The Likelihood of an Above-Average Trade Price in Beijing Real Estate Market from 2011-2017

Luyi, Huang; Mengyang Zheng

March 22nd 2021

1 Introduction

For the final project, we will study the house market in Beijing with house price data from year 2011 to 2017 on Kaggle. The paper will investigate the dispersion of the house market in Beijing and utilize panel data method to find factors that may cause large variances of the house market in Beijing. By turning the house unit price into a binary variable (If higher than the average community price, it will be assigned 1 and if lower or equal, it will be assigned 0), a multinomial model can be constructed to find what factor will affect the binary price variable with the most possibility. To provide unbiased and consistent results, the exogeneity assumptions of both fixed effects model and random effects model need to be testified. Then, we will implement these two models above and conduct comparison experiments to analyze the house market in Beijing.

2 Motivation

The housing price is under spotlight in Beijing for years. The variance of housing price in Beijing has been large over the last two decades. Community infrastructure, house quality, geographical location and other factors may largely influence the house price for a single house unit. Thus, we are highly interested at what variables will make the house prices vary so extensively in Beijing. To capture the variance level, we will use the indicator that whether the house price of the house surpass the average price within that community as the dependent variable. We will further use panel data methodology such as fixed effects regression and random effects regression to analyze what makes the housing price ecosystem variable in Beijing.

3 Literature Review

Previous research of analyzing the Beijing house market has been conducted by a variety of scholars. We will be using credible references to examine our data

results and our conclusion. "Bubble or Not? Beijing Housing Prices"(2017) by Joshua C. Michnowski provide the explanation for the nature of the Beijing house market. Also previous regression results are available on the same data set and here is one of the links $https://www.kaggle.com/eraw0x/house-prices-in-beijing-eda-arima$.

4 Research Design

Create a variable indicating whether the house price is higher/smaller than the average price. We consider usisng fixed effects regression and random effects regression.

- Fixed Effects model
  - Identification requirement for fixed effects model:
    1. Strict Exogeneity: $E[u_{it}|x_i, c_i] = 0$
    2. full rank
  - we can use demean to run fixed effects model.
  - we might use clustered standard errors since house price is sensitive to common shock.Using clustered standard erros will make the result more robust.
  - reference: $https://www.unifyingdatascience.org/html/fixed_effects.html$
- Random effects model:
  - Identification requirement for Random effects model:
    1. $E[u_{it}|x_i, c_i] = 0$ (strict exogeneity)
    2. $E(c_i|x_i) = 0$
    3. full rank
    4. $E[uu_i.T|x_i, c_i] = \delta^2 I_T$
    5. $E[c_i^2|x_i] = 0$
  - In other words,: the differences across groups is unrelated to other explanatory variable and the expectation of it is 0

5 Timeline

- Week 1 3/25-3/31: Reading, cleaning, and reshaping data.
- Week 2 4/1-4/7: Conducting Initial Analysis and report initial results.
- Week 3 4/8-4/14: Validating the model and estimate them with econometric methods.

- Week 4 4/15-4/21: Comparing models and check with references and previous works, summarize introduction and experiment methods for the project.
- Week 5+ 4/22-5/1: Finalizing the output and have the formal report ready.