Econometrics

Lecture 3: Explaining house prices

Dataset "hprice.dta" contains data on an application of the so-called hedonic pricing approach[1] to the housing market. It contains data on $N = 546$ houses sold in Windsor, Canada. The dependent variable, Y, is the sales price of the house in Canadian dollars, while lot size is the explanatory variable.

The price of a house is affected by more than just lot size. Therefore, any serious attempt to explain the determinants of house prices must include more explanatory variables than lot size. In our exercise we focus on four explanatory variables: the lot size of the property (in square feet), the number of bedrooms, the number of bathrooms and the number of storeys (excluding the basement).

1. Create a scatter-plot using the four explanatory variables one at a time. Provide comments.

2. Estimate the following model:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + u_i \tag{1}$$

   where y_i is the house price, x_{1i} is the lot size of the property (in square feet), x_{2i} is the number of bedrooms, x_{3i} is the number of bathrooms, x_{4i} is the number of storeys (excluding the basement) and u_i is the error term. Test hypotheses on disturbances.

3. Test the following hypothesis:

$$H_0 : \beta_j = 0 \text{ vs } H_1 : \beta_j \neq 0, j = 0, 1, 2, 3, 4.$$

4. Provide comments on the estimated values. What is the impact on the price of house if it is added an extra square foot to the size of the property, *ceteris paribus*?

5. Test the following hypothesis:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0 \text{ vs } H_1 : \beta_j \neq 0, j = 0, 1, 2, 3, 4 \text{ for at least one coefficient.}$$

6. Compute the confidence interval for the estimated parameters.

7. Estimate the following model:

$$y_i = \beta_0 + \beta_2 x_{2i} + u_i \tag{2}$$

   where the variables are the same as point 2.

8. Provide comments on the estimated values (especially on $\hat{\beta}_0$). In addition, compare $\hat{\beta}_2$ of equation (1) with $\hat{\beta}_2$ of equation (2). The purpose is to compare the result obtained in the univariate model with the multivariate model. (Hint: bigger houses are more likely to have more bedrooms, more bathrooms, etc.)

[1] Hedonic pricing is a model that identifies price factors according to the premise that price is determined both by internal characteristics of the good being sold and external factors affecting it.