Econometrics_Problem_Set_9.R

Bernardo

Tue Nov 12 09:23:30 2019

```
library(readstata13)
```

```
## Warning: package 'readstata13' was built under R version 3.4.4
```

```
library(sandwich)
```

```
## Warning: package 'sandwich' was built under R version 3.4.4
```

```
library(lmtest)
```

```
## Warning: package 'lmtest' was built under R version 3.4.4
```

```
## Loading required package: zoo
```

```
## Warning: package 'zoo' was built under R version 3.4.4
```

```
##
## Attaching package: 'zoo'
```

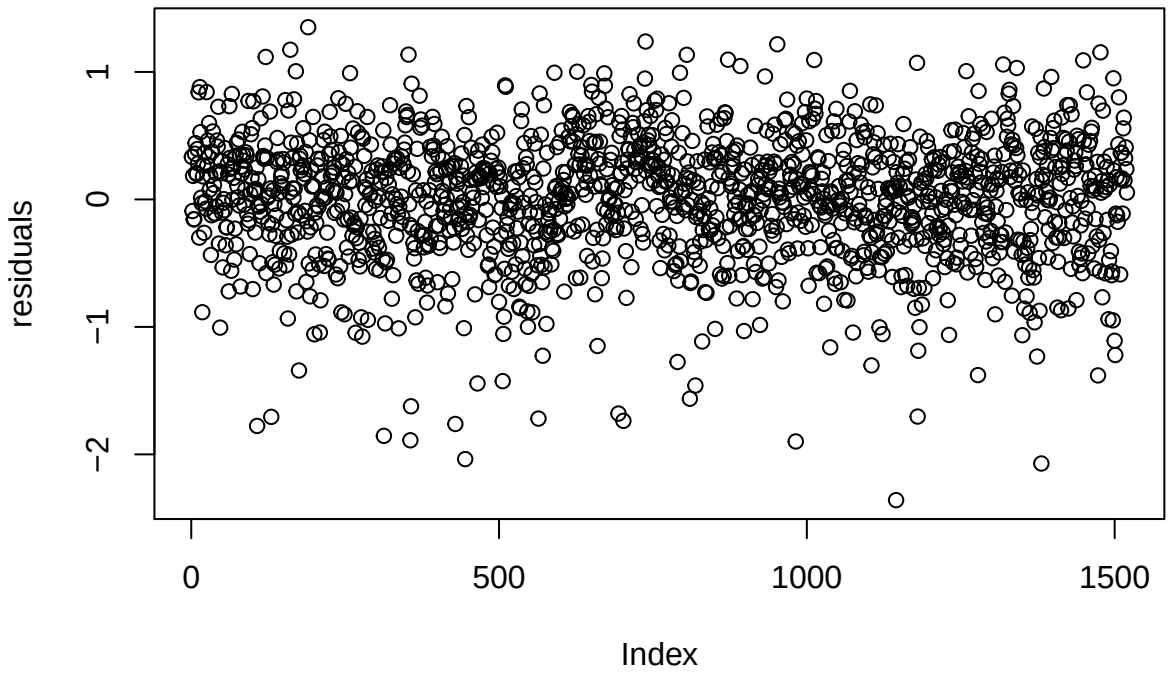
```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
library(plm)


kt <- read.dta13("kt_data.dta")
year10 <- subset(kt, year == 10)

regression <- lm(data = year10, wage ~ educ + abil + I(exper^2) + Fath_ed)

summary(regression)
```

