

The Effectiveness of Biden's American Rescue Plan: Linear Regression Analysis*

Yiran Mei

02/04/2022

Abstract

COVID-19 has dominated people's daily lives across the globe since its emergence at the beginning of 2020. In order to combat the adverse effects of the recession on the United States economy, President Biden announced and signed the American Rescue Plan (ARP) to stimulate the market. The purpose of this study is to investigate the impact of the plan's implementation on a number of economic indicators such as the unemployment rate and inflation rate using simple linear regression and a multiple regression model. Additionally, statistical techniques such as regression analysis and correlation analysis are applied to analyze the regression model's results. We estimate that the ARP is effective in reducing the unemployment rate in the simple linear regression model but has no statistical significance with respect to these economic indicators based on our multiple regression model. Statistical analysis of data and evaluation of statistical results will be conducted using the R statistical programming language, version 4.1.2.

Contents

1. Introduction	2
2. Data	3
2.1 Data Source	3
2.2 Data Characteristics	3
3. Model	5
3.1 Model Assumption Tests	5
3.2 Model Development	5
3.2.1 Simple Linear Model	5
3.2.2 Multiple Linear Model with All Covariates	5
4. Results	6
4.1 Models Summary	6
4.1.1 Simple Linear Regression Model	6
4.1.2 Multiple Linear Regression Model	6

*Code and data are available at: <https://github.com/YiranMei/American-Rescue-Plan>.

4.2 Comparisons of Two Models	7
4.3 Homoscedasticity Assumption Test	8
4.3.1 Simple Linear Regression Model	8
4.3.2 Multiple Linear Regression Model with All Covariates	8
5. Discussion	9
5.1 Discussion and Implications	9
5.2 Ethics and biases	9
Appendix	10
Appendix.A	10
Appendix.B	15
Appendix.C	16
Appendix.D	21
Appendix.E	24
Reference	27

1.Introduction

Since 2020, the outbreak of COVID-19 has adversely affected people's daily lives and the global economy. More than 235 000 cases of infection have occurred in the United States in less than two months since the first case was discovered in Washington on January 20, 2020. As of March 17, the outbreak had spread all over the country, including Washington, New York, and California. The number of COVID-19-related deaths in the United States has exceeded 5000 as of April 2020(Saad B. Omer (2020)). On December 31, 2020, there were more than 20 million infections and more than 346 thousand deaths within the United States, as well as 83,832,334 confirmed cases and 1,824,590 confirmed deaths worldwide(4). Apart from the COVID-19 pandemic, the related economic crisis in the United States has had a detrimental effect on the economic and health well-being of millions of American workers. Over 9.5 million jobs have been lost as a result of COVID-19, with 4 million workers out of work for at least six months(The White House (2021)).

Consequently, in order to stimulate and recover the economy from the foreseeable recession, Biden and the government announced the American Rescue Plan (ARP) on March 11, 2021, under which approximately \$1 trillion in total would be allocated directly to the American people(The White House (2021)). Along with extending PUA (Pandemic Unemployment Assistance) through September 6, 2006(The U.S. Government Publishing Office (2021)), each worker can receive a lump-sum payment of \$1400 and weekly benefits of \$300. A number of provisions were included in the proposed plan, including an extension of unemployment compensation, an increase in the Child Tax Credit, funding for state and local governments, and a subsidy for extra medical expenses resulting from COVID-19(Suzi LeVine (2021)). At this point of reopening and reflecting on the past two years, an examination of the effectiveness of the ARP becomes necessary to assess whether it has boosted the economy.

In light of the foregoing, we conducted this study to investigate whether there is a statistically significant correlation between the growth of the U.S. economy in various aspects and the implementation of ARP using a number of economic indicators from the Federal Reserve Economic Data. We examine the association in this report using a multiple regression model that incorporates covariates such as unemployment rate, population, 10-year breakeven inflation rate, federal funds effective rate, average hourly earnings, as well as civilian labour force level and a dummy variable reflecting the status of before and after the implementation

of the ARP. The implementation of ARP has been shown to effectively reduce unemployment rates in simple linear regression models. Nonetheless, the data do not show a statistically significant association between the unemployment rate and the implementation of the ARP, indicating that the policy has not been effective or that the effect has not been fully exploited and demonstrated by January 2022.

In the remaining elements of this paper, there are three sections: a data section where the data is analyzed and examined, a model section where the regression model is developed, and a discussion section where the limitations and implications are discussed. The findings of this report will provide stakeholders with an insight into how to evaluate ARP policies and make better plans by taking into account findings.

To analyze the data in this report, the R statistical programming language(R Core Team (2021)) will be applied. Moreover, the following packages must be installed in R for cleaning and analysis processes: `openxlsx`(Philipp Schaubberger (2021)), `ggplot2`(Wickham (2016)),`kableExtra`(Zhu (2021)), `gtsummary`(Sjoberg et al. (2021)), `jtools`(Long (2020)), and `modelsummary`(Arel-Bundock (2022)).

2. Data

2.1 Data Source

All data are from the website of FRED, which refers to the Federal Reserve Economic Data. The database is created and maintained by Research Department at the Federal Reserve Bank of St. Louis and contains economic data time series from a range of national, international, public, and private sources. The data were collected and updated by the Federal Reserve on a daily basis. The website tools enables user to combine multiple economic indicators on a single graph/files, and a dataset in excel format is therefore downloaded from the website and imported into R studio directly.

2.2 Data Characteristics

There are nine variables in total on the dataset. The date, unemployment rate, population, 10-year breakeven inflation rate, federal funds effective rate, average hourly earnings, as well as civilian labour force level are from the original dataset downloaded, and the dummy variable reflecting the status of before and after the implementation of the ARP is created through R Markdown. The date is transformed into the time interval collapsed on the date of implementation of the ARP as well to develop the model.

The unemployment rate, which would be the variable of interest in the model, is calculated by dividing the number of unemployed by the number of employed and achieved through the collection of labor force data. It can be seen in Figure 1 that the unemployment rate increase considerably before period -9(2020-06-01), and then rocketed to 14.7%. This may be due to the widespread outbreak of COVID-19 after March 2020. Afterwards, the unemployment rate dropped to 8.83% and then to 5.9% prior to the implementation of the ARP. It is evident that after the implementation of the ARP, the rate is decreasing continuously. The dummy variable of the implementation of ARP is equal to zero if the quarter precedes the ARP and equal to one if the quarter follows.

There are five control covariates in total. The population includes both local residents and armed forces overseas. For inflation rate, we used the breakeven inflation rate which measures the expected inflation rate derived from 10-Year Treasury Constant Maturity Securities. The latest value implies what market participants expect inflation to be in the next 10 years, on average. Additionally, the federal funds rate was used as an interest rate. Basically, it is the overnight rate that depository institutions within the United States use to trade the excess reserves of federal funds with other banks, which is determined by the market, but also monitored by the Federal Reserve as a target rate. Including overtime and late shift payments, the average hourly earnings of all private employees is an indicator of the average hourly earnings of all private employees on a “gross” basis. Benefits, irregular bonuses, retroactive pay, and the employer’s payroll taxes are not included in this measure. The civilian labor force level is restricted to adults aged 16 and older who

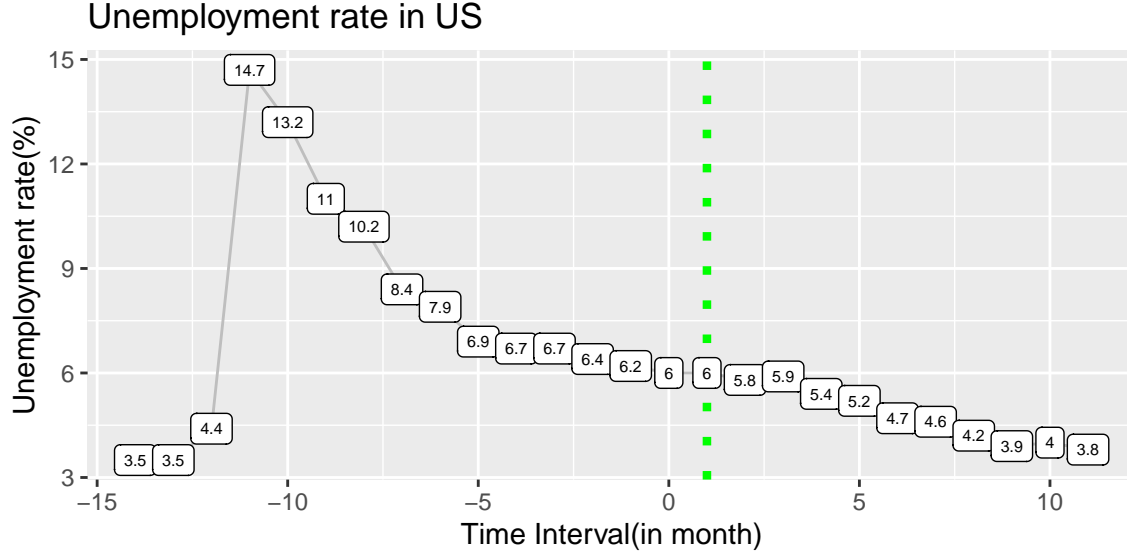


Figure 1: Unemployment rate in US

reside in a state or the District of Columbia, who are not confined to institutions (such as prisons, mental institutions, or retirement homes for the elderly), and who are not on active duty in the armed forces. These data are derived from the Current Population Survey (Household Survey).

An overview of all variables is provided in Table 1, which indicates the scale and range of each variable.

Table 1: Summary of Data

quarter	UnemploymentRate	Population	InflationRate	InterestRate	Laborforce	Wage	ARP
Min. :-14.00	Min. : 3.500	Min. :331443	Min. :0.9868	Min. :0.04864	Min. :156358	Min. :28.43	Min. :0.0000
1st Qu.: -7.75	1st Qu.: 4.450	1st Qu.:331792	1st Qu.:1.6575	1st Qu.:0.07693	1st Qu.:160404	1st Qu.:29.52	1st Qu.:0.0000
Median : -1.50	Median : 5.950	Median :331966	Median :2.1308	Median :0.08000	Median :160906	Median :30.02	Median :0.0000
Mean : -1.50	Mean : 6.508	Mean :332043	Mean :1.9675	Mean :0.21488	Mean :161199	Mean :30.10	Mean :0.4231
3rd Qu.: 4.75	3rd Qu.: 6.850	3rd Qu.:332269	3rd Qu.:2.3493	3rd Qu.:0.09239	3rd Qu.:161997	3rd Qu.:30.74	3rd Qu.:1.0000
Max. : 11.00	Max. :14.700	Max. :332750	Max. :2.6160	Max. :1.58368	Max. :164583	Max. :31.60	Max. :1.0000

The graphs showing the trend of those covariates over the study period are included in Appendix.A. According to Figure 3, the population of the United States has shown an overall increasing trend throughout the period. During periods -4 and -2, it decreased rapidly before increasing slightly. It surged after period 1, reaching more than 332,000,000. During the period illustrated in Figure 4, the inflation rate in the United States fluctuates substantially. For example, it significantly decreased between periods -14 and -12. However, the trend began to accelerate after that, until the period -7. The next increasing period of it occurred between periods -4 and 2 and began to decline one month after the implementation of ARP. According to Figure 5, the effective federal fund rate indicates a stable trend following a decline between periods -13 and -11. Following this decline, the rate remained within the range of 0.7 and 0.9. Figure 6 demonstrates that the average earning in the US increased from period -12 to period -11, reaching \$30/hr. Following the jump, the average earning rises steadily over time and exceeds \$31 in period 7. A fluctuating trend can also be observed in the civilian labor force in the US, as shown in Figure 7. However, after a significant drop and growth from period -13 to period -9, it fluctuates around 161 million, and has continued to rise since the introduction of ARP in period 1, reaching 162.5 million by the end of period 9.

3. Model

The purpose of this subsection is to examine the relationship between observed economic characteristics; the time collapse on the implementation date and the unemployment rate in the United States, using a regression framework. This analysis entails several steps.

3.1 Model Assumption Tests

To begin with, we need to check if our data meet the four main assumptions for the regression.

