Lab5

1512005

6 November 2017

```
require(tibble)
```

```
## Loading required package: tibble
```

```
require(readr)
```

```
## Loading required package: readr
```

```
require(tidyverse)
```

```
## Loading required package: tidyverse
```

```
## Warning in library(package, lib.loc = lib.loc, character.only = TRUE,
## logical.return = TRUE, : there is no package called 'tidyverse'
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

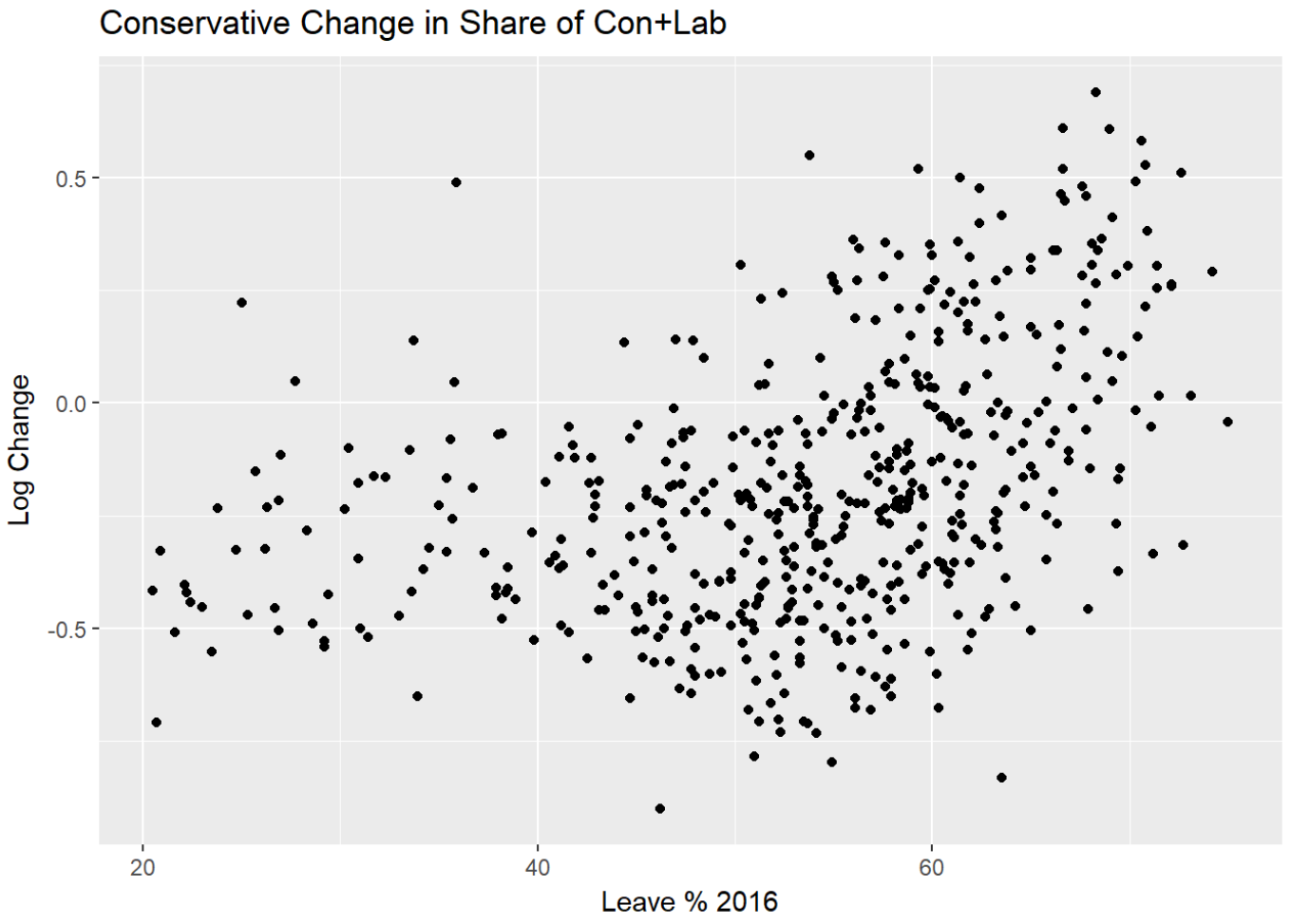
```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
#Remember to load the packages needed

voting <- read.csv("GE-referendum-and-census.csv") #Dataset of votes cast provided
to us
England <- voting %>% filter(!(region %in% c("Wales", "Scotland"))) #Remove Wales
and Scotland from the data

#Calculate lab/con change
England <- England %>% mutate(con_lab_change = log((con17 / lab17) / (con15 / lab1
5))) #Calculate lab/con change

# Calculate libdem change
England <- England %>% mutate(ld_change = log((ld17/ld17+lab17+con17)/(ld15/ld15+c
on15+lab15)))
```

