Problem Set 3 - Applied Stats/Quant Methods 1

Imelda Finn - 22334657

Due: November 20, 2021

Data

```
dat = read_csv("../../../datasets/incumbents_subset.csv")
```

Question 1

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

1. A linear regression was run where the outcome variable was `voteshare` and the explanatory variable was `difflog`. The function call to generate the model is:

   ```
   mod_vote_spend <- lm(voteshare ~ difflog, data = dat)
   ```

   The results are in Table 1.

2. The scatterplot of the two variables, with the regression line is in Figure 1.

3. The residuals from the model were saved as a separate object.

   ```
   resid_vote_spend <- mod_vote_spend$residuals
   ```

4. Prediction Equation

   $$voteshare = 0.579031 + (0.041666) * difflog$$

   $voteshare$ is 0.579031 when $difflog$ is 0 and it increases by 0.041666 for each unit increase in $difflog$, ie the share of the vote received by the incumbent increased when they outspent the challenger.

Table 1: Vote share as a function of Differental Spending

	Dependent variable:
	voteshare
difflog	0.041666***
	(0.000968)
Constant	0.579031***
	(0.002251)
Observations	3,193
R^2	0.367341
Adjusted R^2	0.367143
Residual Std. Error	0.078673 (df = 3191)
F Statistic	1,852.791000*** (df = 1; 3191)
Note:	*p<0.1; **p<0.05; ***p<0.01