Table 2 displays the correlation matrix as a test of the independence of observations. It is evident that none of the correlations display extremely high values, apart from the correlation between the quarter and population and wage. Due to the fact that the quarter is merely a time variable, we can disregard it. Second, we perform a histogram of the dependent variable in order to determine whether it is normal in Figure 8. In light of the small sample size and the presence of an outlier, it is understandable that the histogram of unemployment shows a positive skewness rather than a normal distribution, which violates the assumption. Violating the assumption of normality has few consequences, as it does not result in any bias or inefficiency in the regression model. P values are only affected if the assumption is violated. Following this, we test each covariate for linearity on the time interval. As can be seen from the figures in Appendix.C, the majority of covariates have a linear relationship with the variable of interest except for a few outliers except the dummy variable ARP. For the Homoscedasticity Assumption Test, we would do it after performing the regression.

Table 2: Correlation Matrix of Variables

	quarter	UnemploymentRate	Population	InflationRate	InterestRate	Laborforce	Wage	ARP
quarter	1.00	-0.51	0.96	0.90	-0.52	0.27	0.96	0.86
UnemploymentRate	-0.51	1.00	-0.47	-0.66	-0.35	-0.89	-0.28	-0.49
Population	0.96	-0.47	1.00	0.79	-0.49	0.30	0.95	0.79
InflationRate	0.90	-0.66	0.79	1.00	-0.27	0.40	0.78	0.80
InterestRate	-0.52	-0.35	-0.49	-0.27	1.00	0.59	-0.61	-0.29
Laborforce	0.27	-0.89	0.30	0.40	0.59	1.00	0.09	0.34
Wage	0.96	-0.28	0.95	0.78	-0.61	0.09	1.00	0.83
ARP	0.86	-0.49	0.79	0.80	-0.29	0.34	0.83	1.00

3.2 Model Development

3.2.1 Simple Linear Model

Having completed the tests, we will attempt to construct a simple linear regression in which the dummy variable ARP serves as the running variable and the unemployment rate is the outcome of interest,

$$Y_i = \beta D_i + \mu_i ,$$

where the dummy variable D_i is equal to zero if the quarter precedes the ARP and equal to one if the quarter follows, while μ_i represents the error term. The β indicates the coefficient of the treatment variable, ARP on the outcome variable.

3.2.2 Multiple Linear Model with All Covariates

Consequently, we attempt to build multiple linear regression models for the central equation in order to make a better fit. In essence, the equation is as follows:

$$Y_i = \beta D_i + f(t_i - t') + \beta_i X_i + \mu_i ,$$

taking t_i as the calendar date, and t' as the date on which the ARP was implemented, $f(t_i - t')$ serves as a running variable that represents the relationship between the collapsed date and the outcome of interest. The dummy variable D_i is equal to zero if the quarter precedes the ARP and equal to one if the quarter follows. For bias reduction, we add X_i representing other covariates stated before, while μ_i represents the error term. The β indicates the coefficient of the treatment variable, ARP on the outcome variable while the β_i represents the coefficients of all covariates. We build the model with all of the covariates discussed before in this report.

4. Results

4.1 Models Summary

4.1.1 Simple Linear Regression Model

According to Figure 2, the regression line of the simple regression model does not fit well with the plots since there are no observations between ARP equals zero and one. The coefficient of the simple linear model shown in Table 3 tells that the on average, the unemployment rate became 2.8% lower after the introduction of the ARP, and the result is statistically significant at a 10% significance level but insignificant at 5%. Therefore, the data is evidently sufficient to reject the null hypothesis of the regression and indicating there is a non-zero correlation, which suggests the changes in the independent variable are associated with the changes in the dependent variable at the population level at a significance level of 10%. That is, the implementation of ARP is associated with the reduction of the unemployment rate.

Table 3: Summary of Simple Linear Regression

Characteristic	Beta	95% CI	p-value
ARP	-2.8	-5.0, -0.74	0.010

¹ CI = Confidence Interval

4.1.2 Multiple Linear Regression Model

After performing the multiple linear regression, we examine the regression analysis using the regression summary. In the Table 4, it shows that the coefficient of the variable ARP is 0.37, which suggests a positive relationship with the implementation of the ARP and the unemployment rate. However, this result is not statistically significant as the P-value is 0.6, which is a considerably large value. It implies that there is no sufficient evidence to reject the null hypothesis and find a non-zero correlation. That is, there is no statistical evidence that the implementation of ARP is associated with the reduction of unemployment rate. Moreover, we can see that only two covariates, interest rate and wage variables, produce statistical significant outcomes with a significance level of 5% and 10%.

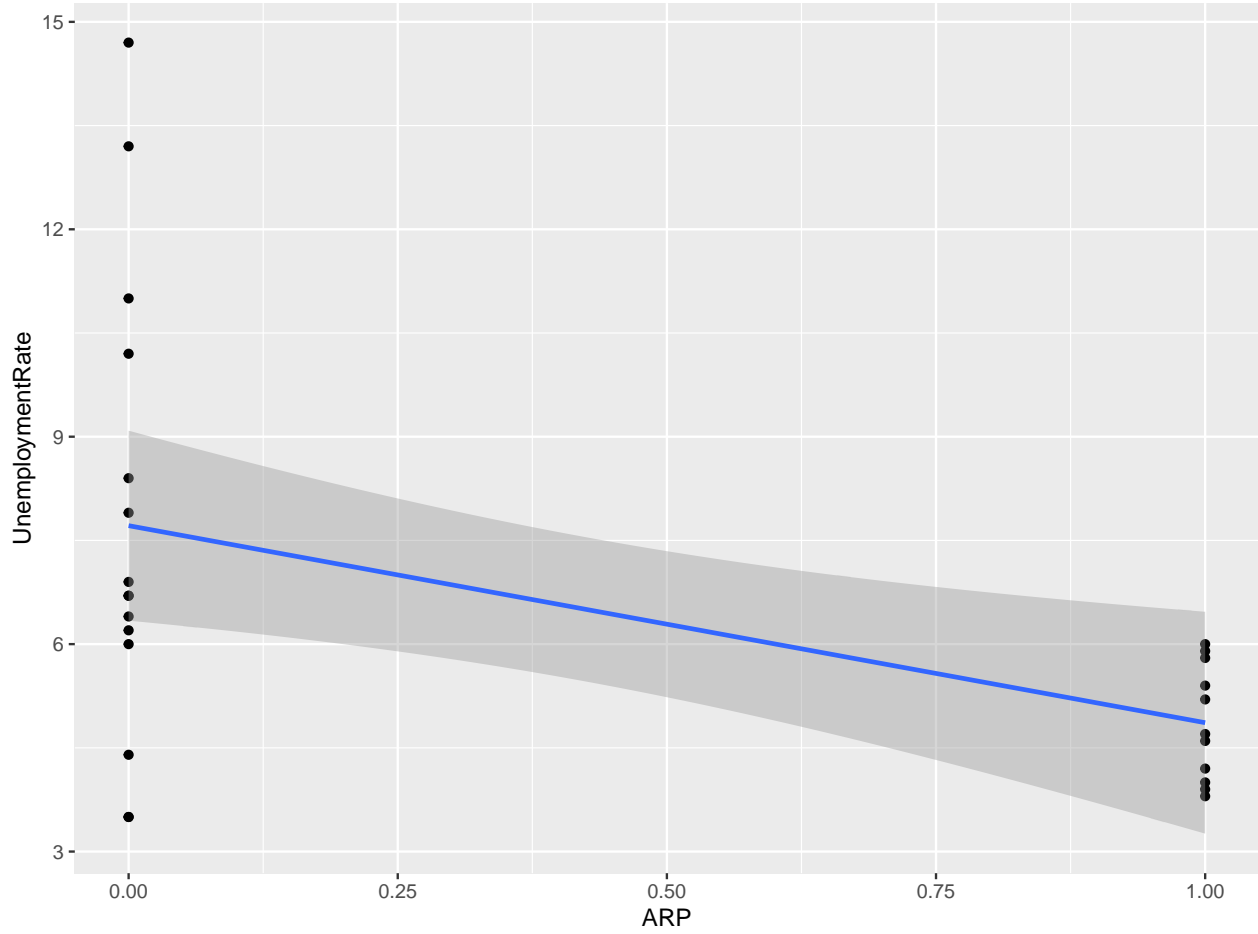


Figure 2: Simple Linear Regression

Table 4: Summary of Multiple Linear Regression

Characteristic	Beta	95% CI	p-value
quarter	-1.0	-1.6, -0.28	0.008
ARP	0.37	-1.2, 1.9	0.6
Population	0.00	-0.01, 0.00	0.7
InflationRate	1.4	-2.2, 4.9	0.4
InterestRate	-4.2	-7.8, -0.52	0.027
Wage	5.4	2.4, 8.5	0.001
Laborforce	0.00	0.00, 0.00	0.7

¹ CI = Confidence Interval

4.2 Comparisons of Two Models

From the comparison summary of two models in Table 5, the R-Squared of simple linear regression model is merely 0.2442, which suggests only 24.42% of the variability observed in the unemployment rate is explained by the regression model. The r-squared of the multiple linear regression model is 0.9497624, which indicates that approximately 94.97% of the data fit the regression model. On the other hand, the standard error of the ARP treatment variable in the multiple linear regression model is also smaller than the simple linear regression model. This suggests a more precise prediction of the multiple linear regression model as well.

Table 5: Comparison of Models

	Model 1	Model 2
(Intercept)	7.713 (0.666)	157.682 (857.630)
ARP	-2.850 (1.023)	0.374 (0.737)
quarter		-0.956 (0.321)
Population		-0.001 (0.003)
InflationRate		1.356 (1.680)
InterestRate		-4.154 (1.728)
Wage		5.414 (1.447)
Laborforce		0.000 (0.000)
Num.Obs.	26	26
R2	0.244	0.950
R2 Adj.	0.213	0.930
AIC	127.0	68.5
BIC	130.7	79.8
Log.Lik.	-60.477	-25.243
F	7.753	48.577
RMSE	2.58	0.77

4.3 Homoscedasticity Assumption Test

4.3.1 Simple Linear Regression Model

Based on the figures in Appendix.D, the “Residuals vs. Fitted” plot clearly indicates an increase in variance in the fitted values. For fitted values below 5, residuals range from -2 to 2. Alternatively, for larger fitted values beyond 7.5, the difference between residuals is greater. The presence of heteroscedasticity is evident from this observation. On the “Scale-Location”-plot, an increase in the square root of the standardized residuals can be observed across the fitted values. Thus, we can conclude, using both the “Residuals versus Fitted” and the “Scale-Location” plots, that a simple linear regression model does not satisfy the homoscedasticity assumption.

4.3.2 Multiple Linear Regression Model with All Covariates

Images from Appendix.E are used to perform the homoscedasticity assumption test for multiple linear regression. Observe that both the “Residual vs. Fitted” and the “Scales-Location” plots are devoid of heteroscedastic residuals. Despite the fact that the lines in these two plots do not appear to be flat, the residuals show a relatively stable variability that is unaffected by the fitted values. Consequently, the multiple linear regression model does not violate the assumption of homoscedasticity.

5. Discussion

5.1 Discussion and Implications

In the two models we developed for this report, the first ignores all other factors that affect the unemployment rate and directly concludes that its implementation led to a reduction. Unfortunately, incomplete data and a limited time frame distort the efficiency of the model itself, which decreases the degree of credibility of the results. In general, based upon the analysis of the models presented in this report, we may only accept the multiple linear regression model. Unexpectedly, this model suggests that the unemployment rate actually increases due to the introduction, but the association is not statistically significant. There is a greater likelihood that the decline in unemployment rate is related to changes in interest rates and wages. Nevertheless, the change in the interest rate and wage may not be independent of the introduction of ARP. COVID-19 definitely results in a monetary shock on the market, resulting in a decrease in the asset market (Khurshid Adnan (2021)). Consequently, in order to achieve full employment, the U.S. government carries out a temporary expansionary fiscal policy which includes the ARP in order to increase government spending and, subsequently, raise aggregate demand and, therefore, the level of production (Galí, Jordi (2007)). Both interest rates and exchange rates will be impacted by the rising demand for money. As a price in the labor market, wages are also implicated in macroeconomic changes as a whole. Accordingly, the effect of the ARP may be transmitted to the unemployment rate through other mediums, such as the change in interest rate, wage and some other omitted variables in the report. The fact is that we do indeed witness a reduction in unemployment in the United States following the implementation of this policy, but the rationale behind it in the overall economic system is far more complicated than the policy itself. Due to this, we cannot draw an arbitrary conclusion that the ARP is completely ineffective at boosting employment and the economy. More analysis and consideration of numerous economic indicators should be completed and combined in order to render a conclusive and comprehensive assessment of the ARP policy.

