

Analyzing COVID-19 Stimulus Checks and their Impact on Consumer Price Index

Qiaojuan Tu, Siting Liu, Mark Kelly

12/3/2021

PROJECT DESCRIPTION

The coronavirus disease 2019 (COVID-19) pandemic is having a profound impact on the United States, especially on its economic aspects. In response to prevent the economy from going into a recession, the United States government passed the Coronavirus Aid, Relief, and Economic Security, known as **the CARES Act**, on March 27, 2020. To provide various means of aid to help American people get over the hard time as well as help the economy to recover, under the CARES Act, starting in March 2020, payments of \$1,200 for each eligible adult and \$500 for each qualifying child under age 17 were disseminated. In late December of 2020, **the Tax Relief Act** of 2020 was enacted, and in early March 2021, **the American Rescue Plan** of 2021 was enacted as pandemic continues evolving. (See Table 1 and sources for details)

Table 1: COVID-19 Economic Payments Details by June 2021

Names	Date	Payment Amount
The Coronavirus Aid, Relief, and Economic Security (CARES Act)	27-Mar-20	\$1200 per adult; \$500 per child
The Tax Relief Act	27-Dec-20	\$600 per adult; \$600 per adult
The American Rescue Plan	14-Jan-21	\$1400 per adult; \$1400 per child

By the end of June 2021, in total of around \$350 billion of the three rounds of Economic Impact Payments or called economic stimulus payments had been issued to state and local governments to distribute. Although the “free” cash may help people to support daily spending, people has been observing that goods they buy at grocery stores get more expensive. From an economic perspective, the hypothesis of extra money flows into the market drives the price of goods to go up is reasonable. Therefore, this project will analyze the relationship between the issuance of Economic Impact Payments and the price of goods. In other words, how the increase of price of goods is associated with the COVID-19 relief payments.

In this project, there are two key variables that we are examining — amount of stimulus payments distributed over time and the Consumer Price Index (CPI) for each month. The CPI is a composite measurement of trends in the prices of a market basket of consumer goods and services and the data for CPI is obtained from the FRED Economic Data website.(Economic) The raw data for the stimulus checks is obtained from the website of official Internal Revenue Service of the United States.(Consumer)

RESEARCH QUESTIONS

The project is targeting the following of the research questions:

Question 1: Does price of goods shows abnormal increase during the COVID-19 pandemic periods?

Question 2: How does Economic Relief Payments during the COVID-19 pandemic impact the price of normal goods.

STATISTICAL QUESTIONS

To answer the research questions, we investigated the following statistical questions:

Question 1: Does CPI change more over the COVID-19 pandemic period than it's previous years?

Question 2: Does issuance of Economic Relief Payments over the COVID-19 pandemic period have a significant impact on the price of goods?

VARIABLES OF INTEREST

We analyzed the two variables that obtained in the project; stimulus is used as explanatory variable, and the CPI is used as response variable. Since the data obtained only contains the CPI for 4 region of the United State, the stimulus is adjusted to be the monthly average of stimulus for each rounds of payments. Table 2 provides the name and a brief description of each variable along with the variable type and units.

Table 2: Variables Attributes

Variables	Description	Type	Comments
CPI	The Consumer Price Index (CPI) is measured monthly for each Region	Numerical	Index 1982 - 1984 = 100, Seasonally Adjusted
Stimulus	The adjusted average relief payments for each U.S. regions within each rounds of payments	Numerical	Adjusted stimulus relief payment for each states within each rounds
Region	The region that the state belongs	Categorical	4 Levels: South, West, Mideast, Northeast
Month	The month that the data is measured	Categorical	Month

EXPLORATORY DATA ANALYSIS (EDA)

The value for the CPI data is obtained for both pre and during COVID-19 pandemic. We looked at the plot of CPI over the 15 months during pandemic (Figure 2) and before the pandemic (Figure 1), the slope for the CPI over time during the pandemic is more steeper than it's previous slope. Further more, we calculated the averaged change in CPI for both periods, and the results shows in table 3 which the results matches the plots of Figure 1 and Figure 2.

