CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Analysis of Short-term Rental Market in Prague

Bachelor's thesis

Author: **Karolína Štollová**

Study program: **Economics and Finance**

Supervisor: **Mgr. Barbara Pertold-Gebicka, M.A., Ph.D.**

Academic year: **2019/2020**

Bibliographic note

ŠTOLLOVÁ, Karolína. Analysis of short-term rental market in Prague. Prague 2020, 41 pages. Bachelor's thesis (Bc.), Charles University, Faculty of Social Sciences, Institute of Economic Studies. Thesis supervisor - Mgr. Barbara Pertold-Gebicka, M.A., Ph.D.

Abstract

This paper investigates determinants of daily rate of Airbnb listings in Prague, Czech Republic. Sample of 13 500 properties was examined to identify a relationship between property attributes and rental price using the ordinary least squares estimation method. The study provides an empirical evidence that twenty-five independent variables describing space, reputational, location, commerciality attributes or management policies significantly impact the average daily rate. According to the analysis the most relevant rental price determinants are property location, its size in terms of number of bedrooms, bathrooms and capacity. Author demonstrated in two robustness checks that the results are stable. The study is conducive to better understanding of the Prague Airbnb market. Insights from the analysis could help hosts in developing a suitable pricing strategy as well as Airbnb or similar platforms in designing a pricing tool to increase hosts efficiency.

Keywords

Airbnb, short-term rentals, sharing economy, price determinants, ordinary least squares

Abstrakt

Tato práce se zabývá determinanty denní sazby Airbnb ubytování v Praze. Vzorek 13 500 ubytovacích kapacit byl zkoumán metodou nejmenších čtverců za účelem identifikace vztahu mezi chakteristikami ubytování a jeho cenou. Studie poskytuje empirické důkazy o tom, že dvacet pět vysvětlujících proměnných popisující prostorové, reputační, lokační komerční nebo management charakteristiky ubytování významně ovlivňuje průměrnou denní sazbu. Dle analýzy jsou proměnné zachycující lokalitu ubytování a jeho velikost z hlediska počtu ložnic, koupelen a kapacity obzvláště důležité. Autor ukázal ve dvou kontrolách robustnosti, že výsledky jsou stabilní. Studie přispívá k lepšímu pochopení pražského Airbnb trhu. Výsledky analýzy by mohly pomoci hostitelům při vývoji vhodné cenové strategie, stejně jako Airbnb nebo podobným platformám při navrhování cenového nástroje ke zvýšení efektivity hostitelů.

Klíčová slova

Airbnb, krátkodobé pronájmy, sdílená ekonomika, determinanty ceny, metoda nejmenších čtverců

Declaration of Authorship

I hereby proclaim that I wrote my bachelor thesis on my own under the leadership of my supervisor and that the references include all resources and literature I have used.

I grant a permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 4, 2020

Signature

Acknowledgments

First, I would like to express gratitude to my supervisor, Mgr. Barbara Pertold-Gebicka, M.A., Ph.D., for supervising this thesis. Namely, for supportive guidance, valuable comments and academic insights throughout the whole writing process.

Second, I would like to thank my family for devoted support during my many years of studies.

Bachelor thesis proposal

Author: Karolína Štollová

Supervisor: Mgr. Barbara Pertold-Gebicka, M.A., Ph.D.

Proposed topic: Analysis of short-term rental market in Prague

---

Research question and motivation

There is vast research on the topic of real estate valuation with most of the papers concentrating on predicting apartments values. In this thesis I would like to analyse a specific part of the real estate market, namely short-term rentals. The goal of the thesis is to find out which apartment characteristics are associated with the highest revenues from short-term rentals. The empirical analysis presented in the thesis will be divided into two parts. The first part is considered to determine which apartments make the greatest amount of money and where are they located. Based on this knowledge second part will be devoted to management of these apartments - which variables are linked with most profitable apartments and how they interact with each other. The determinants that will be subjects of investigation may be location, number of bedrooms, number of bathrooms, average daily rate, number of reviews, overall rating etc.

Contribution

As more and more people take advantage of holiday rentals, the whole industry is growing and has become very lucrative. Several works related to short-term rentals have been published recently. Quattrone et al. (2018) analysed Airbnb's spatial distribution in eight U.S. urban areas, in relation to both geographic, socio-demographic, and economic information. Coyle and Yeung (2016) discussed the structure and the segmentation of the accommodation market in fourteen European cities. Existing literature is mostly focused on American cities which have wider experience with short-term rentals. Since Prague market is young and no study with focus on short term rentals in this city has been published yet, the contribution is fundamental. My thesis aims to reveal market structure, describe consumer behaviour, determine most profitable apartments and propose suitable approaches in their management while modeling both the demand and supply side of the market, what has not been done in the literature yet.. Work of Li, Granados and Netessine (2014) on structural estimation from air-travel industry and demand analysis of accommodation of Masiero, Nicolau and Law (2015) provide several methodological approaches which may be useful to build a model to analyse short-term rentals market in Prague.

Methodology

In empirical analysis I will use two data sets. First data set contains information about nearly 23,000 apartments available for short rental in Prague in 2017. General information about each apartment such as its name, average daily rate, annual revenue, number of reviews, number of photos is provided. Second data set is a list of about 15,000 reservations realised over the period from 2014 to 2019. The data comes from 80 apartments managed by one company. In this data set each reservation contains information like check-in and check-out date, date of creation of particular reservation, total payout, number of guests staying, guest's origin etc. as well as basic apartment characteristics.

First data set will be used to build a model relating yearly revenue generated by an apartment to its characteristics. Simple OLS regressions will be used in this sanalysis. Yearly revenues will be placed on left hand side and apartment characteristics on right hand side. I will experiment with different functional forms (quadratics) and interactions to determine relationships as precisely as possible.

To analyse the second dataset I will first build a theoretical model of supply and demand for short-term rentals. This model will serve as the basis to formulation of a simultaneous equations model capturing the interaction between supply and demand in generating daily rental rates. The model will be estimated by two stage least squares.

Estimation results will help understand consumers behaviour of booking apartments for short-term stay in Prague.

Outline

Abstract

Introduction

A. Introduction to short-term rentals environment

B. Overview of similar studies (what has been already studied in different cities)

Discussion about determinants

A. Determinants - what are they?

B. Expected effect of these determinants and how they are expected to interact with each other

Methodology

A. Description of data sets

B. Description of used methods and test procedure

Results

A. Result of test

B. Comments on results

Conclusion

A. Interpretation of results

B. Possible explanation of what leads to such results

C. Recommendation based on results - what flats is best to invest in, how to manage them

D. Comparison with existing literature

List of Tables

List of Figures

Contents

Acronyms

ADR Average Daily Rate

GWR Geographically Weighted Regression

OLS Ordinary Least Squares

QR Quantile Regression

1 Introduction

Sharing economy or sometimes referred to as peer-to-peer is recently created activity of obtaining, giving, or sharing the access to goods and services through online platforms. This economic-technological phenomenon is a result of the fast-evolving industry of information and communication technologies (Kaplan, Haenlein, 2010) and also of growing concerns over ecological and socio-economic impact (Hamari, Sjöklint, Ukkonen, 2016). The sharing economy as a new way of consumption has become a popular alternative of resource distribution and traditional consumption (Wang, Nicolau, 2017).

Sundararajan (2014) describes four forms (non-exhaustively) of peer-to-peer business. These are rental of owned assets (e.g., Airbnb.com for short-term accommodation and RelayRides.com for cars and vehicles), professional service provision (e.g., Uber for professional drivers and Kitchit for chefs), general-purpose freelance labor provision (e.g., oDesk and FancyHands), and peer-to-peer asset sales (e.g., eBay, Etsy).

The sharing economy in the accommodation sector connects people who currently need short-term accommodation with those who rent out property through internet-based platforms. This sector has enjoyed compelling and sustained growth caused by high demand (Heo, 2016; Qiu, Fan, Liu, 2018). Researchers explain the popularity of this phenomenon by several socio-economic reasons (Heo, 2016; Jung, Yoon, Kim, Park, Lee, Lee, 2016). The strongest motivations tend to involve cost reduction and other practical deliberation since peer-to-peer accommodation is often cheaper than traditional hotels, it has become a popular alternative. Experiential motivations such as cultural exchange or intense social interactions with hosts and locals are in general less substantial than economic motivations driven by generating additional income (Guttentag, Smith, Potwarka, Havitz, 2018).

The sharing economy based accommodation is operated on several digital platforms such as Airbnb, Vrbo or Booking by hosts who meet legal conditions to use the property for rental purposes (Heo, 2016). Such accommodation facilities could be found all over the world, they are rented either entirely or shared with the host.

Properties provide guests unique experiences, their types vary from tents to luxury mansions, which corresponds to a wide range of rental prices.

One of the important practices that need to be revised in order to stay competitive in the accommodation business is pricing. According to Yoo, Lee, Bai (2011) revenue management and pricing have been acknowledged as the two most frequently researched subjects in hospitality marketing. Moreover, Hung, Shang, Wang, (2010) consider pricing to be one of the key practices determining long-term success. Despite the troublesome setting of ideal pricing given by the uniqueness of each property, understanding of pricing is necessary in order to obtain important insights that are essential for improving profits and business management (Gibbs, Guttentag, Gretzel, Morton, Goodwill, 2018). To understand pricing is not crucial just from a practical perspective but also a theoretical one.

Since launching in 2008 Airbnb has experienced massive growth and by now it is one of the leading platforms in the accommodation sector of sharing economy. In February 2020 the platform has about six million listings in about 65 000 cities around the world with about two million guests staying in Airbnb each night (Airbnb Statistics, 2020). This enormous growth creates plenty of challenges, great pressure on host's competitiveness and their business requires more professional and systematic business strategies.

Prague Airbnb market is young and tremendously evolving, however, to the best of author knowledge it still remains academically uninvestigated. This study aims to identify price determinants of the Airbnb accommodation platform in Prague. Hence, the contribution of this work to both academic professionals and experts whose work is closely related to this industry is fundamental. In theory, the study sheds light on the relationship between the daily rate of Airbnb listings and property characteristics and is the first to provide comprehensive insights into rental price determinants in the Prague Airbnb market. Understanding pricing would allow researchers to comprehend this accommodation phenomenon and decisions of hosts and guests behind. Practically, the findings could be applied to invent a pricing tool designed to guide hosts through pricing decisions of their properties in order to help them maximise their revenue stream. Moreover, the study has important implications for rental suppliers to analyse and understand in-depth their market situation which is necessary to substantially improve business decisions and pricing strategies resulting in additional profit gained from the knowledge.

This paper analyses data describing Airbnb listings in Prague by means of ordinary least squares regression, econometric method, to examine whether and how various listing attributes affect the average daily rate. Besides common property characteristics such as layout of property, location, capacity, the model includes also information about managing policies and several relevant hosts information.

The rest of the paper is organised as follows. First, section 2 is devoted to a summary of existing literature related to the topic. Section 3 states hypothesis. Section 4 describes data. Section 5 introduces research model and comments on econometric issues. Section 6 checks robustness of the analysis by repeating analysis on specific subsets and employing quantile regression estimation method. In section 7 results of the analysis and robustness checks are presented and discussed. The last section summarises the findings and concludes the entire thesis.

2 Literature review

2.1 Price determinants of Airbnb rental price

The accommodation industry of sharing economy, specifically Airbnb, has expanded in the past decade, because of the growing demand (Karlsson, Dolnicar, 2016). Plenty of studies were conducted to explore and understand the sharing economy based accommodation. Given the research topic of this thesis, literature review section is mostly concerned with price determinants of Airbnb properties, however, other research topics related to Airbnb properties are presented as well. The first part of this chapter is devoted to findings and outcomes of studies investigating price determinants in the short-term accommodation industry, second part summarises related existing literature from a methodological perspective.

Besides the investigation of price determinants, researchers studied aspects of sharing economy based accommodation such as existence and extent of discrimination (Cheng, Foley, 2018; Cui, Li, Zhang, 2017; Kakar, Voelz, Wu, Franco, 2018; Edelman, Luca, Svirsky, 2015). To create a profile and upload a profile picture make it easy to discriminate due to incomplete profile information, cultural background, minority community affiliation, etc. (Edelman, & Luca, 2014).

