Kindness capital and rating bias. A sentiment analysis on Airbnb reviews *

Laura Abrardi[†], Elisabetta Raguseo[‡], Laura Rondi [§]

September 8, 2022

Abstract

We study whether the personal interactions between buyers and sellers may bias the rating scores left by reviewers on digital platforms and affect consumer demand. Using data and text reviews from Airbnb in Barcelona in 2019, we perform semantic and regression analyses to measure the host's kindness and to quantify the impact of the kindness-related bias on listing ratings and demand. To identify the bias, we exploit the listing's location, an attribute that can be objectively measured through GPS coordinates and cannot be strategically manipulated by the host. We find that kinder hosts receive significantly higher location ratings, given the distance, and higher listing demand, given the overall rating. Furthermore, kindness mitigates the negative impact of a bad location both on location rating and on listing demand. We take advantage of Covid-19 lockdown in early 2020 as a natural experiment to address endogeneity concerns.

Keywords: Rating system, online reviews, kindness, consumer bias, sentiment analysis, sharing economy, Airbnb.

JEL codes: D90, L83

---

*We would like to thank Federico Boffa, Mario Pagliero, Lorien Sabatino, Stefan Trautmann and Lorenzo Zirulia, as well as the participants to the iCare 2021 (Perm, Russia), SIE (2021), SIEPI 2021 (Parma), Cluster seminars in Tourism, Marketing and Regional Development (Bolzano) conferences for the valuable discussions and suggestions made to previous versions of the paper. The authors gratefully acknowledge Edoardo Giacobbe for his invaluable contribution in the construction of the indices of kindness with ML and text-analysis techniques.

[†]Politecnico di Torino, Department of Management, C.so Duca degli Abruzzi 24, Torino, Italy. Email: laura.abrardi@polito.it.

[‡]Politecnico di Torino, Department of Management, C.so Duca degli Abruzzi 24, Torino, Italy. Email: elisabetta.raguseo@polito.it.

[§]Politecnico di Torino, Department of Management, Corso Duca degli Abruzzi 24, Torino, Italy. Email: laura.rondi@polito.it. (Corresponding Author)

1 Introduction

Performance evaluations, online reviews, customer recommendations are increasingly used by consumers and firms, steering their purchasing, business and working decisions. Review systems - rating scores and textual comments - are especially relevant for digital marketplaces, where they provide an indispensable reputational signal, compensating the intrinsic information asymmetry in online transactions.

Unfortunately, the reliability of the rating system is often questioned, due to a substantial bias towards high ratings (Zervas et al., 2021; Dellarocas and Wood, 2008). A considerable effort has been dedicated to understanding the drivers of these skewed distributions, pointing to psychological factors as partly to blame for the distortion (see Magnani, 2020 for a survey). Regret-aversion, as well as the desire to validate the purchasing decision, or even the need to reciprocate the seller's helpfulness may bias to a great extent the ratings left by reviewers. Given the nature of behavioral biases, the aforementioned distortions are likely amplified when a personal relationship is established between the actors involved in a transaction that generates a deeper psychological involvement. For example, in the home-sharing industry, where hosts and guests often meet in person or even live next doors, a good relationship with the host has been found to improve the probability of positive reviews by guests (Fradkin et al., 2021). In this paper, we investigate the role of personal interactions on the bias of online ratings and its impact on market outcomes, specifically on consumer demand.

While the bias of rating systems is plain and its roots clearly identified, its impact on demand outside the realm of experiments has never been investigated before. The reason is that an objective measure of the attribute that is being rated is often unavailable. We exploit data from the Airbnb platform, leveraging on the fact that Airbnb guests rate their sojourn on six dimensions, among which there is its location. The possibility for the guest to directly rate the location of the listing has important implications for analyzing the bias of the rating system. Location has three convenient characteristics. First, the location can be objectively measured through the latitude and longitude data, and this measure can be compared to the subjective evaluation provided by its rating. Second, the rating score of the location should be independent from the host's attitude, if the guest is providing an impartial, unbiased evaluation. Third, location is an exogenous feature of the accommodation that cannot be modified by the host through a kinder service. Therefore, the rating of location provides the instrument for an immediate acid test of whether ratings are influenced by the quality of the personal relationship established with the host. This defines our first research question: May the host's kindness affect the rating score of location provided by the guest?

