Andrew Plum

Dr. Dan Hickman

Econ 453

10/21/2023

Semester Project Proposal

I previously wanted to do something more unique for my project; however, I realized how difficult and time intensive it could be to find and prepare the data I needed for my analysis. Because of this, I decided to change course and do a more straightforward project where the data is more easily obtained. The question that I plan to explore for my project is, “What variables drive the price action of S&P 500?”

I will be using time series data from the S&P 500 stock index as the response variable in the OLS regression model; the data I plan to collect for it goes all the way back to 1983. I want to focus on the S&P 500 because it follows the top 500 US companies, and the index is weighted by market capitalization and not by stock price like the Dow Jones Industrial Average. This means that I would see more or less similar results for most other stock indexes because most stock market indexes are weighted by market capitalization. I plan to obtain the data through Yahoo Finance.

I am considering using several explanatory variables in my model. I hope to cover a variety of factors with them to accurately model the situation, and listed below are ones that I think could be relevant. I plan to collect data from several different sources. Here are the explanatory variables, their data sources, why they are important to consider, and how I think they will impact the price action of the stock market:

1. Real GDP: Provided by FRED; gauges the health of the US economy, if it is good then company earnings are likely good which increases stock prices.
2. Federal Funds Rate: Provided by FRED; indicates the cost of borrowing money, then people are likely spending less in the economy which impacts company earnings negatively which decreases stock prices.
3. Inflation CPI Data: Provided by FRED; inflation impacts people’s spending behavior in the economy, I think cost push and demand-pull inflation will affect the stock market differently; however, I can only account for overall inflation here.
4. Unemployment Rate: Provided by FRED; indicates how productive the economy currently is, if it higher then people are likely cutting back on spending.
5. Real Wages: Provided by FRED; if people earn more in real wages, then consumer spending in the economy will increase which will increase company earnings.
6. Personal Savings Including Consumer Durables: Provided by FRED; reflects how much consumers are saving, if it increases it will mean people are spending less which means company earnings will likely decrease.
7. Consumer Sentiment (University of Michigan): Provided by FRED; reports how good the consumer feels about there situation in the economy, if they report feeling good then they are likely spending more which likely means company earnings are better.
8. Real M2 Money Supply: Provided by FRED; as real M2 money supply increases more money is circulating in the economy which would lead to increased spending, this will likely be positively correlated with CPI inflation.
9. Velocity of M2 Money: Provided by FRED; as the rate that money moves around in the economy increases, consumer spending increases which will likely entail higher company earnings, likely positively correlated with CPI inflation data.
10. Fiscal Federal Government Policies: Provided by US Treasury; if the government spends more than it taxes, then company earnings will increase, and opposite effect for the contrary.
11. Commodity Prices (GSG Index): Provided by Yahoo Finance; data on the GSG Index, which tracks a basket of commodities, will be collected to understand the relationship between commodity prices and the stock market, as it increases.
12. Oil Prices (WTI Index): Provided by Yahoo Finance; the West Texas Intermediate (WTI) crude oil index will be used to factor in the impact of oil prices on the stock market, if the index is higher will see cost push inflation from increased transportation costs which will hurt company earnings.
13. Volatility Index: Provide by Yahoo Finance; indicates how volatile the buying and selling the stock market is; might account for volatile nature of the market.
14. Productivity Data: Provided by Bureau of Labor Statistics; as the economy is more productive, more goods are created more efficiently, savings can be passed onto the consumer, which could entail high consumer spending improving company earnings.
15. Election Turnout Data: Provided by The Federal Election Commission; helps to gauge sentiment of the wellbeing of the economy, if higher election turnout, the economy is likely in a worse place which likely means consumer earnings are hurt. Could also look at things like the presence of elections and its impact on the stock market.

To get an initial feel of the data, I plan on running regressions for each of the explanatory variables where only one explanatory variable is used each time. Next, I plan to run regressions with combinations of the variables and trying to account for all factors of the situation whilst trying to forgo collinear regressors in the model. Given the variation in the length of data and frequency of data collection, I will anticipate running several different regression models to accommodate the data's availability and time span. For example, most data from yahoo finance is daily whereas most FRED data is quarterly, so I will run a regression where the data of the variables is available daily and another where the data is available quarterly (the daily data could be turned into quarterly data so it can be ran in the same model). One thing I should consider when making my model is whether the variable is a leading, lagging, coincident indicator; I should run different regressions based on these groupings.

I am aware this proposal is longer than what you were asking for, but I wanted to give an accurate picture of where I am at with my project.