

Assignment 3

Q1: Heteroskedasticity is one of the assumption for MLR, and is used to find the OLS estimation. Without this, the β_0, β_1, \dots , etc. will not be efficient [$\text{Var}(U|X \dots X_k) = \sigma^2$]. Also, the explanatory variables should be included to make the OLS unbiased, in order to separate them with the error term. If our OLS has sample correlation coefficient of 0.95 between 2 independent variables, it does not wrong since it does not violate any MLR assumption

Q2: (i) Null hypothesis that, after controlling for sales and roe, ros has no effect on CEO salary

$$\rightarrow H_0: \beta_3 = 0$$

The alternative that better stock market performance increase CEO's salary

$$\rightarrow H_1: \beta_3 > 0$$

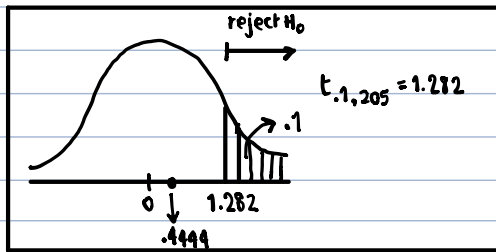
(ii) The proportionate affect of ros on salary if ros increase by 50 points is $.00024(50) = .012$. To get the percentage effect, we can just simply multiply it by 100 and we get 1.2%. Ros does not have a practically large effect on salary as a large change in ROS gives small effect on salary.

(iii) $H_0: \beta_3 = 0$ $d = .1$

$$H_1: \beta_3 > 0$$

$$n = 209, \text{ d.f.} = n - k - 1 = 209 - 3 - 1 = 205$$

$$\rightarrow \text{we can use z-table as d.f.} > 30 \Rightarrow z = \frac{\hat{\beta}_3 - \beta_3}{\text{s.e.} \hat{\beta}_3} = \frac{0.00024 - 0}{0.00054} = 0.4444$$



The z-value does not fall into the rejection region.

\rightarrow Does not reject H_0 at .1 level of significance.

\therefore ros has no effect on salary #

(iv) No, since we already test that ros does not have any effect on salary at .1 level of significance. Eventhough if we add, it will raise r^2 , but it will worsen the variance. We only want to put in what really explained the model #

C1. $\text{vote A} = \beta_0 + \beta_1 \log(\text{expand A}) + \beta_2 \log(\text{expand B}) + \beta_3 \text{prtystrA} + u$

Source	SS	df	MS	Number of obs =	F(3, 169) =	Prob > F =
Model	38405.1096	3	12801.7032	215.23	215.23	0.0000
Residual	10052.1389	169	59.480112		R-squared =	0.7926
Total	48457.2486	172	281.728189		adj R-squared =	0.7889
					Root MSE =	7.7123

voteA	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lexpandA	6.083316	.38215	15.92	0.000	5.328914 6.837719
lexpandB	-6.615417	.3788203	-17.46	0.000	-7.363246 -5.867588
prtystrA	.1519574	.0620181	2.45	0.015	.0295274 .2743873
_cons	45.07893	3.926305	11.48	0.000	37.32801 52.82985

i) $\beta_1 = 6.083316$ means an additional increase in $\log(\text{expand A})$ would increase vote A by 6.083316 units.

ii) $H_0 = \beta_2 = -\beta_1$, $\beta_2 + \beta_1 = 0$
 $H_a = \beta_2 \neq -\beta_1$, $\beta_2 + \beta_1 \neq 0$

iii) $\text{Vote A} = 45.07893 + 6.083316 \log(\text{expand A}) - 6.615417 \log(\text{expand B}) + 0.1519574 \text{prtystrA} + u$

\therefore A's expenditure has a positive effect on vote A. However, B's expenditure has a negative effect on Vote A. \therefore we can't use these result to test.

iv) $H_0: \beta_2 + \beta_1 = 0$

$H_a: \beta_2 + \beta_1 \neq 0$

Let $\hat{\beta}_2 + \hat{\beta}_1 = \hat{\theta} \rightarrow \beta_1 = \theta - \beta_2$
 From $t_{\beta_j} = \frac{\hat{\beta}_j - \beta_j}{\text{s.e.}(\hat{\beta}_j)} = \frac{(\hat{\beta}_2 + \hat{\beta}_1) - 0}{\text{s.e.}(\hat{\beta}_2 + \hat{\beta}_1)} = \frac{\hat{\theta} - 0}{\text{s.e.}(\hat{\theta})}$

