[Frequently Asked Questions]

Specification Errors

Subject:	Business Economics
Course:	B. A. (Hons.), 5th Semester, Undergraduate
Paper No. & Title:	Paper – 531 Elective Paper Q1 – Advanced Econometrics
Unit No. & Title:	Unit – 1 Relaxing the Assumptions of The Classical Linear Model
Lecture No. & Title:	Lecture – 6 Specification Errors

Frequently Asked Questions

Q1. What is meant by specification errors?

A1. The term specification error refers to any mistake in the set of assumptions underlying the model and the associated inference procedure, but it has come to be used particularity for error in specifying the data matrix X. Generally this may be due to omitting some variable or including some variable in the model. It may also be due to the incorrect specification of the model form. This results in specification error and specification bias.

Q2. What is the effect of under fitting the model?

A2. When some explanatory variable in the proposed model is omitted, it is called underfitting the model. This gives biased and inconsistent estimators with larger variance, Due to this usual test of significance do not remain reliable and the conclusions based upon confidence intervals are likely to give misleading results.

Q3. What is the effect of overfitting the model?

A3. When, we include some other variable in the original model, it is called overfitting the model. This gives rise to specification error and specification bias. Here we may get unbiased estimates but if the explanatory variable included in the new model is correlated with the explanatory variable in the original model, then the variance of estimated parameters will be less efficient as compared to that for the original model. Such an inclusion can also lead to the problem of multicollinearity.

Q4. Name some tests for detecting specification errors.

A4. We can use some commonly popular tests like

(a) During wanton d test

(b) Ramsey's regression specific error RESET test

(c) Lagrange's multiplier test etc.

Q5. Suppose that true model is $Y_i = \beta_1 + \beta_2 X_{2i} + U_i \rightarrow (1)$ but we add an irrelevant Variable X_3 in the model and write $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + v_i \rightarrow (2)$ what will be the effect due to new model?

A5. In fact comparing (1) and (2), we find that $U_i = \beta_3 X_{3i} + v_i$ .Thus there will be specification error. Here $\hat{\beta}_1, \hat{\beta}_2$ and $\hat{\beta}_3$ are unbiased estimators as obtained by OLS method. They are also consistent. However the estimated, variance of the regression coefficients will be larger, hence the estimators will be less precise. Also confidence intervals will be larger and hence they will not give prediction very precisely.

Q6. To deal with specification due to measurement errors, we want to apply Bartlett's test. How will you proceed?

A6. Bartlett's test assumes that there are n = 3K sample observations. To apply this test, first these observation are arranged in ascending order as per the X variable and correspondingly observation on Y equal groups of size K each for X and Y variables. Find mean for each group of variables. If we denote them by $\bar{X}_{(1)}, \bar{X}_{(2)}, \bar{X}_{(3)}$ and $\bar{Y}_1^*, \bar{Y}_2^*, \bar{Y}_3^*$ respectively then coefficient β is estimated by $\hat{\beta} = \frac{\bar{Y}_3^* - \bar{Y}_1^*}{\bar{X}_{(3)} - \bar{X}_{(1)}}$ and $\hat{\alpha} = \bar{Y}^* - \hat{\beta}\bar{X}_0$ and where $\bar{X}_0 = \sum_1^n \frac{X_{(i)}}{n}$ (If n=3k + 1,we delete the middle most observation for median and then workout as above.)

Q7. What are measurement errors? What will be its effects when model is estimated?

A7. When making measurement upon the variables, if they are not very accurately measured, there remain errors which are called measurement errors. If there are errors of measurement in the regressand only, the OLS estimators are unbiased as well as consistent, but they are less efficient. If there are errors of measurement in the regressors, the OLS estimators are biased as well as inconsistent.

Q8. What is Instrumental variables method? How for it can be useful?

A8. Instrumental variable method is an approach to deal with measurement errors. Here we devise the IV Z Such that its observations are correlated with X but uncorrelated with the errors terms U and v associated with X and Y respectively. This leads to give consistent estimator of the slope efficient. However it is difficult to obtain such a variable to operate on the model. If such IV is correlated with X then this method is not suitable.

Q9. While dealing with specification error using Ramsey's test, which test statistic is used?

A9. We Compute R^2_{old} and R^2_{new}values from the true and new model, by using the method as indicated by the test, F statistics is computed based upon these values of R^2_{old} and R^2_{new} values and if we find F as significant we can accept the hypothesis that model is misspecified (for more details, see the text given)

Q10. Which test statistic is used in Lagrange's multiplier test to check for specification errors?

A10. Lagrange's method involves first running auxiliary regression and find R^2 then chi square test is used for large samples. If χ^2 is significant we reject the hypothesis about restricted regression. (for more details, see the text given)