Problem set 3 solutions

1. Answer:

(a) False. R^2 never goes down when additional variables are added; it always increases or at worst stays the same. It is not informative about whether variables belong in the regression.

(b) False. OLS residuals are always uncorrelated with explanatory variables x_i. This tells us nothing about the correlation between x_i and the error terms.

(c) False in general. The correct t-statistic is $t = \frac{\hat{\beta}-\hat{\gamma}}{\sqrt{\text{s.e.}(\hat{\beta})^2+\text{s.e.}(\hat{\gamma})^2-2\text{Cov}(\hat{\beta},\hat{\gamma})}}$. The t-statistic formulated in the question is only valid if $\hat{\beta}$ and $\hat{\gamma}$ are uncorrelated, which is generally not the case.

(d) False. The effect of a college diploma on wages for females is given by $\hat{\beta}_2 + \hat{\beta}_3$, so we must test the sum of these estimates to see whether we can reject at the 5% level. $\hat{\beta}_3$ alone only tells us the relative effect of college for females over non-females, not the overall effect of college for females.

2. Answer:

(a) Holding fixed all the other covariates in regression (2), a one point increase in beauty leads to a 0.275 point increase in teaching evaluation scores. A one standard deviation increase in beauty (0.5) leads to a 0.1375 increase in teaching evaluation scores. This effect is statistically significant.

(b) The predicted evaluation score for Professor Mikusheva would be $\widehat{\text{score}} = \underbrace{4.25}_{\text{intercept}} + \underbrace{0.275 \times 0}_{\text{beauty}} + \underbrace{-0.239 \times 1}_{\text{female}} + \underbrace{-0.249 \times 0}_{\text{minority}} + \underbrace{-0.253 \times 1}_{\text{non-native English}} + \underbrace{-0.136 \times 1}_{\text{tenure track}} + \underbrace{-0.046 \times 0}_{\text{intro course}} + \underbrace{0.687 \times 0}_{\text{one-credit}} = 3.622$

(c) The only difference between Professor Mikusheva and the subsequent professor is that the subsequent professor is not female. Therefore a 95% confidence interval for the difference in their course scores would be $[-0.239 - 1.96 \times 0.085, -0.239 + 1.96 \times 0.085] = [-0.4056, -0.0724]$

(d) The difference between regressions (1) and (2) is the inclusion of the one credit course control in the latter. This is clearly an omitted variable in the first regression. Given that the coefficient on beauty drops from regression (1) to (2), this means that there was upwards bias in the first regression. The positive coefficient on one-credit courses indicates that the relationship between beauty and one-credit courses is therefore positive. This makes sense if one believes that those who engage in yoga, aerobics, dance, etc have higher beauty.

(e) We have $t = \frac{0.081}{0.135} = 0.6$. Given that $|t| < 1.96$, we fail to reject that the coefficient on $Beauty \times D_{Beauty>0}$ is non-zero. This is a test of whether the effect of beauty is different for those with above-average beauty compared to those with below-average beauty. The null hypothesis is that the effect of a higher beauty score is the same regardless of whether a person has high or low beauty. The alternative is that beauty has a larger effect for one of the two groups. We find no evidence that the effects of beauty depend on the level of beauty.

(f) The t-statistic is $t = \frac{0.384-0.128}{\sqrt{0.076^2+0.064^2}} = 2.57$. Given that $|t| > 1.96$, we can reject at the 5% level that the effect of beauty is the same for men and women.

3. Answer:

(a) We can see that the R^2 is almost twice as large for the log-specification. Thus, (2) is the more preferable specification as it explains a larger share of the variation in Growth.

(b) The statement is essentially claiming that a linear specification is sufficient for the effect of *TradeShare* on *Growth* and thus we should look at model (5) and test whether the coefficients on $TradeShare^2$ and $TradeShare^3$ are zero. The F-statistic for this test is 1.96 and the p-value is 0.15. Thus, we cannot reject the null hypothesis the two coefficient are zero at the 5% significance level, and the data seems to give support to the statement.

(c) The statement is essentially claiming that coefficient on $TradeShare \times \log YearsSchool$ is positive. We can see that the coefficient on $TradeShare \times \log YearsSchool$ in model (4) is not statistically significantly different from zero at the 5% level. Thus, the data does not give support to the claim that the effect of $TradeShare$ depends on the level of $\log YearsSchool$.

(d) If this claim is true, we should have a zero coefficient on $TradeShare$ once we control for political instability. We may look at model (3). When we control for political instability using Rev_coups and $Assassinations$, we have a much smaller estimate than in the simple specification and it is no longer statistically significantly different from zero. Thus, we may conclude that the claim is supported by the data in the sense that a large proportion of the correlation between $TradeShare$ and $Growth$ can be explained by political instability. (We can also regress $TradeShare$ on Rev_coups and $Assassinations$ to find both coefficients negative.)

(e) Based on parts (a), (b) and (c), model (3) is the preferred model for this analysis. Based on (d), $TradeShare$ does not have a statistically significant effect on $Growth$. In terms of economic significance, increasing $TradeShare$ from 10% to 50% is almost a two standard deviation movement and thus a relatively large change. According to model (3), this change predicts a $1.104 \times 0.4 \approx 0.44$ increase in $Growth$. This does not seem particularly large as the standard deviation of $Growth$ is 1.8. Thus, we may conclude that the claim is supported by the data in the sense that the effect of trade on growth is small in both statistical and economic terms.

Table 1: Effects of openness to trade and education on economic growth

	(1)	(2)	(3)	(4)	(5)
tradeshare	1.898*	1.749*	1.104	1.883	-5.702
	(0.866)	(0.794)	(0.748)	(1.344)	(8.183)
[1em] yearsschool	0.243**				
	(0.0759)				
[1em] lyearsschool		1.016***	2.161***	2.525***	2.133***
		(0.204)	(0.400)	(0.639)	(0.416)
[1em] rev_coups			-2.300*	-2.350*	-2.035*
			(0.921)	(0.921)	(0.938)
[1em] assasinations			0.228	0.224	0.102
			(0.335)	(0.337)	(0.364)
[1em] lrgdp60			-1.621***	-1.641***	-1.584**
			(0.440)	(0.440)	(0.477)
[1em] tradeshare_lyearsschool				-0.690	
				(0.809)	
[1em] tradeshare2					8.488
					(15.83)
[1em] tradeshare3					-2.760
					(8.527)
[1em] _cons	-0.122	-0.186	11.75***	11.50**	12.93***
	(0.691)	(0.566)	(3.297)	(3.345)	(3.055)
r2	0.161	0.287	0.453	0.457	0.471
N	64	64	64	64	64

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$