Paper

目录

[Paper 1](#_Toc88241036)

[Introduction 1](#_Toc88241037)

[Show Data 2](#_Toc88241038)

[Figures 2](#_Toc88241039)

[Deal with data 3](#_Toc88241040)

[VAR MODEL 4](#_Toc88241041)

[Linear Model 4](#_Toc88241042)

[Unit root test 5](#_Toc88241043)

[Multivariate mixing test 5](#_Toc88241044)

[Model diagnosis 5](#_Toc88241045)

[Granger Test 6](#_Toc88241046)

[Model Prediction 7](#_Toc88241047)

## Introduction

Vector Autoregression (VAR) model is the most commonly used joint model of multiple return on assets.

VAR(1):

Where is k dimensions constant vector, is a square matrix of order constant. is a weakly stationary columns with unrelated sequences. , . It is often assumed that obeys multivariate normal distribution

Let , ,

When k=2, the model becomes:

Then the two sequences obey a single AR(1) model. If and are unrelated, the sequence is unrelated to the sequence. Such a separated sequence is said to be uncoupled.

Vars are not built on some economic theory that imposes a theoretical structure on the equations. Every variable is assumed to influence every other variable in the system, which makes a direct interpretation of the estimated coefficients difficult. Despite this, VARs are useful in several contexts:

1. forecasting a collection of related variables where no explicit interpretation is required;
2. testing whether one variable is useful in forecasting another (the basis of Granger causality tests);
3. impulse response analysis, where the response of one variable to a sudden but temporary change in another variable is analyzed;
4. forecast error variance decomposition, where the proportion of the forecast variance of each variable is attributed to the effects of the other variables.

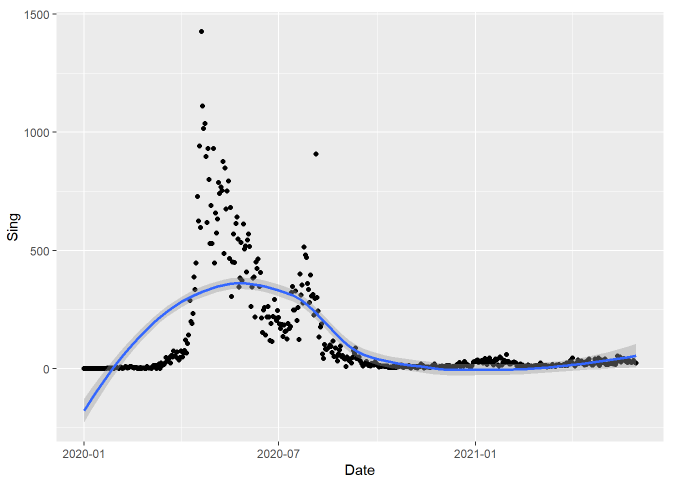
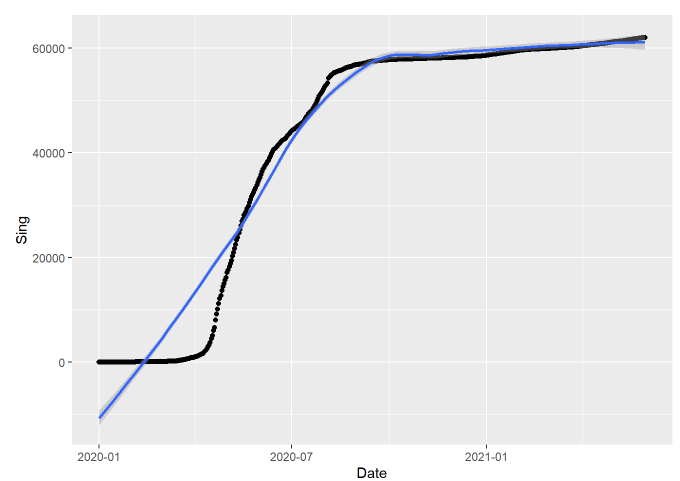
## Show Data