Impact on the traditional hotel industry and competitiveness of Airbnb with respect to the traditional hotels have been recently frequently analysed topics, however, there is no consensus among researchers. Choi, Jung, Ryu, Kim, Yoon (2015) concluded that Airbnb's presence in the market has no significant impact on hotel revenue, supported by findings of Mohamad (2016) which, despite substantial growth of Airbnb over past years, provide evidence that hotel performance is not affected by Airbnb. Mody, Suess, Dogru (2017) argue that hotels and Airbnb have different target customers, however, they acknowledged that there might be a negative impact on hotels, nevertheless, very marginal. Coyle, Yeung (2016) even claimed that Airbnb positively influences hotel revenue and average daily rate. However, the following studies challenge the above stated by contrary evidence. According to Zervas, Proserpio, Byers (2017), Airbnb is an alternative for certain traditional hotels, thus has a negative effect on hotel revenues. The same conclusion was reached by Guttentag, Smith (2017).

Gentrification and impact on housing prices are often discussed in conjunction with Airbnb. Given a greater profitability of Airbnb over long-term rentals, affordability of housing for locals has declined (Yrigoy, 2019). Horn, Merante (2017) claimed there is a positive correlation between a number of Airbnb listings and rental rates. The same conclusion was reached by Wachsmuth, Weisler (2018) who explained that increasing rents occur due to higher demand for use of the land. Barron, Kung, Proserpio (2018) found a positive correlation also between a number of Airbnb listings and housing prices.

A few studies have examined motivation to use short-term rental accommodation, however, results describing consumer's motivation vary to a great extent. According to Guttentag, Smith, Potwarka, Havitz, (2018), in general, guest value practical attributes more than experiential ones. However, Mao, Lyu (2017) place unique travel experiences as one of the top motivations of travellers. Generally, guests are motivated by local interactions, authentic experiences, familiarity. Also saving accommodation costs is important since with cost reduction more expensive destinations and activities come into guest considerations (Tussyadiah, Pesonen, 2016; Möhlmann, 2015). There is no dominant motivation for locals to choose to supply Airbnb accommodation and host guests (Karlsson, Dolnicar, 2016). Hosts participate in the business for both financial and social reasons (Ikkala, Lampinen, 2015). Contrary to the expectations, financial benefits rather strengthen intrinsic motivations than displace them (Lampinen, Cheshire, 2016).

Several researchers have initiated the efforts to examine also the price determinants of accommodation in the sharing economy. However, these studies have a rather limited focus on the impact of certain factor such as location or credibility on rental price than complex insight into a set of determinants affecting price.

Tang and Sangani (2015) in their research focused on the market, where Airbnb was first introduced, San Francisco. They applied a supervised machine learning method using listing's characteristics to identify variables that could sufficiently predict besides property's price also its neighbourhood. Such method allows researchers to understand correlations between neighbourhoods and prices. Their findings imply that higher-priced listings could be in general found in upscale neighbourhoods.

Zhang, Chen, Han, Yang (2017) conducted two models, a general linear model and a geographically weighted regression model, to identify price determinants of listings prices in the metropolitan area of Nashville. The linear model proved that distance to the conference centre, popular sight, number of reviews and rating score are significant determinants of prices. Authors argue that geographically weighted regression suits the analysis better because it potentially allows magnitudes of variables to vary across different city areas and, therefore, it covers spatial heterogeneity. Their findings provide evidence that the prices of properties are more sensitive to the distance from the conference centre in the central area. Even a short distance in the centre can have a greater negative impact on price than a larger distance outside the centre, where the effect of distance is according to the study more negligible.

Because reviews as a form of reputation and trust could be converted into economic value, several researchers dedicated their efforts to examine the role of reviews and rating on the rental price. There is a common strategy among hosts to adjust their pricing strategy by increasing prices as a number of positive reviews grows larger. However, there is a reason why some hosts decide to set prices of their property below market level and that is an opportunity to choose their guests from a larger set of potential visitors (Ikkala, Lampinen, 2014).

Gutt, Hermann (2015) were asking a question how a review system affects prices on Airbnb in New York City. Such system of online reviews reflects experiences of previous guests with their stay and therefore demonstrates a quality of the service. Assuming there is economic motivation behind host’s participation in the sharing economy, the auspicious review should have a positive impact on price and allow hosts to take extra money from the market. Researchers found a robust effect of rating on prices and proved that hosts in NYC transfer credibility, demonstrated by rating, into extra money. More precisely, after a host’s rating is publicly available, that happens as soon as the host receives three reviews, the price could be increased by €2.69 without a drop in occupancy rate.

The topic of economic value of trustworthiness and credibility measures is widely elaborated by Teubner, Hawlitschek, Dann (2017), who focused in their study on the effect of reputational attributes such as rating score, number of reviews, duration of platform membership, number of listing's photos and Airbnb's Superhost status on listings prices. They believe that a financial performance of hosts is determined by their current image on the platform. Their findings imply that a longer duration of platform membership and uploading a greater number of apartment photos are associated with price mark-ups. Despite star rating scale is very subtle and its variation negligible, there is positive, significant and consistent influence of rating score (i.e., more stars) on price. While improving rating results in an increase in price, a greater number of reviews is associated with a decrease in price. Researchers suggest that causal direction in the relationship between price and number of reviews may work also in the opposite direction and lower prices may actually boost demand and generate more reviews.

Ert, Fleischer, Magen (2016) showed in their analysis that rating score of an apartment has no impact on their listing's market price given by the negligible variance in rating score, which is practically indistinguishable since great majority of scores are within 4.5 to 5 stars (with 5 being the best). To find a measure of trustworthiness instead of reviews, researchers investigated an effect of host's photos. In their study concluded that hosts whose photo was evaluated by respondents as more trustworthy charge higher prices, in other words, trustworthy photos result in price markups, moreover, demand for their properties is higher with a greater probability of being booked. The implication of this study is that people are willing to pay more for properties managed by trustworthy hosts rather than properties with excellent rating score, furthermore, this holds even if rating score varies.

Fagerstrøm, Pawar, Sigurdsson, Foxall, Yani-de-Soriano (2017) came up with an even deeper analysis of host's personal photos. Their study aims to determine effect of this photo, more precisely an impression of the photo in general and host's facial expression, on guests behaviour in an online peer-to-peer context. Since the host's picture is in this market used to develop relationship between host and guest and to build trust between them, the picture has a great importance on business performance. Researchers found that negative (angry) facial expression or complete absence of personal photo decrease approach behaviour on Airbnb. On the contrary, a picture of a host with neutral or positive expression has a reverse effect. Study shows that facial expressions have a differential impact on female and male guests. Furthermore, analysis proves that excellent rating score or low price cannot compensate for negative facial expression or absence of the photo.

Pricing is an important skill which has to be mastered by every host to stay competitive in the market (Hung, Shang, Wang, 2010). As both professional and nonprofessional hosts constitute the supply side of Airbnb, affiliation to one of the groups may be important price determinant. Study of Li, Moreno, Zhang (2015) focused on discrepancies in behavioural actions, pricing strategies and their impact on prices among these two groups. Researchers found considerable differences in decision-making. Particularly, professional hosts charge higher prices, achieve greater occupancy and probability of them leaving the market is significantly lower in comparison with their non-professional counterparts. Non-professional hosts rarely adjust prices when a specific date is not rented out and the check-in date is only a few days ahead, or they are more likely to keep same price level even across dates with extraordinary demand such as festivals or holidays. In general, findings are explained by non-professionals being substantially inefficient and suffering from behavioural biases (e.g., loss aversion, limited attention, overconfidence).

Dogru, Pekin (2015) contributed to studies identifying price determinants by several important findings. Results of their work show that space, cleanliness, number of photos, handicap accessibility, family friendliness, free breakfast, location, and unique experiences have a significant effect on pricing. Host, who require guests to pay besides accommodation fee also for cleanliness, charge higher prices in comparison with properties without this fee. Providing free breakfast is according to the study another way how to increase the price. The fact that unique properties such as yachts, caves, treehouses are significantly higher priced may suggest that guests demand remarkable experiences also from their holiday accommodation. Negative impact on price is observed in the case when properties seem to be rented for business purposes.

Dogru, Pekin (2017) studied factors related to the property characteristics, amenities, services and rental rules. Their study reveals several important facts. Even though Airbnb is in general perceived as a social platform, where participants are motivated by social interactions, more than for economic benefits, guests pay for privacy demonstrated by the finding that entire homes and private rooms prices are much higher than shared rooms given by privacy that guests value. Findings are consistent with Teubner, Hawlitschek, Dann (2017) and prove that rating and more apartment pictures have positive influence on prices, whereas a greater number of reviews has an adverse effect. Furthermore, researchers found out that listings managed by a host with Superhost status are higher-priced. They also estimated effect of badges like handicap accessible, family-friendly, suitable for events and business-ready, it turned out that all badges have a positive influence on price with exception of business-ready which has according to the study negative effect on price.

Wang, Nicolau (2017) provides very complex insights into price determinants of prices of Airbnb properties. They studied effects of 31 variables describing properties located in 33 cities across the world on the rental price. From methodological perspective, researchers used ordinary least squares regression and quantile regression which allow them to study differences between these two methods and, more importantly, obtain more comprehensive insights into the distribution of variables and reveal patterns on a different level of rental prices.

Host attributes like Superhost, number of host's listings, host verification of identity lead to significantly higher prices. Whereas host's profile picture is associated with negative impact on listing price. Regarding site & property attributes, variable *distance* has significant negative coefficient, which proves that increasing listing's distance from the central area has negative impact on price. The *Entire home* and *private room* variables have a positive significant effect on listing price Moreover, a daily rate of property is higher if it accommodates more guests and offers additional bathrooms and bedrooms. Amenities like parking, wireless internet, more real beds have auspicious effect on daily rate. However, services like including breakfast in a daily rate or allowing guests to book an accommodation instantly lead to significantly lower listing prices. Rental rules such as non-flexible cancellation policies and phone verification result in price mark-ups. Requirement of profile picture has according to analysis no impact on rental price. Smoking permission leads to significantly lower rental price. Lastly, impact of online review system was examined. Number of reviews received per year yields in a negative effect on price, each additional review obtained causes price to drop given by the fact that lower-priced properties tend to receive more reservations, therefore, more reviews. Regarding the score of the reviews, it positively affects accommodation price.

2.2 Methodological review

From a methodological perspective, researchers investigating price determinants in hospitality industry frequently base their analyses on hedonic price modelling. Hedonic price modelling allows the observed price of product to be formulated as an additive function of various utility-bearing attributes. In accommodation research hedonic pricing model is widely used since this framework can effectively capture effect of different property amenities like free parking, distance to the city centre, availability of swimming pool, kitchen or hairdryer on price. To estimate such models is commonly employed ordinary least squares regression or a similar technique (Sánchez-Ollero, García-Pozo, Marchante-Mera, 2014).

Traditional least squares regression is widely used in the accommodation sector of the sharing economy to quantify price determinants. In the hospitality industry researchers often employ logarithmic transformation of the dependent variable (price) and estimate the log-level model rather than a simple level-level model (Rosen, 1974; Schamel, 2012). Semilogarithmic models are frequently used in literature because in general they reduce heteroskedasticity and simplify interpretation of estimates (Perez-Sanchez, Serrano-Estrada, Marti, Mora-Garcia, 2018). Several researchers employ a rather quantile regression than traditional least squares regression, this is due to limitation of least squares regression which focuses solely on the conditional mean, whereas quantile regression allows researchers to obtain fully representative conditional distribution overview and comprehensive insight into relationship between dependent and explanatory variables (Wang, Nicolau, 2017).

Other methods of evaluation effects of variables on price are rare. However, Zhang, Chen, Han, Yang (2017) employed geographically weighted regression to determine effects of factors affecting price. The dependent variable in their model is explained by multiple independent variables in which coefficients can vary spatially. Some variables which were not significant in the general linear model, showed significance in the GWR model. Also the GWR model has more than two times greater adjusted R squared.

Tang and Sangani (2015) used a supervised machine learning method, more precisely support vector machine with a linear kernel, which was able to successfully predict price and neighbourhood of properties after training the model. However, their model suffered from overfitting given by a limited number of training sets, this issue was at least partially solved by feature selection using recursive feature elimination.

In hospitality industry studies which deal with reputational attributes build their reasoning on signalling theory. This theory describes signals as instruments through which host declares quality of his property. Therefore, their main purpose is to built trust between the two market sides. Signals in this context could be ID verification, rating score or number of photos. Teubner, Hawlitschek, Dann (2017) built reasoning of their study which investigates effects of several reputational attributes on price using signalling theory.