5.2 Ethics and biases

The primary drawback of this research and model is that the data is insufficient to support a reliable and sound conclusion. Since the policy was conducted only a short while ago, there is not enough time for us to collect enough information to predict an accurate statistic. Our model can only be built using 26 periods of data with incomplete distributions. Future data will be needed in order to analyze the policy's long-term effectiveness. Secondly, the factors influencing the unemployment rate may be more extensive than those cited in this report. Consequently, an omitted variable bias may be present. In the absence of accounting for relevant independent variables, the coefficient of the variables in the model is likely to be biased and inaccurate. The economy as a whole involves a large number of factors that influence the rate of unemployment and policy in conjunction, and some of these factors may be unknown or unquantifiable. For future research, it would be beneficial to develop a model with more precise and abundant data in order to eliminate the basis described above.

Appendix

Appendix.A

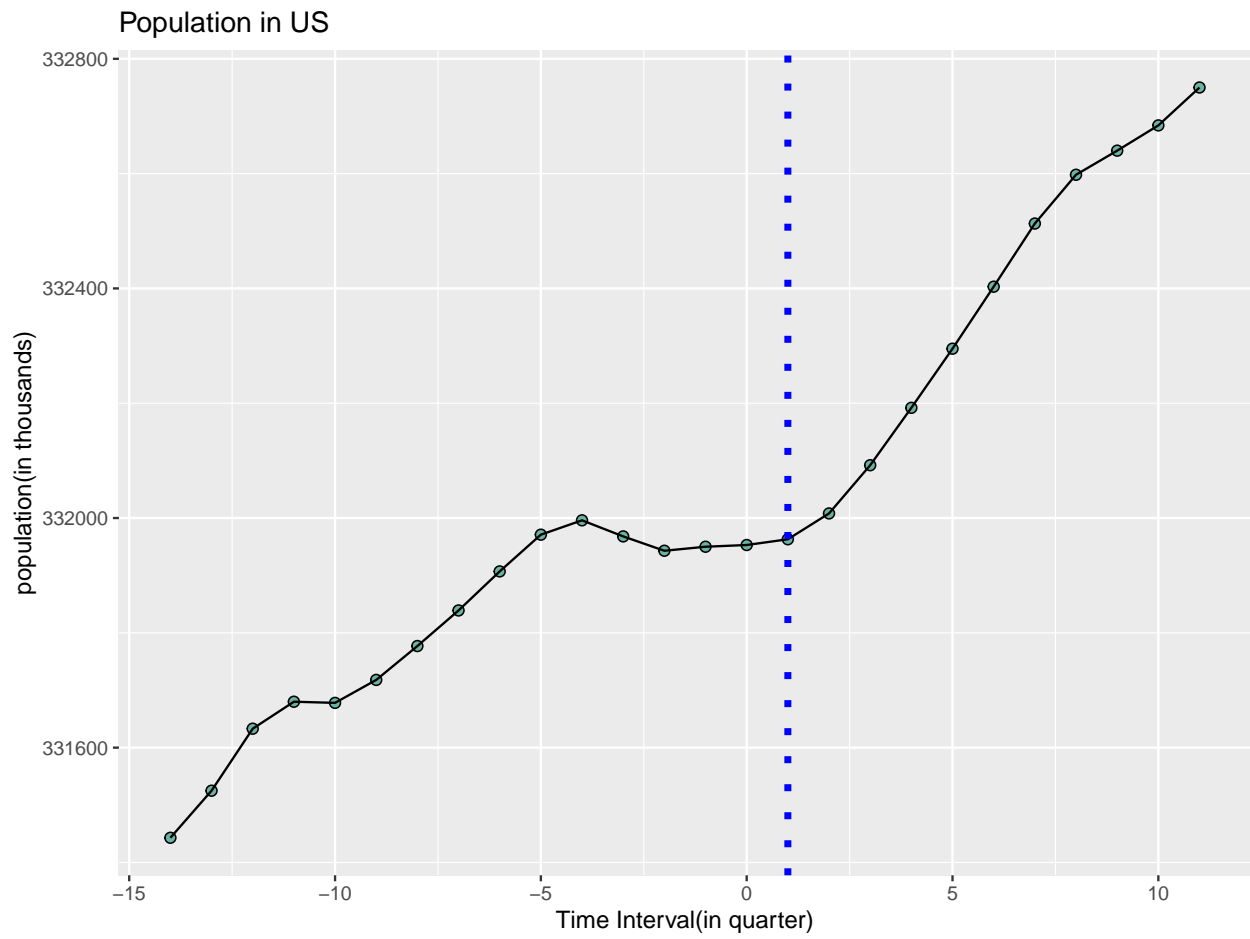


Figure 3: Population in US

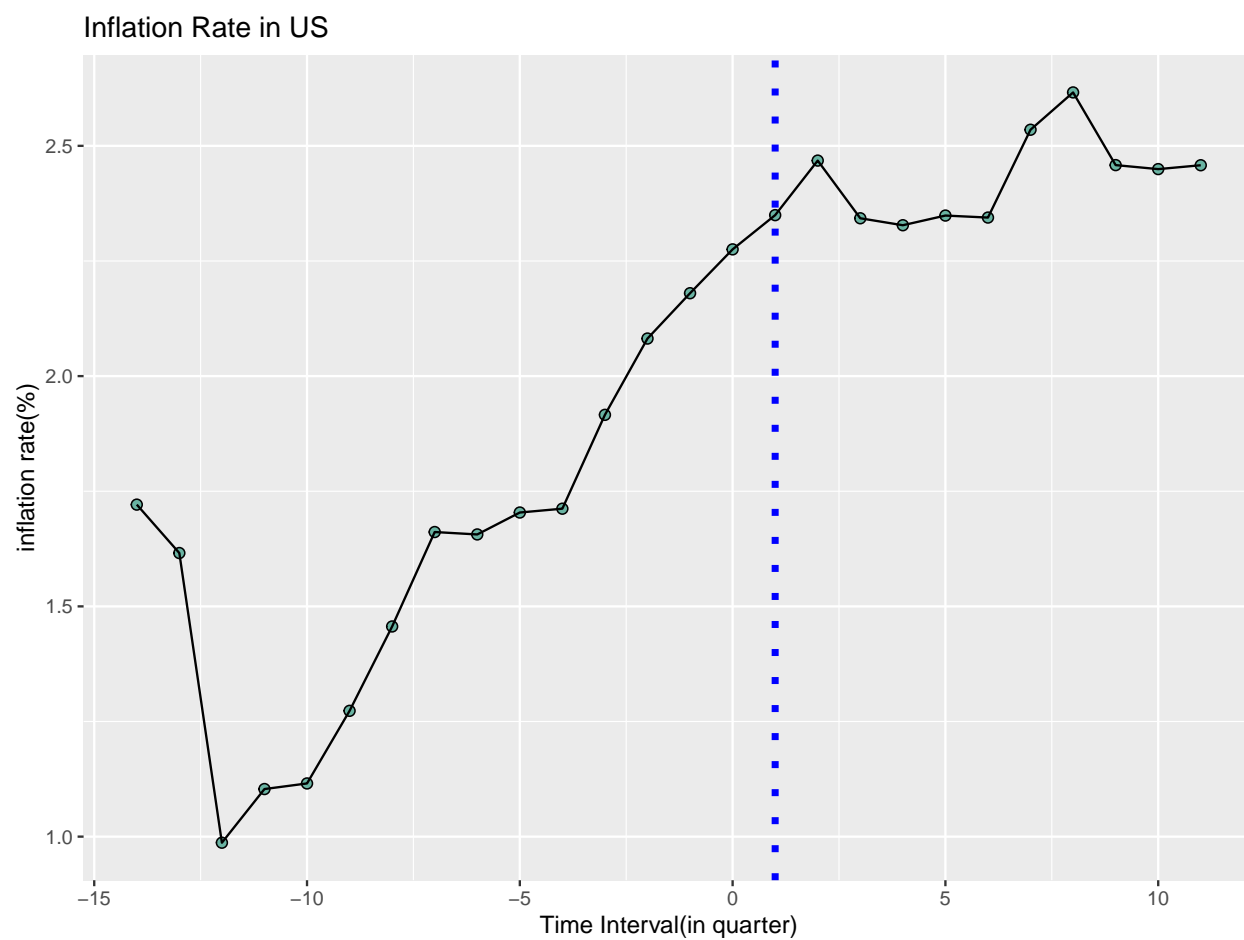


Figure 4: Inflation Rate in US

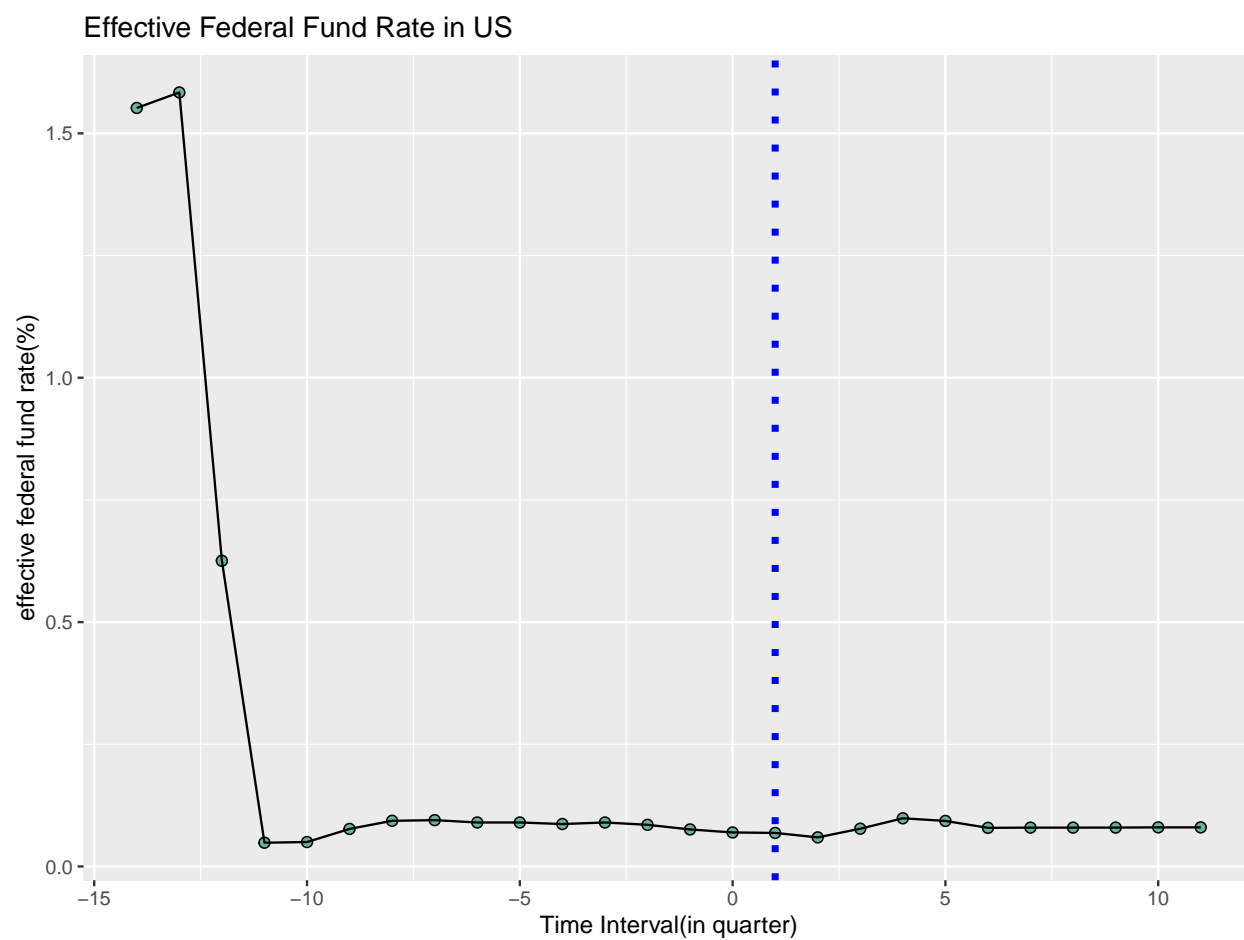


Figure 5: Effective Federal Fund Rate in US

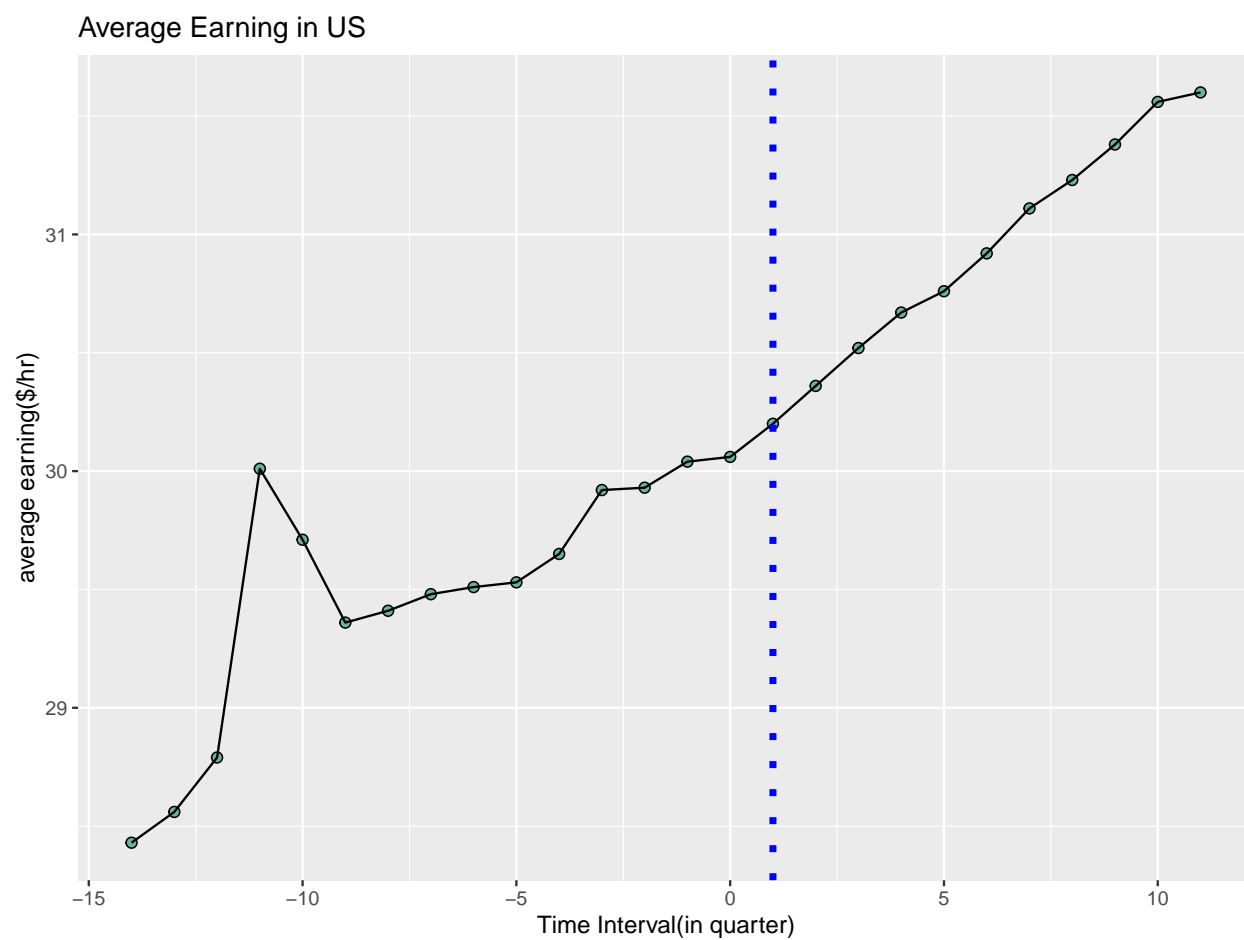


Figure 6: Average Earning in US

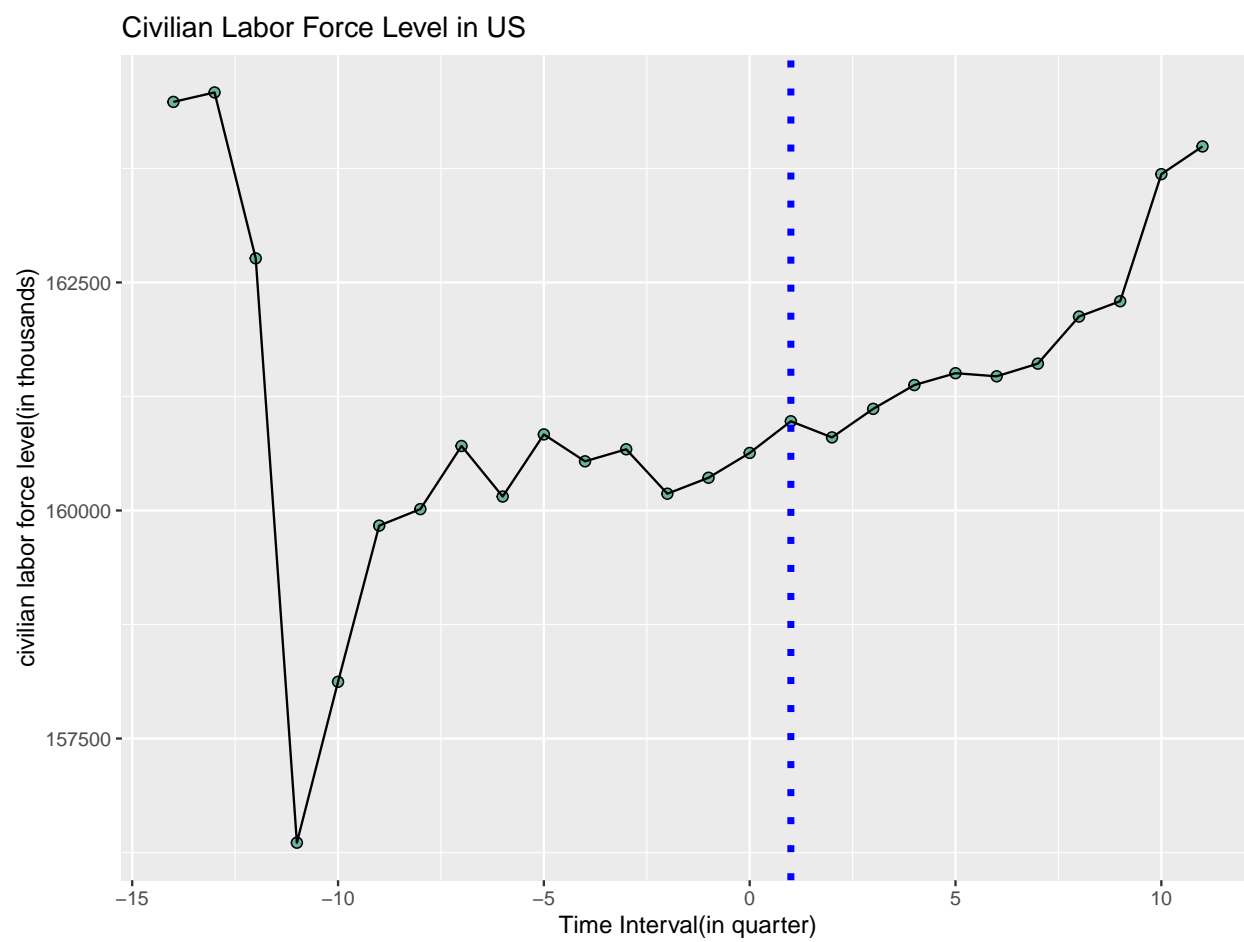


Figure 7: Civilian Labor Force Level in US

Appendix.B

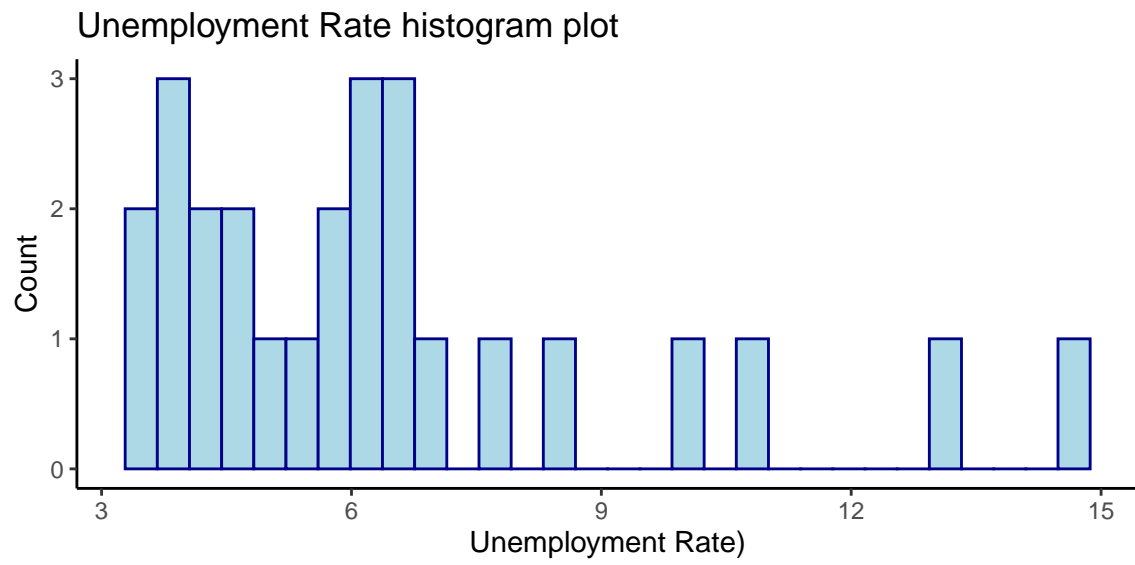
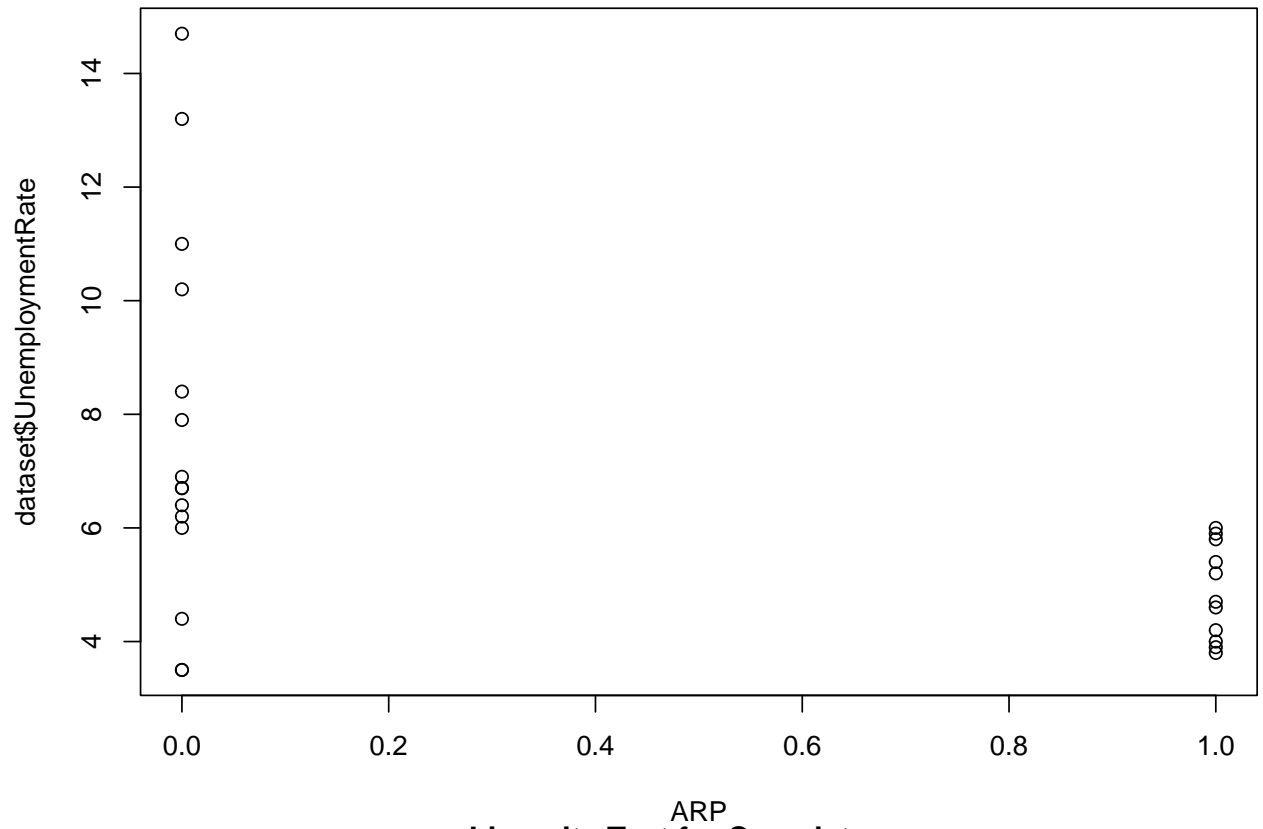