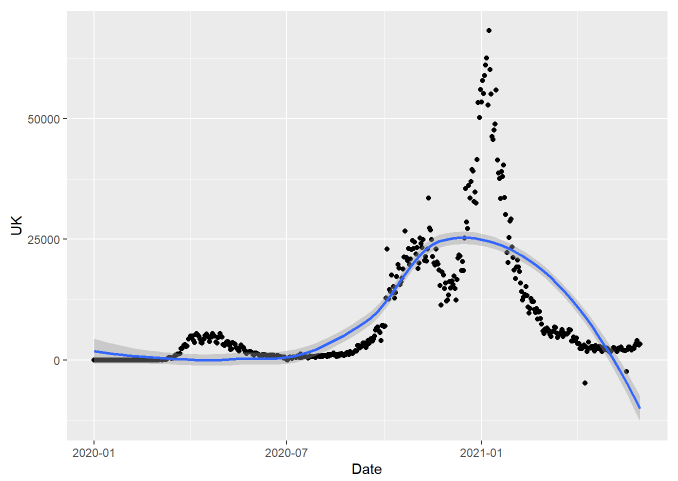
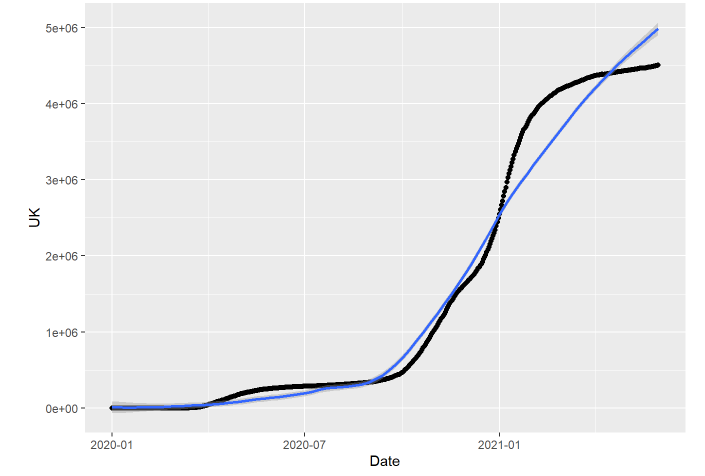
There are six tables which are unemployment, economic, covid, health, response, stringency. All have columns Date, Sing UK, US which represent time and three countries.

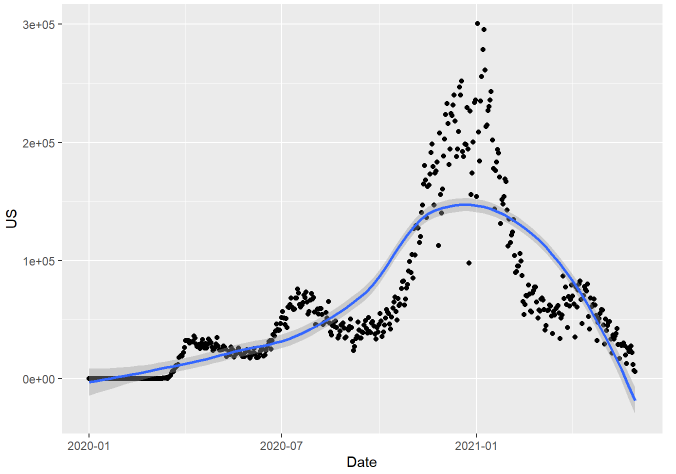
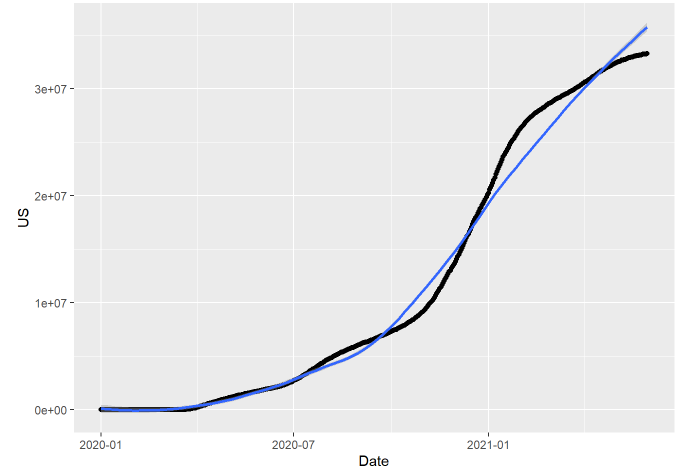
Unemployment is special for different Date in which every element represents a month from 2020-01 to 2021-05.

