BUEC 333 D200, Spring 2012, Final Solutions

1. For simplicity, assume there is only one explanatory variable X. The population regression equation is $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$. (Don't worry about the i subscript, but make sure there are no hats and ϵ and no e.) The estimated regression equation is $Y_i = \hat{Y}_i + e_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + e_i$. We have a sample of observations indexed by $i = 1, \cdots, N$. The OLS estimator minimizes the sum of squared residuals, $\sum_{i=1}^{N} e_i^2$. Crucial ingredients: residuals, not error terms; sum; squares; appearance of the population and/or estimated regression equation

2. An estimator is a recipe, an estimate is what you get by applying the recipe to a specific sample; the estimate is a number, the estimator is a random variable.

3. An answer should be in line with pages 54-57. For example the text after the second box on page 54 and the start of section 2.5.

4. A possible algebraic answer is: assume that $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ with $E(\epsilon_i) = 7$. Then, we can absorb the 7 into the constant term:

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_i + \epsilon_i \\ &= \beta_0 + \beta_1 X_i + \epsilon_i - 7 + 7 \\ &= (\beta_0 + 7) + \beta_1 X_i + (\epsilon_i - 7) \\ &= \beta_0^* + \beta_1 X_i + \epsilon_i^* \end{aligned}$$

   For the new model on the last line, the new error term ϵ_i^* has mean 0, as $E(\epsilon_i^*) = E(\epsilon_i - 7) = 7 - 7 = 0$. So you can force the non-zero mean of the error term to go into the constant term.

5. That it is BUE. You might get some points if you mention that they are normally distributed so that you can do hypothesis testing.

6. (a) BED. $H_0 : \beta_3 <= 0$ versus $H_A : 3 > 0$. The t-score is equal to 1, while $t_c = 1.711$. So, do not reject H_0.
   BEACH. $H_0 : \beta_5 <= 0$ versus $H_A : \beta_5 > 0$. The t-score equals 10, while $t_c = 1.711$. Reject H_0.
   (b) AGE. $H_0 : \beta_2 >= 0$ versus $H_A : \beta_2 < 0$. The t-score is -2, while $t_c = 1.318$, so reject H_0.
   (c) FIRE. $H_0 : \beta_4 = 0$ versus $H_A : \beta_4 \neq 0$. The t-score equals -1, while $t_c = 2.064$, so do not reject H_0.
   (d) One unexpected sign could be β_4 (FIRE), although it is not significantly different from zero. (i) This could be due to an omitted variable, or because there is no effect of FIRE (sampling distribution answer) (ii) Include the omitted variable if you suspect an omitted variable bias is the reason. If you think there is no effect of FIRE on PRICE, then do nothing or drop the variable. [An alternative example could go through AGE. In some cities (Vancouver?), older houses are pricier. So you could accept an answer along these lines. In general, any variable that they single out with some kind of argument is fine. For example, some argument could go against BED, since you are already controlling for LOT size. If the student answers with "No, I have no unexpected signs, you can allocate points if they provide an explanation.

7. See page 183-184, "bias caused by relying on the t-test to choose variables".

8. Causes bias.

9. Christoph/Ebrahim/Soheil

10. *This question is about perfect and imperfect multicollinearity.*

   (a) No explanatory variable is a linear combination of any other variable(s).
   (b) No.
   (c) Some possible answers: (1) even though imperfect multicollinearity does not violate assumption VI, it *almost* violates assumption IV, and it *almost* has the same problems; (2) even though no classical assumption is violated, having a high degree of imperfect multicollinearity leads to high standard errors; (3) idem, low

t-values. It would be great if you mentioned how the estimates are now sensitive to small changes in the data or in the model.

11. (1) combination of a high R^2 with low t-values; (2) correlation coefficient; (3) VIF

12. *Part of this is in the SM*

    (a) See SM, this is Ch.9, q16, (b)

    (b) See SM, this is Ch.9, q16, (c), part i

    (c) See SM, this is Ch.9, q16, (c), part ii

    (d) Bigger, because OLS underestimates the standard errors. Or: smaller, because GLS is minimum variance and OLS is not. (The best answer would be "I cannot say, because there are two effects..." but that is *not* required for full points.)

13. I expect some explanation (possibly very intuitive) that having serial correlation in the error terms leads to the OLS being biased for SE($\hat{\beta}$). Given that the standard errors are biased, t-ratios are going to be biased as well, given that SE is in the denominator. Since hypothesis testing is done using the t-ratio, hypothesis tesing is now unreliable. Additionally, an answer includes the statement that the bias is such that OLS *underestimates* the SE, so *overestimates* the t, leading you to reject to often. (Newey-West SEs would solve this problem in large samples.)

14. If Z_i has a normal distribution, it can take negative values. As a result, if $\sigma_i^2 = \sigma^2 Z_i$ can be negative as well. That is problematic, as the variance should always be nonnegative. Two models that do not have this problem: $\sigma_i^2 = \sigma^2 Z_i^2$; $\sigma_i^2 = \sigma^2 e^{\gamma Z_i}$.