Capstone Two: Project Ideas: Roger Swartz

Project Idea 4

Inflation Rate Modeling

Idea focused on modeling inflation in the United States. Inflation is generally caused by significant fluctuation in the amount of available capital vs. the amount of goods and services being generated. It is fairly well established that increasing the available money supply without a commensurate increase in goods and services produced results in inflation. Economists look at a number of key economic indicators to try and decide if the federal funds rate should be increased to control inflation. Key economic indicators include: nonfarm Payrolls, GDP, capacity utilization, durable goods orders, unemployment rate, building permits, national debt and also quantitative easing and quantitative tightening. Quantitative easing is the process of printing money. The Fed has the capability to simply write checks to the Federal Government without having that money to begin with, without collecting taxes and without issuing bonds. Effectively, the Federal Government can borrow from the Federal Reserve bank at 0% interest while the Federal Government can borrow from the public (the public meaning a source of actual money) at the T-Bill rate. The 6-month T-bill rate is at 4.3% on an annualized basis. One may not even call it borrowing from the Federal Reserve Bank since they do not have to necessarily pay it back.

The general model for how the Federal Reserve Bank controls inflation that the Fed increases the cost to borrow money by setting the Federal Funds Rate (FFR). Increasing the cost to borrow money can make projects substantially more expensive. Consequently, employers tend to lay off employees in response to an increase in the FFR. This in turn reduces the amount of spending and people tend to look for more bargains and buy less and the rate of inflation reduces. COVID is an interesting example of an instance where there was a dramatic introduction of money into the economy and a loss of production. This resulted in a significant amount of inflation that the Fed themselves has admitted is difficult to model. This is because while things such as food might have experience one amount of inflation and things such as experiences and services saw a different amount of inflation. Thus, the Fed themselves was not entirely convinced that the amount of inflation they reported was accurate since the criteria they used previously to model inflation did not fully reflect spending in society.

Thus, this project would seek to make predictions about inflation based on other factors described above and also perhaps other factors such as age distribution amongst the population and Healthcare as a percentage of GDP. Presumably if healthcare cost increases at a pace greater than GDP then people have less to spend.

Project Idea 1

Company Credit Default Risk

Ideas focused on Credit Default Risk for Companies especially those that are subject to macroeconomic cycles this project could also be used to guide Banks to minimize the degree of cyclical risk in their loan portfolios: Ideally, a bank does not want to expose itself to unintended increased probabilities of loan default from business borrowers in different economic cycles without a decreased probability of default from other customers. By building a model that banks could use to determine the maximum business loan portfolio risk exposure in different economic scenarios such a model could help lead banks to healthier economic states. A bank could choose to balance their portfolio for example by selling to a bank (whether for profit or loss) loans from borrowers in one industry and granting loans (or even buying the loans from another bank whether or not at a premium) to borrowers in another industry so as to minimize loan portfolio default exposure.

The search was completed on Google Dataset Search:

The Initial Dataset search Results:

1. **Business and default cycles for credit risk (replication data):**

https://b2find.dkrz.de/dataset/a5450423-d2c0-5fd6-a951-4e7980964af5

Description from the souce: Various economic theories are available to explain the existence of credit and default cycles. There remains empirical ambiguity, however, as to whether these cycles coincide. Recent papers suggest by their empirical research set-up that they do, or at least that defaults and credit spreads tend to co-move with macroeconomic variables. If true, this is important for credit risk management as well as for regulation and systemic risk management. In this paper, we use 1933-1997 US data on real GDP, credit spreads and business failure rates to shed new light on the empirical evidence. We use a multivariate unobserved components framework to disentangle credit and business cycles. We distinguish two types of cycles in the data, corresponding to periods of around 6 and 11-16 years, respectively. Cyclical co-movements between GDP and business failures mainly arise at the longer frequency. At the higher frequency of 6 years, co-cyclicality is less clear-cut. We also show that spreads reveal a positive and negative co-cyclicality with failure rates and GDP, respectively. This pattern disappears, however, if we concentrate on the post World War II period. We comment on the implications of our findings for credit risk management.