Other tables have Date in which every element represents a day from 2020-01-01 to 2021-05-31.

### Figures For covid

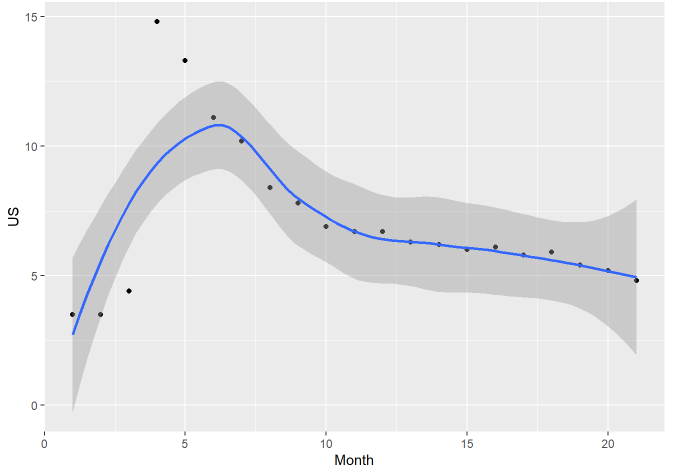
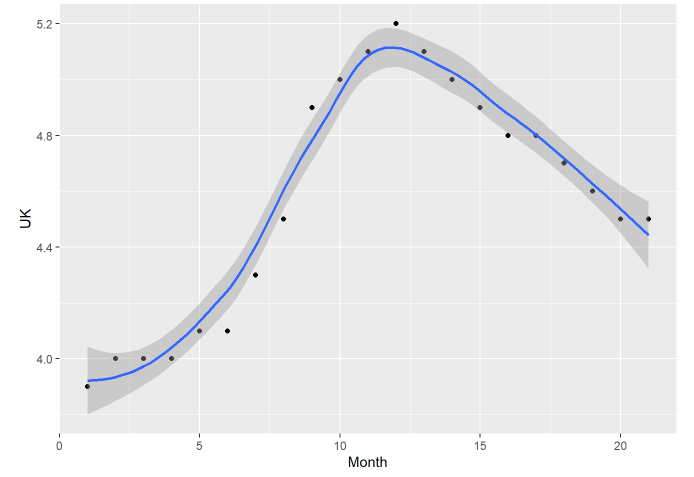
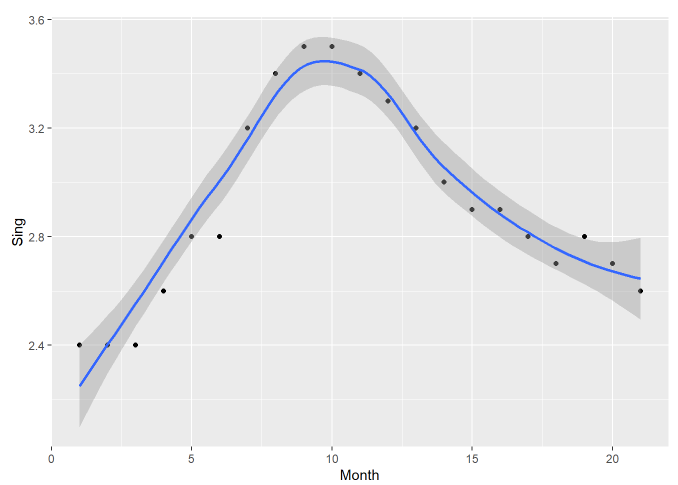
 