3 Hypothesis

This study aims to determine which property characteristics have effect on the average daily rate. If the effect of given characteristic is non-zero then its sign and magnitude are analysed. In this section, hypothesis of the study is stated. The hypothesis is formulated as the predicted impact of a property characteristic on the average daily rate. Namely, what relationship is expected between each property attribute and ADR. The detailed expectations of the relationship are presented in section 4.5. The null hypothesis is defined as:

$$H_0: ßj = 0,$$

where ßj stands for a parameter estimate of a jth property characteristic. The null hypothesis claims that the jth property characteristic does not impact ADR. The alternative hypothesis is:

$$H_a: ßj \neq 0,$$

where ßj is defined exactly as specified above. The alternative hypothesis states that a jth property characteristic impacts ADR. If analysis reveals that estimate of jth property characteristic has the sign as expected, the null hypothesis is rejected in favour of parameter expectation. On the contrary, if sign of the estimate is opposite than expected, the null hypothesis is rejected opposed to parameter expectation.

4 Data

4.1 Data description

In this thesis, Prague, Czech Republic, was chosen as a destination of interest. The cross-sectional data containing information about all properties listed on Airbnb located in Prague at the date of retrieval was obtained from third-party provider of short term rental data, AirDNA, which assembles publicly available data from Airbnb (About AirDNA, 2020).

Originally, the dataset contained 21 767 observations where each observation represents one Airbnb listing located in Prague in October 2017. The dataset provides detailed information about each listing including host identification number, general property information like its title, ID, type, date of creation, further, listing financial figures such as total revenue or average daily rate, property spacial characteristics represented by number of bedrooms, bathrooms or maximum guests that could property accommodate, moreover, GPS coordinates, reputation characteristics like rating score or number of reviews and several management policies like cancellation policy or minimum stay length requirement. Fixed variables such as number of bedrooms or location of property remain constant over time, however, data varying over time like the financial information are calculated for the last twelve months before the date of data retrieval.

4.2 Data cleaning

13 500 observations remain in the dataset after completion of data cleaning. The original dataset was modified for the purpose of this study in the following manner. During data cleaning observations with less than one-day duration of platform membership were deleted because no actual transactions took place, therefore, such listings do not contribute to analysis which is motivated to explain variation in the average daily rate. Since the focus of this paper is on properties preferably targeting short-term stays over long-term stays, properties requiring a minimum stay of more than seven days were excluded. Finally, 5 525 observations in total were dropped from the dataset due to the reasons stated above.

Additionally, another 2 742 observations were deleted due to missing values. 95% of the total missing values were missing information in the *OverallRating* variable describing rating score of a property. When subset of missing values is pulled out of the whole sample, summary statistics on the subset remains very similar to the summary statistics on the whole dataset. Major changes happened in variables capturing number of reviews and property duration on the Airbnb platform. First, average of *Reviews* dropped from 33 to 0.1, second, average of *AirbnbDuration* decreased from 626 to 390 days. According to the characteristics of observations with missing values stated above it could be concluded that they are typically newly created listings on the Airbnb which have not obtain review yet, therefore, they lack rating score. Moreover, when baseline analysis as described later in this study is proceeded without the *OverallRating* variable, that is with the deleted observations, coefficients of OLS estimates change only negligibly, hence it is deduced that newly created listings on Airbnb do not substantially affect the outcome of the analysis.

As stated above, several observations were deleted, however, the process of data preparation included also contemplation over specification of variable forms prompting several other adjustments of the dataset to exploit the character of variables exhaustively. Total number of reviews is included in the final specification in a logarithmic form because this form suits the data the best and maximises adjusted R squared of the model. However, in the dataset are plenty of properties that have never obtained a review, since the real logarithm of zero is not defined this problem had to be solved. The approach of adding a very small number (0,000001) to the zero was used to estimate effect of reviews on rental price. Moreover, several variables (type of property, number of guests per one bedroom, bedrooms, bathrooms, minimum stay length, number of properties managed by one host, cancellation policies) were chosen to be placed in the model in dummy form, therefore, plenty of new dummies were generated to analyse the data. There are two possible reasons why the dummy form was selected for these variables. First, variable is categorical. Second, histogram of numerical variable suggested that its values gather in number of clusters. Hereupon, dummies were specified intuitively and later analytically enhanced to both sign and magnitude of coefficients make sense, moreover, several sensitivity checks were proceeded to maximise adjusted R squared.

4.3 Sample variables

In this study, we would like to understand how the rental price of a property interacts with its characteristics. The explained variable is the rate for renting out a property for one night stay averaged out for the last 365 days before the date of data retrieval. It is expressed in American dollars. ADR was chosen as a dependent variable because the aggregation at yearly level enables to measure relatively homogenous period, moreover, the month-of-year effect is consequently eliminated.

Based on the up-to-date literature and preliminary data inspection 28 variables were chosen as potential determinants of ADR. Control variables capture various characteristics of the property from space, reputation, location, commerciality, and management perspective. Space attributes cover the impact of type of a property, its size, and comfort on ADR. Reputational measures represent effect of trustworthiness and credibility of property on the Airbnb platform on ADR through reviews, rating score, photos, Superhost status, and duration of Airbnb listing membership. Information of listing location is included to capture how distance from popular touristic attractions in Prague to property affects the ADR. Commerciality attributes represent three attributes. First, whether listing is suitable for business travellers in terms of equipment and furnishings provided. Second, whether property could be booked instantly without a required permission of host. Third, how many properties one host manages. The last set of characteristics is related to management policies, that is setting of cancellation policy and whether and how long minimum stay length is required. Table 1 placed in the appendix lists and defines all variables used in the analysis in detail.

4.4 Summary of sample variables

In this section, the sample variables are briefly summarised. To present summary of explanatory variables comprehensibly, they were split into five groups. They are space, reputational, location, commerciality attributes, and management policies.

The dependent variable, average daily rate, ranges from \$4 per night to \$1113 with average value of \$79. Regarding space attributes, 77% of the sample are entire homes suggesting there is great demand for privacy from guests side, private rooms represent 22% of the market and the remaining are shared rooms. Great majority of properties have less than two bedrooms and one bathroom. Regarding convenience of property expressed in number of guests per one bedroom, about half of listings have up to three guests per one bedroom.

The second group of attributes are reputational attributes. Properties have on average 33 reviews. Rating scores of properties vary from one to five, however, after closer examination, 75% of properties have this score between 4.5 to 5 stars signalling remarkably low variance of this variable which is consistent with literature. 25% of hosts meet criteria to be awarded Superhost badge, that means hosting at least 10 days in a year, response to booking requests quickly and in at least of 90% of all cases, 80% of ratings are 5-star ratings and lastly, hosts rarely cancel confirmed reservations. On average, Airbnb hosts posted 18 photos of their properties. Majority of listings have been active on the platform for more than a year with a mean value of nearly two years.

Half of the properties are located up to two kilometres from the Old Town Square, the popular sight in Prague, however, the variable ranges distinctly from 0.02 to 19 kilometres. Median values for the other two well-known tourist attractions Wenceslas Square and the Prague Castle are 1.8 and 2.8 kilometres, respectively. Distances from properties to each of the two attractions vary similarly as to the Old Town Square, which is from 0 to about 20 kilometres.

Commerciality attributes are summarised as follows: 13% of listings are properly equipped with amenities like laptop-friendly workspace or Wi-Fi to meet requirements of guests travelling for business purposes. Further analysis shows that nearly half of the dataset enables guests to book their accommodation instantly. Half of property managers manage up to three listings, however, several hosts manage tens of listings which drives average of number of listings managed by one host up to eight properties.

Regarding cancellation policies, each setting option reaches about one-third of the sample so it could be concluded that hosts who rely on income and are motivated by financial incentives are equally represented in the sample as hosts who do not mind guests to cancel their reservations in the last minute. Nearly 40% of hosts do not put any restriction on minimum stay length, about the same percentage of hosts require at least two days stay.

Table 4.1: Summary statistics of variables

Variable	Min	Max	Mean	Standard deviation	Median	Proportion
ADR	3.9	1112.5	78.6	68.3	60.1	-
Space attributes						
EntireHome			-			76.5%
PrivateRoom			-			21.7%
Bedrooms2			-			22.6%
Bedrooms3			-			6.4%
Bedrooms4			-			1.9%
Bathrooms1,5			-			13.5%
Bathrooms2			-			9.3%
PerBedroom2			-			45.9%
PerBedroom3			-			22.3%
PerBedroom4			-			27 %
Reputational attributes						
Reviews	0	429	33	46	15	-
OverallRating	1	5	4.7	0.42	4.8	-
Superhost			-			25 %
Photos	1	134	18	12	15	-
AirbnbDuration	31	3088	626	478	500	-
Location attributes						
OldTownSquare	0.02	19	2.5	2.2	2	-
PragueCastle	0.1	20.4	3.3	2.2	2.8	-
WenceslasSquare	0.005	18.7	2.4	2.1	1.8	-
Commerciality attributes						
BusinessReady			-			13 %
InstantBook-Enabled			-			52 %
Listings3			-			30.8%
Listings11			-			19 %
Management policies						
Moderate			-			34.2%

Strict	-	36.4%
Stay2	-	45.2%
Stay3	-	13.8%
Stay4	-	3.3%

4.5 Expected impact of determinants on the average daily rate

Reasoning of expectations is based on hedonic price modelling. The approach assumes value or quality of good or service, frequently it is a price, as an additive function of multiple utility-bearing attributes. Price is determined by a set of attributes which are individually evaluated by the market. In this case, the study aims to explain the average daily rate through a set of property characteristics such as its size, capacity, distance from touristic places. If an attribute is expected to be generally considered as auspicious and valuable to guest, the attribute is argued to impact rental price positively and vice versa.

Expectations of price determinants are presented in the same five groups as in the previous section. These are space, reputational, location, commerciality attributes and management policies.

Firstly, regarding space attributes describing size and capacity of the apartment, it is generally expected that properties which are rented entirely without host living in the same apartment or even the same room have on average higher rental price. Next, the more spacious and capacious apartment is, the higher average daily rate is charged. Moreover, the more guests are accommodated in one room, the lower comfort is provided and thus the effect on rental price is negative.

Reputational attributes represent trustworthiness and credibility measures such as number of reviews, overall rating score, number of photos, Superhost status, number of photos and variable concerning listing duration of platform membership. Enhancing reputation through improving or obtaining the previously mentioned credibility variables indicates greater reliability of the service provided, therefore, could be converted into positive economic value. However, based on existing research reverse causality is expected to be present in relationship between number of reviews and rental price. Apart from number of reviews influences daily rate, also price charged by owner is assumed to impact the number of reviews. If a listing is low-priced, it becomes more popular among guests and thus the property is almost constantly occupied resulting in more reviews (Teubner, Hawlitschek, Dann, 2017). Therefore, negative coefficient of the variable is expected.

Location attributes cover distance from property to important city landmark or a popular touristic attraction. It is assumed that guests would like to be accommodated close to a touristic sight, therefore, as distance increases, the average daily rate decreases, thus relationship between these variables is expected to be negative.

The fourth group of attributes are commerciality attributes: Business-travel-ready, instant book enabled and number of listings per host. Business travel ready is expected to have positive impact on ADR since it guarantees extra amenities. Instant booking brings additional convenience for guests during booking, therefore, correlation with ADR is expected to be positive. Lastly, professional hosts represented by greater number of properties managed are more efficient in pricing than their non-professional counterparts (Li, Moreno, Zhang, 2015) and thus positive relationship is expected.

Last, management policies such as cancellation policy and required minimum stay length are possible price determinants. The more strict cancellation policy, the greater ADR is. The intuition behind the idea is following. A host who set cancellation policy as strict does not want guests to cancel their reservations, therefore, he/she is likely to care more about business and is motivated by economic incentives. This economic motivation demands to master pricing techniques and yields into ability to maximise daily rates. On the other hand, if cancellation policy is set as flexible, host arguably does not care about revenue inflow much and is rather motivated by social incentives, hence pricing techniques are not proficient and thus rates charged are on average lower. It is expected that setting minimum stay length on two days positively affects ADR compared to no restriction on minimum stay length as a reference group. This expectation is based on the idea that such restriction could help to achieve higher occupancy rate together with higher rates. However, stronger restrictions throughout a year are expected to affect rate inauspiciously because longer stays are too limiting for guests (Sims, Ameen, Bauer, 2019).

5 Research model

5.1 Hedonic pricing model

Empirical analysis is based on hedonic pricing approach, modelling price of a good or service as an additive function of multiple utility-bearing attributes. This approach allows to capture effect of each attribute and amenity on the explained price variable, this is exactly alike with this thesis that aims to explain average daily rate through set of variables describing property characteristics, therefore, in case of this study this framework is pertinent and apposite. Study adopts ordinary least squares method to estimate the model.