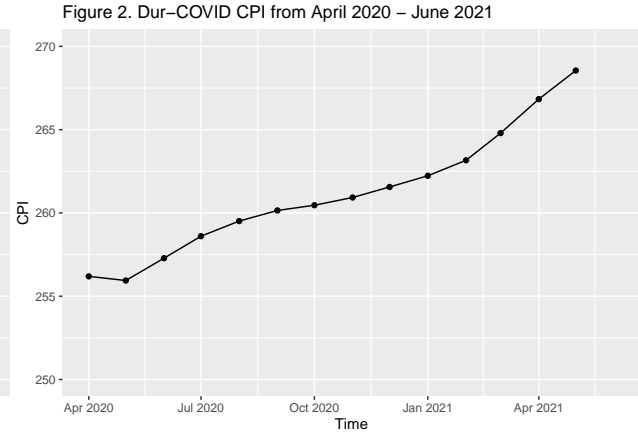
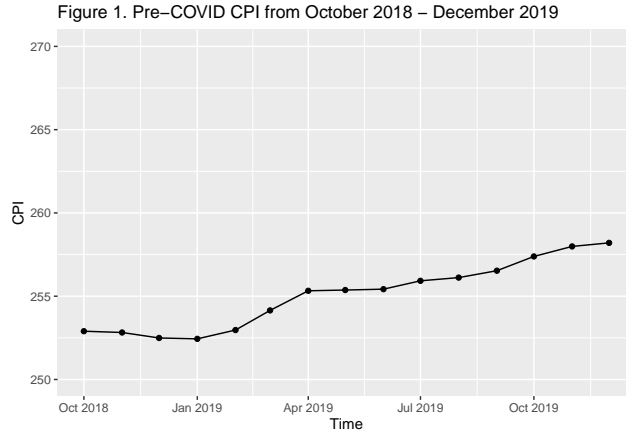


Table 3: Averaged change of CPI in both Pre and During COVID-19

Period	Averaged Monthly Change in CPI
Pre-COVID 19	0.3788571
During-COVID 19	1.0563570

This exploratory data analysis answers the first research question that we conclude the CPI changes nearly three times more over the COVID-19 pandemic period compared to 15 month ahead of it. This observation confirms the rationality of the second research question and further statistical analysis about the potential association of the two variables will be conducted in the section of statistical analysis.

Next, the data that we have lists all of the values for the CPI and the total amount of Stimulus checks in chronological order, starting at April 2020 to June 2021 by the regions. This gives us a hint that we want to look at the CPI and Stimulus check totals by region. We first want to plot the time series for the CPI by the regions of the Untied States.

Figure 3.(next page) gives a detailed look at what the CPI looks like in each region. From here, we can see that each of the region's CPI are non-stationary. Stationarity means that there is a constant mean over the given time period of our data and a variance that does not depend on the time. Since our data is not stationary, we need to transform our data in order to make it stationary. Once our data is stationary, we are able to then make inferences about our data.

In order to make our data stationary, we use time series regression. This process removes the nonstationarity from our model, which in turn will allow us to use these results to answer our questions.

