To what extent do key macroeconomic indicators influence the nominal GDP of the United States?

ECON 103L FINAL PROJECT

Manuela Lozano, Sebastian Farah, Miguel Quiñones

2025-03-13

Contents

1 Introduction

Economic growth is a key indicator of a nation's well-being, influencing policy decisions, investment strategies, and social welfare programs. One of the primary measures of economic performance is **nominal Gross Domestic Product (GDP)**, which captures the total value of goods and services produced within a country. Understanding the factors that drive GDP is critical for policymakers, economists, and financial analysts.

This study investigates the impact of key **macroeconomic indicators**—unemployment rate, inflation, the federal funds rate, money supply, and consumer confidence—on the **nominal GDP of the United States**. These variables are widely recognized as fundamental drivers of economic activity. By analyzing **241 monthly observations from 2004 to 2024**, this study aims to determine the extent to which changes in these factors influence GDP.

The **hypothesis** guiding this study is as follows:
"Changes in macroeconomic variables, specifically the unemployment rate, inflation, the federal funds rate, money supply, and consumer confidence, have a significant impact on the nominal Gross Domestic Product (GDP) of the United States."

To test this hypothesis, we employ **multiple linear regression models** to quantify these relationships. Diagnostic tests such as **Variance Inflation Factor (VIF) analysis, Akaike Information Criterion (AIC) selection, and the RESET test** are used to assess model reliability. The results of this study provide insights into the interactions between **monetary policy, consumer sentiment, and economic performance**, which have important implications for forecasting economic trends and designing effective policy interventions.

