ACTL30007 Actuarial Modelling III

*Assignment 2*

Student Name: *Kelvin Youne Fong Lim Wan*

Student Number: *929715*

Word Count (excluding titles, tables, plots, equations, R outputs and the Appendix): *1087*

***Time Series Analysis on the Australian market in relation to the Chinese market during the Covid-19 pandemic***

June 2, 2020

Introduction:

This report analyses the performance of the Australian financial market in relation to the Chinese financial market during the Covid-19 pandemic. We use the main ‘S&P/ASX 200’ index and focus on its correlation with the main Chinese market index ‘Shanghai Composite’, the number of daily new reported cases in Australia and the number of daily new reported cases in China. We explore these relationships using techniques in time series analysis. This provides us with a better understanding on the movement of the Australian market index during this pandemic.

We have collected data for the daily closing prices of the market indexes and the number of daily new cases in each country for the period between December 31, 2019 and May 28, 2020 inclusive (see Appendix for more details on the dataset structure).

The following notations are used consistently throughout the report:

|  |  |
| --- | --- |
| Symbol | Notation |
|  | Time as an index of the period |
|  | Australian market index ‘S&P/ASX 200’ at time t |
|  | Chinese market index ‘Shanghai Composite’ at time t |
|  | The number of daily new reported cases in Australia at time t |
|  | The number of daily new reported cases in China at time t |
|  | An error term at time t |
|  | The ith regressing parameter |

Exploratory analysis:

From the plots below, it is quite evident that during the month of March, the ‘S&P/ASX 200’ index was devaluing as the cases in Australia was increasing. We also notice a small kink in the ‘S&P/ASX 200’ index at the beginning of February, coinciding with the massive drop in the ‘Shanghai Composite’ index and the start of the increase in the cases in China. We suspect that the cases has a negative relationship with a country’s financial market and that the ‘S&P/ASX 200’ index has some kind of dependency with the ‘Shanghai Composite’ index.

A close up of a map

Description automatically generated

Detrending via first difference:

We detrend the ‘S&P/ASX 200’ index via first difference as

Since the differenced series follow a horizontal linear trend around 0 and its ACF converges quickly to 0, we say that the first difference of the ‘S&P/ASX 200’ index is stationary. This also suggests that the ‘S&P/ASX 200’ index has a linear relationship with time.

A screenshot of a cell phone

Description automatically generated

Detrending via regression (Model 1):

Our first proposed regression model displays a linear relationship between the ‘S&P/ASX 200’ index and the other time series:

The summary of the fitted regression model output suggests that all parameters are significant.

Call:

lm(formula = market\_aus ~ time + market\_chn + cases\_aus + cases\_chn)

Residuals:

Min 1Q Median 3Q Max

-1338.91 -203.86 -6.92 265.97 611.93

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -1.164e+03 1.270e+03 -0.917 0.361

time -8.504e+00 9.435e-01 -9.013 1.08e-15 \*\*\*

market\_chn 2.701e+00 4.140e-01 6.525 1.06e-09 \*\*\*

cases\_aus -1.702e+00 3.155e-01 -5.396 2.71e-07 \*\*\*

cases\_chn 1.412e-01 1.922e-02 7.350 1.33e-11 \*\*\*

---

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

Residual standard error: 318.2 on 145 degrees of freedom

Multiple R-squared: 0.8579, Adjusted R-squared: 0.854

F-statistic: 218.8 on 4 and 145 DF, p-value: < 2.2e-16

The residuals of Model 1 are clearly not stationary since there is still a trend in the residual plot and most bars in the ACF plot stretch beyond the confidence interval (bands).

A screenshot of a social media post

Description automatically generated

From the goodness-of-fit plot below, we see that Model 1 is not very representative of the real data, especially in the first half of the period.

A close up of a map

Description automatically generated

Analysing the relationship between the time series:

To investigate how the ‘Shanghai Composite’ index and the cases in Australia and China affect the ‘S&P/ASX 200’ index, we use a matrix of scatterplots. The ‘S&P/ASX 200’ index has a very high correlation with the ‘Shanghai Composite’ index (0.78) and they seem to have a positive linear relationship. The cases in Australia affects the Australian market negatively (-0.53) and the scatter plot suggests a curvilinear relationship. The correlation between the cases in China and the ‘S&P/ASX 200’ index is the lowest (0.41) but is still significant; there is a possible non-linear positive relationship between the two time series.

