Standard and Poor's Index value over the years*

How do Recessions affect the value of the Index

Mohammad Sardar Sheikh

27 April 2022

Abstract

In this research the Standard and Poor's index was used as a reference to calculate the fluctuations in stock prices at different periods in time and the results were used to estimate the state of the economy. It was found that a positive correlation exists between the price of the index and the value of the dividend along with the overall earnings. Conversely it was also extracted that an inverse relation exists between the interest rate and the value of the dividend. This information is important to predict the general trend in stock prices and investment patterns with changes in economy.

1 Introduction

The Standard and Poor's index, or more commonly referred to as the S&P 500 is a mutual fund or ETF that tracks the Standard and Poor's index of hte 500 largest US companies. The index is widely referred to as a very safe place for investment as the returns are almost unbeatable, even by professional investors. The mutual fund has a great number of investors and it can be useful as a measure of the overall economy, where the health of the index can be used to find the health of the economy. Predictions made based on the index can be extrapolated and applied to many other company stocks and the index is sometimes used as a bench mark. (**citethune_2021?**).

This is why the topic of my research today is the performance of the index over the long run. I use the data that I have to find out the performance of the stock at different times in history, especially during major periods of recession. Throughout the paper, I choose to use real values instead of nominal ones. The price of the index has increased greatly ever since it was founded in 1923, when it was just comprised of 90 of the largest companies in the United States. By 1957, the index had expanded to include 500 of the largest companies and it truly took off. The price seems to have increased exponentially around the 1990's as compared to the long run average growth trend. We can see from the results that whenever there is a recession, the value of the stock takes a big dip. Not only does the price of the stock decrease, the earnings and dividends also decrease. The severity of the decrease in index valuation also depends on the severity of the recession. Recessions such as the Great Depression in 1929-1938, and the great recession in 2007-2008 led to a very large decrease in the value fo the index.

Since the index is so important in investors and portfolio owner's eyes, I also predicted the real value of the stock in the future based on the components of the index such as dividends and earnings. I also added predictors that include the Consumer price index (measure of inflation) and the long term interest rate. When the value of a dividend increases by \$1, the price of the index increases by 5.4%. Similarly, when the earnings increase by \$1, the value of the index increases by 0.5%. When the CPI increases by 1 (%point in this case), the price of the index increases by 0.3%. When the interest rate increases by 1, the value of the index increases by 1.5%.

---

*Code and data are available at: https://github.com/mohammadsardar/S-P500-index

Another very important thing that I want to predict is the value of the real interest rate in the economy. The real interest rate is a very essential tool for investors when making decisions. A high interest rate decreases investment in the economy by making investing and borrowing riskier. A high interest rate also makes saving more appealing as an individual gets a higher rate of return on an savings. Conversely, a low interest rate makes investment very appealing, decreasing the risk associated with doing so and decreasing the realized benefits of savings. (**citepettinger?**) This makes predicting future interest rate more important. Based on the results of the model, when the value of a dividend increases by \$1, the interest rate decreases by 0.8%. Similarly, when the earnings increase by \$1, the interest rate increases by 0.12%. When the CPI increases by 1 (%point in this case), the interest rate increases by 0.2%. When the value of the CAPE ratio increases by 1, the interest rate decreases by 1.5%.

This paper and research was done using R (R Core Team 2020). I used many other packages to help me in the creation of this paper. They include: TidyVerse (Wickham et al. 2019), kableExtra (Zhu 2021) for the informative tables, Patchwork (Pedersen 2020) for the graph layouts, readr (Wickham, Hester, and Bryan 2022) to help load and write data, lubridate (Grolemund and Wickham 2011) to help manage important dates, car (Fox and Weisberg 2019) to help with model assumptions and verification and lme4 (Bates et al. 2015) to help build mixed models.

2 Data

For this paper and analysis, the data that I am working with is collected by the Economist Robert Shiller who is an Economist at Yale, and the dataset is open access at his website (Shiller, n.d.). We download the data directly from his website onto a local machine and then we can use that data. We can also access the dat using a json file, the details of which are provided in the cleaning script. The next step is cleaning and organizing the data so that it can be used effectively in my analysis. All the variables have to be renamed and significant figures have to be checked so that they match. The dataset contains many variables that we do not have an interest in such as bond returns and values of the annualized stock returns that do not add too much information to the report but make it more complex and harder to replicate. The variables that I choose to keep are value of the index, value of the monthly dividends, value of the monthly earnings (all the variables mentioned above have both real and nominal values included in the final dataset). We also utilize information about the long term real interest rate of the economy (this is very important for investment decisions) and the consumer price index (this can be used to calculate inflation in the economy from year to year, or compared to a base year) for which the base year is 1983. We also include the CAPE ratio as one of the variables. This is important as (ADD INFORMATION ABOUT THE CAPE RATIO AND CITE IT IF POSSIBLE).

Another variable that I had to account for was the date. This needed modification as it was not in proper format. effectively utilizing lubridate allowed me to overcome this and create a clean variable. A variable that I created for the sake of the analysis was a year variable. This was made using the original date variable. The data that we had available was from 1871 onwards. However, I chose to start the analysis from the year 1913 as that is the first year that we have accurate data for regarding the consumer price index (CPI). Some of the variables that I considered important are shown in the graphs.

For displaying the data that we have, instead of choosing monthly values, I summarized them into yearly values and then plotted them so that they were not as cluttered and it would be easier to consider them as one off graphs.

We can see from Figure 1 that the value of the index has increased greatly since it was established in 1923. In 1923, it was launched on a very small scale. it started out with only 90 different stocks, but by 1957, it had increased that to 500 of the largest ones.We can see from the figure (which shows both real and nominal prices of the index) that a great proportion of that increase has taken place 1985 onwards, with very steep increases in market valuation for the stocks.