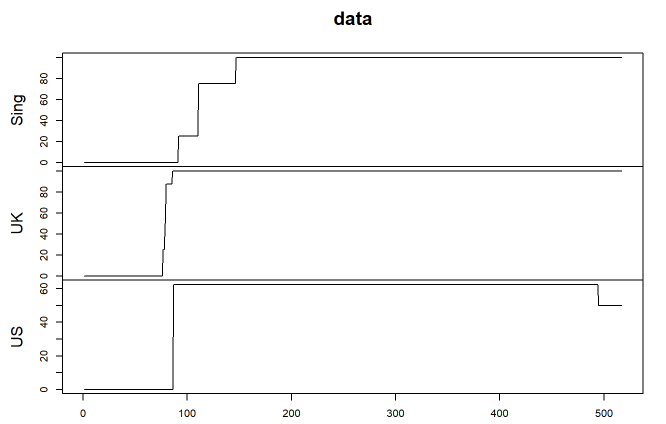
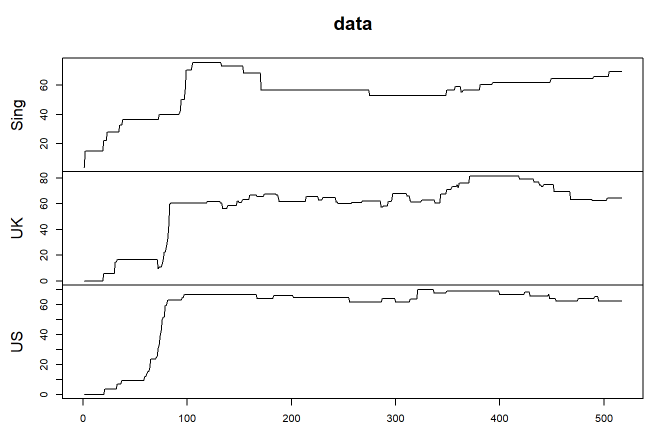
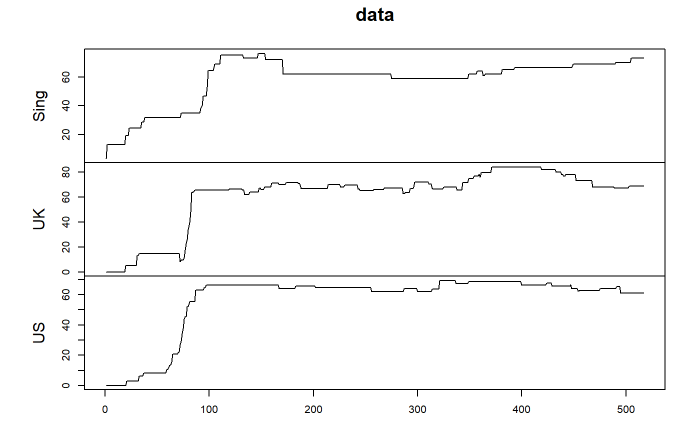
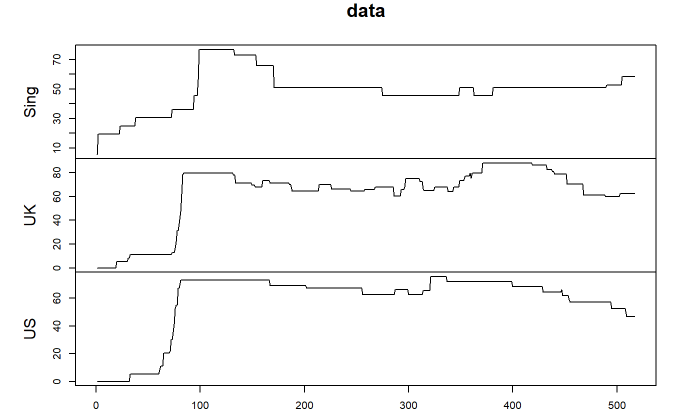
The pictures above represent covid-19 cases in Singapore, the United Kingdoms, the United States. The left three figures represent cases every day in corresponding countries, the right three figures represent cumulated cases at corresponding date in corresponding countries.

