Problem Set Chapter 6

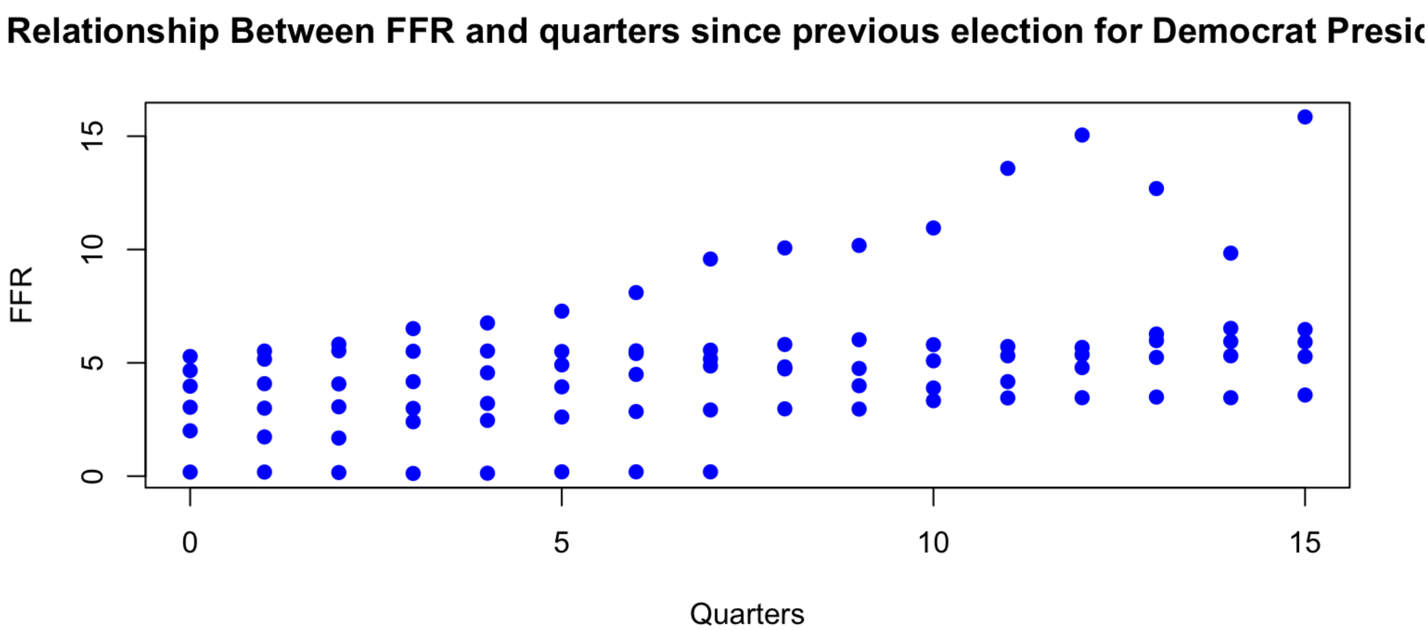
Tianwei Liu

11/1/2019

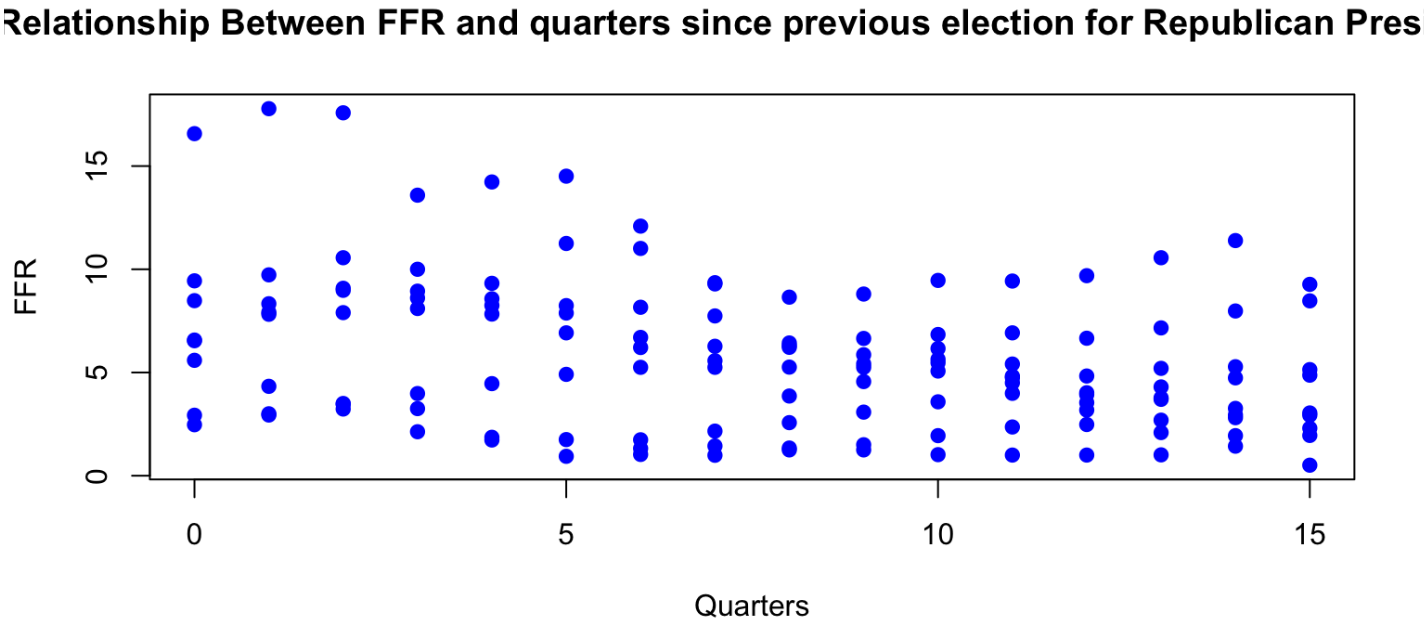
2.These questions are based on “The Fed May Be Politically Independent but It Is not Politically Indifferent,” a paper by William Clark and Vincent Arel-Bundock (2013). The paper explores the relationship between elections and the federal funds rate (FFR). Often a benchmark for financial markets, the FFR is the average interest rate at which federal funds trade in a day. The rate is set by the U.S. central bank, the Federal Reserve, a.k.a. the Fed. Table 6.11 describes the variables from fed_2012.dta that we use in this problem.

(a)Create two scatterplots, one for years in which a Democrat was president and one for years in which a Republican was president, showing the relationship between the FFR and the quarters since the previous election. Comment on the differences in the relationships. The variable Quarters is coded 0 to 15, representing each quarter from one election to the next. For each presidential term, the value of Quarters is 0 in the first quarter containing the election and 15 in the quarter before the next election.

```
dta_dem = dta[dta$democrat == 1,]
plot(dta_dem$election, dta_dem$FEDFUNDS, pch = 19, col = "blue", main = "Relationship Be
tween FFR and quarters since previous election for Democrat President", xlab = "Quarter
s", ylab = "FFR")
```




```
dta_rep = dta[dta$democrat == 0,]
plot(dta_rep$election, dta_rep$FEDFUNDS, pch = 19, col = "blue", main = "Relationship Be
tween FFR and quarters since previous election for Republican President", xlab = "Quarte
rs", ylab = "FFR")
```



