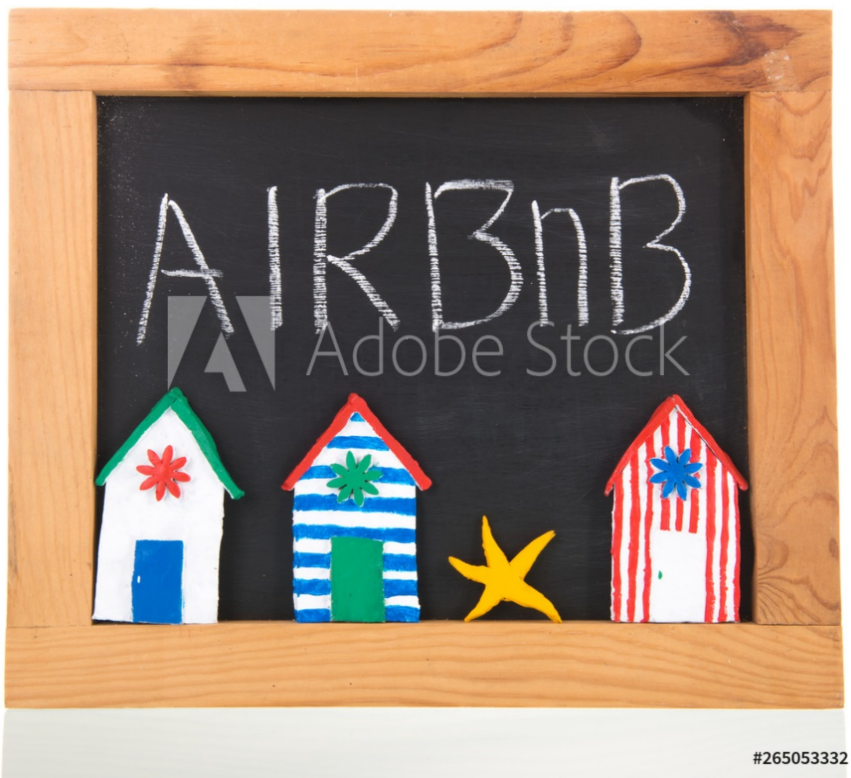
University of Amsterdam
Amsterdam Business School

The Econometrics Game 2021

April 8-9, 2021

The Sharing Economy – Airbnb and The Housing Market



Source: https://stock.adobe.com/, photo number 265053332.

By Martijn I. Dröes, University of Amsterdam, Amsterdam Business School

In cooperation with Brainbay.

The Sharing Economy – Airbnb and The Housing Market

CASE A – Airbnb and house prices

Background and motivation

In the last decade the sharing economy has increased tremendously in both scope and size and it is now affecting all aspects of modern-day life. Some well-known examples include peer-to-peer lending, crowdfunding, ridesharing, and home sharing. A prime example of the sharing economy is Airbnb. There are currently 150 million users worldwide, 6 million listings in over 65,000 cities. In this case we are going to focus on one of these cities: the city of Amsterdam.

The first listings in Amsterdam were in 2008 and have exploded exponentially. Besides having impact on the tourism/hotel industry, it also has a major impact on the housing market. On the one hand, Airbnb raises money for local residents and the (local) economy, on the other hand the increasing flow of tourists also creates nuisance and noise pollution. Which one of these aspects dominates is, a priori, unknown, but we would expect that these forces are affecting house prices (of individual houses) in Amsterdam. So the ***central research question*** for this case is:

To what extent does Airbnb affect house prices in Amsterdam?

Data

For this case microdata (housing transactions/prices, 2000-2018) was made available through Brainbay. Brainbay was established by the Dutch Association of Realtors (NVM) and governs their data. The main datafile is **AIRBNB_Houseprices.dta.**

This is a file which can easily be opened by the statistical software STATA (13 or higher). The dataset is also provided in text (tab delimited) format for those of you who want to work with alternative software. Be aware that we will use an extended version of the dataset for the finals, so keep track of the changes you make to the dataset. Labels for the variables are separately provided in the excel file: Labels_EconometricsGame2021.xlsx. The dataset has been cleaned up a bit, house prices below €25,000 and above €2.5 million are excluded from the data, but some of the other variables might still have some outliers. *Inspect the descriptive statistics carefully.*

An additional QGIS file, AIrbnb_map.qgz, is made available that shows the individual listings (for example to make some maps) as of 2008. DKT refers to transaction layers (raw housing transactions data, available per year and total), 2018 airbnb = active listings in 2018 (raw data). The coordinate system is EPSG:28992 - Amersfoort / RD New.

Modelling house prices

The main dependent variable is house prices (in euros). Distance to the nearest AIRBNB listing (calculated in the year of transaction) as well as the density (nr. of active listings within

250m, you could divide by the area to create listings per km^2) of AIRBNB are two potential independent variables. Note that the distance to AIRBNB as well as the density are set to zero before 2008. You will therefore have at least two sources of identifying variation, the start of AIRBNB and the variation in (active) AIRBNB listings over time after it started. Note that the exact location of AIRBNB listing is only known with 100m accuracy. So there will be some measurement error in the distance and density variables.

House prices can be affected by many other factors, like housing quality (housing characteristics), its location, so it is central to *control for these factors*. Note that many of the variables you have are categorical variables. One way to deal with this is to create dummy variables out of those. Other solutions are, of course, also permitted. Some variables (e.g. parking) might contain a bit too much detail/categories and might need to be aggregated. Some variables are clearly overlapping variables, as they tend to measure the same thing: pc4 & pc6 or size & volume & rooms. Please take this into account when modelling house prices.

The main identification issue is that AIRBNB is endogenous as it might be located or is particularly dense in areas with a particular level or change in house prices. This implies that you will have to think about exogenous variation or an identification strategy including an appropriate benchmark in terms of houses. To give some guidance as to ways how prices are typically modelled see Dröes and Koster (2016), but you are *free to choose your own model, method*. Surprise us.

Writing your report

1. Make sure you write in a clear and concise way. Figures and Tables should be easily readable. More text/tables is not necessarily better.
2. It is allowed to mention in tables that you added control variables, without showing all of the coefficients.
3. Try to avoid too many appendices, if any.
4. YOU ARE NOT ALLOWED TO DEPICT ANY INDIVIDUAL TRANSACTION INFORMATION WITHIN THE REPORT (a spatial plot of the locations of transactions per year is allowed, as well as carefully constructed heat maps).

Potential questions/issues you might want to address could be (bit is not limited to):

- What are you exactly measuring? Is it causal (what is your identification strategy)?
- Is the effect linear?
- To what extent is there selection going on in terms of AIRBNB listings?
- Is the effect on house prices stable over time and the same across types of houses?
- Are there limitations and suggestions for future research?
- In light of your findings what would be the policy implications?

References

Dröes, M.I., Koster H.R.A., 2016. Renewable Energy and Negative Externalities: The Effect of Wind Turbines on House Prices, *Journal of Urban Economics* 96, 121-141.