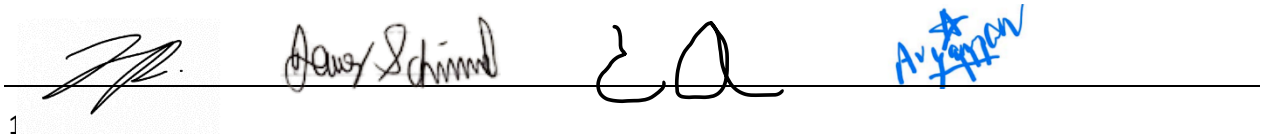
Forecasting United States GDP: A Comprehensive Econometric Analysis Using Linear Regression

James Schimmel, Jung Park, Christopher Ambrus, Aryaman Utkarsh Gangwar

We have neither given nor received help (apart from the instructor) to complete this assignment.

Signature:



Abstract

This econometrics paper explores the intricate dynamics influencing the United States Gross Domestic Product (GDP) through a comprehensive linear regression analysis. Covering the extensive time span from 1966 to April 2023, the study incorporates a diverse set of variables reflecting various aspects of the U.S. economy. The examined variables include demographic factors such as Population (in thousands), economic indicators like Military Spending (in billions), and socio-economic metrics like Noncyclical Unemployment Rate. The analysis further incorporates government-related variables like Government Spending (in billions), Fed Funds Rate, and Inflation modeled using the Consumer Price Index (CPI). Additionally, the study considers external factors such as Net Exports (in billions), Foreign Investment Inflow (in millions), and geopolitical considerations represented by the Net Importer of Oil and the At War categorical variable. The political landscape is accounted for through variables like Presidential Party affiliation, Unified Government status, and the SP500 Index. The findings of the linear regression analysis highlight the significance of specific variables in predicting U.S. GDP. Notably, Population, Military Spending, Noncyclical Unemployment Rate, Government Spending, Net Exports, Government Debt, Labor Force Participation Rate, Net Importer of Oil, Home Ownership Rate, Consumer Sentiment, Total Capital Expenditures, and U.S. Wealth emerged as crucial determinants with a 99% level of significance. This research contributes valuable insights for policymakers, economists, and stakeholders seeking a nuanced understanding of the factors shaping the U.S. economy. The robustness of the model, validated by the significance of key variables, enhances its utility for forecasting and decision-making in the ever-evolving economic landscape.

1. Introduction

This research delves into the intricate task of forecasting the Gross Domestic Product (GDP) of the United States through a comprehensive econometric analysis. Spanning from 1966 to April 2023, the study takes a holistic approach by incorporating a diverse set of economic, demographic, and geopolitical variables. With a keen focus on understanding the multifaceted factors influencing the nation's economic output, the investigation aims to contribute valuable insights to the realm of economic forecasting and decision-making.

The background of this research lies in the recognition of the ever-evolving economic landscape of the United States. As economic conditions continually shift, the need for accurate and nuanced GDP predictions becomes paramount for policymakers, economists, and stakeholders. Against this backdrop, the objective of the study is twofold: to comprehensively analyze the factors influencing U.S. GDP and to develop a robust forecasting model using linear regression.

Key Components and Variable Selection:

The heart of the study lies in its meticulous selection of variables, covering demographic indicators like population, economic metrics such as military spending, and geopolitical factors like the net Importer of oil. The chosen methodology, linear regression, allows for a quantitative analysis of the relationships between these variables and U.S. GDP. This approach provides a nuanced understanding of the specific contributions of each factor to the variations observed in the nation's economic output.

Time Period and Significance:
Analyzing data from 1966 to April 2023, the study embraces a substantial time frame that captures both long-term trends and short-term fluctuations in the U.S. economy. Notably, the research identifies specific variables, including population, military Spending, noncyclical unemployment rate, government spending, net exports, and others, as crucial determinants with a 99% level of significance in predicting U.S. GDP.

Implications and Conclusion:
The outcomes of this research hold significant implications for stakeholders involved in economic forecasting and policymaking. By offering insights into the complex web of factors shaping the U.S. economy, the study contributes to more informed policy formulation and strategic planning. In conclusion, this report provides a comprehensive overview of the background, objectives, and key components of a research endeavor focused on forecasting United States GDP.

2. Description of the Model

Our model is

$$z = \beta_0 + \beta_1 X_1 + \ldots + \beta_{23} X_{23}, \qquad (1)$$

where the variables are listed in Table 1 below. All data corresponds to quarterly observed data from January 1st, 1966 to April 1st, 2023.

3. Discussion of Assumptions

In model 1, for the purpose of a statistical analysis, it is required that the data satisfies the Gauss-Markov assumptions.

(1.) Linearity
Since we are using multivariate linear regression, the model is linear by construction.

(2.) Independence and Randomness of Error Terms
Looking at the residual plots (see fig. 1-23, Appendix), there is a visible amplifier oscillator shape in the residual graph for Population(Fig 1). The other residual plots show no clear shape. Therefore, we conclude that the residuals are independent and random. We recommend rerunning regression omitting Population variable to test for better Independence and Randomness of Error Terms.

Table 1. Description of the Variables

Short Notation	Variable	Description	Source
X_1	Population	The population of the United States in thousands	(13)
X_2	Military Spending	The amount the US government spends on it's military in billions	(19)
X_3	Noncyclical Unemployment Rate	The percentage of people who are unemployed due to not being able to find a job. Equal to Unemployment – Cyclical Unemployment	(18)
X_4	Government Spending	The total expenses of the US government in billions	(14)
X_5	Fed Funds Rate	The rate that banks pay for overnight lending set by the Federal Reserve	(11)
X_6	Inflation:	Change in the value of USD. Modeled using CPI which is a level measure of the price of a basket of goods of a base year (1983) compared to another year	(9)
X_7	Oil Price	Crude Oil spot price	(8)
X_8	Net Exports	The United States' total exports minus its total imports in billions	(16)
X_9	Government Debt	The total public debt of the United States measured in millions of dollars	(15)
X_{10}	Labor Force Participation Rate	The percentage of the total population of the United States that is employed as part of the workforce	(10)
X_{11}	Net Importer of Oil	Categorical variable that is equal to 1 when the US is a net importer of oil and 0 otherwise	(4)
X_{12}	Presidential Party IS Republican	Categorical variable that is equal to 1 if the president's political party is Republican and 0 otherwise	(2)
X_{13}	Home Ownership Rate	The proportion of households that is owner-occupied	(20)
X_{14}	At War	Categorical variable that is equal to 1 when the United States is in an official war or special military operation and 0 otherwise	(22)
X_{15}	SP500 Index	The numerical value of the SP500 on that given day	(23)
X_{16}	Foreign Investment Inflow	The flow of capital into the US from a foreign firm measured in millions of dollars	(17)
X_{17}	Consumer Sentiment	An index that measures how optimistic consumers feel about their finances and the state of the US economy	(12)
X_{18}	Federal Government Tax Revenue	The amount of money the United States Federal government receives in taxes measured in billions	(24)
X_{19}	Industrial Index:	The real output of all relevant establishments located in the United States	(21)
X_{20}	Unified Government	A categorical variable that is equal to 1 when the political party of the presidency matches the party that holds majority in congress and 0 otherwise	(3)
X_{21}	Rental Vacancy Rates	The proportion of the nation's rental inventory that is vacant for rent	(7)
X_{22}	Total Capital Expenditures	The total money invested by US companies in acquiring, maintaining, or improving fixed assets such as property, equipment, and technology measured in millions	(6)
X_{23}	US Wealth	The total net worth of the United States from all sectors measured in millions	(5)
z	US GDP	United States Gross Domestic Product in billions	(1)