A screenshot of a cell phone

Description automatically generated

Detrending via regression (Model 2):

This model takes into consideration the non-linear relationships inferred from the scatterplots, between the ‘S&P/ASX 200’ index and the cases in both Australia and China:

From the summary below, we deduce that all the parameters are relevant to the data. Note that is slightly less relevant as per the \*\* as opposed to the \*\*\* for the other regressors. We also see that the adjusted R-squared (0.9226) is better than the adjusted R-squared in Model 1 (0.854), indicating that Model 2 is better.

Call:

lm(formula = market\_aus ~ time + market\_chn + cases\_aus + cases\_aus\_sq +

cases\_chn + cases\_chn\_sq)

Residuals:

Min 1Q Median 3Q Max

-695.61 -176.19 -27.11 157.53 568.28

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -3.877e+03 1.154e+03 -3.361 0.000997 \*\*\*

time -5.387e+00 8.022e-01 -6.715 4.11e-10 \*\*\*

market\_chn 3.523e+00 3.727e-01 9.451 < 2e-16 \*\*\*

cases\_aus -2.785e+00 7.048e-01 -3.951 0.000122 \*\*\*

cases\_aus\_sq 3.581e-03 1.244e-03 2.879 0.004606 \*\*

cases\_chn 4.197e-01 3.392e-02 12.372 < 2e-16 \*\*\*

cases\_chn\_sq -2.363e-05 2.419e-06 -9.769 < 2e-16 \*\*\*

---

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

Residual standard error: 231.7 on 143 degrees of freedom

Multiple R-squared: 0.9257, Adjusted R-squared: 0.9226

F-statistic: 296.9 on 6 and 143 DF, p-value: < 2.2e-16

The residual plot for Model 2 is choppier than that of Model 1 and does not follow a trend, denoting that the residuals are more uncorrelated to each other. The ACF plot also converges more quickly to zero and more bars lie within the confidence interval. Hence, the residuals are stationary and Model 2 is accurate. Additionally, we note a monthly cycle exhibited in the ACF of residuals, which was masked in the original data.

A screenshot of a cell phone

Description automatically generated

Model 2 follows the ‘S&P/ASX 200’ index better than Model 1, but there is a lot of variation.

A close up of a map

Description automatically generated

Analysing the correlations between the lags of the time series:

Here, we analyse the correlation between the lags of each time series with the ‘S&P/ASX 200’ index. The peak of the ACF plot of the ‘S&P/ASX 200’ index and the peak of the CCF plot with the ‘Shanghai Composite’ are at lag 0, showing an absence of any lead or lag connection.

A screenshot of a cell phone

Description automatically generated

However, the peak of the CCF plot with the cases in Australia is at the negative lag section, with tapering occurring in both directions from that peak. This shows that the ‘S&P/ASX 200’ index leads the cases in Australia. Similarly, the peak of the CCF plot with the cases in China is at a positive lag; the ‘S&P/ASX 200’ index lags the cases in China.

A screenshot of a cell phone

Description automatically generated

Below are grids of scatterplots of the ‘S&P/ASX 200’ index versus the cases in both Australia and China. The first grid indicates that the peak of cross-correlation (-0.54) between the ‘S&P/ASX 200’ index and the cases in Australia occurs at lags of 2, 3 and 4 and the second grid of the cases in China versus the ‘S&P/ASX 200’ index show that the peaks (0.42) occur at lags 5, 6 and 7. Therefore, the ‘S&P/ASX 200’ index leads the cases in Australia by 2, 3, or 4 days and lags the cases in China by 5, 6, or 7 days.

A close up of a map

Description automatically generated

A close up of a map

Description automatically generated

Detrending via regression (Model 3):

Regression model 3 focuses on the cross-correlations of the ‘S&P/ASX 200’ index with the lags described above:

The adjusted R-squared is a little higher than in Model 2 (0.9256 > 0.9226), suggesting that Model 3 is slightly more accurate. However, there are more regressors and some seem to not be as significant (for instance, has only one \*).