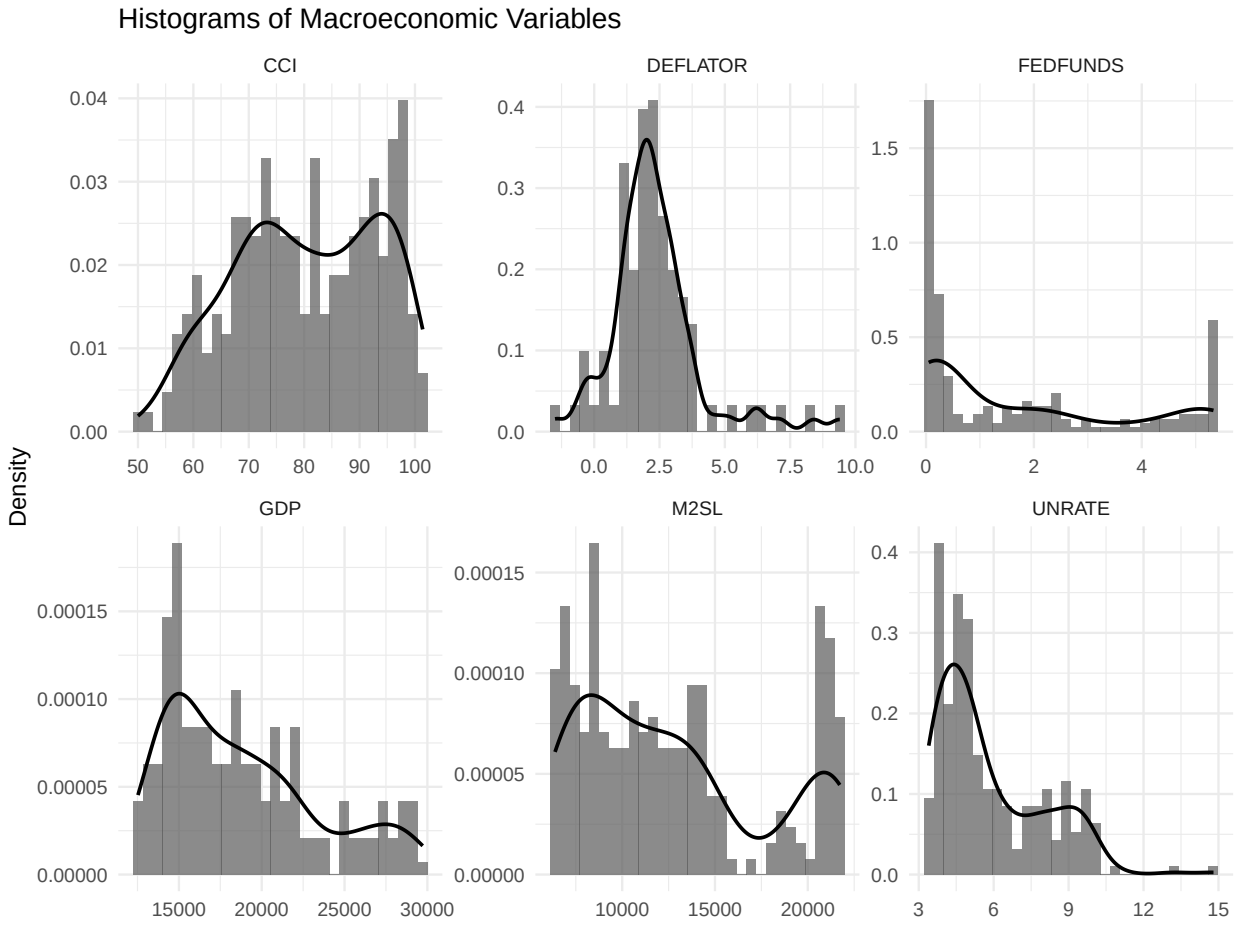
---

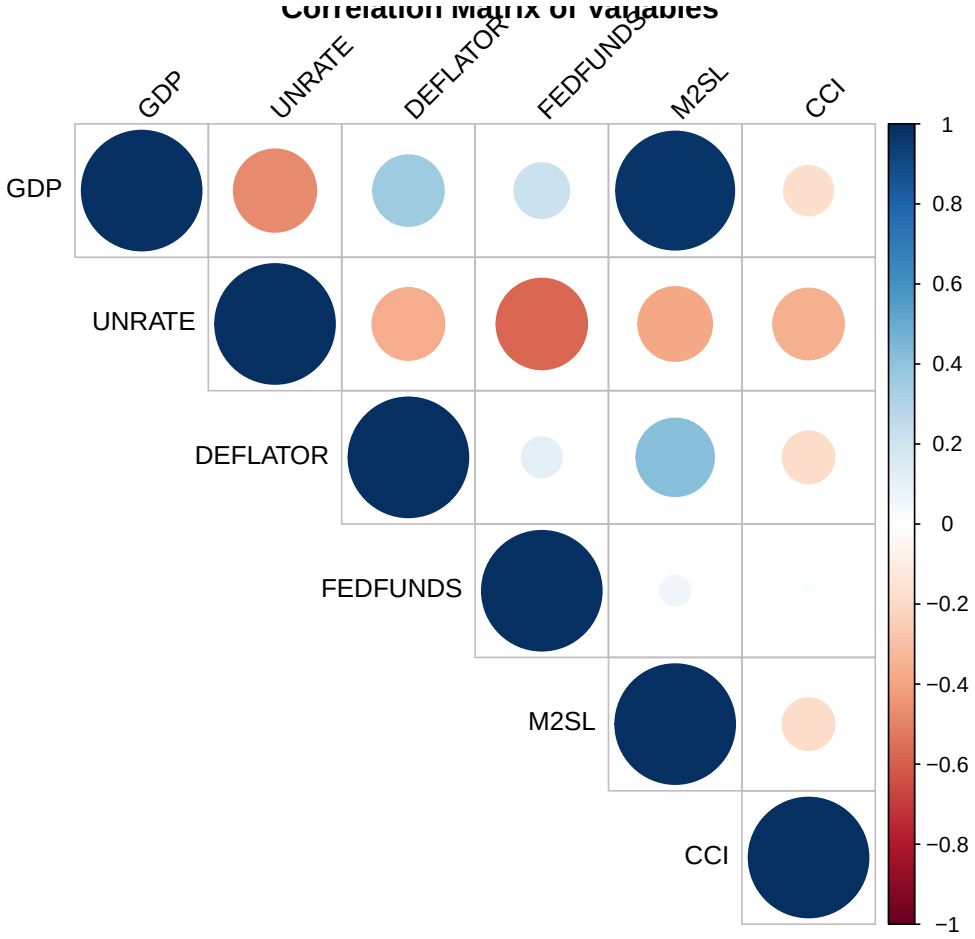
2 Results and Analysis

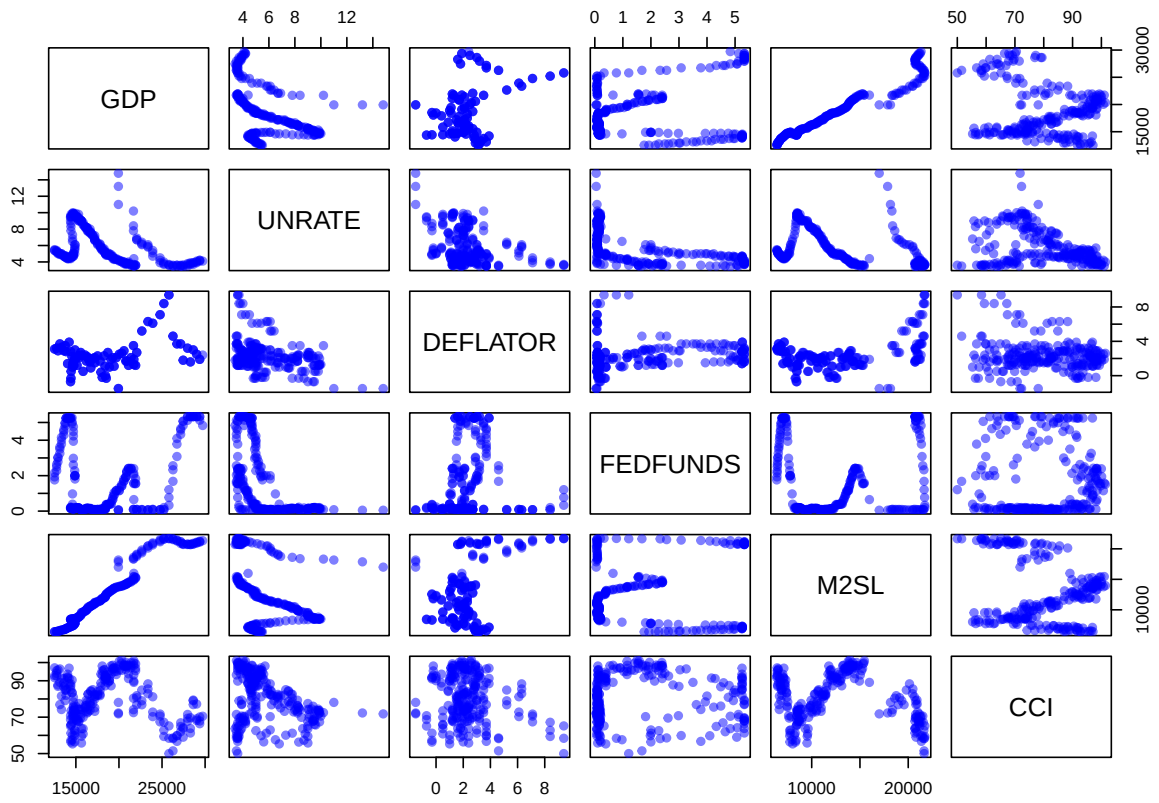
2.1 Understanding the Data

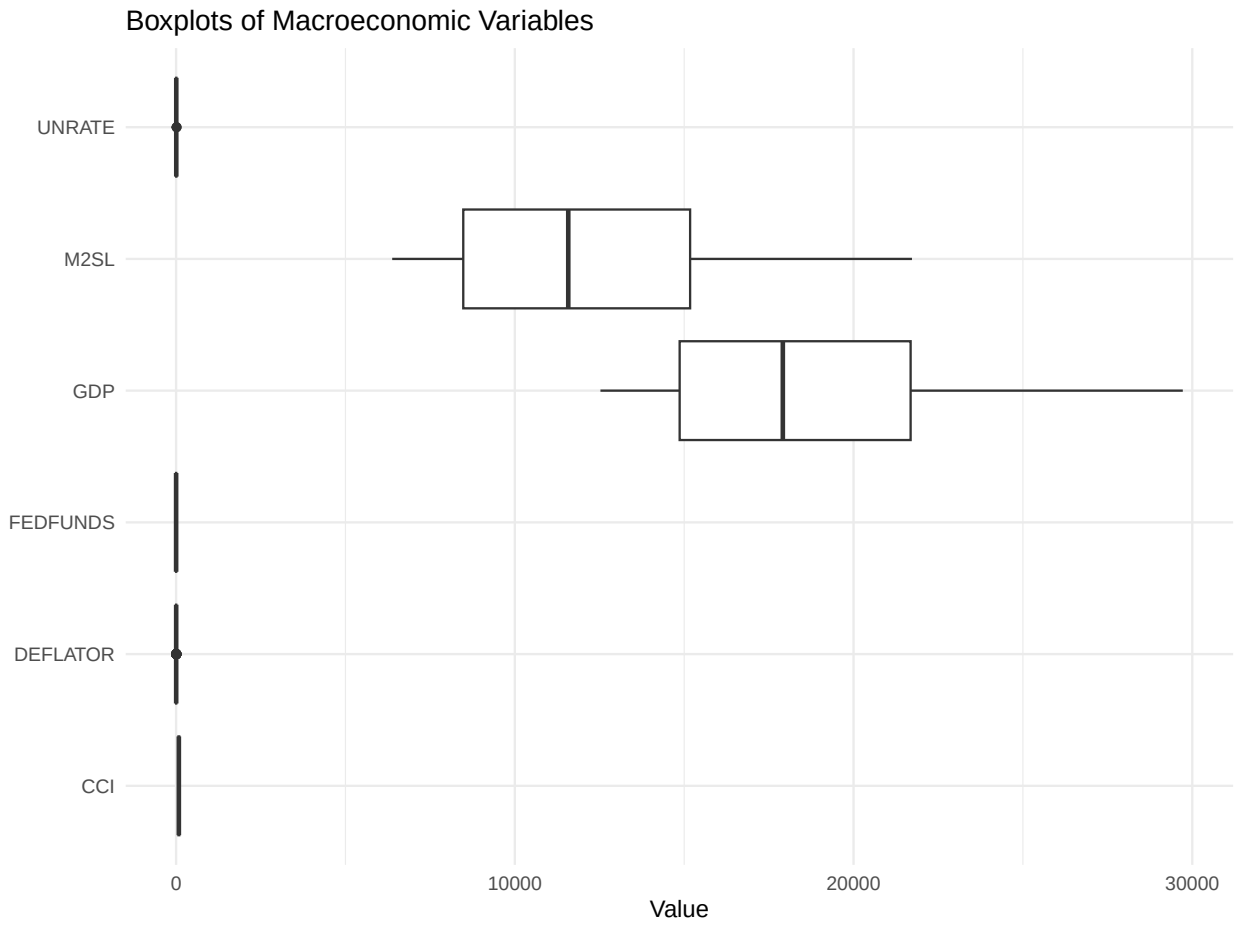
```
## 'data.frame':    241 obs. of  7 variables:
##  $ DATE    : chr  "2004-10-01" "2004-11-01" "2004-12-01" "2005-01-01" ...
##  $ GDP     : num  12527 12527 12527 12767 12767 ...
##  $ UNRATE  : num  5.5 5.4 5.4 5.3 5.4 5.2 5.2 5.1 5 5 ...
##  $ DEFLATOR: num  3.1 3.1 3.1 3.2 3.2 3.2 2.9 2.9 2.9 3.7 ...
##  $ FEDFUNDS: num  1.76 1.93 2.16 2.28 2.5 2.63 2.79 3 3.04 3.26 ...
##  $ M2SL    : num  6380 6406 6425 6431 6439 ...
##  $ CCI     : num  91.7 92.8 97.1 95.5 94.1 92.6 87.7 86.9 96 96.5 ...
```

```
##      DATE                GDP            UNRATE          DEFLATOR      
##  Length:241         Min.   :12527   Min.   : 3.400   Min.   :-1.500  
##  Class :character   1st Qu.:14866   1st Qu.: 4.200   1st Qu.: 1.400  
##  Mode  :character   Median :17912   Median : 5.000   Median : 2.100  
##                     Mean   :18772   Mean   : 5.827   Mean   : 2.337  
##                     3rd Qu.:21685   3rd Qu.: 7.300   3rd Qu.: 2.900  
##                     Max.   :29720   Max.   :14.800   Max.   : 9.400  
##     FEDFUNDS          M2SL            CCI        
##  Min.   :0.050   Min.   : 6380   Min.   : 50.00  
##  1st Qu.:0.120   1st Qu.: 8478   1st Qu.: 70.50  
##  Median :0.400   Median :11578   Median : 80.90  
##  Mean   :1.666   Mean   :12662   Mean   : 80.35  
##  3rd Qu.:2.630   3rd Qu.:15175   3rd Qu.: 91.70  
##  Max.   :5.330   Max.   :21724   Max.   :101.40
```


Histograms of Macroeconomic Variables
CCI
DEFLATOR
FEDFUNDS
GDP
M2SL
UNRATE
Density


Correlation Matrix of Variables
GDP
UNRATE
DEFLATOR
FEDFUNDS
M2SL
CCI
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1


Scatterplot Matrix of Variables
GDP
UNRATE
DEFLATOR
FEDFUNDS
M2SL
CCI



1. **Description of Data:**

Seven variables—nominal GDP, unemployment rate (UNRATE), inflation (GDP deflator), federal funds rate (FEDFUNDS), money supply (M2SL), and consumer confidence index (CCI)—as well as a date variable to monitor changes over time are included in the 241-observation dataset utilized in this study.