For additional testing, we use Durbin-Watson test.

Consider
$E_t = BE_{t-1} + W_t$,

where B is deterministic coefficient and W_t is stochastic process. We test the hypothesis that

H_0: B=0
H_1: B is not equal to 0.

The test statistic is 0.96579, and P-value less than 2.2e-16. With P-Value almost zero we can say there is no evidence that error terms are positively linearly autocorrelated with lag 1.

(3.) No Perfect Multicollinearity
Since R was able to calculate the estimated coefficients without reporting a singularity, we have no perfect collinearity.

(4.) Zero Conditional Mean
By construction, the sum of residuals for the model is 0. Thus, the average of residuals is also 0. Since we use residuals to estimate errors, this assumption is considered to hold for any data set including this one.

(5.) Homoscedasticity of Error Terms
The residual plots (see fig. 1-23, Appendix) shows that variability of error terms is bounded in a corridor except for couple points. Therefore, we do not observe any serious violation of homoscedasticity in error terms.

For additional testing, we perform Breusch-Pagan test to check if there exists conditional heteroscedasticity of linear type in our model.

Within the model
$E^2 = B_0' + B_1'*X_1 + \ldots + B_{23}'*X_{23}$,

where E^2 is the squared residuals, and $X_1, \ldots, X_{23}$ are the original variables.

We test the hypothesis that,
H_0: $B_1', \ldots, B_{23}'$ are all zero
H_1: At least one of $B_1', \ldots, B_{23}'$ is not zero.

With alpha = 0.01, the test statistic is 145.57 with a p-value of less than 2.2e-16. We reject the null hypothesis at significance level 99%. There is evidence of conditional heteroscedasticity of

the linear type in our model. This result means that our model may be unreliable. For the robustness of our testing, we perform the White Test.

Within the model
$E^2 = \beta_0` + \beta_1`X_1 + \ldots + \beta_{23}`X_{23} + \beta_{24}`X_1^2 + \ldots + \beta_{46}`X_{12}^2 + \beta_{47}`X_1X_2 + \ldots + \beta_{322}`X_{11}X_{12}$,

where E^2 is the squared residual and $X_1, \ldots, X_{23}$ are the original variables, we test the hypothesis that,
H_0: $B_1' = \ldots = B_{322}' = 0$
H_1: at least one Bi' is not equal to zero.

With alpha = 0.01, the test statistic is 37.56 with p-value equal to 0. Based on these results there is enough evidence to conclude there is no conditional heteroscedasticity of the quadratic type in the model.

(6.) Normality of Error Terms
The normal probability plot (See figure 25, Appendix) is a straight line apart from a few outliers. With only a few outliers and over 200 data points we conclude that the normality assumption is not seriously violated.

Based on the results of the Gauss-Markov assumptions we see that one residual plot fails the independence and randomness test and the data fails the Breusch-Pagan for conditional heteroscedasticity of the linear type. We recommend the use a time series Garch model to compensate for the heteroscedasticity. Alternative recommendations include the use robust standard errors so that the heteroscedasticity does not affect the estimates, or transforming the data using a logarithmic approach. The model passes all further tests indicating that it is still useful for predictive regression but the results should be considered to be potentially inaccurate.

4. Analysis of Significance of Regressors

Within Model (1), we test the hypothesis that
H_0: $\beta_1 = \ldots = \beta_{23} = 0$
H_1: β_i is not equal to 0, for at least one i = 1, …, 23

to test the join significance of the independent variables.

The p-values, corresponding to the t-statistics and the f-statistic for the independent variables in model 1 can be seen in Table 2 below. Thus, based on these results we reject the null hypothesis

and state that the model is jointly significant. R^2 adjusted for this model is 0.9996, meaning that this model explains 99.06% of variation in US GDP.

Table 2, Output from Model (1) Regression

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-1.954e+02	2.548e+03	-0.077	0.93894	
`Population (thousands)`	2.323e-02	7.626e-03	3.046	0.00262	**
`Military Spending (billions)`	3.789e+00	4.660e-01	8.131	3.93e-14	***
`Noncyclical Unemployment Rate`	-3.005e+02	1.373e+02	-2.188	0.02977	*
`Government Spending (billions)`	-3.317e-01	4.912e-02	-6.754	1.44e-10	***
`Fed Funds Rate`	-7.889e+00	6.346e+00	-1.243	0.21518	
`Inflation (CPI-Level)`	1.128e+01	4.899e+00	2.304	0.02224	*
`Oil Price (WTI Crude Spot Price)`	-1.828e+00	1.530e+00	-1.195	0.23344	
`Net Exports (billions)`	1.806e+00	3.174e-01	5.689	4.35e-08	***
`Government Debt (millions)`	1.165e-04	2.634e-05	4.423	1.58e-05	***
`Labor Force Participation Rate (Percentage)`	-1.520e+02	2.992e+01	-5.079	8.46e-07	***
`Net Importer of Oil`	8.042e+02	1.680e+02	4.788	3.21e-06	***
`Presidential Party IS Republican`	-6.677e+01	3.885e+01	-1.719	0.08718	.
`Home Ownership Rate`	9.532e+01	1.932e+01	4.934	1.66e-06	***
`At War`	1.632e+01	4.334e+01	0.377	0.70689	
`SP500 Index`	1.462e-01	1.033e-01	1.415	0.15860	
`Foreign Investment Inflow (millions)`	1.662e-04	1.144e-04	1.453	0.14784	
`Consumer Sentiment`	-8.481e+00	1.394e+00	-6.086	5.58e-09	***
`Federal Government Tax Revenue (billions)`	-2.203e-01	2.221e-01	-0.992	0.32255	
`Industrial Index`	-5.682e+00	8.350e+00	-0.681	0.49695	
`Unified Government`	1.458e+01	3.238e+01	0.450	0.65288	
`Rental Vacancy Rates (Percentage)`	1.096e+01	2.303e+01	0.476	0.63448	
`Total Capital Expenditures (millions)`	1.618e-03	2.001e-04	8.084	5.27e-14	***
`US Wealth (millions)`	3.723e-05	7.332e-06	5.078	8.53e-07	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 144.6 on 206 degrees of freedom
Multiple R-squared: 0.9996, Adjusted R-squared: 0.9996
F-statistic: 2.403e+04 on 23 and 206 DF, p-value: < 2.2e-16

Testing for individual significance of the variables we test the hypothesis,
H_0: $\beta_i = 0$, for i=1,…,23
H_1: β_i is not equal to 0, for i = 1, …, 23.

