ÇANKAYA UNIVERSITY
FACULTY OF ENGINEERING
COMPUTER ENGINEERING DEPARTMENT

CENG 407
AIRBNB PRICE DETERMINANT

LITERATURE REVIEW REPORT

201627015 MELİSA YILDIZ

201611033 SELİN KARA

201512019 RANA GÜRCÜ DELİÖMEROĞLU

201514039 MERVE KARAKAYA

CONTENTS

Abstract

Airbnb is a platform to share houses for a short stay with those who prefer houses to hotels. Airbnb is also known as sharing economy accommodation service and peer to peer accommodation. The platform puts the responsibility of determining the price of the hosts. Due to Airbnb's unique nature, hosting pricing strategies are very different from the traditional hospitality industry. The prices are not determined randomly. In fact, many features such as location, availability of public transportation, size of the rented house, the existence of shopping places in the neighborhood, and so on, affect the price. In this project, the data from the official site of Airbnb will be used to determine which attributes are the most decisive ones in determining the price. The developed model can be used by hosts and owners of properties to estimate the prices more accurately. Thus, in addition to accurate pricing, it can offer advantages such as a more systematic, easy and helpful approach.

Airbnb, evi otele tercih edenlerle kısa süreli konaklama için evleri paylaşabileceğiniz bir platformdur. Airbnb, aynı zamanda ekonomi konaklama hizmetinin paylaşılması ve eşler arası konaklama olarak da bilinir. Platform, ev sahiplerinin fiyatını belirleme sorumluluğunu üstlenir. Airbnb'nin benzersiz doğası nedeniyle, ev sahipliği fiyatlandırma stratejileri geleneksel konaklama endüstrisinden çok farklıdır. Fiyatlar rastgele belirlenmez. Nitekim lokasyon, toplu taşıma araçlarının kullanılabilirliği, kiralanan evin büyüklüğü, mahallede alışveriş yerlerinin varlığı gibi pek çok özellik fiyatı etkilemektedir. Bu projede Airbnb' nin resmi sitesinden alınan veriler, fiyatın belirlenmesinde hangi özelliklerin en belirleyici olduğunu belirlemek için kullanılacaktır. Geliştirilen model, fiyatları daha doğru tahmin etmek için ev sahipleri ve mülk sahipleri tarafından kullanılabilir. Böylelikle doğru fiyatlandırmanın yanı sıra daha sistematik, kolay ve yardımcı bir yaklaşım gibi avantajlar da sunabilir.

1 Introduction

In the area of sharing economy, Airbnb acts as a novel and alternative business model in the hospitality industry. Founded in 2008, Airbnb has grown into a leading peer-to-peer accommodation platform, with more than 5 million listings located in 81,000 cities worldwide [1]. The essence and rapid success of Airbnb lies on the effective mix of several key factors, including affordable prices and economic advantages authenticity and unique consumer experience, sustainability, perceived attractiveness and responsiveness. Above all these advantages, price and lower cost are frequently reported as one of the most important factors facilitating the rapid diffusion of P2P accommodation sharing phenomenon [2].

Since Airbnb is different from traditional accommodation, many features should be considered in price determination. Each feature should be passed through different evaluation scales, after which the positive and negative impact on the price should be considered. Therefore, pricing should be done in detail.

2 Literature Review

2.1 Airbnb and Sharing Economy

Sharing, as a human activity, is as old as time itself. However, it was not until the booming of Internet technology that “collaborative consumption” and “sharing economy” have become a social phenomenon. In the era of sharing economy or peer economy, individuals rent, lend, or trade goods, services, transportation, or space in a peer-to-peer way. Tourism, which involves all the above-mentioned activities, is one of the areas most affected by the sharing economy. In recent years, sharing economy-based business models have been developed in transportation (such as Uber.com) and accommodation (such as Airbnb.com and Homeaway.com) sectors due to high tourist demand [1].

Regarding sustainability, the emerging sharing economy in the tourism industry has a higher possibility of this than the traditional industry. Through the utilization of under-utilized assets, the sharing economy can increase the amount of service without additional construction or acquisition, which would decrease the environmental footprint of the industry.

Several studies have shown that consumers of the sharing economy have a higher interest in the local community and the environment, so that both consumers and suppliers lead to a reduced consumption of energy and water and reduced waste generation [3].

