Problems with OLS

Considering :

$$Y_i = \alpha + \beta X_i + u_i$$

we assume

$$E(u_i) = 0$$
$$E(u_i^2) = \sigma^2 \text{ or } var(u_i) = \sigma^2$$
$$E(u_i u_j) = 0 \text{ or } cov(u_i, u_j) = 0$$

We have seen that we have to make very specific assumptions about u_i in order to get OLS estimates with the desirable properties.

If these assumptions don't hold than the OLS estimators are not necessarily BLU.

We can respond to such problems by changing specification and/or changing the method of estimation.

First we consider the problems that might occur and what they imply. In all of these we are basically looking at the residuals to see if they are random.

- **1**. The errors are serially dependent ⇒ autocorrelation/serial correlation.
  2. The error variances are not constant ⇒ heteroscedasticity
  3. In multivariate analysis two or more of the independent variables are closely correlated ⇒ multicollinearity
  4. The function is non-linear
  5. There are problems of outliers or extreme values -but what are outliers?
  6. There are problems of missing variables ⇒can lead to missing variable bias

Of course these problems do not have to come separately, nor are they likely to

- Note that in terms of significance things may look OK and even the $R^2$the regression may not look that bad.
- Really want to be able to identify a misleading regression that you may take seriously when you should not.
- The tests in Microfit cover many of the above concerns, but you should always plot the residuals and look at them.