From the results in table 2 we see that for i=1, 2, 3, 4, 8, 9, 10, 11, 13, 17, 22, 23 we reject the null hypothesis and state that these variables are significant.

Based on the results of the individual significance hypothesis testing, we update the model to model 2.

$$z = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6 + \beta_7X_7 + \beta_8X_8 + \beta_9X_9 + \beta_{10}X_{10} + \beta_{11}X_{11} + \beta_{12}X_{12}, \quad \textbf{(2)}$$

where,
z = GDP Forecast (in billions),
X_1 – Population (in thousands)
X_2 – Military Spending (in billions),
X_3 - Noncyclical Unemployment Rate
X_4 - Government Spending (billions)
X_5 - Net Exports (billions)
X_6 - Government Debt (millions)
X_7 - Labor Force Participation Rate
X_8 - Net Importer of Oil (1 if net importer)
X_9 - Home Ownership Rate
X_{10} - Consumer Sentiment
X_{11} - Total Capital Expenditures (millions)
X_{12} - US Wealth (millions).

The p-values for the independent variables can be seen below in table 3. At alpha 0.01 all the variables are statistically significant and R^2 adjusted remains at 0.9996. Due to the benefits of needing less data and keeping all forecasting ability we recommend model 2 for predictions.

Table 3, Output of Model (2) Regression

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-3.713e+03	1.580e+03	-2.350	0.0197	*
`Population (thousands)`	3.106e-02	3.157e-03	9.840	< 2e-16	***
`Military Spending (billions)`	3.970e+00	2.795e-01	14.200	< 2e-16	***
`Noncyclical Unemployment Rate`	-5.559e+02	6.722e+01	-8.270	1.34e-14	***
`Government Spending (billions)`	-3.687e-01	4.686e-02	-7.869	1.68e-13	***
`Net Exports (billions)`	1.711e+00	2.599e-01	6.583	3.42e-10	***
`Government Debt (millions)`	1.641e-04	2.231e-05	7.354	3.89e-12	***
`Labor Force Participation Rate (Percentage)`	-8.580e+01	1.951e+01	-4.397	1.72e-05	***
`Net Importer of Oil`	6.254e+02	1.399e+02	4.471	1.25e-05	***
`Home Ownership Rate`	9.041e+01	1.573e+01	5.749	3.02e-08	***
`Consumer Sentiment`	-6.624e+00	1.212e+00	-5.467	1.25e-07	***
`Total Capital Expenditures (millions)`	1.398e-03	1.440e-04	9.707	< 2e-16	***
`US Wealth (millions)`	4.382e-05	5.569e-06	7.869	1.68e-13	***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 149 on 217 degrees of freedom
Multiple R-squared: 0.9996, Adjusted R-squared: 0.9996
F-statistic: 4.338e+04 on 12 and 217 DF, p-value: < 2.2e-16

5. Conclusion

In our study on the U.S. economy, we aimed to predict the path the United States Gross Domestic Product (GDP) through an linear regression model spanning from 1966 to 2023. Our study incorporated a set of both numerical and categorial variables encompassing demographic, economic, geopolitical, and political factors.

The selection of variables and the use of linear regression facilitated a quantitative analysis of the relationships between these factors and U.S. GDP. Our findings show the significance of specific variables, including Population, Military Spending, Noncyclical Unemployment Rate, Government Spending, Net Exports, Government Debt, Labor Force Participation Rate, Net Importer of Oil, Home Ownership Rate, Consumer Sentiment, Total Capital Expenditures, and U.S. Wealth. These determinants emerged as the statistically significant contributors to U.S. GDP.

While our model had a very high R^2, it is important for us to acknowledge its drawbacks. Mainly, the Breusch-Pagan test revealed evidence of conditional heteroscedasticity of the linear type, raising concerns about the reliability of the model. Further testing was performed, including the White Test, to assess the failure and found no further issues. We recommend the Garch time series model, the use of robust standard errors, or the transformation of the data through a logarithmic process to address this issue. Despite these challenges, the model remains a valuable tool for forecasting, provided users are mindful of its potential inaccuracies.

In conclusion, this research provides a comprehensive overview of our approach, objectives, and key findings in forecasting the United States GDP.

References

1. U.S. Bureau of Economic Analysis, Gross Domestic Product [GDP], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/GDP, December 1, 2023.

2. Wikipedia contributors. "Political parties in the United States." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 28 Nov. 2023. Web. 1 Dec. 2023.

3. Wikipedia contributors. "Divided government in the United States." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 13 Mar. 2023. Web. 1 Dec. 2023.

4. “U.S. Energy Information Administration - EIA - Independent Statistics and Analysis.” U.S. Energy Facts - Data and Statistics - U.S. Energy Information Administration (EIA), www.eia.gov/energyexplained/us-energy-facts/data-and-statistics.php. Accessed 1 Dec. 2023.

5. Board of Governors of the Federal Reserve System (US), All Sectors; U.S. Wealth, Level [BOGZ1FL892090005Q], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/BOGZ1FL892090005Q, December 1, 2023.

6. Board of Governors of the Federal Reserve System (US), All Sectors; Total Capital Expenditures, Transactions [BOGZ1FA895050005Q], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/BOGZ1FA895050005Q, December 1, 2023.

7. U.S. Census Bureau, Rental Vacancy Rate in the United States [RRVRUSQ156N], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/RRVRUSQ156N, December 1, 2023.

8. Federal Reserve Bank of St. Louis, Spot Crude Oil Price: West Texas Intermediate (WTI) [WTISPLC], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/WTISPLC, December 1, 2023.

9. U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items in U.S. City Average [CPIAUCSL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIAUCSL, December 1, 2023.

10. U.S. Bureau of Labor Statistics, Labor Force Participation Rate [CIVPART], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CIVPART, December 1, 2023.

