

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Löw, Franziska; Lüth, Hendrik

Working Paper Quality signals on Airbnb: A hedonic regression approach

Diskussionspapier, No. 189

Provided in Cooperation with: Fächergruppe Volkswirtschaftslehre, Helmut-Schmidt-Universität (HSU)

Suggested Citation: Löw, Franziska; Lüth, Hendrik (2021) : Quality signals on Airbnb: A hedonic regression approach, Diskussionspapier, No. 189, Helmut-Schmidt-Universität - Universität der Bundeswehr Hamburg, Fächergruppe Volkswirtschaftslehre, Hamburg

This Version is available at: https://hdl.handle.net/10419/242492

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WWW.ECONSTOR.EU

Diskussionspapierreihe Working Paper Series



QUALITY SIGNALS ON AIRBNB: A HEDONIC REGRESSION APPROACH

Franziska Löw Hendrik Lüth

Nr./ No. 189 MAY 2021

Department of Economics Fächergruppe Volkswirtschaftslehre

Autoren / Authors

Franziska Löw

Helmut-Schmidt-University Hamburg Department of Economics Holstenhofweg 85, 22043 Hamburg franzi@localyzeapp.com

Hendrik Lüth Helmut-Schmidt-University Hamburg Department of Economics Holstenhofweg 85, 22043 Hamburg hendrik@lueth.eu

<u>Redaktion / Editors</u> Helmut Schmidt Universität Hamburg / Helmut Schmidt University Hamburg Fächergruppe Volkswirtschaftslehre / Department of Economics

Eine elektronische Version des Diskussionspapiers ist auf folgender Internetseite zu finden / An electronic version of the paper may be downloaded from the homepage:

https://www.hsu-hh.de/fgvwl/forschung

Koordinator / Coordinator Ralf Dewenter wp-vwl@hsu-hh.de Helmut Schmidt Universität Hamburg / Helmut Schmidt University Hamburg Fächergruppe Volkswirtschaftslehre / Department of Economics Diskussionspapier Nr. 189 Working Paper No. 189

Quality Signals on Airbnb: A Hedonic Regression Approach

Franziska Löw Hendrik Lüth

Zusammenfassung / Abstract

This study aims at identifying guests' willingness to pay for characteristics of listings on Airbnb, putting a particular emphasis on trust-building information provided by the platform. In order to do so, a hedonic regression model is applied to an extensive dataset that was gathered in 2017 from Airbnb's website and encompasses listings from seven major German cities, namely Berlin, Munich, Hamburg, Cologne, Dresden, Stuttgart and Frankfurt am Main. Our results regarding tangible characteristics are mostly in line with expectations: Additional space and certain amenities increase the value of a listing. The same holds true for an accommodation's distance to the city center, although we find proof for a non-linear relationship. Results for trust-building factors on the other hand are mixed. While favorable review scores and membership duration have a positive effect on prices, we cannot establish such a relationship for "superhost" and "verified ID" badges. In contrast to other studies, which are, however, focused on the US, we cannot find price differences linked to hosts' gender or ethnicity. Using an extended data set that encompasses listings from 2007 to 2008, we furthermore construct hedonic price indices for all seven cities, which suggest supply shifts due to regulatory pressure.

Schlagworte / Keywords: airbnb, sharing economy, hedonic models, GLM, rental markets, price indices, real estate

JEL-Klassifikation / JEL-Classification: L11, L14, L86

1 Introduction

Fourteen years after its launch, Airbnb has become one of the leading marketplaces for the travel industry, connecting travelers with local hosts in more than 100,000 cities and 220 countries (Airbnb, 2021). Similar to online platforms like eBay and Amazon Marketplace, Airbnb enables customers ("guests") to look for offers from other customers ("hosts") and facilitates transactions between them. One of the key factors behind Airbnb's success is the company's ability to increase the platform's value to participants by improving the likelihood and quality of matches. In order to do so, Airbnb offers hosts a variety of tools to assess the true market value of their listings and set prices properly. Additionally, several trust building mechanisms are embedded into the platform, which reduce information asymmetry among hosts and guests.

In traditional markets, repeated transactions and face-to-face interactions can increase trust between market players and therefore reduce the occurrence of inefficiencies caused by information asymmetry. Mechanisms like warranties, advertising and quality reviews, provided by independent agents, also contribute to the reduction of asymmetric information (Dewally and Ederington, 2006). However, digital peer-to-peer (P2P) markets, which allow small businesses or even individuals to engage in transactions with buyers, are prima facie far more anonymous, as the contract parties rarely meet and often play a one-shot game from a game theory point of view (Belk, 2010; Botsman and Rogers, 2010; Hawlitschek, Teubner, and Weinhardt, 2016).

In contrast to online marketplaces like eBay, where physical products are traded, online booking platforms like Airbnb offer a marketplace for a more personal (and usually non-refundable) good. Consequently, the stakes are higher in many cases. At the same time, digital markets can potentially provide an enormous amount of information that can reduce transaction costs induced by information asymmetries, given that the information is wellstructured and presented comprehensively. As trust is a prerequisite for P2P markets in order to function efficiently, platform operators like Uber, Airbnb and eBay, which facilitate trade between large numbers of fragmented buyers and sellers, have to create a frictionless market design to match both sides efficiently (Einav, Farronato, and Levin, 2015).

As mentioned above, the likelihood of matches between hosts and guests also depends on the balance between a listing's price and value. While commercial accommodation providers base their pricing decisions on market data, setting an optimal price is difficult for many hosts, as they often don't know the real value of their accommodation. To support hosts in their pricing decision and therefore improve the likelihood of matching, Airbnb offers a dynamic pricing-tool that recommends a price based on multiple listing attributes. The recommended price is calculated using a classifier algorithm comparing listings with similar attributes.¹ However, the host can still choose the price manually.

Based on a dataset of Airbnb listings in seven German cities,² we use a hedonic regression model to estimate price determinants of listing prices. In contrast to real estate markets, where hedonic models are widely used and where property values are determined by "hard" characteristics like location and size, the price of an Airbnb listing may be affected by additional information. Besides material listing characteristics like the above as well as accommodation type (house, apartment, loft etc.) and room type (entire house, private room, apartment), a listing's value is also affected by trustbuilding factors like review scores and host attributes.³ This is because trust signals like these can reduce uncertainty about the accuracy of an accommodation's description and furthermore reveal information that is not explicitly provided by the host. Our study aims at disentangling the contributory value of a listing's characteristics, thereby giving insight into the relevance of trust signals for price determination. Our regression results underscore the importance of ratings from Airbnb's review system, as good reviews are associated with a host's ability to charge higher prices. The duration of a host's Airbnb membership also contributes to a higher valuation. However, our results don't suggest a positive effect of other presumedly trust-building variables like superhost status and host verification. Personal traits of hosts like gender, name origin and facial expression on the profile picture also don't result in higher valuations.

Building on our results from the hedonic regression, we furthermore construct hedonic price indices for Airbnb markets in each investigated city. With these indices, we can assess price movements on Airbnb while also taking quality-changes into account. Between April 2017 and March 2018, quality-adjusted price levels increased in all investigated cities by 15-25%, in line with the general upward trend for long-term rents in urban areas. Interestingly, our results hint at effects of regulatory intervention on the quality composition of Airbnb's offerings. In cities where the public administration ramped up regulatory pressure on illegal letting of vacation rentals, a decline in median prices accompanied by constant hedonic price levels and a decline in the overall number of offerings suggests market exit of listings at the high-quality end, presumedly owned by professional Airbnb landlords.

¹For more information about the pricing algorithm, see Hill (2015) and Yee and Ifrach (2015).

²Berlin, Munich, Hamburg, Cologne, Frankfurt a.M., Stuttgart, Dresden.

 $^{^{3}}$ We outline our variable selection in Section 3.

2 Literature Review

In the academic literature on the sharing economy, trust has become a prominent field of research (Cheng, 2016). A large strand of the theoretical literature argues that quality signals like rating mechanisms and identity verification can mitigate inefficiencies in markets with asymmetric information, as they provide structured and transparent information about both market sides.⁴ In this context, much research has been conducted on the rating mechanism of the auction platform eBay (McDonald and Slawson, 2002; Resnick and Zeckhauser, 2002), with the general finding that negative ratings lead to a drop in weekly sales (Cabral and Hortaçsu, 2010), and have a negative price effect (Houser and Wooders, 2006; Melnik and Alm, 2002). Klein, Lambertz, and Stahl (2016) use a large dataset in order to estimate a fixed effects model that analyzes the qualitative and quantitative effects of a change in the eBay rating system on adverse selection and moral hazard. They find that an improvement of the rating mechanism decreases the costs linked to information asymmetries. In the hotel industry, online reviews have a big impact on booking decisions, especially for first time users (Panda, Verma, and Mehta, 2015). However, customer reviews can be biased or misleading when customers who submit reviews are not compensated for their efforts (Fradkin, Grewal, and Holtz, 2021). For Airbnb, the sharing economy's leading platform for accommodation rentals, Zervas, Proserpio, and Byers (2021) examined 60,000 listings and found that 95% of them have an average rating of 4.5 or 5 stars (the highest rating) in contrast to approximately half a million hotels worldwide collected on TripAdvisor, where there is a much lower average rating of 3.8 stars, and more variance across reviews. This leads to the assumption that ratings on Airbnb are possibly biased.

Several recent studies have investigated the price formation mechanism on Airbnb using different specifications of the hedonic price model, with the common finding that the average review score has a positive effect on price(Chen and Xie, 2017; Gibbs et al., 2018; Teubner, Hawlitschek, and Dann, 2017; Dan Wang and Nicolau, 2017). Teubner, Hawlitschek, and Dann (2017) estimate a linear model to quantify the economic value of trust on Airbnb using a dataset containing about 16,000 listings. They find a significant positive price effect of several trust-building factors like the average rating score, duration of a host's membership or the number of photos shown for a listing. Chen and Xie (2017) use a dataset of 5,779 Airbnb listings in Austin, Texas, to estimate a quadratic semi-log specification that tests the effects of a group of utility-based attributes on the price of Airbnb listings,

⁴See Bar-Isaac and Tadelis (2008) for a survey.

including the number of hotels in the same census tract, to measure the effect of market competition. In another recent study Dan Wang and Nicolau (2017) apply a quantile regression approach in order to analyse a dataset of 33 different cities. They argue that host attributes are more accurate price predictors than the star ratings traditionally used in the hotel industry. In contrast Ert, Fleischer, and Magen (2016) do not find any price effect of hosts' review scores using Airbnb data from Stockholm, Sweden. Conducting a controlled experiment, they found evidence that the more trustworthy hosts are perceived from their photos, the higher the price of their listings and the more likely they are being booked. Deboosere et al. (2019) applied a hedonic regression to Airbnb listings in New York City and find that locational factors like transit accessibility have a strong influence on prices. Lorde, Jacob, and Weekes (2019) investigate Airbnb accommodations in the Caribbean in a cross-country analysis and find significant effects for accommodation and host properties that are generally in line with expectations, supporting the hypothesis that tangible features are equally important as reputational ones. In contrast to the majority of research on Airbnb, which is concerned with urban rental markets, Moreno-Izquierdo et al. (2019) focus on sun and beach destinations in Valencia.