Call:

lm(formula = market\_aus ~ time + market\_chn + cases\_aus\_lead2 +

cases\_aus\_lead3 + cases\_aus\_lead4 + cases\_chn\_lag5 + cases\_chn\_lag6 +

cases\_chn\_lag7)

Residuals:

Min 1Q Median 3Q Max

-726.74 -149.90 25.42 183.10 416.44

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 4040.11484 971.14174 4.160 5.74e-05 \*\*\*

time -12.80415 0.72222 -17.729 < 2e-16 \*\*\*

market\_chn 1.04499 0.31751 3.291 0.00128 \*\*

cases\_aus\_lead2 -0.80690 0.34444 -2.343 0.02067 \*

cases\_aus\_lead3 -1.20505 0.27630 -4.361 2.61e-05 \*\*\*

cases\_aus\_lead4 -1.06342 0.34552 -3.078 0.00254 \*\*

cases\_chn\_lag5 0.04877 0.01601 3.045 0.00281 \*\*

cases\_chn\_lag6 0.04745 0.01654 2.868 0.00482 \*\*

cases\_chn\_lag7 0.04933 0.01585 3.112 0.00228 \*\*

---

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

Residual standard error: 231.8 on 130 degrees of freedom

(11 observations deleted due to missingness)

Multiple R-squared: 0.9299, Adjusted R-squared: 0.9256

F-statistic: 215.7 on 8 and 130 DF, p-value: < 2.2e-16

The residual plot is smoother than in Model 2 but the ACF plot displays more lines within the blue bands. However, the lowess line indicates a slight trend. The residuals are not stationary. We again notice the seasonal trend as described in Model 2.

A screenshot of a map

Description automatically generated

The goodness-of-fit plot seems to be smoother than the one in Model 2 but is unsurprisingly less close to the real data. The high peak in mid-February is also unrealistic and unrepresentative of the data.

A close up of a map

Description automatically generated

Conclusion:

From our time series analysis, we conclude that during the Covid-19 pandemic, the ‘S&P/ASX 200’ index has had a decreasing trend with time, a highly positive relationship with the ‘Shanghai Composite’ index, a negative relationship with the number of daily new reported cases in Australia and a very small positive relationship with the number of daily new reported cases in China. Furthermore, we have found that the ‘S&P/ASX 200’ index leads the number of daily new reported cases in Australia by 2-4 days and lags the number of daily new reported cases in China by 5-7 days. There is also the possibility of a monthly cycle for the ‘S&P/ASX 200’ index.

***Appendix:***

Excel manipulation:

The data in the columns ‘China New Daily Cases’ and ‘Australia New Daily Cases’ were downloaded from <https://ourworldindata.org/grapher/daily-cases-covid-19>.

The data in the columns ‘Shanghai Composite’ and ‘S&P/ASX 200’ were downloaded from <https://au.investing.com/indices/major-indices>. The closing prices available were only for days when the markets were open (excluding weekends and public holidays). We have arranged the data so that the closing prices correspond to their respective dates. For the missing values on the weekends and public holidays, we set the values as the closing price on the day the market was last open.

R code:

##### ACTL30007 Assignment 2 (929715)

# load packages

install.packages('astsa')

install.packages('readxl')

library(astsa)

library(readxl)

# import data

dataset <- read\_excel('AssignmentDataPath/china\_vs\_australia\_covid-19.xlsx')

# assign the vector variables

time = time(dataset$Date)

market\_aus = dataset$`S&P/ASX 200`

market\_chn = dataset$`Shanghai Composite`

cases\_aus = dataset$`Australia New Daily Cases`

cases\_chn = dataset$`China New Daily Cases`

# visual analysis

month\_vec = c(2, 33, 62, 93, 123)

par(mfrow = c(4,1))

plot(market\_aus~time, xaxt = 'n', type = 'o', col = 'red', ylim = c(4000, 7500), main = 'Australian Market Index (S&P/ASX 200)', xlab = 'Time', ylab = 'Index')

axis(1, at = month\_vec, labels = c('Jan', 'Feb', 'Mar', 'Apr', 'May'))

plot(market\_chn~time, xaxt = 'n', type = 'o', col = 'blue', ylim = c(2600, 3200), main = 'Chinese Market Index (Shanghai Composite)', xlab = 'Time', ylab = 'Index')

axis(1, at = month\_vec, labels = c('Jan', 'Feb', 'Mar', 'Apr', 'May'))

plot(cases\_aus~time, xaxt = 'n', type = 'o', col = 'orange', ylim = c(0, 700), main = 'Australia Daily New Cases', xlab = 'Time', ylab = 'New cases')

axis(1, at = month\_vec, labels = c('Jan', 'Feb', 'Mar', 'Apr', 'May'))

plot(cases\_chn~time, xaxt = 'n', type = 'o', col = 'green', ylim = c(0, 16000), main = 'China Daily New Cases', xlab = 'Time', ylab = 'New cases')

axis(1, at = month\_vec, labels = c('Jan', 'Feb', 'Mar', 'Apr', 'May'))

## detrending via differencing

det\_market\_aus = diff(market\_aus)