Furthermore, the sharing economy increases the employment rate of the local community, as well as the profits of the community. For example, Airbnb's message to tourists, "travelers to your neighborhood," encourages them to have new experiences via sharing the same space with the local community members, which makes the Airbnb tourists engage more in economic activities in the local areas, like eating at a local restaurant, than other tourists. The human capital of the community that provides idle assets to the sharing economy serves as a bridge between these opportunities and helps consumers to respect the culture of the community and to resolve social imbalances [3].

2.2 Differences in Hospitality Industry

Airbnb, a leading platform in peer-to-peer accommodation products, is a "creative destruction" to the traditional hospitality industry. Different from hotels, the host–guest relationship in Airbnb is beyond staff–customer relations. Both Airbnb hosts and guests seek more than functional values. Experiential value, authenticity, sociability, and trust are high on the agenda for Airbnb users. Furthermore, Airbnb affects hotels' performance. Airbnb is found to have a substitution effect on hotels, because Airbnb supply negatively influences hotels' financial performance, and higher Airbnb rating score is related with a decrease in hotel [1].

Sharing economy accommodation has distinctive attributes compared to hotels, mainly due to its residential housing based service. These unique features make it challenging to apply the conventional understanding of the hospitality industry on the value recognition of users. For example, the star rating system based on user reviews has limited influence on the price of Airbnb listings, unlike the hotel industry, while indirect signals such as the service duration and personal information of the hosts are considered as alternative sources for the trust and significant influence on the price [3].

It is generally accepted that the most critical features for hotel pricing are star ratings and location. Star ratings can be perceived as an index signal evaluated by an independent organization that does not benefit from biased reporting of hotel quality and experience. This makes the star rating system extremely effective for customers and hotels as well. Empirical studies have also shown that star ratings of hotels have high explanatory power over the hotel price. Experimental studies have revealed that the sea view, the city center and the transportation center are the most considered location factors in hotel pricing. However, this situation differs for accommodation types such as Airbnb as mentioned above.

2.3 Price Determinants Of Airbnb

First of all, as mentioned in the previous sections, price determination involves quite different approaches due to the unique characteristics of Airbnb.

The difference between Airbnb listings and hotels come from not only real estate, but also service, amenities, the personality of hosts, and local communities. Thus, accumulated knowledge of revenue management in the hospitality industry cannot be applied directly to the sharing economy accommodations. Most pricing studies on Airbnb commonly pointed out that the standardization of business is the key difference between Airbnb and hotels. In the case of hotels, the shape and format of services are quite similar among hotels due to the standardization, so that consumers can easily recognize or compare the values of the service to these industrial service standards. Unlike hotels, each listing in Airbnb has unique characteristics that are hard to compare. However, Airbnb's user interface provides standardized information of listings to customers. In this way, customers recognize the value of listings through the user interface of the platform [3].

Although the reviews of customers are still meaningful indicators of the price, customers have started to look for other signals to eliminate the biases. The empirical studies found that the status of the host, the service length of the listing, and the super host badge are now utilized for customers to measure the quality and value of a listing. The attributes of the given property, and the number, size and types of rooms also showed statistically significant effects on the price of Airbnb listings through product differentiation like hotels [3].

If the above-mentioned Airbnb's pricing features are to be opened in more detail, the table below, figure 1, will help. In addition, the effect of these features on the price is also specified.

Category	Determinants	Effects
Listing attributes	• Room type[a] (reference group: shared room), number of bedrooms[a], number of bathrooms[a], number of accommodations[a], free parking[b], pool, gym, real bed, wireless Internet, number of accommodation photos	positive
	• Accommodation type (reference group: independent properties), free breakfast[b]	negative
	• Months since the listing was established	insignificant
Host attributes	• Hosts' listing count, host verification, host profile picture, response time	positive
	• Race (reference group: white)	negative
	• Gender, marital status, sexual orientation, acceptance rate	insignificant
	• Superhost or not[a], professional (multilisting) or not	mixed
Listing reputation	• Rating on cleanliness, rating on location	positive
	• Number of reviews[a], rating on value	negative
	• Rating on accuracy, rating on check in	insignificant
	• Overall rating score[a], rating on communication	mixed
Rental policies	• Strict cancellation policy[b], guests' phone verification required	positive
	• Instant bookable[b], smoking allowed	negative
	• Guests' profile photo required	insignificant
Listing location	• Neighborhood value, average rental price in the district, number of Airbnb listings in the same district, price of surrounding Airbnb listings and hotels, number of points of interest (POIs) in surrounding area, located within sightseeing, eating or shopping areas, located near a continuous coastal fringe	positive
	• Distance to city center, number of hotels in the same district	negative
	• Distance to the nearest highway	insignificant