There is an upward trend for the FFR when the incumbent is democrat as the Quarters increase, however the trend seems downward sloping for the FFR when the incumbent is republican.

(b)Create an interaction variable between Quarters and Democrat to test whether closeness to elections has the same effect on Democrats and Republicans. Run a model with the FFR as the dependent variable, allowing the effect of the Quarters variable to vary by party of the president.

(i)What change in FFR is associated with a one-unit increase in the Quarters variable when the president is a Republican?

```
dta$duminterac = dta$election * dta$democrat
reg1 <- lm(FEDFUNDS ~ election + democrat + duminterac, data = dta)
summary (reg1)
```

```
##
## Call:
## lm(formula = FEDFUNDS ~ election + democrat + duminterac, data = dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.5060 -1.9970 -0.3652  1.6608 10.3394
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.77030    0.52736  14.734  < 2e-16 ***
## election    -0.26487    0.05879  -4.505 1.07e-05 ***
## democrat    -4.90324    0.81534  -6.014 7.39e-09 ***
## duminterac   0.55819    0.09393   5.943 1.08e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.158 on 222 degrees of freedom
##   (6 observations deleted due to missingness)
## Multiple R-squared:  0.151,  Adjusted R-squared:  0.1395
## F-statistic: 13.16 on 3 and 222 DF,  p-value: 6.083e-08
```

We need to note that the model is Y = B0 + B1* Election + B2* Democrat + B3* Election* Democrat (duminterac) + Error_term Based on this model, when the president is a Republican, our dummy variable Democrat = 0. Therefore, one-unit increase in the Quarters (election) is associated with a 0.265 decrease in the FFR.