A summary statistics analysis was conducted to better understand the distribution of each variable. The nominal GDP values range from 12,527 to 29,720, with a median of 17,912, suggesting significant growth over time. The unemployment rate fluctuates between 3.4% and 14.8%, with an average of 5.83%, reflecting economic cycles, including recessions and recoveries. Inflation, measured using the GDP deflator, varies from -1.5% (indicating deflation) to 9.4%, highlighting periods of both stable prices and inflationary pressures.

With a minimum of 0.05% and a maximum of 5.33%, the federal funds rate, which reflects short-term interest rates set by the Federal Reserve, illustrates changes in monetary policy. With a range of 6,380 to 21,724, the money supply (M2SL) shows significant growth during the investigated period. Last but not least, the consumer confidence index (CCI), which measures changes in consumer sentiment that may affect purchasing decisions, ranges from 50.00 to 101.40.

2. **Description of Histograms:**

The histograms illustrate six key macroeconomic variables: GDP, Unemployment Rate (UNRATE), Inflation (DEFLATOR), Federal Funds Rate (FEDFUNDS), Money Supply (M2SL), and Consumer Confidence Index (CCI). These histograms show the distribution of each variable, with a red density curve overlayed to highlight the overall shape and skewness of the data.

The GDP histogram exhibits a right-skewed distribution, meaning that while some higher values provide a long tail, the majority of GDP values are concentrated at lower levels. With a greater concentration of

lower unemployment rates and a few occurrences of high unemployment (such as during recessions), the unemployment rate (UNRATE) histogram likewise seems to be right-skewed. The somewhat bell-shaped inflation (DEFLATOR) histogram indicates a normal distribution with a small bias toward higher values.

A high left-skew in the federal funds rate (FEDFUNDS) histogram indicates that the majority of values are grouped close to the lower end, most likely representing times when interest rates were low. A right-skewed distribution characterizes the money supply (M2SL) histogram, which shows growth over time as the money supply increases. Lastly, there is a peak in the mid-range values of the consumer confidence index (CCI) histogram, which is comparatively symmetrical and resembles a normal distribution.

3. **Description of Heat Map:**

The correlation matrix of Variables heatmap represents the relationships between the six macroeconomic variables: GDP, Unemployment Rate (UNRATE), Inflation (DEFLATOR), Federal Funds Rate (FEDFUNDS), Money Supply (M2SL), and Consumer Confidence Index (CCI). The color and size of the circles indicate the strength and direction of these relationships, with darker blue shades representing strong positive correlations and red shades indicating negative correlations.

According to the heatmap, GDP shows a strong positive correlation with the money supply (M2SL) and the consumer confidence index (CCI), indicating that economic growth is linked to both an increase in the money supply and a rise in consumer confidence. The projected inverse relationship between economic output and unemployment is further supported by the fact that GDP and the unemployment rate appear to be negatively associated. Similarly, there is a negative association between the CCI and the unemployment rate, suggesting that more unemployment tends to lower consumer confidence. Though it exhibits a minor negative association with GDP and M2SL, the Federal Funds Rate (FEDFUNDS) seems to have lesser relationships with most variables, supporting the notion that rising interest rates might stifle economic growth.

4. **Description of Scatterplot:**

The Scatterplot matrix of Variables allows us to assess potential correlations and nonlinear patterns. In this matrix, the relationship between GDP and the variables shows somewhat linear trends, which is expected since GDP is the dependent variable and should respond to changes in the predictors. For example, the relationship between GDP and M2SL exhibits a clear upward trend, indicating a strong positive correlation, which aligns with economic theory as increased money supply is typically associated with a higher nominal GDP. Similarly, GDP's relation with Unemployment rate shows a downward trend, reflecting the expected inverse relationship between GDP and unemployment, where higher unemployment usually tends to coincide with weaker economic growth. On the other hand, the scatterplots between predictor variables themselves (UNRATE, DEFLATOR, FEDFUNDS, M2SL, CCI), exhibit more scattered and irregular patterns with no clear linear trends, which is a positive outcome. The lack of strong linear relationship between these independent variables reduces the likelihood of multicollinearity, ensuring that each predictor provides distinct and independent information to the regression model. While some variables display mild curvature or clustering, there are no extreme linear patterns that would be indicative of problematic relationships. For example, the scatterplot between UNRATE and FEDFUNDS shows a scattered pattern with no clear linear trend, indicating that changes in the unemployment rate do not move consistently with changes in the federal funds rate, which helps prevent multicollinearity. In a similar way, the relationship between DEFLATOR and CCI appears to be random with no distinct pattern, suggesting that inflation and the consumer confidence index are relatively independent from each other in this dataset.