5.2 The baseline model

The purpose of this study is to understand how the average daily rate from renting out a property interacts with a set of independent variables related to its characteristics and attributes. The model describing such relationship was constructed as follows:

$$\log(ADR_i) = ß_0 + \mathbf{Xß} + u_i, \qquad (1)$$

where $\log(ADR_i)$ is the natural log transformation of the average daily rate for listing i, i = 1,…,n indicates particular observation from up to n, $ß_0$ represents intercept, ß stands for parameter estimate, X represents a vector of explanatory variables and u_i is error term for ith observation.

The average daily rate is calculated by dividing the total revenue earned by the host within the last 365 days before the data retrieval by the number of booked nights within this period. Following researchers like Li, Moreno, Zhang (2015), Wang, Nicolau (2017), Teubner, Hawlitschek, Dann (2017) logarithmic form was chosen as the best form of average daily rate, the response variable, to describe its behaviour thoroughly. Moreover, logarithmic form facilitates interpretation of relationships showing how ADR responds in percentage terms to changes in regressor while keeping all other predictor variables fixed.

Vector of independent variables, denoted by X in equation 1, consists of predictors describing space, reputational, location, commerciality attributes, and management policies related to each property. Explanatory variables are included in the model (1) in several specification forms given the character of each variable (linear, logarithm, quadratic).

To estimate regression parameters of the baseline model (1) ordinary least squares regression technique is adopted as estimation method. Estimates are obtained by minimising the sum of squared residuals. In general, OLS regression allows estimating average response of the dependent variable to the changes in explanatory variables. In this study, OLS is employed to estimate relationships between property and host characteristics and the ADR, which is specified in the model (1).

5.3 Econometric issues

5.3.1 Heteroskedasticity

In this section, the heteroskedasticity issue is formally addressed. The ordinary least squares estimation method assumes homoskedasticity, also known as constant variance. This assumption requires the variance of an error term, conditional on regressor, to be constant. If this assumption is violated, that is the error does not have the same variance given any value of the regressor, then the error is said to exhibit heteroskedasticity. If heteroskedasticity is present in error term, then

$$Var(u_i|x_i) = \sigma_i^2,$$

where i subscript on σ_i^2 denotes that variance of the error u_i is heterogeneous across observations. Even though heteroskedasticity does not cause the OLS estimators to be biased or inconsistent, OLS standard errors are not accurate for calculating t statistics (Wooldridge, 2012). To check whether errors contain heteroskedasticity, the Breusch-Pagan test was run. If the test confirms heteroskedasticity, then standard errors of estimated regression coefficients will be adjusted by using heteroskedasticity-robust standard errors.

5.3.2 Endogeneity

There is a potential threat of endogeneity in the model caused by reversed causality. If disturbance is correlated with particular explanatory variable, the variable is called an endogenous explanatory variable. If endogeneity is present, OLS can produce biased and inconsistent parameter estimates. Furthermore, no causal interpretation of the estimates can be made (Wooldridge, 2012).

For instance, variable describing number of reviews may be determined by probability of apartment being booked which could be affected by price itself. In this case, a downward bias is expected, i.e. the OLS estimate would be lower than the true causal relationship of number of reviews on ADR. This is because the reverse causality goes in this direction with lower price attracting more customers and consequently generating more reviews. Despite the possibility of this threat, including reviews is common practice in hospitality industry analyses (Teubner, Hawlitschek, Dann, 2017; Dogru, Pekin, 2017; Wang, Nicolau, 2017).

The author is aware that endogeneity may be present in the model, unfortunately, is not able to deal with it by any means. Despite endogeneity, to the best of author knowledge OLS is the most appropriate estimation method for the baseline model.

5.3.3 Normality of population errors

One of the classical linear model assumptions is normality. This assumption requires random errors to be independent of explanatory variables and normally distributed with zero mean and variance σ^2. If normality assumption is not satisfied, estimates of independent variables are still BLUE (best, linear, unbiased), nevertheless, t statistics and F statistics do not follow t distribution and F distribution, respectively. In such case, we rely on the central limit theorem to conclude that estimators fulfil asymptotic normality in reasonably large sample sizes. If there are enough observations to satisfy the approximation of the central limit theorem, t testing is proceeded exactly the same way as under the classical linear model assumptions and the analysis remains unchanged (Wooldridge, 2012).

To test normality, the Jarque-Bera test was run. The null hypothesis claims that sample data have the skewness and kurtosis matching a normal distribution. In this case, p-value in the test is below threshold indicating that data are inconsistent with the null hypothesis, therefore, normality in the baseline model does not hold. Given estimators being approximately normally distributed, exact hypotheses of t test and F test can be carried out (Thadewald, Büning, 2007)

6 Robustness checks

OLS method used in the baseline modelling is a linear mean-based model that provides the average estimates by pooling the sample together. It assumes that the relationship between dependent variable and one specific independent variable is the same across all values of other explanatory variables. Moreover, OLS extrapolates the data. Even if a particular combination of control and response variable is not observed in the dataset, OLS predicts the specific relationship for these values.

To test whether relationship between listing characteristics and ADR is stable across observations, two robustness checks were employed. First, the baseline analysis was repeated, however, three specific subsets of properties serve as data inputs. Second, quantile regression was adopted to obtain more detailed insights into the distribution of the dependent variable and to test whether results of the baseline analysis are robust. Detailed motivation to employ quantile analysis of the model is provided later in this chapter.

6.1 The baseline analysis on specific subsets of properties

To test whether results are stable and still valid given by a change in input data, a robustness check was employed. The baseline analysis was proceeded, however, three specific subsets of properties were used as data inputs. Purpose of the check is to examine whether coefficients of independent variables using each subset and the whole dataset as the model inputs are alike. That is to investigate whether results are stable across properties.

The first subset captures professional hosts. Such hosts are expected to manage at least ten listings, have already automatised their booking process, rent entire homes to make the business more profitable, set cancellation policy as strict in order to not lose revenue inflow due to booking cancellations and care about impression and credibility of each property resulting in overall rating score at least 4.5 star. The subset consists of 727 observations.

The second subset is concerned with luxury properties and enables to investigate whether this specific subset of properties has significantly different price determinants. For the purpose of the robustness check are luxury properties characterised as entire homes located up to three kilometres from the Old Town Square offering adequate comfort with a maximum three guests per one bedroom. Moreover, their host is awarded as Superhost to ensure that property has top rating score and provides a convenient reservation process. The subset consists of 2 086 observations.

The last subset consists of seasonal listings that were rented only less than 150 days in the past twelve months before data retrieval. This subset was created to reveal whether listings which are rented only a fraction of the year, often during summer season when daily rates are higher, have different price determinants. Recently created properties with Airbnb membership duration shorter than 365 days were omitted to ensure that potential difference in coefficients will not be due to effect of newly established listings. The subset consists of 3 015 observations.

6.2 Quantile regression

6.2.1 Motivation

The motivation behind employing the quantile regression into this study is to test whether relationships between each listing characteristics and the average daily rate are stable across observations and to obtain more detailed insight into the distribution of the dependent variable. Quantile regression was applied to address the following issues specifically.

Firstly, as argued in the endogeneity section, the relationship between number of reviews and ADR might go in both directions. Intuitively, it is expected that more reviews could be converted into positive economic value because with additional review more trustworthiness and credibility is built. However, there is potential threat of endogeneity which may be caused by reversed causality. The less expensive listing is, the more attractive for bookings becomes, consequently such properties may obtain more reviews. This could be potentially addressed using QR which estimates the relationship for each conditional quantile of the ADR and thus we might observe different relationships for low-priced properties and different for high-priced properties.

Second, in the baseline analysis some coefficients might show as being insignificant not because their effect is actually null, but due to heterogeneity of their effect which consequently cancels out on average. QR estimates the model in a greater detail, therefore, it may reveal conditional quantiles where also variables estimated as insignificant by OLS actually impact ADR. Such heterogeneity is expected to potentially affect those variables which are likely to have different effect on low-priced properties than high-priced ones. Perhaps this may be a case of dummies capturing minimum stay length requirement. The idea behind the expectation is that guests who booked expensive properties typically tend to spend less nights in the property, therefore, adding extra requirement on the minimum stay length could actually lead to lower demand and negative effect on rental price. On the other hand, if guests would like to stay at least couple of days or even a week, they are likely to choose affordable accommodation, thus cheaper properties could be actually better off if apply restriction on the minimum stay length.

6.2.2 Theory

Koenker and Bassett (1978) firstly introduced quantile regression which allows estimation of conditional quantile functions capturing each point in the distribution, therefore, describing whole conditional distribution. Simple regression model for quantile level θ of the response is defined as:

$$Q_\theta(y_i) = \beta_0(\theta) + \beta_1(\theta)x_{i1} + \beta_2(\theta)x_{i2} + \ldots + \beta_k(\theta)x_{k2},$$

where $i = 1,\ldots,n$; $x_{i1},\ldots,x_{ik}$ and $\beta_1(\theta),\ldots,\beta_k(\theta)$ represents particular observation from up to n, the vector of regressors, the vector of estimates relating to each θth quantile, respectively. Coefficients of regressors are estimated by solving following the minimisation problem:

$$min \{ \Sigma\, \theta\, |y_i - \beta_0(\theta) - \Sigma\, x_{ij}\beta_j(\theta)| + \Sigma\, (1-\theta)\, |y_i - \beta_0(\theta) - \Sigma\, x_{ij}\beta_j(\theta)| \} =$$
$$= min\, \Sigma\, \rho_\theta\, (y_i - \beta_0(\theta) - \Sigma\, x_{ij}\beta_j(\theta)\,),$$

where ρ_θ is called check function which gives different weights to positive and negative residuals. Sample median could be defined as a product of minimising a sum of absolute residuals, thus, in the minimisation problem has to be the same number of positive and negative residuals along with the same number of observations above and below the median value. This symmetry results in median, to obtain other quantiles than the 50th one (the median) one should simply weight positive and negative residuals differently so the problem of minimising a sum of absolute residuals yields into minimising a sum of asymmetrically weighted absolute residuals. Such asymmetry could be achieved with the check function (Koenker, Hallock, 2001).

Employing the quantile regression has several useful advantages. By estimating effect of each independent variable on different parts of the distribution of the conditional average daily rate, researchers are able to reveal hidden patterns on different quantile levels including end tails of the distribution (Hung, Shang, Wang, 2010). Moreover, according to Buchinsky (1998) quantile regression produces more efficient estimates than OLS when residuals are not normally distributed. Furthermore, as OLS estimates come from minimising the sum of squared residuals, median estimates (50th quantile estimates) come from minimising the sum of absolute deviations, thus the latter is not affected by outliers or other extreme values.

7 Empirical results

This thesis investigates determinants of the average daily rate of Airbnb properties in Prague by the means of OLS regression. In this chapter results of the baseline analysis are presented. In addition, two robustness checks of the baseline model were conducted, baseline analysis on specific subsets of properties, and quantile regression, their results are discussed later in this section as well.

7.1 The baseline model

In this section are presented results of the baseline analysis. The plain outcome of the baseline analysis could be found in table 2 in the appendix. Unfortunately, heteroskedasticity as discussed in the chapter devoted to research model was detected by the Breusch-Pagan test, therefore, heteroskedasticity-robust standard errors were computed. Propitiously, the significance of explanatory variables did not change under heteroskedasticity-robust standard errors.

Table 7.1 Results of the baseline analysis with heteroskedasticity-robust standard errors

Variable	Estimate	Standard error	Level of statistical significance
Intercept	2.959	0.131	***
Space attributes			
EntireHome	0.971	0.034	***
PrivateRoom	0.577	0.034	***
Bedrooms2	0.342	0.009	***
Bedrooms3	0.642	0.015	***
Bedrooms4	1.07	0.031	***
Bathrooms1,5	0.052	0.009	***
Bathrooms2	0.203	0.013	***
PerBedroom2	0.265	0.017	***
PerBedroom3	0.325	0.017	***
PerBedroom4	0.422	0.018	***
Reputational attributes			
log(Reviews)	-0.99	0.004	***
OverallRating	-0.398	0.061	***

Variable	Estimate	Standard error	Level of statistical significance
OverallRating^2	0.062	0.008	***
Superhost	0.077	0.007	***
log(Photos)	0.087	0.006	***
log(AirbnbDuration)	0.076	0.005	***
Location attributes			
log(OldTownSquare)	-0.146	0.007	***
log(PragueCastle)	-0.075	0.007	***
log(WenceslasSquare)	-0.072	0.006	***
Commerciality attributes			
BusinessReady	0.039	0.008	***
InstantBook-Enabled	0.023	0.007	***
Listings3	0.067	0.008	***
Listings11	0.126	0.009	***
Management policies			
Moderate	0.005	0.008	
Strict	0.065	0.008	***
Stay2	-0.001	0.007	
Stay3	-0.004	0.018	
Stay4	-0.066	0.017	***

Note: '***' p<0.001; '**' p<0.01; '*' p<0.05; '.' p<0.1; ' ' p<1

To present results coherently and systematically independent variables were split into five groups. These are space, reputational, location, commerciality attributes, and management policies.