Other research topics related to Airbnb have addressed regulation (Jefferson-Jones, 2015) and the platform's impact on the hotel industry (Zervas, Proserpio, and Byers, 2017). B. G. Edelman and Luca (2014) find evidence for racial discrimination against Afro-American hosts in New York City as they are forced to charge lower prices than comparable listings of non-black hosts.⁵ Within a field experiment B. Edelman, Luca, and Svirsky (2017) estimate that regarding the hosts' market side, Afro-American guests are less likely to be accepted than white guests.

However, none of the aforementioned studies have calculated a hedonic price index. Such an index is computed from a hedonic function, which describes the relationship between prices of a product's different varieties and their characteristics (Triplett, 2006). The hedonic regression approach dates back to the work of Court (1939), who developed this very concept to evaluate automobiles. In the field of applied economics, the work of Rosen (1974) and Lancaster (1966) laid down the conceptual foundations of consumer valuation of heterogenous products.⁶ The hedonic price model describes the procedure of regressing the price of differentiated goods on it characteristics, where the estimated coefficients represent a characteristic's discrete hedonic price (Dale-Johnson, 1982). The hedonic pricing methods is a prominent approach

⁵A similar effect is estimated for Hispanic and Asian hosts by Kakar et al. (2018). ⁶See also Griliches (2014).

in real estate studies, where housings are differentiated goods that exhibit a bundle of characteristics that completely describe the respective property (Kang and Reichert, 1991; Strand and Vågnes, 2001; Wilhelmsson, 2002).⁷

Other studies applying hedonic regression methods in order to assess price determinants on Airbnb often include the number of reviews as independent variable (Gibbs et al., 2018; Teubner, Hawlitschek, and Dann, 2017; Dan Wang and Nicolau, 2017). This implicitly assumes the number of reviews is exogenous, a crucial assumption for estimates to be consistent. Taking into account that nearly 70% of guests leave a review (Fradkin, Grewal, and Holtz, 2021), one can interpret the number of reviews as a proxy for the demand of the respective listing. Thus including this variable as an independent regressor leads to biased and inconsistent estimators.⁸ Furthermore, existing studies do not adequately control for city-specific differences within their estimation model, ignoring the fact that huge variation exists both in the supply and demand of geographic submarkets that lead to a different structure of prices in each (Straszheim, 1974). Prior studies, which estimate hedonic price functions with a dataset that includes more than one city, try to control for urban heterogeneity using city attributes like the population size and rental price (Teubner, Hawlitschek, and Dann, 2017) or country effects Dan Wang and Nicolau, 2017.⁹ However, one can argue that these structural controls can never sufficiently capture the spatial variation in the market (Dale-Johnson, 1982).

3 Data

Our initial dataset contains information on 18,052 Airbnb listings from seven major German cities, namely Berlin, Munich, Hamburg, Cologne, Dresden, Stuttgart and Frankfurt am Main. Listings were gathered directly from Airbnb's website at first in April 2017¹⁰ using a custom web scraper and then processed by extracting the relevant information from the listings' HTML source code.¹¹ The scraped data entails almost all information that is visible

⁷Malpezzi (2002) presents a review of hedonic pricing literature.

⁸Dan Wang and Nicolau (2017, p. 128) note that "cheaper listings tend to receive more bookings and consequently more reviews", recognizing the dual causality. Nevertheless, they included the number of reviews as an independent variable.

⁹Gibbs et al. (2018) use a dataset of five large urban destinations in Canada to separately estimate the hedonic price function for each of the city.

¹⁰Our dataset includes every visible listing with more than three reviews on the date we scraped the website. The exact dates of scraping include 12 to 16 April 2017.

¹¹We deployed a heavily modified and extended version of a web scraper originally programmed by Tom Slee (http://tomslee.net). The customized web scraper can be

to visitors of an Airbnb listing, including (but not limited to) prices, accommodation features, reviews and host details. A detailed description of the Airbnb user-experience can be found in Gibbs et al. (2018). For our construction of hedonic price indices (see Section 5.3), we additionally scraped Airbnb's website in July and September 2017 as well as in March 2018.

We use several public application programming interfaces (APIs) to further expand our dataset and to obtain structured information from unstructured data like a host's profile picture. More precisely, we use the Google Maps Distance Matrix API¹² to calculate the walking distance from a housing to the respective city center and the Google Maps Geolocation API¹³ to determine the neighborhood an accommodation is located in, based on its geographical coordinates. Furthermore, we make use of Microsoft's Azure Emotion API¹⁴ to analyze facial expressions in a host's profile picture. We thereby obtain eight numerical weights for the different emotions expressed in a picture, but only use the value for "happiness" (in a continuum between 0 and 1) in order to determine if hosts smile on their profile pics.

Additionally, hosts' first names as displayed by Airbnb are analyzed in respect to gender and origin. We access a database of gender predictions provided by genderize.io,¹⁵ which is compiled from social media profiles, to determine the gender of hosts by their first name. Furthermore we use the Behind the Name API¹⁶ to determine host names' ethnic origin with the intent to unveil potential discrimination against hosts from ethnic groups of non-German origin.

As the average rating score of a listing is first displayed when it has at least three reviews, we exclude listings that have not reached this threshold yet. Furthermore, in order to increase the comparability among the investigated listings, we only consider listings with accommodations suitable for less than six guests, with a price less than $\leq 1,000$ and a property type that is defined as apartment, house, loft, townhouse or villa.

Tables 1 and 2 display summary statistics for most of our variables,¹⁷ a correlation matrix for all variables is attached in Appendix A2. Airbnb distinguishes thee different room types that are captured in the the respective variable. Guests can either book an entire home (57.6%), a private room within the accommodation (41.1%) or a shared room (1.3%). As expected

made available by the authors upon request.

¹²https://developers.google.com/maps/documentation/distance-matrix/

¹³https://developers.google.com/maps/documentation/geolocation/

¹⁴https://azure.microsoft.com/en-us/services/cognitive-services/emotion/

¹⁵https://genderize.io

¹⁶https://www.behindthename.com

¹⁷Due to space and readability, we do not include all variables in the tables.

for a developed accommodation market like Germany, which attracts mostly travelers of medium and higher income brackets, shared rooms are rarely offered and only account for a small fraction of total listings. Consistent with the known professionalization of Airbnb landlords, entire homes that are inhabited exclusively by a single renter at a time constitute the majority of supply.

Variable	Levels	n	%
Room Type	Entire Home	10,390	57.6
	Private Room	7,421	41.1
	Shared Room	241	1.3
	all	18,052	100.0
Cancel Policy	Flexible	$5,\!499$	30.5
	Moderate	6,914	38.3
	Strict	$5,\!639$	31.2
	all	18,052	100.0
Smoking Allowed	No	$15,\!271$	84.6
	Yes	2,781	15.4
	all	$18,\!052$	100.0
Verified Host	No	7,312	40.5
	Yes	10,740	59.5
	all	18,052	100.0
Gender	Female	8,305	54.3
	Male	$6,\!978$	45.7
	all	15,283	100.0
Security Deposit	No	11,143	61.7
	Yes	6,909	38.3
	all	$18,\!052$	100.0
TV	No	6,979	38.7
	Yes	$11,\!073$	61.3
	all	18,052	100.0

 Table 1: Descriptive Statistics for Chosen Categorical Variables

At the time when listings were gathered for this research, Airbnb offered its hosts a choice among three different cancellation policies to different degrees of rigour: flexible, moderate and strict. Under the flexible cancellation policy, guests receive a full refund on the reservation cost as long as cancellation is more than 24 hours before check-in time, minus booking fees. When the moderate policy applies and a guest cancels less than 5 days in advance, the first night is non-refundable, but 50% of the reservation cost for the remaining nights will be refunded. A full refund (excluding booking fees) is only given when the cancellation is made more than 5 days before check-in. If a host decides for strict cancellation rules, guests will only be

Variable	$\bar{\mathbf{x}}$	$\widetilde{\mathbf{x}}$	\mathbf{Min}	Max	$\#\mathbf{NA}$
Price	64.8	55.0	9.0	1000.0	0
Accomodates	2.7	2.0	1.0	6.0	0
Reviews	24.7	12.0	4.0	435.0	0
Bathrooms	1.1	1.0	0.0	5.5	1252
Bedrooms	1.2	1.0	1.0	6.0	1197
Minimum Stay (days)	2.2	2.0	1.0	27.0	0
Overall Satisfaction	4.7	5.0	1.0	5.0	4
Pictures	12.7	11.0	1.0	160.0	0
Distance to City Center (km)	4.6	4.0	0.0	29.6	1
Membership (Months)	34.4	33.0	1.0	106.0	0
Happiness	0.5	0.7	0.0	1.0	0

Table 2: Summary Statistics for Chosen Numerical Variables

refunded 50% of the reservation costs when they cancel more than one week in advance and no refund will be given afterwards. Interestingly, the cancel policies are almost evenly distributed between strict (31.2%), moderate (38.3%) and flexible (30.5%).

In most of the listed accommodations smoking is not allowed, mirroring the larger societal trend towards abolishing indoor smoking. Smoking is only permitted in 15.4% of the apartments and houses. Regarding one of the trust building factors, about 60% of hosts are verified by Airbnb, meaning that they have earned a "Verified ID" badge by matching their online identity to offline ID documentation. This can be achieved by scanning a photo ID that corresponds to the personal information entered during signing up for Airbnb. Regarding hosts' gender there is a prevalence of females (54.3%)which exceeds the sex ration seen in the general population. Requesting a security deposit is apparently not the norm on Airbnb's platform as only 38.3% of hosts ask for it. This is an interesting contrast to the traditional practice of letting rental homes, where deposits are prevalent. Exemplary for other amenities Table 1 also depicts the proportion of rentals that offer a TV set to guests (61.3%). Access to wireless internet and cooking facilities are even more common in our dataset, as about 95% provide Wifi and 94% offer a kitchen for either exclusive or shared use. Two thirds of our hosts have a name that is determined as natively German by the API we use, which opens up the opportunity to investigate an effect of ethnic background on perceived trustworthiness that may also be reflected in prices.

Regarding the available property types, our raw dataset contains 26 different property types including unusual ones like caves, tree houses and tents among others. We reduced the dataset to apartment (94%), house (0.04%), loft (0.012%), townhouse (0.003%) and villa (0.001%) and thereby eliminate 513 observations. We constructed a dummy variable that becomes 1, when the property is a house (including house, townhouse and villa), and 0 otherwise. Only 4.8% of all listings represent houses, which is in line with our expectations for the urban rental markets contained in our dataset.

Numerical variables in Table 2 reflect both properties of accommodations and hosts. Listings in our dataset can accommodate two persons in the median and usually entail one bedroom and one bathroom. However, rentals with more lavish layouts with up to 5.5 bathrooms¹⁸ and 6 bedrooms are also available. Median walking distance from the accommodation to the city center is 4 km, ranging from immediate vicinity to 29.6 km.