(ii)What change in FFR is associated with a one-unit increase in the Quarters variable when the president is a Democrat?

When the president is a democrat, our dummy variable Democrat = 1. Therefore, one-unit increase in the quarters (election) is associated with a B1 + B3 = (-0.265 + 0.558) = 0.293 increase in the FFR.

(c)Is the effect of Quarters statistically significant under Republicans? (Easy.) Is the effect of Quarters statistically significant under Democrats? (Not so easy.) How can the answer be determined? Run any additional tests if necessary.

Under Republicans, dummy variable Democrat = 0, so the effect is simply B1. Given the regression result, the p-value associated with B1 is 1.1 * 10^-5 which is smaller than the 0.001 level of significance at alpha = 0.999. Therefore the coefficient on Quarters (Election) is statistically significant.

In order to know if the effect of Quarters (Election) is statistically significant under Democrats, we would run a model of FFR on election controlling for democrats (using the previously created dta_demo dataset).

```
reg1_c <- reg1 <- lm(FEDFUNDS ~ election, data = dta_dem)
summary (reg1_c)
```

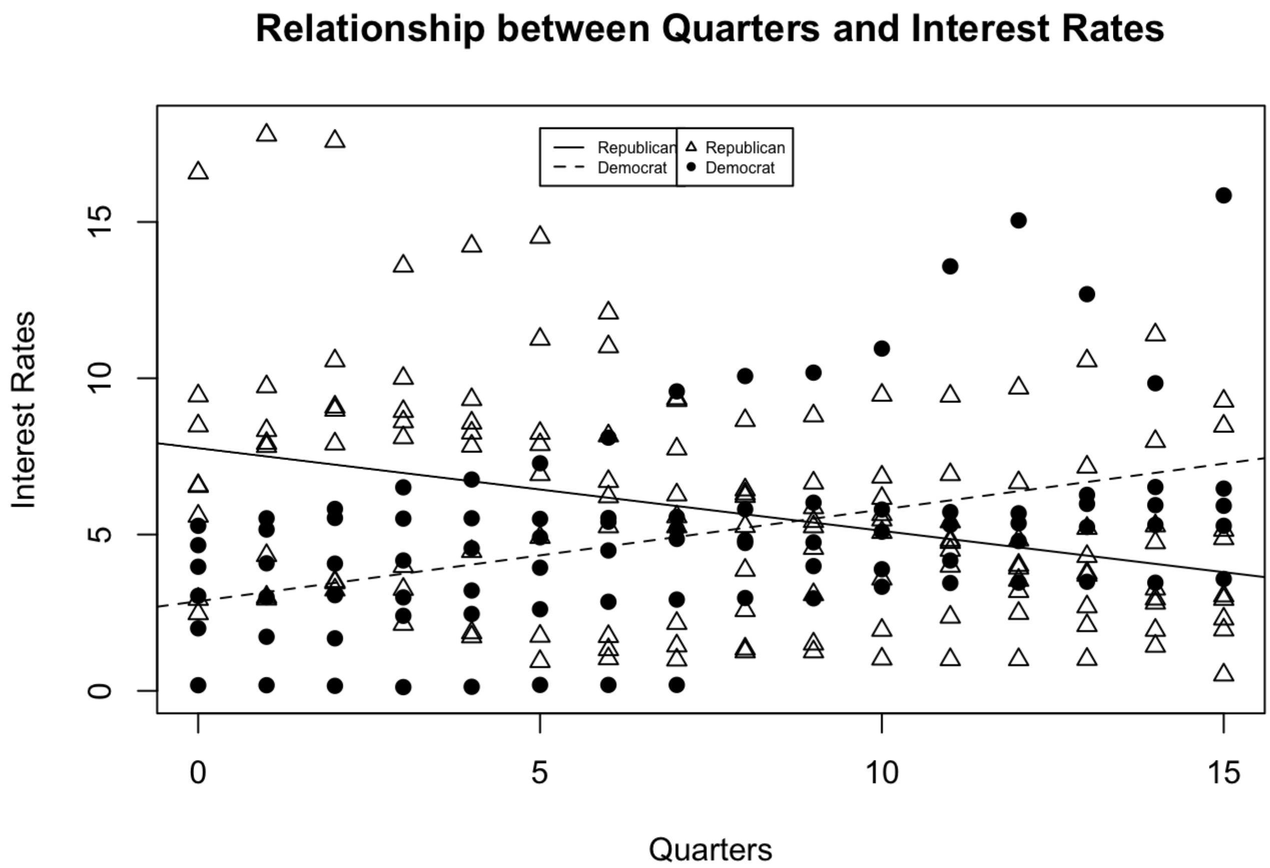
```
##
## Call:
## lm(formula = FEDFUNDS ~ election, data = dta_dem)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -4.730 -1.736 -0.432  1.119  8.663
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.8671     0.5432   5.278 9.66e-07 ***
## election      0.2933     0.0640   4.583 1.54e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.759 on 86 degrees of freedom
## Multiple R-squared:  0.1963, Adjusted R-squared:  0.187
## F-statistic: 21.01 on 1 and 86 DF,  p-value: 1.539e-05
```

