDR Fund seeks to provide investment results that, before expenses, correspond generally to the price and yield performance of the Utilities Select Sector Index

RYH (20% of portfolio) - RYH tracks an equal-weighted index of US health care companies taken from the S&P 500 Index.

VOX (20% of portfolio) - Vanguard Communication Services ETF seeks to track the performance of a benchmark index that measures the investment return of stocks in the communication services sector

Historical Performance of Proposed Assets in the Portfolio compared with the S&P 500 Index.



We took a set of returns of our proposed portfolio and we related them to the benchmark return of the S&P 500 index using the CAPM model. The table below summarizes our findings. With Beta of 0.67, our portfolio is less volatile as compared with the S&P 500 index, while still providing material returns and potentially overperforming broader markets in a short term.

Portfolio Name	Portfolio#1
Alpha	0.0041
Annualized Alpha	0.0507
Beta	0.669
Correlation	0.825
R-squared	0.68

We recommend re-visiting this portfolio allocation at the end of 2023 to re-balance and potentially re-allocate funds to S&P 500 to take advantage of market re-bound. Further, US Treasury securities would be a wise investment in a period obscured with high inflation and the outlook of a bear market.

Works Cited

[1] Liu, C. , Wang, J. , Xiao, D. and Liang, Q. (2016) Forecasting S&P 500 Stock Index Using Statistical Learning Models. Open Journal of Statistics, 6, 1067-1075. doi: 10.4236/ojs.2016.66086.

[2] Estrella, Arturo, and Frederic S. Mishkin. “The Yield Curve as a Predictor of U.S. Recessions.” *SSRN Electronic Journal*, 1996, https://doi.org/10.2139/ssrn.1001228

[3] Gasparėnienė, Ligita, et al. “Modelling of S&P 500 Index Price Based on U.S. Economic Indicators: Machine Learning Approach.” *Engineering Economics*, vol. 32, no. 4, 2021, pp. 362–375., https://doi.org/10.5755/j01.ee.32.4.27985.