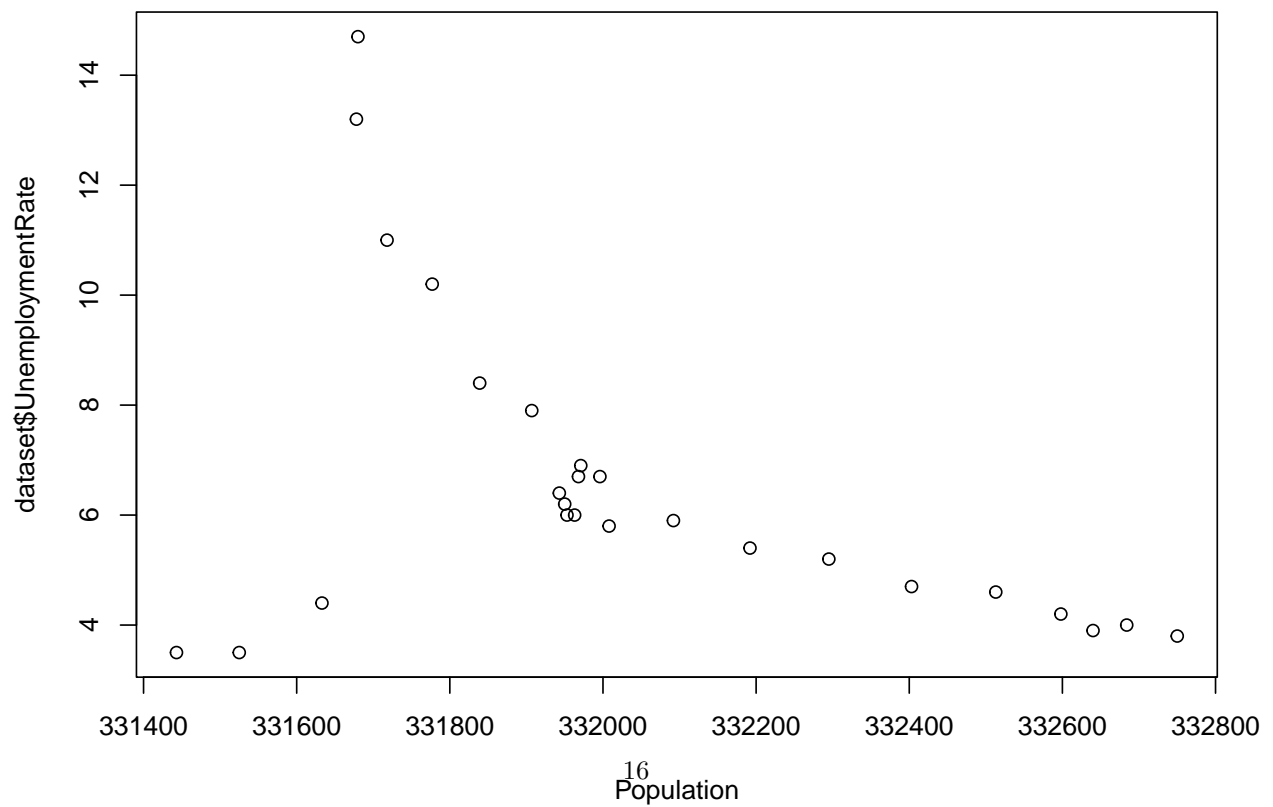
Figure 8: Normality of Unemployment Rate