The p-value associated with the coefficient of election (Quarters) on FFR is 1.43 * 10^-5 which is smaller than the 0.001 level of significance therefore it is very statistically significant.

(d)Graph two fitted lines for the relationship between Quarters and interest rates, one for Republicans and one for Democrats. (In Stata, use the twoway and lfit commands with appropriate if statements; label by hand. In R, use the abline command.) Briefly describe the relationship.

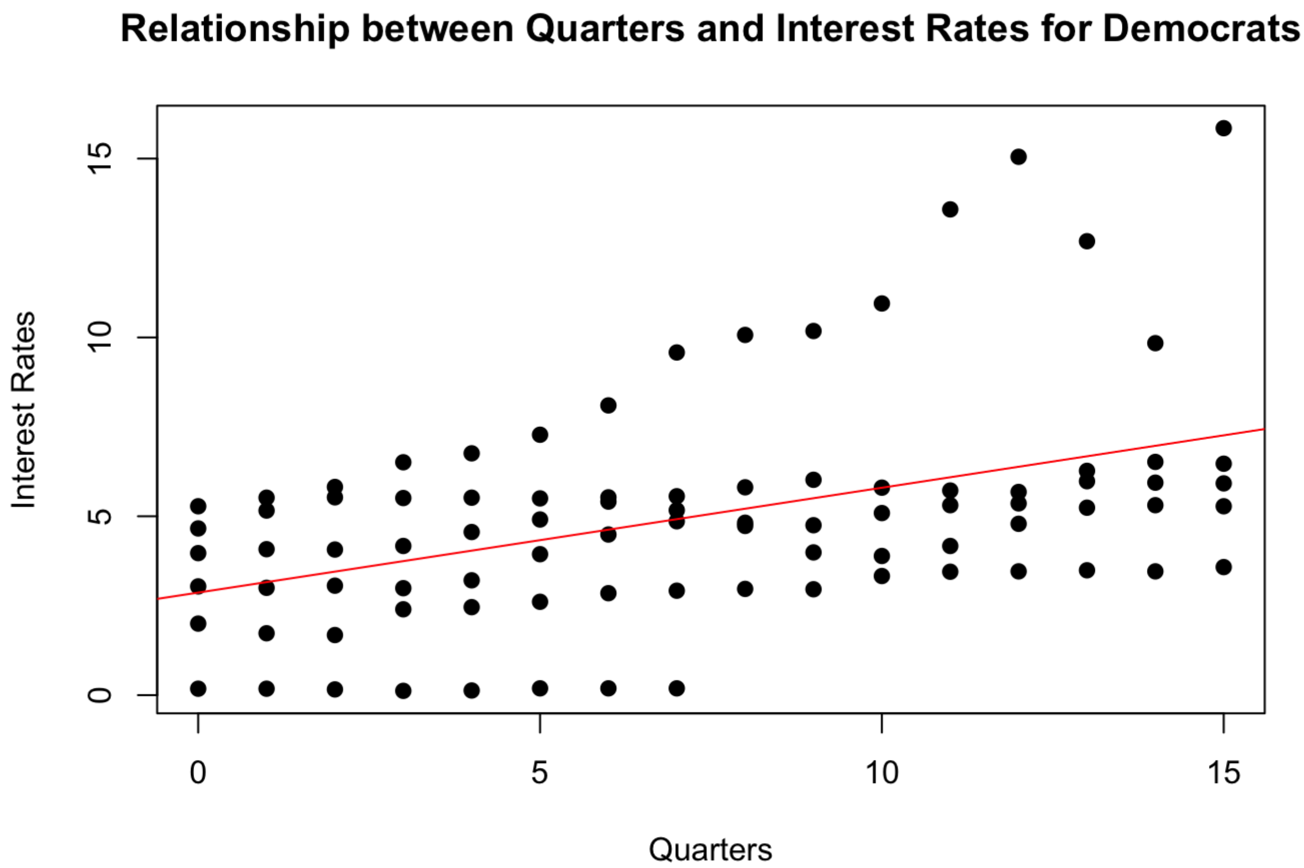
Method 1: Using Dummy Interaction Term

```
reg_test <- lm (FEDFUNDS ~ election + democrat + duminterac, data = dta)
B0 = summary (reg_test)$coefficients[1]
B1 = summary (reg_test)$coefficients[2]
B2 = summary (reg_test)$coefficients[3]
B3 = summary (reg_test)$coefficients[4]
plot (dta_dem$election, dta_dem$FEDFUNDS, pch = 19, main = "Relationship between Quarter
s and Interest Rates", xlab = "Quarters", ylab = "Interest Rates", ylim = c(0,18))
points (dta_rep$election, dta_rep$FEDFUNDS, pch = 2)
abline (B0, B1, lty = 1)
abline (B0+B2, B1+B3, lty = 2)
legend(5,18, legend=c("Republican", "Democrat"),
       lty = 1:2, cex=0.5)
legend(7,18, legend=c("Republican", "Democrat"),
       pch = c(2,19), cex = 0.5)
```

