

Problem Set 3

Applied Stats/Quant Methods 1

Jack Merriman

Question 1

(1)

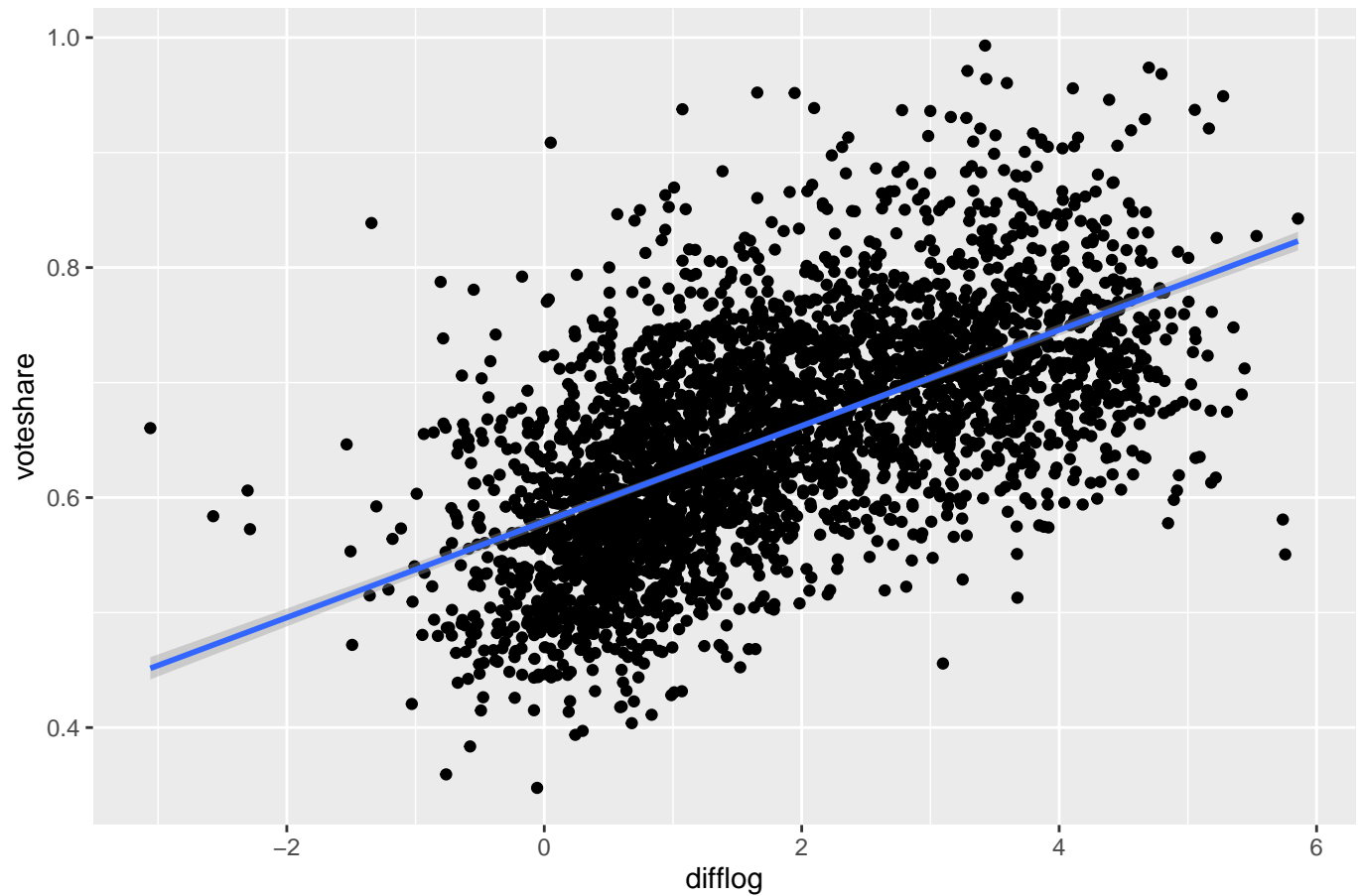
```
1 lm_vd <- lm(formula = voteshare ~ difflog, data = inc.sub)
2 stargazer(lm_vd, title = "Linear Regression: Vote Share – Spending Difference")
```