Figure 1. Price Determinants of Airbnb Listings

Listing Attributes:The first category of Airbnb price determinants is listing attributes. Accommodation type and room type determine Airbnb room rates. Number of bedrooms, number of bathrooms, and number of accommodations also have a positive impact on Airbnb rental price. Facilities such as car parking, swimming pool, and wireless Internet positively affect Airbnb listing price. However, it has been found in several recent studies on hotels that free car parking does not have a significant influence on hotel room rate. This may be because free car parking has already become a standard service in hotels in recent years, whereas in the peer-to-peer rental accommodation sector car parking is part of the hosts' personal property. Including breakfast has a negative effect on Airbnb room rate, which is also inconsistent with findings in hotel research. The number of accommodation photos positively impacts Airbnb price and the number of reviews [1].

Host Attributes: As a member of peer-to-peer accommodation, another notable price determinant of Airbnb is host attributes. When it comes to hosts' race, discovered that, compared with white counterparts, Asian and Hispanic hosts charge a lower price on Airbnb in San Francisco. Existing research has not found a significant effect of the host's gender, marital status, or sexual orientation on Airbnb room rate. Verifying hosts' profile or providing hosts' photo on Airbnb can lead to a price premium, as such action increases the trustworthiness of the host. A "superhost" badge on Airbnb also helps earning additional room rate. Hosts' operation capability also brings more economic income. According to several studies, "Professional hosts" ("professional" means the host has two or more listings on Airbnb) win a higher room rate of each property than their nonprofessional counterparts. However, while professional hosts do not offer a significant higher price than nonprofessionals, they do have a higher daily revenue of each property, as the occupancy rate is higher. The more Airbnb listings a host owns, a higher room rate is charged. Although hosts' acceptance rate does not have a significant impact on price, longer response time is related with a higher room rate [1].

Listing Reputation: The third category is listing reputation. The number of reviews negatively influences listing price but positively affects daily revenue and occupancy rate of each property. A possible explanation is that the quantity of reviews stands for the size of demand and there is more demand for less expensive listings. Customer rating shows a mixed effect on room rate. There is empirical evidence supporting both positive and negative impact of overall rating score. Rating score on accuracy and check-in does not affect price. Higher score on cleanliness, communication effectiveness and location wins a price premium. Rating score on value, however, has a negative impact on Airbnb listing price [1].

Rental Policies: The fourth category is associated with rental policies. Instant bookable listings are less expensive. Refundable cancellation policy is linked with lower rental price, which is contradictory to the findings in hotel research. Permitting smoking also negatively influences Airbnb room rate. Furthermore, requiring guests' phone verification is related with higher room rate, whereas requirement on guests' profile photo does not impact Airbnb listing price [1].

Listing Location: The last category is listing location. Similar to hotels, locational factors are considered in the pricing strategy of Airbnb. Closeness to the city center and coastline contribute to higher room rate. If a listing is located within sightseeing, eating, or shopping area, it also wins a price premium. The number of points of interest (POIs) in the surrounding area is positively related with Airbnb listing price in Tallinn, Estonian. Neighborhood value shows a positive but diminishing return rate on Airbnb listing price. However, a shorter distance to the highway does not exert a significant impact on Airbnb rental price. Market competition is also important, since the number of Airbnb listings in the same district gives rise to rental price whereas the density of hotels negatively affects Airbnb listing price. Higher median gross rent in the district as well as the room rate of surrounding Airbnb listings also raises Airbnb price [1].

2.4 Pricing Methods

Academic inquiries into the pricing of sharing economy business is still in its infancy. Generally speaking, there are two pricing models in sharing economy platforms. The first model is completely based on algorithm, which means the firm has the full control over the prices of each service request. Uber is a typical example of such model. The second pricing model uses algorithm tools only for reference, leaving the pricing decision open to service providers. Peer-to-peer accommodation products such as Airbnb fall into the second category [1].