Empirical results show that effect of space attributes on the average daily rate is positive. Daily rate of entire homes and private rooms is on average greater than rate of shared rooms proving how sensitive are guests about privacy during their stay. As expected, the coefficient of entire home variable is greater than the one of private room variable, thus entire properties charge on average higher rates than private rooms. Specifically, ADR of entire homes and private rooms is on average by 164% and 78% greater than shared rooms, respectively.

Further, reference group for bedroom dummies is properties with less than two bedrooms. In general, guests value large properties, as number of bedrooms in property rises, ADR increases as well. Parameters of bedroom dummies suggest that the effect of increase in number of bedrooms is not linear. Comparing with the reference group, apartments with two, three and at least four bedrooms have on average 41%, 90%, 192% greater average daily rate, respectively.

Regarding bathrooms, findings prove that guests appreciate comfort resulting from additional bathrooms. In comparison with the base group being apartment with one bathroom including toilet only, extra toilet in property increases on average ADR by 5.3%. Property with at least two bathrooms charges on average 23% higher rate than the base group keeping all other variables fixed.

Effect of an area of property is already accounted for, since bedroom dummies serve as a proxy for property size, to cover also effect of comfort provided by each property at given level of property size, influence of number of guests accommodated per one bedroom on ADR is estimated. Unexpectedly, coefficients of *perbedroom* dummies have opposite sign than assumed. With more guests accommodated per one bedroom, ADR increases. Surprisingly, lower level of comfort is associated with higher rate. There are two possible explanations for this unexpected finding. First, supply of listings accommodating large number of guests consists mostly of small size apartments with numerous beds in only a few bedrooms offering sufficient capacity. Despite appalling comfort, hosts are able to charge on average higher rates given the fact that more guests simply mean more people to pay for accommodation resulting in the negative parameter of the variable. The second way to clear up the non-intuitive sign of *perbedroom* dummies is through absence of information about size of apartments in square meters. Since this model takes all bedrooms as homogenous areas and does not take into account its size in square meters, biased parameters may be a result of the model estimation. If researchers have available data about property area, it would be highly recommended to use it in a model to avoid inaccurate estimates in the model.

Reputational attributes are crucial for building trust between host providing own property for short-term accommodation purposes and guest staying by stranger in an unfamiliar environment. The results of the baseline analysis suggest that developing credibility through reviews is not reflected in price markup. Moreover, each review obtained has negative effect on ADR. 1% increase in number of reviews results on average in decrease in ADR by 0.1%. As an example rise from 30 to 33 reviews will according to the analysis cause decline in ADR by 1% on average. As discussed in the endogeneity section, this is a common outcome in research investigating the Airbnb industry. Low-priced properties have higher probability of being booked, therefore, higher probability to actually obtain a review, this causes negative relationship between ADR and number of reviews. Moreover, Dogru & Pekin (2015) suggested that more reviews may also signal that hosts use the Airbnb platform for business purposes rather than experiential and social motivations which guests often seek.

Variable *OverallRating* is included in the model in quadratic form. Up to 3.2 stars there is negative effect connected with improving rating score, further, improvement is associated with price mark-ups. The inauspicious effect from 1 star to 3.2 stars is given by the distribution of the rating variable. The distribution is skewed to the left with mean 4.7 stars and median 4.8 stars. Even though the scale is from one to five stars, median 4.8 stars suggests half of the observations has rating between 4.8 and 5 star which is remarkably negligible variance. The initial dubious effect is given by having only a few observations in the dataset with a such low rating score. To demonstrate how *OverallRating* affects ADR, let mean value 4.7 stars increase up to 4.8 stars, this yields in rise in ADR by 1.9%.

As a consequence of quality guarantee and reliability, being a Superhost in Prague significantly leads to higher prices. Hosts, who are labeled with this status, charge on average 8% greater rates than their counterparts without the badge.

Regarding variable capturing effect of number of apartment photos on rental price, the impact is approximately insignificant. Increasing number of pictures by 1% is associated with 0.087% price markup. So guests appreciate more photos, however, their willingness to pay for it is limited. For instance, uploading four extra photos of apartment to 16 already displayed photos will on average increase ADR by 2.2%.

With respect to the length of Airbnb membership, the duration since the listing was established affects the average daily rate positively. If duration of membership prolongs by 1%, ADR responses by 0.08% increase. For instance, just because of extending Airbnb membership from year to year and a half, ADR rises by 3.8%. The parameter is in line with intuition manifesting that longer duration of Airbnb membership emphasises credibility and trustworthiness of listing.

Distance from an accommodation to the city centre or to popular touristic sight is important price determinant. Intuitively all three variables capturing effect of listing location on ADR have negative coefficient. The findings indicate that the farther from the particular touristic sight is listing located, the lower ADR host charge. However, the impact is not as eminent as expected. Regarding the Old Town Square which has the most influential effect on the average daily rate, doubling distance from property to the Old Town Square corresponds to a 14% drop in ADR. This relatively low influence could be explained by character of the variable data. 75% of properties are located within 3.2 kilometres from Old Town Square. Therefore, the ADR is not that sensitive to changes in location. Similar holds for other two Prague touristic attractions, Wenceslas Square and Prague Castle. Doubling distance from Prague Castle and Wenceslas Square yields in 7.4% and 7.2% decrease in price, respectively.

Regarding the fourth category— commerciality attributes, average daily rates for properties offering business-friendly workspace and including suitable equipment for working are about 4% higher than those without any customisation.

If a booking process is automatised and property could be booked instantly on Airbnb without need to contact a host, guests are willing to pay on average 2.4% more for daily rate than if booking process consists of direct communication between guest and host.

Room rates associated with professional Airbnb management or in other words associated with properties managed by host with multiple properties are on average greater than those managed by host with one or two properties only. It is worth pointing out that average daily rates of listings managed by property manager who has in portfolio between three to ten listings are on average 7% greater. Whereas impact of managing more than ten properties on price is even more remarkable, 13.4%. Li, Moreno, Zhang (2015) claimed in their study on discrepancies in behavioural actions and pricing strategies among professional and non-professional hosts expressed in number of listings managed that differences among the two groups which were found also in this thesis are given by inefficiency and behavioural biases carried out by non-professionals.

Last, impact of several property management policies on daily rate was examined. Listings whose cancellations policy is set as moderate do not on average experience significantly different ADR than those with flexible cancellation policy keeping all other variables fixed. However, substantial impact has been recognised for strict cancellation policy. Analysis implies that this policy setting is related to 6.7% greater ADR in comparison with properties with flexible cancellation policy. There are two plausible explanations of different impact of moderate and strict setting on price. First, there is no distinct difference between flexible and moderate setting. Under flexible and moderate cancellation policy guest must cancel at least 24 hours and 5 days before check-in in order to receive a full refund, respectively. Whereas under strict cancellation policy guest must cancel at least 14 days before check-in and within two days after booking was made in order to receive a full refund. Clearly, moderate setting is related to a greater extent to flexible setting rather than the strict one, therefore, impact on price is observed only by strict cancellation policy and not by moderate one. Second, strict cancellation policy is commonly set by hosts who are motivated by economic incentives and who master pricing techniques, thus they set strict policies in order to prevent guests to cancel reservations and to maximise inflow of money.

In order to maximise revenue streams, some hosts put a restriction on minimum stay length. Hosts who allow guests to stay even only one night serve as the reference group. Requirements to stay at least two or three nights do not show statistically significant effect on ADR. In contrast, restriction solely on bookings for at least four nights impacts price negatively. Listings with this requirement have on average 6.4% lower ADR. The estimate is in line with intuition arguing that such requirement represents an excessive limitation for guests, therefore, properties without any restriction charge on average higher rates.

7.2 The baseline analysis on specific subsets of properties

The baseline analysis on specific subsets of properties was run to examine whether results are stable when the input data distinctly shrinks and modifies. Three tested subsets were specified in the previous chapter. These are professional hosts, luxury properties, and seasonal properties. Overall results of the robustness check are stated in table 3 in the appendix.

The first subset captures professional hosts. Overall, the analysis resulted in practically identical outcome as observed on the whole sample data. The analysis reveals insignificance of the following variables *OverallRating, Superhost, Hostslistings, Bathroom1.5* due to lowering their variance in the sample or omitting relating regressor. Divergently from the outcome of the baseline analysis, any restriction of minimum stay length has negative impact on the average daily rate, therefore, professional hosts as characterised for purpose of this study should allow one-night stays in order to charge higher rates. Surprisingly properties advertised as a working-friendly charge on average 13% lower daily rates. This coefficient has unintuitive sign potentially due to focus of professional hosts on vacation and holiday rentals, therefore, if rented for business purposes, hosts are not able to maximise daily rates.

The second subset is concerned with luxury properties and enables to investigate whether this specific group of properties has significantly different price determinants. Generally, results of the analysis are very similar to the results of the whole market, however, there are several differences in estimates given by characteristics of luxury properties. In comparison with the whole market, variable *InstantBookEnabled* becomes insignificant suggesting that automatised reservation process is not an important feature for guests accommodating in luxury properties. Coefficient of dummy variable *perbedroom2* has turned from positive to negative indicating that fitting more guests into one bedroom has on average negative impact on daily rate of luxury properties which is in line with a general perception of luxury properties. Moreover, impact of strict cancellation policy on daily rate of luxurious properties is more than doubled, thus setting strict policy is especially important for such properties in order to maximise ADR. Furthermore, luxury properties managed by professional hosts with more than ten properties in portfolio charge significantly higher rates than non-luxury ones.

The last subset consists of seasonal listings that were rented only a fraction of a year. Coefficients of space, reputational and location attributes remain practically unchanged, however, some marked changes occur in parameters of commerciality attributes and management policies especially in heretofore insignificant variables. The largest difference that this change makes to the results of the analysis is the coefficient of *AirbnbDuration* variable which has almost tripled. The possible explanation for this significant change might be given by the fact that seasonal listings are advertised only a fraction of a year, therefore, it takes more time for host to become proficient on Airbnb (advertising, communication with guests). Hence increasing duration of Airbnb membership has more profound impact on seasonal listings in comparison with who rent out majority of a year.

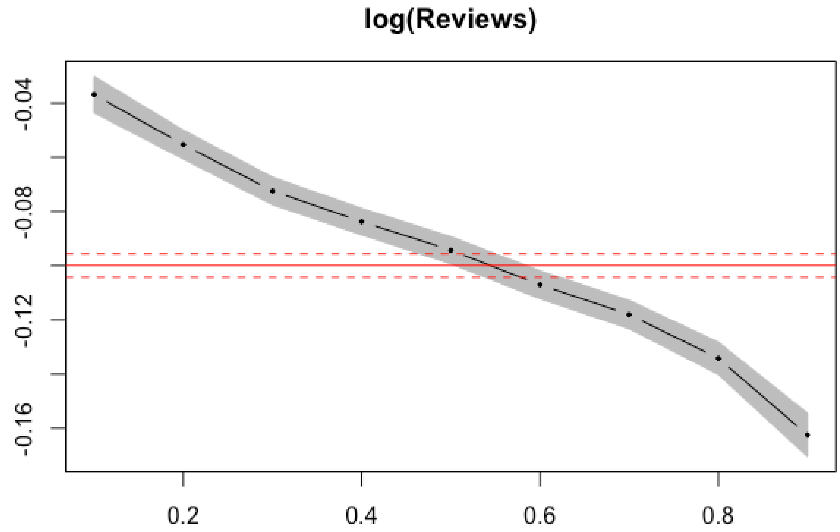
In conclusion, the baseline analysis run on the three subsets provides similar outcomes as the same analysis run on the whole dataset, in addition, only subtle changes in models parameters were observed by alterations of the input data, therefore, the baseline analysis is robust with analogous performance on the subsets.