Table 1: Linear Regression: Vote Share and Spending Difference

	<i>Dependent variable:</i>
	voteshare
difflog	0.042*** (0.001)
Constant	0.579*** (0.002)
Observations	3,193
R ²	0.367
Adjusted R ²	0.367
Residual Std. Error	0.079 (df = 3191)
F Statistic	1,852.791*** (df = 1; 3191)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

(2)

```
1 inc.sub %>%  
2   ggplot(aes(x = difflog, y = voteshare))+  
3   geom_point()+  
4   geom_smooth(method = lm)
```



(3)

```
1 res_vd <- unlist(residuals(lm_vd))
```

(4)

Where Y is `voteshare` and x is `difflog` the prediction equation is as follows:

$$Y = 0.58 + 0.04x$$

Question 2

(1)

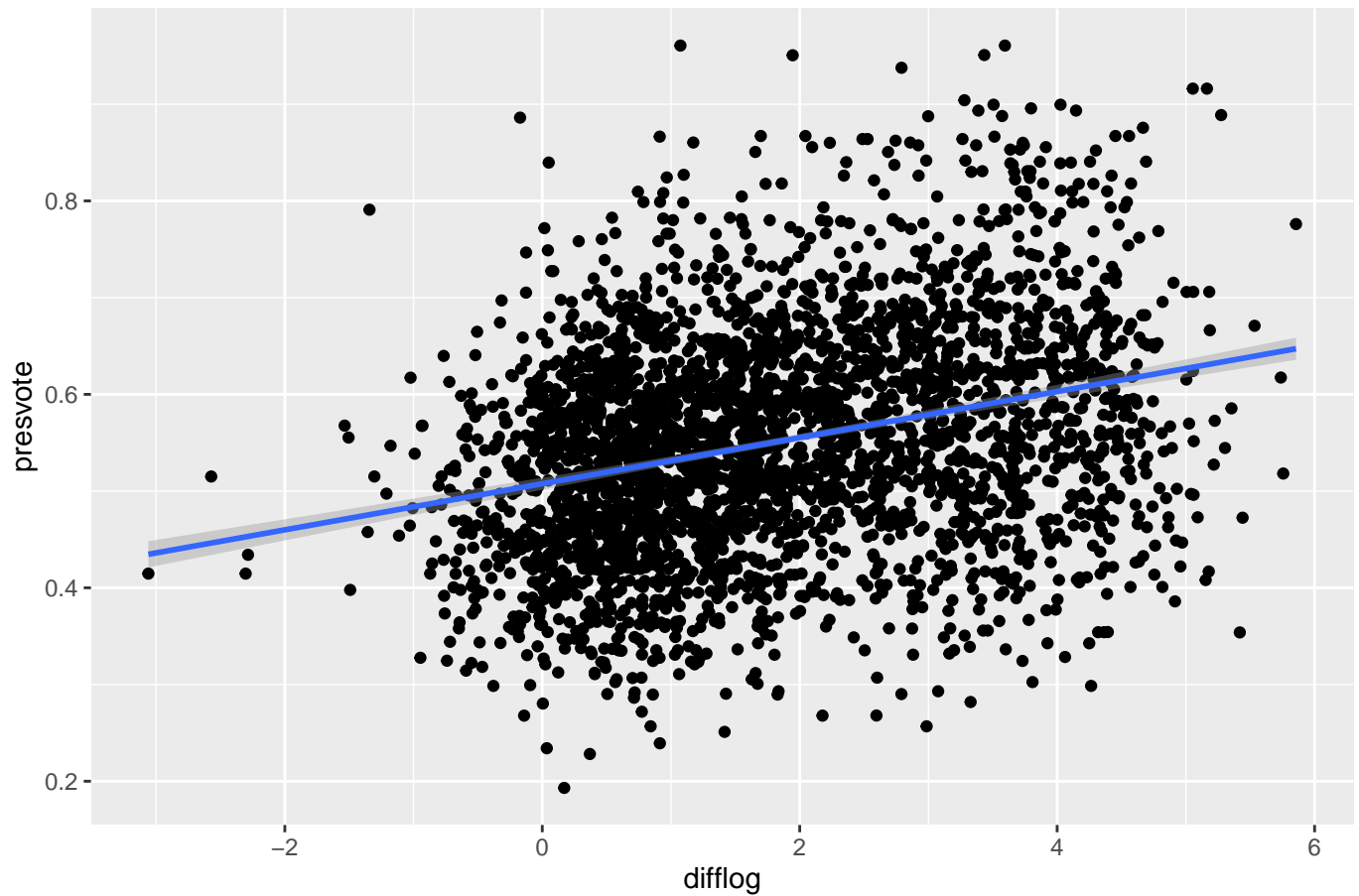
```
1 lm_pd <- lm(formula = presvote ~ difflog, data = inc.sub)
2 stargazer(lm_pd, title = "Linear Regression: Presidential Vote Share –
  Spending Difference")
```