Figure 2 shows us the real value of the dividends for the index (A dividend is the distribution of some of a company's earnings to a class of its shareholders). We can see that while the value of the dividends rises
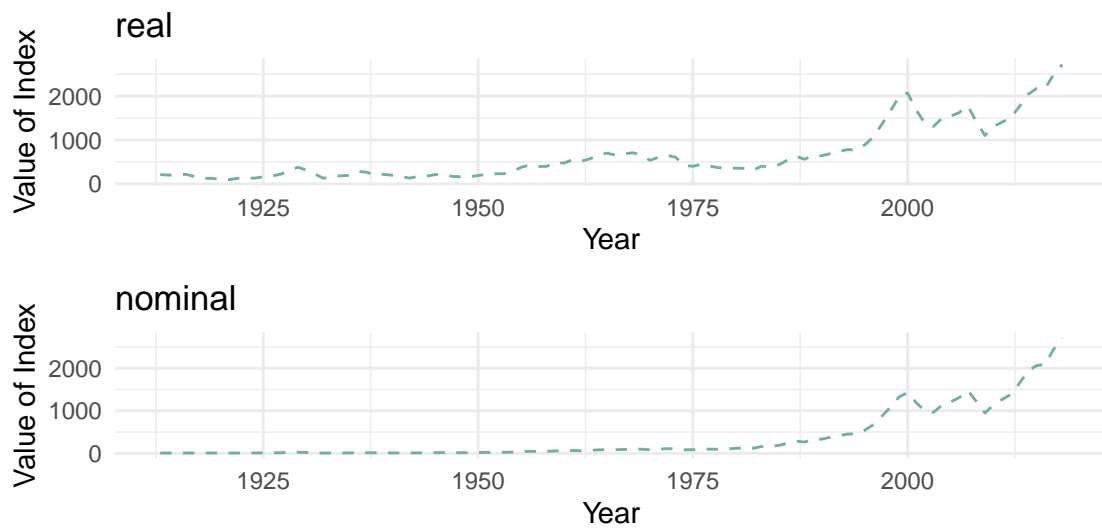
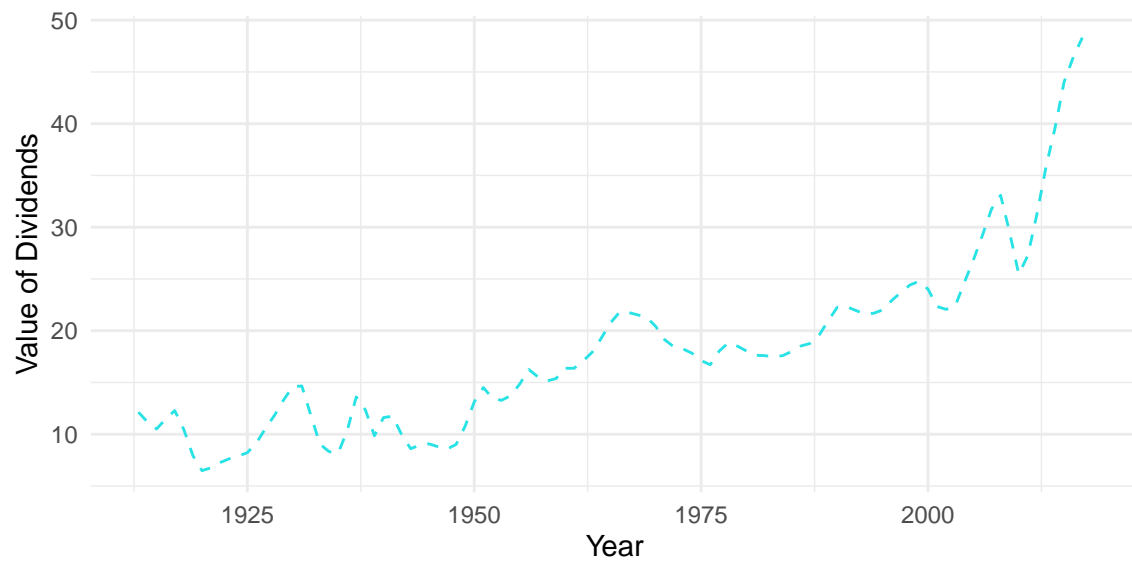Figure 1: Average yearly S&P500 index value

Figure 2: Average yearly S&P500 dividends (real)

just like the valuation of the index, it does so at a much slower rate and it is a lot less smoother, with many peaks and troughs.

Figure 3: Average yearly S&P500 earnings (real)

Using Figure 3 we can see that the earnings increase steadily over a great period of time. We can also see near the end how the earnings take a huge plunge. This can most likely be attributed to the great recession of 2007-2008 and we will look deeper into this.
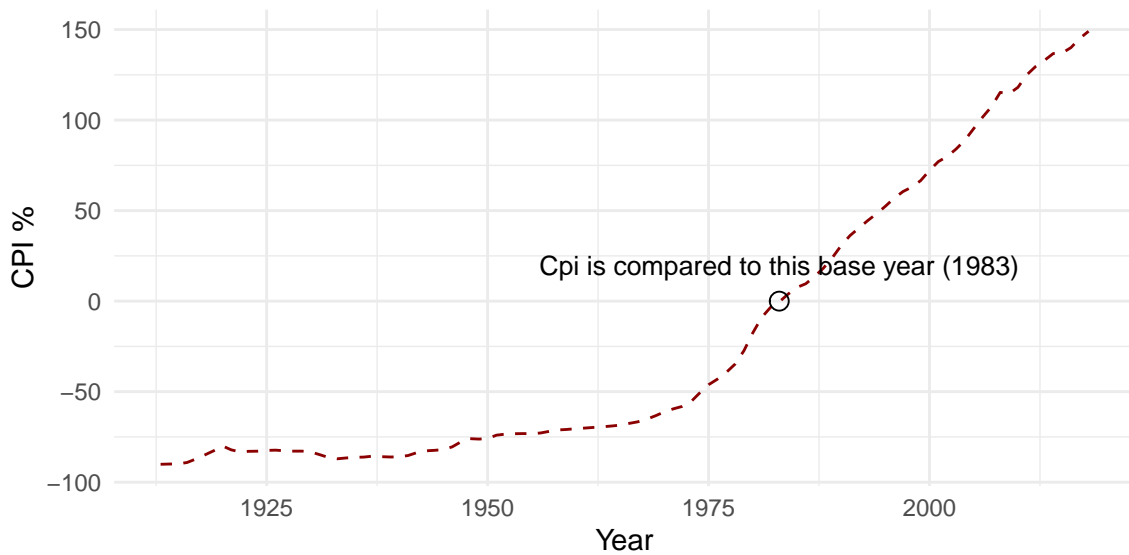