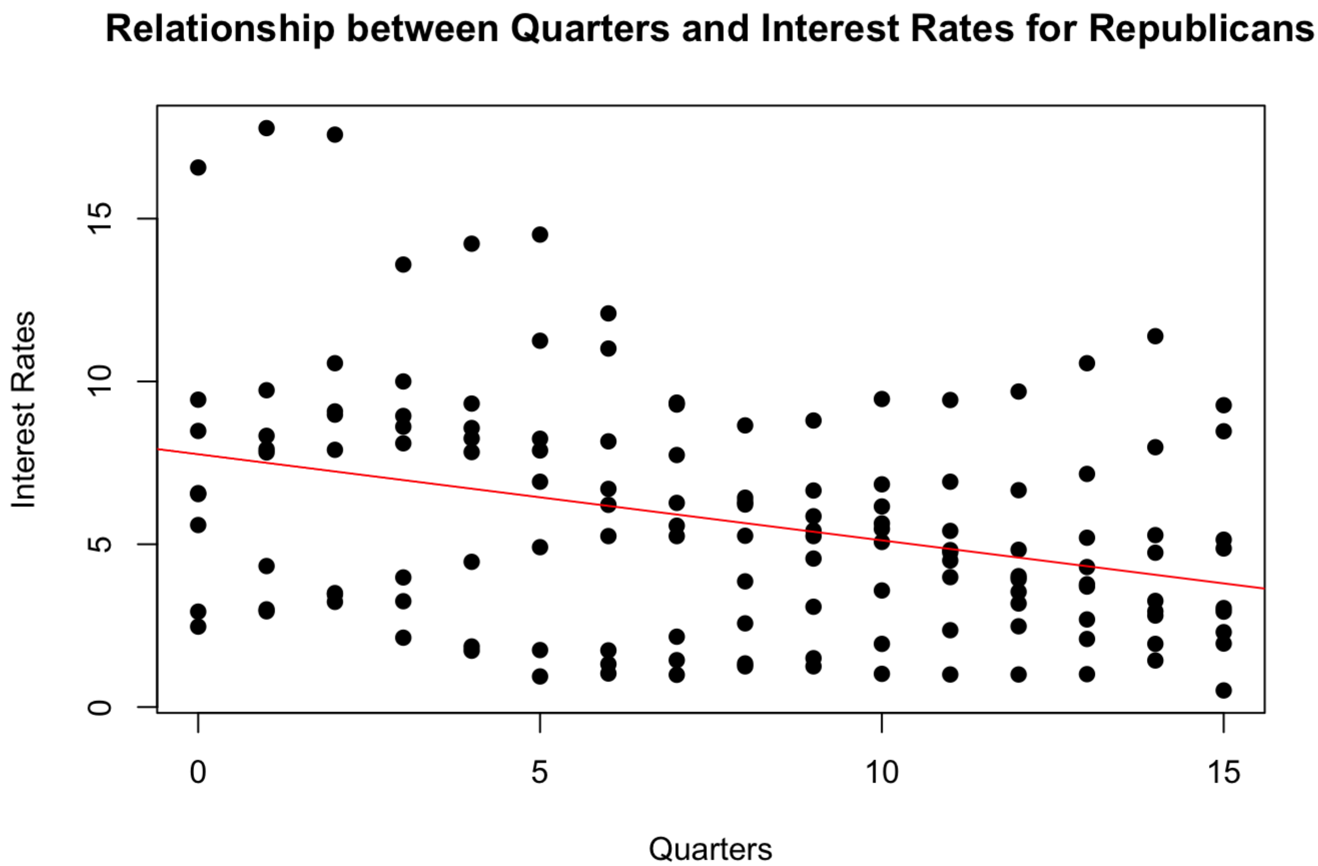


Method 2: Using Data Subset

```
reg2 <- lm (FEDFUNDS ~ election, data = dta_dem)
a1 = summary(reg2)$coefficients[1]
b1 = summary(reg2)$coefficients[2]
plot (dta_dem$election, dta_dem$FEDFUNDS, pch = 19, main = "Relationship between Quarter
s and Interest Rates for Democrats", xlab = "Quarters", ylab = "Interest Rates")
abline (a1, b1, col = "red")
```




```
reg3 <- lm (FEDFUNDS ~ election, data = dta_rep)
a2 = summary(reg3)$coefficients[1]
b2 = summary(reg3)$coefficients[2]
plot (dta_rep$election, dta_rep$FEDFUNDS, pch = 19, main = "Relationship between Quarter
s and Interest Rates for Republicans", xlab = "Quarters", ylab = "Interest Rates")
abline (a2, b2, col = "red")
```



As we see in the graphs above, when the president is Democrat, inflation rates tend to go up as the their terms continue. In contrast, when the president is Republican, inflation rates go down sligtly as their terms continue.

(e)Rerun the model from part (b) controlling for both the interest rate in the previous quarter (lag_FEDFUND) and inflation b and discuss the results, focusing on (i) effect of Quarters for Republicans, (ii) the differential effect of Quarters for Democrats, (iii) impact of lagged FFR, and (iv) inflation. Simply report the statistical significance of the coefficient estimates; don't go through the entire analysis from part (c).

```
reg4 <- lm(FEDFUNDS ~ election + democrat + duminterac + lag_FEDFUNDS + inflation, data
= dta)
summary (reg4)
```

```
##
## Call:
## lm(formula = FEDFUNDS ~ election + democrat + duminterac + lag_FEDFUNDS +
##     inflation, data = dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0728 -0.3618  0.0197  0.4032  5.2985
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.24630    0.20557   1.198   0.2322
## election     -0.02300    0.01679  -1.370   0.1720
## democrat     -0.11165    0.24253  -0.460   0.6457
## duminterac    0.04699    0.02757   1.705   0.0897 .
## lag_FEDFUNDS  0.88927    0.02369  37.541  < 2e-16 ***
## inflation     0.11648    0.02468   4.719 4.23e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8649 on 219 degrees of freedom
##   (7 observations deleted due to missingness)
## Multiple R-squared:  0.9367, Adjusted R-squared:  0.9353
## F-statistic: 648.2 on 5 and 219 DF,  p-value: < 2.2e-16
```