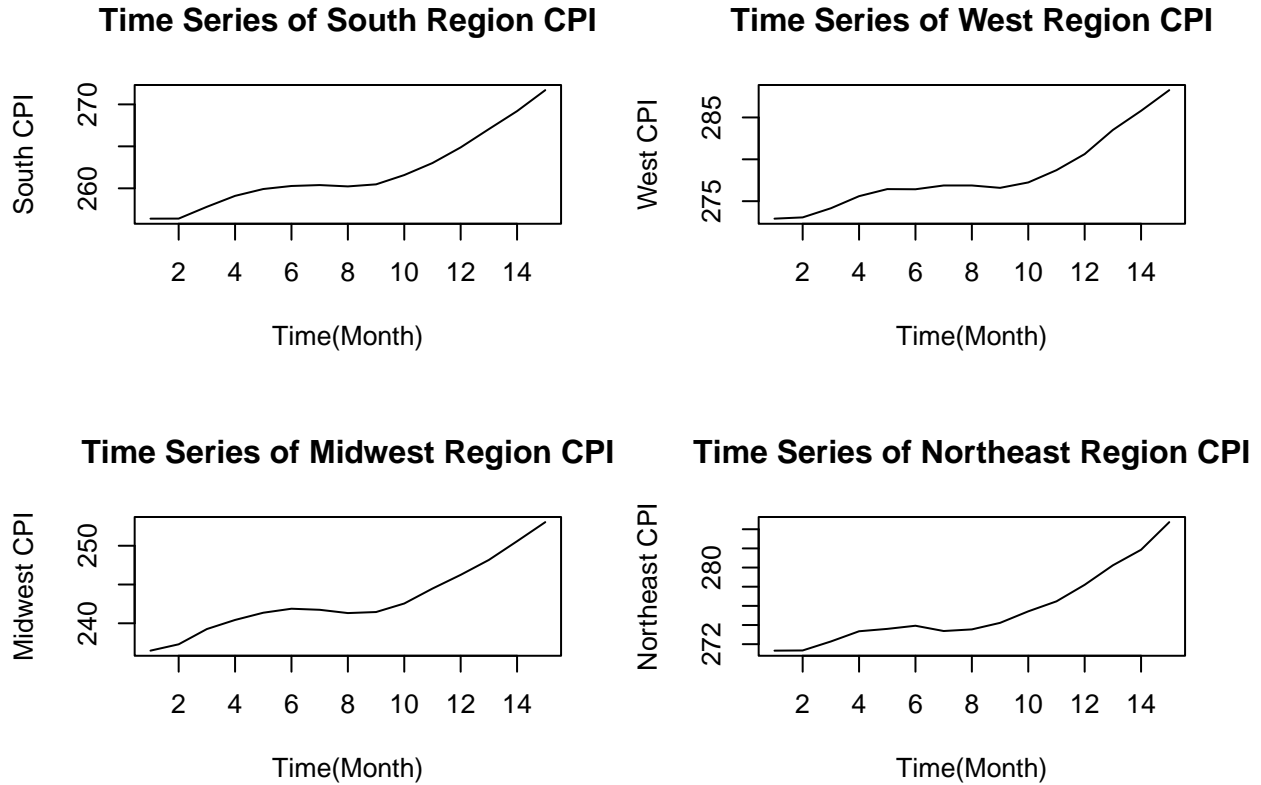


Figure 3. Times Series for CPI by regions of the U.S.

STATISTICAL ANALYSIS

In order to use regression to manipulate our data, we must first find out which of the time series models best fits the data. This process is done by simulating ARMA(p,q) models with the data as a basis and looking at the results of the output. The coefficient(s) that are given will then be used to create the regression model and contribute the final results.

After checking all possible ARMA(p,q) models, we found that the ARMA(1,0) (also called AR(1)) model provides the most beneficial results in which output shows in Table 4.

Table 4: Results of the ARMA(p,q) Model

Model	Coefficient	Standard.Error	AIC.output
ARMA(1,0)	0.4647	0.1186	308.45

From the output, we can see that the AR(1) model with our data gives us an AIC value equal to 308.45, which was the lowest of all the possible models. The coefficient that we will use for our regression model will be 0.4647. The standard error given in the output tells us that there is no possibility of the value 0 for our coefficient.

Next, the regression model for the data will use the coefficient generated from the AR(1) model. The detailed output of the regression model is provided in the Technical Appendix under “Regression Model Output” and the assumption for the regression model is checked under the Technical Appendix as well.