Figure 4: Average yearly Consumer Price Index

By the aid of 4, we can see that the average price of a bundle of goods in the economy has increased greatly since the measure was introduced in 1913. The CPI is used to determine the level of inflation between years. The US considers the base year for the CPI to be 1983. This means that all valuation of bundles and percentages are relative to 1983 dollars and compared to the price of the same bundle of goods in 1983. For example, we can see from the figure that the value of the CPI in 2005 was around 100%. We can interpret

this as: The same bundle of goods in 2004 costs us twice as much compared to if it were purchased in 1983. So the inflation is effectively 100% between the two periods.

3   Model

I want to create 2 models as I want to be able to predict 2 things. the first one is the long term interest rate. the second model that i create is the real value of the index Now we will start thinking about what sort of model to use for our predictions. we can use any model that we prefer abut it has to be well explained. i think that for my purposes, a linear model will be the best with just a few predictors (real values instead of nominal ones)

We want to be able to predict the value of the S&P 500 index based on the predictors that we have. Or can we do something about inflation? as that has an effect on the economy as a whole.... WE WILL USE long term real interest rate as the response variable. what i have to do is justify why i think this is an important thing to create a model for. Do I have to do any calculations for this?

the model that i chose has real interest rate as the response variable. now I have to justify why I think that a linear model is suitable and why interest rate is a suitable choice for the model

There are some things that I would like to predict based on the data that we have. Firstly, I would like to predict what the value of the index will be in the future. The S&P500 is one of the most popular indexes in the world, and its valuation is closely followed by millions. We can predict the state of the economy based on the value of the stock at a current period in time. based on the data exploration that has been done, and fulfilling model assumptions, the most appropriate predictors to use for the model would be the value of the dividends, value of earnings, the long term interest rate and the consumer price index (measure of inflation). We also take the natural log for the response (in this case it is the real value of the stock), to fulfill model assumptions. We will fit a linear regression model which has the following formula:

$$log(\gamma_i) = \beta_0 + \beta_1\alpha_1 + \beta_2\alpha_2 + \beta_3\alpha_3 + \beta_4\alpha_4$$

The elements of the model can be interpreted as:

- β_0 is the intercept term.
- γ_i is the real price of the stock (we take the log).
- α_1 is the long term real interest rate.
- α_2 is the value of the dividends that investors get from the index.
- α_3 is the value of the earnings (profits) from the index for a given period.
- α_4 is the the the consumer price index (base year is 1983).

Another model that is important to fit is to predict the value of the long term interest rate. The interest rate has been used by investors to figure out if a specific period is worth investing in. High interest rates dry up the investment in the economy as they increase the risk associated with investing, as there is a greater interest to be paid on the principal amount. During periods of high interest rates, people choose to save instead and this slows economic growth. The model that seems appropriate to fit includes the predictors such as real value of dividends, real value of earnings, consumer price index and the pe10 ratio (CAPE ratio). For the sake of fulfilling model assumptions, we take the logarithm of the response variable (interest rate). We also make sure that there isn't too much collinearity between the predictors and that the VIF (variance inflation factor) of the model is not too high. the final model can be represented as:

$$log(\gamma_i) = \beta_0 + \beta_1\alpha_1 + \beta_2\alpha_2 + \beta_3\alpha_3 + \beta_4\alpha_4$$

The elements of the model can be interpreted as:

- β_0 is the intercept term.
- γ_i is the real interest rate that we get from the output of the model (we take the log for this to make it normal).
- α_1 is the value of the dividends that investors get from the index.
- α_2 is the value of the earnings (profits) from the index for a given period.
- α_3 is the the consumer price index (base year is 1983).
- α_4 is the CAPE ratio.

A linear model seems appropriate to predict both of these values.

4   Results

I want to isolate the periods in the history of the United States when there were major recessions and then analyze how the price of the stock varied in those times. As explained, a recession is a period when the overall GDP growth rate of the economy falls as compared to the long run average growth (can go negative). In all of the recessions that I analyze, the GDP decreased by a huge margin and I want to analyze if this had any significant effects on the price of the S&P500 index value or on the value of the earnings for the index. I also want to predict both the value of the long term interest rate based on the components of the S$P500, and i want to be able to predict the value of the index in the future. To do this, I have created two models which will be interpreted.
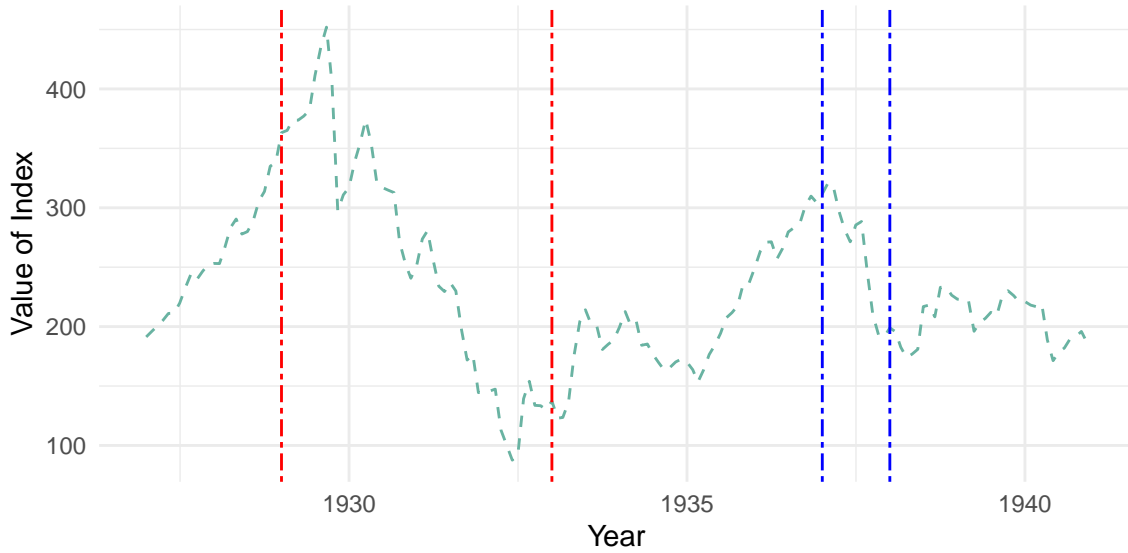