On average each listing has 24.7 reviews while the median is at only 12 reviews, meaning the distribution of reviews is positively skewed. The variable overall satisfaction is at the core of Airbnb's user review system and reflects satisfaction of guests who booked the respective accommodation. Manifestations of this variable can range between 1 and 5 stars, but are predominantly in the upper range with a median (and mode) of 5 stars. The observation that buyers and sellers give exceptionally high ratings to each other is not uncommon for P2P markets like Airbnb and has also been noticed on Ebay (Bolton, Greiner, and Ockenfels, 2012).

Ert, Fleischer, and Magen (2016) discuss two possible explanations for biased review scores on these platforms. Firstly, mutual feedback systems might deter platform users from posting negative reviews because they fear retaliation from their counterpart. This potentially leads to a selection bias where guests with an unfavorable opinion refrain from posting a fitting review. However, Airbnb changed its review mechanism in 2014 in order to counter this potential source of bias. From then onwards reviews and review scores were revealed simultaneously to hosts and guests not until both have submitted their assessment or alternatively a 14-day period has passed after checkout. As a consequence hosts could not retaliate against unfavorable reviews any longer, which potentially encouraged guests to share negative experiences more often. Ert and Fleischer (2019) investigated this assumption and indeed found evidence that implementing this change to the underlying review mechanism significantly decreased overall review scores on Airbnb. As data collection for our empirical analysis started in 2017 and about three years after the revised review mechanism was introduced, we can assume that the change has been sufficiently absorbed in guests' reviewing practices and fear of retaliation is rarely cause for overenthusiastic (or withheld negative) reviews anymore.

An additional potential source for an upward bias discussed by Ert, Fleis-

¹⁸Half-baths are bathrooms with a sink and toilet, but no shower or bathtub.

cher, and Magen (2016) is the fact that Airbnb users experience personal contact in context of the market place transaction. Due to personal acquaintance with the host before and during their stay, guests might be more reluctant to give a negative review and be more understanding towards any shortcomings in general. This is underlined by the finding of Ert, Fleischer, and Magen (2016) that guests in Airbnb reviews often refer to their hosts on a first-name basis in contrast to reviews on Booking.com, where reviewers refer to anonymous staff members. However, even under the assumption that submitted review scores trump the actual experience, this would not lead to biased estimation results in our empirical setting. Since we want to explicitly estimate the effect of posted ratings (and not the effect of former guests' experiences) on prices, deviations between posted reviews and privately held opinions do not present an issue for our hedonic valuation approach.

An additional variable that may reassure hosts is the membership duration of the respective host, measured in months. Our dataset includes listings by new-fledged hosts that have just signed up for Airbnb as well as hosts with longstanding experience and reputation with a maximum membership duration of almost 9 years. Hosts are registered with Airbnb for 33 months in the median. Providing pictures of the accommodation is another feasible way to build trust among potential guests and according to our data many hosts do so accordingly. While 11 pictures are shown in the median, some hosts provide much more and uploaded up to 160 pictures for a single accommodation. As we also look out for more subtle influences on trustworthiness and subsequently prices, facial expression are evaluated. The majority of hosts smiles on their profile pictures, i.e. our happiness variable derived from Microsoft's Azure Emotion API, which ranges between 0 and 1 per profile picture, is much greater than zero on average (0.5) and in the median (0.7).

A graphical representation of a selection of these variables can be seen in Figures 1a - 1d. The average listing price¹⁹ within our dataset is $\in 64.8$ with a median of $\in 55$. However, as can be seen in Figure 1a, the median differs between cities. The biggest difference is between Munich ($\in 65$) and Stuttgart and Dresden (both $\in 45$). Furthermore, the walking distance to the city center from Airbnb accommodations varies between cities. For example the median walking distance of listings in Berlin is 5.1 km, whereas it is 2.5 km in Frankfurt. Additional data visualizations on the city level can be found in Appendices A3 and A4.

Figure 1c shows that the distribution of the star-ratings within our sample is left-skewed, over 90% of all listings have a star rating of 4.5 or higher.

¹⁹The displayed price of an Airbnb listing for one person and one night without specified booking dates.

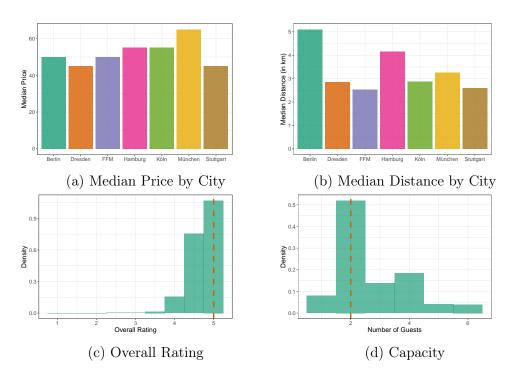


Figure 1: Median and Density Plots of Selected Variables

The phenomenon of ratings that are overwhelmingly positive and skewed to maximum scores can not only be found for Airbnb (Chen and Xie, 2017), but for other online platforms like eBay as well (Nosko and Tadelis, 2015). After eliminating all listings with a capacity higher than six guests, nearly 52% of all listings are suited for two guests.

4 Empirical Strategy

Following the standard hedonic price model as described above, we observe a set of listing prices at time t, which we denote as p_{it} and where i is the index of a particular listing. As the listing prices differ due to differences in their properties, we can write p_{it} as a function of a set of characteristics \boldsymbol{x} .

$$E(p_{it}) = f_t(\boldsymbol{x}'_{it}\boldsymbol{\beta}) \tag{1}$$

The form of the function f has a direct bearing on the interpretation of a hedonic price index. As there is no a priori reason to expect price and attributes to be related in any particular fixed fashion, the selection of an adequate functional form is an empirical question. Our dependent

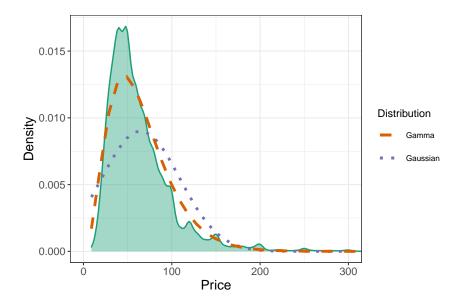


Figure 2: Density of Price

variable is continuous, but is right-skewed and always positive (see Figure 2). The most common way to analyze such data and to address the problem of non-linearity, and mitigate the problem of heteroscedasticity, is to log transform the outcome variable (Malpezzi, 2002). Alternatively, we follow an approach that builds on the generalized linear model (GLM), where the response variable p_{it} is assumed to follow an exponential family distribution with mean μ_i , which is assumed to be some function of $\mathbf{x}'_{it}\boldsymbol{\beta}$. Thus, even with non-normal errors, our outcome variable is linear in parameters (McCullagh and Nelder, 1989). As f is monotone, there is a unique function g, called the inverse function of f, such that we can rewrite Equation 1 as

$$g\{E(p_{it})\} = \boldsymbol{x}'_{it}\boldsymbol{\beta} \tag{2}$$

Function g is also called the link function, as it relates the linear predictor to the predicted mean of the response. For a classical linear model, the link function is the identity function. In our case, as suggested by the distribution of our price variable, we choose a log-link function, such that we estimate the following model:

When regressing the price of a listing on its characteristics, the estimated coefficients can be interpreted as the consumers' implicit valuations of these

characteristics. Such an implicit price is termed as the hedonic price, or shadow price, of an attribute.

Our hedonic price function in Equation 2 can subsequently be computed for multiple periods. The implicit prices we obtain will then depend on the particular period. If the weight pattern of the periods are not too different and the equation holds well enough in all our periods, we can estimate the average price change between periods directly by estimating Equation 4 and assessing a period's respective time dummy coefficient τ_i (Griliches, 1961):

$$log\{E(p_{it})\} = \beta_0 + \beta_1 x_{1it} + \beta_2 x_{2it} + \dots + \sum_{j=1}^{T-1} \tau_j T_j + \epsilon_{it},$$
(4)

where ϵ_{it} is *iid* and τ_j is the coefficient of time dummy T.

As in any regression model, the selection of independent variables is a matter of the empirical problem at hand. We divide our set of variables into several categories that we add subsequently to our base model in order to test the validity of our results. Within our base model, we include variables of which we assume that they will have the biggest impact on listing price. The correlation matrix in Appendix A2 does not hint at any multicollinearity issues within our dataset.²⁰ As stated above, we argue that the number of reviews is not an exogenous variable, and therefore including it would cause biased estimates of other coefficients (Frölich, 2008). Excluding it means that the number of reviews is part of the error term, which would be an issue if it is correlated with the remaining regressors. As no significant correlation can be found, we decided to exclude the the number of review from our model apart from a robustness check in Section 5.2.

A crucial assumption made by recent studies (Gibbs et al., 2018; Teubner, Hawlitschek, and Dann, 2017; Dan Wang and Nicolau, 2017) is that the price of a listing reflects the market equilibrium price even though we do not observe real market transactions. We assume rational hosts who maximize their income by setting prices that reflect the value of their listing. This assumption is supported by the fact that we only analyze listings with more than two reviews, as the average rating score of a listing is first displayed if it has at least three reviews. Of course listing prices will occasionally deviate from their actual value and consequently impact occupancy rates. But there is no evidence that this would occur in a systematic way, which would lead to biased regression results with biased hedonic prices. In further support of using ask prices for hedonic regressions, Ye, Law, and Gu (2009) confirm

 $^{^{20}}$ The number of bedrooms and the capacity are correlated as expected, but their correlation coefficient (0.59) remains under any threshold that would be cause for concern (Dormann et al., 2013).

a significant relationship between online consumer reviews and hotel-room sales, indicating that reviews reflect real transactions.

In contrast to real estate markets, where ask prices from online platforms will often differ from actual transaction prices, rental rates on Airbnb are not negotiated between hosts and customers. Subsequently a general upward bias in offered room prices, as it could be reasonably assumed for offer prices on real estate platforms, is unlikely in the case of Airbnb. It is conceivable that newer hosts might overestimate the value of their property and therefore set prices higher than in the long-term equilibrium. However, we do not find any evidence in our data that suggests this is the case, as prices do not negatively correlate with membership duration. Furthermore, as Airbnb is competing with the hotel industry, the platform will try its best to nudge hosts to set competitive prices.²¹ Although we can reasonably assume offer prices translate into transaction prices and no bias results from the chosen approach, we should nevertheless bear in mind that we do not observe transaction prices and conclusions rely on the assumption that ask prices overall reflect the actual value of listings on Airbnb. If, for example, hosts indifferent to depressed occupancy rates indeed set systematically higher prices than guests are willing to pay, hedonic prices resulting from our regression are potentially upward biased compared to actual implicit prices.