# visualising the difference plots

par(mfrow = c(2, 1))

plot(det\_market\_aus~time[-1], xaxt = 'n', type = 'o', main = 'Australian Market Index (S&P/ASX 200) \ndetrended via first difference', xlab = 'Time', ylab = 'Difference')

axis(1, at = month\_vec-1, labels = c('Jan', 'Feb', 'Mar', 'Apr', 'May'))

lines(lowess(det\_market\_aus, f = 2/3), lwd = 2, lty = 2, col = 'red')

abline(0, 0, lty = 2)

legend('bottomright', legend = 'Lowess with f = 2/3', fill = 'red')

# acf plot

acf(det\_market\_aus, lag.max = 50, main = 'ACF Plot of Australian Market Index (S&P/ASX 200) \ndetrended via first difference')

## fitting regression model 1

# model fit

model1 = lm(market\_aus~time+market\_chn+cases\_aus+cases\_chn)

summary(model1)

# plots to check stationarity

# residual plot

par(mfrow = c(2, 1))

plot(residuals(model1)~time, xaxt = 'n', type = 'o', col = 'chocolate4', main = 'Residual Plot (Model 1)', xlab = 'Time', ylab = 'Residual')

axis(1, at = month\_vec, labels = c('Jan', 'Feb', 'Mar', 'Apr', 'May'))

lines(lowess(residuals(model1), f = 2/3), lwd = 2, lty = 2, col = 'red')

abline(0, 0, lty = 2)

legend('bottomright', legend = 'Lowess with f = 2/3', fill = 'red')

# acf plot of residuals

acf(residuals(model1), lag.max = 50, col = 'chocolate4', main = 'ACF Plot of Residuals (Model 1)')

# goodness-of-fit plot

par(mfrow = c(1, 1))

plot(market\_aus~time, xaxt = 'n', type = 'o', ylim = c(4000, 7500), main = 'Australian Market Index (S&P/ASX 200)', xlab = 'Time', ylab = 'Index')

axis(1, at = month\_vec, labels = c('Jan', 'Feb', 'Mar', 'Apr', 'May'))

lines(lowess(market\_aus, f = 1/3), lwd = 2, lty = 2, col = 'red')

lines(model1$coefficients[1]+model1$coefficients[2]\*time+model1$coefficients[3]\*market\_chn+model1$coefficients[4]\*cases\_aus+model1$coefficients[5]\*cases\_aus, lwd = 2, col = 'chocolate4')

legend("topright", legend = c('Lowess with f = 1/3', 'Model 1'), fill = c('red', 'chocolate4'))

## scatterplot matrix

panel <- function(x, y){

usr <- par('usr')

on.exit(par(usr))

par(usr = c(0, 1, 0, 1))

r <- round(cor(x, y), digits = 2)

corr <- bquote(rho == .(r))

text(0.5, 0.5, corr, col = 'violetred4')

}

pairs(cbind(Australian\_Market = market\_aus, Chinese\_Market = market\_chn, Australia\_New\_Cases = cases\_aus, China\_New\_Cases = cases\_chn), lower.panel = panel)

## fitting regression model 2

# square vectors

cases\_aus\_sq <- cases\_aus^2

cases\_chn\_sq <- cases\_chn^2

# model fit

model2 <- lm(market\_aus~time+market\_chn+cases\_aus+cases\_aus\_sq+cases\_chn+cases\_chn\_sq)

summary(model2)

# plots to check stationarity

par(mfrow = c(2, 1))

# residual plot

plot(residuals(model2)~time, xaxt = 'n', type = 'o', col = 'turquoise4', main = 'Residual Plot (Model 2)', xlab = 'Time', ylab = 'Residual')

axis(1, at = month\_vec, labels = c('Jan', 'Feb', 'Mar', 'Apr', 'May'))

lines(lowess(residuals(model2), f = 2/3), lwd = 2, lty = 2, col = 'red')

abline(0, 0, lty = 2)

legend('bottomright', legend = 'Lowess with f = 2/3', fill = 'red')

# acf plot of residuals

acf(residuals(model2), lag.max = 50, col = 'turquoise4', main = 'ACF Plot of Residuals (Model 2)')