### Figures for unemployment



The three figures above represent unemployment data in corresponding country for every month from 2020-01 to 2021-05. And there are confident intervals that mainly convers the point and fitted curves that can predict the future points.

### Figures for economic, health, response, stringency

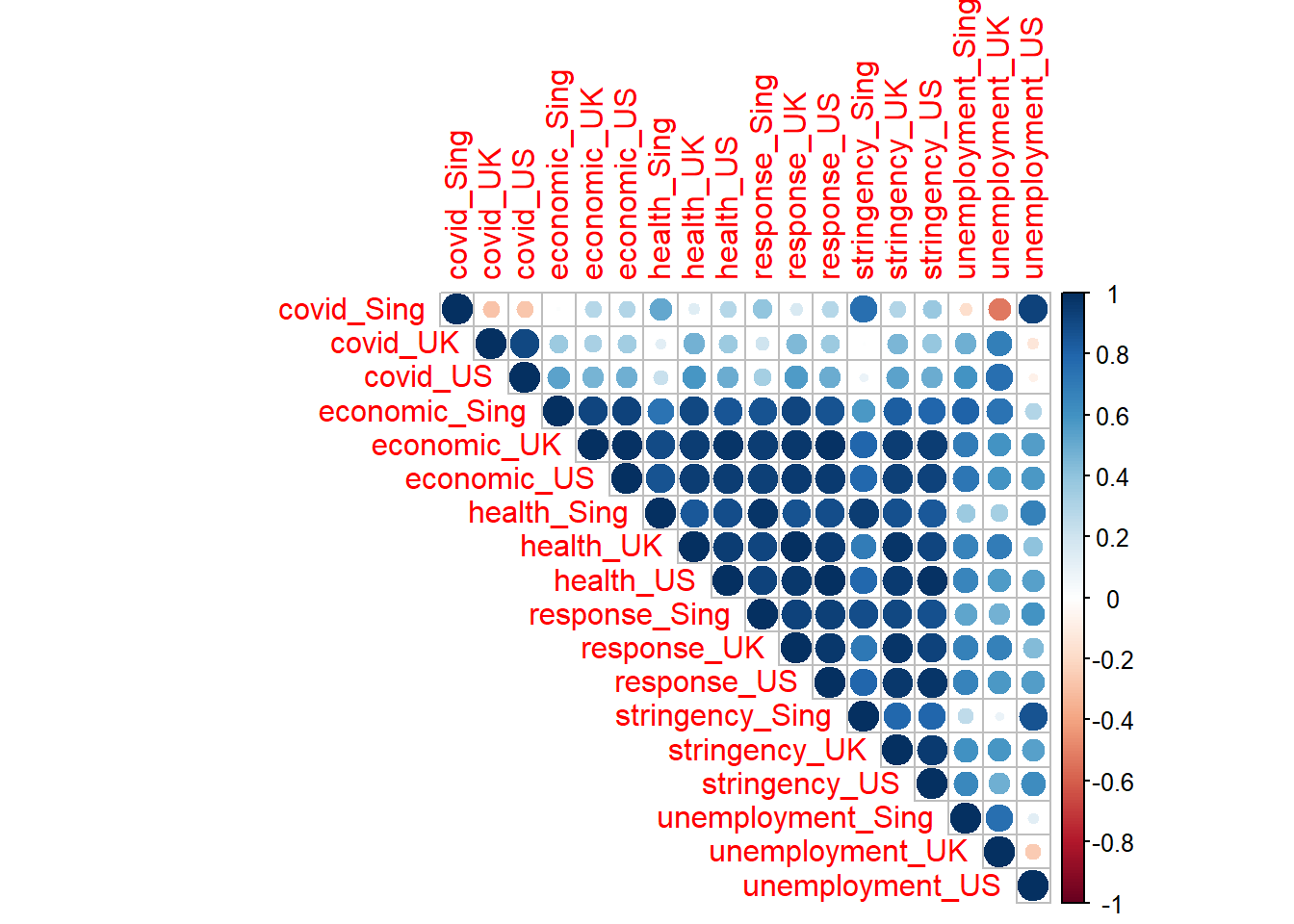
   

The four figures are for economic, health, response, stringency correspond to the orders.

## Deal with data

First we combine covid, economic, health, response, stringency by column Date. Each column are renamed to combination of source table and corresponding country. We name the table total.

Sum the table total grouped by the Month. We get 17 rows which represent 2020-01 to 2021-05, then combine it and unemployment by corresponding Month. We name it totalbymonth. We get the correlation matrix of totalbymonth. Then we plot the heatmap of the correlation matrix.



