UNIVERSITY OF CAMBRIDGE
Judge Business School

Updated: November 11, 2011

Lecturer: Thilo Klein
Contact: tk375@cam.ac.uk

Contest Quiz 6

Question Sheet

In this quiz we will review non-linearity and model transformations covered in lectures 6 and 7.

Question 1: Logarithms

(i) The interpretation of the slope coefficient in the model $Y_i = \beta_0 + \beta_1 ln(X_i) + u_i$ is as follows:

(a) a 1% change in X is associated with a β_1% change in Y.
(b) a 1% change in X is associated with a change in Y of 0.01 β_1.
(c) a change in X by one unit is associated with a $\beta_1 100$% change in Y.
(d) a change in X by one unit is associated with a β_1 change in Y.

(ii) The interpretation of the slope coefficient in the model $ln(Y_i) = \beta_0 + \beta_1 X_i + u_i$ is as follows:

(a) a 1% change in X is associated with a β_1% change in Y.
(b) a change in X by one unit is associated with a $100\beta_1$% change in Y.
(c) a 1% change in X is associated with a change in Y of $0.01\beta_1$.
(d) a change in X by one unit is associated with a β_1 change in Y.

(iii) The interpretation of the slope coefficient in the model $ln(Y_i) = \beta_0 + \beta_1 ln(X_i) + u_i$ is as follows:

(a) a 1% change in X is associated with a β_1% change in Y.
(b) a change in X by one unit is associated with a β_1 change in Y.
(c) a change in X by one unit is associated with a $100\beta_1$% change in Y.
(d) a 1% change in X is associated with a change in Y of $0.01\beta_1$.

(iv) To decide whether $Y_i = \beta_0 + \beta_1 X + u_i$ or $ln(Y_i) = \beta_0 + \beta_1 X + u_i$ fits the data better, you cannot consult the regression R^2 because

(a) $ln(Y)$ may be negative for $0 < Y < 1$.
(b) the TSS are not measured in the same units between the two models.
(c) the slope no longer indicates the effect of a unit change of X on Y in the log-linear model.
(d) the regression R^2 can be greater than one in the second model.

(v) The exponential function

(a) is the inverse of the natural logarithm function.

(b) does not play an important role in modeling nonlinear regression functions in econometrics.

(c) can be written as $exp(e^x)$.

(d) is e^x, where e is 3.1415...

(vi) The following are properties of the logarithm function with the exception of

(a) $ln(1/x) = -ln(x)$.

(b) $ln(a + x) = ln(a) + ln(x)$.

(c) $ln(ax) = ln(a) + ln(x)$.

(d) $ln(xa) = aln(x)$.

(vii) In the log-log model, the slope coefficient indicates

(a) the effect that a unit change in X has on Y.

(b) the elasticity of Y with respect to X.

(c) $\Delta Y/\Delta X$.

(d) $\frac{\Delta Y}{\Delta X} \times \frac{Y}{X}$

(viii) In the model $ln(Y_i) = \beta_0 + \beta_1 X_i + u_i$, the elasticity of $E(Y|X)$ with respect to X is

(a) $\beta_1 X$

(b) β_1

(c) $\frac{\beta_1 X}{\beta_0 + \beta_1 X}$

(d) Cannot be calculated because the function is non-linear

(ix) Consider the following least squares specification between testscores and the student-teacher ratio:

$$\widehat{TestScore} = 557.8 + 36.42ln(Income).$$

According to this equation, a 1% increase income is associated with an increase in test scores of

(a) 0.36 points

(b) 36.42 points

(c) 557.8 points

(d) cannot be determined from the information given here

Question 2: Interactions

(i) In the case of regression with interactions, the coefficient of a binary variable should be interpreted as follows:

(a) there are really problems in interpreting these, since the $ln(0)$ is not defined.

(b) for the case of interacted regressors, the binary variable coefficient represents the various intercepts for the case when the binary variable equals one.

(c) first set all explanatory variables to one, with the exception of the binary variables. Then allow for each of the binary variables to take on the value of one sequentially. The resulting predicted value indicates the effect of the binary variable.

(d) first compute the expected values of Y for each possible case described by the set of binary variables. Next compare these expected values. Each coefficient can then be expressed either as an expected value or as the difference between two or more expected values.