Model in 2020:

Existing Airbnb pricing studies have limitations considering spatial variances and crucial geographic information for estimating the influence of the pricing variables. It is well known that the influence of pricing variables may have local variations, indicating different dynamics among factors across space . To explicitly reflect this spatial heterogeneity, the geographically weighted regression (GWR) has been utilized for pricing research into hotels and Airbnb hosts and the provisions of accommodations in the sharing economy. An issue here is that GWR does not consider a varying scale of spatial heterogeneity, treating all variables as having an identical scale of operation . Pricing variables are spatial processes that determine the price, and each of them will operate under a different scale, from local to global. However, to date, no research effort has been made that considers this multiscale regression approach for Airbnb pricing analysis [3].

The second issue is negligence on a crucial decision criterion for accommodation: distance to tourism destinations. Since tourists are the most frequent customers of Airbnb , the distance criterion needs to be included in Airbnb pricing research. A few previous studies have considered the distance variable in their pricing model. However, they used surrogate points such as a city center and highway exits for tourism destinations , which is not an accurate representation. A tourist will consider accessibility to multiple tourism destinations to make accommodation decisions. Therefore, distance to multiple tourism destinations must be explicitly considered in the Airbnb pricing research [3].

GWR captures spatial heterogeneity in pricing strategies by conducting a localized estimation of the influence of the pricing variables. The core idea of GWR is directly capturing geographical variance in regression estimates across space. It computes localized regression coefficients of given explanatory variables for a given location, only using its neighboring locations, to show the varying influence of pricing variables. The neighborhood is defined by search bandwidth, a reflection of the concept of the scale of operation. Airbnb pricing research has adopted the concept of locally varying relationships among variables, introducing GWR to the research [3].

However, the scale of operation is likely to vary for each pricing variable, depending on its nature. For example, management variables such as cancellation policy and superhost status are likely to show a similar tendency at the regional or global level. Other variables, such as the number of bedrooms, guests, and distance to tourism destinations, would be the more localized scale of operation, as the influence of them would be more sensitive to the local environment. A critical limitation of GWR is that it applies an identical scale of operation for all the explanatory variables, ignoring the possibility of varying scales . If we assume that all the pricing variables have an identical scale of operation, it will result in creating false spatial variances in global variables. Differences in the scale of operation among local variables will be ignored as well [3].

To analyze the influence of pricing variables on Airbnb price accurately, a different regression model needs to be utilized to reflect the different scale of operation [3].

Multiscale Geographically Weighted Regression: A regression model estimates relationships between explanatory and response variables. Conventional models like OLS assume globally identical relationships in a given region. Therefore, they estimate the relationships, expressed as correlation coefficients, using the entire cases with the same level of weight. If we apply this logic to the Airbnb pricing strategy, we would assume that all the hosts in the region use the same logic and criteria to determine their price, independently to each other. However, this is not entirely true [3].

To overcome this limitations, Fotheringham et al. (2017) propose another variation of GWR, multiscale GWR (MGWR). The fundamental idea of MGWR is using different search bandwidth, the scale of operation, for each explanatory variable to estimate its coefficient. Bandwidths are determined by the data rather than relying on external factors [3].

$$y_i = \sum_{j=0}^{m} \beta_{bwj}(\mu_i, v_i)x_{ij} + \epsilon_i, i = 1, \ldots, n$$

Figure 2. MGWR Function

Here, indicates the scale of operation (bandwidth) of the jth explanatory variable. To deter mine the bandwidth, MGWR requires a different approach than GWR as it needs to derive multiple different bandwidths for given variables at the same time, so it uses a back-fitting approach. This back-fitting is initialized from ordinary GWR estimation, and tests goodness-of-fit for each variable to find the most suitable bandwidth to reflect the scale of operation indicated in the given dataset [3].

MGWR results revealed that the scale of operation varies widely among variables; and even for the identical variable, the scale of operation differs between cities. Based on the bandwidth and the statistical distribution of the coefficients across each city, we classified the pricing variables as local, regional, and global variables. These global variables can be interpreted in two ways: (1) Airbnb hosts in each city use similar pricing strategies regarding these variables, and/or (2) the hosts are influenced by other competitors' pricing strategy related to these variables' city-wide scale [3].