According to the output, we can see that there is a significant relationship between the CPI and Stimulus checks by the regions of the United States using a 95% confidence level since the p-value for the stimulus checks (called x.new in output) is 0.00958. This number gives the evidence that there is an association

between the payments and the CPI, or the Economic Relief Payments is a significant factor that affects the price of goods during the COVID-19 pandemic period. In addition, the output checked the factor Region and output indicates small p-values for Northeast, South, and West, which shows these three region also contributes the change of CPI during the examined period. In other words, the goods' price of Mideast Region was not impacted significantly by the issuance of relief payments.

RECOMMENDATIONS

This project intends to answer two main research question:

- Does CPI change more over the COVID-19 pandemic period than it's previous years?
- Does the issuance of Economic Relief Payments over the COVID-19 pandemic period have a significant impact on the price of goods?

Since the averaged monthly CPI increased almost three times than it's pre-COVID period, we conclude that there are factors that are contributing to the abnormal change.

According to the conducted times series analysis, we show that there is a significant relationship between the CPI during the COVID-19 pandemic and the COVID Stimulus checks and the regions of the United States. When factoring the regions in the model, we see that the Stimulus checks has a significant effect to United States CPI. Therefore, there is a relationship between the COVID-19 Economic Payments and the increase of price of goods.

CONSIDERATIONS

Although this project has showed that there is a significant relationship between the issuance of COVID-19 pandemic relief payments and the price of goods over this period, distributing the money by the government may not be the only reason that drives the price of goods to go up. Considering the economic concepts behind this, price of goods goes up may simply be the reason of the high demand or shortage of a certain product.

Additionally, this study assumes that people would spend all of the payments that they get within each rounds of payments, and we expect the monetary effect reflects in the CPI in the month that person spends. We all know it's never the case in real world. We suspect there are other factors that contribute the increase of CPI over the COVID-19 pandemic period, and further complicated methods could be used to extend this study.

We would like to give a special "thank you" to Dr. Conor B Ryan, who is an Assistant Professor in the Department of Economics at Penn State, for the help and support to this project.

RESOURCES

The following are sources used to complete this projects:

Economic impact payments. Internal Revenue Service. (n.d.). Retrieved December 7, 2021, from <https://www.irs.gov/coronavirus/economic-impact-payments>.

Consumer price index for all urban consumers: All items in U.S. city average. FRED. (2021, November 10). Retrieved December 7, 2021, from <https://fred.stlouisfed.org/series/CPIAUCSL>.

TECHNICAL APPENDIX

Assumptions Checked for Time Series Model

```
## (Intercept)
##      235.1106
```