As the overall rating of a listing constitutes a significant quality signal, we expect higher ratings to have a positive impact on the price. The same is true for the number of pictures, as pictures can offer visual information about the "product" a guest is about to pay for. Other variables within the base model describe the characteristics of a listing and are therefore assumed to have an impact on the price: room type, capacity, number of bedrooms, number of bathrooms, if a listing is a house or some other type of building, and if a listing offers a real bed (as opposed to e.g. pull-out sofa).

The definition of geographical submarkets is another important task when estimating a hedonic regression model, as demand and supply is likely to diverge between these geographical submarkets, because types of housing as well as the demand for housings of a given type at a given location vary. If observations are affected by variables that operate at a higher level (e.g. region), the assumption of random sampling, i.e. error terms are i.i.d., might be invalid. Straszheim (1974) shows that by estimating separate hedonic price functions for different geographic areas of the San Francisco Bay area, the sum of squared errors (SSE) in predicting prices across the entire sample were significantly reduced. A usual attempt to control for these submarket differences is the inclusion of area specific variables like the population size

²¹The intelligent price setting mechanism provided by Airbnb supports this assumption.

or the rent price level (Teubner, Hawlitschek, and Dann, 2017).

However, it is likely that these variables not only affect the intercept terms, but also the coefficients of the characteristics that describe the individual unit. The resulting regression could then yield biased estimates of the guests' marginal willingness to pay for key listing attributes. We use city dummies to control for the variation between cities and distance to center variables to control for within-city heterogeneity. Besides distances derived from the Google Maps Distance Matrix API, we also included distance measures calculated with the haversine formula²² in our different model specifications. We also recognize the possibility that localization in different districts in a city may have an impact on the value of a listing, regardless of their distance to the city center. In Figure 3 we can observe that for the larger cities (Berlin, Hamburg, Munich, Cologne) Airbnb listings are centered around two or more areas.²³ To control for the impact of a certain district, we also estimate an alternative model including district dummies.

The first category we add to our base model include dummy variables about additional amenities a listing offers (Wifi, TV, a Kitchen, Breakfast) and rules guests have to obey during their stay. Regarding the amenities we assume that each of them increases the value of a listing and therefore its price. In respect to the apartment policies, direction of effects is not that clear a priori. Allowing guests to smoke or to bring their pets may reduce the value of the listing as it reduces the expectations regarding cleanliness. The claim of a security deposit on the other hand may be seen as an indicator for the value of a listing: accommodations that are well equipped with high quality furniture are more likely to urge a security deposit.

The second category of variables added to the above set of regressors provide a description of the host himself. Direct interactions between seller and buyer make the hosts' (as well as guests') attributes an important factor when entering a contract. Airbnb offers a wide range of information on hosts that can build trust and therefore supposedly has a positive impact on price: The duration of membership, if the host is a superhost²⁴ and if a host's ID is verified.

Lastly, we convert the profile picture as well as the name of hosts into quantifiable information using the methods and APIs described in Section 3. As Ert, Fleischer, and Magen (2016) show that hosts' photos have an impact on listings' prices, we assume that the happier hosts are perceived from their

 $^{^{22}{\}rm The}$ have rsine formula gives the great-circle distances between two points on a sphere from their longitudes and latitudes.

²³Enhanced maps can be found in Appendix A1.

²⁴A superhost has to satisfy several requirements such as considerable experience, a high response rate and 5-star reviews.

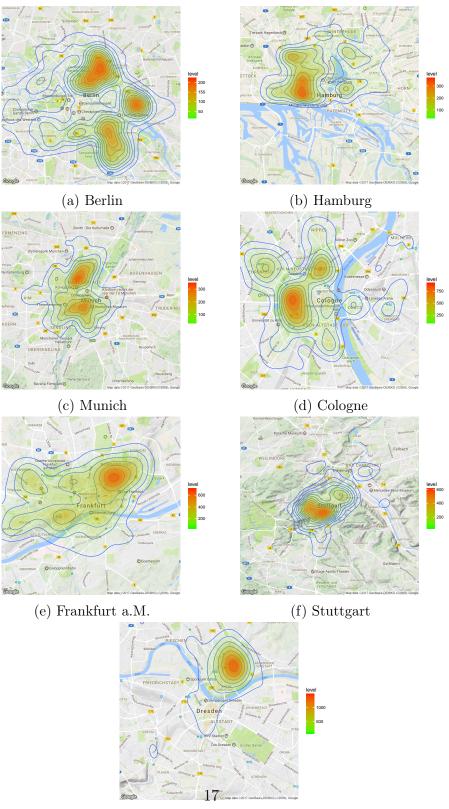


Figure 3: Spatial Distribution of Airbnb Listings

(g) Dresden

photo, the higher the price of their listings. Furthermore, we control for hosts' gender as well as for host names' ethnic origin in order to measure the impact of non-German names on price formation.

5 Results

5.1 Hedonic Regression Results

Estimating six different specifications of Model 4 yields the regression results shown in Table 3 with price per one person and one night as dependent variable. A parsimonious model with only fundamental properties already shows the expected signs for regression parameters, which are all significant at the 0.01 level with the exception of bed type only. The overall rating has a significant impact on price, which is increased by 12.3% ($e^{0.117} - 1=0.1229$) for every rating increase of 0.5 stars. As discussed in Section 3, a potential upward bias in review scores, meaning guests' posted reviews overrate the actual experiences, is irrelevant for the validity of our hedonic regression results. We can nevertheless hypothesize that in a review system where no upward bias in rating scores is present and grades are more evenly distributed, guests would be more tolerant towards less than perfect scores, resulting in a weaker relationship between overall rating and price.

While prices increase in a housing's capacity (11.4% per person), private rooms and shared rooms display high discounts in comparison to entire apartments (-34% and -40% respectively). The number of bedrooms (17.4%) and bathrooms (18.2%) have significant positive coefficients just as the provision of real beds (2.8%) in contrast to pull-out sofas, airbeds or similar. Prices are also significantly higher (9.3%) for accommodations that are entire houses. An increase in the number of photos a listing provides is associated with higher prices, for every additional photo we observe a price increase of 0.6%. Consequently, guests spend more when hosts reduce the uncertainty about a listing's quality by providing more photos.

Consistent with empirical findings for hotel markets (Bull, 1994; Hung, Shang, and F.-C. Wang, 2010; Lee and Jang, 2012; Monty and Skidmore, 2003; Thrane, 2007), we find a negative effect of the distance to city center variable. As touristic attractions are usually located in close proximity to a city center, guests are willing to pay a premium if a listing is in the vicinity of the center. Our estimation results suggest that prices decrease by -4.2% per kilometer.²⁵

 $^{^{25}}$ In Section 5.2 we estimate a model variation assuming a non-linear relationship between price and distance to the city center.

Our results remain stable with changes in the model specification and coefficients keep the expected signs when control dummies for amenities and house rules are added in model (2). Having a TV set in the accommodation allows hosts to charge a premium of 8.8%, while the positive effect of wifi access is surprisingly weaker with only 2.7% surcharge and significance at the 0.1 level. If breakfast is provided at an accommodation, prices are 2.9% higher.

Our findings for the impact of certain house rules that hosts impose on their guests are somewhat surprising at first glance, as more liberties tend to lower prices in some cases. Accommodations that allow smoking within the apartment are cheaper by -3% and allowing pets in the apartment leads to a price decrease of -2.6%. Hence, although smokers and pet owners might be pleased by these permits, the majority of guests seem to be deterred by the prospect of cigarette smell and animal hair, resulting in the observed negative coefficients. Allowing children in an accommodation in contrast increases the value of an accommodation, as the estimated parameter suggests 3% higher prices.

As expected coefficient estimates for city dummies show a pronounced heterogeneity in price levels among the seven cities in our dataset (see Figure 4), which cannot be explained by other regressors. Within model specification (2), Dresden's city dummy coefficient estimate (-0.369) suggests 31% lower prices for otherwise equal listings in comparison to Berlin, whereas hosts in Munich and Frankfurt can charge a premium of 28% and 11% respectively. Hamburg (-1.5%) and Cologne (-4%) on the other hand show only a minor negative price differential when compared to our baseline Berlin.

Adding additional information on hosts as in model (3) to (6), which might influence if guests perceive a host as trustworthy, shows mostly inconclusive and negligible interdependencies with accommodation prices. Hosts that are longtime members of Airbnb can charge slightly more for their listings as the estimated semi-elasticity for an additional month of membership is consistently 0.2%. Negative parameter estimates for superhost status and ID verification on the other hand seem less straightforward. A conceivable explanation is that since obtaining superhost status from Airbnb is linked to the number of reviews hosts receive over a certain time span, hosts that attract more guests by setting lower prices are more likely to be superhosts. A negligible factor for accommodation prices is the number of friends a host has accumulated on Airbnb, p-values remain above the 0.1 threshold for all models.

Extending the model with more detailed and personal variables on hosts as done in model (4), does not reveal additional determinants for accommodation prices. A host's gender does not explain differences in prices, which is

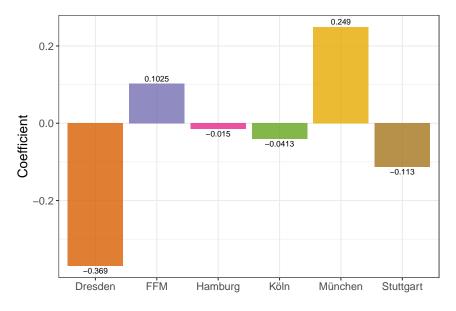


Figure 4: City Dummy Coefficients with Berlin as Baseline - Model (2)

in line with results obtained by other authors (Kakar et al., 2018). The coefficient of our name heritage variable, which indicates German host names, is insignificant as well. This is somewhat in contrast to findings from other authors (B. G. Edelman and Luca, 2014; Gilheany, David Wang, and Xi, 2015; Kakar et al., 2018), who reveal that hosts from ethnical minorities must charge significantly lower prices than their caucasian counterparts when controlling for other factors.

However, there are multiple possible explanations for these differences. First and foremost, we analyze listings only from German cities, whereas the aforementioned studies that detect racial bias are investigating Airbnb offerings from the US, where ethnicity-related heterogeneity is more common, as it is already apparent from simple summary statistics. The median price in our sample is $\in 55$ for hosts with German names as well as for hosts with foreign names, whereas B. G. Edelman and Luca (2014) show that black Airbnb hosts' median prices are \$37 lower in comparison to other ethnicities, even before taking other factors into account. Owing to our large sample set, the construction of our ethnic origin variable however differs significantly from other approaches. In contrast to studies that rely on manual visual inspection of host profile pictures and subsequently very small numbers of observations (B. G. Edelman and Luca, 2014; Gilheany, David Wang, and Xi, 2015; Kakar et al., 2018), we derive ethnicity from host names algorithmically using a name origin database.