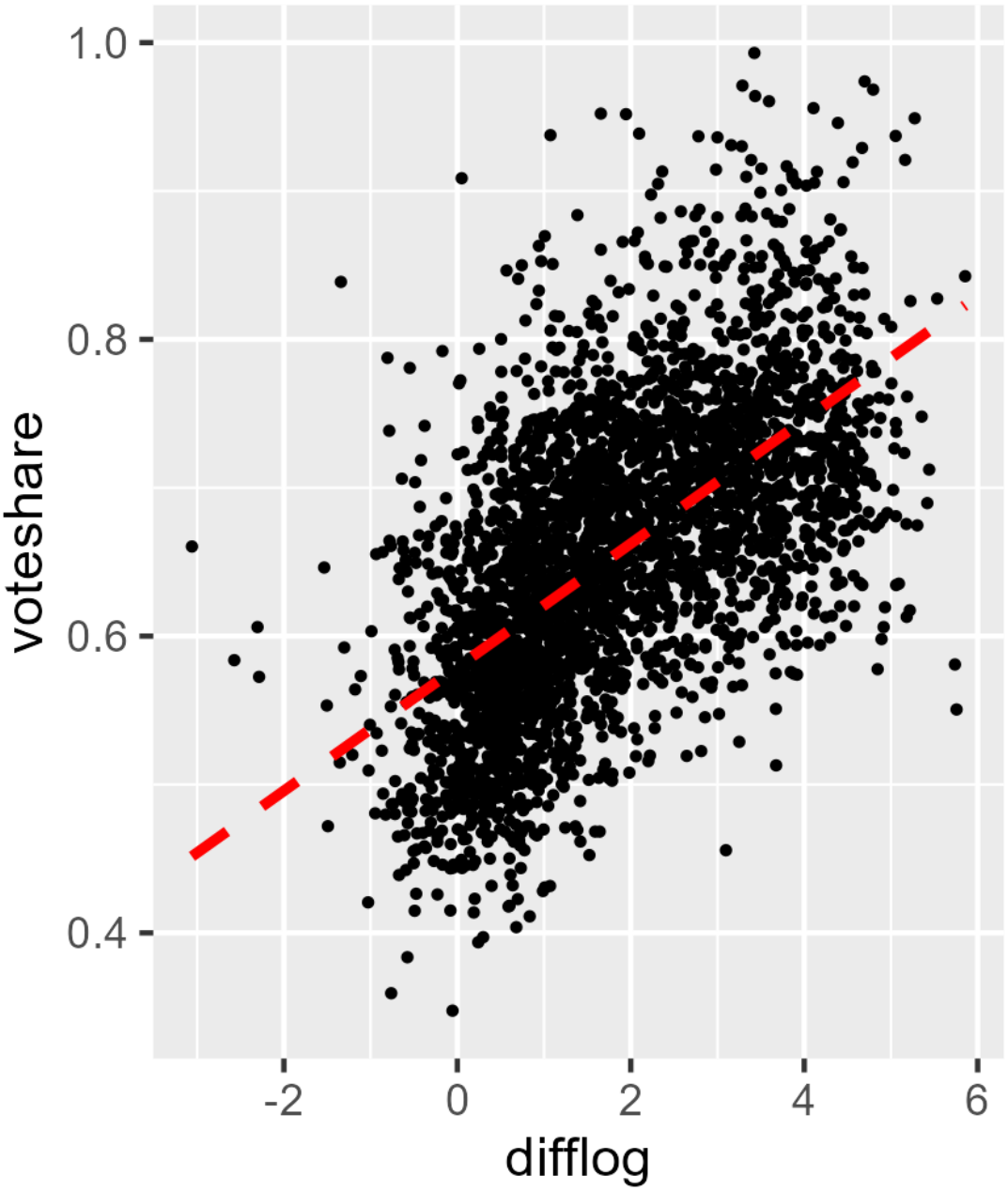

Figure 1: Incumbent's vote share as a function of differental spending

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. A linear regression was run where the outcome variable was `presvote` and the explanatory variable was `difflog`.

   The function call to generate the model is:

   ```
   mod_pres_spend <- lm(presvote ~ difflog, data = dat)
   ```

   The results are in table 2

2. Scatterplot of the two variables, with the regression line in Figure 2.

3. The residuals from the model were saved as a separate object.

   ```
   resid_pres_spend <- mod_pres_spend$residuals
   ```

4. Prediction Equation

   $$presvote = 0.507583 + (0.023837) * difflog$$

   `presvote` is 0.507583 when `difflog` is 0, and it increases by 0.023837 for each unit increase in `difflog` (ie the share of the vote received by the incumbent's presidential candidate increased when the incumbent outspent the challenger.)

Table 2: Presidential vote share as a function of Differental Spending

	Dependent variable:
	presvote
difflog	0.023837***
	(0.001359)
Constant	0.507583***
	(0.003161)
Observations	3,193
R^2	0.087951
Adjusted R^2	0.087665
Residual Std. Error	0.110442 (df = 3191)
F Statistic	307.715400*** (df = 1; 3191)
Note:	*p<0.1; **p<0.05; ***p<0.01

