

Specification Error

1. NATURE

One of the assumptions which we have made is that the model should be correctly specified. A model is said to be correctly specified if it includes all the relevant variables and if its functional form is correct *i.e.* linear or curvilinear. If a model is not correctly specified it is called mis-specified model. In other words, the relevant variables may be excluded and irrelevant variables may be included. Specification of model presupposes the knowledge of economic theory.

Specification involves three things :

1. Determination of dependent and explanatory variables which will be included in the model.
2. Determination of a priori theoretical expectations about the magnitude and sign of the parameters of the function.
3. Determination of the mathematical form of the model, *e.g.*, linear, non-linear, simultaneous or single equation model.

Reason for incorrect specification of economic models are :

- (a) Limitation of our knowledge of the variables which are operative.
- (b) Non-availability of required data.
- (c) Imperfection of economic theory itself which forms the basis for the specification of the model.

2. TYPES OF THE SPECIFICATION ERROR

Specification error results from :

- (i) Omission of relevant variables from the function.
- (ii) Inclusion of irrelevant variables in the function.
- (iii) Omission of the equations from the model, and
- (iv) Incorrect mathematical form of the model.

1. Omission of relevant variables:

- The estimators are biased and inconsistent.
- As a result the hypothesis tests do not hold.
- Even choosing a larger sample size does not make the estimators unbiased or consistent.
- The inconsistency of estimators is generated by a lower than normal variance in regression analysis.

Eg: True model is

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i \quad \text{--- ①}$$

But due to ignorance or because of non-availability of data on X_2 , the following model is estimated.

$$Y_i = \beta_0 + \beta_1 X_{1i} + u_i \quad \text{--- ②}$$

2. Inclusion of irrelevant explanatory variable:

- This is the least serious problem that leads to specification error.
- The hypothesis tests of a model which has included an irrelevant variable are still valid.

c) The inclusion of irrelevant variable does not affect the relationship between other variables and the dependent variable because the estimator for such a variable turns out to be zero. Page 2

d) The estimators violate the BLUE (Best Linear Unbiased Estimator) concept of regression because they are inefficient.

Eq: Let the true model be

$$Y_i = \beta_0 + \beta_1 X_{1i} + u_i$$

But the equation estimated includes X_2 explanatory variable :-

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$$

3) Incorrect mathematical form of the model:

a) When one chooses the wrong functional form of the regression model, the model will have a specification error.

b) Eq: If one chooses a double-log model for the analysis instead of log-linear model, the model will suffer from specification bias.