11. Board of Governors of the Federal Reserve System (US), Federal Funds Effective Rate [FEDFUNDS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/FEDFUNDS, December 1, 2023.

12. University of Michigan, University of Michigan: Consumer Sentiment [UMCSENT], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/UMCSENT, December 1, 2023.

13. US. Bureau of Economic Analysis, Population [B230RC0Q173SBEA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/B230RC0Q173SBEA, November 30, 2023.

14. U.S. Bureau of Economic Analysis, Federal Government: Current Expenditures [FGEXPND], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/FGEXPND, December 1, 2023.

15. U.S. Department of the Treasury. Fiscal Service, Federal Debt: Total Public Debt [GFDEBTN], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/GFDEBTN, December 1, 2023.

16. U.S. Bureau of Economic Analysis, Net Exports of Goods and Services [NETEXP], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/NETEXP, December 1, 2023.

17. Board of Governors of the Federal Reserve System (US), Rest of the World; Foreign Direct Investment in U.S.; Asset (Current Cost), Transactions [ROWFDIQ027S], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/ROWFDIQ027S, November 30, 2023.

18. U.S. Congressional Budget Office, Noncyclical Rate of Unemployment [NROU], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/NROU, December 1, 2023.

19. U.S. Bureau of Economic Analysis, Federal Government: National Defense Consumption Expenditures and Gross Investment [FDEFX], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/FDEFX, December 1, 2023.

20. U.S. Census Bureau, Homeownership Rate in the United States [RHORUSQ156N], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/RHORUSQ156N, December 1, 2023.

21. Board of Governors of the Federal Reserve System (US), Industrial Production: Total Index [INDPRO], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/INDPRO, December 1, 2023.

22. Wikipedia contributors. "List of wars involving the United States." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 27 Nov. 2023. Web. 1 Dec. 2023.

23. S&P 500 Index (^SPX) Historical Data." *Yahoo! Finance*, Yahoo!, 6 Dec. 2023, finance.yahoo.com/quote/%5ESPX/history?period1=1075334400&period2=1701734400&interval=1mo&filter=history&frequency=1mo&includeAdjustedClose=true.

24. U.S. Bureau of Economic Analysis, Federal government current tax receipts [W006RC1Q027SBEA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/W006RC1Q027SBEA, December 5, 2023.

APPENDIX

Fig. 1

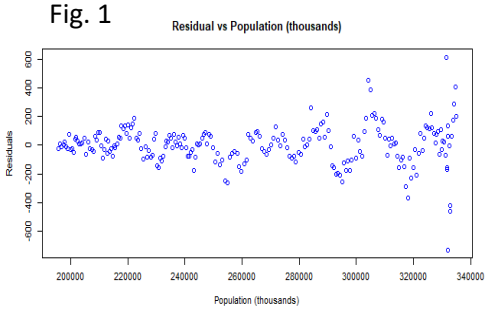


Fig. 2

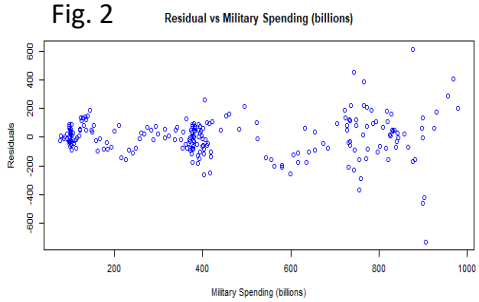


Fig. 3

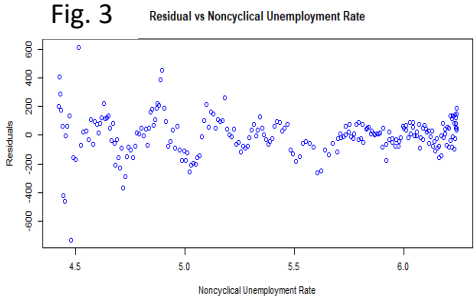


Fig. 4

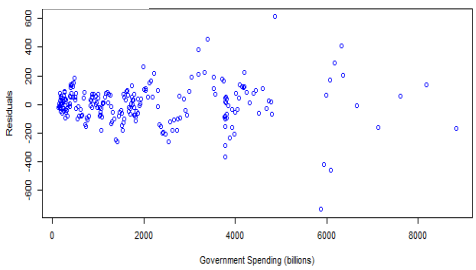


Fig. 5

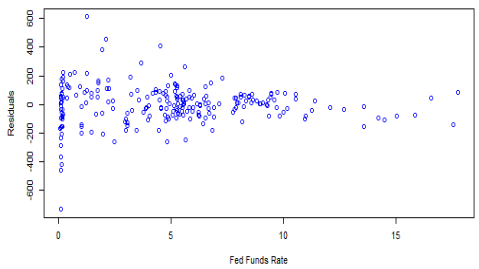


Fig. 6

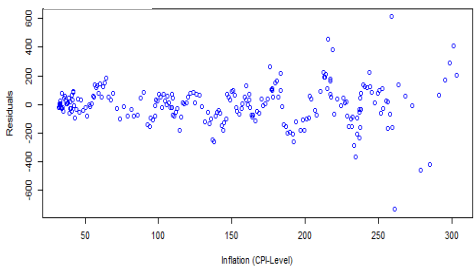


Fig. 7

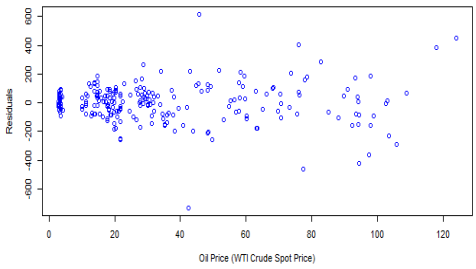


Fig. 8

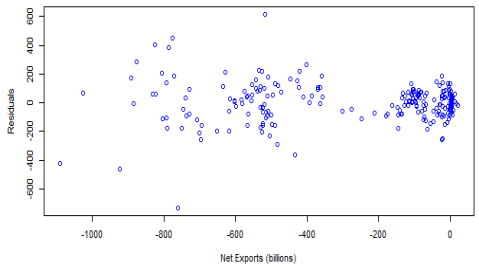


Fig. 9

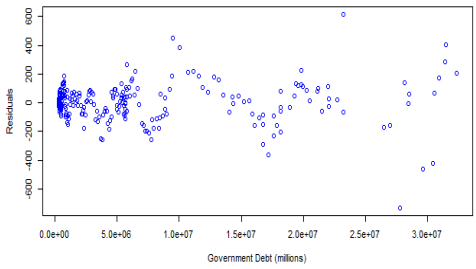


Fig.