```
library(ggplot2)
#Use ggplot to make simple plot of Conservative and Labour change against leave vot
e.
con_lab_change.map <- England %>%
                      ggplot(aes(x = leave16sh, y= con_lab_change)) +
                      geom_point()+
                      ggtitle("Conservative Change in Share of Con+Lab")+
                      xlab("Leave % 2016")+
                      ylab("Log Change")
print(con_lab_change.map)
```

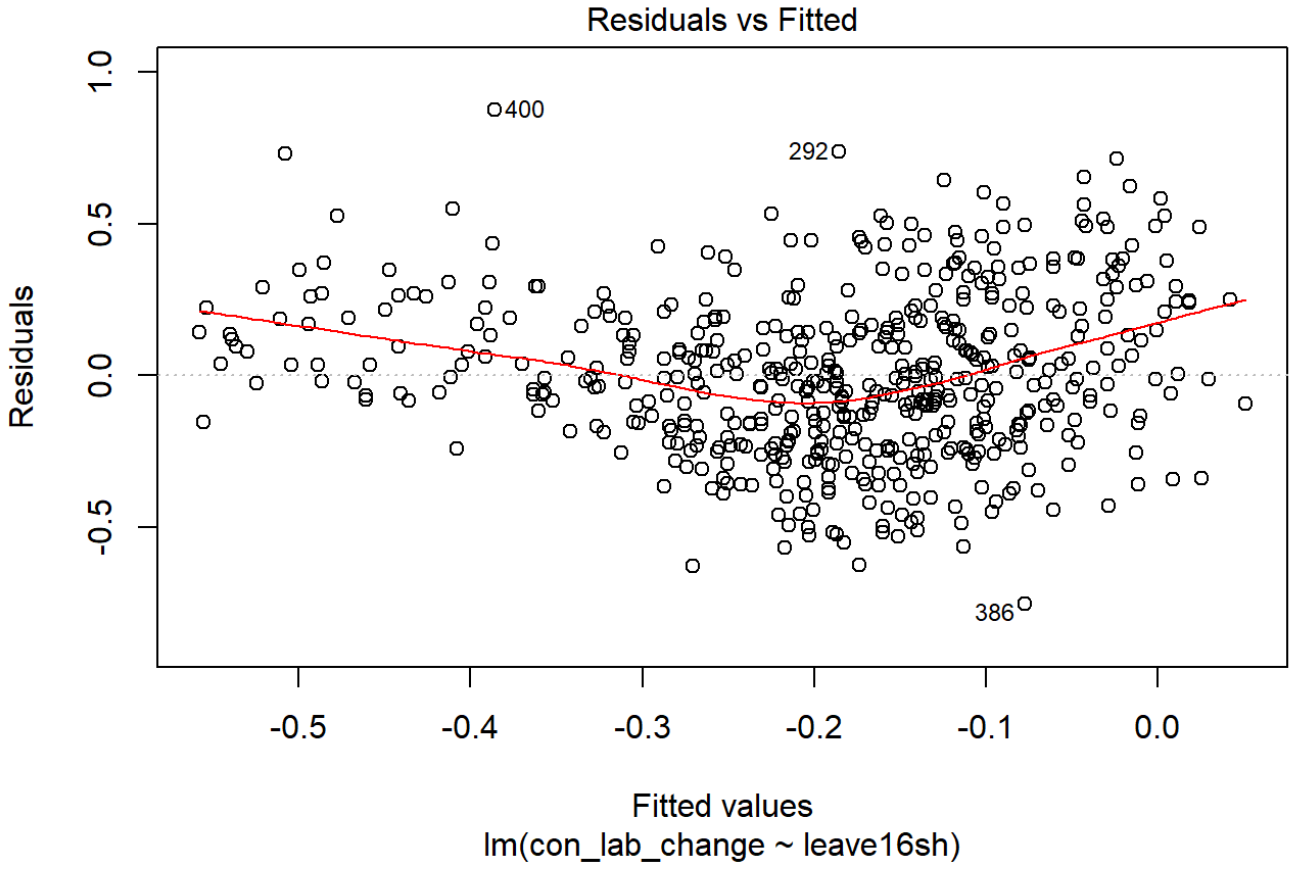
```
## Warning: Removed 1 rows containing missing values (geom_point).
```

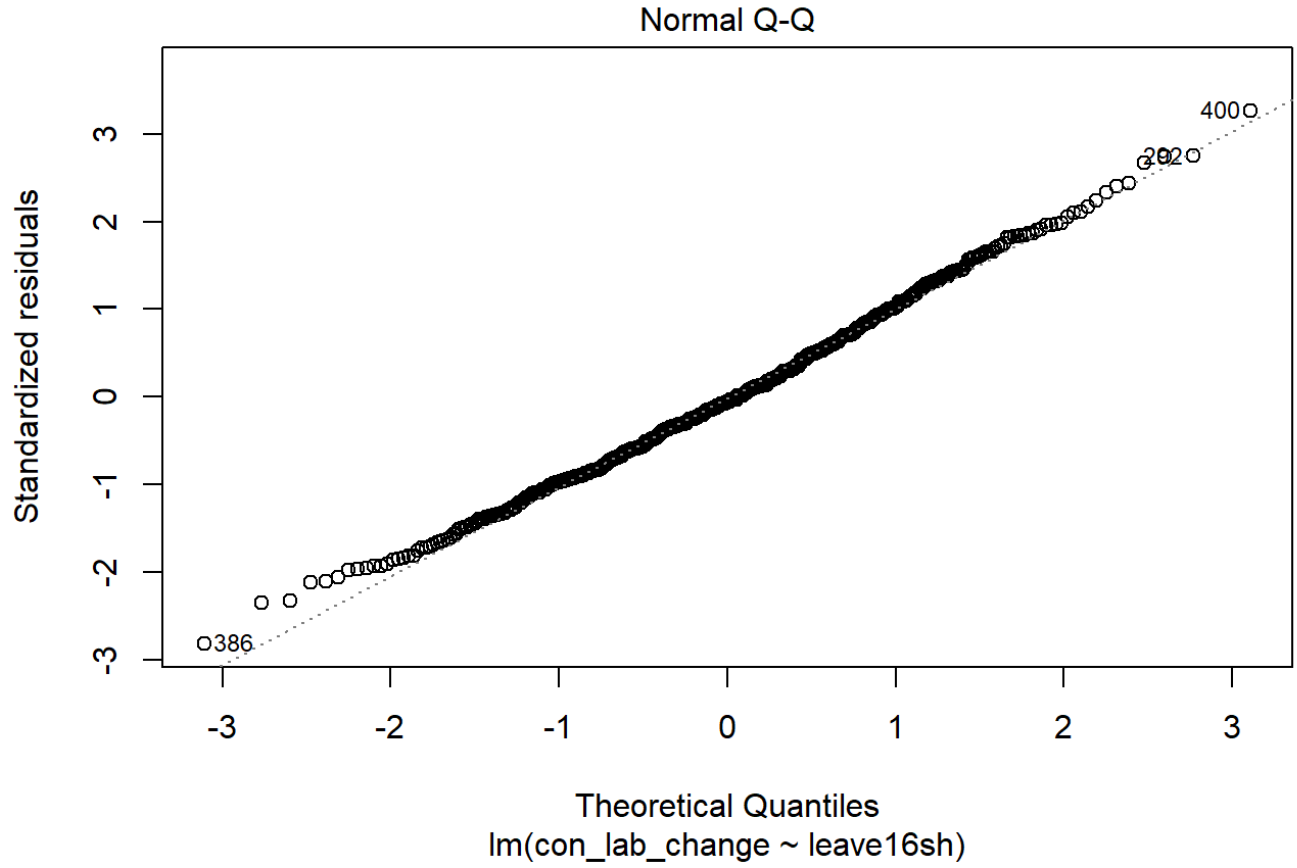


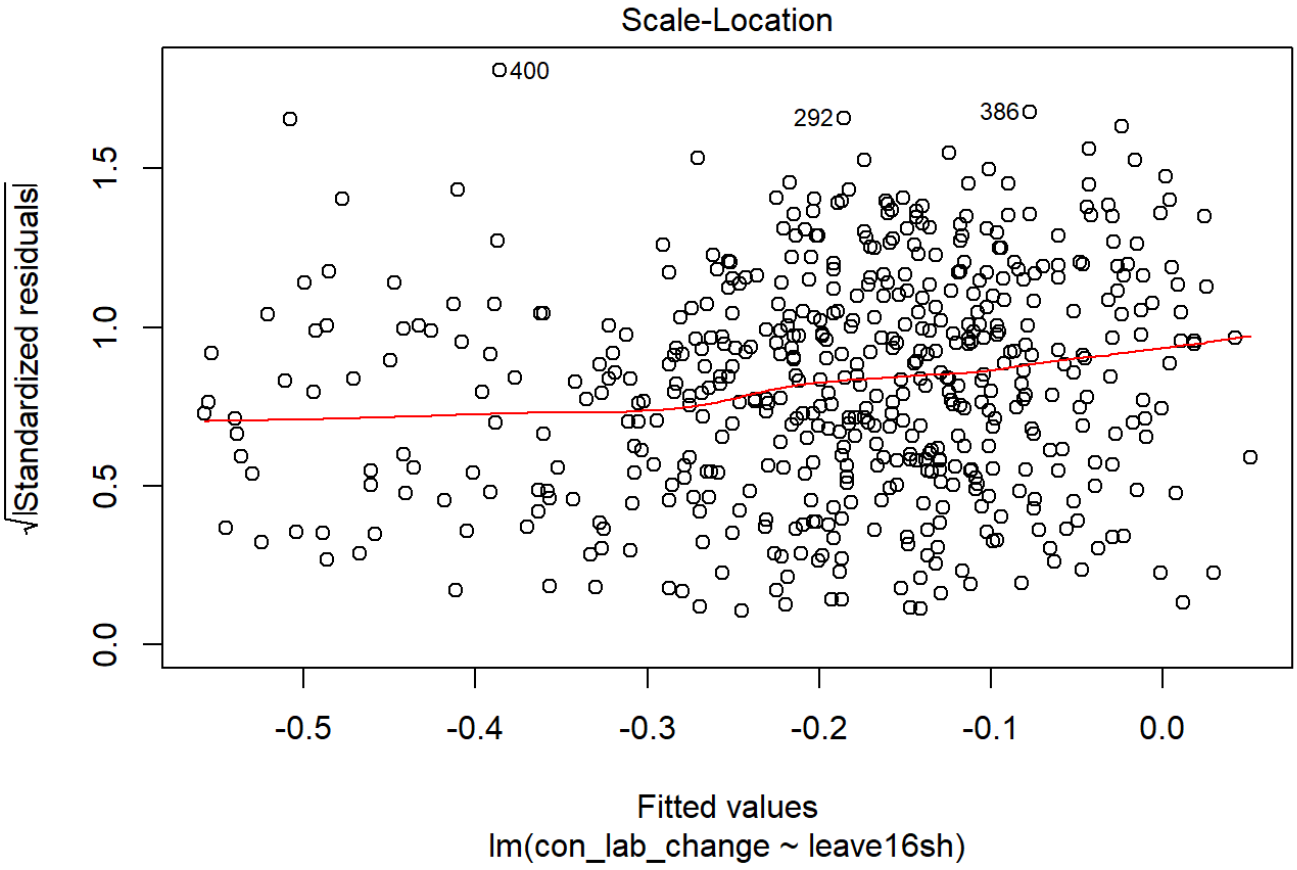

