Problem Set 3

Applied Stats/Quant Methods 1

Due: November 12, 2021

Instructions

- Please show your work! You may lose points by simply writing in the answer. If the problem requires you to execute commands in R, please include the code you used to get your answers. Please also include the .R file that contains your code. If you are not sure if work needs to be shown for a particular problem, please ask.

- Your homework should be submitted electronically on GitHub in .pdf form.

- This problem set is due before class on Friday November 12, 2021. No late assignments will be accepted.

- Total available points for this homework is 80.

In this problem set, you will run several regressions and create an add variable plot (see the lecture slides) in R using the incumbents subset.csv dataset. Include all of your code.

Question 1

We are interested in knowing how the difference in campaign spending between incumbent and challenger affects the incumbent's vote share.

1. Run a regression where the outcome variable is voteshare and the explanatory variable is difflog.

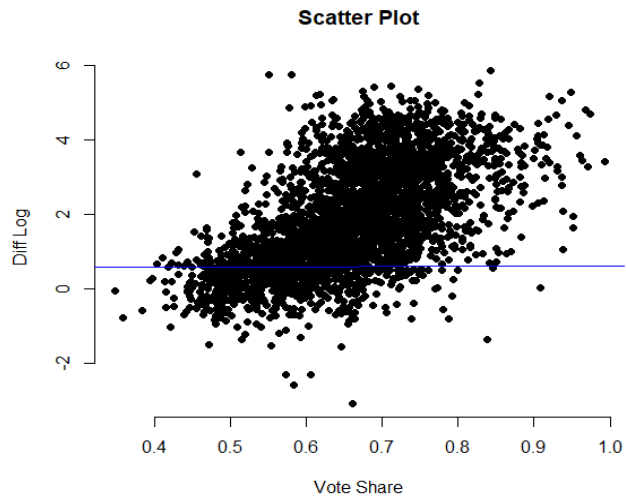   > regression1 <- lm(voteshare~difflog, data)

   > regression1

   Call:

     lm(formula = voteshare ~ difflog, data = data)

   Coefficients:

     (Intercept)   difflog

   0.57903   0.04167

2. Make a scatterplot of the two variables and add the regression line.

Scatter Plot

3. Save the residuals of the model in a separate object.

residual1 <- regression1$residuals

4. Write the prediction equation.

Voteshare = 0.57903 + 0.04167 * difflog

Question 2

We are interested in knowing how the difference between incumbent and challenger's spending and the vote share of the presidential candidate of the incumbent's party are related.

1. Run a regression where the outcome variable is presvote and the explanatory variable is difflog.

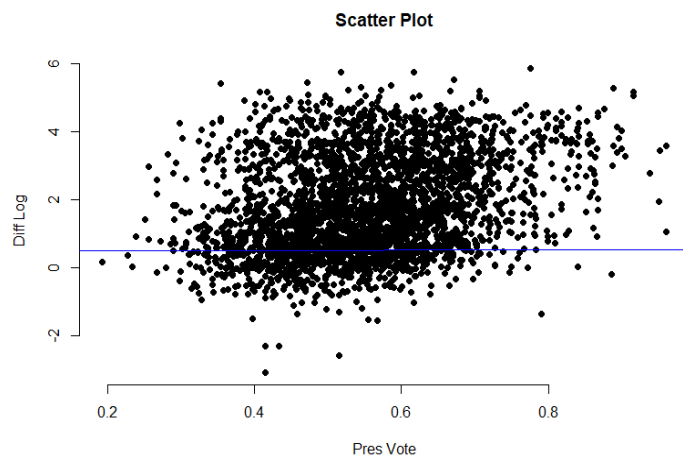> regression2 <- lm(presvote~difflog, data)

> regression2

Call:

lm(formula = presvote ~ difflog, data = data)

Coefficients:

 (Intercept)   difflog

0.50758   0.02384

2. Make a scatterplot of the two variables and add the regression line.

3. Save the residuals of the model in a separate object.

   residual2 <- regression2$residuals

4. Write the prediction equation.

   Presvote = 0.50758 + 0.02384 * difflog

Question 3

We are interested in knowing how the vote share of the presidential candidate of the incumbent's party is associated with the incumbent's electoral success.