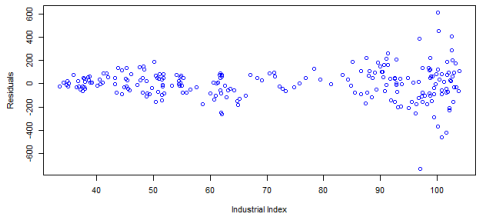


Fig. 11

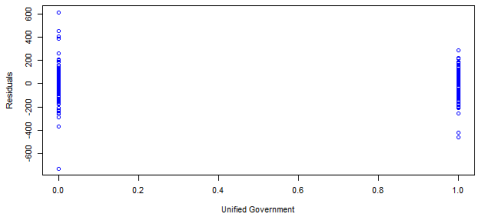


Fig. 12

Fig. 13

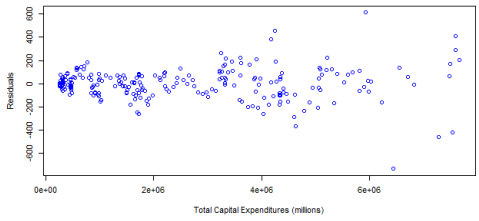


Fig. 14

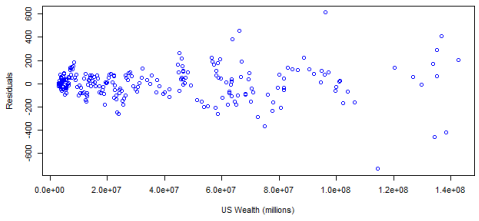


Fig. 15

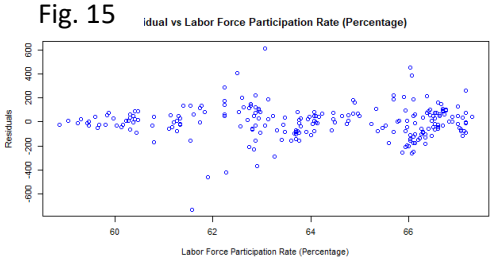


Fig. 16

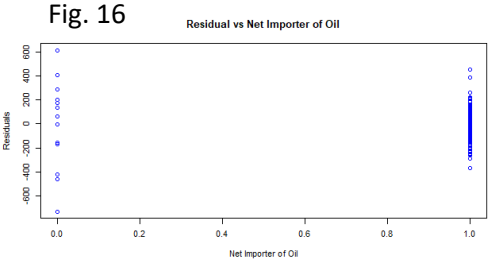


Fig. 17

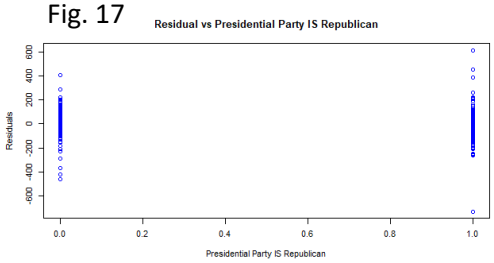


Fig. 18

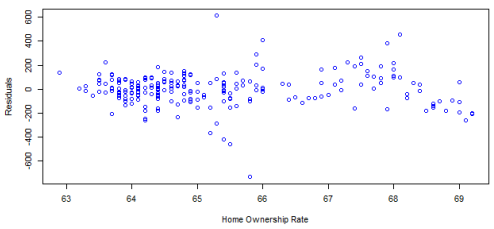


Fig. 19

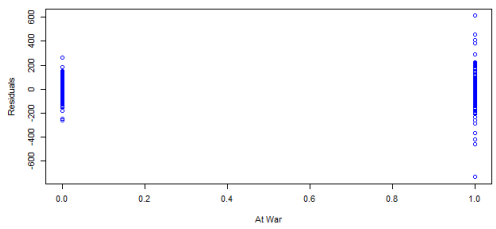


Fig. 20

Fig. 21

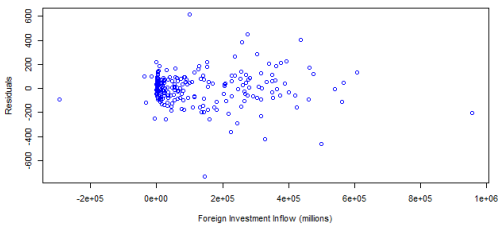


Fig. 22

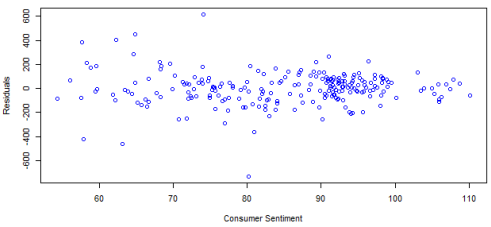


Fig. 23

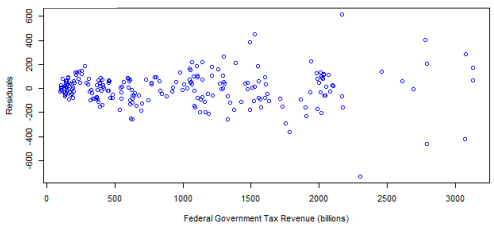


Fig. 24

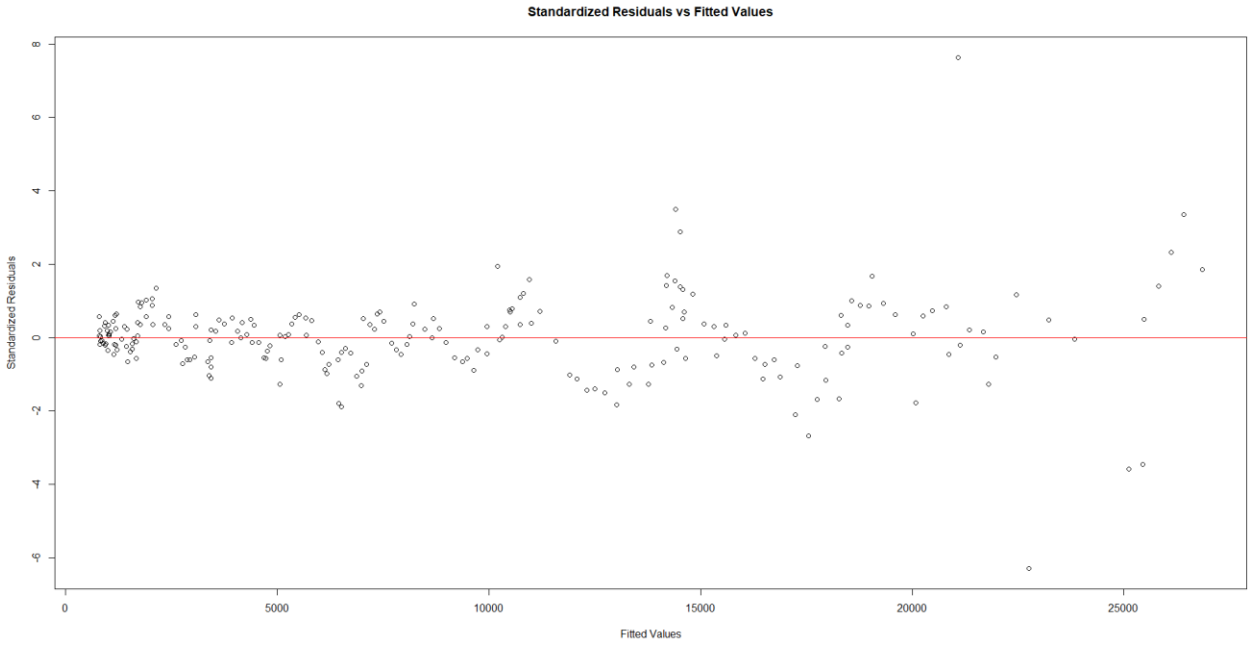
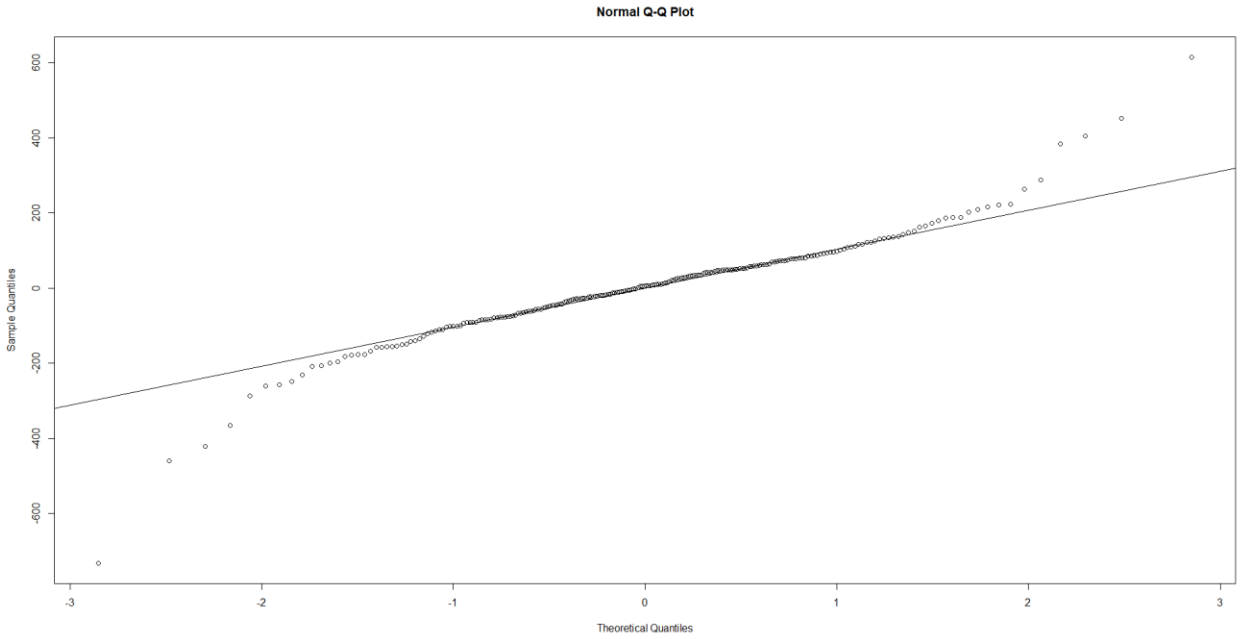


Fig. 25



R Code

```
library(readxl)
library(lmtest)
library(whitestrap)
library(car)


excel_file <- "C:/Users/James/Documents/RBS/Econometrics/project/GDP_Forecast_Econometrics.xlsx"
data <- read_excel(excel_file)

#model usually, 3 for updated numbers
model <- lm(`GDP (billions)` ~ `Population (thousands)` +
              `Military Spending (billions)` + `Noncyclical Unemployment Rate`
            + `Government Spending (billions)` + `Fed Funds Rate`
            + `Inflation (CPI-Level)` + `Oil Price (WTI Crude Spot Price)`
            + `Net Exports (billions)` + `Government Debt (millions)` +
              `Labor Force Participation Rate (Percentage)` +
              `Net Importer of Oil` + `Presidential Party IS Republican` +
              `Home Ownership Rate` + `At War` +
              `SP500 Index` + `Foreign Investment Inflow (millions)` +
              `Consumer Sentiment` + `Federal Government Tax Revenue (billions)` +
              `Industrial Index` + `Unified Government` +
              `Rental Vacancy Rates (Percentage)` + `Total Capital Expenditures (millions)`
            + `US Wealth (millions)`, data = data)

summary(model)

#Now Remove Insignificant Variables
model2 <- lm(`GDP (billions)` ~ `Population (thousands)` +
```

```
        `Military Spending (billions)` + `Noncyclical Unemployment Rate`