The effect of quarters for republicants (democrat = 0) has a p-value 0.23 which is greater than the conventional 0.05 level of significance therefore is not statistically significant. The differential effect of Quarters for Democrats (democrat = 1) is the coefficient for the dummy interaction term, which has a p-value of 0.09. The p-value associated with the coefficient is greater than conventional 0.05 level of significance. Therefore, the differential effect of Quarters for Democrats is also not significant. The effect of logged FFR has a large t statistic and a very small p-value, which is much smaller than 0.001 level of significance therefore it is highly significant. The effect of inflation also has a small p-value (4.23e-06) which is much smaller than 0.001 level of significance and therefore it is also highly statistically significant.

4.In this problem we continue analyzing the speeding ticket data first introduced in Chapter 5 (page 175). The variables we use are in Table 6.12.

(a)Implement a simple difference of means test that uses OLS to assess whether the fines for men and women are different. Do we have any reason to expect endogeneity? Explain.

```
load("Ch6_Exercise4_Speeding_tickets.RData")
```

```
reg5 <- lm(Amount ~ Female, data = dta)
summary (reg5)
```

```
##
## Call:
## lm(formula = Amount ~ Female, data = dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -121.67  -49.67   -6.73   30.33  600.33
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 124.6654     0.3857  323.23   <2e-16 ***
## Female       -7.9394     0.6698  -11.85   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 56.12 on 31672 degrees of freedom
##   (36683 observations deleted due to missingness)
## Multiple R-squared:  0.004416,   Adjusted R-squared:  0.004385
## F-statistic: 140.5 on 1 and 31672 DF,  p-value: < 2.2e-16
```

Average fines for men (Female = 0) is 124.67 while average finds for women (Female = 1) is 124.67 - 7.94 = 116.73. The difference is the coefficient on Female dummy variable. As the coefficient on this variable is highly statistically significant (p-value much smaller than 0.001 level of significance), we reject the null that the difference is zero and we can say that the fines for men and women are different.