Sub in the equation

$\text{Vote A} = \beta_0 + \beta_1 \log(\text{expand A}) + \beta_2 \log(\text{expand B}) + \beta_3 \text{prtystrA} + u$

$\text{Vote A} = \beta_0 + (\theta - \beta_2) \log(\text{expand A}) + \beta_2 \log(\text{expand B}) + \beta_3 \text{prtystrA} + u$

$= \beta_0 + \theta \log(\text{expand A}) - \beta_2 \log(\text{expand B}) + \beta_3 \text{prtystrA} + u$

$\beta_0 + \theta \log(\text{expand A}) - \beta_2 (\log(\text{expand A}) - \log(\text{expand B})) + \beta_3 \text{prtystrA} + u$

Do regression again

$t = \frac{\hat{\theta} - 0}{\text{s.e.}(\hat{\theta})} = \frac{0.532101}{0.5330858} = 0.99815$



Source	SS	df	MS	Number of obs =	F(3, 169) =	Prob > F =
Model	38405.1097	3	12801.7032	215.23	215.23	0.0000
Residual	10052.1388	169	59.4801115		R-squared =	0.7926
Total	48457.2486	172	281.728189		adj R-squared =	0.7889
					Root MSE =	7.7123

voteA	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lexpandA	6.083316	.38215	15.92	0.000	5.328914 6.837719
lexpandB	-6.615417	.3788203	-17.46	0.000	-7.363246 -5.867588
prtystrA	.1519574	.0620181	2.45	0.015	.0295274 .2743873
_cons	45.07893	3.926305	11.48	0.000	37.32801 52.82985

So, we do not reject H_0 b/c P-value is more than 0.005 (significant level)

\therefore Increase in an additional unit in $-\beta_2$ would cause an additional negative value to β_1 and vice versa

C6. i) $\log(\text{wage}) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{exper} + \beta_3 \text{tenure} + u$
 $H_0: \beta_2 = \beta_3$

from $H_0: \beta_2 = \beta_3, \beta_2 - \beta_3 = 0$
 $H_1: \beta_2 \neq \beta_3, \beta_2 - \beta_3 \neq 0$
 Let $\theta = \hat{\beta}_2 - \hat{\beta}_3 \rightarrow \beta_2 = \theta + \beta_3$

Source	SS	df	MS	Number of obs	F(3, 931)	Prob > F
Model	25.6953242	3	8.56510806	935	56.97	0.0000
Residual	139.960959	931	.150334005		R-squared	0.1551
Total	165.656283	934	.177362188		Adj R-squared	0.1524
					Root MSE	.38773

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
educ	.0748638	.0065124	11.50	0.000	-.062083 .0876446
exper	.0019537	.0047434	0.41	0.681	-.0073554 .0112627
expten	.0133748	.0025872	5.17	0.000	.0082974 .0184522
_cons	5.496696	.1105282	49.73	0.000	5.279782 5.713609

sub in equation, we get

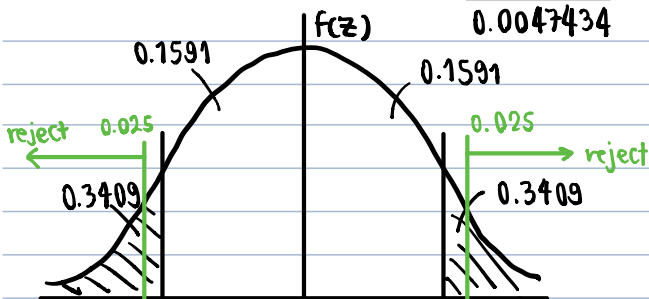
$$\begin{aligned} \log(\text{wage}) &= \beta_0 + \beta_1 \text{educ} + \beta_2 \text{exper} + \beta_3 \text{tenure} + u \\ \log(\text{wage}) &= \beta_0 + \beta_2 \text{educ} + (\theta + \beta_3) \text{exper} + \beta_3 \text{tenure} + u \\ &= \beta_0 + \beta_2 \text{educ} + \theta \text{exper} + \beta_3 \text{exper} + \beta_3 \text{tenure} + u \\ &= \beta_0 + \beta_2 \text{educ} + \theta \text{exper} + \beta_3 (\text{exper} + \text{tenure}) + u \end{aligned}$$

new V

Source	SS	df	MS	Number of obs	F(3, 931)	Prob > F
Model	25.6953242	3	8.56510806	935	56.97	0.0000
Residual	139.960959	931	.150334005		R-squared	0.1551
Total	165.656283	934	.177362188		Adj R-squared	0.1524
					Root MSE	.38773