Table 2: Linear Regression: Presidential Vote Share - Spending Difference

<hr/> <hr/>	
	<i>Dependent variable:</i>
	<hr/> presvote
difflog	0.024*** (0.001)
Constant	0.508*** (0.003)
<hr/>	
Observations	3,193
R ²	0.088
Adjusted R ²	0.088
Residual Std. Error	0.110 (df = 3191)
F Statistic	307.715*** (df = 1; 3191)
<hr/> <hr/>	
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

(2)

```
1 inc.sub %>%  
2   ggplot(aes(x = difflog , y = presvote))+  
3   geom_point()+  
4   geom_smooth(method = lm)
```



(3)

```
1 res_pd <- unlist(residuals(lm_pd))
```

(4)

Where Y is `presvote` and x is `difflog` the prediction equation is as follows:
 $Y = 0.51 + 0.02x$

Question 3

(1)

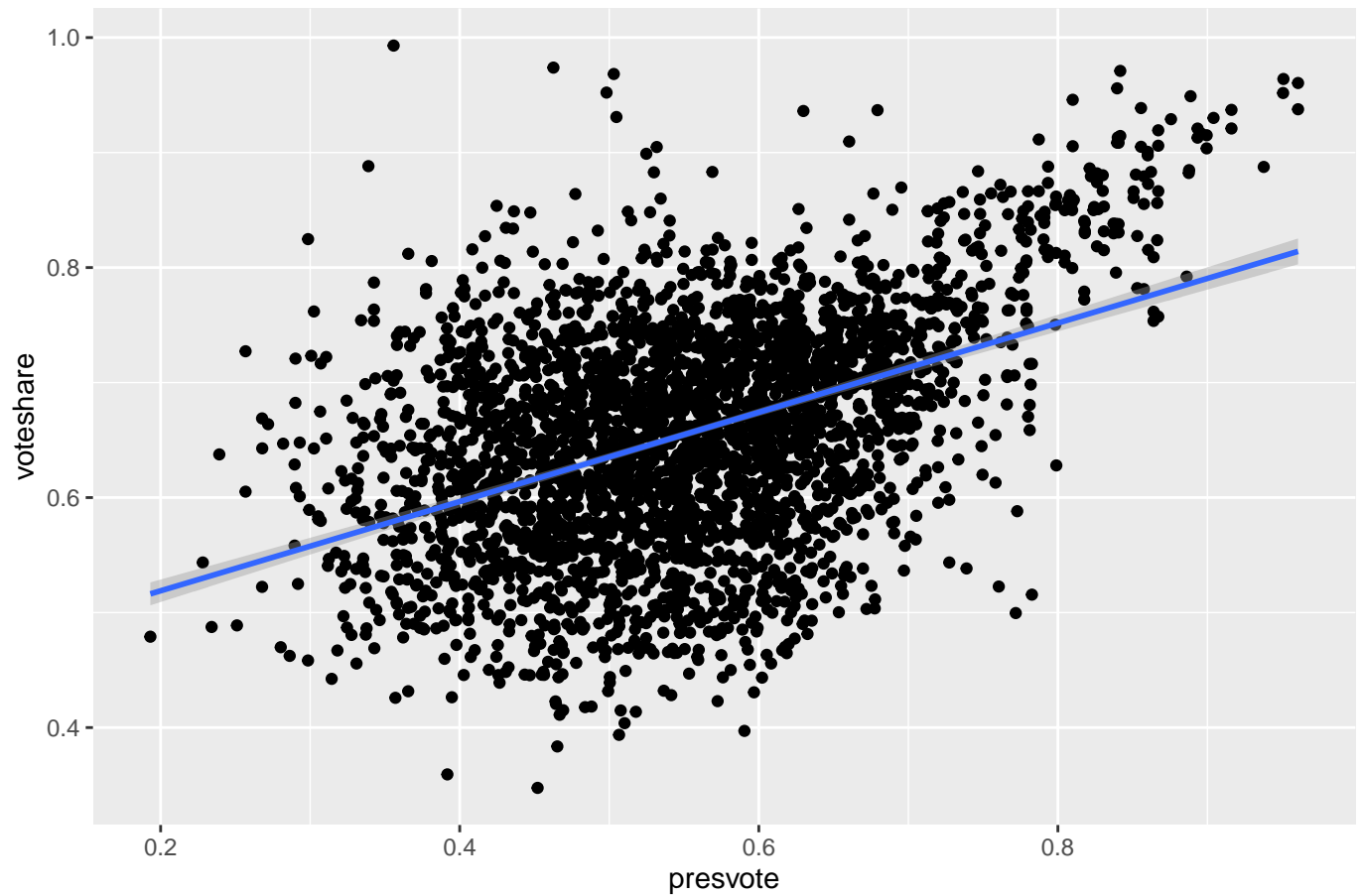
```
1 lm_vp <- lm(formula = voteshare ~ presvote, data = inc.sub)
2 stargazer(lm_vp, title = "Linear Regression: Vote Share - Presidential Vote
  Share")
```