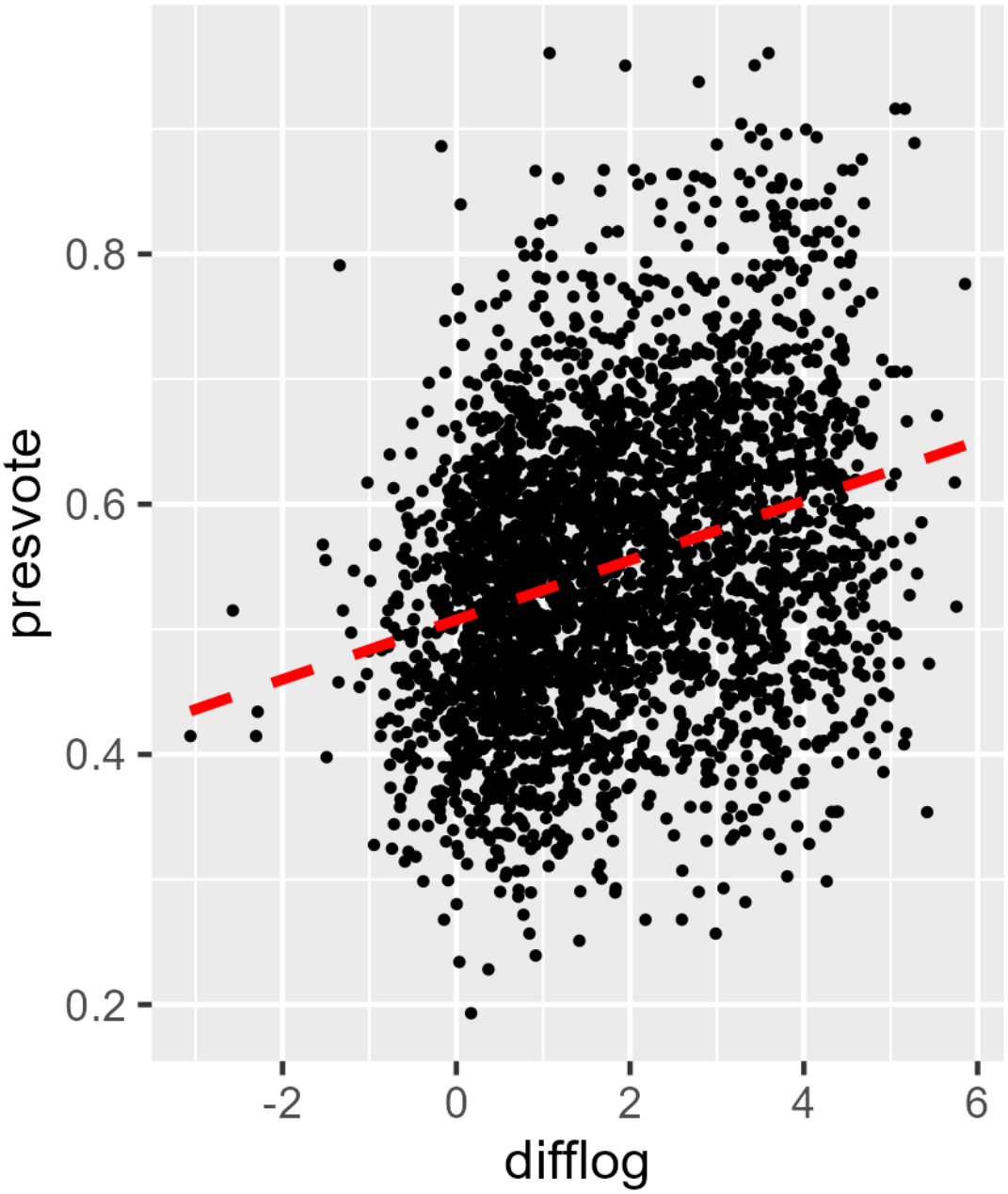

Figure 2: Presidential vote share as a function of incumbent's differental spending

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. A linear regression was run where the outcome variable was `voteshare` and the explanatory variable was `presvote`.

   The function call to generate the model is:

   ```
   mod_vote_pres <- lm(voteshare ~ presvote, data = dat)
   ```

   and the results are in Table 3

2. Scatterplot of the two variables, with the regression line in Figure 3.

3. Prediction Equation

$$voteshare = 0.441330 + (0.388018) * presvote$$

   `voteshare` is 0.441330 when `presvote` is 0 and it increases by 0.388018 for each unit increase in `presvote`, ie the incumbent's share of the vote increased when their presidential candidate received a higher share of the vote.

