Econ Assignment - 7

Piyush Kulkarni
pxk161130

Q 8.4 a) Graph of residuals of income suggests that as income increases, the error variance increases → hetroskedasticity. Graph of age residuals doesn't seem to have any pattern. error here is not varying with age

c)

i) Vacation miles traveled increases as income & age increases → +ve correlation
vacation miles traveled decreases as no. of kids increases → -ve correlation

ii) Both std errors are approximately same. White's std errors are a little bit higher for variables than that of least square. we can say that white's std errors will give us wider confidence interval compared to least square. All are significat at 5% in both models.

iii) Genralized least squares to have smallest std errors of all. therefor they will have a narrow confidence interval of all. they will give most precise estimates of all. ∴ They are ~~the~~ better than others in this case.

8.12 a) Rich countries having more money have multiple options to spend their money. they can spend only on basic education & also can spend more than basic eduction. → more variability

Countries with less GDP have very less money so they can not afford spending a lot of money on education. they can spend only on basic. → less variability

b) $y = \frac{EE}{P}$ $x = \frac{GDP}{P}$

$\hat{y} = -1.24 + 0.073\,x$

plots shows that there is hetroskedasticity

c) After plotting the models

$\chi^2 = N \times R^2 = 34 \times 0.293$

$= 9.962$

$\chi^2_{(0.95,2)} = 5.991$

$\chi^2 > \chi^2_{critical}$

H_0: $\alpha_2 = \alpha_3 = 0$

H_1: at least one of $\alpha_1, \alpha_2 \neq 0$

∴ We reject the null hypothesis

∴ Heteroskedasticity exists

d) test sq. std error for $b_2 = 0.0051$

Robust std error for $b_2 = 0.0063$

$0.0731 \pm 2.0369 \times 0.0051$

$= (0.0627, 0.0834)$

Robust

$0.0731 \pm 2.0369 \times 0.0063$

$(0.0602, 0.0859)$

Robust std error gives wider interval.

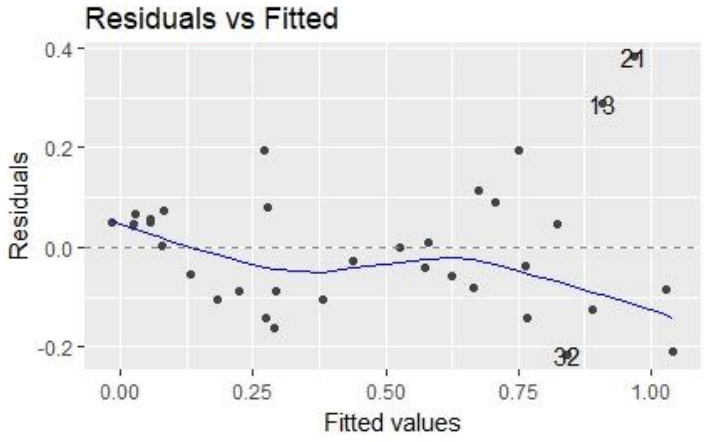
e) ~~$\hat{y} = 0.0929 + 0.069$~~.

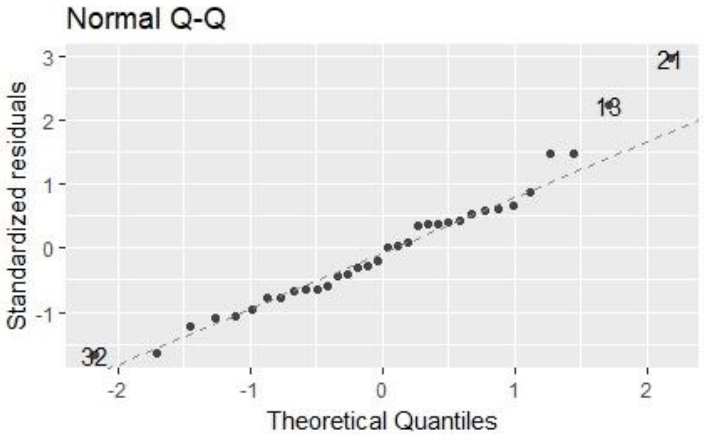
$\hat{y} = 0.0929 + 0.0693$ x

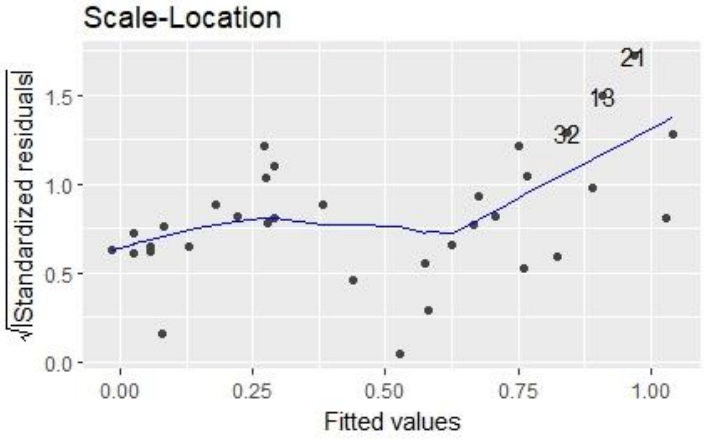
95% confint $= 0.069 \pm 2.036 \times 0.00412$