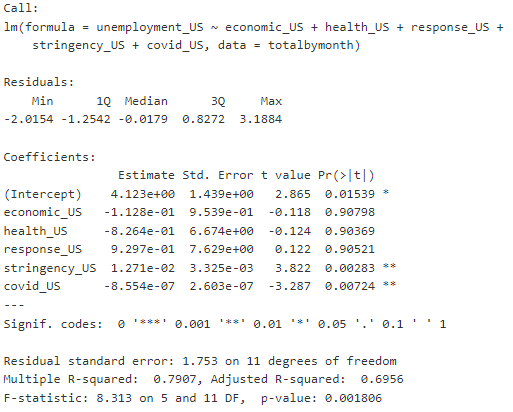
We can see the correlation between each variables directly. For example, we can see that economic have a larger correlation with many other variables.

## VAR MODEL

### Linear Model

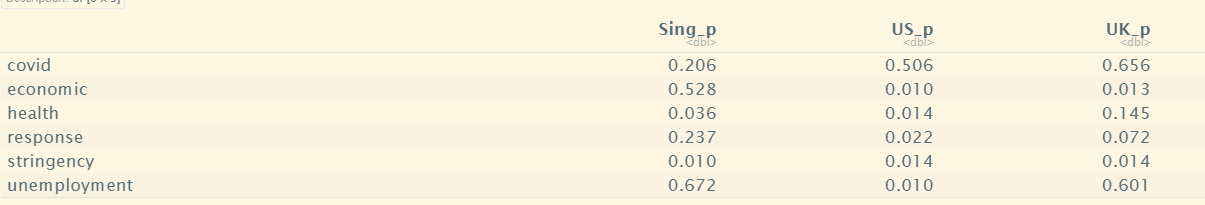
This is a small try to find out whether there are linear relations between unemployment and other variables. In all the countries, the linear models are not quite good with small R-squared and small q values.

Such as linear model for data in US



### Unit root test

We use adf test function to get the unit root which is directly shown by p value.



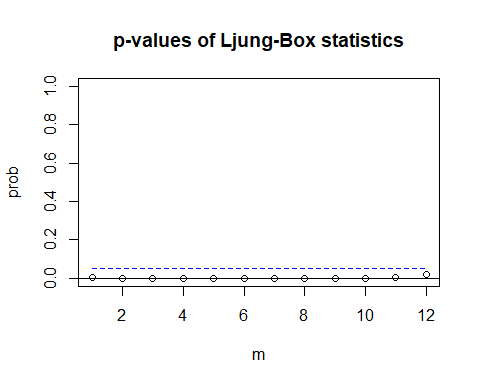
For Singapore, covid, economic, response, unemployment are unstable.

