Real Estate Modelling

AT 2023

Prof. Dr. Kathleen Kürschner Rauck

Tutorial 3: Regression I – Hedonic

Learning Outcomes	R Functions	Packages	Data
Be able to plot kernel density estimates	`max()`; `density()` `polygon()`		Lucas_County_data
Know how to add a curve to a plot	`curve()`; `dnorm()` `mean()`; `sd()`		
Be able to perform the Jarque-Bera test in R	`jarque.bera.test()`	`tseries` `quadprog`	
Be able to run an OLS regression using R & assess regression output	`lm()` `summary()`		
Be able to produce a two-way scatter plot & remove scientific notation from axes & add a fitted regression line to the plot	`plot()` `options()` `abline()`		
Understand/recap common functional forms			
Know how to include & interpret a quadratic term in a linear regression model	`I()`		
Be able to plot a non-linear fitted line & determine minimum/maximum point	`data.frame()`; `seq()` `min()`; `max()` `nrow()`; `predict()` `as.data.frame()` `lines()`; `abs()`		
Know how to interpret coefficient estimates for categorical variables	`as.factor()` `levels()`		
Know how to produce contingency tables	`table()`		

Practical 1: Jarque-Bera Test

a. Generate a variable for the age of buildings in the **Lucas_County_data**. Compute and plot the kernel density of variable *age*.

b. Add the density function of the normal distribution to visually inspect if *age* is normally distributed.

c. Run a Jarque-Bera test for normality of variable *age*.

Practical 2: The Vintage Effect in Real Estate

a. *Simple linear regression model:* run an ordinary least squares (OLS) regression of the dependent variable, *price*, on one explanatory variable, *age*, and a constant term. Store the results from this regression in an R object, called **fit1**.

b. Produce a scatter plot of *price* against *age*, which shows both the actual data points and the fitted line from the simple regression.

c. *Functional form:* extend the model to allow for a potential non-linear relationship between the price and age of a house, including *age* and *age squared* as explanatory variables. Store the results from this regression in an R object, called **fit2**.

d. Plot the regression results analogously to the ones in part b. Where is the tipping point?

Practical 3: Hedonic Regression Analysis

a. Hedonic regression analysis involves a multitude of regressors. Add three further hedonic items to the model: the number of bed- and bathrooms (*beds, baths*) as well as information on wall material (*wall*).

b. Which price would you predict for a ten year old, three bed-, two bathroom house with full brick exterior?

c. Can we improve the prediction by adding the total living area (*tla*) to the model? Why is the effect of *beds* now negative?

d. Which price would you predict for a ten year old, three bed-, two bathroom house with full brick exterior and a total living area of 130 square feet?