EE325 Homework 4

6104640138 Jidapa Chandhasrisawad

① Heteroskedasticity (i) and omitting an important explanatory variable (iii) can cause the usual OLS t statistics to be invalid because it is breaking the assumptions on the multiple regression at the beginning. Therefore, the OLS estimation method cannot be used.

For part (ii), it is acceptable to have a correlation between two independent variable in the model in order to generate regression function by using OLS method and calculate t statistics.

② $\log(salary) = \beta_0 + \beta_1 \log(sales) + \beta_2 roe + \beta_3 ros + u$

(i) $H_0: \hat{\beta}_2 = 0$ and $\hat{\beta}_3 = 0$

H_a: otherwise

(ii)

$$\widehat{\log(salary)} = \underset{(.32)}{4.32} + \underset{(.035)}{.280} \log(sales) + \underset{(.0041)}{.0174}\, roe + \underset{(.00054)}{.00024}\, ros$$

$n = 209$, $R^2 = .283$.

$$\frac{d \log(salary)}{d\, ros} = 0.00024$$

$$100\frac{\left(\frac{1}{salary}\, d\, salary\right)}{d\, ros} = 100(0.00024) = 0.024$$

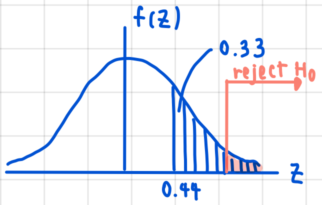
$\%\Delta$ in salary $= 0.024\ \%$ increase

(iii) $H_0: \beta_3 = 0$

$H_a: \beta_3 > 0$

d.f. $= 209 - 3 - 1 = 205$

$$z = \frac{0.00024 - 0}{0.00054} = 0.44 \rightarrow \text{p-value} = 0.5 - 0.17 = 0.33$$



Since p-value is greater than level of significance, we accept H_0 at 10 % significant level. Therefore, ros has no effect on salary.

(iv) From (iii), we should not include ros in a final model explaining CEO compensation in terms of firm performance.

C1 (i) β_1 is the impact of log(expend A) on vote A.

(ii) $\dfrac{d\log(\text{expend}A)}{d\,\text{vote}A} = \beta_1$ $\qquad$ $\dfrac{d\log(\text{expend}B)}{d\,\text{vote}A} = \beta_2$

$100\left(\dfrac{\frac{1}{\text{expend}A}\,d\,\text{expend}A}{d\,\text{vote}A}\right) = (\beta_1)\,100$ $\qquad$ $100\left(\dfrac{\frac{1}{\text{expend}B}\,d(\text{expend}B)}{d\,\text{vote}A}\right) = (\beta_2)\,100$

$\%\Delta$ in expend A $= 100\beta_1$ $\qquad$ $\%\Delta$ in expend B $= 100\beta_2$

$H_0: 100\hat{\beta}_1 = -100\hat{\beta}_2$
$H_a: 100\hat{\beta}_1 \neq -100\hat{\beta}_2$
$\Rightarrow$
$H_0: \hat{\beta}_1 + \hat{\beta}_2 = 0$
$H_a: \hat{\beta}_1 + \hat{\beta}_2 \neq 0$

(iii)

```
regress voteA lexpendA lexpendB prtystrA
```

Source	SS	df	MS			
Model	38405.1096	3	12801.7032	Number of obs	=	173
Residual	10052.1389	169	59.480112	F(3, 169)	=	215.23
				Prob > F	=	0.0000
				R-squared	=	0.7926
				Adj R-squared	=	0.7889
Total	48457.2486	172	281.728189	Root MSE	=	7.7123

voteA	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lexpendA	6.083316	.38215	15.92	0.000	5.328914	6.837719
lexpendB	-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

$\text{vote}A = 6.083\log(\text{expend}A) - 6.615\log(\text{expend}B) + 0.152\,\text{prtystr}A + u$

A's expenditure and B's expenditure affect the outcome which is the percentage of the vote received by candidate A. While A's expenditure has a positive impact, B's expenditure has a negative impact on vote A.

We cannot use this result to test the hypothesis in part (ii) because we need the standard error of $\beta_1 + \beta_2$.

(iv) $\theta_1 = \beta_1 + \beta_2$
$\beta_1 = \theta_1 - \beta_2$

$\text{vote}A = \beta_0 + (\theta_1 - \beta_2)\log(\text{expend}A) + \beta_2\log(\text{expend}B) + \beta_3\,\text{prtystr}A + u$
$\text{vote}A = \beta_0 + \theta_1\log(\text{expend}A) + \beta_2(\log(\text{expend}B) - \log(\text{expend}A)) + \beta_3\,\text{prtystr}A + u$

```
. regress voteA lexpendA diff_lexB_lexA prtystrA
```

Source	SS	df	MS			
Model	38405.1097	3	12801.7032	Number of obs	=	173
Residual	10052.1388	169	59.4801115	F(3, 169)	=	215.23
				Prob > F	=	0.0000
				R-squared	=	0.7926
				Adj R-squared	=	0.7889
Total	48457.2486	172	281.728189	Root MSE	=	7.7123

voteA	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lexpendA	-.532101	.5330858	-1.00	0.320	-1.584466	.5202638
diff_lexB_lexA	-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

$\hat{\theta}_1 = -0.532$

Since p-value of $\hat{\theta}_1$ is equal to 0.320 which is more than 0.05 which is 5% level of significant, we accept H_0 at 5% significant level in this case. Therefore, one percent increase in A's expenditure is not offset by one percent increase in B's expenditure.