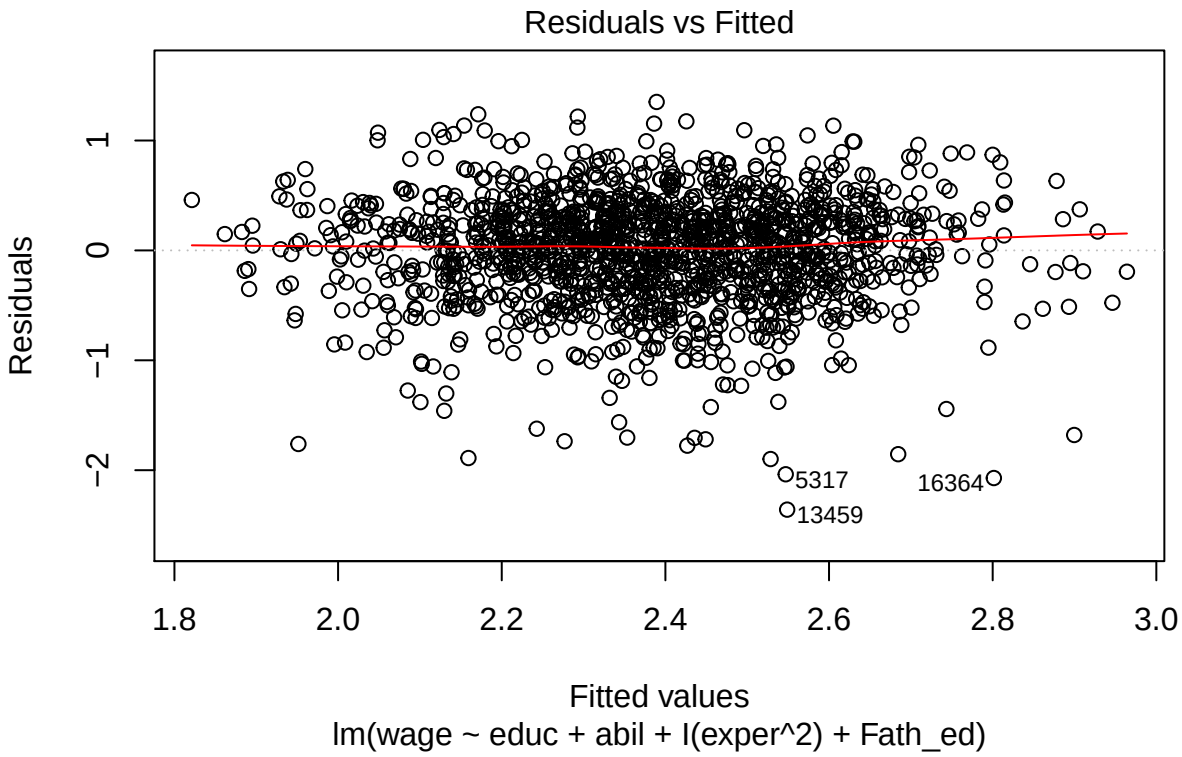
```
##
## Call:
## lm(formula = wage ~ educ + abil + I(exper^2) + Fath_ed, data = year10)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.35884 -0.27775  0.04841  0.31240  1.35068
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1304059  0.1271802   8.888  < 2e-16 ***