```
#Looking at the plot it appears the conservatives generally gained votes in constit
uencies with higher leave votes.
#ggplot automatically omits any 0's in the data. This is fine as a 0 value in the d
ata implies a candidate did not stand for, in this analysis, LD or for Con.

con_lab_change.model<-lm(con_lab_change ~ leave16sh, data = England)
summary(con_lab_change.model)
```

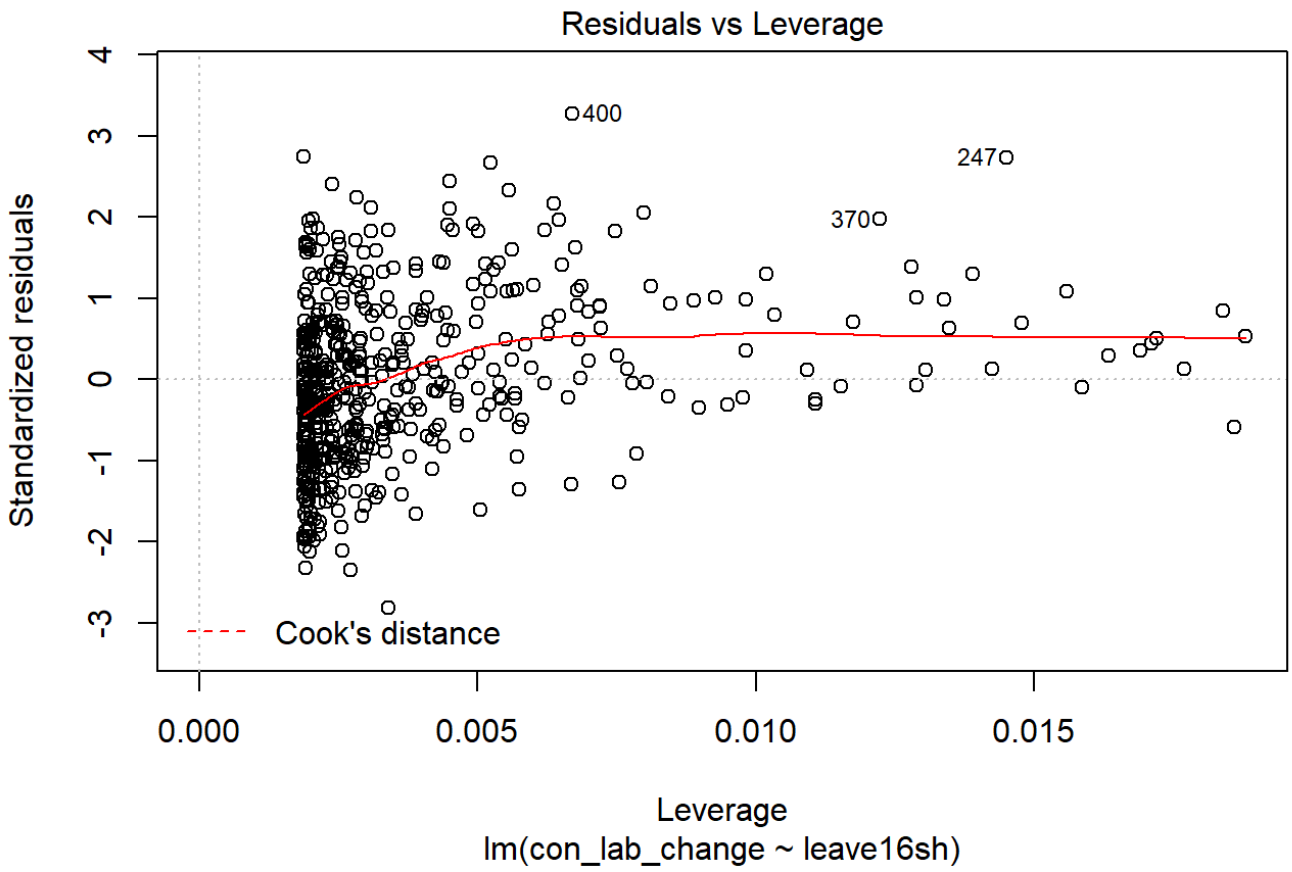
```
##
## Call:
## lm(formula = con_lab_change ~ leave16sh, data = England)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.75573 -0.18860 -0.01465  0.17872  0.87550
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.786977   0.057774  -13.62   <2e-16 ***
## leave16sh    0.011173   0.001056   10.58   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2686 on 530 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.1743, Adjusted R-squared:  0.1727
## F-statistic: 111.9 on 1 and 530 DF,  p-value: < 2.2e-16
```

