Economics 471: Econometrics

University of Alabama

Department of Economics, Finance and Legal Studies

Fall 2017

Midterm II – Answers

1. (a) (1) $y = \alpha + \beta x + u$ (the model is correctly specified); (2) $E(u) = 0$ (mean zero error); (3) $E(u|x) = 0$ (errors are not correlated with the regressors); (4) no perfect collinearity; (5) $V(u|x) = \sigma^2$ (homoskedasticity); (6) $u \overset{iid}{\sim} N(0, \sigma^2)$

   (b) (1) Examples include a missing regressor, irrelevant regressor or missing nonlinearity; (2) Examples include a non-mean zero error or one of the cases above where the additional terms end up in the error and are non-mean zero; (3) Examples include the three cases from part (1); (4) two or more variables that are perfectly correlated; (5) A heteroskedastic variance; (6) Examples include cases where the errors are not normally distributed, are not iid or are heteroskedastic

   (c) These answers will depend upon the example in part (b), but typically any violation of (1-4) leads to biased estimators, a violation of (5) leads to inefficient estimators (and biased estimates of the variance of the estimators) and a violation of (6) if only on the distribution leads to incorrect inference (distribution of the test statistics)

2. (a) SSR_R: residual sum of squares from the restricted model. SSR_U: residual sum of squares from the unrestricted model. q: numerator degrees of freedom (number of restrictions in the null). n: sample size. $k+1$: number of parameters estimated

   (b) $F_{q,(n-k-1)}$ and the the F-distribution, the plot is that of a F-distribution which is a non-negative skewed distribution from 0 to infinity - as the restricted sum of squares can never be smaller than the unrestricted version

   (c) $F = \frac{(SSR_R - SSR_U)/q}{SSR_U/(n-k-1)} = \frac{\left(\frac{SSR_R}{SST} - \frac{SSR_U}{SST}\right)/q}{\frac{SSR_U}{SST}/(n-k-1)} = \frac{\left[\left(1-R_R^2\right)-\left(1-R_U^2\right)\right]/q}{\left(1-R_U^2\right)/(n-k-1)} = \frac{\left(R_U^2 - R_R^2\right)/q}{\left(1-R_U^2\right)/(n-k-1)}$

   (d) $SST = \sum_{i=1}^{n} (y_i - \overline{y})^2$ and does not depend upon the model (restricted or unrestricted)

   (e) The simple example given in class had three regressors $y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u$ whereby two were deemed irrelevant $H_0 : \beta_1 = \beta_2 = 0$. For a given sample size, $q = 2$ and $n - k - 1 = n - 3 - 1$.

3. (a) In the linear model, the partial effect $\frac{\partial y}{\partial x} = \beta$, is the coefficient on x. In the quadratic model, the partial effect $\frac{\partial y}{\partial x} = \beta + 2\gamma x$ is a function of x. Plugging in the values from the table we see that $\widehat{\frac{\partial y}{\partial x}} = 4.0113$ in the linear model and $\widehat{\frac{\partial y}{\partial x}} = 9.5326 + 2(-1.6915)x = 9.5326 - 3.3832x$.

   (b) In the linear model a one-unit increase in homework will bring about the same change in test score regardless of the value of homework. Given that we know the maximum score is 100, we need to find the value of homework that gives the predicted test score equal to 100. $100 = 49.838 + 4.0113 * homework^* \Rightarrow homework^* = (100 - 49.838)/4.0113 = 12.50$ (insane) hours per day. For the quadratic model with $\beta > 0$ and $\gamma < 0$ we know that the function achieves its maximum when

$\frac{\partial y}{\partial x} = \beta + 2\gamma x = 0 \Rightarrow 9.5326 - 3.3832 * homework^* = 0 \Rightarrow homework^* = 9.5326/3.3832 = 2.8176 \Rightarrow$ the maximum test score is equal to $47.232 + 9.5326 * 2.8176 - 1.6916 * 2.8176^2 = 60.6616$.

(c) The easiest way to do this is to test the null: $H_0 : \gamma = 0$ in the quadratic model. The t-statistic for this test is -9.0473 and its respective $p - value = 0.0000$. Therefore we reject the model is linear. A more complicated test would be to take the RSS from both the linear and quadratic models, construct the F-statistic$= \frac{(332301.3-325165.7)/1}{325165.7/(3733-2-1)} = 81.835$ and compare them to the critical value from the F-table$= 3.84$ which would also suggest to reject the null. Note that $t^2 = F$ in this case of a single restriction.

(d) In this case we are confident that the sign of γ is negative and given that homework is non-negative, the correlation between homework and homework squared is positive. Therefore the bias of the estimate of the coefficient on homework in the linear model is negative.

(e) Here we list all five selection criteria available from the EViews output. Note that the quadratic model performs better in each. It has a higher value for each of the first two selection criteria and a smaller value for the remaining selection criteria.

	Selection Criteria				
	R^2	$\overline{R}^2$	$\widehat{\sigma}$	AIC	SC
linear model	0.0271	0.0268	9.4374	7.3278	7.3311
quadratic model	0.0479	0.0475	9.3368	7.3066	7.3116