# goodness-of-fit plot

par(mfrow = c(1, 1))

plot(market\_aus~time, xaxt = 'n', type = 'o', ylim = c(4000, 7500), main = 'Australian Market Index (S&P/ASX 200)', xlab = 'Time', ylab = 'Index')

axis(1, at = month\_vec, labels = c('Jan', 'Feb', 'Mar', 'Apr', 'May'))

lines(lowess(market\_aus, f = 1/3), lwd = 2, lty = 2, col = 'red')

lines(model2$coefficients[1]+model2$coefficients[2]\*time+model2$coefficients[3]\*market\_chn+model2$coefficients[4]\*cases\_aus+model2$coefficients[5]\*cases\_aus\_sq+model2$coefficients[6]\*cases\_chn+model2$coefficients[7]\*cases\_chn\_sq, lwd = 2, col = 'turquoise4')

legend("topright", legend = c('Lowess with f = 1/3', 'Model 2'), fill = c('red', 'turquoise4'))

## cross-correlation and lag plots

par(mfrow = c(4, 1))

acf(market\_aus, lag.max = 30, main = 'ACF Plot of Australian Market Index (S&P/ASX 200)')

ccf(market\_aus, market\_chn, lag.max = 30, main = 'CCF Plot of Australian Market Index (S&P/ASX 200) \nand Chinese Market Index (Shanghai Composite)')

ccf(market\_aus, cases\_aus, lag.max = 30, main = 'CCF Plot of Australian Market Index (S&P/ASX 200) \nand Australia Daily New Cases')

ccf(market\_aus, cases\_chn, lag.max = 30, main = 'CCF Plot of Australian Market Index (S&P/ASX 200) \nand China Daily New Cases')

lag2.plot(market\_aus, cases\_aus, max.lag = 11)

lag2.plot(cases\_chn, market\_aus, max.lag = 11)

## fitting regression model 3

# lead and lag functions

func\_lag <- function(x, lag){

stop <- length(x) - lag

c(rep(NA,lag), x[(1:stop)])

}

func\_lead <- function(x, lead){

c(x[-(1:lead)], rep(NA, lead))

}

# lead and lag vectors

cases\_aus\_lead2 <- func\_lead(cases\_aus, 2)

cases\_aus\_lead3 <- func\_lead(cases\_aus, 3)

cases\_aus\_lead4 <- func\_lead(cases\_aus, 4)

cases\_chn\_lag5 <- func\_lag(cases\_chn, 5)

cases\_chn\_lag6 <- func\_lag(cases\_chn, 6)

cases\_chn\_lag7 <- func\_lag(cases\_chn, 7)

# model fit

model3 <- lm(market\_aus~time+market\_chn+cases\_aus\_lead2+cases\_aus\_lead3+cases\_aus\_lead4+cases\_chn\_lag5+cases\_chn\_lag6+cases\_chn\_lag7)

summary(model3)

# plots to check stationarity

par(mfrow = c(2, 1))

# residual plot

plot(residuals(model3)~time[-c(1:7, 147:150)], xaxt = 'n', type = 'o', col = 'chartreuse4', main = 'Residual Plot (Model 3)', xlab = 'Time', ylab = 'Residual')

axis(1, at = month\_vec-8, labels = c('Jan', 'Feb', 'Mar', 'Apr', 'May'))

lines(lowess(residuals(model3), f = 2/3), lwd = 2, lty = 2, col = 'red')

abline(0, 0, lty = 2)

legend('bottomright', legend = 'Lowess with f = 2/3', fill = 'red')

# acf plot of residuals

acf(residuals(model3), lag.max = 50, col = 'chartreuse4', main = 'ACF Plot of Residuals (Model 3)')

# goodness-of-fit plot

par(mfrow = c(1, 1))

plot(market\_aus~time, xaxt = 'n', type = 'o', ylim = c(4000, 7500), main = 'Australian Market Index (S&P/ASX 200)', xlab = 'Time', ylab = 'Index')

axis(1, at = month\_vec, labels = c('Jan', 'Feb', 'Mar', 'Apr', 'May'))

lines(lowess(market\_aus, f = 1/3), lwd = 2, lty = 2, col = 'red')

lines(model3$coefficients[1]+model3$coefficients[2]\*time+model3$coefficients[3]\*market\_chn+model3$coefficients[4]\*cases\_aus\_lead2+model3$coefficients[5]\*cases\_aus\_lead3+model3$coefficients[6]\*cases\_aus\_lead4+model3$coefficients[7]\*cases\_chn\_lag5+model3$coefficients[8]\*cases\_chn\_lag6+model3$coefficients[9]\*cases\_chn\_lag7, lwd = 2, col = 'chartreuse4')

legend("topright", legend = c('Lowess with f = 1/3', 'Model 3'), fill = c('red', 'chartreuse4'))