      + `Government Spending (billions)`
      + `Inflation (CPI-Level)`
      + `Net Exports (billions)` + `Government Debt (millions)` +
        `Labor Force Participation Rate (Percentage)` +
        `Net Importer of Oil` + `Presidential Party IS Republican` +
        `Home Ownership Rate` +
        `Consumer Sentiment`  + `Total Capital Expenditures (millions)`
      + `US Wealth (millions)`, data = data)

summary(model2)

#Remove 90% significance variables (only 99% confidence variables remain)

model99Only <- lm(`GDP (billions)` ~ `Population (thousands)` +
        `Military Spending (billions)` + `Noncyclical Unemployment Rate`
      + `Government Spending (billions)`
      + `Net Exports (billions)` + `Government Debt (millions)` +
        `Labor Force Participation Rate (Percentage)` +
        `Net Importer of Oil` +
        `Home Ownership Rate` + `Consumer Sentiment`  +
        `Total Capital Expenditures (millions)`
      + `US Wealth (millions)`, data = data)

summary(model99Only)

# Test for Independence of Residuals
dw_test <- dwtest(model)
print(dw_test)

# Homoscedasticity test
bp_test <- bptest(model)
print(bp_test)

#Heteroscedasticity test
white_test <- white_test(model)
print(white_test)

#Standardized residuals vs. fitted values
std_residuals <- rstandard(model)
plot(model$fitted.values, std_residuals,
    main = "Standardized Residuals vs Fitted Values",
    xlab = "Fitted Values", ylab = "Standardized Residuals")
abline(h = 0, col = "red")

#Normal Probability Plot
qqnorm(resid(model))
qqline(resid(model))

#Extracting residuals and predictors
residuals_model <- resid(model)
predictor_variables <- model$model[, -1]
num_predictors <- ncol(predictor_variables)
```

```
# Creating residual vs variable plots for each predictor variable
par(mfrow = c(3, 3))

for (i in 1:num_predictors) {
  plot(predictor_variables[, i], residuals_model,
       main = paste("Residual vs", colnames(predictor_variables)[i]),
       xlab = colnames(predictor_variables)[i],
       ylab = "Residuals",
       col = "blue")
}

par(mfrow = c(1, 1))
```