Table 3: Vote share as a function of Presidential vote share

	Dependent variable:
	voteshare
presvote	0.388018***
	(0.013493)
Constant	0.441330***
	(0.007599)
Observations	3,193
R^2	0.205814
Adjusted R^2	0.205565
Residual Std. Error	0.088146 (df = 3191)
F Statistic	826.950200*** (df = 1; 3191)
Note:	*p<0.1; **p<0.05; ***p<0.01

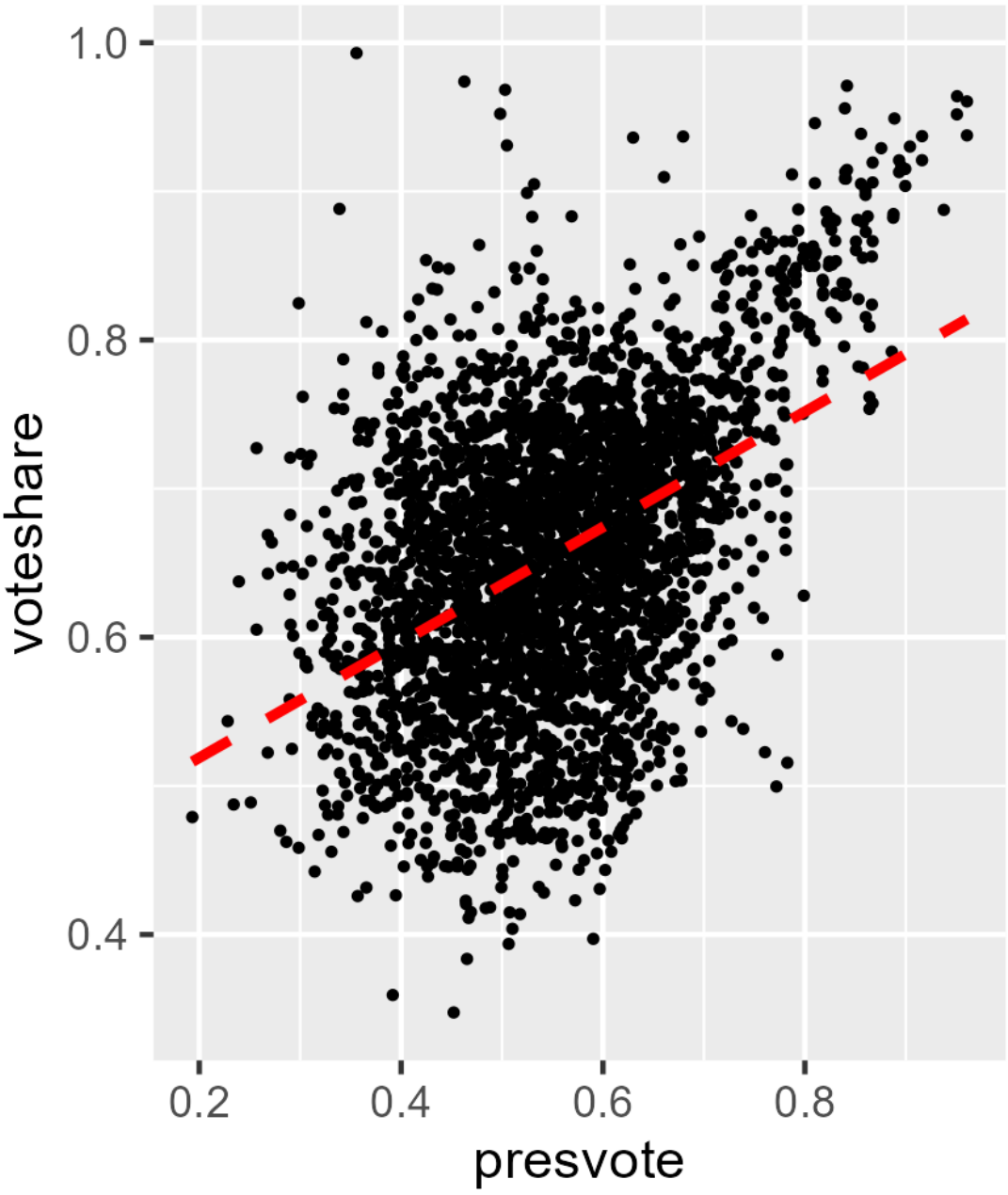

Figure 3: Incumbent vote share as a function of presidential vote share

Question 4

The residuals from Question 1 tell us how much of the variation in `voteshare` is *not* explained by the difference in spending between incumbent and challenger. The residuals in Question 2 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. A linear regression was run where the outcome variable was the residuals from Question 1 and the explanatory variable was the residuals from Question 2.