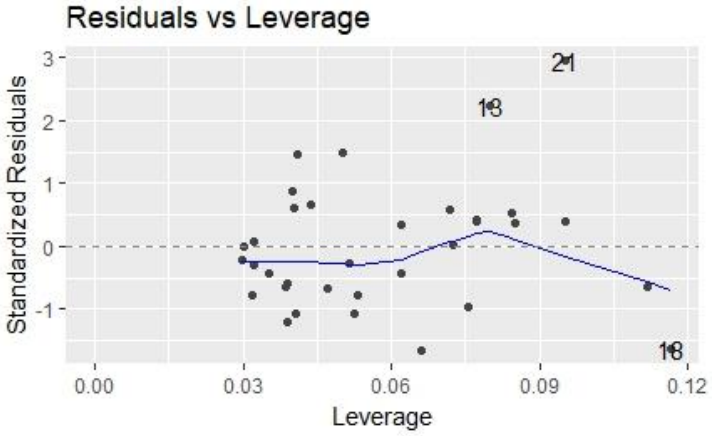
$= (0.0603, 0.07830)$

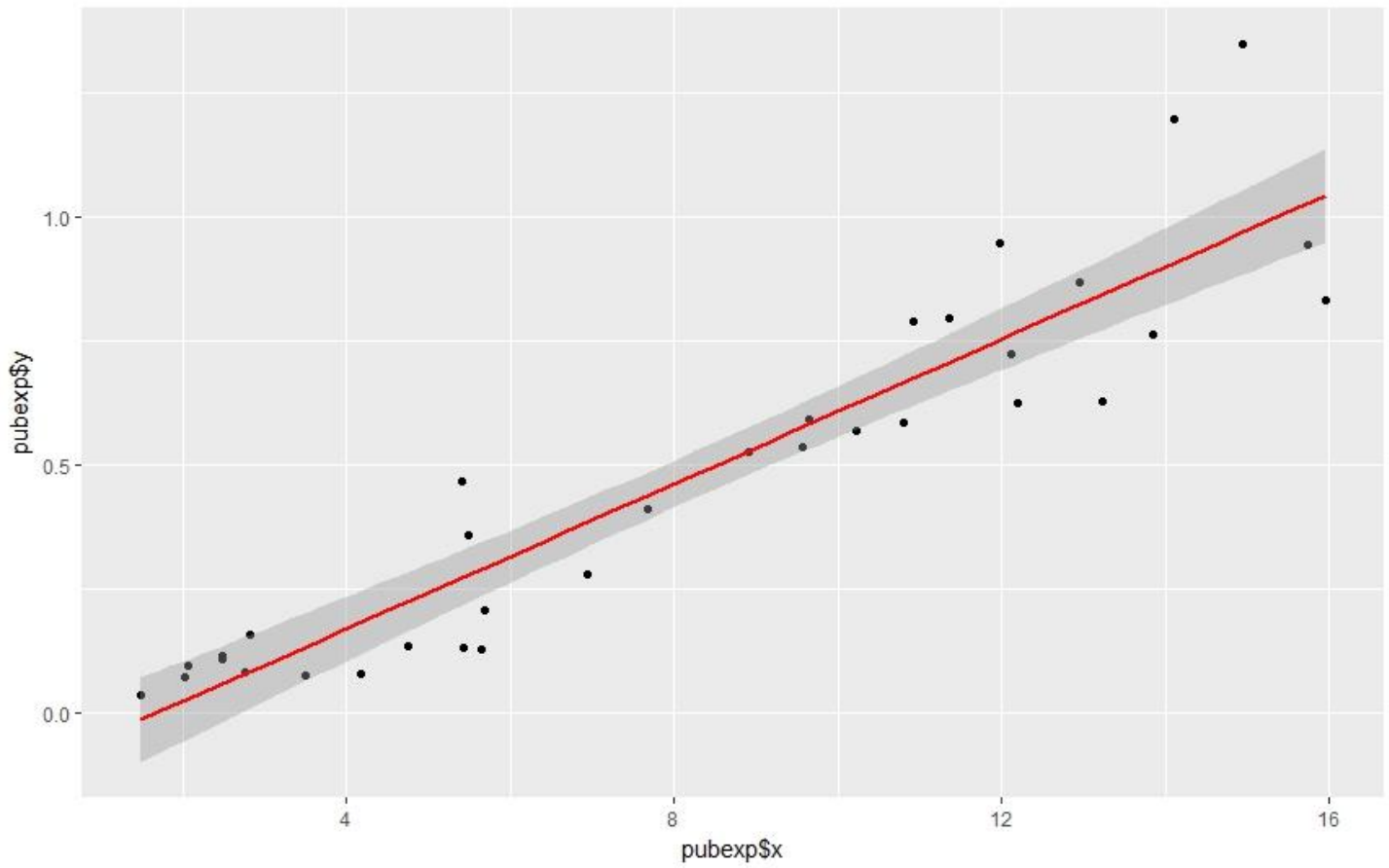
This confidence interval is narrower than both of the intervals obtained in part d.