7.3 Quantile regression

Quantile regression was employed to address two issues specified in the previous chapter. Plain results of the quantile regression could be found in table 4 in the appendix, where estimates of the 10th, 25th, 50th, 75th, and 90th quantiles are provided.

First, as estimated by OLS, the coefficient of number of reviews is unintuitive most likely due to endogeneity issues specified in the endogeneity section as well as the previous chapter. Concisely, low-priced listings tend to receive more bookings, hence more reviews. Therefore, additional review leads to lower price. QR was applied to detect whether this is the case also of higher-priced listings. According to QR, coefficients of observed independent variables experience decreasing trend across the distribution which implies that even for high-priced listings is the relationship negative. Figure 7.1 shows that high-priced properties experience on average even more pronounced negative effect on ADR when obtaining additional review than low-priced ones. Even though high-priced properties are more expensive, they could be at the same time still underpriced due to their outstanding characteristics and increase their probability of being booked and obtain a review. Moreover, large number of reviews may be a sign that property is managed by professional host with economic motivations which may not be desirable for guests seeking cultural exchange resulting in lower willingness to pay for property with plenty of reviews. In addition, some guests may understand many reviews as a signal of property being worn out rather than trustworthy.

Figure 7.1: Impact of number of reviews on ADR

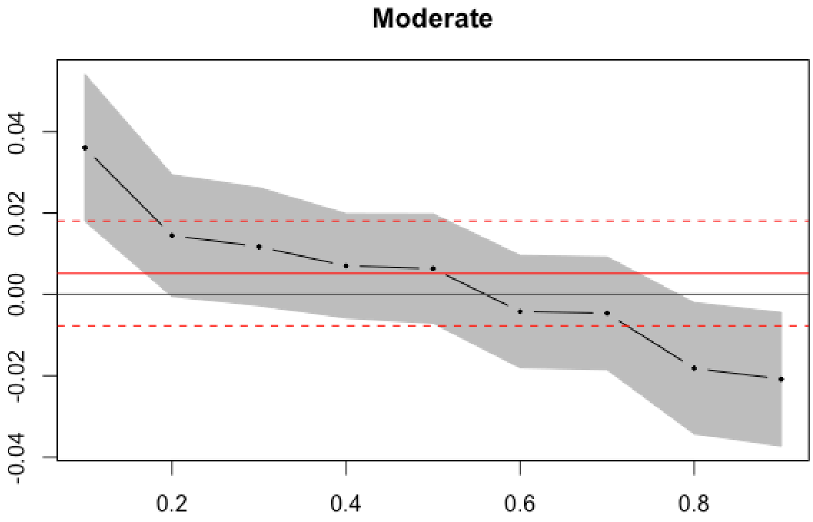


Note: The x-axis depicts quantiles of the dependent variable, while y-axis shows variation in the estimates of the independent variable. Each black dot represents coefficient estimate for the quantile stated on the x-axis. The grey area along is confidence intervals of coefficient estimates. The red lines are the OLS estimate and corresponding confidence intervals.

Further reason to employ QR is to investigate whether variables estimated as insignificant by OLS have actually zero effect on ADR or their effect is heterogenous across quantiles and thus cancels out on average. The baseline analysis argues that setting moderate cancellation policy as well as requirements of minimum stay length of two and three days do not impact the rental price. In the next paragraphs is examined whether results of QR are in line with OLS findings.

Although according to the baseline analysis setting moderate cancellation policy does not affect ADR in comparison with flexible cancellation policy, QR reveals that in some quantiles there is actually impact of the policy on ADR. Figure 7.2. suggests great heterogeneity in the effect of variable *Moderate.* Especially for the very low-priced and high-priced properties is the effect significant. Across the middle quantiles is the effect very mild and even insignificant. The analysis discovers that properties that charge low rental price could increase it by setting more restrictive cancellation policy. The positive impact on ADR is likely to happen due to diminishing cancelled bookings with moderate cancellation policy setting. The effects of strict cancellation policy estimated by QR and OLS are alike.

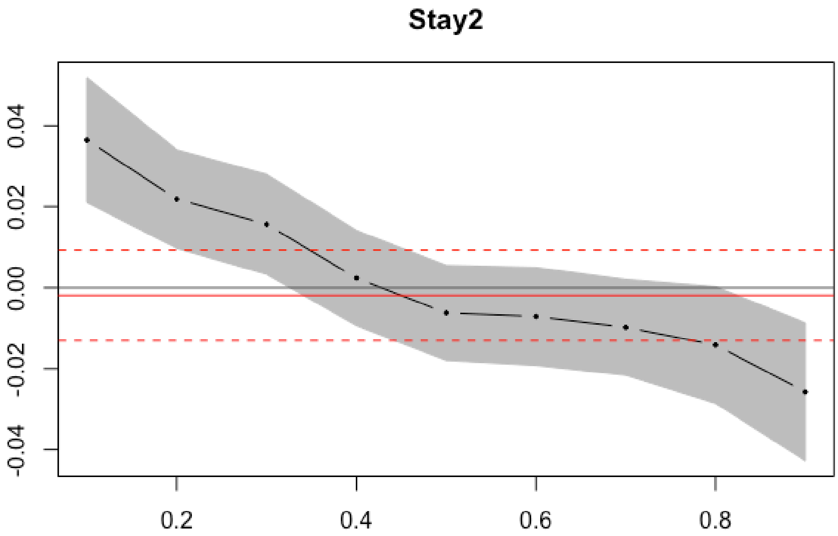
Figure 7.2: Impact of moderate cancellation policy on ADR



Note: The axes and lines represent the same as in the figure 7.1

The baseline analysis revealed that requirement of minimum of two or three long stay does not significantly impact ADR compared with the reference group (properties without restriction on minimum stay length), however, QR found statistically significant effects in the majority of quantiles. Regarding the properties with minimum of two nights stay, observed pattern in the quantile estimates is decreasing with positive effect on lower-priced properties and contrary impact on higher-priced properties. Given by very moderate effect across the middle quantiles and ambiguous effect in outer quantiles, the effect of minimum of two days long stay variable on ADR estimated by OLS cancels out.

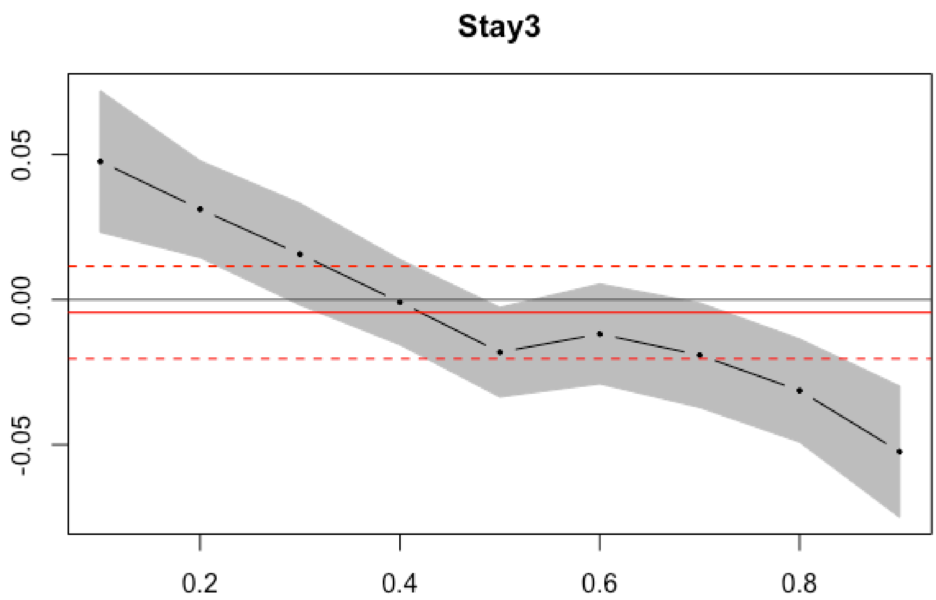
Figure 7.3: Impact of requirement of minimum two nights long stay on ADR



Note: The axes and lines represent the same as in the figure 7.1

Regarding the requirement of staying at least three days, the effect on ADR is even more pronounced than the one of variable *Stay2* discussed above. Impact of the restriction is even more positive on lower-priced properties and even more negative on high-priced properties (the beta coefficients are higher in the absolute value). Daily rates of low-priced properties with this restriction are nearly 5% higher than if the same properties would not set the restriction. On the contrary, daily rates of high-priced properties with this restriction are about 5% lower than if the same properties would not set any restriction.

Figure 7.4: Impact of requirement of minimum three nights long stay on ADR



Note: The axes and lines represent the same as in the figure 7.1

Even more restricting requirement of minimum at least four nights long stay impacts ADR negatively across all quantiles. Such results suggest that this restriction is too limiting for guests and it does not yield in price markup. To conclude all three dummy variables capturing effect of restricting minimum stay length impact ADR even though only on specific intervals of the conditional distribution. In general, restriction of minimum stay length of low-priced properties brings positive or at least neutral impact on ADR, however, the same restriction for higher-priced properties results in decrease in ADR. Moreover, as the restriction gets more strict, the impact on ADR is even more negative. This is in line with expectations that requirement of the minimum stay length could actually lead to negative effect on rental price of pricy properties and positive effect on rental price of cheaper properties.

The remaining independent variables were estimated as significant by the means of OLS. Moreover, they experience homogenous effect across quantiles. Therefore, they were not subject of further investigation in this section.

8 Conclusion

The thesis states and evaluates price determinants of average daily rates of Airbnb, sharing economy based accommodation platform, properties in Prague. Since launching in 2008 Airbnb has experienced massive growth and by now it is in many aspects comparable and competitive provider of short-term accommodation to traditional hotels.

OLS regression method was used to investigate impact of rental price predictors. We found that 25 out of the 28 explanatory variables to be statistically significant and felicitous price determinants for the purpose of the study.

This study confirms that space, reputational, location, commerciality attributes, and management policies have significant influence on rental price. The results imply that entire homes and private rooms charge higher daily rates than shared rooms. Spacious properties in terms of number of bedrooms and bathrooms tend to be more expensive. Variable *perbedroom* describing number of guests accommodated per one bedroom is identified as unique in the context of Airbnb price determinants analyses. Surprisingly, the study reveals that the more guests are accommodated per one bedroom, the higher price is on average charged. In line with existing literature, more reviews result in price discount due to potential threat of endogeneity. Improving overall rating score has positive effect on price, as well as to carry Superhost status or to be suited for business travellers. More property photos are associated with price mark-ups. Regarding location, increasing distance from property to the Old Town Square, Prague Castle and Wenceslas Square significantly lower average daily rate of the property. Allowing guests to book accommodation instantly results in price increase. The more listings one host manages, the higher average daily rates are observed. Properties with strict cancellation policy have higher rates than with flexible one. However, rental price is unaffected by setting moderate cancellation policy. Any restrictions on minimum stay length up to three nights have no impact on daily rates in comparison with reference properties having no requirement on stay length, however, prices of Airbnb accommodations decrease if minimum stay length of at least four days is required compared to no restriction. Rental price responds positively to extension of duration of Airbnb platform membership.

Stability of findings of the baseline analysis was examined in two robustness checks, firstly, the same baseline analysis was repeated on specific subsets and second, quantile regression was employed. The baseline analysis on the subsets discovered that findings are robust when the data input distinctly shrinks and modifies. Estimates of property characteristic are alike regardless of data inputs. QR focused on variables with unclear impact on rental price. Its findings proves that unintuitive sign of variable *Reviews* is observed across all quantiles, therefore, it is not just case of low-priced properties. Moreover, variables estimated as insignificant by OLS do not have null effect on rental price but heterogenous impact which on average cancels out, QR provided detailed picture of the impact of these variables on ADR.

This study examines factors determining the average daily rate on Airbnb platform in Prague, even though the topic has been studied before, another location was picked as a destination of interest and different set of control variables was chosen to explain variance in rental price. Obtained results are in line with existing literature.

The findings provide an in-depth comprehension of average daily rate determinants. The analysis identifies determinants and evaluates their impact. This study contributes to the existing literature by analysing academically unexplored Prague Airbnb market and employing the unique set of independent variables. On the theoretical level, analysis reveals novel results and provides complex insights into Airbnb determinants in the Prague market. From a practical perspective, with the findings rental suppliers could achieve a better understanding of the Prague Airbnb market and accordingly develop suitable pricing strategy resulting in enhancement of business situation and additional profit. Moreover, insights from the analysis could serve as one of the starting points for Airbnb or similar platforms when designing a pricing tool for hosts to improve inefficiency in pricing strategies.