5. **Description of Box Plots:**

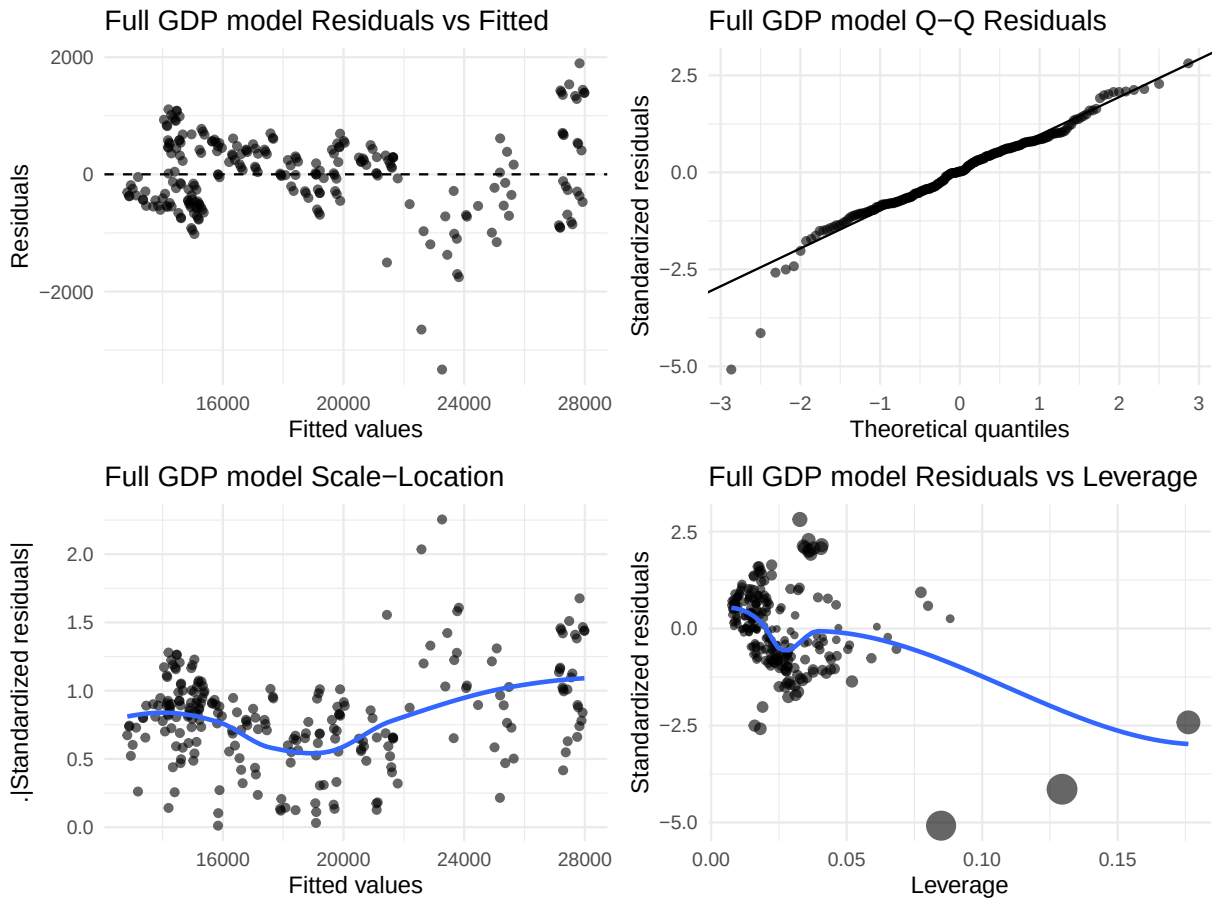
The boxplots provide a visual representation of the distribution, central tendency, and dispersion of each of the variables in the dataset, highlighting potential outliers and the overall spread of the data. The GDP boxplot highlights a relatively right-skewed distribution with no apparent outliers, suggesting that GDP values fall within a reasonable range throughout the dataset. The median GDP value is near the center of the interquartile range, indicating that GDP growth over the time the observations were collected has followed a relatively consistent trajectory, with no pronounced skewness or extreme deviations within the majority of the observations. The boxplot for the unemployment rate shows a somewhat right-skewed distribution

with outliers above the maximum. These outliers are important and represent periods of abnormally high unemployment, likely tied to major economic crises such as the 2019 financial crisis or the COVID pandemic. Furthermore, the boxplot for the inflation deflator displays a right-skewed distribution and appears to be a compacted distribution with various outliers above the maximum. These outliers are reflective of periods of inflation spikes, which may coincide with economic crises or rapid demand fluctuations. The presence of the outliers translate to a skewness to the right, suggesting that inflation may have an unpredictable effect on nominal GDP growth. Moving forward, the boxplot for Federal Funds displays a right-skewed distribution with most data points concentrated near the lower end and gradually extending toward higher values. This pattern is reflective of the FED's typical monetary policy behavior, where interest rates are kept low for extended periods, with occasional increases in response to rising inflation. The absence of outliers in this boxplot suggests that even the highest interest rates fall within an expected range and are not considered extreme enough to be considered outliers. Given the skewness, the regression model may still be influenced by the concentration of low values and occasional higher rates. The M2SL boxplot shows a right-skewed distribution with no apparent outliers. The right-skewness reflects the natural tendency for the money supply to increase over time as the economy expands, resulting in an upward trend that aligns with economic patterns where the supply of money growth is influenced by factors such as monetary policy adjustments, and economic development. Finally, the boxplot for the Consumer Confidence Index (CCI) shows a slightly left-skewed distribution with most data points concentrated toward the higher end of the range and gradually extending toward lower values. The median is above the center of the box, with approximately 75% of the values falling between 70 and 100, while the lower 25% spans a wider range down to about 50. This pattern suggests a tendency for CCI values to cluster at higher levels, yet, the absence of outliers indicates that even the lowest and highest values of CCI fall within an expected range and are not enough to be considered outliers.

2.2 Creating Multi-Linear Regression

```
##
## Call:
## lm(formula = GDP ~ UNRATE + DEFLATOR + FEDFUNDS + M2SL + CCI,
##     data = gdp_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3331.9  -453.3    13.3   441.5  1894.2
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.064e+04  6.461e+02  16.466  < 2e-16 ***
## UNRATE      -2.365e+02  3.626e+01  -6.521 4.24e-10 ***
## DEFLATOR    -3.003e+02  2.986e+01 -10.054  < 2e-16 ***
## FEDFUNDS     2.382e+02  3.112e+01   7.655 4.98e-13 ***
## M2SL         8.915e-01  1.115e-02  79.974  < 2e-16 ***
## CCI         -1.830e+01  4.463e+00  -4.100 5.68e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 685.3 on 235 degrees of freedom
## Multiple R-squared:  0.9782, Adjusted R-squared:  0.9777
## F-statistic:  2108 on 5 and 235 DF,  p-value: < 2.2e-16
```

```
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
```



1. (A) **Analysis of Multilinear Model:**

The initial multiple linear regression model was constructed to analyze the relationship between GDP and five key macroeconomic indicators: unemployment rate (UNRATE), inflation (DEFLATOR), the federal funds rate (FEDFUNDS), money supply (M2SL), and consumer confidence (CCI).

The residuals exhibit a non-random pattern, suggesting potential heteroscedasticity or model misspecification. The presence of clusters and deviations from randomness indicates that the model may not fully capture the underlying relationships between GDP and the predictors. This warrants further investigation through diagnostic tests such as the Breusch-Pagan test for heteroscedasticity and the RESET test for functional form misspecification.