   The function call to generate the model is:

```
resid_dat = tibble(resid_pres_spend, resid_vote_spend)
mod_resid_vote_pres <- lm(resid_vote_spend ~ resid_pres_spend, data =
    resid_dat)
```

   and the results are in Table 4.

2. Scatterplot of the two sets of residuals, with the regression line in Figure 4.

3. Prediction Equation

$$voteshareresiduals = -5.207e-18 + (0.2569) * presvoteresiduals$$

   The `voteshare` residual value is 0 when `presvote` residual value is 0; it increases by 0.2569 for each unit increase in `presvote` residuals

   The value of the incumbent vote share not accounted for by the difference in incumbent spending increases by 0.2569 for each unit increase in the value of the factors which cause an increase in presidential vote share, excluding incumbent spending.

Table 4: Incumbent's vote share residuals as a function of Presidential vote share residuals

	Dependent variable:
	resid_vote_spend
resid_pres_spend	0.256877***
	(0.011762)
Constant	−0.000000
	(0.001299)
Observations	3,193
R^2	0.130038
Adjusted R^2	0.129765
Residual Std. Error	0.073380 (df = 3191)
F Statistic	476.974700*** (df = 1; 3191)
Note:	*p<0.1; **p<0.05; ***p<0.01

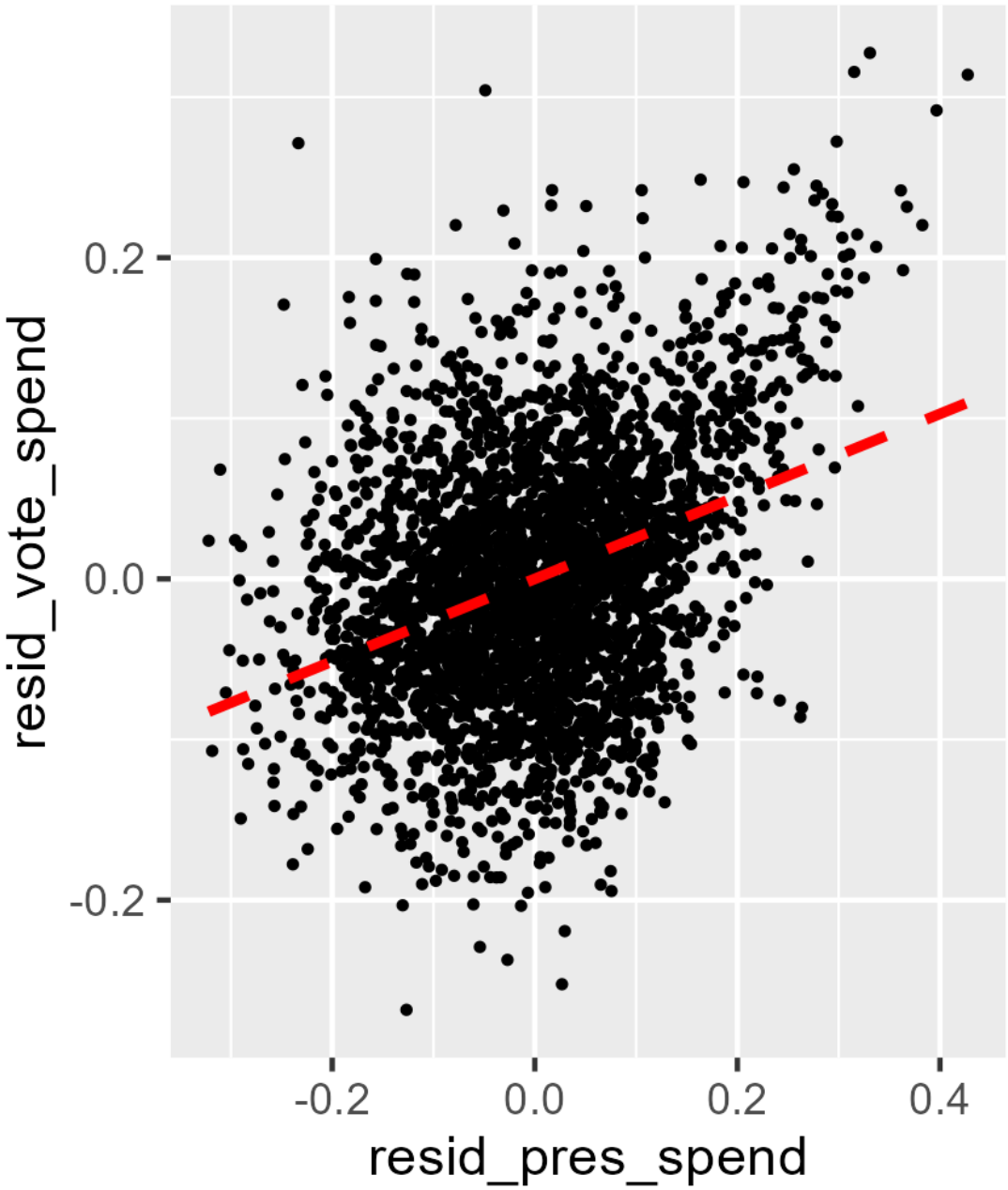

Figure 4: Incumbent's vote share residuals as a function of Presidential vote share residuals

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. A linear regression was run where the outcome variable was `voteshare` and the explanatory variables were `presvote` and `difflog`.

   The function call to generate the model is:

   ```
   mod_vote_spend_pres <- lm(voteshare ~ presvote + difflog, data = dat)
   ```

