Investigating Macroeconomic Conditions during Bank Failures

McKenna Fobes and Max Taylor

2023-11-30

I. Introduction

Topic: What are the macroeconomic conditions when a bank fails.

Why it matters: It is important to see what the macroeconomic conditions are when a bank fails. This will allow for you to have an idea of what happens to the conditions when the bank does fail. This is not a guarantee that a bank will fail but it will allow for you to have a better idea of what could happen when these conditions can be found.

Data: We pulled data from two different sources. This allowed for us to get a more broad range of data.

Bank Failure data: https://banks.data.fdic.gov/bankfind-suite/failures

- Name
- Location
- Date
- Costs and assets

Economic Indicators data: https://fred.stlouisfed.org/graph/?g=3obN#0

- Unemployment rate
- CPI
- Discount rate
- Real GDP.

The economic factors are monthly, GDP is quarterly and the bank failures just happen when the bank failed. The data that we have spans from 1948 till now. The data sets are related to each other because they both have to deal with the financial markets. They will both have an impact on the other. All of the macroeconomic indicators data is in percentages and the Bank data is either categorical or there's numbers that will either be the ID of the bank or in dollars to show the assets that the bank lost.

II. Data Processing

In order to make the data that we had more meaningful we did a few things such as merging based on the months, changing the column names, and removing all of the rows that didn't have all of the macroeconomic data. In order to have one data set instead of two different data sets, we merged based on the months. This allows for everything to be in one row, all of the needed information for when that bank failed would be together. We had problems because there was more than one bank that failed each year in some cases so we merged based on the month and year that the bank failed. In order to fix this we took off the day in the dates and then merged. This allowed for the macroeconomic data to go to the correct banks by the month and year. By renaming all of the column names it allowed for not only us but everyone else who looks at our work to be able to easily know what all of the columns mean. The last thing that we did to make sure that we had a good data set that we can use is that we took out all of the rows that didn't have all of the macroeconomic data. The points that didn't have all of the data weren't going to be useful in the analysis because we are looking at how all of the different macroeconomic factors change when a bank fails.

To ensure thorough data cleaning, we meticulously followed a structured data cleaning checklist. The initial data set was sourced from Excel and read into R. Subsequently, we removed columns that offered little relevance for our deeper analysis.

We took the time to understand the meaning of each variable, which will be discussed in later sections. As the data originated from official government sources, it arrived complete and devoid of any missing entries.

As we brought everything together, we made sure to switch the important columns to dates and numbers. Also, when we looked at the text-based info, we thought it would be helpful to split up the state and city names. This way, we could easily check how many banks failed in each state.

Looking further into the bank data we looked at everything that could be found in the different columns. There's 4 different categorical columns in the data set, they are the insurance fund, the bank type, the transaction type, and the resolution. In the insurance fund there's bank insurance fund (BIF), resolution trust corporation (RTC), federal savings and loan insurance (FSLIC), savings association insurance fund (SAIF), deposit insurance fund (DIF) and FDIC. Within the bank type there's national member banks (N), state member banks(SM), state nonmember banks(NM), savings associations (SA) and saving banks and savings and loans (SB). In the transaction types there's assistance transactions (A/A), reprivatization (REP), purchase and assumptions (PA), insured deposit transfer (IDT), consignment program institution (MGR) and payout. In the resolution column it will tell you if the bank failed or if there was assistance meaning that the bank was bailed out. This allows for us to have a better idea of what the bank failure looked like, not every bank failed in the same way and some of the banks that did fail got bailed out.

The data that we got from the FRED is more intuitive, we changed everything to be a percent which makes looking at the macroeconomic data easier to understand. We wanted to try and get the most broad idea of the market. We felt like these different indicators gave us the best idea to what the market looked like at the national level.

III. Transformations

In order to keep the results easier to understand we decided to not do any transformations. This allows for anyone to look at the graphs and the data and have an idea of what it is telling you. It would be hard to intuitively look at a graph that has the log of an interest rate and know what it is telling you. The only time that we would transform the data would be if we needed it for a graph to show the relationship better. This was a tool that we tried to limit the amount of times that we used. We decided to take the log in some graphs because the data in the graph was left skewed, by taking the log we were able to condense the graph. With there being areas such as California and Texas that had really large numbers and a few other areas that had a very small number.