For US, covid are unstable for p value is much larger than 0.05 which is consistent with the breakout in US.

For UK, covid, health and unemployment are unstable.

### Multivariate mixing test

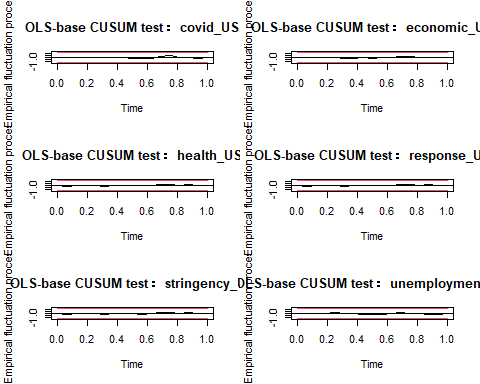
We use p value of Ljung-Box statistics. The function mq is used to get the figures of p values of different lags. This can test the null hypothesis that multivariate time series is wide white noise, that is, is weak stationary column and has no sequence autocorrelation, and can have synchronous intercomponent correlation.



### Model diagnosis

To see if the VAR model root is inside the unit circle. Follow-up analysis can only be carried out if the roots of VAR model are all in the unit circle.

In the curve generated by this test, the residual accumulation and curve take time as abscess, and two critical boundaries are drawn in the figure. If the accumulation and exceed these two critical boundaries, it indicates that the parameter is not stable.



From the figures, all are inside the bottom and up lines.

Residual tests and tests show that the model results are stable.

### Granger Test

From table totalbymonth, we separate tables of each countries of six variables. Then we do the granger test with the for loop which compare each two different variables.

for (i in 2:7){  
 for(j in 2:7){  
 if (j!=i){  
 GX=colnames(totalbymonth.Sing)[i]  
 GY=colnames(totalbymonth.Sing)[j]  
 EGtest <- grangertest(totalbymonth.Sing[,j] ~ totalbymonth.Sing[,i],order = 1, data = totalbymonth.Sing)  
 if (EGtest$`Pr(>F)`[2]<0.1){  
 print(paste(GX," is a cause for",GY))  
 }  
 }  
 }  
}

We get the result.

## [1] "covid\_Sing is a cause for economic\_Sing"

## [1] "covid\_Sing is a cause for unemployment\_Sing"

## [1] "economic\_Sing is a cause for covid\_Sing"

## [1] "economic\_Sing is a cause for health\_Sing"

## [1] "economic\_Sing is a cause for stringency\_Sing"

## [1] "health\_Sing is a cause for economic\_Sing"

## [1] "response\_Sing is a cause for economic\_Sing"

## [1] "response\_Sing is a cause for health\_Sing"

## [1] "stringency\_Sing is a cause for economic\_Sing"

## [1] "unemployment\_Sing is a cause for covid\_Sing"

From the for loops of three countries tables, we find that economic is a cause of many variables in terms of granger test which is consistent with the correlation heatmap.

### Model Prediction

We use the table total, separate for each countries and set the train size as the 75% of the rows. Then use var select to get AIC, HQ, SC,FPE of the train set for every countries

## AIC(n) HQ(n) SC(n) FPE(n)

## 7 2 1 7

USE SC which is also BIC. Get the VAR(1) model and use Portmanteau Test to testify the model

For Singapore, we get result:

