

## MOOC Econometrics

Lecture 3.4 on Model Specification:  
Evaluation

Michel van der Wel

## RESET

Instead, add fitted  $y$ -values  $\hat{y} = Xb = X(X'X)^{-1}X'y$  to the model:

$$y_i = x_i'\beta + \sum_{j=1}^p \gamma_j (\hat{y}_i)^{j+1} + \varepsilon_i,$$

and test for joint significance of  $\gamma$ 's. Under null of correct specification,  $H_0: \gamma_j = 0$  for all  $j$ , test distribution approximately  $F(p, n - k - p)$ .

## Test

For  $p = 1$ , compute the number of extra parameters in the alternative specification as compared to the total number of parameters in the RESET specification.

Answer: Above model with  $p = 1$  has  $k + 1$  parameters. Model with squares and cross-terms has  $k + (k - 1) + \frac{1}{2}(k - 2)(k - 1)$  coefficients. For example, if  $k = 6$ , then this is 7 compared to 21.

## RESET

Extend linear model

$$y_i = \beta_1 + \sum_{j=2}^k \beta_j x_{ji} + \varepsilon_i,$$

to non-linear model

$$y_i = \beta_1 + \sum_{j=2}^k \beta_j x_{ji} + \sum_{j=2}^k \gamma_{jj} x_{ji}^2 + \sum_{j=2}^k \sum_{h=j+1}^k \gamma_{jh} x_{ji} x_{hi} + \varepsilon_i.$$

Test for linearity by testing significance of  $\gamma$  coefficients.

Challenge: Nonlinear model contains many parameters.

## Chow break test

In case of a possible break, split the sample and test for constancy of parameters.

$$\begin{aligned} y_1 &= X_1 \beta_1 + \varepsilon_1 & (n_1 \text{ observations}) \\ y_2 &= X_2 \beta_2 + \varepsilon_2 & (n_2 = n - n_1 \text{ observations}) \end{aligned}$$

Combine:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

Test  $H_0: \beta_1 = \beta_2$  against this unrestricted set-up.

## Chow break test

F-test for null hypothesis of no break:

$$F = \frac{(e'_R e_R - e'_U e_U)/k}{e'_U e_U/(n-2k)}.$$

Here:

- Have  $e_U = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$ , with  $e_j$  the OLS residuals of each group.
- Thus  $e'_U e_U = e'_1 e_1 + e'_2 e_2 \equiv S_1 + S_2$ .
- Get  $F = \frac{(S_0 - S_1 - S_2)/k}{(S_1 + S_2)/(n-2k)}$ , with  $S_0 = e'_R e_R$ .



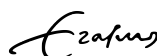
Lecture 3.4, Slide 5 of 8, Erasmus School of Economics

## Test for normality of error terms

- Model misspecification may appear in the error terms.
- Normality of  $\varepsilon$  can be tested by distribution of residuals.
- Jarque-Bera test evaluates skewness  $S$  and kurtosis  $K$ :

$$JB = \left( \sqrt{\frac{n}{6}} S \right)^2 + \left( \sqrt{\frac{n}{24}} (K - 3) \right)^2,$$

which approximately has  $\chi^2(2)$  distribution if  $H_0: \varepsilon_i \sim NID(0, \sigma^2)$  holds true.



Lecture 3.4, Slide 7 of 8, Erasmus School of Economics

## Chow forecast test

A variation on the Chow break test is based on

$$y_i = x'_i \beta + \sum_{j=n_1+1}^{n_1+n_2} \gamma_j D_{ji} + \varepsilon_i,$$

test  $H_0: \gamma_j = 0$  for all  $j$ .

### Test

What is the number of parameters in the above specification?

Answer: The model contains the usual  $k$  variables and  $n_2$  dummy-variables (one for each observation in group 2), so in total  $k + n_2$  parameters.

- Perfect fit in second sample, thus  $e_2 = 0$ .
- Thus  $F = \frac{(S_0 - S_1)/n_2}{S_1/(n_1 - k)}$ .



Lecture 3.4, Slide 6 of 8, Erasmus School of Economics

## TRAINING EXERCISE 3.4

- Train yourself by making the training exercise (see the website).
- After making this exercise, check your answers by studying the webcast solution (also available on the website).



Lecture 3.4, Slide 8 of 8, Erasmus School of Economics