```
#Our hypothesis from the plot appears true here, as leave is significant at the 1% level, with a positive valued coefficient


plot(con_lab_change.model) #Residuals seem a little skewed, ie not random but we will ignore for now.
```



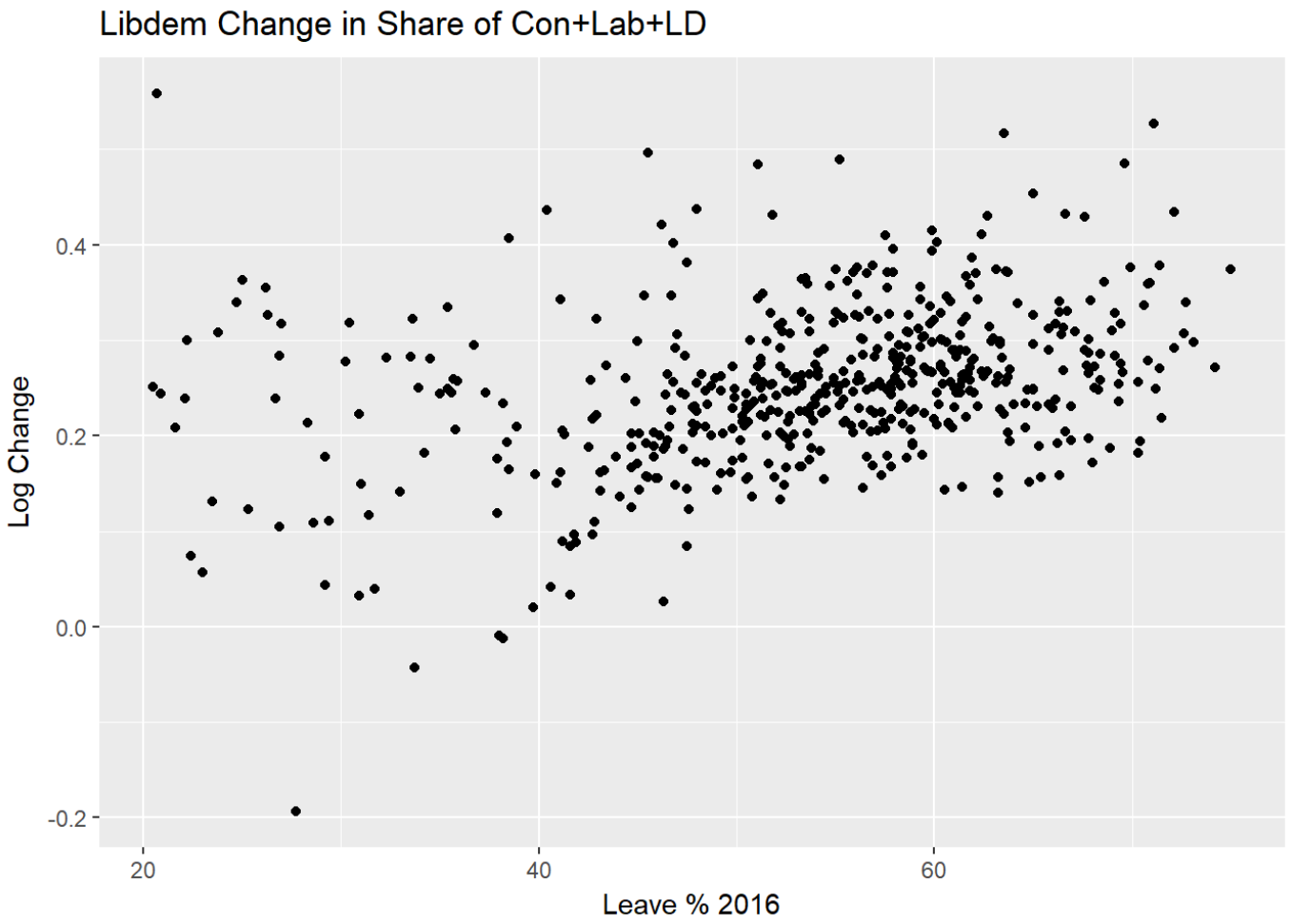

Normal Q-Q
Standardized residuals
Theoretical Quantiles
lm(con_lab_change ~ leave16sh)
400
292
386
-3
-2
-1
0
1
2
3


Scale-Location
√|Standardized residuals|
Fitted values
lm(con_lab_change ~ leave16sh)
400
292
386
0.0
0.5
1.0
1.5
-0.5
-0.4
-0.3
-0.2
-0.1
0.0