1. (B) **Overall Significance of the Model:**

The overall significance of the regression model was assessed using the F-test, which evaluates whether the predictors collectively explain a significant portion of the variation in GDP. The F-statistic is highly significant ($p < 0.001$), confirming that the model provides meaningful insights into GDP fluctuations based on macroeconomic indicators. Furthermore, the adjusted R^2 value of 0.9777 indicates that nearly 98% of the variation in GDP is explained by the predictors, suggesting a strong fit.

However, the exceptionally high R^2 value raises concerns about potential multicollinearity among predictors, which could inflate the explanatory power of the model artificially. This issue is addressed through Variance Inflation Factor (VIF) analysis in subsequent sections.

2. **Interpretation of Coefficients:**

- Money Supply (M2SL, = 0.8915, $p < 0.001$): The positive and highly significant coefficient indicates

that an increase in the money supply is strongly associated with higher GDP levels. This aligns with economic theory, as increased money circulation typically stimulates economic growth by boosting investment and consumption.

- Federal Funds Rate (FEDFUNDS, = 238.2, $p < 0.001$): Surprisingly, the federal funds rate has a positive relationship with GDP, contrary to conventional economic expectations, which suggest that higher interest rates should slow economic growth. This anomaly may stem from periods where interest rate hikes coincided with strong economic expansion rather than contraction.
- Unemployment Rate (UNRATE, = -236.5, $p < 0.001$): The negative coefficient supports economic theory, indicating that higher unemployment rates are associated with lower GDP due to reduced consumer spending and lower overall production levels.
- Inflation (DEFLATOR, = -300.3, $p < 0.001$): Inflation negatively affects GDP, likely due to reduced purchasing power and increased costs of production. This result aligns with expectations, where excessive inflation leads to economic instability and slower growth.
- Consumer Confidence Index (CCI, = -18.3, $p < 0.001$): The negative coefficient is somewhat counterintuitive, as higher consumer confidence is generally expected to increase GDP through higher consumption. However, this result might suggest that consumer confidence fluctuations do not have a straightforward impact on GDP when accounting for other macroeconomic variables.

Despite the model's high explanatory power, diagnostic tests suggest several potential issues, which are explored further.

3. **Residual vs. Fitted Values Graph Analysis:**

The Residuals vs. Fitted plot reveals several potential concerns regarding the regression model's performance and validity. Most notably, the residuals display a curved pattern, indicating that the relationship between the predictors and GDP is not linear. This curvature suggests potential model misspecifications, meaning that the current model is overlooking important non-linear effects. Another concerning feature is the presence of heteroscedasticity, which is evident from the increasing spread of residuals as fitted values increase (the residuals appear tightly clustered at lower fitted values and become more dispersed as fitted values grow). This pattern violates the assumption of constant variance and homoscedasticity, conditions that must hold to ensure reliable standard errors and efficient coefficient estimates in linear regression. In addition, the plot highlights potential outliers (observations 188, 189, and 241), which significantly deviate from the rest of observations.

4. **Q-Q Plot Graph Analysis:**

The Q-Q plot provides further valuable insights into the distribution of the regression model's residuals and reveals additional potential issues with the model's assumptions. In this plot, most data points, particularly in the middle range, align well with the regression line, suggesting that the residuals follow a roughly normal distribution for typical GDP values. This alignment indicates that the model performs reasonably well in predicting average observations, which is positive in terms of the overall model fit. However, the plot reveals noticeable deviations at both tails of the distribution, raising concerns about non-normality in the residuals.These deviations are particularly evident in the tails, where points deviate more significantly from the regression line. For instance, on the left tail, observations such as 188 and 189 fall significantly below the regression line, indicating outliers in negative residuals suggesting that the model significantly underestimated GDP for these observations which can be connected to periods of economic decline. On the other hand, the observation of 241 on the right tail deviates significantly from the rest, suggesting that the model overestimated GDP for this observation, indicating that the fitted value was considerably higher than the actual GDP. The presence of these influential points in the residual distribution are key indicators that that model may struggle to accurately predict extreme GDP values, potentially reducing the reliability.

5. **Scale-Location Graph Analysis:**

The scale-location plot reveals concerning signs of heteroskedasticity, as indicated by the clear upward slope of the red trend line, which suggests that the spread of residuals increases as fitted values grow larger. While

the residuals appear tightly clustered at lower fitted values, particularly around the 15,000 to 18,000 range, they become increasingly dispersed as fitted values increase, forming a noticeable pattern that reflects a clear non-constant variance. This pattern indicates that the model's residuals exhibit greater variability for higher predicted GDP values, meaning that the model may struggle to provide accurate and stable predictions for larger GDP observations. Additionally, the presence of influential points such as 188, 189, and 241 is concerning, as these points deviate significantly from the rest of observations and are clear outliers in the model. Since these observations coincide with areas where the residual variance is highest, their impact on the regression coefficients is substantial, further reducing the reliability of the model.

6. **Residuals vs. Leverage Graph Analysis:**

We apply the residuals vs leverage plot to identify influential observations that may have an excessive effect on the regression's model's coefficients. In this plot, leverage measures how far an observation's values for the independent variables are from the average, while standardized residuals indicate how far the predicted value is from the actual value. As seen graphically, the majority of points are tightly grouped near the left side, indicating low leverage and relatively small residuals, suggesting that most observations are not exerting an excessive influence on the model. However, there are notable exceptions that have already been discussed previously, particularly observations 187, 188, and 189 which stand out as influential points. These points have higher leverage values and, in some cases, large residuals, positioning them closer to the dashed Cook's distance curves, which mark the threshold for identifying influential data points.Observation 187 appears to be the most concerning, as it combines both relatively high leverage and a large residual, increasing the likelihood that it is disproportionately affecting the model's coefficients. The presence of these influential points is significant since such observations can distort the regression results, making the model's estimates less reliable and potentially skewing the interpretation of the predictor variables' effects.