Results
Regression
Table 3:

	$\begin{array}{c} (1) \\ 0.117^{***} & (0.010) \\ 0.006^{***} & (0.0004) \\ -0.414^{***} & (0.008) \\ -0.509^{***} & (0.008) \\ -0.509^{***} & (0.0131) \\ 0.108^{***} & (0.004) \\ 0.106^{***} & (0.014) \\ 0.009^{***} & (0.017) \\ 0.008^{***} & (0.017) \\ 0.008^{***} & (0.017) \\ 0.008^{***} & (0.017) \\ 0.008^{***} & (0.017) \\ 0.008^{***} & (0.017) \\ 0.008^{***} & (0.017) \\ 0.008^{***} & (0.017) \\ 0.008^{***} & (0.017) \\ 0.008^{***} & (0.017) \\ 0.008^{***} & (0.017) \\ 0.008^{***} & (0.018) \\ 0.008^{***} $	(2) 0 100*** (0 010)	price (3)	ce (4)	Ĺ	
es () () () () () () () () () ($\begin{array}{c} (1) \\ & & & & \\ & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\$	(2) 0.100*** (0.010)	(3)	(4)	(1)	
es tes tes tes tes tes tes tes t	*** (0.010) ** (0.0004) *** (0.008) *** (0.008) *** (0.004) *** (0.004) *** (0.009) *** (0.017) *** (0.017)	0 100*** (0 010)	(~)	(-)	(c)	(9)
es) tes to Center (km)	$\sum_{i=1}^{n} (0.004)$ $\sum_{i=1}^{n} (0.008)$ $\sum_{i=1}^{n} (0.004)$ $\sum_{i=1}^{n} (0.004)$ $\sum_{i=1}^{n} (0.004)$ $\sum_{i=1}^{n} (0.014)$ $\sum_{i=1}^{n} (0.017)$		0.116^{***} (0.010)	$0.129^{***} (0.012)$	0.120^{***} (0.012)	0.115^{***} (0.010)
se to Center (km)	$\begin{array}{c} \begin{array}{c} \begin{array}{c} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} $	0.006^{***} (0.0004)	$0.005^{***}(0.0004)$	$0.005^{***}(0.0005)$	$0.005^{***}(0.0005)$	0.005^{***} (0.0004)
e to Center (km)		-0.392^{***} (0.008)	-0.391^{***} (0.008)	-0.386^{***} (0.010)	-0.371^{***} (0.010)	-0.391^{***} (0.008)
s as (D) (D) (D) (D) (C) (C) (C) (C) (C) (C) (C) (C) (C) (C		-0.501^{***} (0.030)	-0.498^{***} (0.030)	-0.471^{***} (0.036)	-0.434^{***} (0.035)	-0.497^{***} (0.030)
) (D) (D) (D) (C) (C) (C) (C) (C) (C) (C) (C) (C) (C		$0.100^{***} (0.004)$	$0.101^{***} (0.004)$	$0.099^{***} (0.004)$	$0.094^{***} (0.004)$	$0.101^{***} (0.004)$
) (D) (D) (D) (C) (C) (C) (C) (C) (C) (C) (C) (C) (C		$0.162^{***} (0.009)$	$0.160^{***} (0.009)$	$0.168^{***} (0.010)$	$0.170^{***} \ (0.010)$	$0.159^{***} (0.009)$
) (D) 5 Distance to Center (km) -	~ `	0.162^{***} (0.014)	$0.166^{***} (0.013)$	$0.164^{***} (0.015)$	0.152^{***} (0.015)	0.166^{***} (0.013)
(D) Distance to Center (km) –		$0.081^{***}(0.017)$	$0.079^{***}(0.017)$	$0.090^{***}(0.019)$	0.069^{***} (0.020)	0.082^{***} (0.017)
I	(0.013)	0.033^{**} (0.013)	0.037^{***} (0.013)	0.034^{**} (0.014)	0.025^{*} (0.014)	0.036^{***} (0.013)
	-0.043^{***} (0.001)	-0.043^{***} (0.001)	-0.042^{***} (0.001)	-0.042^{***} (0.002)	-0.047 ^{***} (0.003)	10007 000000000
Walking Distance to Center (km)						-0.039^{***} (0.001)
TV (D)		0.084^{***} (0.007)	$0.085^{***} (0.007)$	$0.081^{***} (0.008)$	0.083^{***} (0.008)	0.084^{***} (0.007)
Wifi (D)		$0.027^{*} (0.016)$	$0.028^{*} (0.016)$	$0.025\ (0.018)$	$0.031^{*} (0.018)$	0.027^{*} (0.016)
Kitchen (D)		$-0.024\ (0.015)$	$-0.024\ (0.015)$	-0.018(0.017)	-0.021(0.017)	$-0.024\ (0.015)$
Breakfast (D)		$0.029^{**} (0.014)$	$0.031^{**} (0.014)$	$0.039^{**} (0.016)$	$0.034^{**} (0.016)$	$0.031^{**} \ (0.014)$
Smoking Allowed (D)		-0.033^{***} (0.010)	-0.030^{***} (0.010)	-0.019^{*} (0.011)	-0.012(0.011)	-0.030^{***} (0.010)
Children Allowed (D)		$0.030^{***} (0.009)$	0.024^{***} (0.008)	$0.033^{***} (0.010)$	$0.029^{***} (0.010)$	0.024^{***} (0.008)
Pets Allowed (D)		-0.026^{**} (0.010)	-0.025^{**} (0.010)	-0.031^{***} (0.011)	-0.039^{***} (0.011)	-0.025^{**} (0.010)
Minimum Stay (days)		-0.011^{***} (0.002)	-0.012^{***} (0.002)	-0.014^{***} (0.002)	-0.014^{***} (0.002)	-0.012^{***} (0.002)
Security Deposit (D)		$0.060^{***} (0.007)$		0.052^{***} (0.008)	0.060^{***} (0.008)	
Membership Duration (months)			$0.002^{***} (0.0002)$	$0.002^{***} (0.0002)$	$0.002^{***} (0.0002)$	$0.002^{***} (0.0002)$
Superhost (D)			-0.017^{*} (0.010)	-0.024^{**} (0.011)	-0.029^{***} (0.011)	-0.016^{*} (0.010)
Verified Host (D)			-0.023^{***} (0.007)	-0.023^{***} (0.008)	-0.013^{*} (0.008)	-0.023^{***} (0.007)
Number of Friends			-0.003(0.005)	-0.002(0.005)	-0.002(0.005)	-0.003 (0.005)
Host Happiness (API)				-0.017^{**} (0.009)	-0.013 (0.009)	
Male Host Gender (API, D)				$0.001 \ (0.008)$	0.001 (0.008)	
Host Name (API, D)				$0.004 \ (0.008)$	$0.002 \ (0.008)$	
Constant 3.074**	$3.074^{***} (0.051)$	$3.075^{***} (0.054)$	$2.999^{***} (0.055)$	$2.944^{***} (0.065)$	3.043^{***} (0.078)	$3.009^{***} (0.055)$
City Dumnies	Yes	Yes	Yes	Yes	No	Yes
Neighborhood Dummies	No	No	No	No	Yes	No
	16,782	16,782	16,782	13,348	13,348	16,781
	-74,793.690	-74,595.150	-74,522.660	-59,353.910	-59,135.220	-74,504.160
Akaike Inf. Crit. 149,6	149,621.400	149,242.300	149,105.300	118,773.800	118,682.400	149,068.300

Note: *p<0.1; **p<0.05; ***p<0.01

Nevertheless, within the limitations of our name-based approached, we cannot find any hints to ethnic discrimination on Airbnb in our Germany-focused dataset. This holds true when we control for subgroups of hosts with non-German names, e.g. hosts with Arabic names.

Contrary to our expectations, the parameter estimate for the host happiness variable, which assesses if a host smiles in his profile picture, is negative. Taking into account that experimental evidence concludes that guests attribute a positive monetary value to the trustworthiness they sense from a host's profile picture (Ert, Fleischer, and Magen, 2016), showing a smile does not seem to be sufficient to establish a perception of trust. However, the coefficient is rendered insignificant in other specifications (see Section 5.2).

5.2 Robustness Checks

Although our parameter estimates are remarkably stable for model specifications (1) to (4), we conduct further robustness checks to substantiate our results. As city dummies cannot account for heterogeneity in location value within a particular city, we replace city dummies with neighborhood dummies, which we obtained by reverse geocoding the coordinates of every accommodation to the respective city district using Google's Geolocation API. As the lower AIC score of specification (5) in comparison to model (4) show, this increases model quality in comparison. However, accounting for heterogeneity on the city district level does not change any sign of our parameter estimates.

Since all our estimations identify distance to the city center as an important driver for accommodation prices, we seek to find a more precise measure for our distance variable. Hence we replace the simple geodesic distance between housing and city center with walking distance obtained from Google Maps Distance Matrix API, as this distance measure accounts for obstacles like rivers and motorways. As anticipated, when compared to models (1) to (3), the more exact measure increases model quality in terms of AIC score and yields a lower parameter estimate for the distance variable.²⁶

Figure 5 shows residuals vs. fitted plots for all six specifications in Table 3 with residuals on the y-axis and fitted values on the x-axis. Ideally, residuals should be scattered randomly around the zero line. Besides a few outliers that are highlighted in the respective figures, residuals in all specifications are well-behaved and don't show a pattern that suggests a non-linear

²⁶The decrease of the coefficient seems plausible, when keeping in mind that the walking distance obviously grows overproportionally with geodesic distance. Subsequently, every additional kilometer measured in terms of geodesic distance impacts the accommodation value more negatively than an additional kilometer of walking distance.

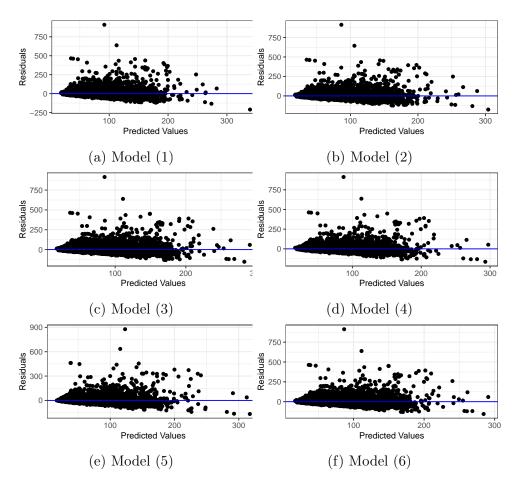


Figure 5: Residuals vs. Fitted for Models (1) - (6)

relationship we don't account for in our model. However, further inspection of the relationship between price and distance to city center suggests that the effect of the distance variable might change direction when distance to the city center reaches a certain threshold. Figure 6 draws a loess smoothed curve in green and the linear regression line in orange.²⁷ As long as the loess smoother roughly approximates the linear regression line, the assumption of linearity is verified. As we can see in Figure 6, this assumption does not hold for higher distance values in our dataset as the loess smoother's slope changes from negative to positive somewhere between 10 and 15 km.

In order to capture this non-linear relationship in our model, we add the squared value of the distance variable to our hedonic regression model

 $^{^{27}\}mathrm{The}$ shading around the loess smoother and the linear regression line corresponds to standard errors.