To answer the question, we examine the content of 668,824 online reviews posted on the Airbnb platform in 2019 in Barcelona and develop a text analytics algorithm and semantic analysis based on neural networks to infer the quality of each host's attitude – i.e., "kindness" - towards her guests. We then use regression analysis to test whether "kindness" may bias the listing's location rating.

We find that the host's kinder attitude is associated to a higher rating of location, after controlling for the actual position of the listing, i.e., its distance from the main tourist attractions. This finding suggests that the host's behavior generates a psychological bias on the rating left by the guest, who may feel indebted to reciprocate the host's kindness with an overly generous score. Moreover, we find that kindness moderates the negative effect of distance on the rating score of location, implying that the bias related to kindness is stronger for listings in a worse geographical position.

The host's kindness might influence not only reviewers, but also prospective guests who make their choice by reading past reviews and looking at ratings. This is our second research question. In principle, the host's kindness should have no direct effect on demand, as the comprehensive rating system of Airbnb could capture not only the quality of the apartment and the comfort of the stay, but also the host's attitude. Yet, we find that a kinder behaviour is associated with significantly higher listing demand, after controlling for the overall rating score. This finding suggests that text reviews convey information that the rating score system does not fully capture, and that this information about the host's behavior has an impact on its own on the listing's demand. Furthermore, we find evidence that the kindness-related bias mitigates the negative impact that an inconvenient location of the apartment has on its demand, hence it has a positive effect also on the listing's market performance. If not for this bias, poorly positioned listings would have received a lower rating of location and a lower number of renting requests. Hence, kindness can be viewed as a modern capital asset that affects the demand for listings.

Being aware that the relationship between kindness and demand is potentially endogenous, our econometric strategy includes many guest, listing, and contract characteristics to account for omitted variables problems. Moreover, it addresses simultaneity and reverse causality concerns by exploiting the discontinuity caused by the Covid-19 pandemic as a natural experiment.

Although several studies have identified the host-guest interaction as one of the key factors for a positive Airbnb experience (Sthapit and Jiménez-Barreto, 2018; Alsudais, 2017; Cheng and Jin, 2019), and despite the fact the true sentiment of the guest might be expressed in text reviews, few studies have analyzed the words of the review to infer the real value experienced. Previous research focuses on the general sentiment of the text review, studying its relationship with the star rating and the effect on the price of the listing (Lawani et al., 2019).

We thus contribute to the existing literature in several ways. First, we use machine learning to develop a measure of the host's behavior. To the best of our knowledge, no previous study has attempted to measure the quality of the interaction with the host through machine-learning techniques. Second, from a management perspective, we quantify the impact of the hosts' attitude on ratings as well as on an important metric of economic performance, such as the listing's demand. Kindness is an indispensable intangible asset in the modern, highly competitive markets (Tillquist, 2008).[1] We find that the positive impact of a kind behavior on demand materializes not only via higher ratings, but also through the value judgment that the potential guests form by reading past reviews. This result has important implications on the role and impact of online reviews, which adds to the measure provided by the rating, and suggests that the reviews may imply information in addition to what just conveyed by the rating score system. Third, from a theoretical point of view, this analysis sheds light on the reasons for the bias of the rating system, highlighting the role of personal interactions on the gap between an un-enthusiastic review and its "5-star" rating.

The rest of the paper is organized as follows. In Section 2 we present the theoretical

[1]Be nice. Might airlines consider kindness as a business strategy? The Economist, Mar 22nd, 2012.

framework and the pertinent literature. In Section 3 we describe the data and variables used for the empirical analysis. Section 4 describes our empirical and identification strategy. Results are presented in Section 5, including a number of robustness checks on alternative definitions and shapes of demand and extensions to single rooms rentals. Section 6 concludes.