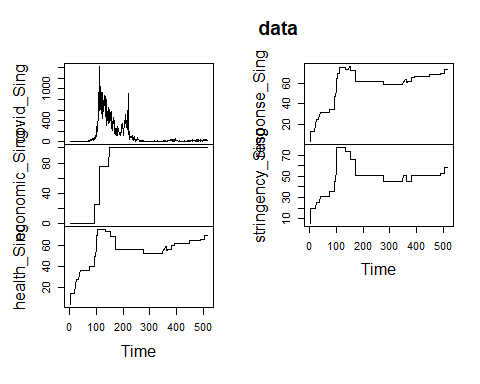
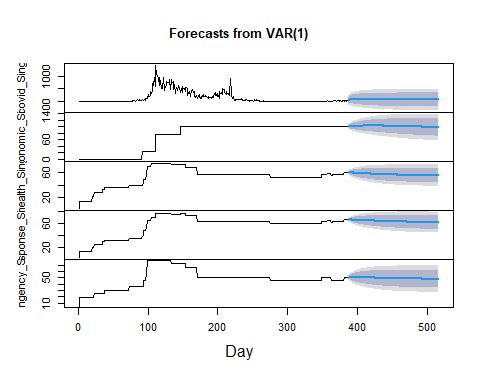
## Portmanteau Test (asymptotic)

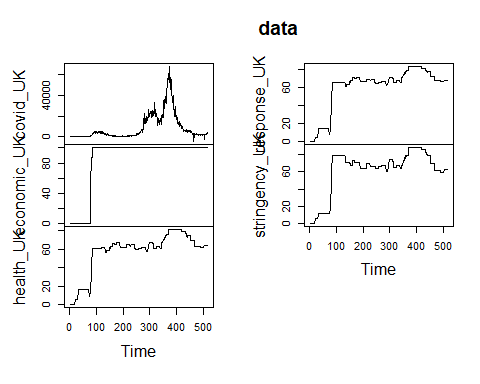
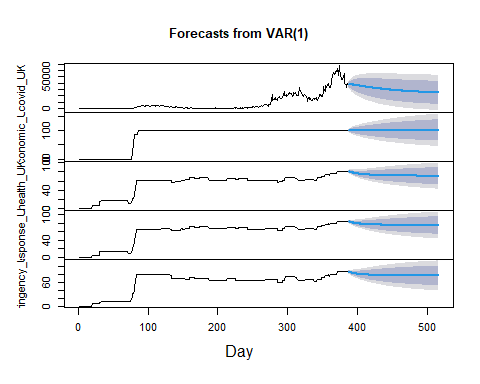
##

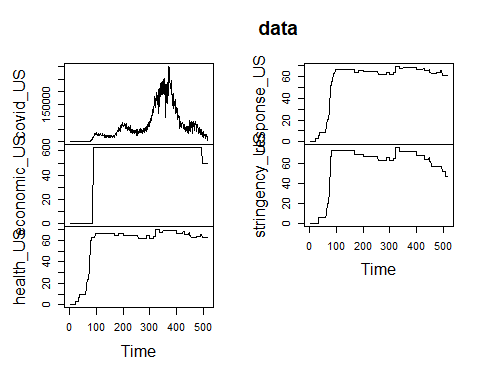
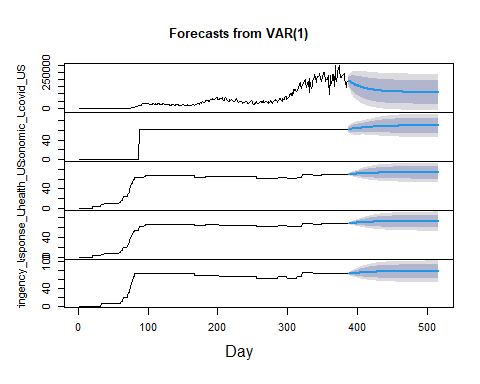
## data: Residuals of VAR object total.Sing.var

## Chi-squared = 467.57, df = 225, p-value < 2.2e-16

Plot the predictions for test size which is the total size minus train size.







It can be seen from the results that the prediction effect of VAR model is similar to that of unitary AR model, which is relatively smooth and tends to a certain value in the later period.