(ii) The following interactions between binary and continuous variables are possible, with the exception of

(a) $Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + \beta_3 (X_i \times D_i) + u_i$.

(b) $Y_i = \beta_0 + \beta_1 X_i + \beta_2 (X_i \times D_i) + u_i$.

(c) $Y_i = (\beta_0 + D_i) + \beta_1 X_i + u_i$.

(d) $Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + u_i$.

(iii) Including an interaction term between two independent variables, X_1 and X_2, allows for the following except:

(a) the interaction term lets the effect on Y of a change in X_1 depend on the value of X_2.

(b) the interaction term coefficient is the effect of a unit increase in X_1 and X_2 above and beyond the sum of the individual effects of a unit increase in the two variables alone.

(c) the interaction term coefficient is the effect of a unit increase in $\sqrt{X_1 \times X_2}$.

(d) the interaction term lets the effect on Y of a change in X_2 depend on the value of X_1.

(iv) The binary variable interaction regression

(a) can only be applied when there are two binary variables, but not three or more.

(b) is the same as testing for differences in means.

(c) cannot be used with logarithmic regression functions because $ln(0)$ is not defined.

(d) allows the effect of changing one of the binary independent variables to depend on the value of the other binary variable.

(v) In the regression model $Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + \beta_3 (X_i \times D_i) + u_i$, where X is a continuous variable and D is a binary variable, β_3

(a) indicates the slope of the regression when $D = 1$.

(b) has a standard error that is not normally distributed even in large samples since D is not a normally distributed variable.

(c) indicates the difference in the slopes of the two regressions.

(d) has no meaning since $(X_i \times D_i) = 0$ when $D_i = 0$.

(vi) In the regression model $Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + \beta_3 (X_i \times D_i) + u_i$, where X is a continuous variable and D is binary, to test that the two regressions are identical, you must use the

(a) t-statistic separately for $\beta_2 = 0, \beta_2 = 0$.

(b) F-statistic for the joint hypothesis that $\beta_0 = 0, \beta_1 = 0$.

(c) t-statistic separately for $\beta_3 = 0$.

(d) F-statistic for the joint hypothesis that $\beta_2 = 0, \beta_3 = 0$.

(vii) In the model $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_1 \times X_2) + u_i$, the expected effect $\frac{\Delta Y}{\Delta X_1}$ is

(a) $\beta_1 + \beta_3 X_2$.

(b) β_1.

(c) $\beta_1 + \beta_3$.

(d) $\beta_1 + \beta_3 X_1$.

Question 3: Polynomials

(i) You have estimated the following equation:

$$\widehat{TestScore} = 607.3 + 3.85 Income - 0.0423 Income^2,$$

where *TestScore* is the average of the reading and math scores on the Stanford 9 standardized test administered to 5th grade students in 420 California school districts in 1998 and 1999. Income is the average annual per capita income in the school district, measured in thousands of 1998 dollars. The equation

(a) suggests a positive relationship between test scores and income for most of the sample.

(b) is positive until a value of *Income* of 610.81.

(c) does not make much sense since the square of income is entered.

(d) suggests a positive relationship between test scores and income for all of the sample.

(ii) To test whether or not the population regression function is linear rather than a polynomial of order r,

(a) check whether the regression R^2 for the polynomial regression is higher than that of the linear regression.

(b) compare the *TSS* from both regressions.

(c) look at the pattern of the coefficients: if they change from positive to negative to positive, etc., then the polynomial regression should be used.

(d) use the test of $(r - 1)$ restrictions using the F-statistic.

(iii) The best way to interpret polynomial regressions is to

(a) take a derivative of Y with respect to the relevant X.

(b) plot the estimated regression function and to calculate the estimated effect on Y associated with a change in X for one or more values of X.

(c) look at the t-statistics for the relevant coefficients.

(d) analyze the standard error of estimated effect.

(iv) Misspecification of functional form of the regression function

(a) is overcome by adding the squares of all explanatory variables.

(b) is more serious in the case of homoskedasticity-only standard error.

(c) results in a type of omitted variable bias.

(d) requires alternative estimation methods such as maximum likelihood.