Yes, I suspect endogeneity in this model. Let's consider MPHover. MPHover is correlated with gender (probably on average men drive faster than women) and MPHover has a big say in the fine amount.

(b)Implement a difference of means test for men and women that controls for age and miles per hour. Do we have any reason to expect endogeneity? Explain.

```
reg6 <- lm(Amount ~ Female + Age + MPHover, data = dta)
summary (reg6)
```

```
##
## Call:
## lm(formula = Amount ~ Female + Age + MPHover, data = dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -309.392  -20.369    6.841   25.496  225.611
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.89857    0.97164   5.042 4.64e-07 ***
## Female      -3.55524    0.47371  -7.505 6.30e-14 ***
## Age          0.02578    0.01758   1.466    0.143
## MPHover      6.87692    0.03871 177.670  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 39.63 on 31670 degrees of freedom
##   (36683 observations deleted due to missingness)
## Multiple R-squared:  0.5035, Adjusted R-squared:  0.5034
## F-statistic: 1.071e+04 on 3 and 31670 DF,  p-value: < 2.2e-16
```

Holding age and MPHover constant, the difference in fine amount between women and men is 3.555. As the coefficient on Female dummy variable has a p-value of 6.3 * 10^-14, which is much smaller than the 0.001 level of significance, we reject the null that the difference is zero. Therefore, there is difference between find amounts men and women get holding all else constant, and the difference of means is 3.555.

Yes, racism may play a part in speeding tickets! Black people or ethnic minority are more likely to be fined a higher amount even when their MPHover is relatively small.

(c)Building from the model just described, also assess whether fines are higher for African-Americans and Hispanics compared to everyone else (non-Hispanic whites, Asians and others). Explain what the coefficients on these variables mean.

```
reg7 <- lm(Amount ~ Female + Age + MPHover + Black + Hispanic, data = dta)
summary (reg7)
```

```
##
## Call:
## lm(formula = Amount ~ Female + Age + MPHover + Black + Hispanic,
##     data = dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -309.477  -20.231    6.834   25.457  225.559
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.84512    0.97546   4.967 6.83e-07 ***
## Female      -3.54327    0.47457  -7.466 8.47e-14 ***
## Age          0.02682    0.01761   1.523   0.1278
## MPHover      6.87850    0.03874 177.547  < 2e-16 ***
## Black       -2.02930    1.01789  -1.994   0.0462 *
## Hispanic     1.93201    1.05784   1.826   0.0678 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 39.63 on 31668 degrees of freedom
##   (36683 observations deleted due to missingness)
## Multiple R-squared:  0.5036, Adjusted R-squared:  0.5035
## F-statistic:  6426 on 5 and 31668 DF,  p-value: < 2.2e-16
```

For black: the coefficient on black is -2.029. It means that holding all else equal, blacks are, on average, fined 2.029 less than non-blacks. This coefficient is statistically significant at alpha = 0.95 (p-value smaller than 0.05). For Hispanic: the coefficient on hispanic is 1.93. It means that holding all else equal, hispanic drivers are on average fined 1.93 more than non-hispanic drivers. This coefficient is not statistically significant at conventional alpha = 0.95 because the associated p-value is greater than 0.05.

(d)Look at standard errors on coefficients for the Female, Black, and Hispanic variables. Why they are different?

Standard errors are used to measured the average distance between actual observations and the estimates (fitted values), The standard errors on coefficients for the Female, Black, and Hispanic variables are different because the spread caused by different factors are different.

