**Name: Khoa Tran**

**Assignment 8**



Text, letter

Description automatically generated

2.

Text, letter

Description automatically generated

3)

(1)

> #OLS model relating colGPA to hsGPA, ACT, skipped, and PC

> ols = lm(colGPA~hsGPA+ACT+skipped+PC,data=data)

> summary(ols)

Call:

lm(formula = colGPA ~ hsGPA + ACT + skipped + PC, data = data)

Residuals:

Min 1Q Median 3Q Max

-0.84006 -0.20392 -0.03352 0.25346 0.74558

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1.35651 0.32750 4.142 6.01e-05 \*\*\*

hsGPA 0.41295 0.09243 4.468 1.65e-05 \*\*\*

ACT 0.01334 0.01044 1.278 0.20353

skipped -0.07103 0.02625 -2.706 0.00768 \*\*

PC 0.12444 0.05731 2.171 0.03165 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3251 on 136 degrees of freedom

Multiple R-squared: 0.2593, Adjusted R-squared: 0.2375

F-statistic: 11.9 on 4 and 136 DF, p-value: 2.553e-08

(2)

> #Breusch-Pagan test for heteroskedasticity

> bptest(ols, data=data)

studentized Breusch-Pagan test

data: ols

BP = 11.874, df = 4, p-value = 0.01831

> #white's test for heteroskedasticity

> data$residualSQ<-ols$residuals^2

> data$fit<-ols$fitted.values

> # regression of the squared OLS residuals

> WhiteOLS<-lm(residualSQ ~fit+I(fit^2), data=data)

> # k+1: the number of explanatory variables + 1 (for intercept)

> Wkplus1<-nrow(summary(WhiteOLS)$coef)

> # obsn: the number of observations

> Wobsn<-WhiteOLS$df.residual+Wkplus1

> Wobsn

[1] 141

>

> # LM statistics for White test

> WhiteLM<-Wobsn\*summary(WhiteOLS)$r.squared

> WhiteLM

[1] 6.957569

>

> # Calculating p-value from Chi-squared distribution

> pchisq(WhiteLM, df=2, lower.tail=FALSE)

[1] 0.03084489

Discussion:

Based on BP test, we can reject the null hypothesis at 5% level -> the error is likely to be heteroskedastic at 5% level.

Based on the White’s test, we can also reject the null hypothesis at 5% level -> the error is likely to be heteroskedastic at 5% level.

Thus, it is likely that the error has linear relationship with our dependent variables.

(3)

> #Calculating robust standard error for ols

> robustse<-coeftest(ols, vcov = vcovHC(ols, "HC0"))

> robustse

t test of coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1.356509 0.336211 4.0347 9.075e-05 \*\*\*

hsGPA 0.412952 0.097018 4.2565 3.845e-05 \*\*\*

ACT 0.013344 0.010528 1.2675 0.207151

skipped -0.071034 0.025937 -2.7387 0.006996 \*\*

PC 0.124439 0.058913 2.1123 0.036492 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Discussion: While the standard error and the p-value change between the ols model and the robust estimation, it does not affect our conclusions regarding the statistical significances of the dependent variables in this case. Specifically, hsGPA and skipped are statistically significant at 1% level, PC is statistically significant at 5% level and ACT is statistically insignificant.

(4)

> #Calculating robust standard error for ols

> ## FWLS/FGLS

> #step 1. save u\_hat from the original OLS model

> u\_hat<-ols$residuals

> #step 2. generate log u\_hat

> lu\_hat\_sq<-log(u\_hat^2)

> #step 3. regress log u\_hat on x

> ols2<-lm(lu\_hat\_sq~ hsGPA+ACT+skipped+PC, data=data)

> fv<-ols2$fitted.values

> #step 4. exponentiate the fitted values

> efv<-exp(fv)

> #step 5. WLS

> fwls<-lm(colGPA~hsGPA+ACT+skipped+PC, weights=1/(efv), data=data)

> summary(fwls)

Call:

lm(formula = colGPA ~ hsGPA + ACT + skipped + PC, data = data,

weights = 1/(efv))

Weighted Residuals:

Min 1Q Median 3Q Max

-4.4444 -1.1274 -0.2533 1.1320 4.6901

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1.454363 0.287174 5.064 1.31e-06 \*\*\*

hsGPA 0.369504 0.076541 4.828 3.67e-06 \*\*\*

ACT 0.016098 0.009364 1.719 0.0879 .

skipped -0.085803 0.021318 -4.025 9.42e-05 \*\*\*

PC 0.124991 0.060211 2.076 0.0398 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.677 on 136 degrees of freedom

Multiple R-squared: 0.3005, Adjusted R-squared: 0.2799

F-statistic: 14.61 on 4 and 136 DF, p-value: 5.978e-10

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 1.35651 0.32750 4.142 6.01e-05 \*\*\*

hsGPA 0.41295 0.09243 4.468 1.65e-05 \*\*\*

ACT 0.01334 0.01044 1.278 0.20353

skipped -0.07103 0.02625 -2.706 0.00768 \*\*

PC 0.12444 0.05731 2.171 0.03165 \*

Discussion: The difference between the conclusion of the OLS estimation and the FGLS estimation is that under the FGLS, ACT is statistically significant at 10% level while in the OLS estimation ACT is statistically insignificant at the 10% level.