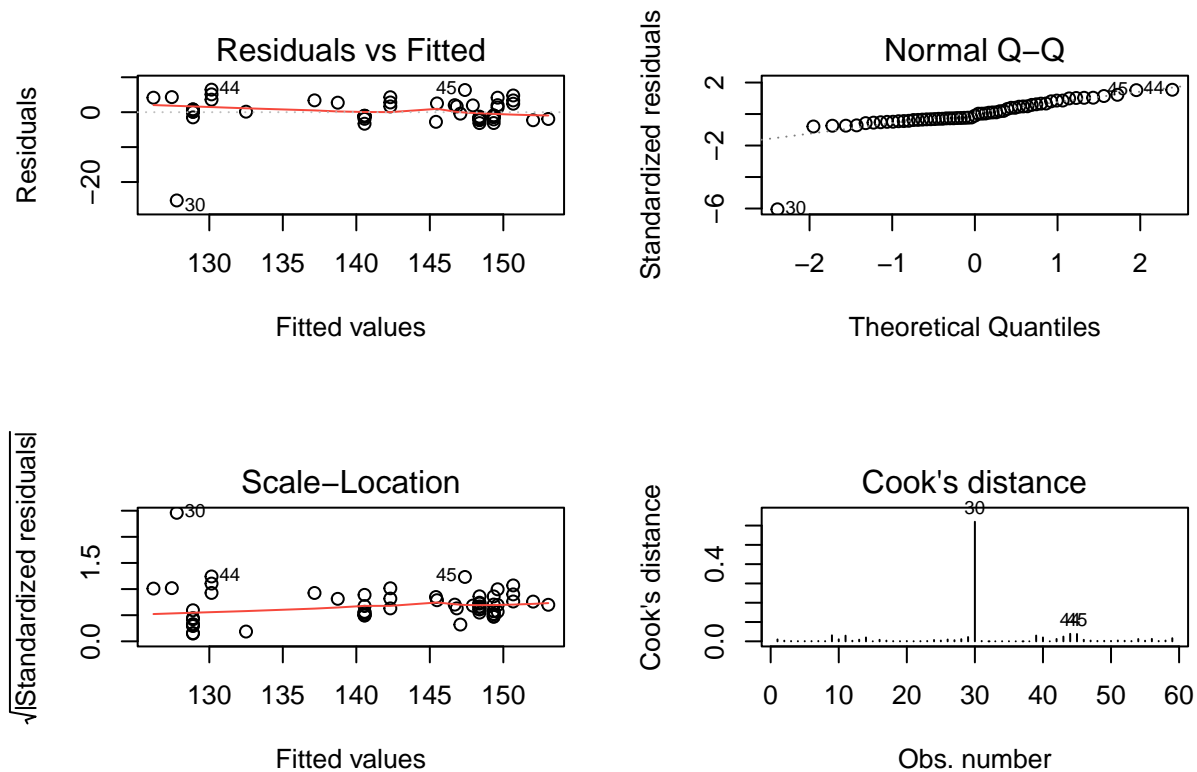
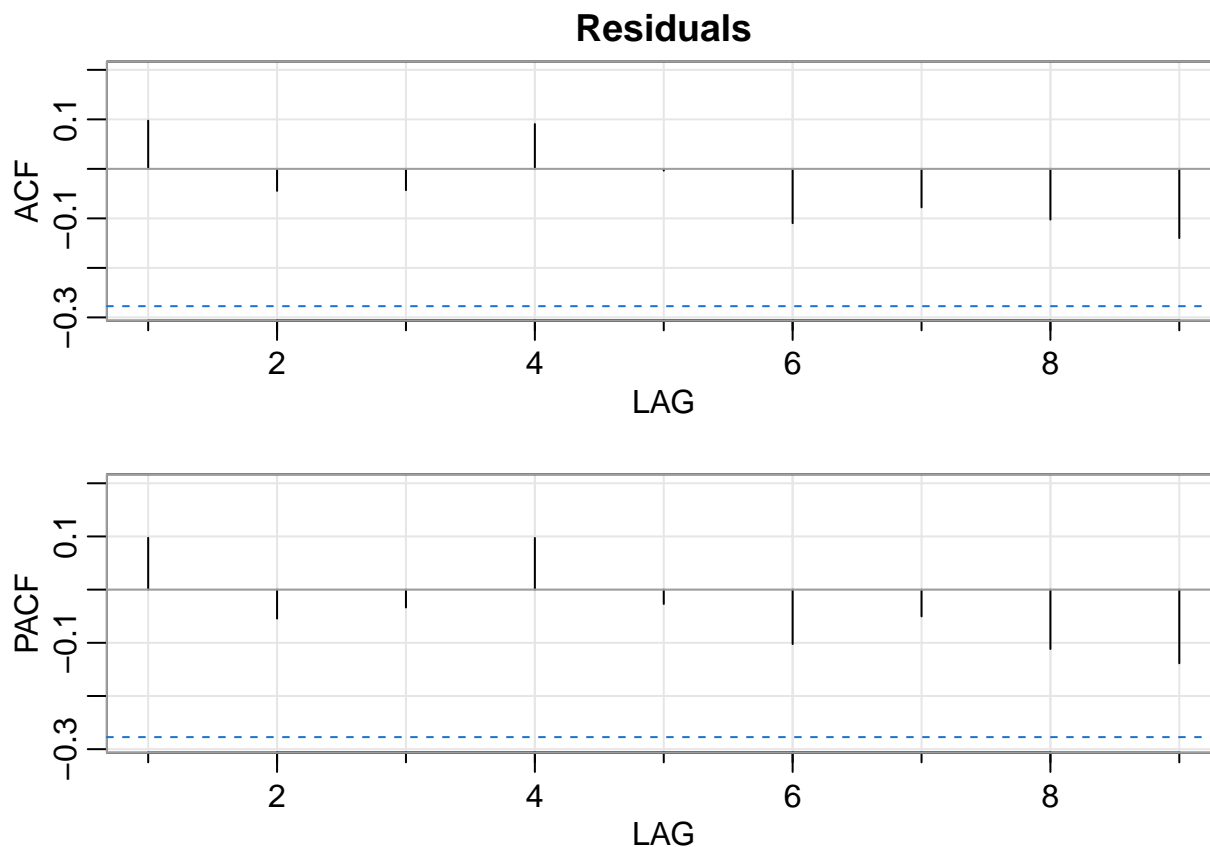


