Weekly assignment 5

Jon Marius Vaag Iversen

SØK 1005 - Spring 2024

This assignment is meant to guide your revision of the last week of lectures and is about non-linearities, dummy variables and interaction terms.

1 Wages - practice OLS with Python

Use the dataset wage1 which you find using the wooldridge package on blackboard.The variable to be used are:

wage=average hourly earnings
educ=years of education
exper=years of potential experience
expersq=squared years of experience
female=1 if female
nonwhite=1 if nonwhite
northcen=1 if live in north central US
west=1 if live in wester region
south=1 if live in south region
lwage=log(wage)

(a) Find basic summary statistics (mean, st.dev, min, max) for wage, educ and exper, and the correlation between them. Briefly comment the results.
(b)Estimate the model below using OLS and give an interpretation of the results:

$$wage_i = \beta_0 + \beta_1 educ + \beta_2 exper + u_i$$

(c)Include the squared of experience. Interpret the results and calculate the turning point.
(d)Test whether experience has an effect on wages. Perform the test by hand.

(e)Test whether experience has a non-linear effect on w ages. Perform the test by hand.
(f)Does experience have different effects for males and females? Estimate a model that analyses this question and and use the p-values in the python output to perform a test to test the relationship.
(g)Estimate the same model as in (b), but use *lwage* as dependent variable. Interpret the results
(h)Include *south*, *west* and *northcen* in the model in (g) and interpret the coefficients for these variables.

2 House prices

An analyse of the determinants of house prices is presented in table 1 below. The analyse use data in the period 1978 to 1981. The variables used in the table are age (number of years), rooms (number of rooms in the property), area (square feet of the house), baths (number of bathrooms), agesq (squared age), nearinc (distance from a potentian incinerator), y81(year 81).Average price in the period (in 1978 prices) is 83 721.

(a) Interpret all the coefficients in column 1
(b) Is there a non linear relationship between age of the house and the house price. Interpret the coefficients in column 2, perform relevant tests and calculations.
(c) In 1981, decisionmakers decided to build a waste-to-energy incinerator in the area. nearinc is a dummy variable taking the value 1 if the distance to the incinerator is below 15 000 feet. Interpret the effect of the decision to build the incinerator based on the results in column 3.
(d)In column 4 and 5 the dependent variable is the logarithm of house prices. In column 5 the logarithm of area replace the other area-variable. Interpret the effect of area in both columns (e)Discuss whether the exogenity assumption is violated when analysing the effect of area in this table.

VARIABLES	(1) rprice	(2) rprice	(3) rprice	(4) lrprice	(5) lrprice
age	-216.7***	-595.8***	-667.5***	-0.00337***	-0.00316***
	(41.13)	(123.7)	(131.8)	(0.000428)	(0.000430)
rooms	3,761**	2,464	3,130*	0.0689***	0.0681***
	(1,719)	(1,740)	(1,715)	(0.0179)	(0.0182)
area	20.66***	20.92***	18.44***	0.000206***	
	(2.406)	(2.372)	(2.374)	(2.50e-05)	
baths	11,305***	7,856***	8,011***	0.153***	0.151***
	(2,522)	(2,703)	(2,655)	(0.0262)	(0.0274)
agesq		2.580***	3.149***		
		(0.796)	(0.828)		
y81nrinc			-13,334***		
			(5,121)		
nearinc			6,967*		
			(3,886)		
y81			13,588***		
			(2,817)		
larea					0.412***
					(0.0547)
Constant	-7,129	12,203	6,767	10.08***	7.386***
	(8,993)	(10,679)	(10,602)	(0.0936)	(0.366)
Observations	321	321	321	321	321
R-squared	0.595	0.608	0.635	0.680	0.671

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1