R Code Output

```
> library(readxl)
> library(lmtest)
> library(whitestrap)
> library(car)
>
>
> excel_file <- "C:/Users/James/Documents/RBS/Econometrics/project/GDP_Forecast_Econometrics.xlsx"
> data <- read_excel(excel_file)
>
> #model usually, 3 for updated numbers
> model <- lm(`GDP (billions)` ~ `Population (thousands)` +
+               `Military Spending (billions)` + `Noncyclical Unemployment Rate`
+             + `Government Spending (billions)` + `Fed Funds Rate`
+             + `Inflation (CPI-Level)` + `Oil Price (WTI Crude Spot Price)`
+             + `Net Exports (billions)` + `Government Debt (millions)` +
+               `Labor Force Participation Rate (Percentage)` +
+               `Net Importer of Oil` + `Presidential Party IS Republican` +
+               `Home Ownership Rate` + `At War` +
+               `SP500 Index` + `Foreign Investment Inflow (millions)` +
+               `Consumer Sentiment` + `Federal Government Tax Revenue (billions)` +
+               `Industrial Index` + `Unified Government` +
+               `Rental Vacancy Rates (Percentage)` + `Total Capital Expenditures (millions)`
+             + `US Wealth (millions)`, data = data)
>
> summary(model)

Call:
lm(formula = `GDP (billions)` ~ `Population (thousands)` + `Military Spending (billions)` +
    `Noncyclical Unemployment Rate` + `Government Spending (billions)` +
    `Fed Funds Rate` + `Inflation (CPI-Level)` + `Oil Price (WTI Crude Spot Price)` +
    `Net Exports (billions)` + `Government Debt (millions)` +
    `Labor Force Participation Rate (Percentage)` + `Net Importer of Oil` +
    `Presidential Party IS Republican` + `Home Ownership Rate` +
    `At War` + `SP500 Index` + `Foreign Investment Inflow (millions)` +
    `Consumer Sentiment` + `Federal Government Tax Revenue (billions)` +
    `Industrial Index` + `Unified Government` + `Rental Vacancy Rates (Percentage)` +
    `Total Capital Expenditures (millions)` + `US Wealth (millions)`,
    data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-732.44  -70.58    5.75   69.48  614.74

Coefficients:
                                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                             -1.954e+02  2.548e+03  -0.077  0.93894
`Population (thousands)`                 2.323e-02  7.626e-03   3.046  0.00262 **
`Military Spending (billions)`           3.789e+00  4.660e-01   8.131 3.93e-14 ***
`Noncyclical Unemployment Rate`         -3.005e+02  1.373e+02  -2.188  0.02977 *
`Government Spending (billions)`        -3.317e-01  4.912e-02  -6.754 1.44e-10 ***
`Fed Funds Rate`                        -7.889e+00  6.346e+00  -1.243  0.21518
`Inflation (CPI-Level)`                  1.128e+01  4.899e+00   2.304  0.02224 *
`Oil Price (WTI Crude Spot Price)`      -1.828e+00  1.530e+00  -1.195  0.23344
`Net Exports (billions)`                 1.806e+00  3.174e-01   5.689 4.35e-08 ***
```

```
`Net Exports (billions)`                      1.806e+00  3.174e-01   5.689 4.35e-08 ***
`Government Debt (millions)`                  1.165e-04  2.634e-05   4.423 1.58e-05 ***
`Labor Force Participation Rate (Percentage)` -1.520e+02  2.992e+01  -5.079 8.46e-07 ***
`Net Importer of Oil`                         8.042e+02  1.680e+02   4.788 3.21e-06 ***
`Presidential Party IS Republican`           -6.677e+01  3.885e+01  -1.719  0.08718 .
`Home Ownership Rate`                         9.532e+01  1.932e+01   4.934 1.66e-06 ***
`At War`                                      1.632e+01  4.334e+01   0.377  0.70689
`SP500 Index`                                 1.462e-01  1.033e-01   1.415  0.15860
`Foreign Investment Inflow (millions)`        1.662e-04  1.144e-04   1.453  0.14784
`Consumer Sentiment`                         -8.481e+00  1.394e+00  -6.086 5.58e-09 ***
`Federal Government Tax Revenue (billions)`  -2.203e-01  2.221e-01  -0.992  0.32255
`Industrial Index`                           -5.682e+00  8.350e+00  -0.681  0.49695
`Unified Government`                          1.458e+01  3.238e+01   0.450  0.65288
`Rental Vacancy Rates (Percentage)`           1.096e+01  2.303e+01   0.476  0.63448
`Total Capital Expenditures (millions)`       1.618e-03  2.001e-04   8.084 5.27e-14 ***
`US Wealth (millions)`                        3.723e-05  7.332e-06   5.078 8.53e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 144.6 on 206 degrees of freedom
Multiple R-squared:  0.9996,    Adjusted R-squared:  0.9996
F-statistic: 2.403e+04 on 23 and 206 DF,  p-value: < 2.2e-16

>
> #Now Remove Insignificant Variables
> model2 <- lm(`GDP (billions)` ~ `Population (thousands)` +
+                `Military Spending (billions)` + `Noncyclical Unemployment Rate`
+              + `Government Spending (billions)`
+              + `Inflation (CPI-Level)`
+              + `Net Exports (billions)` + `Government Debt (millions)` +
+                `Labor Force Participation Rate (Percentage)` +
+                `Net Importer of Oil` + `Presidential Party IS Republican` +
+                `Home Ownership Rate` +
+                `Consumer Sentiment`  + `Total Capital Expenditures (millions)`
+              + `US Wealth (millions)`, data = data)
>
> summary(model2)

Call:
lm(formula = `GDP (billions)` ~ `Population (thousands)` + `Military Spending (billions)` +
    `Noncyclical Unemployment Rate` + `Government Spending (billions)` +
    `Inflation (CPI-Level)` + `Net Exports (billions)` + `Government Debt (millions)` +
    `Labor Force Participation Rate (Percentage)` + `Net Importer of Oil` +
    `Presidential Party IS Republican` + `Home Ownership Rate` +
    `Consumer Sentiment` + `Total Capital Expenditures (millions)` +
    `US Wealth (millions)`, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-828.75  -65.74   -3.08   70.17  589.85
```

```
Coefficients:
                                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                                  -1.413e+03  1.843e+03  -0.767   0.4441
`Population (thousands)`                      2.686e-02  4.037e-03   6.652 2.36e-10 ***
`Military Spending (billions)`                3.882e+00  3.409e-01  11.386  < 2e-16 ***
`Noncyclical Unemployment Rate`              -4.849e+02  7.195e+01  -6.740 1.43e-10 ***
`Government Spending (billions)`             -3.517e-01  4.682e-02  -7.511 1.55e-12 ***
`Inflation (CPI-Level)`                       5.170e+00  3.053e+00   1.693   0.0918 .
`Net Exports (billions)`                      1.762e+00  2.678e-01   6.581 3.52e-10 ***
`Government Debt (millions)`                  1.390e-04  2.448e-05   5.679 4.36e-08 ***
`Labor Force Participation Rate (Percentage)` -1.197e+02  2.356e+01  -5.080 8.16e-07 ***
`Net Importer of Oil`                         6.238e+02  1.400e+02   4.455 1.35e-05 ***
`Presidential Party IS Republican`           -4.326e+01  2.477e+01  -1.747   0.0821 .
`Home Ownership Rate`                         9.019e+01  1.574e+01   5.731 3.34e-08 ***
`Consumer Sentiment`                         -6.327e+00  1.211e+00  -5.225 4.10e-07 ***
`Total Capital Expenditures (millions)`       1.399e-03  1.466e-04   9.541  < 2e-16 ***
`US Wealth (millions)`                        4.529e-05  5.591e-06   8.099 4.10e-14 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 147.3 on 215 degrees of freedom
Multiple R-squared:  0.9996,    Adjusted R-squared:  0.9996
F-statistic: 3.804e+04 on 14 and 215 DF,  p-value: < 2.2e-16

>
> #Remove 90% significance variables (only 99% confidence variables remain)
>
> model99Only <- lm(`GDP (billions)` ~ `Population (thousands)` +
+                 `Military Spending (billions)` + `Noncyclical Unemployment Rate`
+               + `Government Spending (billions)`
+               + `Net Exports (billions)` + `Government Debt (millions)` +
+                 `Labor Force Participation Rate (Percentage)` +
+                 `Net Importer of Oil` +
+                 `Home Ownership Rate` + `Consumer Sentiment`  +
+                 `Total Capital Expenditures (millions)`
+               + `US Wealth (millions)`, data = data)
>
> summary(model99Only)

Call:
lm(formula = `GDP (billions)` ~ `Population (thousands)` + `Military Spending (billions)` +
    `Noncyclical Unemployment Rate` + `Government Spending (billions)` +
    `Net Exports (billions)` + `Government Debt (millions)` +
    `Labor Force Participation Rate (Percentage)` + `Net Importer of Oil` +
    `Home Ownership Rate` + `Consumer Sentiment` + `Total Capital Expenditures (millions)` +
    `US Wealth (millions)`, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-919.05  -75.36   -4.37   78.96  539.68
```

```
Coefficients:
                                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                                   -3.713e+03  1.580e+03  -2.350   0.0197 *
`Population (thousands)`                       3.106e-02  3.157e-03   9.840  < 2e-16 ***
`Military Spending (billions)`                 3.970e+00  2.795e-01  14.200  < 2e-16 ***
`Noncyclical Unemployment Rate`               -5.559e+02  6.722e+01  -8.270 1.34e-14 ***
`Government Spending (billions)`              -3.687e-01  4.686e-02  -7.869 1.68e-13 ***
`Net Exports (billions)`                       1.711e+00  2.599e-01   6.583 3.42e-10 ***
`Government Debt (millions)`                   1.641e-04  2.231e-05   7.354 3.89e-12 ***
`Labor Force Participation Rate (Percentage)` -8.580e+01  1.951e+01  -4.397 1.72e-05 ***
`Net Importer of Oil`                          6.254e+02  1.399e+02   4.471 1.25e-05 ***
`Home Ownership Rate`                          9.041e+01  1.573e+01   5.749 3.02e-08 ***
`Consumer Sentiment`                          -6.624e+00  1.212e+00  -5.467 1.25e-07 ***
`Total Capital Expenditures (millions)`        1.398e-03  1.440e-04   9.707  < 2e-16 ***
`US Wealth (millions)`                         4.382e-05  5.569e-06   7.869 1.68e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 149 on 217 degrees of freedom
Multiple R-squared:  0.9996,    Adjusted R-squared:  0.9996
F-statistic: 4.338e+04 on 12 and 217 DF,  p-value: < 2.2e-16

>
> # Test for Independence of Residuals
> dw_test <- dwtest(model)
> print(dw_test)

        Durbin-Watson test

data:  model
DW = 0.96579, p-value < 2.2e-16
alternative hypothesis: true autocorrelation is greater than 0

>
> # Homoscedasticity test
> bp_test <- bptest(model)
> print(bp_test)

        studentized Breusch-Pagan test

data:  model
BP = 145.57, df = 23, p-value < 2.2e-16

>
> #Heteroscedasticity test
> white_test <- white_test(model)
> print(white_test)
White's test results

Null hypothesis: Homoskedasticity of the residuals
Alternative hypothesis: Heteroskedasticity of the residuals
Test Statistic: 37.56
P-value: 0
```

```
>
> #Standardized residuals vs. fitted values
> std_residuals <- rstandard(model)
> plot(model$fitted.values, std_residuals,
+      main = "Standardized Residuals vs Fitted Values",
+      xlab = "Fitted Values", ylab = "Standardized Residuals")
> abline(h = 0, col = "red")
>
> #Normal Probability Plot
> qqnorm(resid(model))
> qqline(resid(model))
>
> #Extracting residuals and predictors
> residuals_model <- resid(model)
> predictor_variables <- model$model[, -1]
> num_predictors <- ncol(predictor_variables)
>
> # Creating residual vs variable plots for each predictor variable
> par(mfrow = c(3, 3))
>
> for (i in 1:num_predictors) {
+   plot(predictor_variables[, i], residuals_model,
+        main = paste("Residual vs", colnames(predictor_variables)[i]),
+        xlab = colnames(predictor_variables)[i],
+        ylab = "Residuals",
+        col = "blue")
+ }
>
> par(mfrow = c(1, 1))
```