4.1   Great Depression

Figure 5: Monthly value of the S&P500 during the Great Depression

The Great Depression was the worst economic downturn in the history of the industrialized world, lasting from 1929 to 1939. It began after the stock market crash of October 1929, which sent Wall Street into a panic and wiped out millions of investors.(**citedepression_2009?**). the great depression was caused by a massive influx of wealth into the stock market, which over inflated it. The price of the stocks reached record high levels and investors started doubting the authenticity of the price. Massive waves of selling took place where millions of shares were traded and the value of the stock market plummeted and with it, millions of investors wealth. The country was already experiencing a decrease in both agricultural and industrial production and

millions of jobs were lost, with over 20% of Americans being left unemployed. wages dropped and bank runs started to take place. Newly elected president FDR worked hard to reverse the recession effects.

When the great depression hit (the largest recession in american history, where GDP fell by 27%), the value of the S&P500 fell by a great deal. We can see from Figure 5 the relative value of the real price of the index. The great depression hit the economy in two stages. The first was during 1929-1933 and we can see the huge plunge that the value of the stock took, where the price of the stock dropped from a local high of $450 to around $100. This is a huge decrease of around 350%. This can be seen from the vertical red lines in the figure. This was only just the first wave of the recession. We can see that in the few years after the end of the first wave, the value of the index start to slowly rise and recover, but then the second wave of the recession hit. This is represented by the blue lines. while the recession only lasted for one year, the value of the index yet again took a steep dive, learning its lesson and staying down.

4.2 World War 2 recession

But with the surrender of both Germany and Japan in 1945, military contracts were slashed and soldiers started coming home, competing with civilians for jobs. As government spending dried up, the economy dipped into a serious recession with GDP contracting by a whopping 11 percent. this was actually a fiscal cliff. (**citehistory_2019?**)
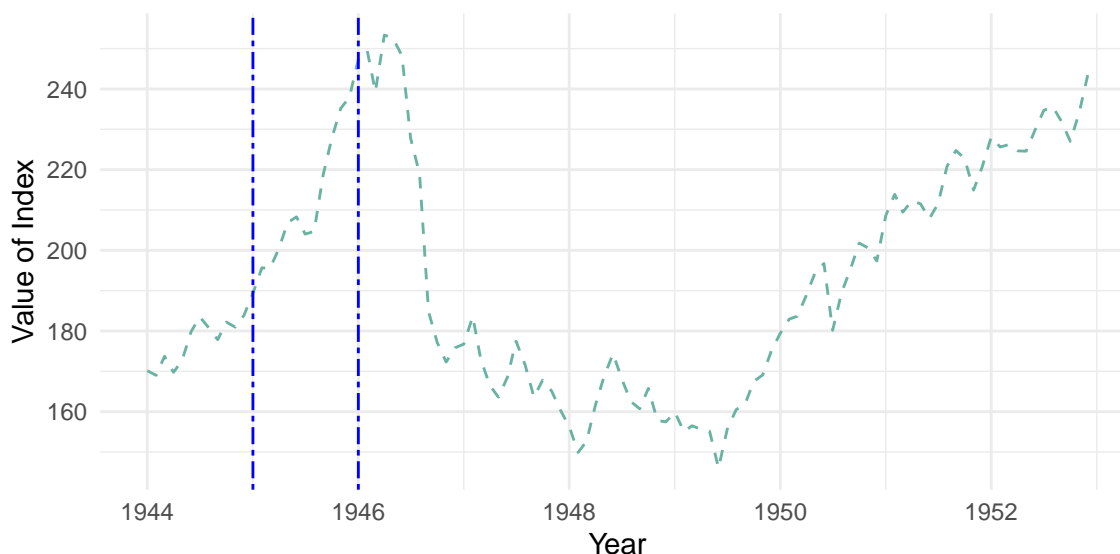

Figure 6: Value of the S&P500 during the Recession of 1945

From Figure 6 we can see that the value of the index actually increased during the recession of 1945. This seems surprising as this recession in particular hit the United States very hard and there was a decrease of 12.7% in GDP. yet we can see that the value of the index remained strong. We have a counter argument since the effects of a regression are visible for a long period even after the recession has subsided. we can see from the figure above that this seems to be the case. since the regression was officially over within 8 months, the index did not have time to adjust to the fluctuations in the market and we can see this as a lagged effect since we observe a huge and long drop in the value of the index which took a few years just to recover and reach its original level. The after war policy was expansionary however and we can see from the figure that within the span of 10 years, the value of the stock was back to its original level. We can also see that the years leading up to the recession were very beneficial for the value of the index. This is because the government was investing billions into the military and defense industry, and many of these companies were part of the index.