This missing data was talked about earlier, this was dealt with by taking out the rows that didn't have everything that was needed to complete the analysis. There was also no extreme values that we had to take care of.

The data didn't require much transformations in order to make the data useful and easy to understand. If there was too many transformations, understanding the graphs would've been more difficult, especially without fully understanding what the transformations do to the scale.

IV. Findings and visualizations

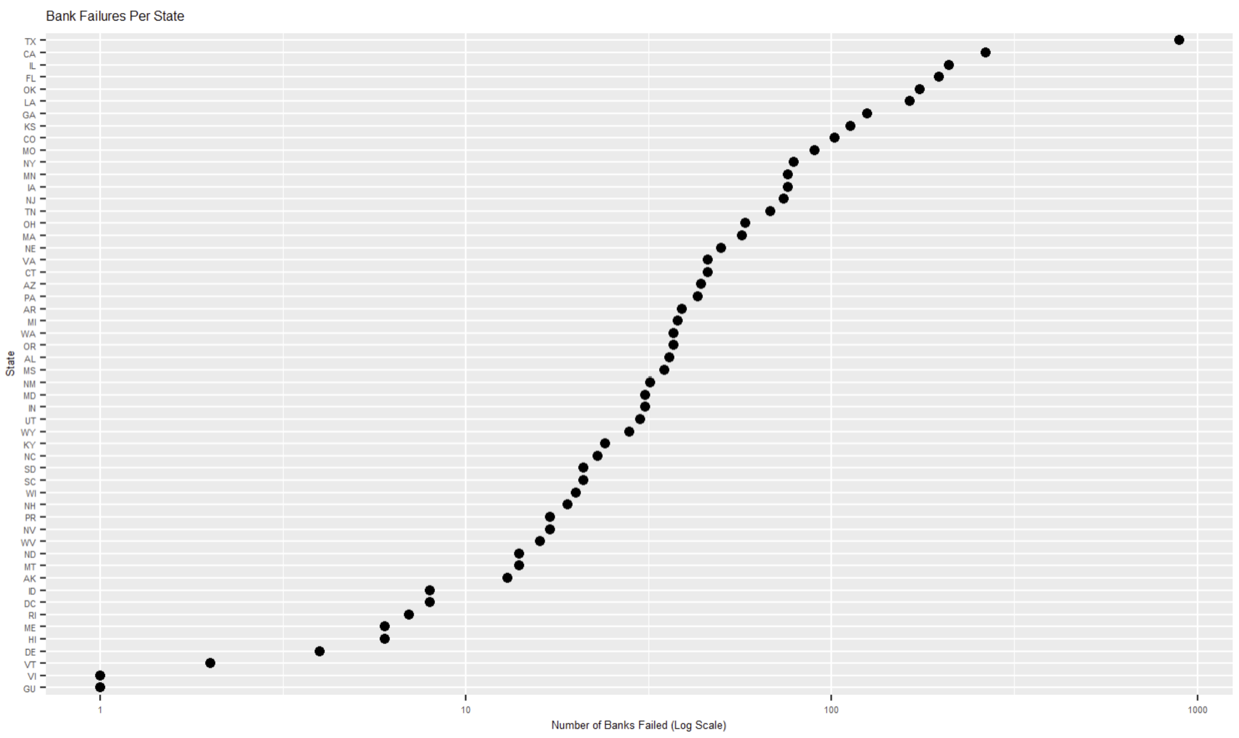

Figure 1: Bank Failures Per State

This graph was used to see if there was any relationship between the banks that failed and the states that the banks failed in. This will give us a better idea of the banks that failed in each of the states. We took the log to show the data in a more meaningful way. The data that was not log transformed didn't allow for us to find a meaningful relationship. The log allowed for the graph to be more condensed, with places like Texas and California that had the most banks that failed, the orginal graph was skewed to the left. From this graph you can see that top 5 state with the most bank failures is Texas, California, Illinois, Florida and Oklahoma.

