Project 1: “The Relationship Between Commercial Banking Assets and the Level of Employment in a State”

Analysis and Folder Organization Completed by: Ryan Lund on 8/13/23

Research Questions:
1) Is there a relationship between level of employment in a State and the amount of commercial banking assets in that State?
2) Is there a relationship between level of employment Nationally and the amount of commercial banking assets Nationally?

Hypothesis:
1) There is a positive correlation between the level of employment in a State and Nationally and the amount of commercial banking assets in a State and Nationally. The reasoning for the positive correlation is straightforward: If there are more employed people in State A than State B, those employed people in State A will generate more assets collectively than the people in State B. The assets generated by employed people will find their way into the banking system and be reflected as commercial banking assets. For example, an employment person buys a home (asset for the individual) and procures a mortgage loan to finance the home purchase from a bank. The mortgage loan for the bank is held on the bank’s balance sheet and classified as an asset.

Data Sources:
1) FRED ('Federal Reserve Economic Data') for total commercial banking assets for Arizona (‘AZ’), New York (‘NY’), Texas (‘TX’), Florida (‘FL’) and Nationally (‘US’) and for level of employment in each respective State and Nationally.
2) Data is in CSV file format and stored in the ‘Resources’ folder in Github.

Methodology:
1) Performed straight line regressions for total commercial banking assets and level of employment in each State and Nationally for men and women separately.
2) Performed ANOVA test to determine if the difference in the mean of commercial banking assets and difference in the mean of level of employment for each State and Nationally is significant.
3) Calculate the distribution of total commercial banking assets as measured by standard deviations from the mean total commercial banking assets for all 5 States.

Analysis Results:
1) Straight lines regressions on the total commercial banking assets and the total level of employment for the 5 States was performed as well as scatter plots that color coded plots by time period (two-year periods beginning from 2003 and ending in 2020) was performed. The regression results showed that the States of CA, TX and FL all reported strong positive correlations (R values ranging from 0.80 up to 0.91) with statistically significant values (P values less than 0.05).

2) The State of AZ showed a week correlation between the variables and NY State effectively showed no correlation with no significance.

3) The colorful scatterplots revealed that all 3 States that showed positive correlations (CA, TX & FL) had lowest commercial banking assets at the lowest level of employment in 2003/2004 and highest commercial banking assets and highest levels of employment in the 2017/2020 time period.

4) Both AZ and NY reported highest commercial banking assets in the 2003/2004 time period and a drop in commercial banking assets in 2007-2010 time period. The visualization suggests that commercial banking assets were highest in NY and AZ at lowest levels of employment in 2003/2004 and dropped in the years around the Great Recession of 2007/2009. Both AZ and NY appear to show an upward trend of commercial banking assets as employment increases beginning in approximately 2014.

5) Straight line regressions on the total commercial banking assets of the US and the total level of employment for Women and total employment for Men both showed a positive correlation between the variables (as demonstrated by the R values of 0.85 and 0.89 respectively) that was statistically significant (as demonstrated by the P values of less than 0.05). The colorful scatterplot also showed that as employment level rose so did commercial banking assets, just as the three States did that showed positive correlations (CA, TX & FL).

6) The trend of the both colorful scatterplots showed that there was a disruption in 2008/2010 which correlates with the Great Recession in which commercial banking assets declined and employment levels dropped.

7) Box plots were created for all 5 States and measured Commercial Banking Assets and Level of Employment. The Commercial Banking Assets box plot shows that NY had the greatest amount of assets while AZ had the least amount of assets. The Employment Level box plot shows that CA has the highest level of employment and AZ has the lowest level of employment. Furthermore, the ranges of the level of employment for the 3 States that showed positive correlations between commercial banking assets and level of employment show similar ranges.

8) Two ANOVA tests on the means of each State for Commercial Banking Assets and the means of each State for Level of Employment showed that the differences between the means are statistically significant.

9) To further investigate the differences between the States, we calculated and compared the amount commercial banking assets calculated as percentage of employment for each State. The resulting bar graph showed that the three States that demonstrated positive correlations between commercial banking assets and level of employment were similar: CA: 3.77%, TX: 4.18% and FL: 7.23%.

10) NY State was the lowest at 1.28%, which combined with the knowledge that NY had billions in commercial banking assets (as compared to millions for the other four States) and did not show a significant correlation between commercial assets and level of employment, the 1.28% appears to be an outlier. AZ State reported the highest at 14.32% which given that that State showed the lowest level of employment and lowest level of commercial banking assets, appeared also to be an outlier.

11) To investigate the potential outliers, we then created a bell curve to show the distribution of the mean Commercial Banking Assets of each State compared the mean for all 5 States. The distribution shows that the three States that report a positive correlation between Commercial Banking Assets and Level of Employment (CA, TX & FL) all were within 2 standard deviations of the mean. The two States that did not report positive correlations were AZ and NY: AZ was close to 3 standard deviations from the mean and NY was not on the bell curve (reminder: NY reported nearly a 0% R value).

Conclusion:

1) In conclusion, our study found support for our hypothesis that commercial banking assets and the level of employment in State is positively correlated in 3 of 5 States: CA, TX and FL. The two States that did not show a positive correlation both showed that commercial banking assets were highest pre-Great Recession at low levels of employment and post-Great Recession showed partial recovery in commercial banking assets against rising level of employment.
2) Commercial banking assets nationally and the level of employment of women and men supported our hypothesis. Further research could be done to expand the scope of the study to include all States in the analysis and consider other variables that may be explanatory including geographic data to isolate concentrations of commercial banking assets and employment in States. A separate study could be performed to regress the relationship between unemployment levels and unemployment rate against commercial banking assets to explore potentially an inverse relationship between the variables.

References:

1) TA – Andrew Smigaj on 7/15/23
2) FRED - https://fred.stlouisfed.org/ on 7/14/23 and 7/18/23