Residuals vs Fitted
Residuals
Fitted values
0.4
0.2
0.0
-0.2
0.00
0.25
0.50
0.75
1.00
21
18
32


Normal Q-Q
Standardized residuals
Theoretical Quantiles
3
2
1
0
-1
-2
21
18
32


Scale-Location
√|Standardized residuals|
Fitted values
1.5
1.0
0.5
0.0
0.00
0.25
0.50
0.75
1.00
21
18
32


Residuals vs Leverage
Standardized Residuals
Leverage
3
2
1
0
-1
0.00
0.03
0.06
0.09
0.12
21
18
18



R code

```
> library(haven)
> pubexp <- read_dta("D:/Class Notes/Fall 17 Classes/ECON/Data_sets/pubexp.dta")
> View(pubexp)
> # creating y and x variables
> y <- pubexp$ee/pubexp$p
> x <- pubexp$gdp/pubexp$p
> x2 = x^2
> data1 <- cbind(pubexp, y)
> data <- cbind(data1, x)
>
> pubexp <- data
```

```
> # Running least square regression
>
> leastsq <- lm (y ~ x, data = pubexp)
> summary(leastsq)

Call:
lm(formula = y ~ x, data = pubexp)

Residuals:
    Min      1Q  Median      3Q     Max
-0.21682 -0.08804 -0.01401  0.06517  0.38156

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.124573   0.048523  -2.567   0.0151 *
x            0.073173   0.005179  14.128 2.65e-15 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1359 on 32 degrees of freedom
Multiple R-squared:  0.8618,      Adjusted R-squared:  0.8575
F-statistic: 199.6 on 1 and 32 DF,  p-value: 2.65e-15

> # getting residuals
> leastsq.res <- resid(leastsq)
> # Plotting residuals
>
> library(ggfortify)
> library(ggplot2)
> autoplot(leastsq, data=pubexp, color = 'blue')
>
> ggplot(pubexp, aes(x = pubexp$x, y = pubexp$y)) + geom_point() +
```

```
+   stat_smooth( method = "lm", col = 'red')
> # White formula
>
> res = residuals(leastsq)
> ressq = res^2
> #yhat = leastsq$fitted.values
> #yhat2 = yhat^2
>
> #combining all into one dataset
> #m4 = data.frame(cbind(ressq, yhat, yhat2))
>
> m3 = data.frame(cbind(ressq, x, x2))
>
> white <- lm(ressq ~ x + x2, data = m3)
> summary(white)

Call:
lm(formula = ressq ~ x + x2, data = m3)

Residuals:
     Min       1Q   Median       3Q      Max
-0.048773 -0.011263 -0.005239  0.005350  0.097664

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.0176770  0.0161120   1.097   0.2810
x           -0.0052062  0.0045479  -1.145   0.2611
x2           0.0004840  0.0002638   1.835   0.0762 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02513 on 31 degrees of freedom
```

```
Multiple R-squared:  0.293,        Adjusted R-squared:  0.2474
F-statistic: 6.423 on 2 and 31 DF,  p-value: 0.004636


>
> # critical chai-square
> qchisq(.95, 2)
[1] 5.991465
> #Robust std errors
> library(sandwich)
> #coef
>
> #coeftest(leastsq, vcov = vcovHC(leastsq,type = "HC1"))
>
> SE_robust <- sqrt(diag(vcovHC(leastsq, type="HC2")))
>
> model2 <- summary(leastsq)
> SE_robust
(Intercept)           x
0.040783211 0.006307339
>
> model2$coefficients[,2] <- SE_robust
> model2


Call:
lm(formula = y ~ x, data = pubexp)


Residuals:
    Min      1Q  Median      3Q     Max
-0.21682 -0.08804 -0.01401  0.06517  0.38156


Coefficients:
            Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept) -0.124573   0.040783  -2.567   0.0151 *
x           0.073173   0.006307  14.128 2.65e-15 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1359 on 32 degrees of freedom
Multiple R-squared:  0.8618,      Adjusted R-squared:  0.8575
F-statistic: 199.6 on 1 and 32 DF,  p-value: 2.65e-15

> qt(.975,32)
[1] 2.036933
> #Weightd least squares
>
> model <- lm(y ~ x, data = pubexp, weights = (1/pubexp$x))
> summary(model)

Call:
lm(formula = y ~ x, data = pubexp, weights = (1/pubexp$x))

Weighted Residuals:
     Min       1Q   Median       3Q      Max
-0.072028 -0.038561 -0.008488  0.027706  0.105415

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.092921   0.028904  -3.215  0.00298 **
x           0.069321   0.004412  15.713  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.04451 on 32 degrees of freedom
Multiple R-squared:  0.8853,      Adjusted R-squared:  0.8817
```

F-statistic: 246.9 on 1 and 32 DF, p-value: < 2.2e-16

>

> confint(model, 'x', level = 0.95)

      2.5 %   97.5 %

x 0.06033514 0.07830786