```
ld_change.map <- England %>%
                      ggplot(aes(x = leave16sh, y= ld_change))+
                      geom_point()+
                      ggtitle("Libdem Change in Share of Con+Lab+LD")+
                      xlab("Leave % 2016")+
                      ylab("Log Change")
print(ld_change.map)
```

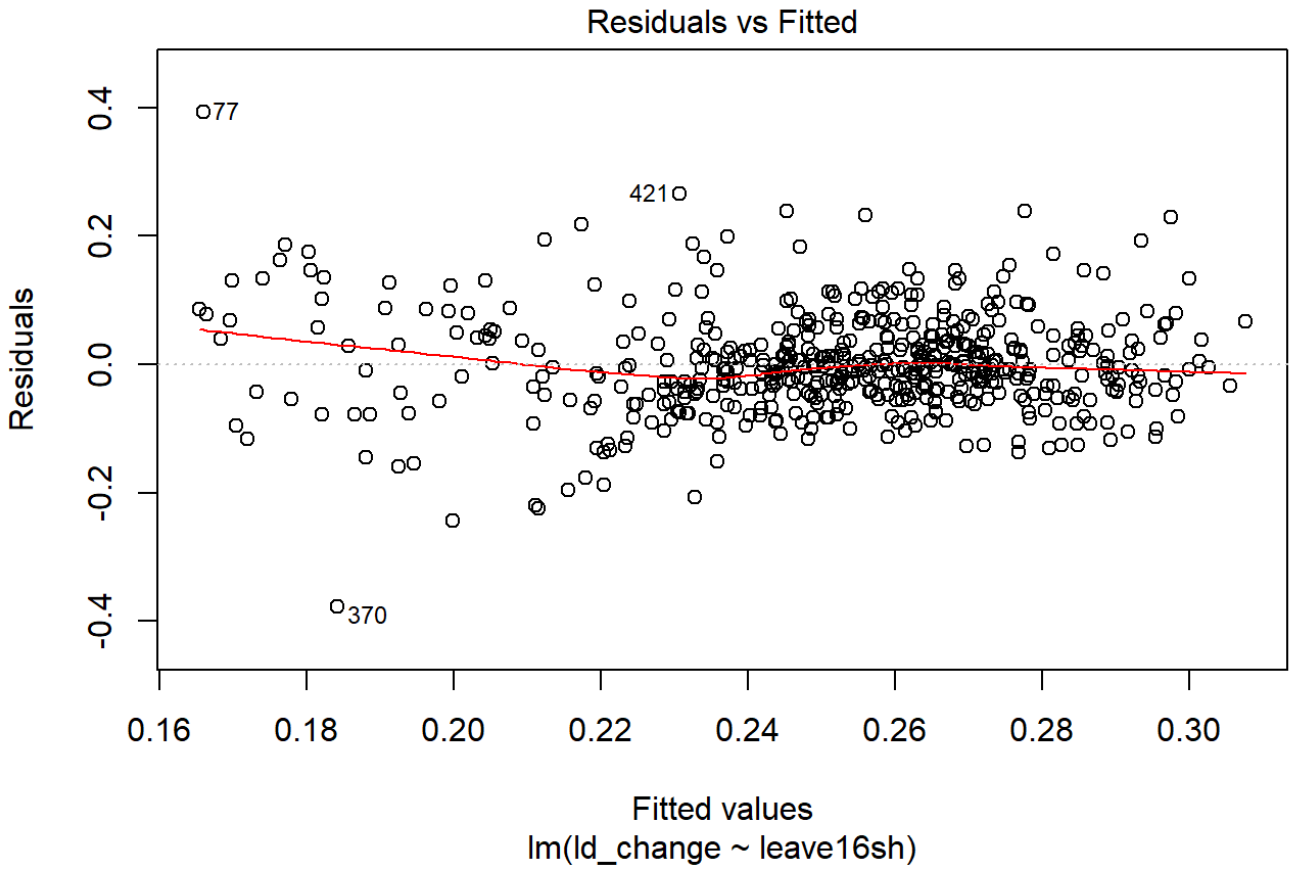
```
## Warning: Removed 3 rows containing missing values (geom_point).
```

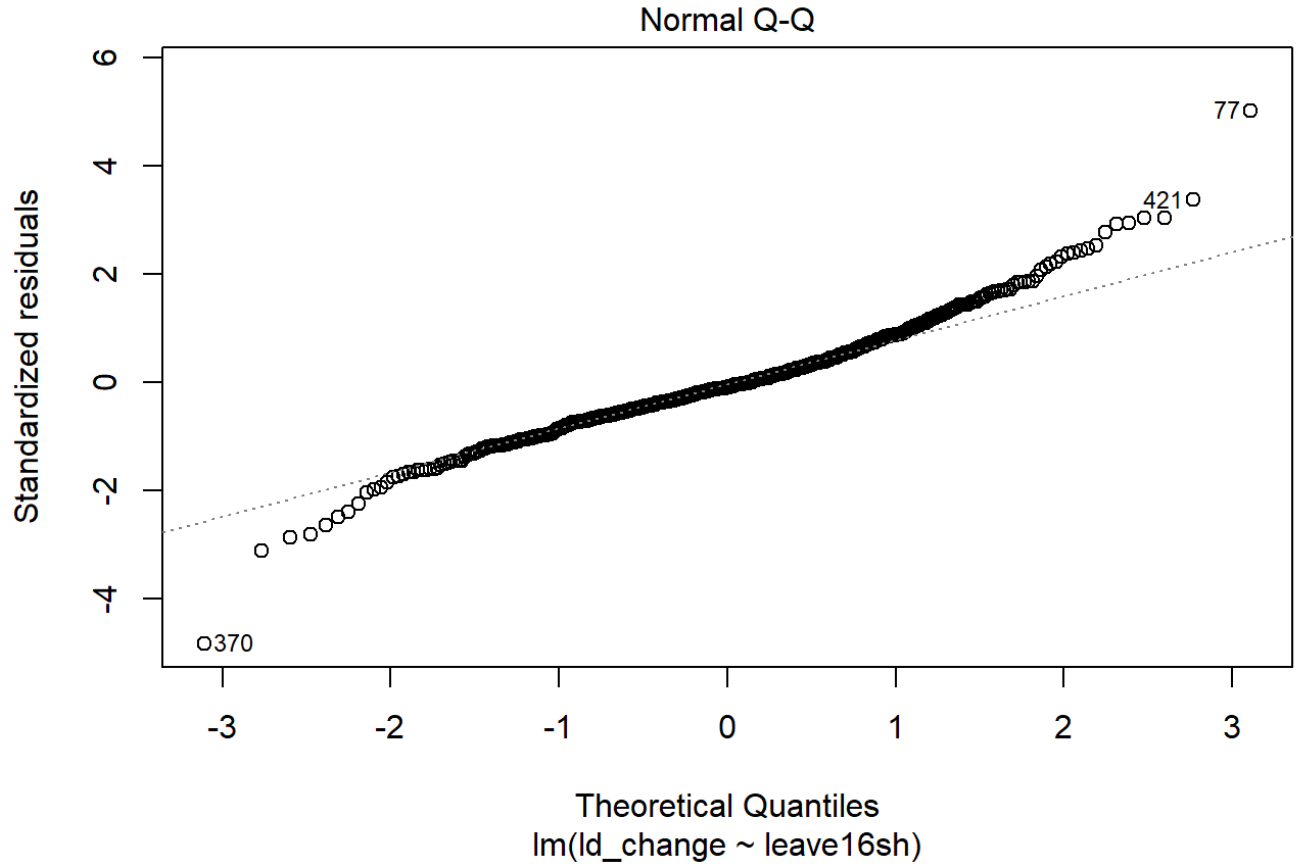


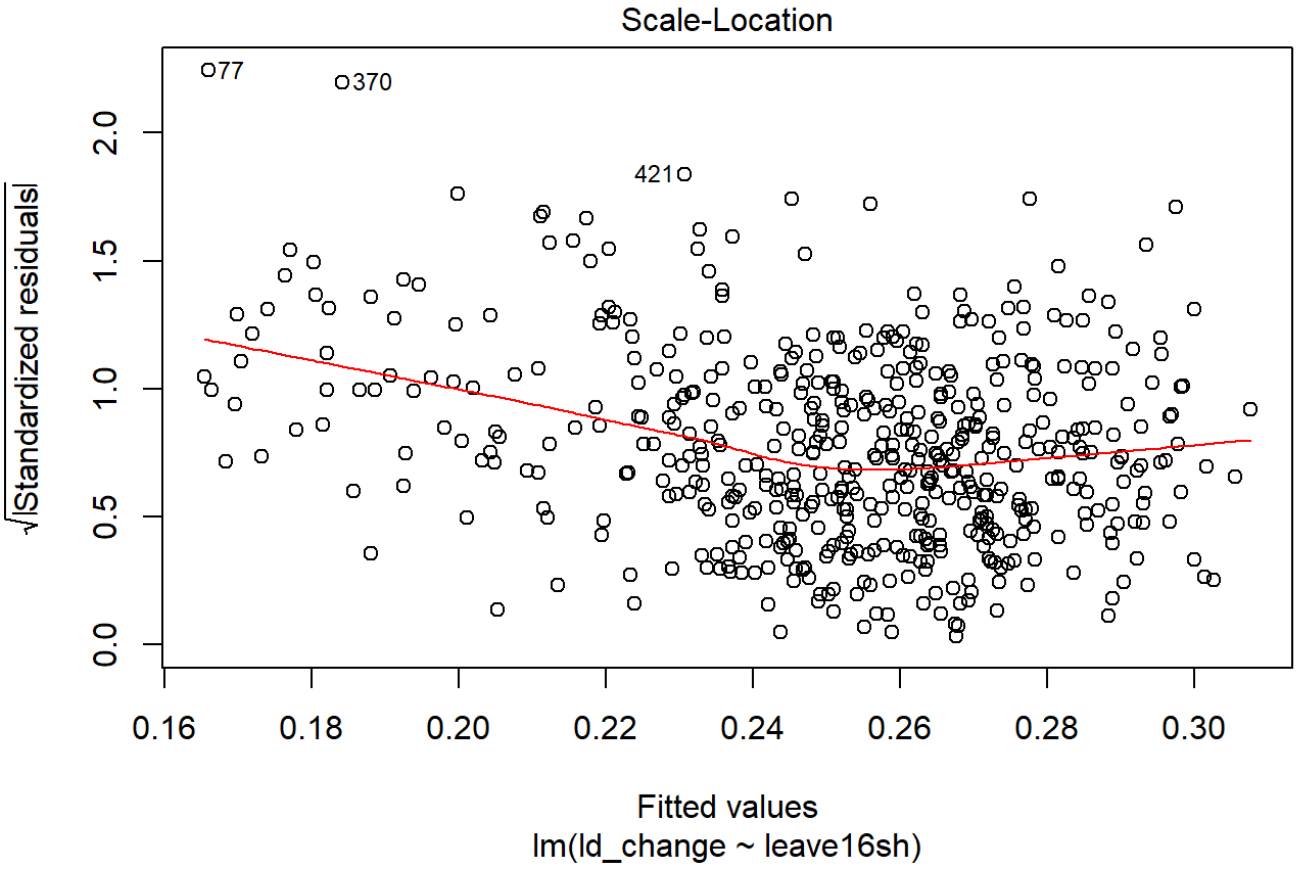

