Homework 02

Your Name

2024-02-09

# 1. Problem 1.

*This problem was adapted from: Chapters 6 and 7 problems 6.17, 6.20, and 7.26 (Principles of Econometrics. John Wiley & Sons; 2018).*

The data file data/br.csv contains data on 1080 house sales in Baton Rouge, Louisiana, during July and August 2005. The variables are defined in data/br.def. Example 6.16 uses this data to estimate the following model:

In this exercise, we will investigate whether we can improve on this model by adding various characteristics of the homes.

### 1.0.1 A. Compare the following three models.

* mod1:
* mod2:
* mod2:

#### 1.0.1.1 Load the data into a data frame named house

# YOUR ANSWER HERE

*Hold out 200 observations as a test set and use remaining observations as a training set.*

# YOUR ANSWER HERE

*Use the training set to estimate each model.*

# YOUR ANSWER HERE

*Output the summary of each model along with* *, AIC, BIC by ussing the jtools::export\_summs function and setting the statistics argument as follows, statistics = c(N = "nobs",R2 = "r.squared",AIC = "AIC",BIC = "BIC").*

# YOUR ANSWER HERE

*Calculate and display the out-of-sample RMSE, MAPE and*

# YOUR ANSWER HERE

### 1.0.2 B. Evaluate the models.

*Based on results from part A, which model would you choose? Why?*

*Use the plot.fitted() function defined in the chunk below to analyze the residuals from the “winning” model.*

# YOUR ANSWER HERE

### 1.0.3 C. Add categorical data to the model.

#### 1.0.3.1 Create a new variable named traditional as follows

house$traditional <- ifelse(house$style == 1,1,0)

# YOUR ANSWER HERE

*Add the indicator variables WATERFRONT, TRADITIONAL and the interaction between WATERFRONT and TRADITIONAL to your winning model from parts A and B. Display a summary of the model.*

# YOUR ANSWER HERE

*Interpret the coefficients of each of the new independent variables.*

### 1.0.4 D. Conduct a Chow test on the model from part C.

*Test the model from part C to determine if houses with and without pools can be pooled together. Store the unrestricted model in an object named mod5.*

# YOUR ANSWER HERE

*Draw the residual vs fitted values plot for mod5. Based on this plot and the results from the Chow test, can we group together homes with and without pools?*

# YOUR ANSWER HERE

*Using the training set, re-estimate mod5 omitting all insignificant variables. Store this model in an object named mod6.*

# YOUR ANSWER HERE

*Calculate the simulated out of sample performance of mod5 and mod6.*

# YOUR ANSWER HERE

*Provide a bullet-point summary of your analysis and conclusions.*

# 2. Problem 2.

*This problem was adapted from: Chapter 7 problem 7.13 and 7.24 (Principles of econometrics. John Wiley & Sons; 2018).*

Many cities in California have passed Inclusionary Zoning policies (also known as below-market housing mandates) as an attempt to make housing more affordable. These policies require developers to sell some units below the market price on a percentage of the new homes built. For example, in a development of 10 new homes each with market value $850,000, the developer may have to sell 5 of the units at $180,000. Means and Stringham (2012) examine the effects of such policies on house prices and number of housing units available using 1990 and 2000 census data on California cities. Use the data file data/means.csv for this problem.

### 2.0.1 A. Consider the following two models.

* mod1:
* mod2:

*Load the data into a data frame named zoning*

# YOUR ANSWER HERE

*Estimate difference-in-difference statistic for mod1 and mod2*

# YOUR ANSWER HERE

*Interpret the difference-in-difference statistic for mod1 and mod2. Is the estimate of the treatment effect significant and of the expected sign?*

### 2.0.2 B. Add the log of median household income, lmedhhinc, to mod1 and mod2, and name the resulting models mod3 and mod4.

# YOUR ANSWER HERE

*Interpret the estimates of the new variables, including sign and significance.*

*How does this addition affect the estimates of the treatment effect?*