2.3 Multicollinearity Test (VIF)

```
##   UNRATE DEFLATOR FEDFUNDS     M2SL      CCI
## 3.048832 1.438530 1.829024 1.563454 1.679881
```

Analysis Result of VIF Test: Why we should take out UNRATE

The VIF results indicate that some predictors exhibit moderately high values (3), suggesting multicollinearity may be present. For example:

- Money Supply (M2SL): High VIF values suggest that this variable may be correlated with other predictors, such as inflation or interest rates.
- Federal Funds Rate (FEDFUNDS): Moderate VIF values indicate potential overlap with other monetary policy-related variables.

Multicollinearity can lead to unreliable coefficient estimates and inflated R^2 values. To address this issue, alternative modeling approaches such as principal component analysis (PCA) or stepwise regression may be considered.

2.4 Adjusted Model Without UNRATE

```
##
## Call:
## lm(formula = GDP ~ DEFLATOR + FEDFUNDS + CCI, data = gdp_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7691.1 -3169.3   387.5  2671.0  9137.5
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 19553.24    1863.66  10.492  < 2e-16 ***
```

```
## DEFLATOR      795.61     157.29   5.058 8.49e-07 ***
## FEDFUNDS      426.28     142.76   2.986  0.00312 **
## CCI           -41.69      21.61  -1.929  0.05488 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4221 on 237 degrees of freedom
## Multiple R-squared:  0.1656, Adjusted R-squared:  0.1551
## F-statistic: 15.68 on 3 and 237 DF,  p-value: 2.454e-09
```

Analysis of New Model Without UNRATE:

After removing UNRATE, the revised model included only DEFLATOR, FEDFUNDS, and CCI:

- Adjusted R^2 drops from 0.9777 to 0.1551, showing a significant loss of explanatory power.
- Inflation (DEFLATOR) reverses its coefficient from negative (-300.3) to strongly positive (795.6), indicating that multicollinearity was influencing the original model.
- Consumer confidence (CCI) becomes statistically insignificant ($p = 0.05488$).
- Residual standard error increases dramatically from 685.3 to 4221, indicating a weaker model fit.
- This comparison highlights the trade-off between removing multicollinearity and retaining predictive power.

2.5 AIC or Shwarts Criteria

```
## AIC for different models:
```

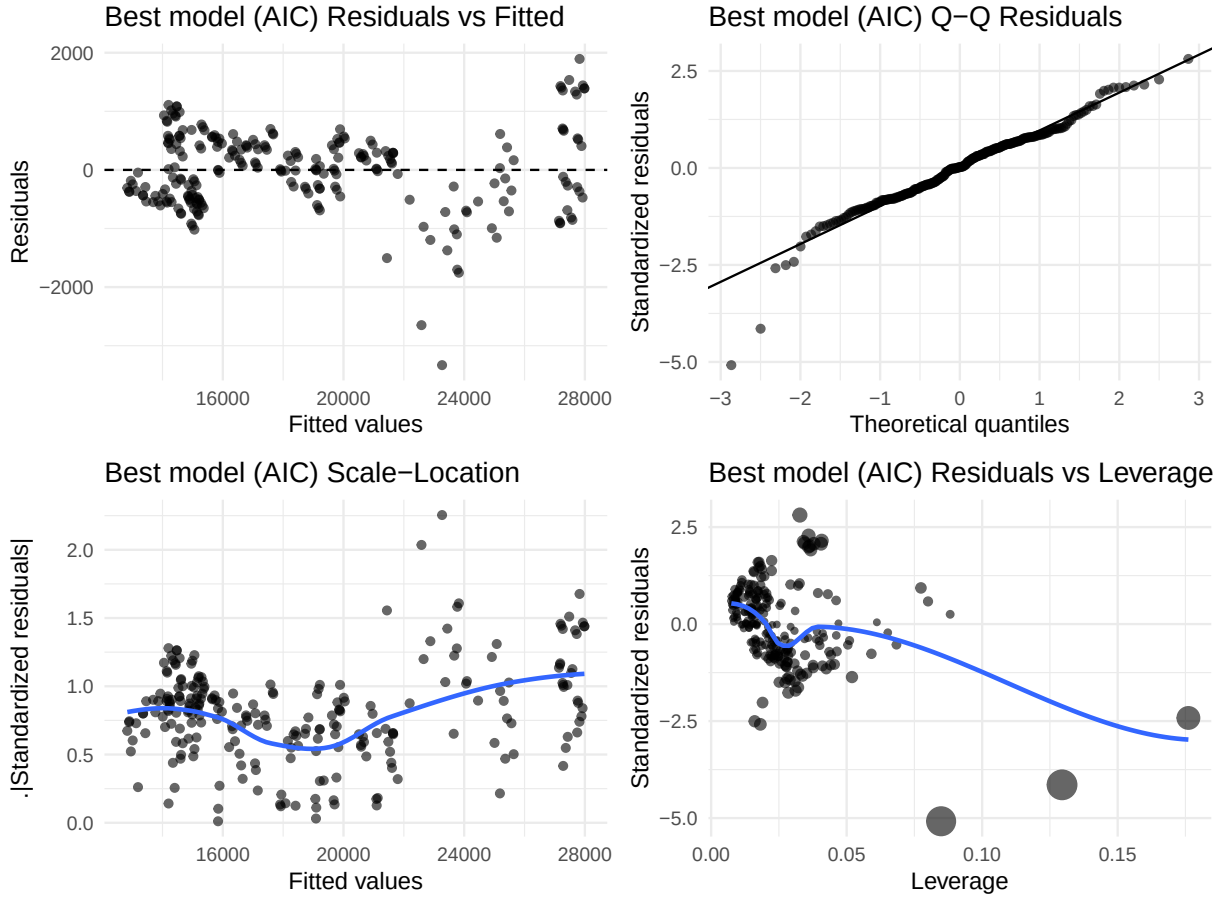
```
##                                                Model      AIC
## 6 GDP ~ UNRATE + DEFLATOR + FEDFUNDS + M2SL + CCI 3839.225
## 4                                     GDP ~ M2SL 4050.332
## 1                                   GDP ~ UNRATE 4691.635
## 7                 GDP ~ DEFLATOR + FEDFUNDS + CCI 4713.513
## 2                                GDP ~ DEFLATOR 4721.625
## 3                                GDP ~ FEDFUNDS 4742.102
## 5                                     GDP ~ CCI 4745.923
```

```
##
## Best model by AIC:
```

```
##
## Call:
## lm(formula = GDP ~ UNRATE + DEFLATOR + FEDFUNDS + M2SL + CCI,
##     data = gdp_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3331.9  -453.3    13.3   441.5  1894.2
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.064e+04  6.461e+02  16.466  < 2e-16 ***
## UNRATE      -2.365e+02  3.626e+01  -6.521 4.24e-10 ***
## DEFLATOR    -3.003e+02  2.986e+01 -10.054  < 2e-16 ***
## FEDFUNDS     2.382e+02  3.112e+01   7.655 4.98e-13 ***
## M2SL         8.915e-01  1.115e-02  79.974  < 2e-16 ***
## CCI         -1.830e+01  4.463e+00  -4.100 5.68e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 685.3 on 235 degrees of freedom
## Multiple R-squared:  0.9782, Adjusted R-squared:  0.9777
## F-statistic:  2108 on 5 and 235 DF,  p-value: < 2.2e-16
```

```
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
```