```
#Again ggplot omits 3 values that are missing, this is okay.
#The plot appears to suggest that the liberal democrats generally increased their v
ote share regardless of whether the constituency voted leave, in fact they only see
med to lose voteshare in places that were very low in leave vote.
ld_change.model<-lm(ld_change ~ leave16sh, data = England)
summary(ld_change.model)
```

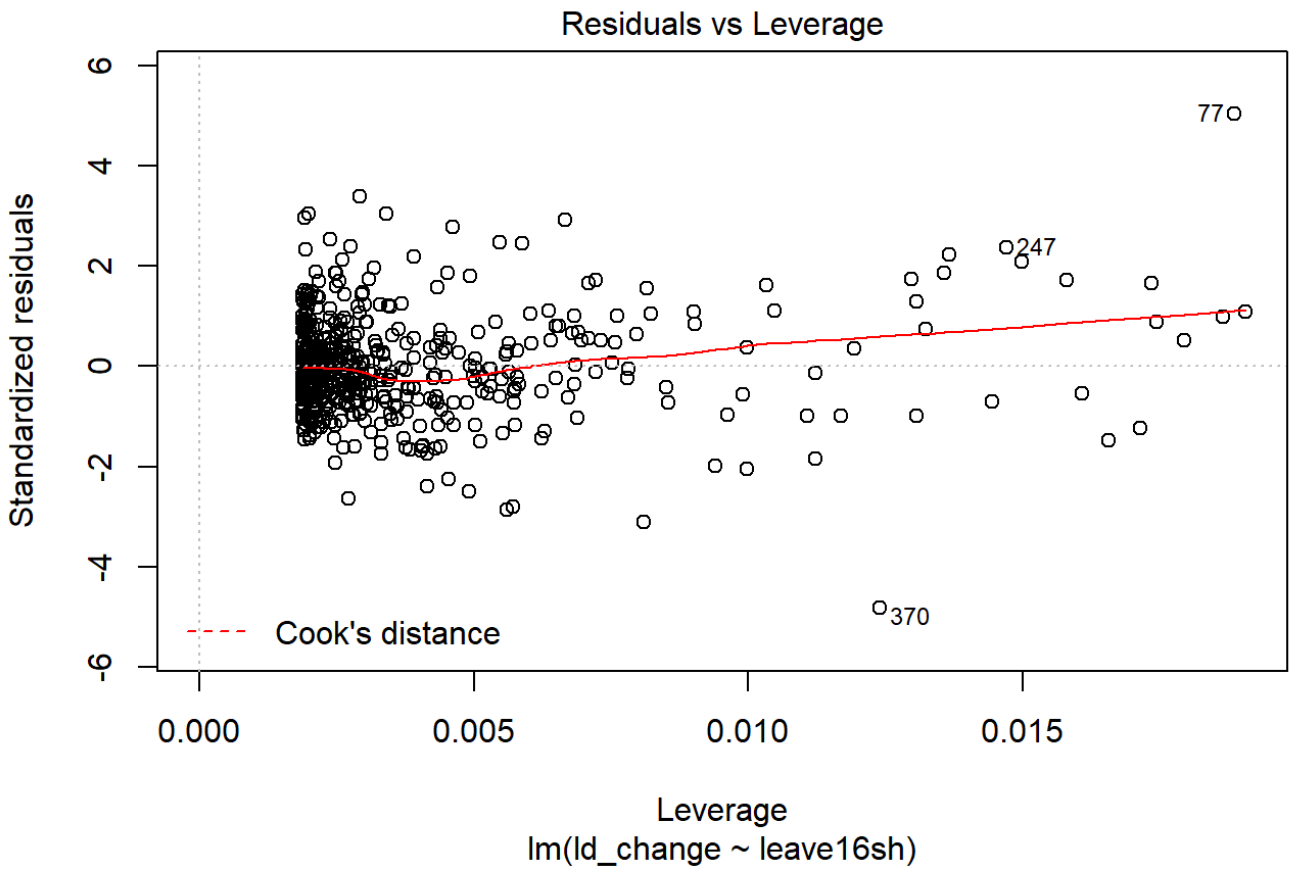
```
##
## Call:
## lm(formula = ld_change ~ leave16sh, data = England)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.37782 -0.04582 -0.00664  0.04059  0.39273
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.1119580  0.0170642   6.561 1.28e-10 ***
## leave16sh   0.0026097  0.0003117   8.371 5.14e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07881 on 528 degrees of freedom
##   (3 observations deleted due to missingness)
## Multiple R-squared:  0.1172, Adjusted R-squared:  0.1155
## F-statistic: 70.08 on 1 and 528 DF,  p-value: 5.14e-16
```