1. Run a regression where the outcome variable is voteshare and the explanatory variable is presvote.

   > regression3 <- lm(voteshare~presvote, data)

   > regression3

   Call:
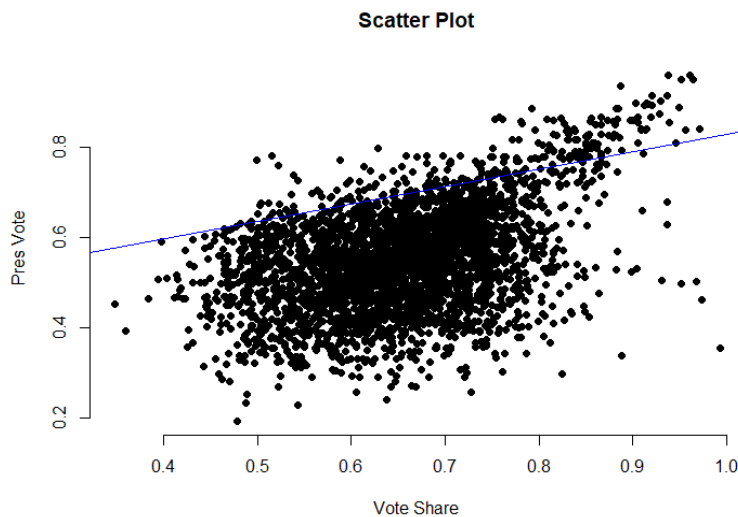
     lm(formula = voteshare ~ presvote, data = data)

   Coefficients:

(Intercept)   presvote

0.4413   0.3880

2. Make a scatterplot of the two variables and add the regression line.

Scatter Plot

3. Write the prediction equation.

Voteshare = 0. 4413 + 0.3880 * presvote

Question 4

The residuals from part (a) tell us how much of the variation in voteshare is *not* explained by the difference in spending between incumbent and challenger. The residuals in part (b) tell us how

much of the variation in presvote is *not* explained by the difference in spending between incumbent and challenger in the district.

1. Run a regression where the outcome variable is the residuals from Question 1 and the explanatory variable is the residuals from Question 2.

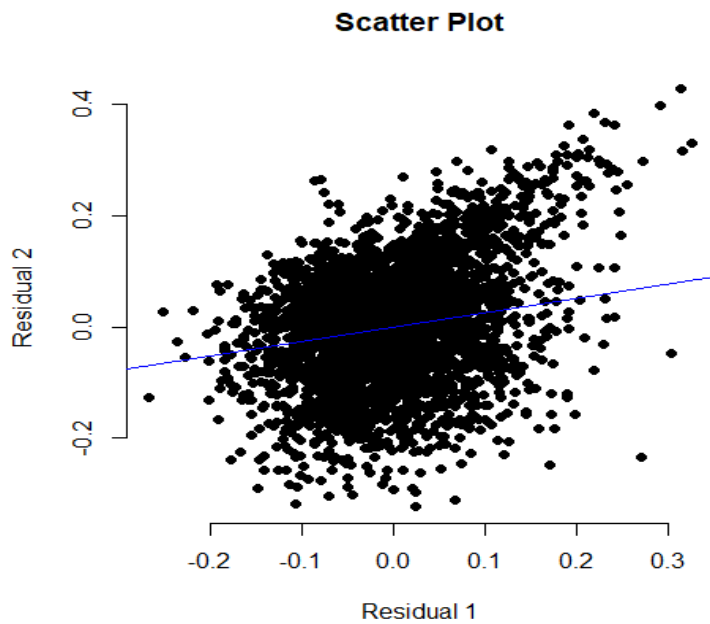> regression4 <- lm(residual1~residual2, data)

> regression4

Call:

lm(formula = residual1 ~ residual2, data = data)

Coefficients:

(Intercept)   residual2

-4.860e-18   2.569e-01

2. Make a scatterplot of the two residuals and add the regression line.

Scatter Plot

3. Write the prediction equation.

Residual1 = -4.860e-18 + 2.569e-01 * residual2

Question 5

What if the incumbent's vote share is affected by both the president's popularity and the difference in spending between incumbent and challenger?

1. Run a regression where the outcome variable is the incumbent's voteshare and the explanatory variables are difflog and presvote.

```
> regression5 <- lm(voteshare~presvote+difflog, data)
```

> regression5

Call:

lm(formula = voteshare ~ presvote + difflog, data = data)

Coefficients:

(Intercept)   presvote   difflog

0.44864   0.25688   0.03554

2. Write the prediction equation.

Voteshare = 0.44864 + 0.25688 * presvote + 0.03554 * difflog

3. What is it in this output that is identical to the output in Question 4? Why do you think this is the case?

The difference in spending between incumbent and challenger in the district does not influence the voteshare. The amount of money spent has no impact on voteshare unlike the president's popularity. **CODE in the file below:**

Project 3.R.txt