Analysis of AIC Test:

The Akaike Information Criterion (AIC) test is a statistical method used for model selection, and in this case, it is applied to multiple regression models to determine which one provides the best trade-off between goodness of fit and model complexity.

From the AIC comparison table shown, we see that different models with varying combinations of predictors were assessed. The model with the lowest AIC value is considered the best in terms of balancing fit and simplicity. Here, the model that includes **UNRATE, DEFLATOR, FEDFUNDS, M2SL, and CCI** has the lowest AIC value (3839.225), making it the preferred model.

The **R-squared value is 0.9782**, indicating a strong explanatory power, and the **adjusted R-squared is 0.9777**, suggesting that the model performs well even after adjusting for the number of predictors. The coefficients of the predictors are statistically significant, as seen in the small p-values. This confirms that each variable contributes meaningfully to predicting GDP.

Finally, the residuals vs. fitted values plot for the best model is shown. The residuals are spread somewhat evenly around zero, though some patterns and deviations exist, which might suggest mild heteroscedasticity or other non-linearity. The **red horizontal line at zero** represents an ideal residual distribution, and the

blue smoothing line helps visualize the trend in residuals. While the model is strong, slight deviations in residuals suggest potential areas for further refinement or transformation.

Overall, the AIC analysis confirms that the original model is optimal among the tested alternatives. However, further diagnostic checks might be needed to ensure robustness.

2.6 RESET TEST

```
##
##  RESET test
##
## data:  best_model
## RESET = 55.694, df1 = 10, df2 = 225, p-value < 2.2e-16
```

Analysis of RESTE Test:

The RESET (Regression Equation Specification Error Test) results provide strong evidence of model misspecification in the "best_model." With a remarkably high test statistic of 55.694 (df1 = 10, df2 = 225) and an extremely significant p-value (< 2.2e-16), the test firmly rejects the null hypothesis that the current linear functional form is correctly specified. This suggests that important non-linear relationships exist between the predictors and GDP that aren't captured by the current specification. The test was conducted using powers 2 and 3 (as indicated by power=2:3), meaning that adding squared and cubed terms of the existing predictors would likely significantly improve the model's fit. Despite the high R-squared values observed in the original model, this RESET test reveals fundamental limitations in the linear approach, indicating that a more complex functional form incorporating non-linear transformations would be more appropriate for modeling GDP relationships.

---

3 Findings / Conclusion

The results of this study provide several key insights into the relationship between macroeconomic variables and nominal GDP:

1. Significance of Predictors:

- The Initial regression Model showed that all five macroeconomic indicators had a statistically significant relationship with GDP
- Money Supply (M2SL) and the Federal Funds Rate (FEDFUNDS) exhibited a strong positive impact on GDP
- Unemployment Rate (UNERATE) and Inflation (DEFLATOR) had a negative impact on GDP, as expected from economic theory

2. Model Refinement & Assumption Violations:

- Diagnostic tests revealed issues with heteroscedasticity and non-linearity, indicating that a simple linear model may not fully capture GDP dynamics.
- The RESET test confirmed model misspecification, suggesting that additional non-linear terms or interaction effects should be considered.
- Multicollinearity tests (VIF) indicated that UNRATE should be removed, but doing so significantly reduced explanatory power (R^2 dropped from 0.9782 to 0.1551).

3. Model Selection via AIC:

- The best model, based on AIC, retained all five predictors, despite potential collinearity concerns, as it provided the best trade-off between complexity and predictive power.

4. Policy Implications:

- The strong relationship between money supply and GDP underscores the role of monetary policy in influencing economic growth.
- The negative impact of unemployment and inflation on GDP suggests that economic downturns and high inflationary periods significantly reduce economic output.
- The findings highlight the need for policymakers to consider macroeconomic stability measures when designing fiscal and monetary policies.

4 Final Thoughts

Overall, this study confirms that macroeconomic indicators play a crucial role in determining GDP fluctuations in the U.S.. However, the analysis suggests that a more complex model incorporating non-linear effects might improve predictions. Future research could explore alternative modeling techniques, such as time-series analysis or machine learning approaches, to enhance predictive accuracy.

---

5 References

- GDP Data: https://fred.stlouisfed.org/series/GDP
- Unemployment Rate Data: https://fred.stlouisfed.org/series/UNRATE
- GDP Deflator Data: https://fred.stlouisfed.org/series/A191RI1Q225SBEA
- Federal Funds Rate Data: https://fred.stlouisfed.org/series/DFF
- Money Supply Data: https://fred.stlouisfed.org/series/WM2NS
- Consumer Confidence Index Data: https://www.oecd.org/en/data/indicators/consumer-confidence-index-cci.html?oecdcontrol-b2a0dbca4d-var3=2001-01&oecdcontrol-b2a0dbca4d-var4=2024-11