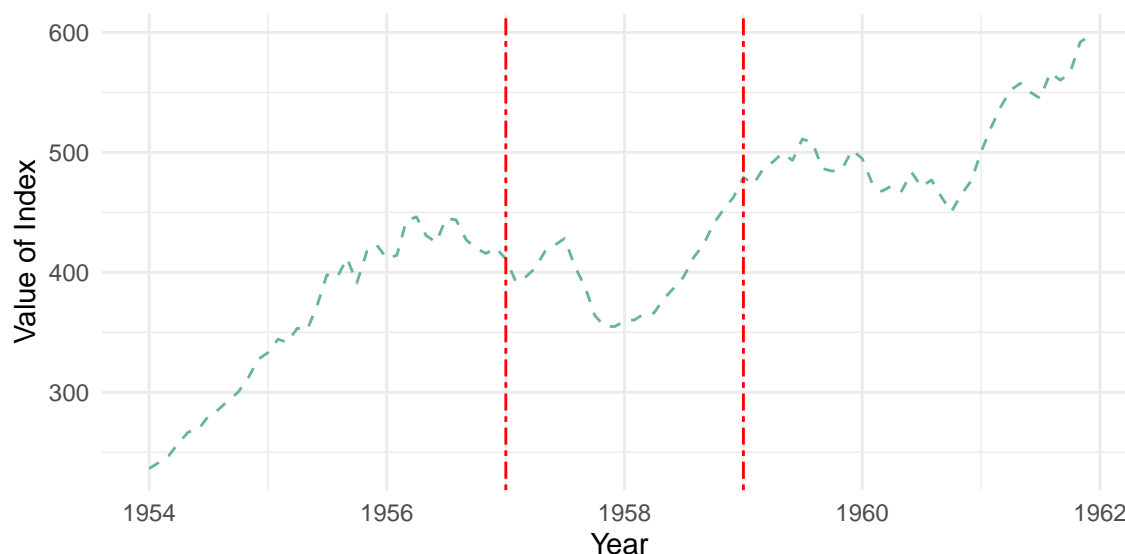## 4.3 Recession of 1957-1958 / Eisenhower recession.

Figure 7: Monthly value of the S&P500 during the Eisenhower Recession

The recession of 1957-1958 is also known as the Eisenhower recession. It was for a short time period, but corporate profits fell by a great deal. Using Figure 7, we can see during this regression as well that the value of the index also fell somewhat. This fall is not as significant as the other recessions. This recession was significant in the post world war 2 period and unemployment rose greatly while corporate profits fell by as much as 25%. A weak fiscal and monetary policy contributed to this and there was significant stagflation as well as just a reduction in output. Stagflation occurs when the economy deteriorates and there is high inflation at the same time. this leads to a very bad situation for the country and it is what was seen in this particular recession.
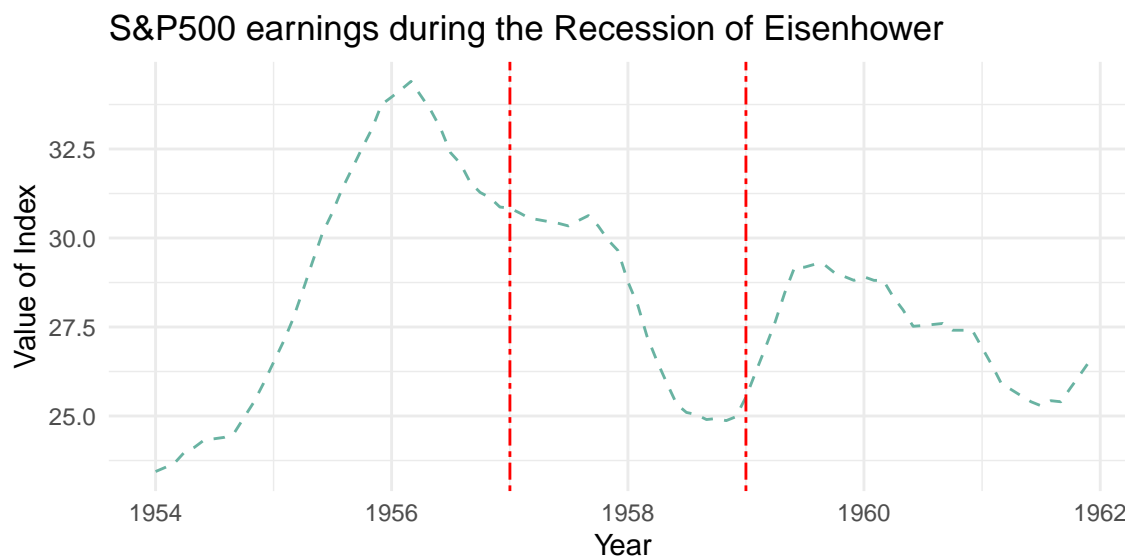

Figure 8: S&P500 earnings during the Eisenhower Recession

Looking at Figure 8 we can see that the earnings/profit of the index fell greatly during the recession. This

Table 1: Coefficients from Interest Rate Model

term	estimate	standard error	statistic	p-value
(Intercept)	1.651	0.044	37.321	0.000
real_dividend	-0.008	0.004	-2.040	0.042
real_earnings	0.001	0.001	0.603	0.547
consumer_price_index	0.002	0.000	5.601	0.000
pe10	-0.015	0.002	-6.103	0.000

seems in line with what the research says as profits were more affected that GDP. (**citepike?**)
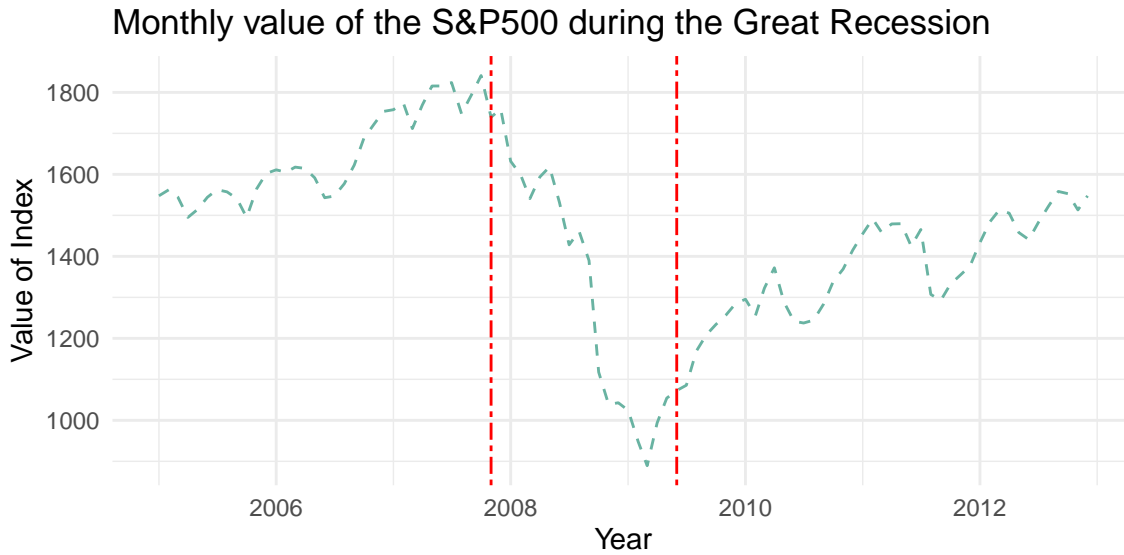
4.4 The Great Recession:

Figure 9: Monthly value of the S&P500 during the Great Recession

The Great Recession was similar to the Great Depression in the sense that the cause of the recession was similar. There was a rapid increase in the valuation of stocks and companies while profits did not increase linearly and these stocks were predicted to be in a bubble. Housing was a huge contributor to the Great Recession. With the mass of wealth creation, many people wanted to own homes. Banks were giving out sub-prime mortgages to increasingly risky individuals, claiming that the increase in the value of the property would make up for it. It did not. Unemployment grew rampant and the economy plunged. Not only was this specific recession in the United States, its effects spread worldwide, effecting so many other countries and continents. We can see just how devastating this regression was in Figure 9, for the value of the stocks that were traded in the index as we can see a huge dip in the value of the standard and poor's index in just the span of one year. Going from a local maximum of 1800 to a local minimum of 900 dollars in real terms.

4.5 Long Term Interest Rate model

Interpreting the results of the regression in Table 1, we can see that the interest rate depends on a few predictors. When the value of a dividend increases by $1, the interest rate decreases by 0.8%. Similarly, when the earnings increase by $1, the interest rate increases by 0.12%. For the Consumer Price Index, we observe a similar relationship. when the CPI increases by 1 (%point in this case), the interest rate increases

Table 2: Coefficients from Price Model

term	estimate	standard error	statistic	p-value
(Intercept)	4.614	0.035	133.724	0
long_interest_rate	0.013	0.004	3.522	0
real_dividend	0.054	0.003	18.365	0
real_earnings	0.005	0.001	5.739	0
consumer_price_index	0.003	0.000	10.641	0

by 0.2%. finally, when looking at the pe10 ratio (CAPE ratio), when the value of the ratio increases by 1, the interest rate decreases by 1.5%.

4.6 S&P500 index value model

Interpreting the results of the regression in Table 2, we can see that the value of the index depends on a few predictors. When the value of a dividend increases by $1, the price of the index increases by 5.4%. Similarly, when the earnings increase by $1, the value of the index increases by 0.5%. For the Consumer Price Index, we observe a similar relationship. when the CPI increases by 1 (%point in this case), the price of the index increases by 0.3%. Finally, when looking at the long term interest rate, when the interest rate increases by 1, the value of the index increases by 1.5%.

5 Discussion

5.1 Value of the Index during Major Recessions

A recession is a period of time when economic activity slows down and output in the country drops as compared to the average. There are many adverse effects of a recession, the unemployment increases and wages tend to decrease. Instability rises in a country and the population is miserable. Recessions are bad for businesses as well as for individuals and it is always the central banks policy to avoid a deep recession. Since the S&P 500 index is a aggregation of the 500 largest companies in the United States, it is a good predictor for what the rest of the economy is experiencing and the performance of the economy as a whole. Based on the results that I get, we can see that during every major recession, the value of the index has decreased significantly. In the case of the Great Depression, the drop in value is drastic. The value of the index decreases by 350% from its short term maximum and it takes a very long time for the value to return to the industry average for that period.

The recession at the end of world was 2 lasted for a very short period of time, yet it was still devastating for the value of the index. The recession took place because soldiers returned from war and they were competing with civilians for jobs that just were not available. While the price did not actually fall during the recession period (8 months), it did just after the recession, appearing as a lagged effect. The market recovered relatively quickly after some time due to the expansionary monetary policy that was put in place.

Another major recession was the recession of 1957-1958, that is also known as the Eisenhower recession. This recession was also due to the after effects of world war 2 and the effects were exasperated due to weak management of the economy and bad monetary policy. Unemployment rose greatly while corporate profits fell by as much as 25%. The economy was also subject to stagflation, a period where there is a recession so the economy is stagnating, and this is accompanied by rising prices, inflation. Stagflation makes the economy extremely weak and is to be avoided at all costs. During this particular recession, the value of the index did not fall greatly, yet the value of the earnings of the index plummeted.

The final major recession that I looked at was the Great recession. The recession was caused by a bubble in the stock and housing market, and due to very risky investments and loans from the banking and shadow banking

industries. Sub-prime mortgages were authorized and when investors realized that the actual situation was dire, everyone tried pulling out of the market at the same time, causing a larger fall in market valuation and a loss of billions of dollars of wealth in the economy. unemployment was very high and it took a long time for the economy to recover. We can see once again that the value of the index fell greatly and it took a long time for it to recover to the original level.

5.2 The relationship between Long term interest rates and the economy

Long term interest rates are very important while making investment decisions in the economy. A higher interest rate slows investment in the economy and it makes the act of investing and borrowing riskier. individuals and businesses would rather save and earn the interest than take the risk and face a larger liability. This is why trying to predict the long term interest rate is of great importance in the economy. Investors can try to make decisions about the future based on current events and long term predictions. We can see from the results of the regression that there a few things that have an effect on the long term interest rate. An increase in the value of the index dividends decreases the interest rate. An increase in the value of the earnings increases the interest rate. An increase in the CPI increases the interest rate. Finally, an increase in the CAPE ratio for the index decreases the interest rate. Since we get information about these values monthly, we can use them to predict the interest rate.

5.3 The relationship between price of the stock and how it appears to be important (can include how the other stocks and the strength of the economy can be indicated by the performance of the S$P500 index and so this can be expanded into a well written point.)