Figure 4. Checking Assumptions

From our regression output in Figure 4, the assumptions are all met. The only concern is for the data point 30 in our data which has a high Cook's distance, however, it is the only point within our data with that high of a Cook's distance. We also do not want to remove said data point since we're working with limited data.

Last, the Autocorrelation Function (ACF) and Partial-Autocorrelation Function (PACF) are checked for the residuals of our model.

Looking at Figure 5, we can see that there is no autocorrelation within the model that we created. This means that none of the data points influence the outcome or results of data points before or after it. In conclusion, we have a regression model that fits all assumptions.



```
##      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9]
## ACF   0.1 -0.04 -0.04 0.09  0.00 -0.11 -0.08 -0.10 -0.14
## PACF  0.1 -0.05 -0.03 0.10 -0.03 -0.10 -0.05 -0.11 -0.14
```

Figure 5. ACF and PACF of the regression

Results of ARMA(p,q) Model

```
##
## Call:
## arima(x = residuals(regmodel2), order = c(1, 0, 0), include.mean = F)
##
## Coefficients:
##      ar1
##      0.4647
## s.e.  0.1186
##
## sigma^2 estimated as 9.636:  log likelihood = -153.22,  aic = 308.45
```

Regression Model Output

```
## (Intercept)
##      235.1106
##
## Call:
## lm(formula = y.new ~ x.new + Master$Region[-1])
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -25.2678  -1.5233  -0.4215   2.2237   6.4449
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.259e+02  1.631e+00  77.161 < 2e-16 ***
## x.new          1.168e-06  4.347e-07   2.686 0.00958 **
## Master$Region[-1]Northeast 1.937e+01  1.589e+00  12.190 < 2e-16 ***
## Master$Region[-1]South    1.088e+01  1.659e+00   6.556 2.17e-08 ***
## Master$Region[-1]West     2.052e+01  1.589e+00  12.914 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.35 on 54 degrees of freedom
## Multiple R-squared:  0.8021, Adjusted R-squared:  0.7874
## F-statistic: 54.71 on 4 and 54 DF,  p-value: < 2.2e-16
```