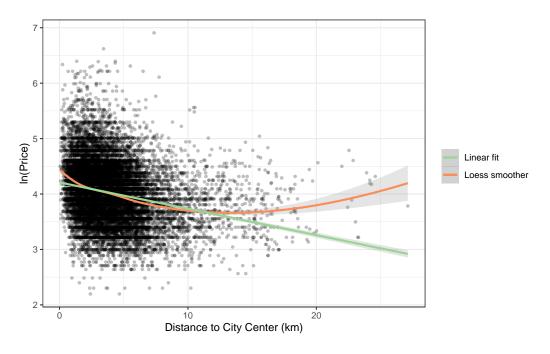


Figure 6: Distance from City Center and Price

in Equation 3 in order to obtain β_1 and β_2 from regression results in Table 4. Subsequently, we then calculate the marginal effect of distance (x) on price (y):

$$\frac{\Delta y_t}{\Delta x_t} = \beta_1 + 2\beta_2 x_t \tag{5}$$

Since the coefficient estimate for our distance variable $(\hat{\beta}_1)$ is negative, prices fall with increasing distance to the city center at first. However, as the coefficient estimate for its quadratic transformation $(\hat{\beta}_2)$ is positive, the aforementioned effect changes direction when a certain threshold is reached. We calculate the extreme point for $\beta_1 < 0$ and $\beta_2 > 0$ by setting Equation 5 equal to zero:

$$x^* = -\hat{\beta}_1/(2\hat{\beta}_2)$$

$$x^* = 0.08123338/(2*0.00282588)$$

$$x^* = 14.3731$$
(6)

	1 Person, 1 Night
	price
Overall Rating	0.118^{***} (0.010)
Number of Pictures	0.005^{***} (0.0004)
Private Room (D)	-0.388^{***} (0.008)
Shared Room (D)	-0.491^{***} (0.030)
Capacity	0.101^{***} (0.004)
Bedrooms	0.160^{***} (0.009)
Bathrooms	0.165^{***} (0.013)
House (D)	0.070^{***} (0.017)
Realbed (D)	0.033^{***} (0.013)
Distance to CC (km)	-0.081^{***} (0.003)
Squared Distance to CC (km)	0.003^{***} (0.0002)
TV (D)	$0.081^{***}(0.007)$
Wifi (D)	0.030^{*} (0.016)
Kitchen (D)	-0.028^{*} (0.015)
Breakfast (D)	0.032^{**} (0.014)
Smoking Allowed (D)	-0.028^{***} (0.010)
Children Allowed (D)	0.025^{***} (0.008)
Pets Allowed (D)	-0.024^{**} (0.010)
Minimum Stay (days)	-0.012^{***} (0.002)
Security Deposit (D)	0.056^{***} (0.007)
Membership Duration (months)	0.002^{***} (0.0002)
Superhost (D)	-0.018^{*} (0.010)
Verified Host (D)	-0.020^{***} (0.007)
Number of Friends	$-0.003\ (0.005)$
Constant	3.097^{***} (0.056)
City Dummies	Yes
Observations	16,782
Log Likelihood	-74,407.870
Akaike Inf. Crit.	148,877.700
Note:	*p<0.1; **p<0.05; ***p<0.01

 Table 4: Non-Linear Regression Results

_

_

This result implies that after a threshold of about 14km distance to the city center is reached, guests value locations higher the more remote they are. This can be explained by the fact that guests who book accommodations in a city's outskirt likely have different preferences compared to those who book apartments close to the center. Listings that are located in a city's "bacon belt" allow easy access to nature attractions and therefore offer higher recreational value than accommodations closer to the center. Guests that book accommodations more than 14km away from the center tend to be more interested in these surrounding excursion destinations than in urban tourist attractions. Inclusion of the quadratic term however does not substantially change the remaining parameter estimates in Table 4.

As discussed in Section 4, we do not include the number of reviews in our model specifications in Table 3 as basic supply and demand theory suggests that lower prices will cause more bookings and consequently more reviews. The number of reviews then is a proxy for demand and not exogenous for a listing's price. For the sake of illustration and as further robustness check, we introduce the number of reviews as an additional regressor in our model, obtaining the expected negative coefficient (-0.001) significant at the 1 percent level (see Appendix A5). Other parameter estimates keep mostly in line with our prior results and signs of coefficients do not change direction. However, the coefficient for kitchen access gains significance in contrast to model (1)-(6) specifications in Table 3.

5.3 HPI for Seven German Cities

With the results from the hedonic regression as conducted in Section 5, we can disentangle the contributory value of each characteristic to a listing's price. Building on that approach and by including the time dimension (see Equation 4), we can use our results to also construct hedonic price indices (HPIs) for each city in our dataset. These hedonic indices are useful measures, as the apartments and houses on Airbnb are obviously very heterogenous in their characteristics and changes in rental prices can reflect pure price changes as well as quality changes. Hence, an exemplary increase in a city's overall price level on Airbnb can be purely driven by an influx of high-quality listings. If this is the case, the level of a hedonic price index would remain constant, although mean (and median) prices go up.

In order to construct hedonic price indices for each city, we repeatedly scraped Airbnb's website for listings at different points in time and estimated Equation 4 for each scraping's results, using the same specification as in model (3) in Table 3. Specifically, our scraping process covered a time span of one year, gathering listings from Airbnb in April, July, September 2017 and

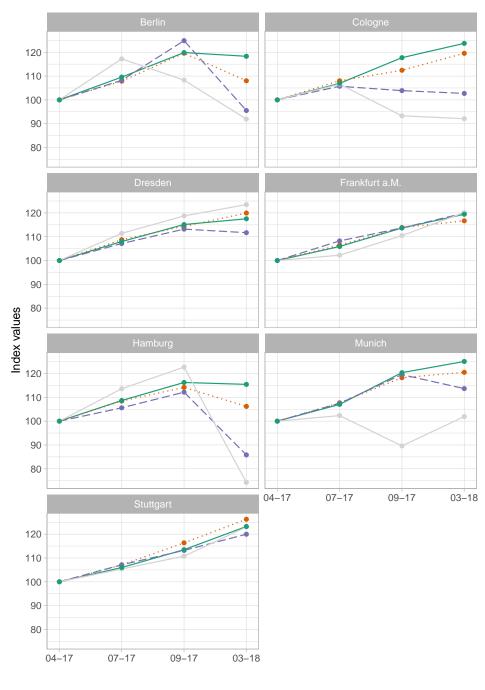


Figure 7: Hedonic and Other Price Indices (s.a.)

-- Hedonic · • · Mean -- Median -- Observations

March 2018. As we need to take seasonal price fluctuations into account, we conduct a seasonal adjustment with seasonality index values obtained from the German CPI for "Accommodation services of hotels, inns or the like".²⁸

Figure 7 shows the evolution of the hedonic price index in seven German cities over time in contrast to mean and median price levels. All cities exhibit a remarkable upward trend in their price indices, resulting in about 20% higher mean prices in most cities within a year. While the figures for Dresden, Frankfurt, Munich and Stuttgart only show minor deviations between hedonic price levels and mean as well as median prices, other cities like Berlin, Hamburg and Cologne exhibit a different pattern.

In Hamburg, we can observe a steep decline in the mean and median price level between September 2017 and March 2018, while the hedonic price index remained relatively constant. Apparently, in this time span an overproportional amount of high-quality accommodations left the market, leading to the observed drop in the average and median price in Hamburg. This finding is supported by a decline in total Airbnb listings in Hamburg by about 45 index points between September 2017 and March 2018. An apparent explanation for this finding is the legal decree that came into effect on the 20th of March 2018, in which the sufficient provision of housing space was declared as endangered in the whole city of Hamburg until 2028 according to the city's law to protect and preserve long-term housing.²⁹ The legal process was accompanied by a public debate and reports in the local press, attracting attention to the legal restrictions and menacing fines threatening landlords conducting illegal short-term letting. As the legal restrictions predominantly apply to professional landlords, who rent out whole apartments with high quality characteristics, this is a possible explanation for the decline in the median price level, while the quality-adjusted HPI value remained constant.

Similar patterns can be observed for Berlin, Cologne and – for a lesser extent – Munich. Although median price levels declined, the quality-adjusted HPI increased or remained constant. As it was the case in Hamburg, the turning public mood and the pushback of municipalities against vacation homes in crowded urban rental markets and subsequent sensitization of landlords might be a factor in this development.³⁰

In summary, the development of hedonic price levels on Airbnb mirrors

²⁸In detail, we decompose the monthly time series of SEA-CPI 11201 between 2015 and 2019, using a multiplicative model of the form $Y_t = T_t S_t e_t$, where T is the trend and S the seasonality component. Subsequently, a seasonality index with April=1 is constructed and multiplied with the values from our Airbnb price indices.

 $^{^{29}\}mathrm{See}$ § 9 Abs. 1 HmbWoSchG.

³⁰See Reichel (2017) for Munich's threat to imprison illegal Airbnb landlords and Schönball (2017) for Berlin's move to tighten regulation for short-term rentals.

the upward trend in urban long-term housing markets. Within a year, the quality-adjusted price level in all investigated cities climbed by 15-25%. However, our results hint at the effectiveness of municipal regulation countering misappropriation of housing space.

6 Conclusion

Housing markets in general are characterized by a high degree of heterogeneity and price formation depends on a vast set of characteristics, which are not always easily observed. The same holds true for short-term vacation rentals as they can be found on Airbnb's platform. The hedonic regression approach allows to identify each characteristic's contribution to a property's overall value. In respect to Airbnb, hedonic regression results are of particular interest, as they give insights into the relevance of quality signals that are build into the platform's trust-buildings mechanisms. These are designed to reduce customer's uncertainty regarding the vacation rental and thereby reduce information asymmetry between host and guest. Our hedonic regression model therefore includes a set of variables that is not formed by hard facts about the listing, but includes information on other guests' experiences as well as personal traits of hosts.

Regarding characteristics of the vacation rental, our results are predominantly in line with expectations: Whole apartments achieve higher prices than private and shared rooms, capacity and size improves a rental's valuation. Listings in close proximity to the city center and amenities like a TV set and provided breakfast can obtain price premiums as well. However, results for variables that reflect trust-worthiness are mixed. As expected, good reviews allow hosts to set higher prices on the platform, but superhost and verified status have no significant positive effect. In contrast to findings for Airbnb in the US, we cannot confirm implicit racial discrimination, as we cannot find price differences linked to the ethnicity of hosts. Neither name origin or gender of hosts has a significant effect on prices. In conclusion, our results suggest that Airbnb's review system and the host's membership duration are the most important factors in building trust, which in turn translates into hosts' ability to set higher prices for their rental.

In respect to the evolution of (quality-adjusted) prices, we observe an overall increase in hedonic price index levels in the investigation period for all seven cities, ranging between 15 and 25%. Market exit of high-quality listings in cities with extensive regulatory pressure on short-term vacation rentals suggests effectiveness of these measures. However, this impression needs to be properly evaluated in causal inference studies.