Model in 2019:

Based upon hedonic regression models, these studies focus on the algorithm perspective of pricing. Driven from previous research on Airbnb pricing, the authors identify five groups of explanatory variables, which are presented in Table 1: (1) listing attributes; (2) host attributes; (3) listing reputation; (4) rental policies; (5) listing location [1].

To estimate the effects of explanatory variables on price, the authors employed semilogarithm hedonic price models. Hedonic price model was developed by Rosen (1974) to quantify the marginal effects of a product's attributes on price. The original function was written as [1]:

$$P = F(Z, \varepsilon) \qquad (1)$$

Figure 3. The Original Hedonic Price Function

where P is the observed price of a product, Z is a vector of the product's attributes or utilities, and ε is the standard error [1].

This study used the semilog form of hedonic price function, which is specified as:

$$\ln P = \beta_0 + \sum_{i=1}^{n} \beta_i X_i + \varepsilon \qquad (2)$$

Figure 4. Semilog Form of Hedonic Price Function

where lnP is log form of the observed price of a sample listing, β_0 is the constant term, X_i is the ith explanatory variable, β_t is the semielasticity of P with respect to X_i , and ε is the error term [1].

According to Wooldridge (2014), in semilogarithm functions, the exact percentage change in y caused by a change in x_i (controlling all other factors) is predicted as [1]:

$$\Delta\hat{y} = \left(e^{\hat{\beta}_i \Delta x_i} - 1\right) \times 100\% \qquad (3)$$

Figure 5. Exact Percentage Change in y

Model in 2018:

In the 2018's study set out to test the following hypothesis: the determinants of Airbnb prices in the context of Spanish Mediterranean cities are strongly related to the location of the listings: the closer to the city center, the more expensive the accommodation.[5]

This study adopts a multivariable analysis technique which consists of estimating a hedonic price model that is often used in market studies of heterogeneous goods . The method is used for determining and quantifying the relation between Airbnb accommodation attributes and price. To estimate how these characteristics influence price, the model is estimated by using ordinary least squares—OLS—adopting the renowned method of successive steps to select the variables . [5]

Taking into account the aforementioned limitations, this study adopts the quantile regression method—QRM, whose advantage over the OLS estimation is that the importance of the dependent variable determinants can be explained at any point of the distribution. Thus, the model is less affected since it is not necessary to establish an aleatory perturbation hypothesis. Therefore, the results obtained from the estimation of both models—OLS and QRM, allow identification of the shadow price component of the listing price, which is generated by the accommodation characteristics. [5]

The estimation for these types of models often requires the logarithmic transformation of the dependent variable, the listing price. The literature highlights several reasons as to why the use of semilogarithmic models is rather frequent. For instance, the goodness-of-fit criteria applied to data for avoiding heteroskedasticity and facilitating interpretation of the coefficients. Due to the logarithmic transformation, the goodness-of-fit criteria shows how the dependent variable—the price—presents percentual variables when there are unitary changes of the independent variable (pp. 193–194). Moreover, Sirmans et al. indicate that hedonic models are estimated with logarithmic forms given the great probability that the implicit prices obtained for each characteristic are not the same for all price ranges. [5]

The proposed OLS model is the following:

$$ln(P_i) = \alpha + \sum_{j=1}^{n} \beta_j X_{ij} + \sum_{k=1}^{m} \gamma_k D_{ik} + \varepsilon_i \quad (1)$$

where:

ln (Pi) is the neperian logarithm of *daily_price* for the property "i"
α is the fixed component, which is independent from the market.
β_j is the parameter to be estimated, related to the characteristic "j".
X_{ij} is the continuous variable that considers the characteristic "j" of observation "i".
γ_k is the parameter to be estimated, related to the characteristic "k".
D_{ik} is the fictitious variable that considers the characteristic "k" of observation "i".
ε_i is the term of the error associated with the observation "i".