   The results of the linear model are in Table 5(1).

   The additional variables are plotted in Figure 5.[1]

2. Prediction Equation

   $$voteshare = 0.4486442 + (0.0355431) * difflog + (0.2568770) * presvote$$

   `voteshare` is 0.4486442 when `difflog` and `presvote` are 0; it increases by 0.0355431 for each unit increase in `difflog` (holding `presvote` constant); it increases by 0.2568770 for each unit increase in `presvote` (holding `difflog` constant).

3. The coefficient for residual presidential vote share in Q4 is the same as the coefficient for presidential vote share in Q5, ie 0.256877.

   In model 5, the coefficient for `presvote` is a partial predictor, with `difflog` held constant. The residuals from model 2 represent the variation in the value of the presidential vote, excluding `difflog` (which was specifically accounted for as a predictor). In both models, we are getting a predictive value for `voteshare` based on `presvote`, with `difflog` excluded/controlled.

[1] https://stackoverflow.com/questions/59150905/is-there-a-ggplot2-analogue-to-the-avplots-function-in-r

Table 5: Vote share as a function of Presidential vote share and differential spending

	Dependent variable:	
	voteshare	resid_vote_spend
	(1)	(2)
presvote	0.256877***	
	(0.011764)	
difflog	0.035543***	
	(0.000946)	
resid_pres_spend		0.256877***
		(0.011762)
Constant	0.448644***	−0.000000
	(0.006330)	(0.001299)
Observations	3,193	3,193
R^2	0.449610	0.130038
Adjusted R^2	0.449265	0.129765
Residual Std. Error	0.073391 (df = 3190)	0.073380 (df = 3191)
F Statistic	1,302.947000*** (df = 2; 3190)	476.974700*** (df = 1; 3191)
Note:		*p<0.1; **p<0.05; ***p<0.01

