Class 7. Choosing the functional form
Advanced Econometrics I

Lectures: Elena Kotyrlo
Classes: Yuri Trifonov, Elena Semerikova

Fall 2022

Problem 1 "Titanic"

Jack and Rose want to build a regression model that aims to explain the safety rating of the ship depending on its age and ticket price. The sample includes 42 ships made by a White Star Company and 30 ships that were developed by other companies.

First, they estimated two different models:

$$safety_i = \alpha_0 + \alpha_1 \times age_i + \alpha_2 \times t\_price_i + \varepsilon_i \tag{1}$$

$$\ln safety_i = \beta_0 + \beta_1 \times \ln age_i + \beta_2 \times \ln t\_price_i + \xi_i \tag{2}$$

Further, they constructed two artificial models, but Rose forgot which variables they have included in both models to run a PE test:

$$\widehat{safety}_i = \underset{(2.1)}{10.1} + \underset{(0.3)}{2.5} \times age_i + \underset{(1.7)}{6.8} \times t\_price_i + \underset{(0.7)}{0.8} \times (\ldots) \tag{3}$$

$$\widehat{\ln safety}_i = \underset{(0.3)}{1.6} + \underset{(0.1)}{0.6} \times age_i + \underset{(0.4)}{1.2} \times t\_price_i + \underset{(0.5)}{1.5} \times (\ldots) \tag{4}$$

1. Help Rose to remember which variables they have included in artificial models.

2. Determine the most appropriate functional form via the PE test.

After that, Rose argued that maybe the model should also include squared terms besides linear ones. Thus, she insisted to run a RESET test to check this. After getting predicted values from the initial regression, she constructed and estimated the following artificial model:

$$\widehat{safety}_i = \underset{(1.7)}{8.1} + \underset{(0.2)}{2.3} \times age_i + \underset{(1.5)}{6.5} \times t\_price_i + \underset{(0.1)}{0.78} \times \widehat{safety^*}_i^2 \tag{5}$$

3. Help Rose to determine whether the model suffers from the misspecification of the functional form or not.

Finally, Jack assumed that maybe separate models should be used for ships developed by a White Star Company and all other manufacturers (and he is really concerned about it). He decided to check this by conducting a Chow test. The values of RSS for the pooled model and models estimated for each of the two types of ships are as follows:

$$RSS_{pooled} = 247.1,$$
$$RSS_{WS} = 105.6,$$
$$RSS_{other} = 110.4.$$

4. Help Jack to determine whether his concerns are right.

Problem 2

Using the dataset data.xlsx estimate a linear and a full log form of the following model:

$$price\_msq_i = \alpha_0 + \alpha_1 \times dist_i + \alpha_2 \times livesp_i + \alpha_3 \times kitsp_i + \alpha_4 \times metrdist_i + \alpha_5 \times floor_i + \alpha_6 \times brick + \varepsilon_i$$

(a) Choose between log and log-linear functional forms via the PE test;

(b) Conduct a RESET test to check for the misspecifications of the functional form;

(c) Check whether one model can be estimated for apartments made with bricks and with other materials.