Appendix.C

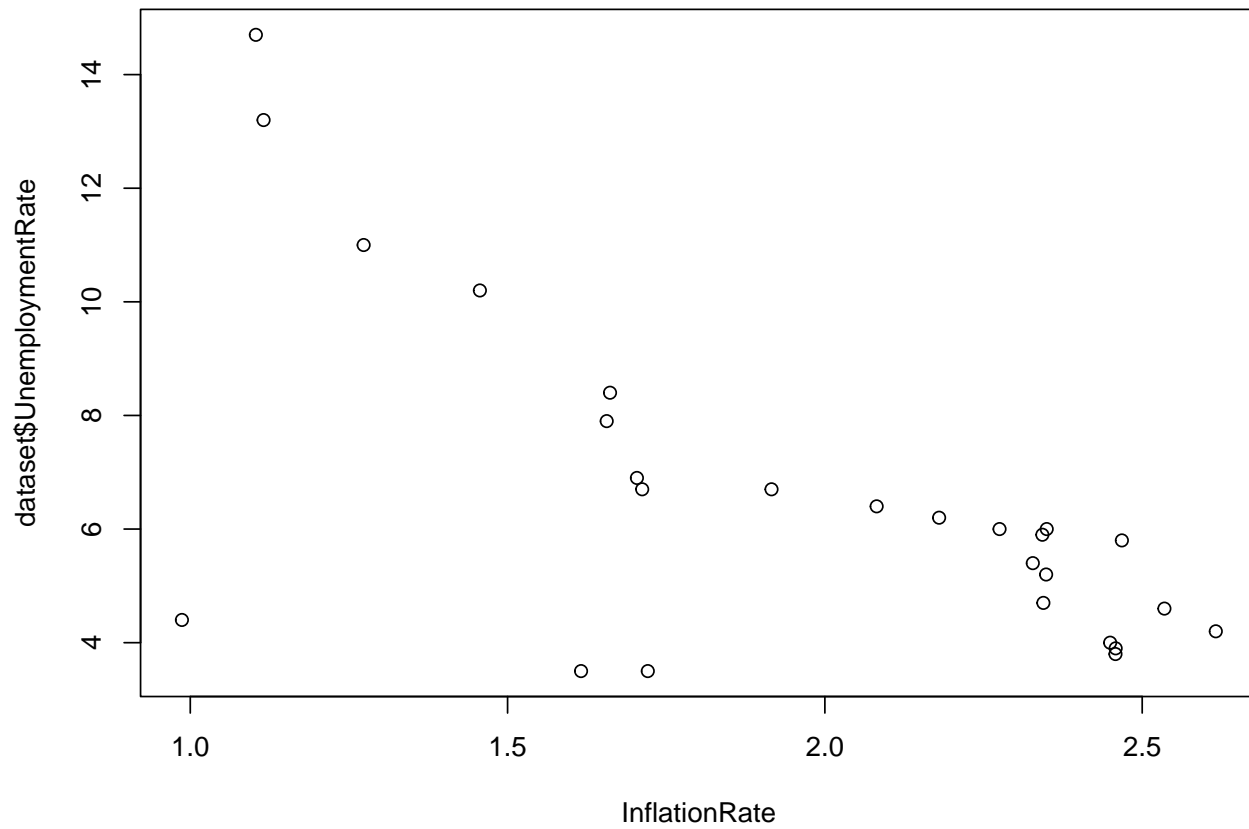
Linearity Test for Covariates



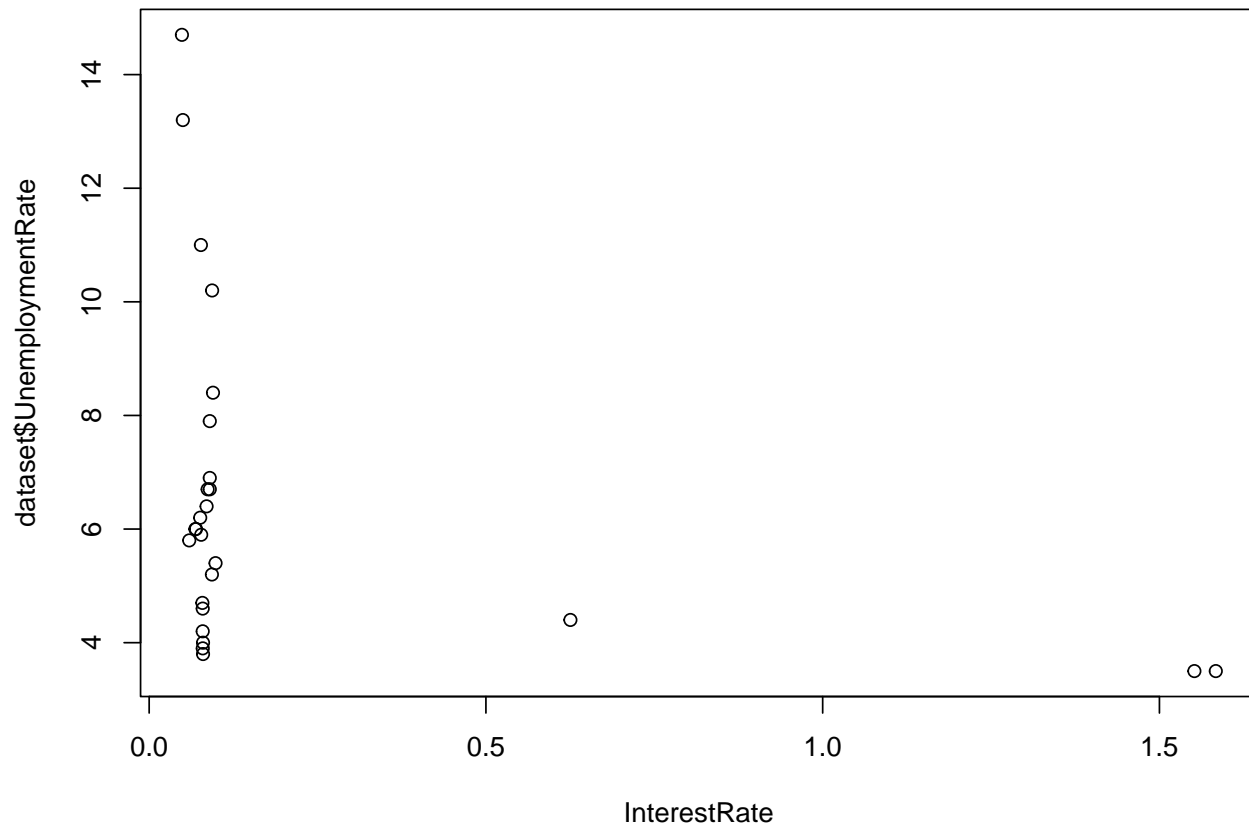
Linearity Test for Covariates



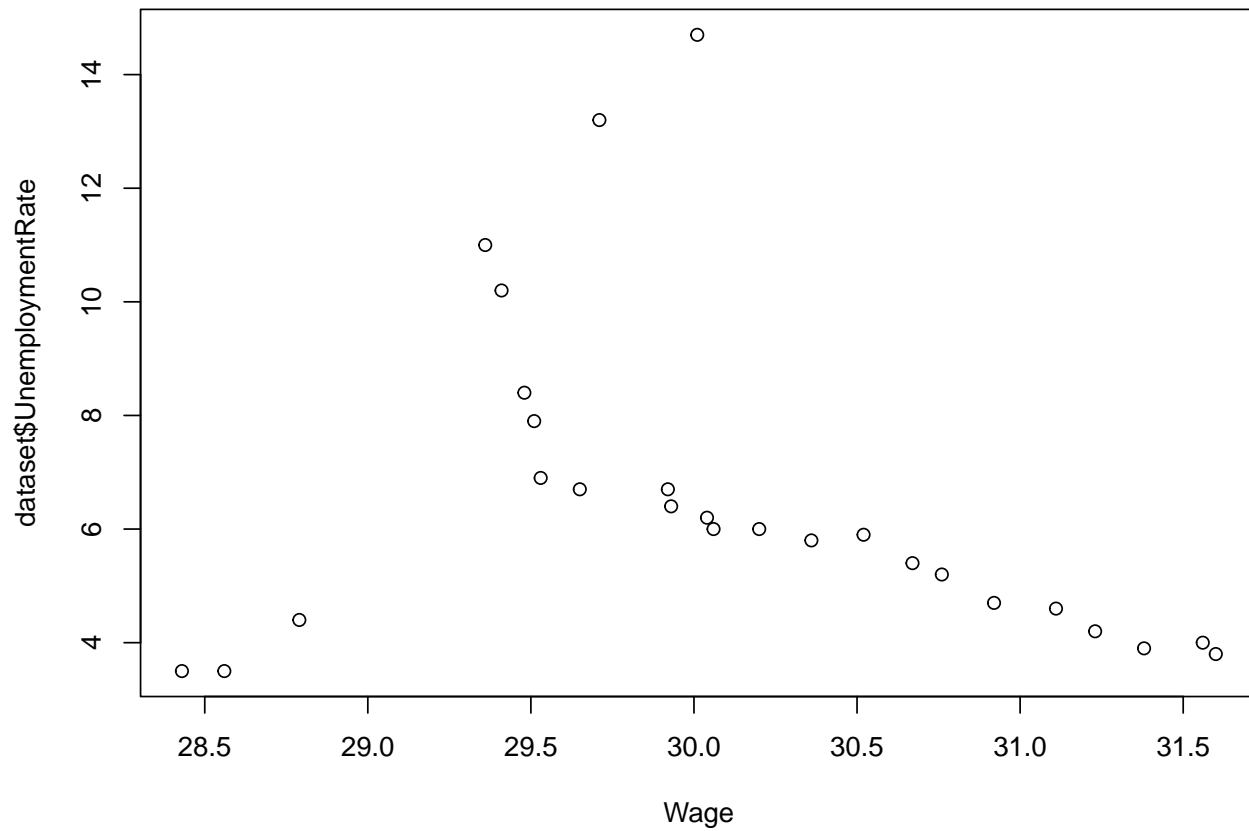
Linearity Test for Covariates



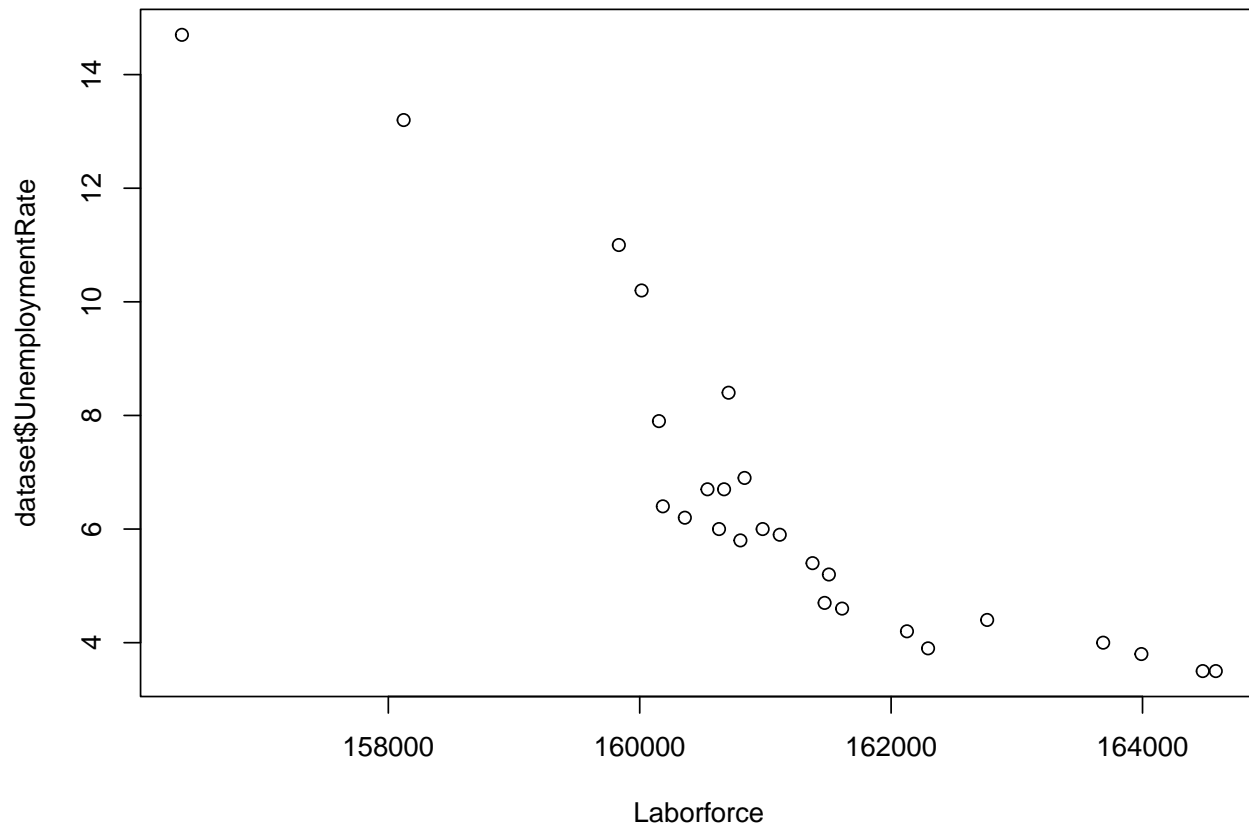
Linearity Test for Covariates



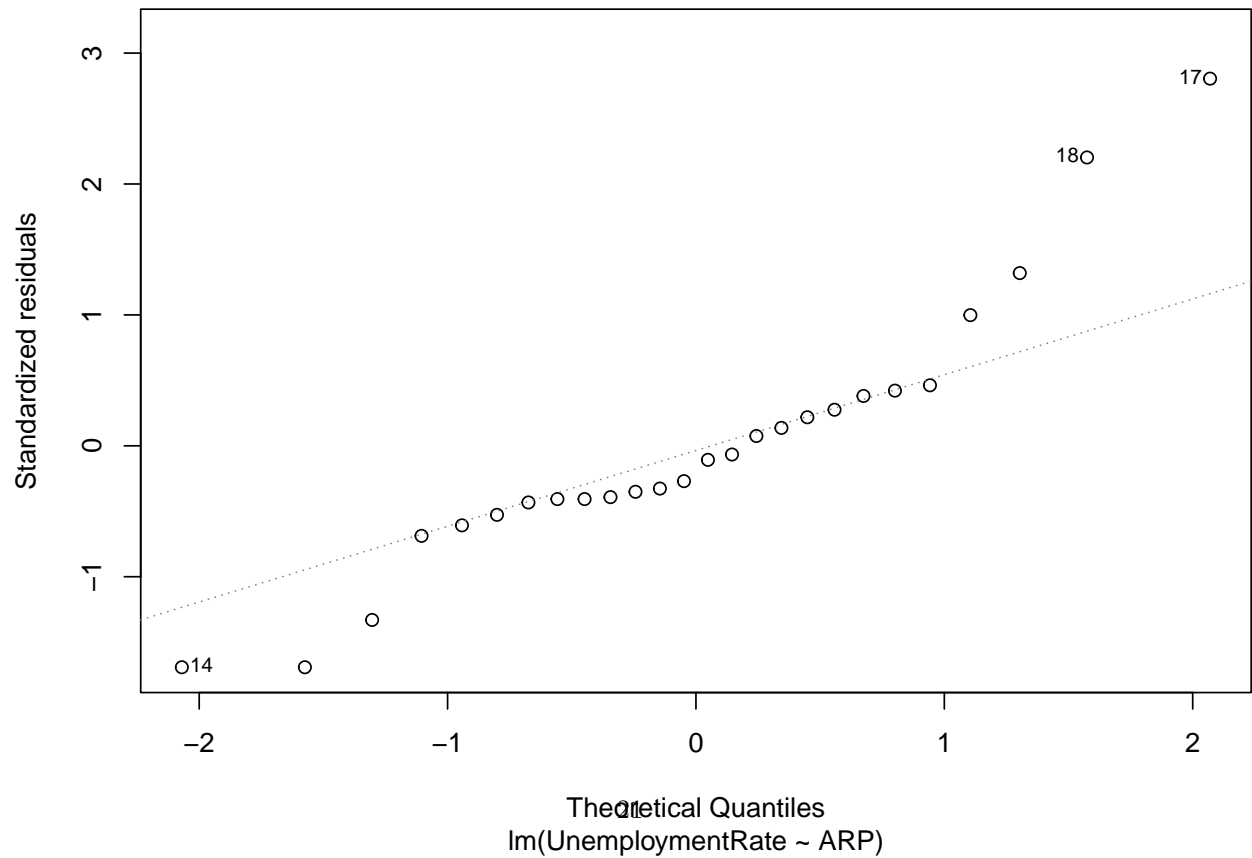
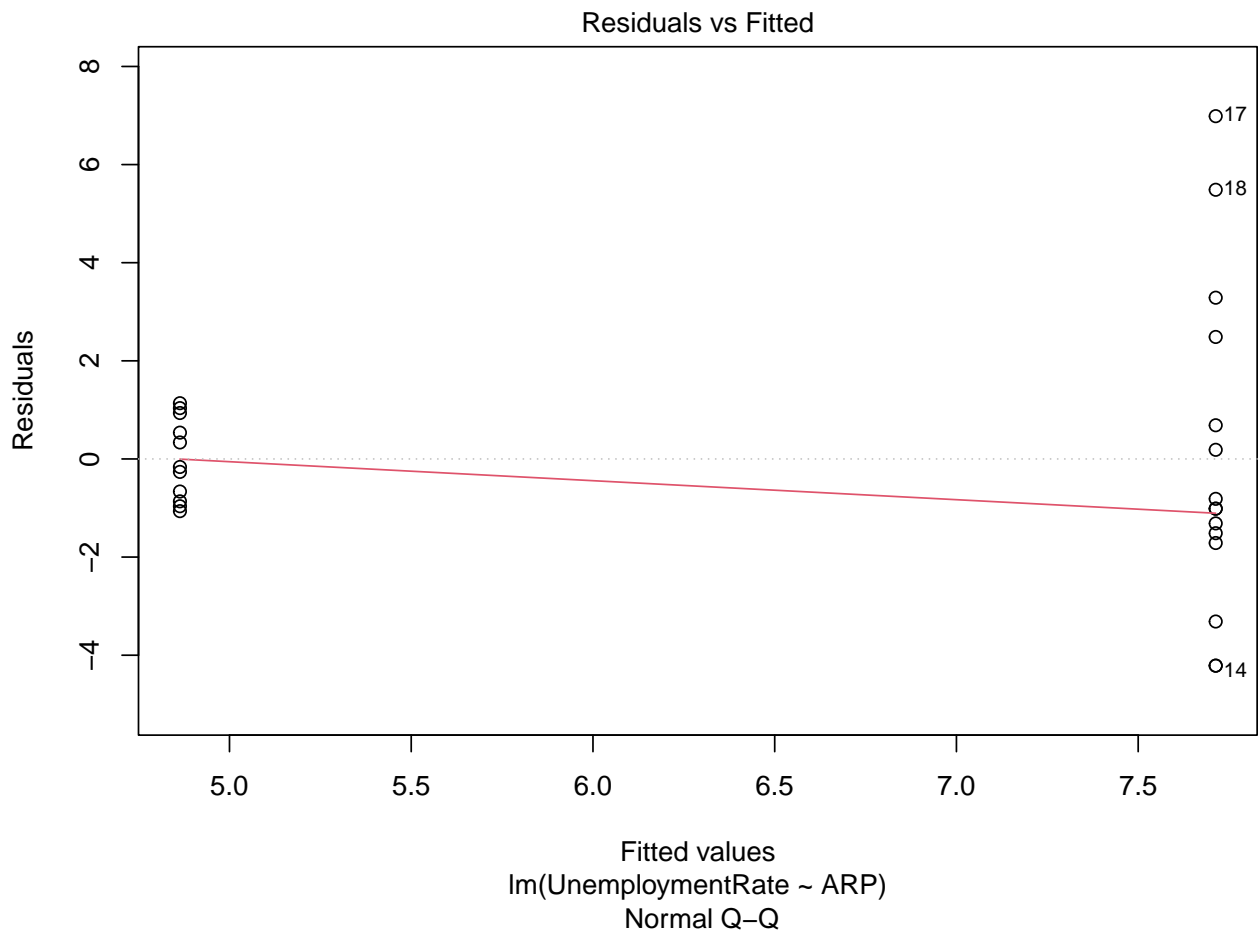
Linearity Test for Covariates