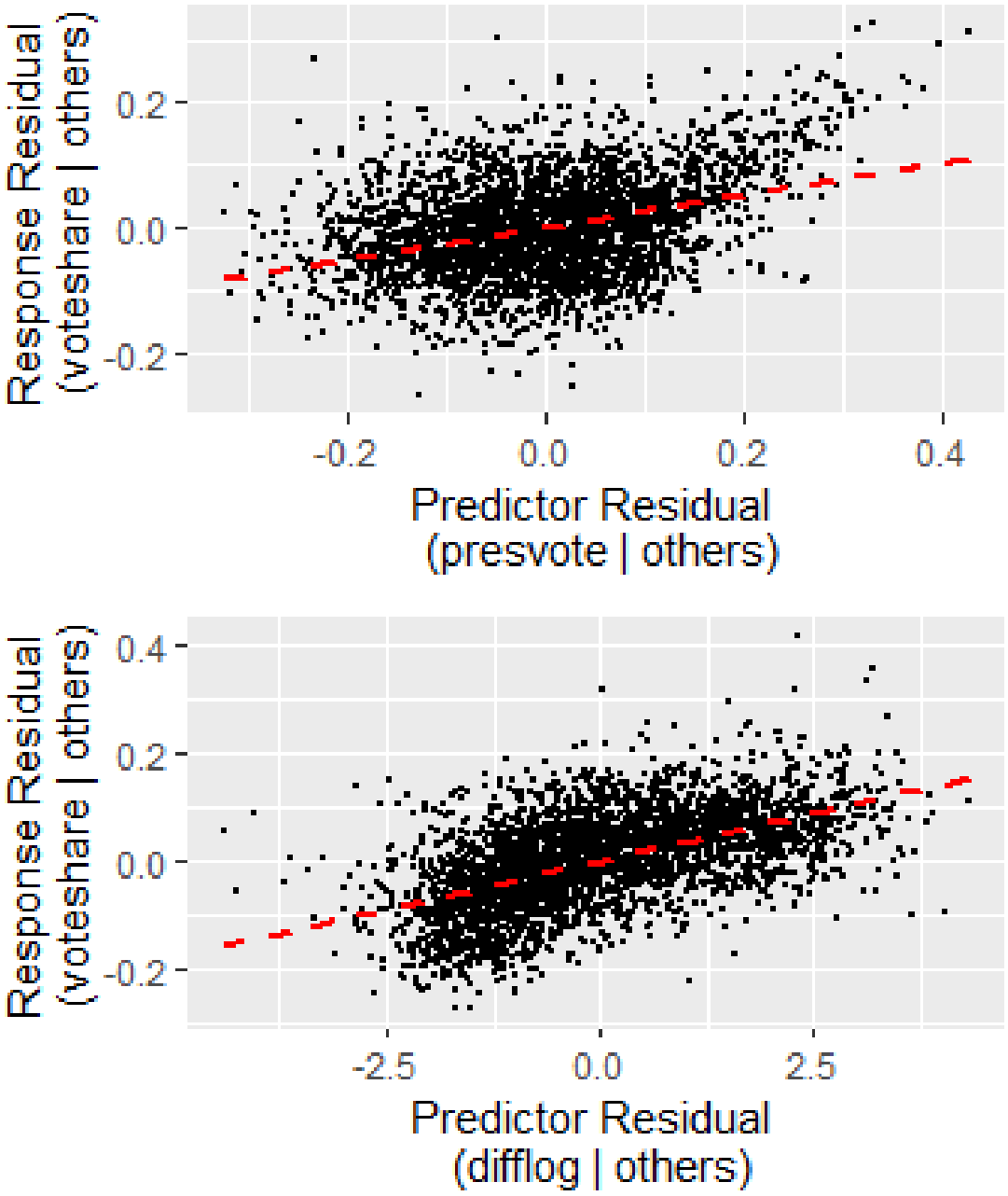

Figure 5: Added variable plots for differential spending and presidential vote share

Appendix - Code
Code in `PS03_ImeldaFinn.R`