References

- Airbnb (2021). Fast Facts. Airbnb Newsroom. URL: https://press.atairbnb. com/fast-facts/.
- Bar-Isaac, Heski and Steven Tadelis (2008). "Seller Reputation". In: Foundations and Trends in Microeconomics 4.4, pp. 273–351.
- Belk, Russell (2010). "Sharing". In: Journal of Consumer Research 36.5, pp. 715–734.
- Bolton, Gary, Ben Greiner, and Axel Ockenfels (2012). "Engineering Trust: Reciprocity in the Production of Reputation Information". In: *Management Science* 59.2, pp. 265–285.
- Botsman, Rachel and Roo Rogers (2010). What's Mine Is Yours: The Rise of Collaborative Consumption. 1st ed. New York: Harper Business. 304 pp.
- Bull, Adrian O. (1994). "Pricing a Motel's Location". In: International Journal of Contemporary Hospitality Management 6.6, pp. 10–15.
- Cabral, Luís and Ali Hortaçsu (2010). "The Dynamics of Seller Reputation: Evidence from Ebay". In: *The Journal of Industrial Economics* 58.1, pp. 54–78.
- Chen, Yong and Karen Xie (2017). "Consumer Valuation of Airbnb Listings: A Hedonic Pricing Approach". In: International Journal of Contemporary Hospitality Management 29.9.
- Cheng, Mingming (2016). "Sharing economy: A Review and Agenda for Future Research". In: International Journal of Hospitality Management 57, pp. 60–70.
- Court, A. T. (1939). *Hedonic Price Indexes with Automotive Examples*. The Dynamics of Automobile Demand. New York: General Motors.
- Dale-Johnson, David (1982). "An Alternative Approach to Housing Market Segmentation Using Hedonic Price Data". In: *Journal of Urban Economics* 11.3, pp. 311–332.
- Deboosere, Robbin et al. (2019). "Location, Location and Professionalization: A Multilevel Hedonic Analysis of Airbnb Listing Prices and Revenue". In: *Regional Studies, Regional Science* 6.1, pp. 143–156.
- Dewally, Michaël and Louis Ederington (2006). "Reputation, Certification, Warranties, and Information as Remedies for Seller-Buyer Information Asymmetries: Lessons from the Online Comic Book Market". In: The Journal of Business 79.2. Publisher: The University of Chicago Press, pp. 693-729. URL: https://www.jstor.org/stable/10.1086/499169.
- Dormann, Carsten F. et al. (2013). "Collinearity: A Review of Methods to Deal with It and a Simulation Study Evaluating Their Performance". In: *Ecography* 36.1, pp. 27–46.

- Edelman, Benjamin, Michael Luca, and Dan Svirsky (2017). "Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment". In: *American Economic Journal: Applied Economics* 9.2, pp. 1–22.
- Edelman, Benjamin G. and Michael Luca (2014). Digital Discrimination: The Case of Airbnb.com. SSRN Scholarly Paper ID 2377353. Rochester, NY: Social Science Research Network. URL: https://papers.ssrn.com/ abstract=2377353.
- Einav, Liran, Chiara Farronato, and Jonathan Levin (2015). *Peer-to-Peer Markets*. Working Paper 21496. National Bureau of Economic Research. URL: http://www.nber.org/papers/w21496.
- Ert, Eyal and Aliza Fleischer (2019). "The Evolution of Trust in Airbnb: A Case of Home Rental". In: Annals of Tourism Research 75, pp. 279–287.
- Ert, Eyal, Aliza Fleischer, and Nathan Magen (2016). "Trust and reputation in the sharing economy: The role of personal photos in Airbnb". In: *Tourism Management* 55, pp. 62-73. URL: http://www.sciencedirect. com/science/article/pii/S0261517716300127.
- Fradkin, Andrey, Elena Grewal, and Dave Holtz (2021). "Reciprocity and Unveiling in Two-sided Reputation Systems: Evidence from an Experiment on Airbnb". In: *Marketing Science*, Advance Online Publication. URL: https://andreyfradkin.com/assets/reviews_paper.pdf.
- Frölich, Markus (2008). "Parametric and Nonparametric Regression in the Presence of Endogenous Control Variables". In: *International Statistical Review* 76.2, pp. 214–227.
- Gibbs, Chris et al. (2018). "Pricing in the Sharing Economy: A Hedonic Pricing Model Applied to Airbnb Listings". In: Journal of Travel & Tourism Marketing 35.1, pp. 46–56.
- Gilheany, John, David Wang, and Stephen Xi (2015). The Model Minority? Not on Airbnb.com: A Hedonic Pricing Model to Quantify Racial Bias against Asian Americans. Technology Science. URL: https://techscience. org/a/2015090104/.
- Griliches, Zvi (1961). "Hedonic Price Indexes for Automobiles: An Econometric of Quality Change". In: The Price Statistics of the Federal Government. National Bureau of Economic Research, pp. 173–196.
- (2014). Price Indexes and Quality Change: Studies in New Methods of Measurement. Cambridge: Harvard University Press.
- Hawlitschek, Florian, Timm Teubner, and Christof Weinhardt (2016). "Trust in the Sharing Economy". In: *Die Unternehmung* 70.1, pp. 26–44.
- Hill, Dan (2015). "How Much is Your Spare Room Worth?" In: *IEEE Spectrum* 52.9, pp. 32–58.

- Houser, Daniel and John Wooders (2006). "Reputation in Auctions: Theory, and Evidence from eBay". In: Journal of Economics & Management Strategy 15.2, pp. 353–369.
- Hung, Wei-Ting, Jui-Kou Shang, and Fei-Ching Wang (2010). "Pricing Determinants in the Hotel Industry: Quantile Regression Analysis". In: International Journal of Hospitality Management 29.3, pp. 378–384.
- Jefferson-Jones, Jamila (2015). "Airbnb and the Housing Segment of the Modern 'Sharing Economy': Are Short-Term Rental Restrictions an Unconstitutional Taking?" In: Hastings Constitutional Law Quarterly 42.
- Kakar, Venoo et al. (2018). "The Visible Host: Does race guide Airbnb rental rates in San Francisco?" In: *Journal of Housing Economics*. Race and the City 40, pp. 25-40. URL: https://www.sciencedirect.com/science/ article/pii/S1051137716301309.
- Kang, Han-Bin and Alan K. Reichert (1991). "An Empirical Analysis of Hedonic Regression and Grid-Adjustment Techniques in Real Estate Appraisal". In: *Real Estate Economics* 19.1, pp. 70–91.
- Klein, Tobias J., Christian Lambertz, and Konrad O. Stahl (2016). "Market Transparency, Adverse Selection, and Moral Hazard". In: Journal of Political Economy 124.6, pp. 1677–1713. URL: https://www.journals. uchicago.edu/doi/full/10.1086/688875.
- Lancaster, Kelvin J. (1966). "A New Approach to Consumer Theory". In: Journal of Political Economy 74.2, pp. 132–157.
- Lee, Seul Ki and SooCheong (Shawn) Jang (2012). "Premium or Discount in Hotel Room Rates? The Dual Effects of a Central Downtown Location". In: Cornell Hospitality Quarterly 53.2, pp. 165–173.
- Lorde, Troy, Jadon Jacob, and Quinn Weekes (2019). "Price-setting Behavior in a Tourism Sharing Economy Accommodation Market: A Hedonic Price Analysis of AirBnB Hosts in the Caribbean". In: *Tourism Management Perspectives* 30, pp. 251–261.
- Malpezzi, Stephen (2002). "Hedonic Pricing Models: A Selective and Applied Review". In: *Housing Economics and Public Policy*. Ed. by Tony O'Sullivan and Kenneth Gibb. Blackwell Science Ltd, pp. 67–89.
- McCullagh, P. and John A. Nelder (1989). *Generalized Linear Models*. 2nd ed. Chapman and Hall/CRC. 532 pp.
- McDonald, Cynthia G. and V. Carlos Slawson (2002). "Reputation in an Internet Auction Market". In: *Economic Inquiry* 40.4, pp. 633–650.
- Melnik, Mikhail I. and James Alm (2002). "Does a Seller's Ecommerce Reputation Matter? Evidence from eBay Auctions". In: The Journal of Industrial Economics 50.3, pp. 337–349.

- Monty, Ben and Mark Skidmore (2003). "Hedonic Pricing and Willingness to Pay for Bed and Breakfast Amenities in Southeast Wisconsin". In: *Journal of Travel Research* 42.2, pp. 195–199.
- Moreno-Izquierdo, Luis et al. (2019). "Tourist Environment and Online Reputation as a Generator of Added Value in the Sharing Economy: The Case of Airbnb in Urban and Sun-and-Beach Holiday Destinations". In: Journal of Destination Marketing & Management 11, pp. 53–66.
- Nosko, Chris and Steven Tadelis (2015). The Limits of Reputation in Platform Markets: An Empirical Analysis and Field Experiment. Working Paper 20830. National Bureau of Economic Research.
- Panda, Rasananda, Surbhi Verma, and Bijal Mehta (2015). "Emergence and Acceptance of Sharing Economy in India: Understanding through the Case of Airbnb". In: International Journal of Online Marketing (IJOM) 5.3, pp. 1–17.
- Reichel, Reiner (2017). "Airbnb, Wimdu & Co.: München nimmt Vermieter in Zwangshaft". In: Handelsblatt Online. URL: https://www.handelsblatt. com/finanzen/immobilien/airbnb-wimdu-und-co-muenchen-nimmtvermieter-in-zwangshaft/20272076.html.
- Resnick, Paul and Richard Zeckhauser (2002). "Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay' s Reputation System". In: *The Economics of the Internet and E-commerce*. Ed. by Michael R. Baye. Vol. 11. Advances in Applied Microeconomics. Emerald Group Publishing Limited, pp. 127–157.
- Rosen, Sherwin (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition". In: Journal of Political Economy 82.1, pp. 34–55.
- Schönball, Ralf (2017). "Berliner Zweckentfremdungsgesetz wird verschärft". In: Tagesspiegel Online. URL: https://www.tagesspiegel.de/berlin/ airbnb-in-berlin-berliner-zweckentfremdungsgesetz-wird-verschaerft/ 20472786.html.
- Strand, Jon and Mette Vågnes (2001). "The Relationship Between Property Values and Railroad Proximity: A Study Based on Hedonic Prices and Real Estate Brokers' Appraisals". In: *Transportation* 28.2, pp. 137–156.
- Straszheim, Mahlon (1974). "Hedonic Estimation of Housing Market Prices: A Further Comment". In: The Review of Economics and Statistics 56.3, pp. 404–406.
- Teubner, Timm, Florian Hawlitschek, and David Dann (2017). "Price Determinants on Airbnb: How Reputation Pays Off in the Sharing Economy". In: Journal of Self-Governance and Management Economics 5.4.