```
#Again, leave is significant at the 1% level, with a real valued coefficient. This
is a surprising result, as it implies the liberal democrats improved more in consti
tuencies with high leave vote than those with low leave vote. Considering their str
ong anti-brexit stance it would be normal to assume the opposite.
#However, it is possible that in areas with high leave votes, the lib dems scooped
up those who did vote remain in the constituency.

plot(ld_change.model) #similar to before, the residuals may show some signs of not
being random.
```

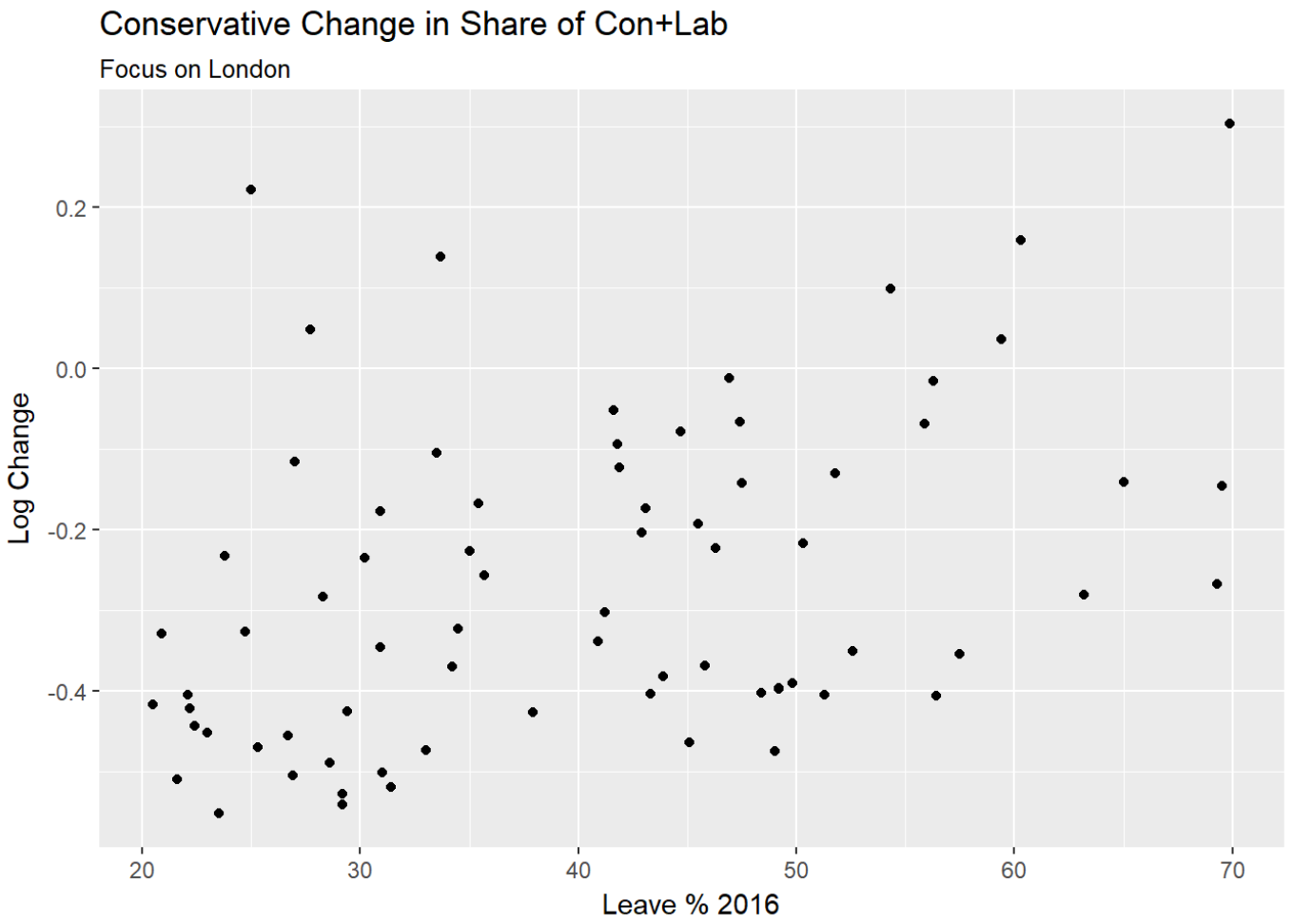



Normal Q-Q
Standardized residuals
Theoretical Quantiles
lm(ld_change ~ leave16sh)
77
421
370


Scale-Location
√|Standardized residuals|
Fitted values
lm(ld_change ~ leave16sh)
77
370
421



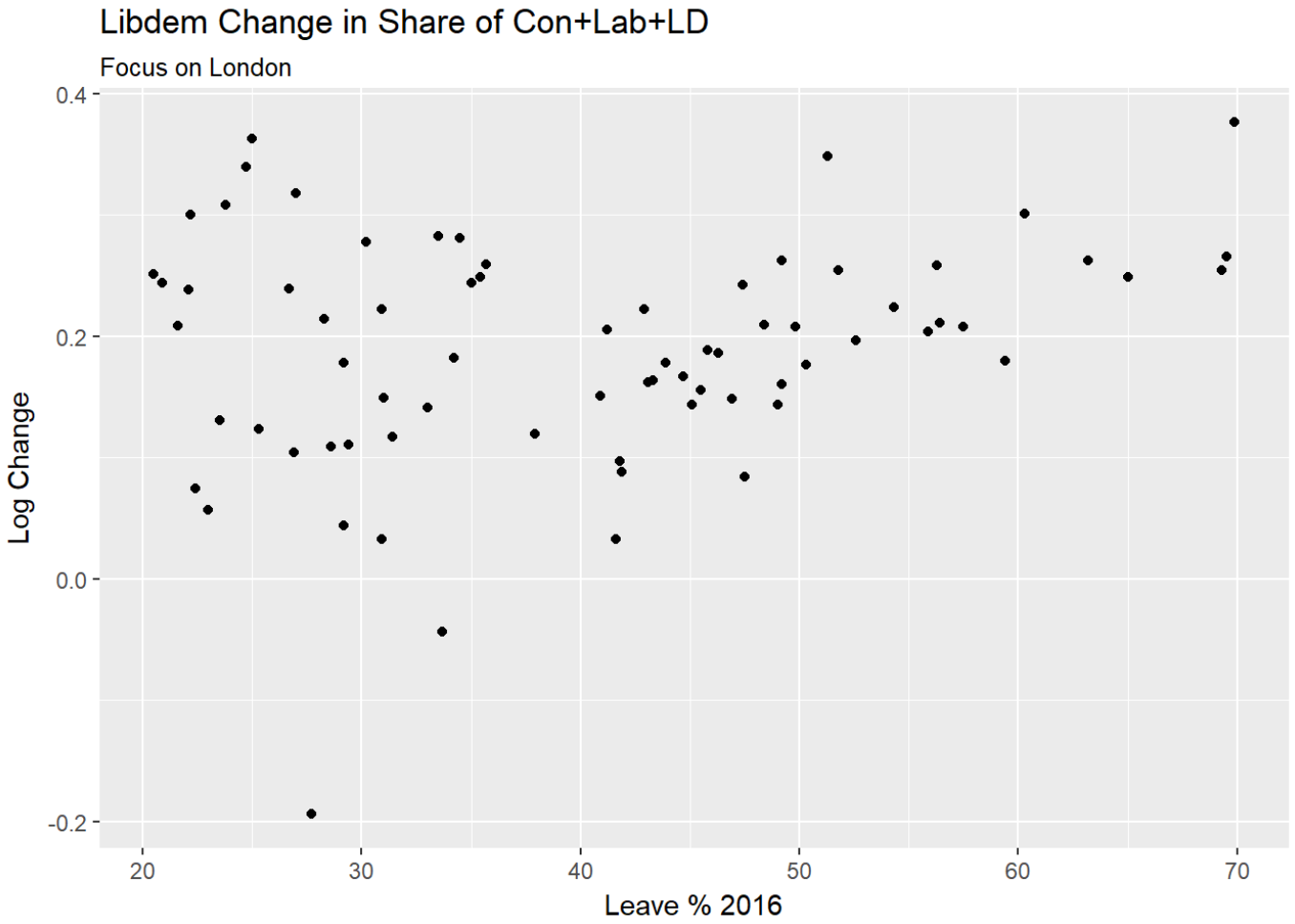