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
educ	.0748638	.0065124	11.50	0.000	-.062083 .0876446
exper	.0019537	.0047434	0.41	0.681	-.0073554 .0112627
newV	.0133748	.0025872	5.17	0.000	.0082974 .0184522
_cons	5.496696	.1105282	49.73	0.000	5.279782 5.713609

from $t_{\theta_j} = \frac{\hat{\beta}_i - \beta_j}{\text{s.e. } \hat{\beta}_i} = \frac{\hat{\theta} - 0}{\text{s.e. } \hat{\theta}}$
 $= \frac{0.0019537}{0.0047434} = 0.41178 \approx 0.41$



Since, P-value of $0.41 = 0.3409$ is greater than significant level of 5% (0.025). Therefore, we can conclude that we will not reject null hypothesis cause P-value is not fall into rejection area.

C8.

```

> use "C:\Users\6104641177\AppData\Local\Temp\Temp1_130527010X_522192.zip\Data Sets- STATA\401ksub
> s.dta"
. sum nettfa if fsize == 1
. do "C:\Users\610464-1\AppData\Local\Temp\STD1474_000000.tmp"
. reg nettfa inc age if fsize == 1

```

Variable	Obs	Mean	Std. Dev.	Min	Max
nettfa	2,017	13.59498	47.59058	-143.5	1134.098

```

. reg nettfa inc age if fsize == 1

```

Source	SS	df	MS	Number of obs	F(2, 2014)	Prob > F	R-squared	Adj R-squared	Root MSE
Model	544916.989	2	272458.495	2,017	136.46	0.0000	0.1193	0.1185	44.683
Residual	4021048.06	2,014	1996.54819						
Total	4565965.05	2,016	2264.86361						

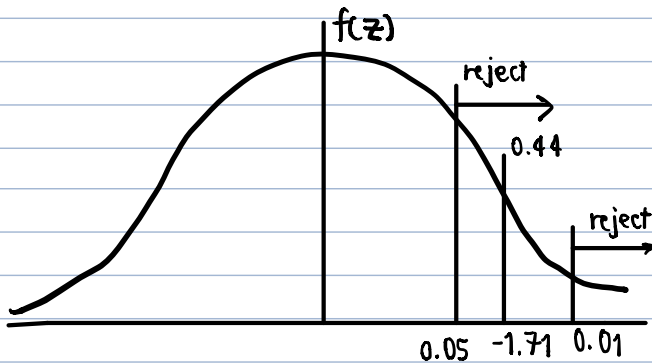
	nettfa	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
β_1 inc		.7993167	.0597307	13.38	0.000	.6821762 .9164572
β_2 age		.8426563	.0920169	9.16	0.000	.6621982 1.023115
β_0 _cons		-43.03981	4.080393	-10.55	0.000	-51.04204 -35.03758

i) 2017

ii) $\hat{\beta}_1$ means a \$1 thousand dollars increase in income corresponds to \$799 increase in nettfa. $\hat{\beta}_2$ as 1 year increase in age corresponds to a \$842 increase in net total financial wealth.

iii) $\beta_0 = -43.04$ \therefore when their income and age equal to 0, so this is the net financial wealth of newborn babies.

iv) $t_{\beta_2} = \frac{\hat{\beta}_2 - \beta_2}{s.e. \hat{\beta}_2} = \frac{(0.843 - 1)}{0.092017} = -1.71$ \therefore The P-value is $P(T < -1.71) \approx 0.044 \rightarrow$ from the table



\therefore We can reject the null hypothesis at 5% level significant but not 1% significant

```

. reg nettfa inc fsize == 1
== invalid name
r(198);
end of do-file
r(198);
. reg nettfa inc if fsize ==1

```

Source	SS	df	MS	Number of obs	F(1, 2015)	Prob > F	R-squared	Adj R-squared	Root MSE
Model	377482.064	1	377482.064	2,017	181.60	0.0000	0.0827	0.0822	45.592
Residual	4188482.98	2,015	2078.6516						
Total	4565965.05	2,016	2264.86361						

	nettfa	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
inc		.8206815	.0609	13.48	0.000	.7012479 .940115
_cons		-10.57095	2.060678	-5.13	0.000	-14.61223 -6.529671

\rightarrow we will get coef of $\hat{\beta}_1 = 0.8206815$ which is not so different from $\hat{\beta}_1 = 0.799$ in the previous regression fn. Since, there is an omitted variable in this regression fn. which we found to be 0.039. That's why the coefficient does not change much.