- Thrane, Christer (2007). "Examining the Determinants of Room Rates for Hotels in Capital Cities: The Oslo Experience". In: Journal of Revenue and Pricing Management 5.4, pp. 315–323.
- Triplett, Jack (2006). Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes. Paris: OECD Publishing.
- Wang, Dan and Juan L. Nicolau (2017). "Price Determinants of Sharing Economy Based Accommodation Rental: A Study of Listings from 33 Cities on Airbnb.com". In: International Journal of Hospitality Management 62, pp. 120–131.
- Wilhelmsson, Mats (2002). "Spatial Models in Real Estate Economics". In: Housing, Theory and Society 19.2, pp. 92–101.
- Ye, Qiang, Rob Law, and Bin Gu (2009). "The Impact of Online User Reviews on Hotel Room Sales". In: International Journal of Hospitality Management 28.1, pp. 180–182.
- Yee, Hector and Bar Ifrach (2015). Aerosolve: Machine learning for humans. Medium. URL: https://medium.com/airbnb-engineering/aerosolvemachine-learning-for-humans-55efcf602665.
- Zervas, Georgios, Davide Proserpio, and John W. Byers (2017). "The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry". In: *Journal of Marketing Research* 54.5, pp. 687–705.
- (2021). "A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average". In: *Marketing Letters* 32.1, pp. 1–16.

Appendix

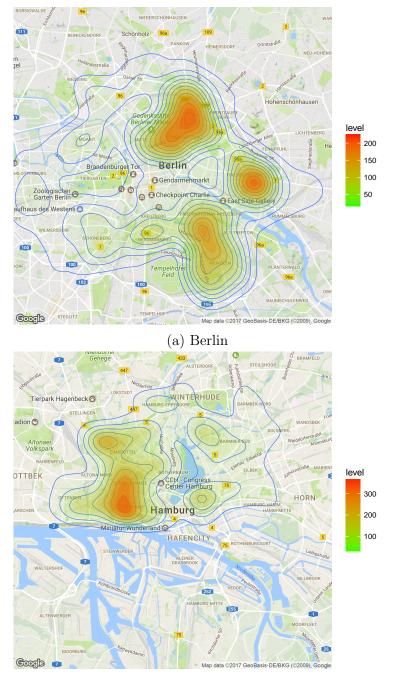


Figure A1: Spatial Distribution of Airbnb Listings

(b) Hamburg

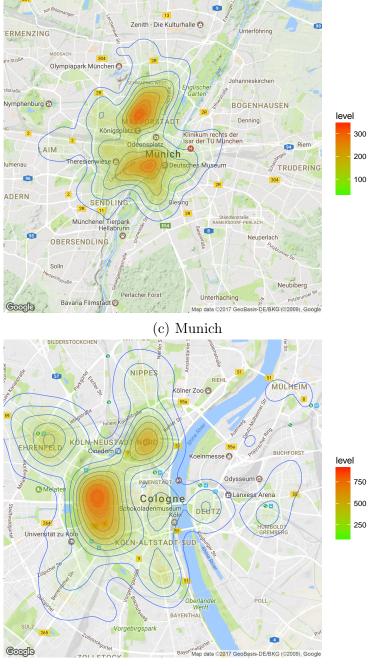


Figure A1: Spatial Distribution of Airbnb Listings (cont.)

(d) Cologne

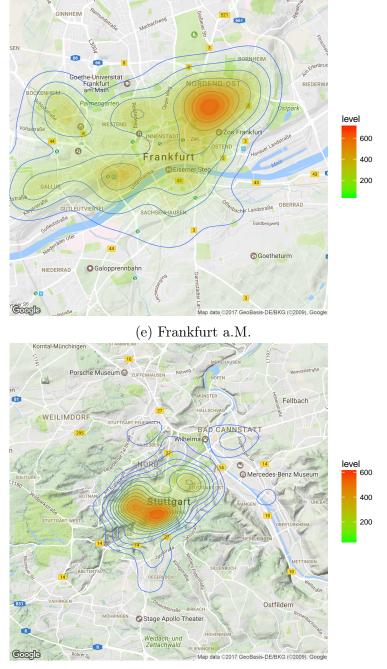


Figure A1: Spatial Distribution of Airbnb Listings (cont.)

(f) Stuttgart

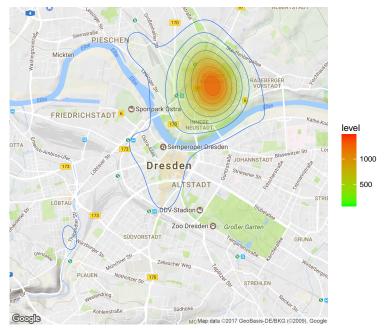


Figure A1: Spatial Distribution of Airbnb Listings (cont.)

(g) Dresden

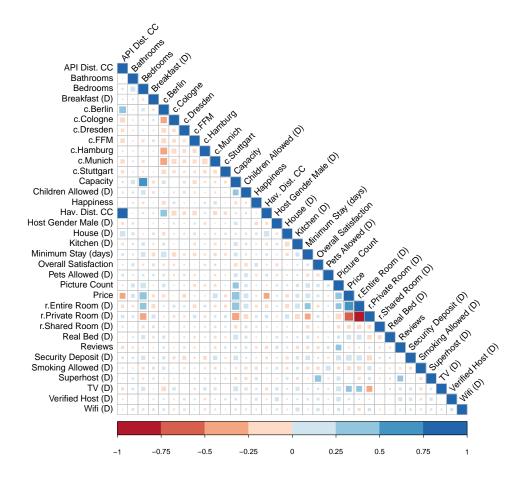


Figure A2: Correlation Matrix

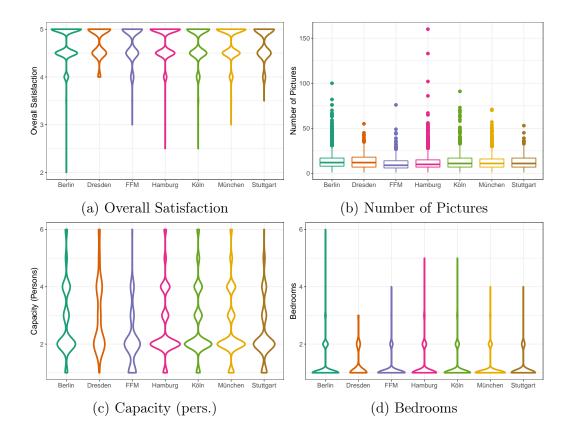


Figure A3: Distribution of Chosen Numerical Variables by City



Figure A4: Distribution of Chosen Categorical Variables by City

	1 Person, 1 Night
	price
Overall Rating	0.106^{***} (0.010)
Number of Reviews	-0.001^{***} (0.0001)
Number of Pictures	0.007^{***} (0.0004)
Private Room (D)	-0.388^{***} (0.008)
Shared Room (D)	-0.503^{***} (0.030)
Capacity	$0.101^{***} (0.004)$
Bedrooms	0.158^{***} (0.009)
Bathrooms	0.156^{***} (0.013)
House (D)	0.082^{***} (0.017)
Realbed (D)	0.031^{**} (0.013)
Distance to CC (km)	-0.045^{***} (0.001)
TV (D)	0.087^{***} (0.007)
Wifi (D)	0.030^{*} (0.016)
Kitchen (D)	-0.036^{**} (0.015)
Breakfast (D)	0.026^{*} (0.014)
Smoking Allowed (D)	-0.037^{***} (0.010)
Children Allowed (D)	0.032^{***} (0.008)
Pets Allowed (D)	-0.022^{**} (0.010)
Minimum Stay (days)	-0.012^{***} (0.002)
Security Deposit (D)	0.058^{***} (0.007)
Constant	3.129^{***} (0.054)
City Dummies	Yes
Observations	16,782
Log Likelihood	-74,515.460
Akaike Inf. Crit.	149,084.900
Note:	*p<0.1; **p<0.05; ***p<0.01

Table A5: Regression Results with Number of Reviews

2021	
188	Tran, Thi Xuyen: Typhoon and Agricultural Production Portfolio. Empirical Evidence for a Developing Economy, May 2021
187	Richau, Lukas; Follert, Florian; Frenger, Monika; Emrich, Eike: The Rainmaker?! The impact of investors on transfer fees in the English Premier League, January 2021
2020	
186	Beckmann, Klaus: Konzept für eine Militärökonomik, November 2020
185	Beckmann, Klaus: Endogenisierung der Politikreaktion im SIR-Modell einer Epidemie, November 2020
184	Bernhardt, Lea: Common factors of withdrawn and prohibited mergers in the European Union, October 2020
183	Bernhardt, Lea; Dewenter, Ralf; Thomas, Tobias: Watchdog or Loyal Servant? Political Media Bias in US Newscasts, August 2020
2019	
182	Ross, Harm Hauke: Second-hand price volatility of green ships: an empirical analysis across main shipping segments, November 2019
2018	
181	Wenzel, Daniela: Droughts and Corruption, September 2018
180	Linder, Melissa; Muijs, Matthias: A new price test in geographic market definition – an application to german retail gasoline market, August 2018
179	Dewenter, Ralf; Linder, Melissa; Thomas, Tobias: Can Media Drive the Electorate? The Impact of Media Coverage on Party Affiliation and Voting Intentions, April 2018
2017	
178	Beckmann, Klaus: Bounded rationality in differential games, December 2017
177	Herzer, Dierk; Nagel, Korbinian: The effects of adult and non-adult mortality on long-run econo- mic development: Evidence from a heterogeneous dynamic and cross-sectionally dependent panel of countries between 1800 and 2010, July 2017
176	Dewenter, Ralf; Heimeshoff, Ulrich; Löw, Franziska: Market Definition of Platform Markets, March 2017
2016	
175	Dewenter, Ralf; Dulleck, Uwe; Thomas, Tobias: Does the 4th estate deliver? Towards more direct measure of political media bias, November 2016
174	Luik, Marc-André: Child Health, Human Capital and Adult Financial Behavior, November 2016
173	Michael Berlemann; Marc-André Luik: Institutional Reform and Depositors' Portfolio Choice - Evidence from Bank Account Data, November 2016
172	Lauenstein, Philipp; Küster Simic, André: Information Processing in Freight and Freight Forward Markets: An Event Study on OPEC Announcements, September 2016
171	Nagel, Korbinian: A Life Course Perspective on the Income-to-Health Relationship: Macro- Empirical Evidence from Two Centuries, July 2016
170	Dluhosch, Barbara; Horgos, Daniel: International Competition Intensified - Job Satisfaction Sa- crified?, June 2016
169	Beckmann, Klaus; Dewenter, Ralf; Thomas, Tobias: Can news draw blood? The impact of media coverage on the number and severity of terror attacks, May 2016
168	Afflatet, Nicolas: Deficit Policy within the Framework of the Stability and Growth Pact - Empirical Results and Lessons for the Fiscal Compact, April 2016
167	Werner, Max: Evaluating Prediction Markets for Internal Control Applications, May 2016
166	Werner, Max; Eißing, Klaus; Langton, Sebastian: Shared Value Potential of Transporting Cargo via Hyperloop, May 2016
165	Werner, Max; Vianelli, A.; Bodek, Mariusz C.: Monitoring Venture Capital Investments through Internal Control Prediction Markets, May 2016
164	Jahn, Vera; Steinhardt, Max Friedrich: Innovation and Immigration - Insights from a Placement Policy, February 2016