Several limitations of this analysis are about to be acknowledged. First, the analysis investigated daily rate determinants based on available data, however, related social or psychological factors affecting hosts pricing-decision-making were not part of the examination. To obtain more accurate results such factors should be included in research. Moreover, this study uses the dataset with an absence of information about specific amenities and services provided in the accommodation such as elevator, hot tub, free breakfast, kitchen, Wi-Fi. It is assumed that adding this kind of information would increase the explanatory power of the model. Besides quantity of property pictures, variables capturing quality of the property photos and interior design would add information on impact of design of property on price decisions. Furthermore, host profile characteristics such as host verification or profile picture as already used in previous research were not available, nevertheless, they are encouraged to use if available.

Bibliography

Airbnb Statistics, 2020. IPropertyManagement [online]. Chicago: iPropertyManagement.com, February 2020 [cit. 2020-02-25]. Available from: https://ipropertymanagement.com/research/airbnb-statistics

About AirDNA, 2020. AIRDNA [online]. Denver: AirDNA [cit. 2020-04-27]. Available from: https://www.airdna.co/about-2

Barron, K., Kung, E., & Proserpio, D. (2018). The Sharing Economy and Housing Affordability: Evidence from Airbnb. In EC (p. 5).

Buchinsky, M. (1998). Recent advances in quantile regression models: a practical guideline for empirical research. Journal of human resources, 88-126.

Chen, C. F., & Rothschild, R. (2010). An application of hedonic pricing analysis to the case of hotel rooms in Taipei. Tourism Economics, 16(3), 685-694.

Cheng, M., & Foley, C. (2018). The sharing economy and digital discrimination: The case of Airbnb. International Journal of Hospitality Management, 70, 95-98.

Choi, K. H., Jung, J., Ryu, S. Y., Kim, S. D., & Yoon, S. M. (2015). The relationship between Airbnb and the hotel revenue: in the case of Korea. Indian Journal of Science and Technology, 8(26), 1-8.

Coyle, D., & Yeung, T. (2016). Understanding AirBnB in fourteen European cities. The Jean-Jacques Laffont Digital Chair Working Papers, 7088, 1-33.

Cui, R., Li, J., & Zhang, D. (2017). Discrimination with incomplete information in the sharing economy: Evidence from field experiments on Airbnb. Harvard Business School, 1-35.

Dogru, T., & Pekin, O. (2015). Pricing Airbnb accommodations: A hedonic pricing approach. Researchgate. Net, 1-13.

Dogru, T., & Pekin, O. (2017). What do guests value most in Airbnb accommodations? An application of the hedonic pricing approach.

Edelman, B. G., & Luca, M. (2014). Digital discrimination: The case of Airbnb. com. Harvard Business School NOM Unit Working Paper, (14-054).

Edelman, B. G., Luca, M., & Svirsky, D. (2015). Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment. Harvard Business School NOM Unit Working Paper, 16-069.

Ert, E., Fleischer, A., & Magen, N. (2016). Trust and reputation in the sharing economy: The role of personal photos in Airbnb. Tourism Management, 55, 62-73.

Espinet, J. M., Saez, M., Coenders, G., & Fluvià, M. (2003). Effect on prices of the attributes of holiday hotels: a hedonic prices approach. Tourism Economics, 9(2), 165-177.

Fagerstrøm, A., Pawar, S., Sigurdsson, V., Foxall, G. R., & Yani-de-Soriano, M. (2017). That personal profile image might jeopardize your rental opportunity! On the relative impact of the seller's facial expressions upon buying behavior on Airbnb™. Computers in Human Behavior, 72, 123-131.

Gibbs, C., Guttentag, D., Gretzel, U., Morton, J., & Goodwill, A. (2018). Pricing in the sharing economy: a hedonic pricing model applied to Airbnb listings. Journal of Travel & Tourism Marketing, 35(1), 46-56.

Gutt, D., & Herrmann, P. (2015). Sharing Means Caring? Hosts' Price Reaction to Rating Visibility. In ECIS (Vol. 54).

Guttentag, D. A., & Smith, S. L. (2017). Assessing Airbnb as a disruptive innovation relative to hotels: Substitution and comparative performance expectations. International Journal of Hospitality Management, 64, 1-10.

Guttentag, D., Smith, S., Potwarka, L., & Havitz, M. (2018). Why tourists choose Airbnb: A motivation-based segmentation study. Journal of Travel Research, 57(3), 342-359

Hamari, J., Sjöklint, M., & Ukkonen, A. (2016). The sharing economy: Why people participate in collaborative consumption. Journal of the association for information science and technology, 67(9), 2047-2059.

Heo, C. Y. (2016). Sharing economy and prospects in tourism research. Annals of Tourism Research, 58, 166-170.

Horn, K., & Merante, M. (2017). Is home sharing driving up rents? Evidence from Airbnb in Boston. Journal of Housing Economics, 38, 14-24.

Hung, W. T., Shang, J. K., & Wang, F. C. (2010). Pricing determinants in the hotel industry: Quantile regression analysis. International Journal of Hospitality Management, 29(3), 378-384.

Ikkala, T., & Lampinen, A. (2014). Defining the price of hospitality: networked hospitality exchange via Airbnb. In Proceedings of the companion publication of the 17th ACM conference on Computer supported cooperative work & social computing (pp. 173-176).

Ikkala, T., & Lampinen, A. (2015). Monetizing network hospitality: Hospitality and sociability in the context of Airbnb. In Proceedings of the 18th ACM conference on computer supported cooperative work & social computing (pp. 1033-1044).

Jung, J., Yoon, S., Kim, S., Park, S., Lee, K. P., & Lee, U. (2016). Social or financial goals? Comparative analysis of user behaviors in couchsurfing and Airbnb. In Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems (pp. 2857-2863).

Kakar, V., Voelz, J., Wu, J., & Franco, J. (2018). The visible host: Does race guide Airbnb rental rates in San Francisco?. Journal of Housing Economics, 40, 25-40.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. Business horizons, 53(1), 59-68.

Karlsson, L., & Dolnicar, S. (2016). Someone's been sleeping in my bed. Annals of Tourism Research, 58, 159-162.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. Econometrica: journal of the Econometric Society, 33-50.

Koenker, R., & Hallock, K. F. (2001). Quantile regression. Journal of economic perspectives, 15(4), 143-156.

Lampinen, A., & Cheshire, C. (2016, May). Hosting via Airbnb: Motivations and financial assurances in monetized network hospitality. In Proceedings of the 2016 CHI conference on human factors in computing systems (pp. 1669-1680).

Lee, C. G. (2011). The determinants of hotel room rates: Another visit with Singapore's data. International Journal of Hospitality Management, 30(3), 756-758.

Lee, D. (2016). How Airbnb short-term rentals exacerbate Los Angeles's affordable housing crisis: Analysis and policy recommendations. Harv. L. & Pol'y Rev., 10, 229.

Li, J., Moreno, A., & Zhang, D. J. (2015). Agent behavior in the sharing economy: Evidence from Airbnb. Ross School of Business Working Paper Series, 1298, 2015.

Mao, Z., & Lyu, J. (2017). Why travelers use Airbnb again?. International Journal of Contemporary Hospitality Management.

Mody, M., Suess, C., & Dogru, T. (2017). Comparing apples and oranges? Examining the impacts of Airbnb on hotel performance in Boston. Boston Hospitality Review, 5(2), 1-15.

Mohamad, H. (2016). Estimating the impact of Airbnb on hotels in Toronto (Doctoral dissertation, Massachusetts Institute of Technology).

Monty, B., & Skidmore, M. (2003). Hedonic pricing and willingness to pay for bed and breakfast amenities in Southeast Wisconsin. Journal of Travel Research, 42(2), 195-199.

Möhlmann, M. (2015). Collaborative consumption: determinants of satisfaction and the likelihood of using a sharing economy option again. Journal of Consumer Behaviour, 14(3), 193-207.

Perez-Sanchez, V. R., Serrano-Estrada, L., Marti, P., & Mora-Garcia, R. T. (2018). The what, where, and why of Airbnb price determinants. Sustainability, 10(12), 4596.

Qiu, R. T., Fan, D. X., & Liu, A. (2018). Exploring the booking determinants of the airbnb properties: an example of the listings of London. In Information and Communication Technologies in Tourism 2018 (pp. 44-51). Springer, Cham.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. Journal of political economy, 82(1), 34-55.

Sánchez-Ollero, J. L., García-Pozo, A., & Marchante-Mera, A. (2014). How does respect for the environment affect final prices in the hospitality sector? A hedonic pricing approach. Cornell Hospitality Quarterly, 55(1), 31-39.

Schamel, G. (2012). Weekend vs. midweek stays: Modelling hotel room rates in a small market. International Journal of Hospitality Management, 31(4), 1113-1118.

Sims, J., Ameen, N., & Bauer, R. (2019). Dynamic pricing and benchmarking in AirBnB.

Sundararajan, A. (2014). Peer-to-peer businesses and the sharing (collaborative) economy: Overview, economic effects and regulatory issues. Written testimony for the hearing titled The Power of Connection: Peer to Peer Businesses.

Tang, E., & Sangani, K. (2015). Neighborhood and price prediction for San Francisco Airbnb listings.

Teubner, T., Hawlitschek, F., & Dann, D. (2017). Price determinants on AirBnB: How reputation pays off in the sharing economy. Journal of Self-Governance & Management Economics, 5(4).

Teubner, T., Saade, N., Kawlitschek, F., & Weinhardt, C. (2016). It's only pixels, badges, and stars: On the economic value of reputation on Airbnb.

Thadewald, T., & Büning, H. (2007). Jarque–Bera test and its competitors for testing normality–a power comparison. *Journal of applied statistics*, *34*(1), 87-105.

Thrane, C. (2005). Hedonic price models and sun-and-beach package tours: the Norwegian case. Journal of Travel Research, 43(3), 302-308.

Thrane, C. (2007). Examining the determinants of room rates for hotels in capital cities: The Oslo experience. Journal of revenue and Pricing Management, 5(4), 315-323.

Tussyadiah, I. P., & Pesonen, J. (2016). Impacts of peer-to-peer accommodation use on travel patterns. Journal of Travel Research, 55(8), 1022-1040.

Varma, A., Jukic, N., Pestek, A., Shultz, C. J., & Nestorov, S. (2016). Airbnb: Exciting innovation or passing fad?. Tourism Management Perspectives, 20, 228-237.

Wachsmuth, D., & Weisler, A. (2018). Airbnb and the rent gap: Gentrification through the sharing economy. Environment and Planning A: Economy and Space, 50(6), 1147-1170.

Wang, D., & Nicolau, J. L. (2017). Price determinants of sharing economy based accommodation rental: A study of listings from 33 cities on Airbnb. com. International Journal of Hospitality Management, 62, 120-131.

White, P. J., & Mulligan, G. F. (2002). Hedonic estimates of lodging rates in the four corners region. The Professional Geographer, 54(4), 533-543.

Wooldridge, Jeffrey M. (2012). Introductory econometrics : a modern approach. Mason, Ohio :South-Western Cengage Learning, Yang, Y., Luo, H., & Law, R. (2014). Theoretical, empirical, and operational models in hotel location research. International Journal of Hospitality Management, 36, 209-220.

Yoo, M., Lee, S., & Bai, B. (2011). Hospitality marketing research from 2000 to 2009: topics, methods, and trends. International Journal of Contemporary Hospitality Management, 23(4), 517-532.

Yrigoy, I. (2019). Rent gap reloaded: Airbnb and the shift from residential to touristic rental housing in the Palma Old Quarter in Mallorca, Spain. Urban Studies, 56(13), 2709-2726.

Zervas, G., Proserpio, D., & Byers, J. W. (2017). The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. Journal of marketing research, 54(5), 687-705.

Zhang, Z., Chen, R. J., Han, L. D., & Yang, L. (2017). Key factors affecting the price of Airbnb listings: A geographically weighted approach. Sustainability, 9(9), 1635.

Zhang, Z., Ye, Q., & Law, R. (2011). Determinants of hotel room price. International Journal of Contemporary Hospitality Management.