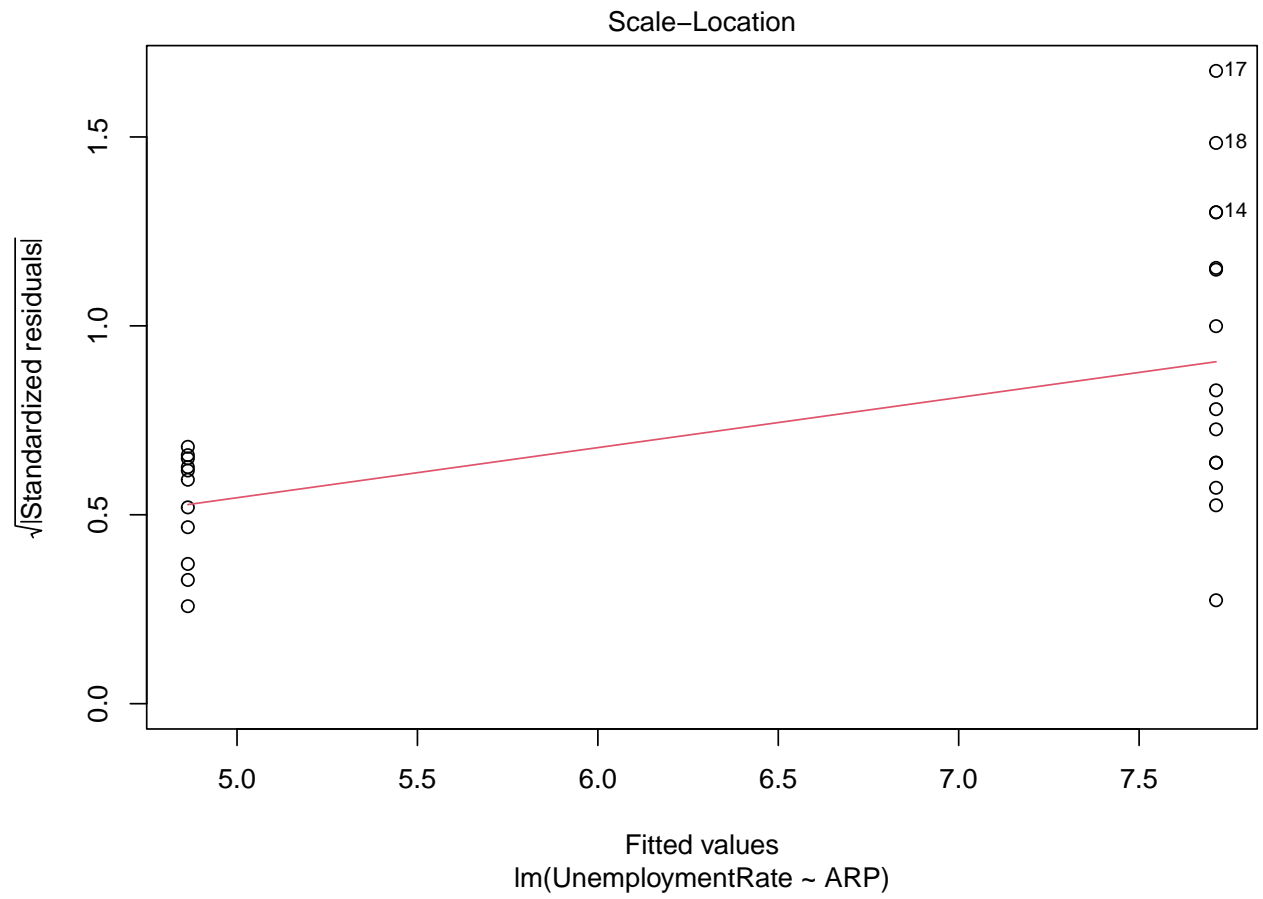


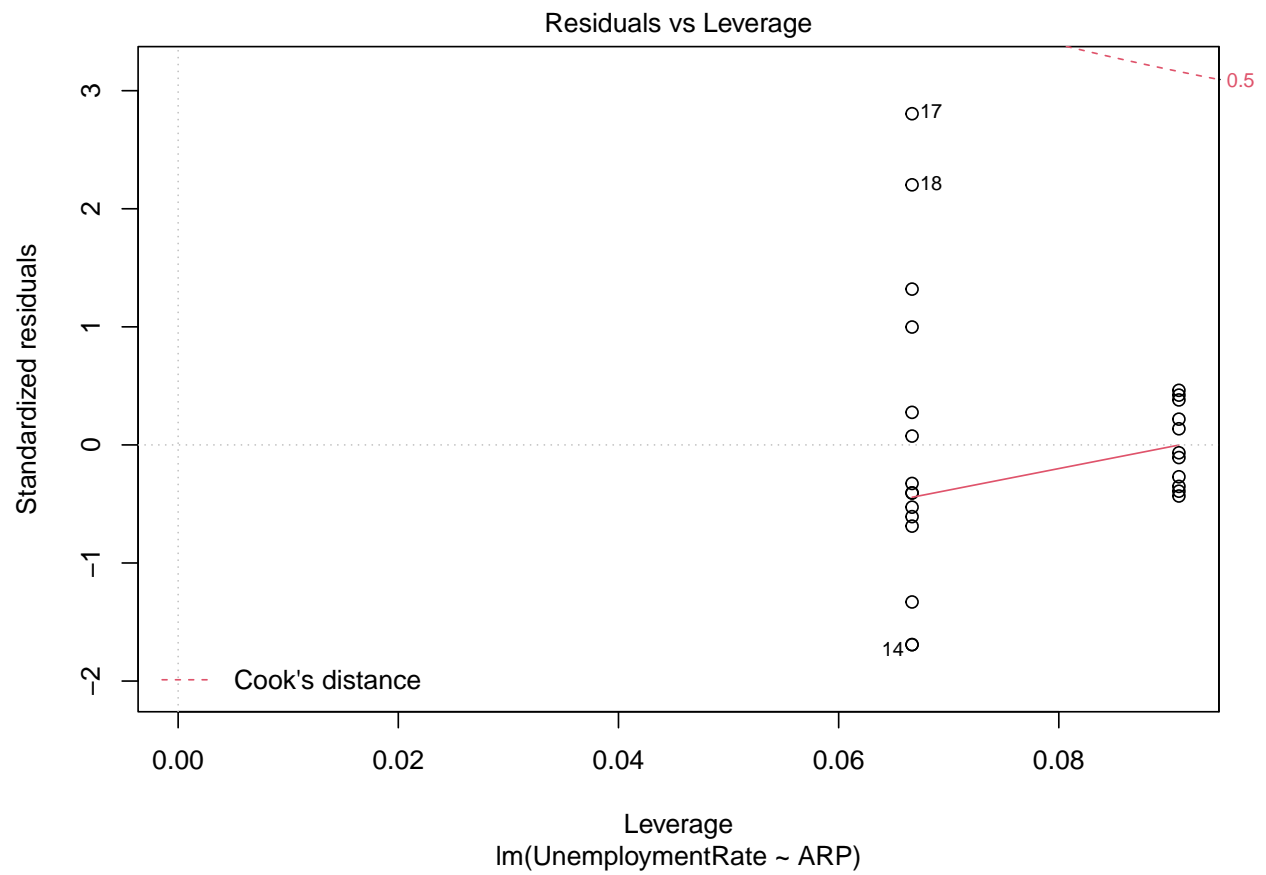
Linearity Test for Covariates



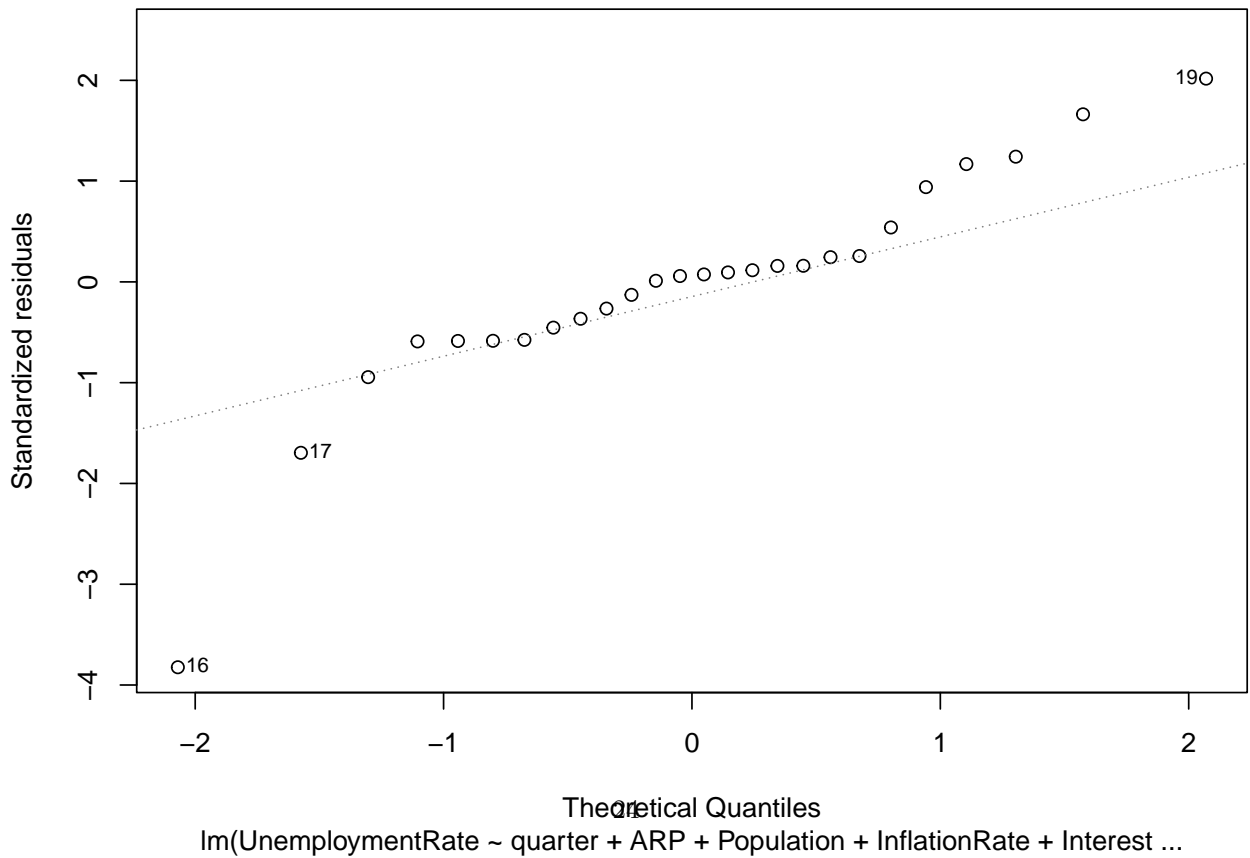
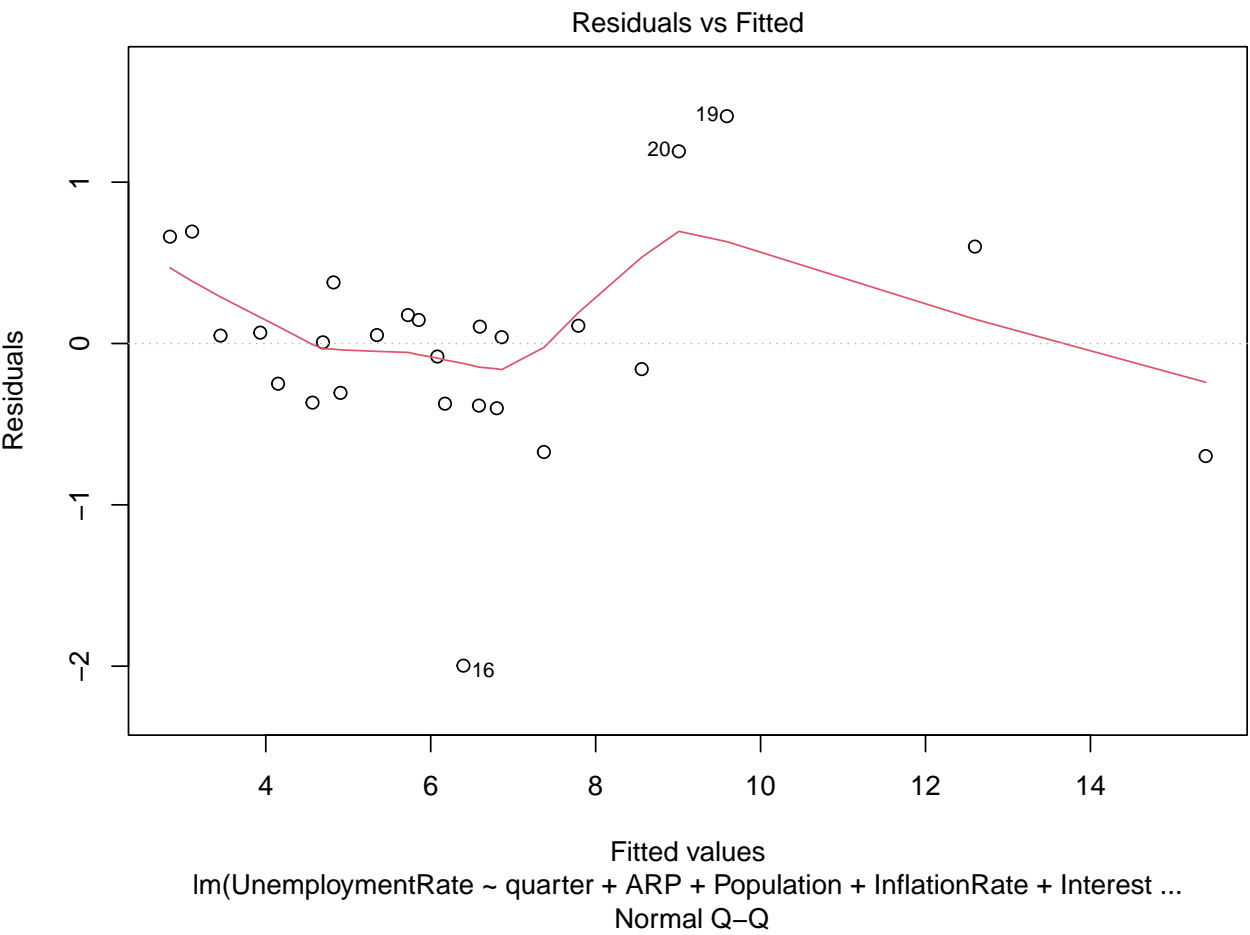
Appendix.D