C6 (i) $H_0: \beta_2 - \beta_3 = 0$

$H_a: \beta_2 - \beta_3 \neq 0$

(ii) d.f. $= 935 - 3 - 1 = 931$

$$z = \frac{(\hat{\beta}_2 - \hat{\beta}_3) - 0}{s.e.(\hat{\beta}_2 - \hat{\beta}_3)}$$

Let $\hat{\theta}_1 = \hat{\beta}_2 - \hat{\beta}_3$

$H_0: \hat{\theta}_1 = 0$

$H_a: \hat{\theta}_1 \neq 0$

$\hat{\beta}_2 = \hat{\theta}_1 + \hat{\beta}_3$

$$\log(wage) = \beta_0 + \beta_1 educ + (\theta_1 + \beta_3) exper + \beta_3 tenure + u$$

$$\log(wage) = \beta_0 + \beta_1 educ + \theta_1 exper + \beta_3 (exper + tenure) + u$$

. regress lwage educ exper combined_exper_tenure

Source	SS	df	MS
Model	25.6953242	3	8.56510806
Residual	139.960959	931	.150334005
Total	165.656283	934	.177362188

Number of obs	=	935
F(3, 931)	=	56.97
Prob > F	=	0.0000
R-squared	=	0.1551
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	$\hat{\theta}_1$ .0019537	.0047434	0.41	0.681	-.0073554	.0112627
combined_exper_tenure	.0133748	.0025872	5.17	0.000	.0082974	.0184522
_cons	5.496696	.1105282	49.73	0.000	5.279782	5.713609

$\hat{\theta}_1 = 0.00195$, $s.e.(\hat{\theta}_1) = 0.00474$

Since p-value of $\hat{\theta}_1$ is 0.681 greater than 0.05 level of significant, we accept H_0 at 5% significant level. So, year of experience and year of tenure do not generate the same effect on log(wage).

The 95% CI for $\hat{\theta}_1 = [0.00195 - 2.06(0.00474), 0.00195 + 2.06(0.00474)]$

$= [-0.0078, 0.0117]$

The value between -0.0078 and 0.0117 captures 95% chance.

C8 (i) 2,017 single-person households

(ii)

```
regress nettfa inc age if fsize == 1
```

Source	SS	df	MS
Model	544916.989	2	272458.495
Residual	4021048.06	2,014	1996.54819
Total	4565965.05	2,016	2264.86361

Number of obs	=	2,017
F(2, 2014)	=	136.46
Prob > F	=	0.0000
R-squared	=	0.1193
Adj R-squared	=	0.1185
Root MSE	=	44.683

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

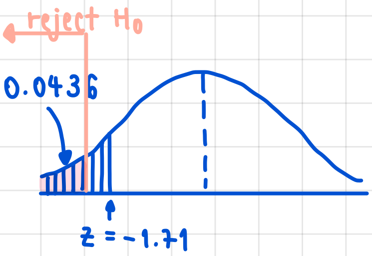
Slope coefficient of family income is 0.7993 which can be interpreted that when family income increases by one, net financial wealth will increase by 0.7993. Slope coefficient of age of the survey respondent is 0.8427 which means that as respondent age increases by one, it affects net financial wealth to be increase by 0.8427.

(iii) The intercept of this regression function is -43.0398 meaning that as age and income are zero, the net financial wealth will be negative number

(iv) $H_0 : \beta_2 = 1$

$H_a : \beta_2 < 1$

$d.f. = 2017 - 2 - 1 = 2014$, $z = \dfrac{0.8427 - 1}{0.092} = -1.71$



P-value = 0.5 - 0.4564 = 0.0436

Since p-value is greater than significance level, we accept H_0 at 1% level of significance. Therefore, one unit change in age causes one unit change in nettfa.

(v.)

```
regress nettfa inc if fsize == 1
```

Source	SS	df	MS
Model	377482.064	1	377482.064
Residual	4188482.98	2,015	2078.6516
Total	4565965.05	2,016	2264.86361

Number of obs	=	2,017
F(1, 2015)	=	181.60
Prob > F	=	0.0000
R-squared	=	0.0827
Adj R-squared	=	0.0822
Root MSE	=	45.592

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

The results are not much different because it is omitted only one variable which is age.