

Untitled

Paulo Kalkhake & Dan Murphy

27 October 2016

Contents

Proposal Name	1
Abstract	1
I. Motivation & Research Questions	1
II. Literature Review	1
III. Data	1
IV. Methodology identify data sources/methodologies that you can access to help answer your question	2
Bibliography	3

Proposal Name

Abstract

I. Motivation & Research Questions

justifying why the topic of study is interesting

II. Literature Review

providing a basic literature review (properly cited using BibTeX)

III. Data

Our investigation is based on data from three different sources:

First, we use monthly tourism survey data from the Statistical Information System Berlin Brandenburg (PLACEHOLDER CITATION: StatIS-BBB). The data provides reliable information about the current

situation and especially the short-term development of tourism in Germany. The results are organized regionally according to districts, municipalities and also by region. This allows us to have specific data for each of the ten districts in Berlin and thus have a closer look at the geographically specific relationship between hotel occupancy rate and Airbnb supply. The surveys are carried out at the beginning of each month and refer to the reporting period of the previous month.

Second, we use

Third, our paper uses data downloaded from *InsideAirbnb.com*. The data behind *Inside Airbnb* is extracted from publicly available information from the Airbnb site between 18 July 2015 and 6 January 2016. It includes listings from more than 15 cities in 16 countries, among them Berlin. Using data directly extracted from Airbnb’s website or API for our research would of course be preferable, but that data is not disclosed. The extracted data is an appropriate alternative. We have detailed information on all Listings for Berlin in the mentioned time period, such as Calendar Data, Review data, and Listing ID, which allows us to conduct time based analytics.

IV. Methodology identify data sources/methodologies that you can access to help answer your question

Having discussed the data source and the basic properties of the datasets, we propose to approach our research questions using a regression discontinuity model. Contrary to Texas Model, our analysis cannot take advantage of a treatment and control region accounting for the different Airbnb market entry patterns. Mainly, because

Instead, we propose the use of a Regression Discontinuity Model (RDM). By introducing a binary variable which is one for all observation after Airbnb’s market entry and completely interacted with all regressors in our model, we account for all the changes in hotel occupancy rates before and after Airbnb’s market entry in Berlin, i. e. the discontinuity at market entry.

Taking the arguments above into account, our main specification for our proposal would like this:

$$\log OccupRate_{it} = \beta_i * \log Abb_{it} + \tau_t + X + \varepsilon_{it}$$

where $\log OccupRate$ is the occupancy rate for all hotels in district i at time t . $\log Abb$ is the total number of Airbnb listings in \log , and T is a month-year time dummy. Further, we control for economic conditions (unemployment rate and GDP per capita at federal level), a district-specific linear time trend to account for unobserved heterogeneous variation across districts.

Bibliography