The value of the index is another important metric that I want to predict. Since the index is comprised of 500 of the largest companies, it represents a huge market share and it serves as a benchmark for portfolios and investors with private investment. The strength of the index can be used to gauge the strength of the economy as a whole and any trends in the economy. Since many people have invested in the index, it is also informative and beneficial for them to figure out the factors that effect the value of the stock. From the results of the regression, we can see that an increase in the value of the index dividend increases the value of the index. An increase in index earnings also increase the value of the index. An increase in the Consumer price index and the long term interest rate also increase the value fo the index. These results are easy to predict. the value of the index, and its constituents have been increasing since the index was established, so it is easy to predict that an increase in the value of any one of its constituents would increase the value of the index.

5.4 Weaknesses and next steps

One of the next steps that we can take is to include a new variable that codes for recessions. this can be a binary variable, and we can include this in the models to give us information about what happens to the relative price of the stock and the interest rate during recessions. We can and should also include information from other indexes such as the DOW Jones and NASDAQ as the combination of all these would result in us getting more accurate predictions.
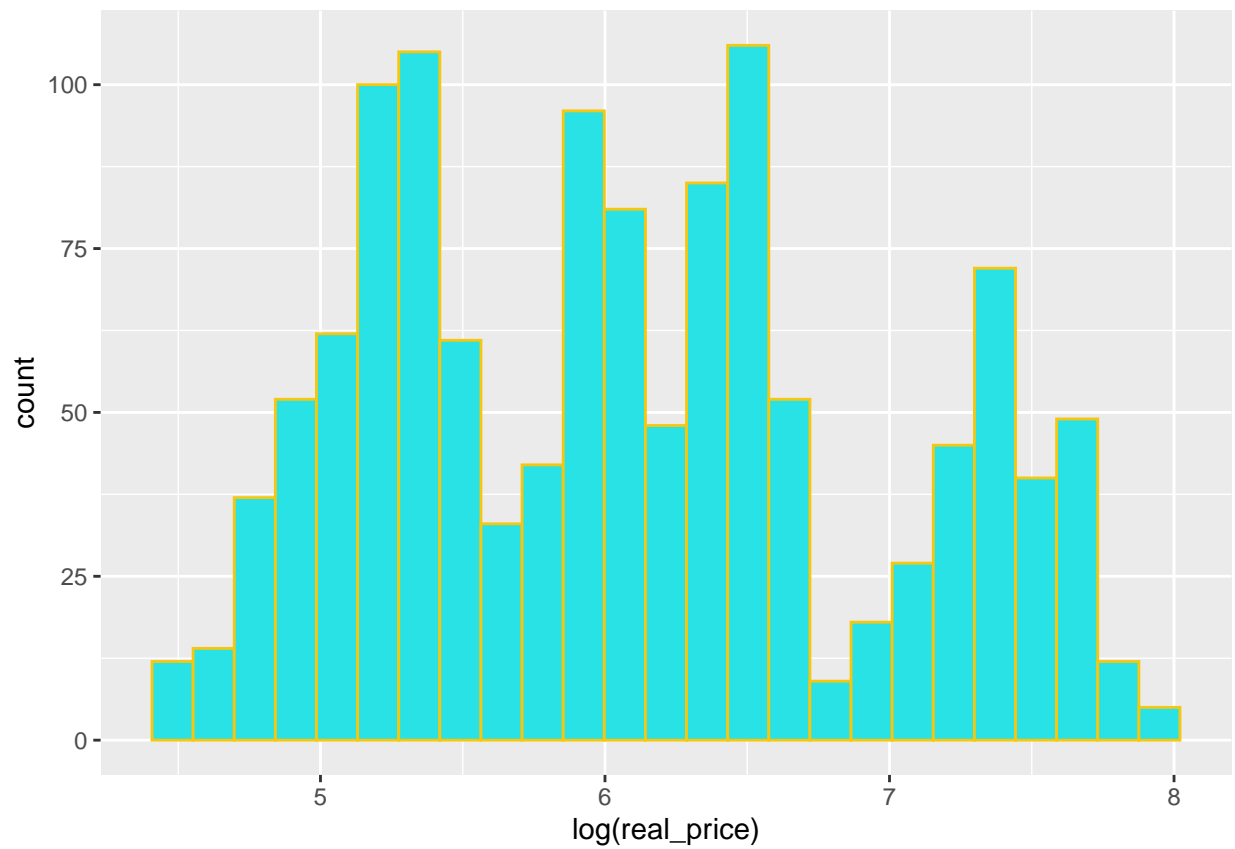
Appendix

This is just some of the work to justify the model and check for some of the model assumptions