(e)Within a single OLS model, assess whether miles over the speed limit has a differential effect on the fines for women, African-Americans, and Hispanics.

```
dta$mphfemale = dta$MPHover*dta$Female
dta$mphblack = dta$MPHover*dta$Black
dta$mphhis = dta$MPHover*dta$Hispanic
reg8 <- lm(Amount ~ MPHover + Female + Black + Hispanic + mphfemale + mphblack + mphhis,
data = dta)
summary (reg8)
```

```
##
## Call:
## lm(formula = Amount ~ MPHover + Female + Black + Hispanic + mphfemale +
##     mphblack + mphhis, data = dta)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -314.213  -19.565    7.415   25.358  220.864
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.36455    0.87847   3.830 0.000128 ***
## MPHover       7.01543    0.04832 145.194  < 2e-16 ***
## Female        8.96813    1.51645   5.914 3.38e-09 ***
## Black        -7.39776    3.00078  -2.465 0.013696 *
## Hispanic    -12.60802    3.07825  -4.096 4.22e-05 ***
## mphfemale    -0.74441    0.08556  -8.700  < 2e-16 ***
## mphblack      0.29812    0.15627   1.908 0.056442 .
## mphhis        0.81435    0.16282   5.001 5.72e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 39.56 on 31666 degrees of freedom
##   (36683 observations deleted due to missingness)
## Multiple R-squared:  0.5053, Adjusted R-squared:  0.5052
## F-statistic:  4621 on 7 and 31666 DF,  p-value: < 2.2e-16
```

The interaction terms are measures of differential effects on the fines for women, blacks, and hispanics. From the regression result, we can see that the different effect of MPHover for women and for hispanics are statistically significant (both with p-values smaller than 0.001 level of significant and therefore highly significant). We reject the null that there is no differential effect. The differential effect of MPHover for blacks has a p-value above conventional level 0.05, so it is not statistically significant. As a result, we fail to reject null that there is no differential effect for MPHover for blacks.

Question 3

Relationship between civility and evaluation of President TrumpQUESTION: We want to understand the relationship between people's views of civility in politics and their views of President Trump. In particular, do views about civility seem to affect how Democrats and Republicans view the president differently?You will need to develop a statistical model that addresses this question using the data in Battleground-65-Final-Dataset.xlsx. The questionnaire is available in Civility-Questionnaire-.pdf. To keep everyone on the same page, use the variable *CIVIL2* to measure civility, *DTID* to measure approval of Donald Trump. Convert the *PARTYID* variable into a dummy variable that equals 1 for Republicans and 0 for non-Republicans. You should control for other variables; while this could get quite involved, for the purpose of this question, you need not control for more than two other variables.

```
## Read in dataset
df <- read.xlsx("Battleground-65-Final-Dataset.xlsx",1)
```

```
## Create a dummy variable for Republican
df$Republican = as.numeric(df$PARTYID == 1 | df$PARTYID == 2 | df$PARTYID == 3)
## Create a dummy interaction variable for Republican * Civility
df$repciv = df$Republican * df$CIVIL2
```

```
reg9 <- lm (DTID ~ CIVIL2 + Republican + repciv, data = df)
summary (reg9)
```

```
##
## Call:
## lm(formula = DTID ~ CIVIL2 + Republican + repciv, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.8078 -0.5192  0.1922  0.1922  3.5648
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.03826    0.06622  60.982  < 2e-16 ***
## CIVIL2      -0.23048    0.03881  -5.939 3.95e-09 ***
## Republican  -2.49107    0.09546 -26.095  < 2e-16 ***
## repciv       0.20248    0.05099   3.971 7.66e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8616 on 1007 degrees of freedom
## Multiple R-squared:  0.626,  Adjusted R-squared:  0.6249
## F-statistic: 561.8 on 3 and 1007 DF,  p-value: < 2.2e-16
```

Model : Trump_Approval (DTID) = B0 + B1* Civility (CIVIL2) + B2* Republican + B3* Republican * Civility + Error_term

Intepretation:

When a voter is not republican, the average views of Donald Trump is around 4.03. Which means that people somewhat disagree with Trump. For non-republicans, an one-unit increase in the civility measurement (that people value civility less) tend to decrease measurement of views of Donald Trump by 0.23 unit (that people disagree less/agree more with Trump). This coefficient is highly statistically significant as the p-value associated with the coefficient is much smaller than 0.001 level of significance so that we are 99.9% confident that we reject the null that civility has no effect on views of Trump.

When a voter is republican, the average views of Donald Trump is around 1.547 (B0 + B2), meaning that most republicans somewhat agree with Donald Trump. For republicans, an one-unit increase in the civility (that people value civility less) tend to dcrease the measurement of views of Donald Trump by 0.028 unit, meaning people disagree slightly less/agree slightly more with Trump. The differential effect of civility for republicans is statistically significant because the p-value associated is less than 0.001 level of significance. So we reject the null that there is no differential effect and accept the alternative hypothesis.

In general, republicans agree more with Trump then democrats - this is obvious because partisanship plays a vital role in American politics. Also, civility matters but civility matters more for democrats than for republicans. More civil democrats have more mild political attitudes even when facing such an unusual politician.