List of appendices

Appendices

Table 1. Description of variables

Variable	Description
ADR	Average daily rate of a property expressed in American dollars calculated for the last 365 days before date of data retrieval
Space attributes	
EntireHome	Listing type, whole property is rented for guest use (dummy variable, shared rooms serve as reference group)
PrivateRoom	Listing type, whole room out of a property is rented for guest use (dummy variable, shared rooms serve as reference group)
Bedrooms2	Two bedroom property (dummy variable, properties with less than two bedrooms serve as reference group)
Bedrooms3	Three bedroom property (dummy variable, properties with less than two bedrooms serve as reference group)
Bedrooms4	Property with at least four bedrooms (dummy variable, properties with less than two bedrooms serve as reference group)
Bathrooms1,5	Property with bathroom and extra toilet (dummy variable, properties with one bathroom including a toilet serve as reference group)
Bathrooms2	Property with at least two bathrooms (dummy variable, properties with one bathroom including a toilet serve as reference group)
PerBedroom2	Number of guest accommodated per one bedrooms is between two and three (dummy variable, properties from one to two guests per bedroom serve as reference group)
PerBedroom3	Number of guest accommodated per one bedrooms is between three and four (dummy variable, properties from one to two guests per bedroom serve as reference group)
PerBedroom4	Number of guest accommodated per one bedrooms is at least four (dummy variable, properties from one to two guests per bedroom serve as reference group)
Reputational attributes	
Reviews	Total number of reviews received
OverallRating	Overall rating score expressed in stars from one to five
Superhost	Host is awarded Superhost status (dummy variable)
Photos	Number of property photos presented on Airbnb platform
AirbnbDuration	Duration (in days) of property being offered on Airbnb platform
Location attributes	
OldTownSquare	Distance from property to the Old Town Square in kilometres
PragueCastle	Distance from property to the Prague castle in kilometres

Variable	Description
WenceslasSquare	Distance from property to the Wenceslas Square in kilometres
Commerciality attributes	
BusinessReady	Property meets condition to be evaluated as suitable for business travellers (dummy variable)
InstantBookEnabled	Properties could be booked without approval of host (dummy variable)
Listings3	One host manages from 3 to 10 listings (dummy variable, to manage from one to two listings serves as reference group)
Listings11	At least 11 listings managed by one host (dummy variable, to manage from one to two listings serves as reference group)
Management policies	
Moderate	Cancellation policy is set as moderate (dummy variable, reference group are properties with flexible cancellation policy)
Strict	Cancellation policy is set as strict (dummy variable, reference group are properties with flexible cancellation policy)
Stay2	Guests are required to stay two days at minimum (dummy variable, properties with no restriction on minimum stay length serve as reference group)
Stay3	Guests are required to stay three days at minimum (dummy variable, properties with no restriction on minimum stay length serve as reference group)
Stay4	Guests are required to stay four days at minimum (dummy variable, properties with no restriction on minimum stay length serve as reference group)

Table 2. Results of the baseline analysis

Variable	Estimate	Standard error	Level of statistical significance
Intercept	2.959	0.101	***
Space attributes			
EntireHome	0.971	0.024	***
PrivateRoom	0.577	0.024	***
Bedrooms2	0.342	0.009	***
Bedrooms3	0.642	0.014	***
Bedrooms4	1.07	0.025	***
Bathrooms1,5	0.052	0.009	***
Bathrooms2	0.203	0.012	***
PerBedroom2	0.265	0.015	***
PerBedroom3	0.325	0.016	***
PerBedroom4	0.422	0.016	***
Reputational attributes			
log(Reviews)	-0.99	0.003	***
OverallRating	-0.398	0.048	***
OverallRating^2	0.062	0.006	***
Superhost	0.077	0.008	***
log(Photos)	0.087	0.005	***
log(AirbnbDuration)	0.076	0.004	***
Location attributes			
log(OldTownSquare)	-0.146	0.007	***
log(PragueCastle)	-0.075	0.007	***
log(WenceslasSquare)	-0.072	0.006	***
Commerciality attributes			
BusinessReady	0.039	0.01	***
InstantBook-Enabled	0.023	0.007	***
Listings3	0.067	0.007	***
Listings11	0.126	0.009	***
Management policies			
Moderate	0.005	0.008	
Strict	0.065	0.008	***

Variable	Estimate	Standard error	Level of statistical significance
Stay2	-0.001	0.007	
Stay3	-0.004	0.01	
Stay4	-0.066	0.017	***
Observations		13 500	
R^2		0.7	
Adjusted R^2		0.7	

Note: '***' p<0.001; '**' p<0.01; '*' p<0.05; '.' p<0.1; ' ' p<1

Table 3. Results of the baseline analysis on subsets
(standard errors are in parentheses)

Variable	Estimate		
	Professional hosts	**Luxury properties**	**Seasonal properties**
Intercept	-13.202 (10.513)	6.582*** (0.818)	2.163*** (0.223)
Space attributes			
EntireHome	-	-	0.965*** (0.053)
PrivateRoom	-	-	0.596*** (0.052)
Bedrooms2	0.388*** (0.032)	0.223*** (0.016)	0.373*** (0.022)
Bedrooms3	0.602*** (0.045)	0.427*** (0.028)	0.654*** (0.036)
Bedrooms4	0.893*** (0.068)	0.876*** (0.053)	1.015*** (0.066)
Bathrooms1,5	-0.023 (0.033)	0.125*** (0.018)	0.106*** (0.024)
Bathrooms2	0.218*** (0.042)	0.323*** (0.025)	0.148*** (0.029)
PerBedroom2	0.354*** (0.091)	-0.071*** (0.013)	0.245*** (0.033)
PerBedroom3	0.415*** (0.091)	-	0.3*** (0.036)
PerBedroom4	0.477*** (0.093)	-	0.396*** (0.037)
Reputational attributes			
log(Reviews)	-0.119*** (0.012)	-0.09*** (0.007)	-0.102*** (0.007)
OverallRating	7.06 (4.444)	-1.278*** (0.361)	-0.467*** (0.094)
OverallRating^2	-0.734 (0.47)	0.153*** (0.04)	0.066*** (0.012)
Superhost	0.04 (0.053)	-	0.045* (0.023)
log(Photos)	0.085** (0.026)	0.066*** (0.013)	0.083*** (0.012)
log(AirbnbDuration)	0.095*** (0.017)	0.066*** (0.01)	0.222*** (0.184)

Variable	Estimate		
Location attributes			
	Professional hosts	**Luxury properties**	**Seasonal properties**
log(OldTownSquare)	-0.068** (0.023)	-0.145*** (0.012)	-0.166*** (0.02)
log(PragueCastle)	-0.185*** (0.036)	-0.122*** (0.013)	-0.044* (0.018)
log(WenceslasSquare)	-0.094*** (0.022)	-0.115*** (0.014)	-0.053** (0.017)
Commerciality attributes			
BusinessReady	-0.142*** (0.03)	0.072*** (0.014)	0.07** (0.027)
InstantBook-Enabled	-	0.01 (0.014)	-0.015 (0.017)
log(HostsListings)	0.007 (0.021)	-	-
Listings3	-	0.036* (0.015)	0.031. (0.019)
Listings11	-	0.172*** (0.021)	0.173*** (0.023)
Management policies			
Moderate	-	0.043* (0.018)	0.016 (0.019)
Strict	-	0.14*** (0.019)	0.098*** (0.02)
Stay2	-0.127*** (0.034)	-0.012 (0.015)	-0.035* (0.017)
Stay3	-0.235*** (0.043)	0.003 (0.019)	0.00004 (0.023)
Stay4	-0.128 (0.187)	-0.119** (0.042)	-0.027 (0.033)
	Professional hosts	**Luxury properties**	**Seasonal properties**
Observations	727	2 086	3 015
R^2	0.613	0.686	0.665
Adjusted R^2	0.601	0.682	0.662

Note1: Each column represents estimates of rental price determinants on specific subsets of properties

Note2: '***' p<0.001; '**' p<0.01; '*' p<0.05; '.' p<0.1; ' ' p<1

Table 4. Quantile regressions estimates of the baseline model
(standard errors are in parentheses)

Variable	**0.1**	**0.25**	**0.5**	**0.75**	**0.9**
Intercept	2.719*** (0.14)	2.771*** (0.142)	2.923*** (0.129)	3.423*** (0.351)	4.017*** (0.315)
Space attributes					
EntireHome	1.128*** (0.021)	1.109*** (0.03)	1.045*** (0.045)	0.893*** (0.076)	0.72*** (0.104)
PrivateRoom	0.669*** (0.02)	0.663*** (0.03)	0.634*** (0.045)	0.524*** (0.076)	0.379*** (0.104)
Bedrooms2	0.269*** (0.011)	0.294*** (0.009)	0.344*** (0.009)	0.363*** (0.011)	0.419*** (0.013)
Bedrooms3	0.553*** (0.022)	0.588*** (0.016)	0.651*** (0.017)	0.68*** (0.019)	0.718** (0.022)
Bedrooms4	0.888*** (0.061)	0.953*** (0.06)	1.086*** (0.044)	1.148*** (0.027)	1.186*** (0.04)
Bathrooms1,5	0.077*** (0.011)	0.052*** (0.009)	0.041*** (0.008)	0.048*** (0.011)	0.055** (0.019)
Bathrooms2	0.17*** (0.021)	0.194*** (0.015)	0.199*** (0.016)	0.239*** (0.017)	0.234*** (0.019)
PerBedroom2	0.236*** (0.02)	0.242*** (0.015)	0.26*** (0.015)	0.241*** (0.023)	0.285*** (0.03)
PerBedroom3	0.267*** (0.021)	0.296*** (0.016)	0.325*** (0.015)	0.309*** (0.024)	0.329*** (0.03)
PerBedroom4	0.34*** (0.021)	0.381*** (0.016)	0.424*** (0.016)	0.41*** (0.024)	0.462*** (0.03)
Reputational attributes					
log(Reviews)	-0.037*** (0.004)	-0.065*** (0.003)	-0.094*** (0.003)	-0.126*** (0.004)	-0.163*** (0.005)
OverallRating	-0.466*** (0.067)	-0.396*** (0.066)	-0.414*** (0.058)	-0.521* (0.159)	-0.686*** (0.139)
OverallRating^2	0.071*** (0.009)	0.062*** (0.008)	0.065*** (0.007)	0.077*** (0.018)	0.094*** (0.016)
Superhost	0.076*** (0.01)	0.081*** 0.008)	0.07*** (0.007)	0.069*** (0.009)	0.056*** (0.01)
log(Photos)	0.06*** (0.007)	0.073*** (0.006)	0.078*** (0.006)	0.096*** (0.006)	0.099*** (0.008)
log(AirbnbDuration)	0.032*** (0.006)	0.047*** (0.005)	0.073*** (0.004)	0.096*** (0.005)	0.131*** (0.006)

Variable	0.1	0.25	0.5	0.75	0.9
Location attributes					
log(OldTownSquare)	-0.16*** (0.009)	-0.152*** (0.008)	-0.151*** (0.007)	-0.139*** (0.006)	-0.144*** (0.007)
log(PragueCastle)	-0.046*** (0.009)	-0.071*** (0.008)	-0.082*** (0.008)	-0.087*** (0.009)	-0.091*** (0.012)
log(WenceslasSquare)	-0.096*** (0.008)	-0.088*** (0.006)	-0.072*** (0.006)	-0.07*** (0.007)	-0.05*** (0.008)
Commerciality attributes					
BusinessReady	0.044*** 0.012	0.035*** (0.009)	0.034*** (0.009)	0.044*** (0.01)	0.048*** (0.011)
InstantBook-Enabled	0.028* (0.009)	0.022** (0.007)	0.023** (0.007)	0.016* (0.008)	0.036*** (0.008)
Listings3	0.067*** (0.01)	0.07*** (0.008)	0.061*** (0.008)	0.061*** (0.009)	0.061*** (0.011)
Listings11	0.144*** (0.011)	0.137*** (0.009)	0.102*** (0.009)	0.108*** (0.01)	0.081*** (0.011)
Management policies					
Moderate	0.036** (0.011)	0.016. (0.008)	0.006 (0.008)	-0.009 (0.01)	-0.021* (0.01)
Strict	0.086*** (0.011)	0.054*** (0.009)	0.056*** (0.009)	0.052*** (0.01)	0.077*** (0.012)
Stay2	0.036*** (0.009)	0.018* (0.007)	-0.006 (0.007)	-0.013 (0.008)	-0.026* (0.01)
Stay3	0.048** 0.015	0.02* (0.01)	-0.018*** (0.009)	-0.029. (0.01)	-0.052*** (0.014)
Stay4	-0.04 0.027)	-0.046* (0.02)	-0.066*** (0.017)	0.096*** (0.005)	-0.139*** (0.015)

Note1: '***' p<0.001; '**' p<0.01; '*' p<0.05; '.' p<0.1; ' ' p<1

Note2: Column names stand for quantiles of the dependent variable