`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
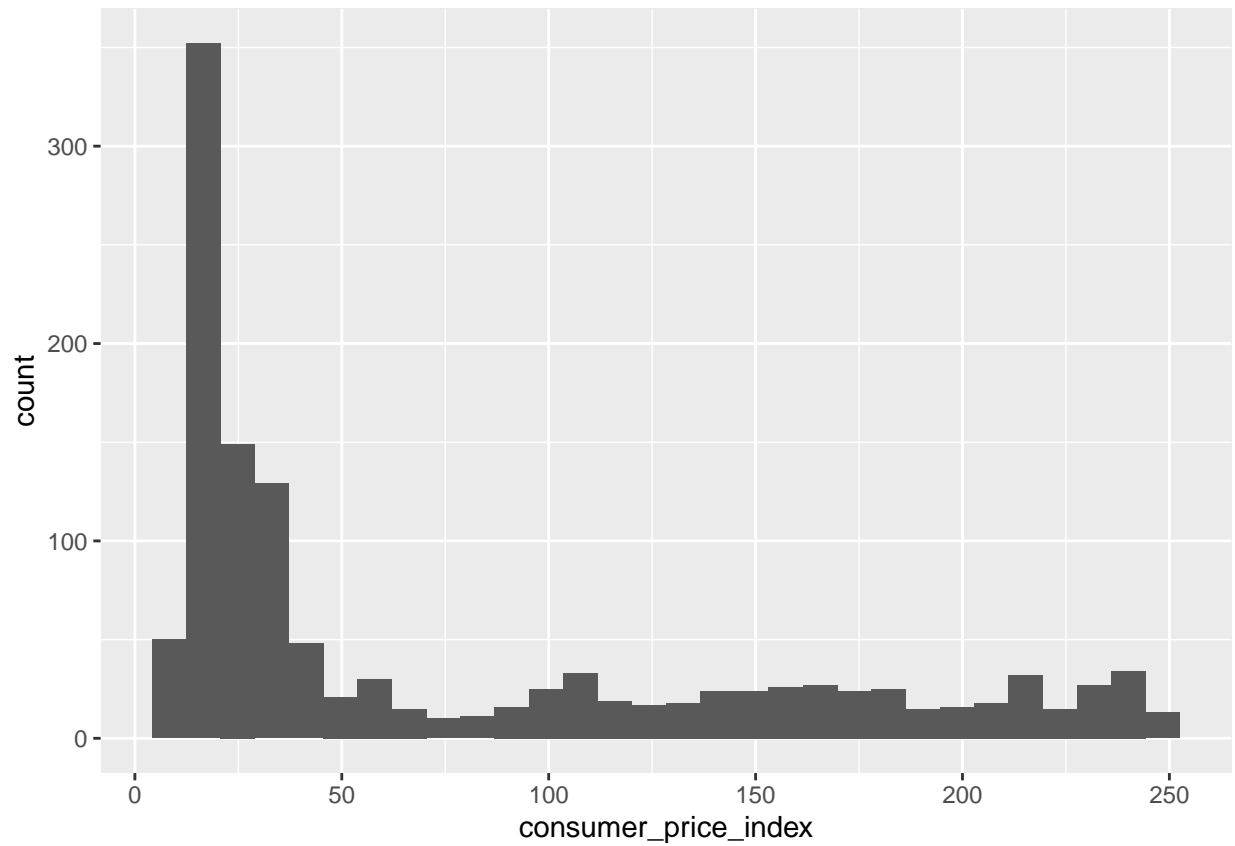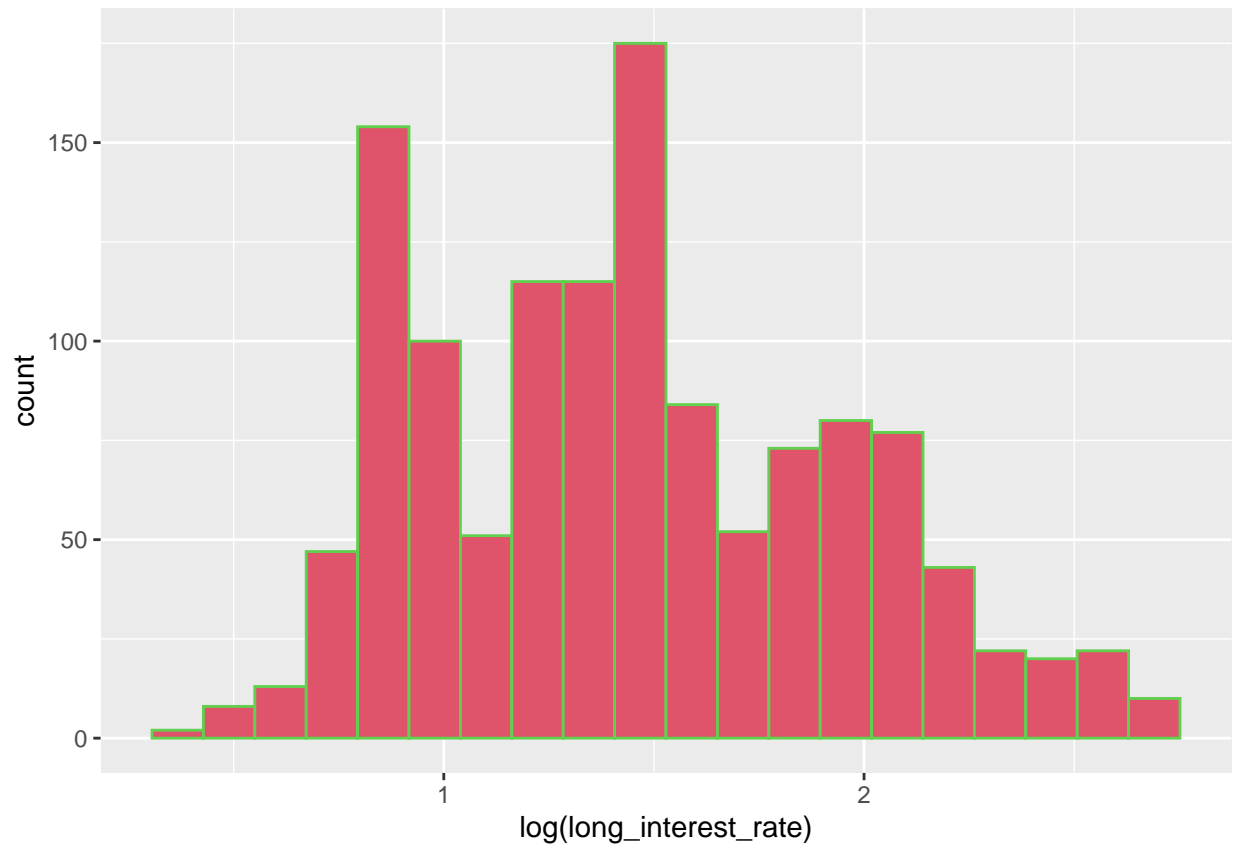
A   Additional details

References

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67 (1): 1–48. https://doi.org/10.18637/jss.v067.i01.

Fox, John, and Sanford Weisberg. 2019. *An R Companion to Applied Regression.* Third. Thousand Oaks CA: Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Grolemund, Garrett, and Hadley Wickham. 2011. "Dates and Times Made Easy with lubridate." *Journal of Statistical Software* 40 (3): 1–25. https://www.jstatsoft.org/v40/i03/.

Pedersen, Thomas Lin. 2020. *Patchwork: The Composer of Plots.* https://CRAN.R-project.org/package=patchwork.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Shiller, Robert. n.d. "S&p 500 Data." *Online Data - Robert Shiller.* http://www.econ.yale.edu/~shiller/data.htm.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2022. *Readr: Read Rectangular Text Data.* https://CRAN.R-project.org/package=readr.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax.* https://CRAN.R-project.org/package=kableExtra.