R code

The following R code is used to conduct this study, and if wish to replicate, a git hub repository is available here at

```
# clean up & set default chunk options
rm(list = ls())
knitr::opts_chunk$set(echo = FALSE)

# packages
library(data.table)
library(readr)
library(knitr)
library(dplyr)
library(ggplot2)
library(kableExtra)
library(lubridate)
library(TSA)
library(astsa)

# inputs
Payment <- read.csv("Pmt_detail.csv")
setnames(Payment, c("Names", "Date", "Payment Amount"))

Variables <- read.csv("Variables_Table.csv")
setnames(Variables, c("Variables", "Description", "Type", "Comments"))

Diff <- read.csv("CPI_Diff.csv")
setnames(Diff, c("Period", "Averaged Monthly Change in CPI "))

AR_Results <- read.csv("ARMA_Results.csv")
#setnames(AR_Results, c(""))
CPI_PreCovid <- read.csv("CPIAUCSL_2018_2019.csv")
CPI_PreCovid$DATE <- mdy(CPI_PreCovid$DATE)

CPI_DurCovid <- read.csv("CPIAUCSL_2020April_2021_June.csv")
```



```

CPI_DurCovid$DATE <- ymd(CPI_DurCovid$DATE)

CPI_Pre <-
  ggplot(CPI_PreCovid, aes(x = CPI_PreCovid$DATE, y = CPI_PreCovid$CPIAUCSL)) +
  labs(x = "Time", y = "CPI") + geom_point() + geom_line() +
  ggtitle("Figure 1. Pre-COVID CPI from October 2018 - December 2019") +
  ylim(250, 270)

CPI_Pre

CPI_Dur <- ggplot(CPI_DurCovid,
  aes(x = CPI_DurCovid$DATE, y = CPI_DurCovid$CPIAUCSL)) +
  labs(x = "Time", y = "CPI") + geom_point() + geom_line() +
  ggtitle("Figure 2. Dur-COVID CPI from April 2020 - June 2021") +
  ylim(250, 270)
CPI_Dur

#calculating the average change in CPI for both pre and during covid
Diff_Pre <- data.frame(diff(CPI_PreCovid$CPIAUCSL))
setnames(Diff_Pre, c("Value"))
mean(Diff_Pre$Value)

Diff_dur <- data.frame(diff(CPI_DurCovid$CPIAUCSL))
setnames(Diff_dur, c("Value"))
mean(Diff_dur$Value)
Master <- fread("./Master.csv")
plot.ts(Master$CPI, main="Time Series of CPI")

South_Final <- fread("./South_Final.csv")
West_Final <- fread("./West_Final.csv")
Midwest_Final <- fread("./Midwest_Final.csv")
Northeast_Final <- fread("./Northeast_Final.csv")
par(mfrow=c(2,2))
plot.ts(South_Final$CPI, main="Time Series of South Region CPI",
  xlab = "Time(Month)", ylab = "South CPI")
plot.ts(West_Final$CPI, main="Time Series of West Region CPI",
  xlab = "Time(Month)", ylab = "West CPI")
plot.ts(Midwest_Final$CPI, main="Time Series of Midwest Region CPI",
  xlab = "Time(Month)", ylab = "Midwest CPI")
plot.ts(Northeast_Final$CPI, main="Time Series of Northeast Region CPI",
  xlab = "Time(Month)", ylab = "Northeast CPI")
x=ts(Master$Stimulus)

y=ts(Master$CPI)
y.new<-y[-1]-0.4647*y[-length(y)]

## new xt
x.new<-x[-1]-0.4647*x[-length(x)]

```

```

regmodel2.new <- lm(y.new ~ x.new + Master$Region[-1])
coef(regmodel2.new)[1]/(1-0.4647)

## look at the residuals
## assumptions
par(mfrow=c(2,2))
plot(regmodel2.new,which=1:4)
x=ts(Master$Stimulus)

y=ts(Master$CPI)
y.new<-y[-1]-0.4647*y[-length(y)]

## new xt
x.new<-x[-1]-0.4647*x[-length(x)]
regmodel2.new <- lm(y.new ~ x.new + Master$Region[-1])
acf2(residuals(regmodel2.new), main = "Residuals")
## transform our data
y.new<-y[-1]-0.4647*y[-length(y)]

## new xt
x.new<-x[-1]-0.4647*x[-length(x)]

regmodel2.new <- lm(y.new ~ x.new + Master$Region[-1])
coef(regmodel2.new)[1]/(1-0.4647)

summary(regmodel2.new)

```