Forecasting, Homework 3

Robby Jeffries

2/22/2022

Set working directory

```
rm(list = ls(all.names = TRUE))

setwd("~/MSEA2022/Spring 2022/ECON 5753, Forecasting")
```

Import packages

```
library(tidyverse)
library(lmtest)
```

Import data

```
data <- readr::read_csv("Data/saving_inc.csv")

# Name the second column 'y'
names(data)[2] <- "y"
# Name the third column 'x'
names(data)[3] <- "x"
```

(a) Run the regression

```
reg <- lm(formula = y~x+WAR, data = data)
summary(reg)
```

```
##
## Call:
## lm(formula = y ~ x + WAR, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.3505  -0.3178   0.8660   1.5431   7.3836
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.14131    2.50424  -1.254    0.227
## x            0.07632    0.01279   5.965 1.54e-05 ***
## WAR         20.16531    2.37522   8.490 1.61e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.531 on 17 degrees of freedom
## Multiple R-squared:  0.8455, Adjusted R-squared:  0.8273
## F-statistic: 46.51 on 2 and 17 DF,  p-value: 1.278e-07
```

(b) Test to determine whether knowledge of the war years makes a significant contribution to the prediciton of personal savings beyond that provided by personal income.

ANSWER: Yes, knowledge of the war years makes a significant contribution. The t-statistic for X_2 = 8.490, which indicates strong statistical significance.

(c) Calculate the standard error of estimation

```
std_error <- sqrt(deviance(reg)/df.residual(reg))
std_error
```

```
## [1] 4.53134
```

(d) Test the significance of the regression

ANSWER: For this, I will look at the F-statistic and its associated p-value. The F-statistic is 46.51 and its p-value is very close to zero (1.278e-07). Therefore, I conclude that the regression is statistically significant.

(e) Test for autocorrelation based on Durbin-Watson test and Breusch and Godfrey LM test

Durbin-Watson Test

```
lmtest::dwtest(y~x+WAR, data = data)
```

```
##
##  Durbin-Watson test
##
## data:  y ~ x + WAR
## DW = 2.0103, p-value = 0.3258
## alternative hypothesis: true autocorrelation is greater than 0
```

CONCLUSION: I fail to reject the null hypothesis that the residuals from the OLS regression are not autocorrelated.

Breusch and Godfrey LM Test

```
bgtest(y~x+WAR, order = 3, data = data) # if small p-value, reject null hypothesis
```

```
##
##  Breusch-Godfrey test for serial correlation of order up to 3
##
## data:  y ~ x + WAR
## LM test = 1.7846, df = 3, p-value = 0.6183
```

CONCLUSION: I fail to reject the null hypothesis that there is no serial correlation of any order up to 3.

(f) What is the model for the saving model for war years? Compare with the model for not war year, how do we explain the difference between those two models?

Model with war years

```
summary(reg)
```

```
##
## Call:
## lm(formula = y ~ x + WAR, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.3505  -0.3178   0.8660   1.5431   7.3836
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.14131    2.50424  -1.254    0.227
## x            0.07632    0.01279   5.965 1.54e-05 ***
## WAR         20.16531    2.37522   8.490 1.61e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.531 on 17 degrees of freedom
## Multiple R-squared:  0.8455, Adjusted R-squared:  0.8273
## F-statistic: 46.51 on 2 and 17 DF,  p-value: 1.278e-07
```

Model without war years

```
reg2 <- lm(formula = y~x, data = data)
summary(reg2)
```

```
##
## Call:
```

```
## lm(formula = y ~ x, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.9938 -6.2507 -3.3966 -0.1049 22.5058
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.97793    5.14865   0.967   0.3464
## x            0.05767    0.02804   2.057   0.0545 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.08 on 18 degrees of freedom
## Multiple R-squared:  0.1903, Adjusted R-squared:  0.1454
## F-statistic: 4.231 on 1 and 18 DF,  p-value: 0.05447
```

The model with war years shows that the dummy variable significantly affects the dependent variable, personal savings. That is, during wartime, people save more money than during peacetime. When the wartime variable is not included, the R-squared value decreases from 0.8455 to 0.1903, which indicates that personal income only explains 19% of the variation in personal savings. Together, the wartime and personal income variables explain 85% of the variation in personal savings. Therefore, I concluded that the dummy variable was a great addition, and it strengthened the overall model.