Pros: This report focuses on an important idea the cyclical nature of credit risk. Looks like missing data could be obtained from public sources.

Cons: Dataset is limited to 1933 – 1997 and there may be substantial changes in the present conditions. Especially since companies tend to have larger bank accounts enabling them to weather storms. Although, the missing data could just be retrieved from various sources.

The authors of this report use two main datasets:

Table 1: We use three data series in our analysis of TABLE 1: YEARLY log real GDP, credit spreads, and business failure rates. The third series, real GDP, is taken from the data base of the Federal Reserve Bank of St.Louis (FRED). The series contains GDP in chained 1996 dollars. From the same site, we also obtain Moody's yields on Baa corporate bonds and the yield on government bonds with a maturity exceeding 10 years. These are used to construct annual credit spreads, defined as the difference between the two yields. Our third series is from Dun and Bradstreet (1998) and contains U.S. business failure rates per 10,000 companies over the period 1927-1997.

Table 2: We use three data series in our analysis of TABLE 2: QUARTERLY log real GDP, credit spreads, and business failure rates. The third series, real GDP, is taken from the data base of the Federal Reserve Bank of St.Louis (FRED). The series contains GDP in chained 1996 dollars. From the same site, we also obtain Moody's yields on Baa corporate bonds and the yield on government bonds with a maturity exceeding 10 years. These are used to construct quarterly credit spreads, defined as the difference between the two yields. Our third series is from Dun and Bradstreet (1998) and contains quarterly U.S. business failure rates per 10,000 companies over the period.

Project Idea 2

Credit Card Default Prediction

This search was completed on Kaggle:

Link: https://www.kaggle.com/code/ybifoundation/credit-card-default

Credit Card Default Prediction

The data set consists of 2000 samples from each of two categories. Five variables are: 1. Income, 2. Age, 3. Loan, 4. Loan to Income (engineered feature) and 5. Default

Pros: the dataset includes 2000 samples:

Cons: Credit score is not one of the categories, There is no basic assumption on 401K savings. Thus, there is limited information on the history of the borrower.

Project Idea 3:

Copper Price Prediction

The search was completed on Google Dataset search

https://datasetsearch.research.google.com/search?src=0&query=copper%20price%20prediction&docid=L2cvMTFsY205bDFzZg%3D%3D

Copper is an important industrial metal that has seen quite a bit of price volatility over the last 65 years. It is one of the most volatile commodities in history and further its price volatility is not dependent on events that are difficult to model such as weather patterns or droughts in particular regions. The events that led to price volatility in copper have some degree of predictability and repeatability such as global infrastructure demand, pandemic recovery, technological advancements, geopolitical risk, industrial booms, Global supply imbalance, green technology demand, global recessions, national stockpiling and global economic cycles.

The question would be to what degree can we predict the beginning or end of a price peak or price trough. Also, to what degree can we predict absolute maximum peaks and absolute minimum bottoms in a cycle. We could use a weighted historical behavior analysis to do that. Likely different analytical methods including weighted regression analysis or machine learning methods could be used to tie the influence global events to the price cycles of copper. Data could be broken into cycles and each cycle could be treated as an indivisible element of the data set. We could then have a training dataset and a test dataset. To the degree that global events are unique they may need to be grouped together with some probability of any one of them occurring. For example, there may be one example in history of copper increasing in price due to a war. There may be one example in history of copper increasing in price due to national stockpiling. We can certainly have a probability that each one of these individual events occur to drive volatility. Alternatively, we can group events together and say that the probability of any one of these events occurring in some time window to be of some value determined by out simulation. Arguably, we can likely say there is a certainty that an event that causes volatility will occur in some time window and perhaps other economic factors and/or geopolitical instability could be used to guide which event is most likely to occur next. Finally, we could factor in the threat of substitutes and future outlooks to guide the degree with which some alternative lower cost material could be used in its place in response to increasing price. Furthermore, we could compare our own probabilistic map to using futures to hedge copper prices.

Project Idea 4