Table 3: Linear Regression: Vote Share - Presidential Vote Share

<hr/> <hr/>	
	<i>Dependent variable:</i>
	<hr/> voteshare <hr/>
presvote	0.388*** (0.013)
Constant	0.441*** (0.008)
<hr/>	
Observations	3,193
R ²	0.206
Adjusted R ²	0.206
Residual Std. Error	0.088 (df = 3191)
F Statistic	826.950*** (df = 1; 3191)
<hr/> <hr/>	
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

(2)

```
1 inc.sub %>%  
2   ggplot(aes(x = presvote, y = voteshare))+  
3   geom_point()+  
4   geom_smooth(method = lm)
```



(3)

Where Y is voteshare and x is presvote the prediction equation is as follows:

$$Y = 0.44 + 0.39x$$

Question 4

(1)

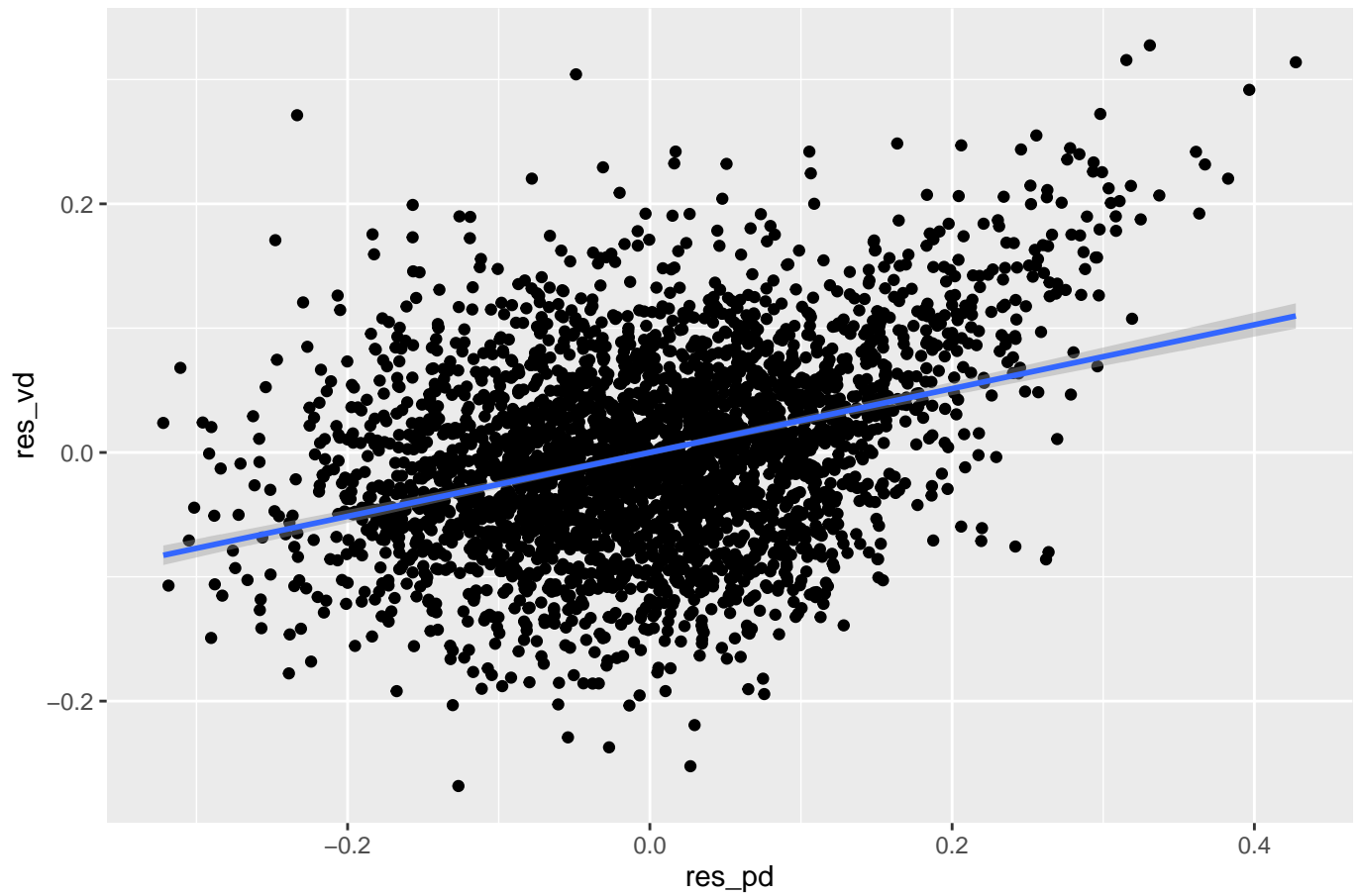
```
1 lm_res <- lm(formula = res_vd ~ res_pd)
2 stargazer(lm_res, title = "Linear Regression: Residual Model")
```

Table 4: Linear Regression: Residual Model

	<i>Dependent variable:</i>
	res_vd
res_pd	0.257*** (0.012)
Constant	−0.000 (0.001)
Observations	3,193
R ²	0.130
Adjusted R ²	0.130
Residual Std. Error	0.073 (df = 3191)
F Statistic	476.975*** (df = 1; 3191)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

(2)

```
1 tib_res <- tibble(res_vd, res_pd)
2 tib_res %>%
3   ggplot(aes(x = res_pd, y = res_vd))+
4   geom_point()+
5   geom_smooth(method = lm)
```



(3)

Where Y is `res_vd` and x is `res_pd` the prediction equation is as follows:

$$Y = 0.44 + 0.39x$$

Question 5

(1)

```
1 final_model <- lm(formula = voteshare ~ difflog + presvote, data = inc.sub)
2 stargazer(final_model, title = "Linear Regression: Vote Share - Spending
  Difference and Presidential Vote")
```

Table 5: Linear Regression: Vote Share - Spending Difference and Presidential Vote

	<i>Dependent variable:</i>
	voteshare
difflog	0.036*** (0.001)
presvote	0.257*** (0.012)
Constant	0.449*** (0.006)
Observations	3,193
R ²	0.450
Adjusted R ²	0.449
Residual Std. Error	0.073 (df = 3190)
F Statistic	1,302.947*** (df = 2; 3190)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

(2)

Where Y is **voteshare**, x_1 is **difflog** and x_2 is **presvote** the prediction equation is as follows:

$$Y = 0.45 + 0.04x_1 + 0.26x_2$$

(3)

When using R's **summary()** command, we can see the residuals of the two models are the same. The residual model investigated the confounding effects between **presvote** and **difflog**'s effect on **voteshare**, so a model including them both as input variables will have the same residuals as the residual model