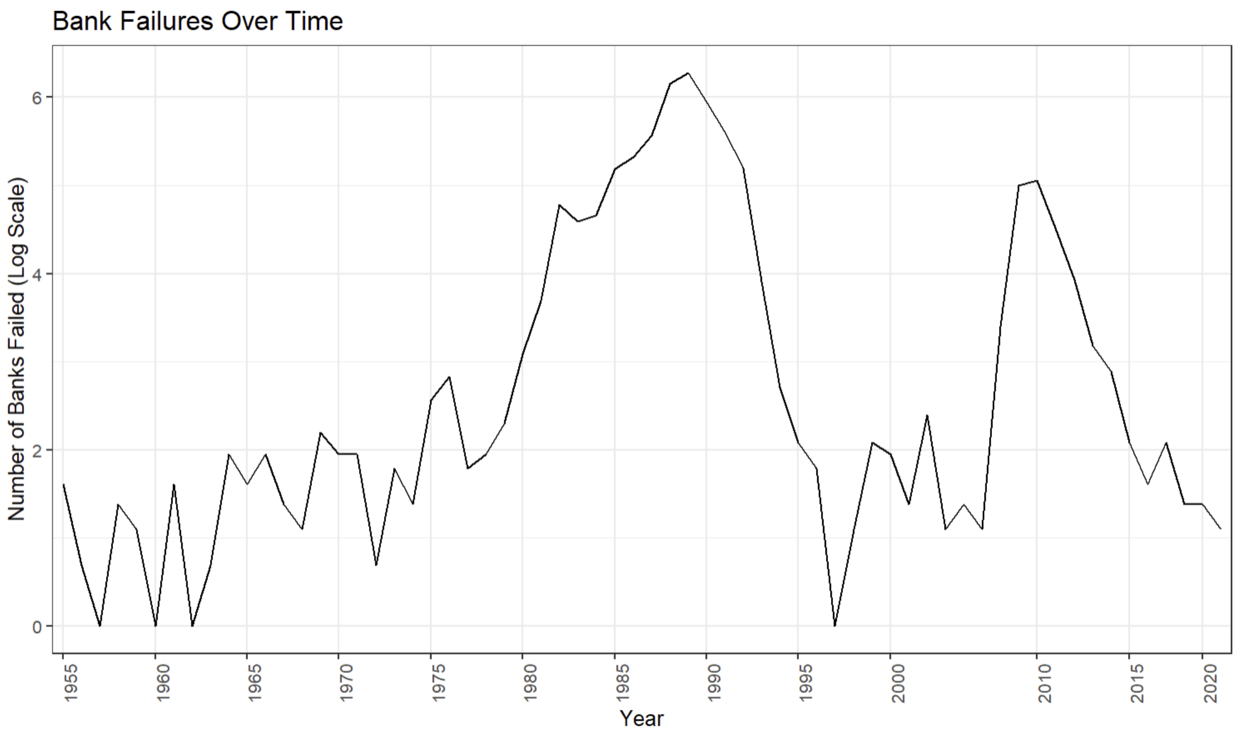

Figure 2: Bank Failures Over Time

We wanted to see if there was a time where more banks failed, so we wanted to plot the data in a way that shows the number of banks over the time. This will allow for us to see if there is a certain time period that we should be looking into closer. Once again the number of banks that failed is in a log scale. The log scale allows for us to look at the trends in the number of bank failures. This graph shows two different peaks where there was significantly more bank failures than all of the other years around 1988 and 2010 have the largest number of bank failures.