Given the specification of the model, the impact on the price with a change ranging from 0 to 1 in a dummy variable, while keeping all other independent variables constant, can be calculated using the expression (2), as Kennedy suggests .[5]

$$\hat{p} = 100\left[exp\left(\hat{c} - \frac{1}{2}\hat{V}(\hat{c})\right) - 1\right] \quad (2)$$

The quantile regression method does not require the fulfillment of certain criteria as in the case of the OLS model. The QRM allows controlling non-linearity; non-normality due to asymmetries and outliers; and heteroskedasticity. OLS models are based on the conditional mean of the dependent variable, given certain values of the predictor variables, but do not provide information for other conditional quantiles of the dependent variable, whereas QRM allows modeling different conditional quantiles of the dependent variable, overcoming the limitations of OLS models. Thus, it is possible to estimate the implicit value of each accommodation characteristic for different price ranges—or quantiles—as they may vary. Then, the impact of each characteristic, depending on the sales price ranges, can be estimated. As suggested by, a QRM can be defined using a multiple linear regression model as [5]

$$Y_i = X_i\beta_\theta + u_{\theta i} \tag{3}$$

where

Y_i is the dependent variable
X_i is the matrix of independent variables
β_θ is the vector of parameters to be estimated for quantile θ
$u_{\theta i}$ is the aleatory perturbation that corresponds to quantile θ.

As explained by Koenker and Bassett, in this regression model, quantiles θ are defined for the dependent variable Yi , given the Xi dependent variables, where Quant(Yi|Xi) = Xiβθ . Each quantile regression θ, with 0 < θ < 1, is defined as the solution to the minimization problem, which is solved with a simplex linear programming model [5]

$$\min_{\beta \in R^k} \left\{ \sum_{i \in \{i: Y_i \geq X_i\beta\}} \theta |Y_i - X_i\beta_\theta| + \sum_{i \in \{i: Y_i < X_i\beta\}} (1-\theta)|Y_i - X_i\beta_\theta| \right\} \tag{4}$$

The parameter estimation for the QRM is formulated through minimization of the mean absolute deviation with asymmetric weights, whereas in the OLS is obtained by minimizing the sum of squared residuals. The median regression is a particular case of the quantile regression where θ = 0.5, being the only case in which the weighs are symmetric.[5]

The quantile regression has been estimated using the statistical package R “quantreg”— version 5.36. The algorithmic method used to compute the fit is the modified version of Barrodale and Roberts , and is described in detail by Koenker and d’Orey. The goodness-of-fit of quantile models has been calculated from the absolute deviations using the pseudo-R1. It is useful to compare quantile models; however, they are not comparable to the determination coefficient obtained through OLS as it is based on the variance of the square deviations. The pseudo-R1 is obtained as 1 minus the ratio between the sum of absolute deviations in the fully parameterized models and the sum of absolute deviations in the null (non-conditional) quantile model.[5]

Model in 2017:

The purpose of 2017's study is to identify the price determinants of sharing economy based accommodation offers in the digital marketplace (specifically Airbnb.com). A sample of 180,533 accommodation rental offers from 33 cities listed on Airbnb.com is examined. Ordinary least squares (OLS) analysis and quantile regression (QR) analysis are used to investigate price determinants in five categories: host attributes, site and property attributes, amenities and services, rental rules, and online review ratings. The findings have important implications for the design of pricing-suggestion systems for sharing economy based accommodation service providers, such as the price-recommendation tool recently launched by Airbnb. [4]

Linear QR models and linear OLS regression models are used to detect linear relationships between a dependent variable and a set of explanatory variables. The main difference between the model types is that OLS regression models are based on the conditional mean of the dependent 13 variable, whereas QR models are based on the conditional τth quantile of the dependent variable, where τ (0, 1). Therefore, QR goes beyond the analysis of the conditional mean of a dependent variable, providing a more comprehensive description of the conditional distribution. In other words, rather than estimating the average response of the dependent variable to changes in the explanatory variables, QR measures the effects of individual explanatory variables on the whole distribution of the dependent variable. This allows the analyst to uncover hidden price-response patterns that exist depending on the level of prices. [4]