## educ        0.0743388  0.0083968   8.853  < 2e-16 ***
## abil        0.0907044  0.0169678   5.346 1.04e-07 ***
## I(exper^2)  0.0018780  0.0002808   6.689 3.15e-11 ***
## Fath_ed     0.0082859  0.0036598   2.264   0.0237 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

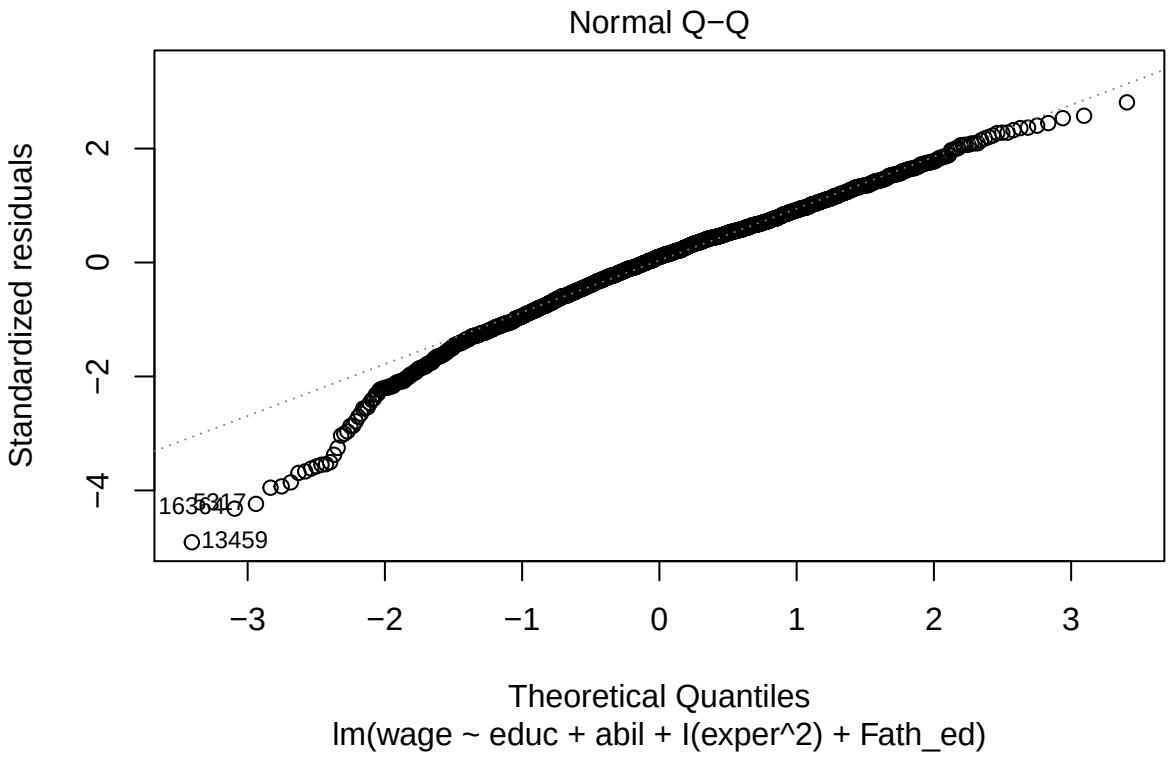
```