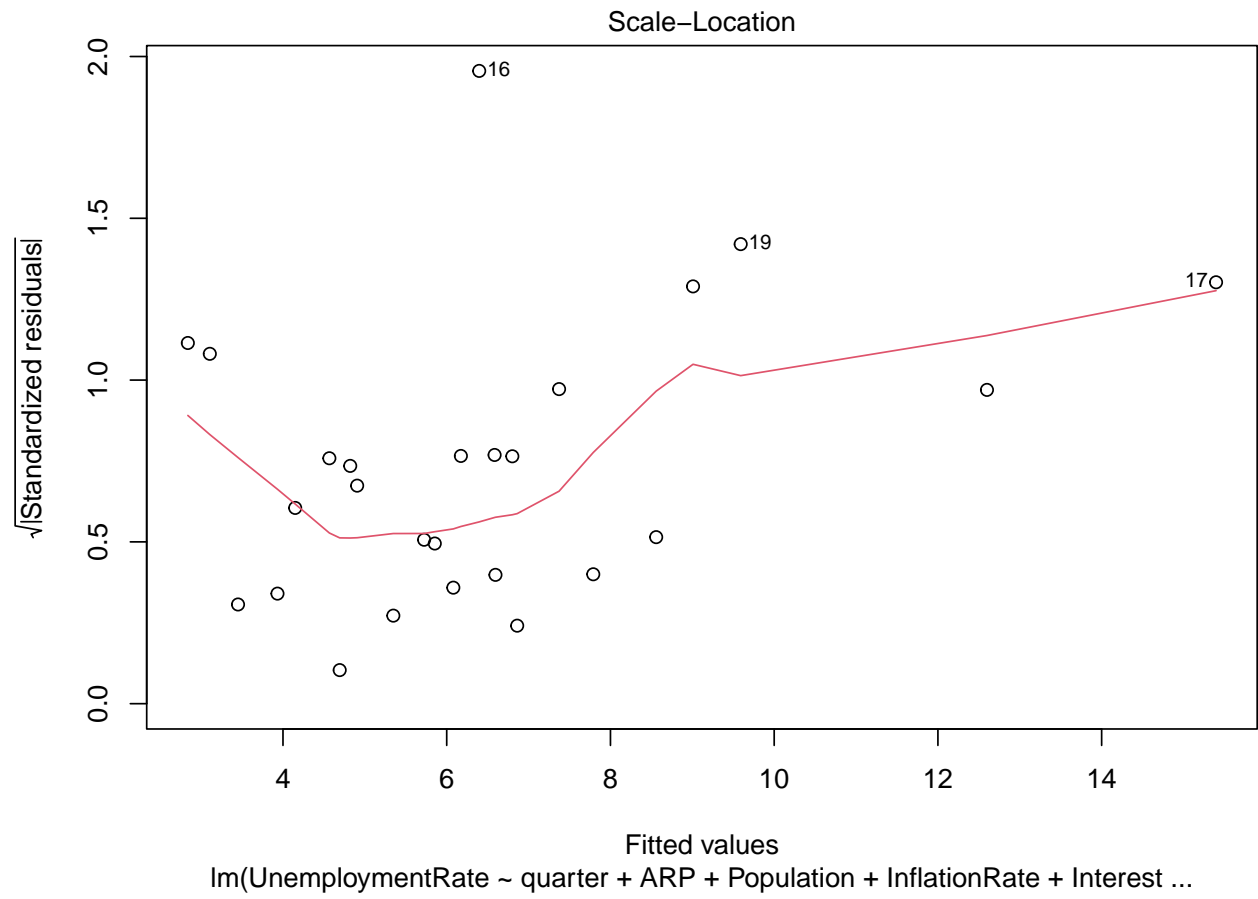


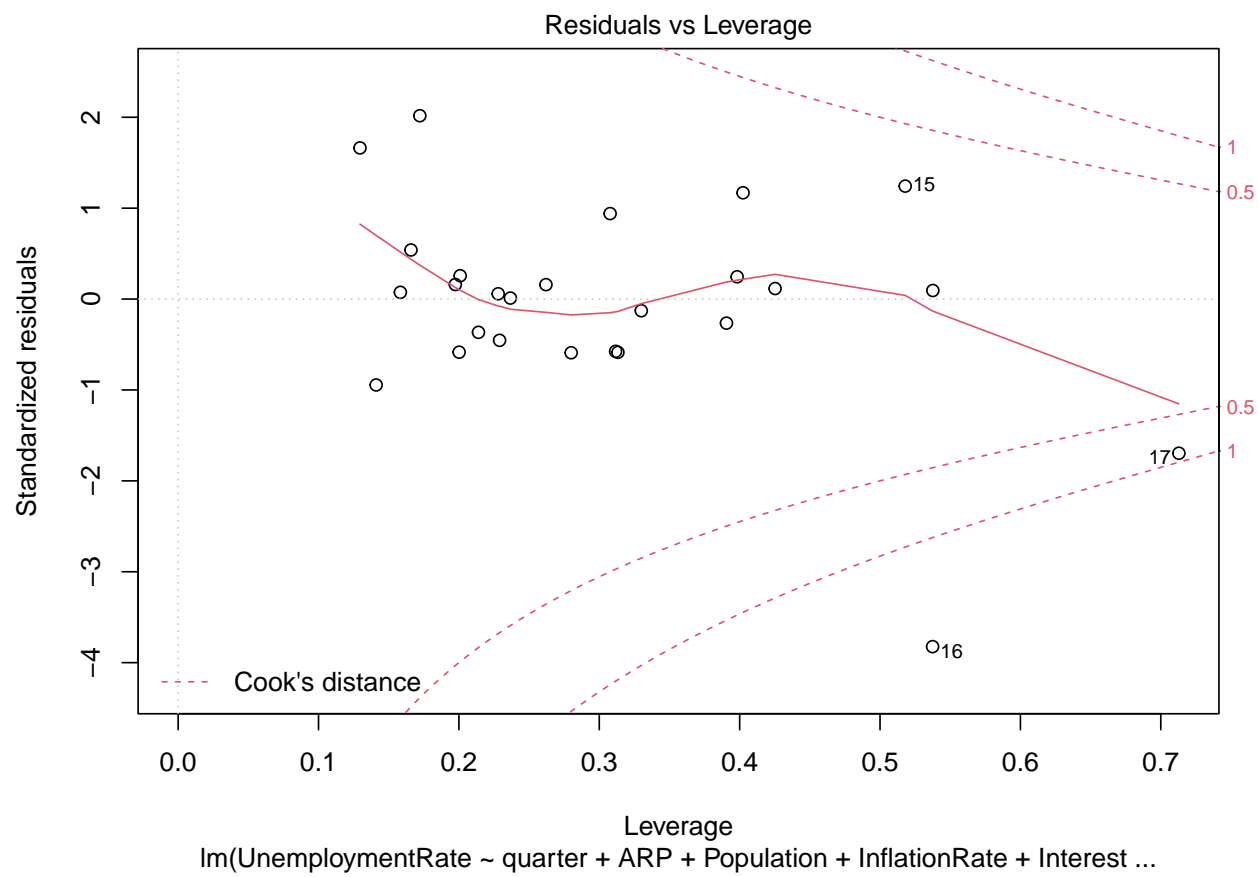




Appendix.E







Reference

- Arel-Bundock, Vincent. 2022. *Modelsummary: Summary Tables and Plots for Statistical Models and Data: Beautiful, Customizable, and Publication-Ready*. <https://CRAN.R-project.org/package=modelsummary>.
- Gali, Jordi, J. David, López-Salido. 2007. “Understanding the Effects of Government Spending on Consumption.” *Journal of the European Economic Association*. <https://doi.org/10.1162/JEEA.2007.5.1.227>.
- Khurshid Adnan, Khan Khalid. 2021. “How COVID-19 Shock Will Drive the Economy and Climate? A Data-Driven Approach to Model and Forecast.” *Environmental Science and Pollution Research*. <https://doi.org/10.1007/s11356-020-09734-9>.
- Long, Jacob A. 2020. *Jtools: Analysis and Presentation of Social Scientific Data*. <https://cran.r-project.org/package=jtools>.
- Philipp Schauburger, Luca Braglia, Alexander Walker. 2021. *Package ‘Openxlsx’*. <https://github.com/ycphs/openxlsx>.
- R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Saad B. Omer, Carlos del Rio, Preeti Malani. 2020. “The COVID-19 Pandemic in the US: A Clinical Update.” <https://jamanetwork.com/journals/jama/article-abstract/2764366>.
- Sjoberg, Daniel D., Karissa Whiting, Michael Curry, Jessica A. Lavery, and Joseph Larmarange. 2021. “Reproducible Summary Tables with the Gtsummary Package.” *The R Journal* 13: 570–80. <https://doi.org/10.32614/RJ-2021-053>.
- Suzi LeVine. 2021. “Building Better Benefits Solutions for the Future.” https://blog.dol.gov/2021/03/12/building-better-benefits-solutions-for-the-future?_ga=2.99927510.93452783.1649147194-1007578703.1649147194.
- The U.S. Government Publishing Office. 2021. “AMERICAN RESCUE PLAN ACT OF 2021.” <https://www.congress.gov/117/plaws/publ2/PLAW-117publ2.pdf>.
- The White House. 2021. “The American Rescue Plan.” <https://www.whitehouse.gov/wp-content/uploads/2021/03/American-Rescue-Plan-Fact-Sheet.pdf>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Zhu, Hao. 2021. *kableExtra: Construct Complex Table with ‘Kable’ and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.