QR is specified as follows (Koenker & Bassett, 1978). Assuming a random variable Y with a probability-distribution function $F(y) = Prob(Y \leq y)$, the τth quantile of Y can be defined as the smallest value of y satisfying $F(y) \geq \tau$: $Q(\tau) = \inf\{y: F(y) \geq \tau\}$, where $0 < \tau < 1$. For n observations of Y, the empirical distribution function is given as $Fn(y) = \sum 1(Yi \leq y)$, where 1(z) is an indicator function that takes the value of 1 if the argument z is true and 0 otherwise. Accordingly, the empirical quantile is defined as follows: [4]

$$Q_n(\tau) = \inf\{y: F(y) \geq \tau\}.$$

This expression is given as an optimization problem below:

$$Q_n(\tau) = \arg\min_{\xi} \left\{ \sum_{i:Y_i \geq \xi} \tau |Y_i - \xi| + \sum_{i:Y_i < \xi} (1-\tau)|Y_i - \xi| \right\} = \arg\min_{\xi} \left\{ \sum_i \rho_\tau |Y_i - \xi| \right\},$$

where rt(u) = u(τ -1(u < 0)) is the so-called check function, which weights positive and negative values asymmetrically. A linear specification of the conditional quantile of the dependent variable gives Q(τ|Xi,β(τ)) = Xi'β(τ), where Xi is the vector of the explanatory variables and β(τ) is the vector of the coefficients associated with the tth quantile. Under these conditions, the previous optimization problem is as follows: [4]

$$\hat{\beta}_n(\tau) = \arg\min_{\beta(\tau)} \left\{ \sum_i \rho_\tau (Y_i - X_i^{'}\beta(\tau)) \right\}.$$

Intuitively, the parameters of QR are estimated by considering different weights of the absolute residuals. To analyze listing prices, the variable of price per person per night (in logarithmic form) is selected as the dependent variable. As the resulting expression is a semi-logarithmic specification, the coefficient values represent semi-elasticities, namely the percentage change in price when an explanatory variable varies by 1, having in mind that the effect of a dummy independent variable on a log dependent variable is measured by eβ -1. [4]

As a result of this work, OLS analysis reveals that 24 of the 25 variables under study are good predictors of price, while QR analysis indicates that all of the variables have significant effects on price, but these effects are often dependent on price range. The findings thus offer insights into the complexities of the price-determinant relationship in sharing economy based accommodation rentals. [4]

3 Aim of The Project

As can be seen in these literature studies, the issue of Airbnb price determination is a very detailed and extensive subject. There are many different price determining variables. There are many methods with different features, advantages, limitations etc. With this project, the most effective attributes in price determination will be determined. A model will be developed for a more effective and more accurate pricing. This model will be accessible by host and owners. Thus, they will be provided with the necessary informations to provide a more helpful, effective and accurate pricing option.

4 Conclusion

As a result of this study, we examined similar studies conducted in 2017, 2018, 2019 and 2020 for our Airbnb price determinant project. We grasped why, when and for what Airbnb, a Sharing economy business model, is used. We conclude that a price determinant system is required for these platform users. We have observed that regression was the method with the most results in price determinant projects that were previously worked for Airbnb and projects in the field of sharing economy. We decided to use the regression model in our own project.

References

[1] Yuan Caı,Yongbo Zhou, Jıanyu (Jenny) Ma, Noel Scott. Price Determinants of Airbnb Listings: Evidence From Hong Kong. May 2019.

[2] Gábor Dudás, Tamás Kovalcsik, György Vida, Lajos Boros, Gyula Nagy. Price Determinants of Airbnb Listing Prices in Lake Balaton Touristic Region, Hungary. February 2020.

[3] Insu Hong, Changsok Yoo. Analyzing Spatial Variance of Airbnb Pricing Determinants Using Multiscale GWR Approach. June 2020.

[4] Dan Wang*, Juan L. Nicolau. Price Determinants of Sharing Economy Based Accommodation Rental: A Study of Listings from 33 Cities on Airbnb.com. 2017.

[5]V. Raul Perez-Sanchez , Leticia Serrano-Estrada * , Pablo Marti and Raul-Tomas Mora-Garcia. The What, Where, and Why of Airbnb Price Determinants. December 2018.

[6]Magno, F., Cassia, F., Ugolini. "Accommodation prices on Airbnb: effects of host experience and market demand". May 2018.

[7]Zhihua Zhang 1 , Rachel J. C. Chen 2,*, Lee D. Han 3,* and Lu Yang 1. Key Factors Affecting the Price of Airbnb Listings: A Geographically Weighted Approach. September 2017.

[8]Juan L Eugenio-Martin, José M Cazorla-Artiles, Christian González-Martel. On the Determinants of Airbnb Location and its Spatial Distribution. January 2019.