## Residual standard error: 0.4815 on 1515 degrees of freedom
## Multiple R-squared:  0.1323, Adjusted R-squared:   0.13
## F-statistic: 57.73 on 4 and 1515 DF,  p-value: < 2.2e-16
```

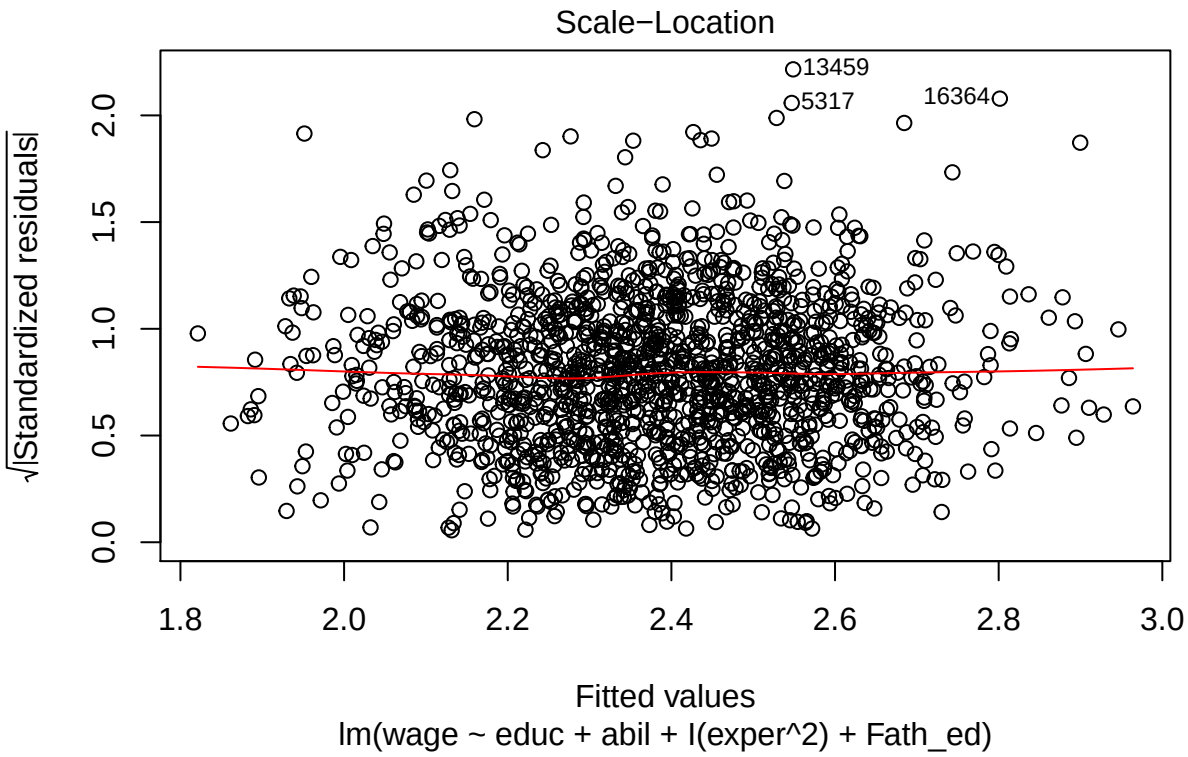
```
residuals <- resid(regression)
plot(residuals)
```

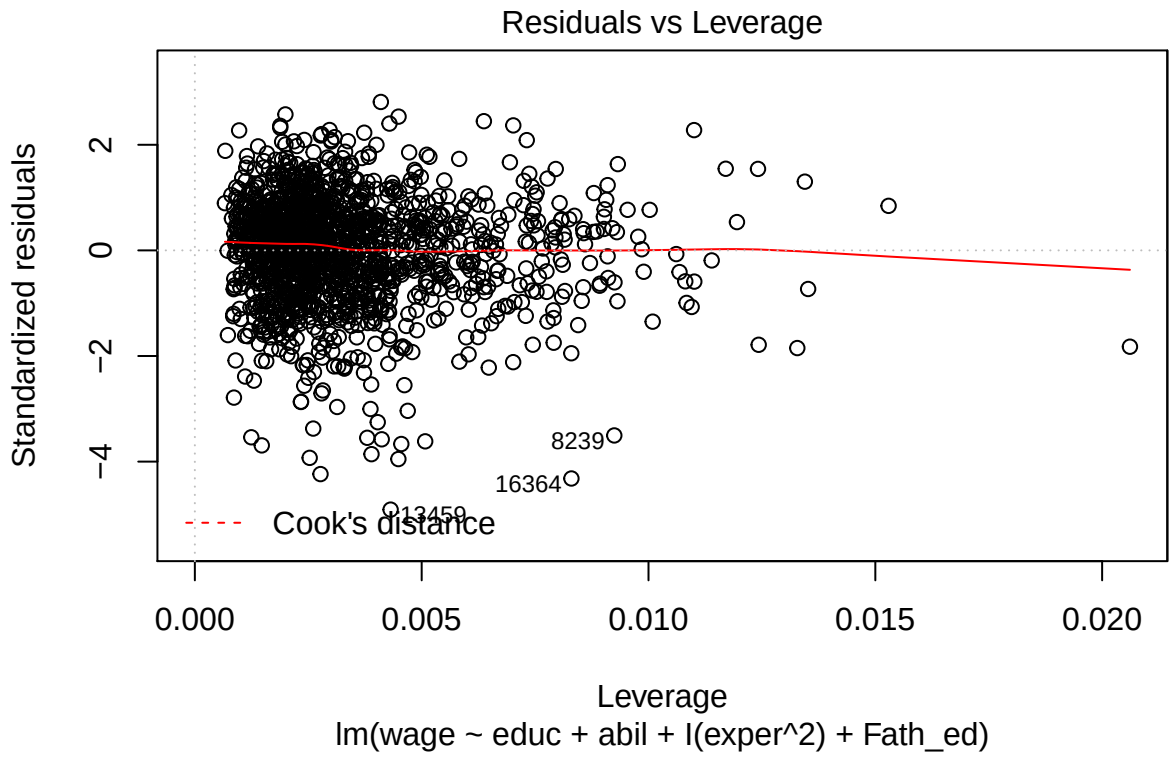



