Tanaphat C. 6104649990

1. Which of the following can cause the usual OLS *t* statistics to be invalid (that is, not to have *t* distributions under $\mathbf{H}_0$)?

   i. Heteroskedasticity.

   ii. A sample correlation coefficient of .95 between two independent variables that are in the model.

   iii. Omitting an important explanatory variable.

Ans : (i) and (iii) cause the t-statistics not to have a distribution under H_0
Homoskedasticity is one of the classical linear model assump. OLS doesn't work
(iii) important omitted variable violates assumption MLR3

2. Consider an equation to explain salaries of CEOs in terms of annual firm sales, return on equity (*roe*, in percentage form), and return on the firm's stock (*ros*, in percentage form):

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

   i. In terms of the model parameters, state the null hypothesis that, after controlling for *sales* and *roe*, *ros* has no effect on CEO salary. State the alternative that better stock market performance increases a CEO's salary.

$H_0 : \beta_3 = 0$
$H_a : \beta_3 > 0$

   ii. Using the data in CEOSAL1, the following equation was obtained by OLS:

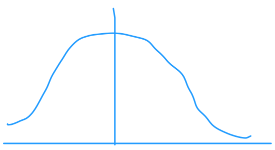
$$\widehat{\log(salary)} = \underset{(.32)}{4.32} + \underset{(.035)}{.280} \log(sales) + \underset{(.0041)}{.0174}\, roe + \underset{(.00054)}{.00024}\, ros$$
$$n = 209,\ R^2 = .283.$$

By what percentage is *salary* predicted to increase if *ros* increases by 50 points? Does *ros* have a practically large effect on *salary*?

the proportionate effect on salarry = 0.00024 (50) = .012 = 1.2%
Therefore, a 50 point ceteris paribus increase in ros is predicted to rise salary by 1.2%

   iii. Test the null hypothesis that *ros* has no effect on *salary* against the alternative that *ros* has a positive effect. Carry out the test at the 10% significance level.

$H_0 : \beta_3 = 0$
$H_a : \beta_3 > 0$



sig. level 10% = 0.1
d.f. = 209 - 3 - 1 = 205 > 30 use Z-scores
$t_{cri} = 1.282$
$t_{cal} = \dfrac{\hat{\beta}_3 - 0}{SE(\hat{\beta}_3)} = .44$
$0.44 < 1.282$
we can't reject H_0 at 10% sig level

C1. The following model can be used to study whether campaign expenditures affect election outcomes:

$$voteA = \beta_0 + \beta_1 \log(expendA) + \beta_2 \log(expendB) + \beta_3 prtystrA + u,$$

where *voteA* is the percentage of the vote received by Candidate A, *expendA* and *expendB* are campaign expenditures by Candidates A and B, and *prtystrA* is a measure of party strength for Candidate A (the percentage of the most recent presidential vote that went to A's party).

   i. What is the interpretation of β_1 ?

$\Delta vote A = \beta_1 \Delta \log(Expand A)$
$= (\beta_1/100)\,[100 \times \Delta \log(Expand A)]$
$\approx (\beta_1/100)\,[\% \Delta \log(Expand A)]$

$\therefore \beta_1/100$ is ceteris paribus percentage point change of vote recieved when campaign expenditure by candidate A increases by 1%

ii. In terms of the parameters, state the null hypothesis that a 1% increase in A's expenditures is offset by a 1% increase in B's expenditures.

$H_0: \beta_2 = -\beta_1$

$H_1: \beta_2 \neq -\beta_1$

iii. Estimate the given model using the data in VOTE1 and report the results in usual form. Do A's expenditures affect the outcome? What about B's expenditures? Can you use these results to test the hypothesis in part (ii)?

```
. reg voteA lexpendA lexpendB prtystrA

      Source |       SS           df       MS      Number of obs   =       173
-------------+----------------------------------   F(3, 169)       =    215.23
       Model |  38405.1096         3  12801.7032   Prob > F        =    0.0000
    Residual |  10052.1389       169   59.480112   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
------------------------------------------------------------------------------
```

regression model in usual form:

$\text{Vote}A = (45.1) + 6.08 \log(\text{Expand}\,A) - 6.62 \log(\text{Expand}\,B) + 0.152\,(\text{prtystr}A)$

v. Estimate a model that directly gives the *t* statistic for testing the hypothesis in part (ii). What do you conclude? (Use a two-sided alternative.)

rewrite hypothesis $\theta = \beta_1 + \beta_2 \Rightarrow H_0: \theta_1 = 0 \quad H_a: \theta_1 \neq 0$

rearrange equation $\widehat{\text{vote}}\,A = \beta_0 + \theta \log(\text{Expand}A) + \beta_2 [\log(\text{Expand}B - \log(\text{Expand}A)] + \beta_3\, \text{prtystr}A$

when estimalimate equation we obtain $\hat{\beta}_1 \approx .533$

the tcal $= \dfrac{-.532 - 0}{.533} \approx -1 \;\therefore$ can't reject $H_0: \beta_2 = -\beta_1$

C6. Use the data in WAGE2 for this exercise.

i. Consider the standard wage equation

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure + u.$$

State the null hypothesis that another year of general workforce experience has the same effect on $\log(wage)$ as another year of tenure with the current employer.

ii. Test the null hypothesis in part (i) against a two-sided alternative, at the 5% significance level, by constructing a 95% confidence interval. What do you conclude?

```
. reg lwage educ exper tenure

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

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0748638   .0065124    11.50   0.000      .062083    .0876446
       exper |   .0153285   .0033696     4.55   0.000     .0087156    .0219413
      tenure |   .0133748   .0025872     5.17   0.000     .0082974    .0184522
       _cons |   5.496696   .1105282    49.73   0.000     5.279782    5.713609
------------------------------------------------------------------------------
```

C8. The data set 401KSUBS contains information on net financial wealth (*nettfa*), age of the survey respondent (*age*), annual family income (*inc*), family size (*fsize*), and participation in certain pension plans for people in the United States. The wealth and income variables are both recorded in thousands of dollars. For this question, use only the data for single-person households (so $fsize = 1$).

i. How many single-person households are there in the data set?

ii. Use OLS to estimate the model

$$nettfa = \beta_0 + \beta_1 inc + \beta_2 age + u,$$

and report the results using the usual format. Be sure to use only the single-person households in the sample. Interpret the slope coefficients. Are there any surprises in the slope estimates?

iii. Does the intercept from the regression in part (ii) have an interesting meaning? Explain.

iv. Find the *p*-value for the test $\mathbf{H_0}: \beta_2 = 1$ against $\mathbf{H_1}: \beta_2 < 1$. Do you reject $\mathbf{H_0}$ at the 1% significance level?

v. If you do a simple regression of *nettfa* on *inc*, is the estimated coefficient on *inc* much different from the estimate in part (ii)? Why or why not?

```
. reg nettfa inc age if fsize ==1
```

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