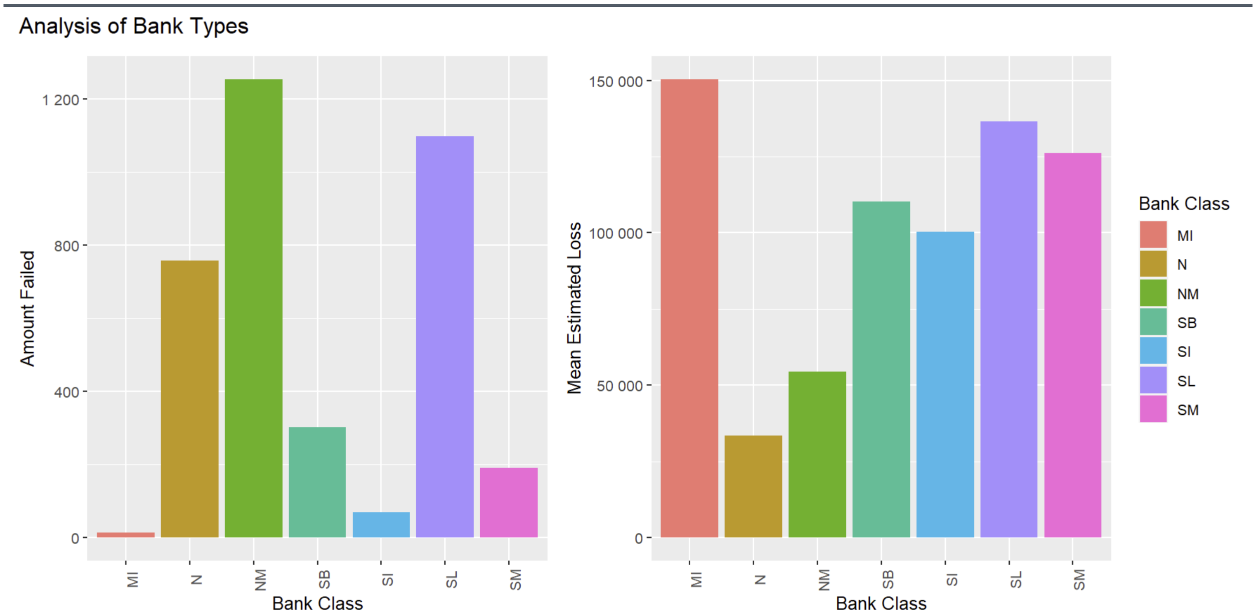

Figure 3: Analysis of Bank Types

In order to get a better idea of what causes the banks to fail, we looked at all of the different kind of banks to see if there was a relationship between the type of bank and how many times the bank failed. The bank that had the lowest number of failures had the highest amount for the expected loss, this could mean that the bank that failed was a larger bank possibly a national level bank. With that type of bank having the largest expected loss this could mean that more was done in order to keep the bank from failing. The type of bank that had the most failures had the lowest expected loss, this means that it most likely was smaller banks. A smaller bank wouldn't have a large expected loss if the bank failed, which could mean that less will be done in order to keep the bank from failing.