```
plot(regression)
```


Residuals vs Fitted
Residuals
1
0
−1
−2
5317
16364
13459
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Fitted values
lm(wage ~ educ + abil + I(exper^2) + Fath_ed)


Normal Q−Q
Standardized residuals
2
0
−2
−4
16364
53117
13459
−3
−2
−1
0
1
2
3
Theoretical Quantiles
lm(wage ~ educ + abil + I(exper^2) + Fath_ed)


Scale−Location
13459
5317
16364
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Fitted values
lm(wage ~ educ + abil + I(exper^2) + Fath_ed)




```
bptest(regression)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  regression
## BP = 9.046, df = 4, p-value = 0.05996
```

```
## Non-robust Errors
summary(regression)
```

```
##
## Call:
## lm(formula = wage ~ educ + abil + I(exper^2) + Fath_ed, data = year10)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.35884 -0.27775  0.04841  0.31240  1.35068
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1304059  0.1271802   8.888  < 2e-16 ***
## educ        0.0743388  0.0083968   8.853  < 2e-16 ***
## abil        0.0907044  0.0169678   5.346 1.04e-07 ***
## I(exper^2)  0.0018780  0.0002808   6.689 3.15e-11 ***
## Fath_ed     0.0082859  0.0036598   2.264   0.0237 *
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4815 on 1515 degrees of freedom
## Multiple R-squared:  0.1323, Adjusted R-squared:   0.13
## F-statistic: 57.73 on 4 and 1515 DF,  p-value: < 2.2e-16
```

```
## Robust HC1 Errors
coeftest(regression, vcov. = vcovHC(regression, type = "HC1"))
```

```
##
## t test of coefficients:
##
##               Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 1.13040589 0.13409134  8.4301 < 2.2e-16 ***
## educ        0.07433879 0.00946234  7.8563 7.444e-15 ***
## abil        0.09070444 0.01694254  5.3536 9.946e-08 ***
## I(exper^2)  0.00187805 0.00028551  6.5778 6.556e-11 ***
## Fath_ed     0.00828590 0.00366331  2.2619   0.02385 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Robust HC3 Errors
coeftest(regression, vcov. = vcovHC(regression, type = "HC3"))
```

```
##
## t test of coefficients:
##
##               Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 1.13040589 0.13455699  8.4009 < 2.2e-16 ***
## educ        0.07433879 0.00950231  7.8232 9.586e-15 ***
## abil        0.09070444 0.01699787  5.3362 1.093e-07 ***
## I(exper^2)  0.00187805 0.00028625  6.5608 7.327e-11 ***
## Fath_ed     0.00828590 0.00367958  2.2519   0.02447 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
clustered <- plm(data = year10, wage ~ educ + abil + I(exper^2) + Fath_ed, model = "pooling", index = c

# clustered SE method by Richard Bluhm
# compute State like df-adjustment
G <- length(unique(year10$loc))
N <- length(year10$loc)
dfa <- (G/(G - 1)) * (N - 1)/clustered$df.residual

#display with VCE and df-adjustment
loc_c_vcov <- dfa * vcovHC(clustered, type = "HC0", cluster = "group")
coeftest(clustered, vcov. = loc_c_vcov)
```

```
##
## t test of coefficients:
##
##               Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 1.13040589 0.09182886 12.3099 < 2.2e-16 ***
## educ        0.07433879 0.00830324  8.9530 < 2.2e-16 ***
## abil        0.09070444 0.01710717  5.3021 1.314e-07 ***
## I(exper^2)  0.00187805 0.00037938  4.9503 8.238e-07 ***
```

```
## Fath_ed     0.00828590 0.00417019  1.9869   0.04711 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```