```
#Filter region to London
con_lab_change.lnd <- England %>%
                      filter(region == "London") %>%
                      ggplot(mapping = aes(x = leave16sh, y = con_lab_change))+
                      geom_point()+
                      ggtitle("Conservative Change in Share of Con+Lab", subtitle
= "Focus on London")+
                      xlab("Leave % 2016")+
                      ylab("Log Change")
print(con_lab_change.lnd)
```




```
#Places that votes remain appear to be more likely to move towards Labour in 2017 (
in London)


con_lab_change_model.lnd <- lm(con_lab_change ~ leave16sh, data = England, subset
= (region == "London"))
summary(con_lab_change_model.lnd)
```

```
##
## Call:
## lm(formula = con_lab_change ~ leave16sh, data = England, subset = (region ==
##     "London"))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.26144 -0.15545 -0.02777  0.12834  0.58153
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.511516   0.068443  -7.474 1.58e-10 ***
## leave16sh    0.006082   0.001618   3.758 0.000348 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1822 on 71 degrees of freedom
## Multiple R-squared:  0.1659, Adjusted R-squared:  0.1542
## F-statistic: 14.13 on 1 and 71 DF,  p-value: 0.0003479
```

```
#Leave again is significant at 1% level with positive coefficient. So in London, ju
st as in the rest of the UK, Conservatives gained votes in areas which had a higher
percentage of leave vote.

ld_change.lnd <- England %>%
                 filter(region == "London") %>%
                 ggplot(mapping = aes(x = leave16sh, y = ld_change)) +
                 geom_point()+
                 ggtitle("Libdem Change in Share of Con+Lab+LD", subtitle = "Focus
on London")+
                 xlab("Leave % 2016")+
                 ylab("Log Change")

print(ld_change.lnd)
```




```
#The libdems gained in all but 2 London constituencies, and it does not look like i
t correlates with leave vote.

ld_change_model.lnd <- lm(ld_change ~ leave16sh, data = England, subset = (region
== "London"))
summary(ld_change_model.lnd)
```

```
##
## Call:
## lm(formula = ld_change ~ leave16sh, data = England, subset = (region ==
##     "London"))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.36320 -0.04329 -0.00447  0.05869  0.19781
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.125224   0.034849   3.593 0.000598 ***
## leave16sh   0.001603   0.000824   1.945 0.055691 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09277 on 71 degrees of freedom
## Multiple R-squared:  0.0506, Adjusted R-squared:  0.03723
## F-statistic: 3.784 on 1 and 71 DF,  p-value: 0.05569
```

```
#Indeed, as suspected from the graph, leave vote is not significant for this linear model. Again this is interesting as the libdems tried to paint themselves as the most pro EU party, and they did not seem to do significantly better in those constituencies (in London).
```