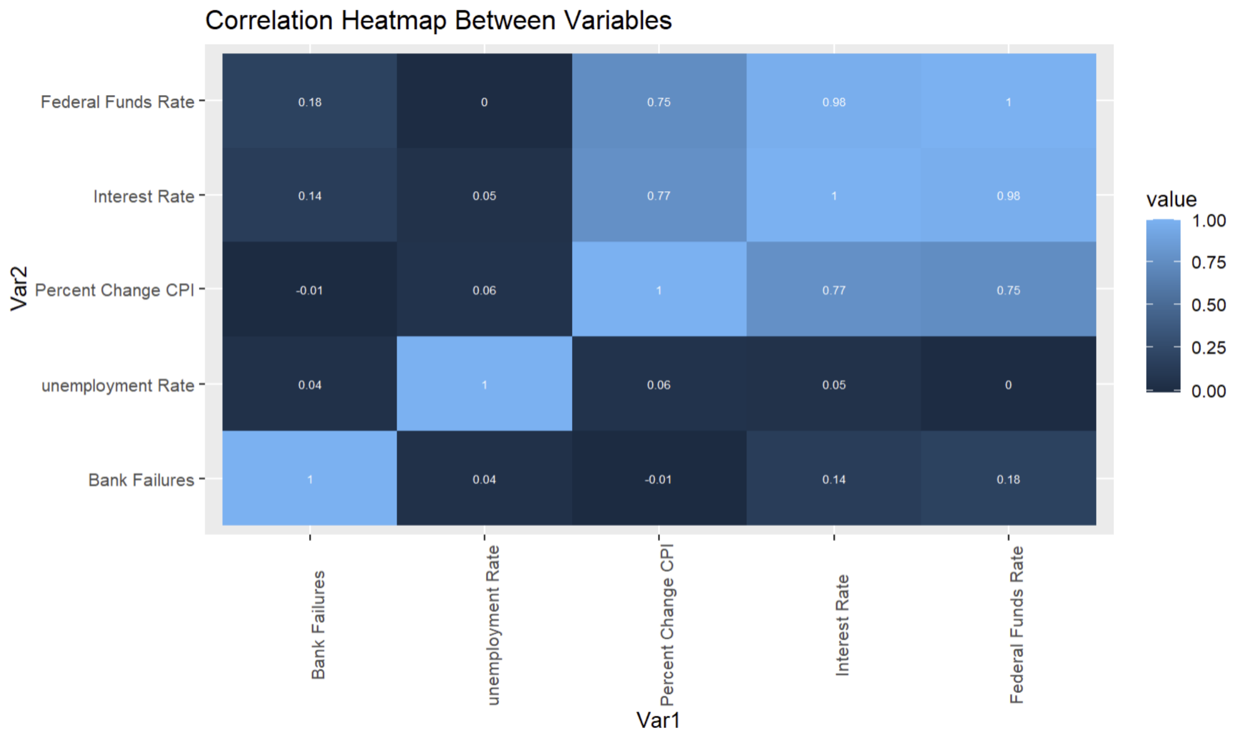

Figure 4: Correlation Plot of All Variables

By looking at the correlation plot this allows us to look at how all of the other variables affect each other. The lighter the box is the more correlated that the variables are with each other. This will allow for us to get a better idea of what macroeconomic factors we should be looking at closer. When looking at the correlation plot, you can see that the highest correlation was between the interest rate and the Federal Funds rate, this had a correlation of 0.98. The percent change CPI and the Interest rate have a correlation of 0.77, the Federal Funds rate and the percent change CPI have a correlation of 0.75 so there is a very similar correlation between the two. The percent change CPI has a very low correlation with the unemployment rate and the Bank failures. The percent change CPI has a 0.06 with the unemployment rate and a -0.01 correlation with the bank failures. This means that these factors have very little relationship between each other, when one changes the other won't change. The macroeconomic condition that has the highest correlation with the Bank Failures is the Federal Funds rate with a correlation of 0.18. This is a very low correlation so there is not a strong relationship between the two.

V. Conclusion

From looking at the data and the graphs that were shown above we were able to see that there was not a macroeconomic condition that has a big impact on when a bank fails. When looking at a national level it is hard to tell when a bank in any of the states would fail. In order to get a better idea of what the true macroeconomic conditions in the state was when a bank fails you would have to pull some of the macroeconomic conditions in the state and then look at the correlation.

A bank failing has other factors than just the macroeconomic conditions that are found nationwide. Some of these factors that are found within the states would not be found in the nationwide macroeconomic conditions. If we were to this again we would look into the specific states or we would look at just tha nationwide banks that would have a stronger relationship with the nationwide macroeconomic conditions. Additionally our analysis, revealed an interesting trend with the characteristic of the failing bank. The expected loss and the type of bank seem to be correlated, this suggests that the larger banks might use different measures to prevent failure.

The visualizations provided insights into the distribution of the bank failures across the states and over time. Texas, California, Illinois, Florida and Oklahoma were the top 5 states with the highest number of bank failures. The peaks in 1988 and 2010 indicated that there was periods of greater bank failure. This would be an area where you could dig in deeper and look at various different conditions.

The correlation plot highlighted relationship between various macroeconomic indicators, with the Federal Funds rate showing the highest correlation with bank failures. However, the overall correlations were relatively low which emphasizes the complexity of the factors of bank failures. This makes predicting when a bank will fail difficult when only looking at the national level macroeconomic conditions.

In conclusion, even though our analysis provides a broad perspective on macroeconomic conditions surrounding bank failures. It shows a need for a more nuanced approach. Understanding the specific state-level economic indicators and the correlations with the bank failures may lead to a more accurate prediction and insights.