2 Theoretical framework and hypotheses

The efficient functioning of digital platforms depends on the possibility for consumers to review suppliers on the basis on the level of satisfaction achieved after the purchase, both through predefined rating systems (stars), and through the option of posting a detailed feedback, that freely reviews various aspects of the product or service. This system is essential both on the demand and on the supply side: consumers reduce the information asymmetry on the demand side, while suppliers are able to reap the benefit of reputation.

What in theory is an efficient system, in practice is vulnerable to psychological biases that may skew ratings on the high-end tail of the distribution (Zervas et al., 2021). Regret-aversion, for example, makes consumers more likely to remember the positive aspects after the purchase, thus minimizing the negative aspects (Lind et al., 2017). Moreover, the desire to validate the purchasing decision once it is sunk may explain the higher propensity to leave a positive review (the so-called purchasing-bias highlighted by Hu et al., 2009). The dimension of informality that characterizes the service provided by the Airbnb platform also contributes to make users more tolerant and understanding (Arcidiacono et al., 2016). As a consequence, guests using the Airbnb platform are more likely to feel satisfaction from the experience, have their expectations surpassed, and thus leave a positive review (Bridges and Vásquez, 2018).

The aforementioned behavioral biases are likely to be exacerbated when a personal relationship is established between hosts and guests, causing a deeper psychological involvement of the reviewer. The interaction with the host is indeed a key attribute used by guests to evaluate their experience (Sthapit and Jiménez-Barreto, 2018; Cheng and Jin, 2019), and nearly 80% of reviews contain a mention to the host (Alsudais, 2017). If a personal relationship is established, the guest might be reluctant to reveal his true opinion, when it comes to "giving bad news" (Dellarocas and Wood, 2008). Furthermore, the personal relationship exacerbates the reciprocity bias, which involves the tacit expectation of receiving mutually positive evaluations and leads to the omission of information that may be unpleasant (Fradkin et al., 2021; Proserpio et al., 2018). Notably, this phenomenon is aggravated by the lack of anonymity of review systems, as the reviews are linked to the user profile –else they would be considered unreliable. These results support the conjecture that a more satisfactory personal interaction between hosts and guests may bias upwards the ratings, even on those dimensions –such as the score given to the location– that should not be related to the quality of the communication with the host. This "halo effect" (Leuthesser et al., 1995) - an individual's tendency to bias his responses about an attribute by his predisposition toward another attribute - has been already detected in the hospitality industry for hotel stays. Nicolau et al. (2020) find that the quality of the interaction with the hotel staff influences the valuation of location. The halo effect might influence reviewers also on the Airbnb platform, where the personal interaction with the host might induce a psychological bias on reviewers and their ratings. Accordingly, we make the following hypothesis.

Hypothesis 1 *The host's kindness is positively related to the rating score location, regardless of the location of the listing.*

The guest's unwillingness to giving "bad news" is likely to be stronger, the worse is the news, as the host is not responsible for the listing's bad location. Hence, the behavioral bias induced by the host's attitude on the location rating is likely higher, the more decentralized is the location of the listing, thus resulting in a higher skewedness of the rating. Thus, we suppose the following:

Hypothesis 2 *The host's kindness positively moderates the negative impact of the distance from the focal points on the rating of the location.*

By observing the impact of the host's kindness on ratings, Hypotheses 1 and 2 focus on the potential bias induced on reviewers, i.e. past guests. However, the host's attitude might have an effect also on prospective guests, who could be influenced to book an apartment by reading about the host's behavior in past reviews. Indeed, the literature studying the role of textual reviews on demand suggests that consumers respond to the content of online reviews, in addition to customer ratings (Lawani et al., 2019). Archak et al. (2011) also find that review textual comments influence consumer decisions even when a star rating system is available. This evidence suggests that the host's behavior might produce a direct effect on demand, beyond the effect passing via the rating channel. We therefore formulate the following hypothesis:

Hypothesis 3 *The host's kind attitude has a positive effect on demand, controlling for the rating.*

Hypothesis 3 suggests that potential guests rely not only on the information summarized by the rating scores, but also on the textual reviews left by previous guests. Such comments provide information on the host's attitude which influences their decision to book an apartment. In the empirical analysis, quadratic and cubic effects of kindness are tested, as prior work reported decreasing returns to kindness (see, e.g. Becker et al., 2012 for tipping behavior). Having verified that kindness affects demand, we could expect that kindness can moderate the negative effect of an inconvenient location on the demand for a listing. Accordingly, we formulate the following hypothesis:

Hypothesis 4 *The host's kindness mitigates the negative effect of a more decentralized location on demand.*

3 Data

We choose Airbnb data in the empirical analysis for many reasons. As already noted, Airbnb data provide the unique possibility to obtain an objective measure of quality to which ratings can be compared. Moreover, there are also specific reasons for which the platform represents an ideal research case for investigating the role of personal relationships in online services. First, there is no intermediary between the parties involved that might further influence the ratings. Second, Airbnb adopts a double-blind rating system, where hosts and guests submit their review before having the possibility to read each other comments. A large majority of trips thus result in a guest's review (Fradkin et al., 2021). Finally, personal interactions are intrinsic features of the Airbnb service, given that hosts and guests often meet or live next door. The connections people make during their stay are deeply human and personal. The platform itself is the perfect incarnation of the philosophy of the sharing economy, building a community of people sharing spaces and experiences and characterized by a strong social vocation. Airbnb specifically attracts users who prefer an accommodation from Airbnb, rather than a hotel, due to the added value provided by social interaction with their landlord (Guttentag et al., 2018).

3.1 The dataset

The data collection involved 8,758 Airbnb listings located in Barcelona in 2019 and the corresponding 668,824 reviews posted by guests on Airbnb website. We choose the city of Barcelona as our research setting because it is one of the cities with the highest touristic inflows in Europe. The unit of observation of our empirical analysis is the listing (the entire apartment), which is run buy a host. For each listing/host, we measure the average *Kindness* as conveyed by the online reviews left by guests since the beginning of the activity up to the end of 2019. We construct three alternative measures of kindness using semantic analysis and machine-learning techniques. We focus on entire apartments, rather than on single rooms, to keep the relationship between host and guest on a similar basis.[2] Given this choice, the number of Airbnb listings in the empirical analysis is 4,150. We adopt a multi-method approach by complementing text analysis with econometric analysis, and we triangulate the data and the reviews sourced from the publicly available Inside Airbnb database (http://insideairbnb.com) and data of AirDNA, a data analytics company that provides information about Airbnb properties (https://www.airdna.com/). The following Section describes in detail how we apply machine learning to measure kindness.

[2]The alternative solution to a private accommodation is to rent a shared or a single room in an apartment or a house, a choice that, in our view, identifies a visitor with a stronger focus on the price point rather than on other aspects of the stay, such as the presence of amenities and the quality of the relationship with the host, in other words guests with a more elastic demand function. As part of our robustness tests, in Section 5.5 the empirical analysis is also performed with hosts who rent shared or single rooms.

3.2 Measuring kindness: A Machine Learning approach

In this study, we used Python to operationalize the measures of kindness by employing instruments of semantic analysis. We elaborated the text of 668,824 reviews posted on Airbnb and available on Inside Airbnb, by following four main steps.

First, we used Python to obtain reviews with a consistent formatting, in order to have the single sentences in the same review divided by only a "full stop" and divide clearly the different phrases in the same review.

Second, all the reviews were translated in English, since the algorithm of machine learning used for analyzing the review text was based on the English vocabulary. In order to make the translation and achieve high levels of accuracy, we used Google Translate API based on Natural Language Processing (NLP), which is able to identify directly the original language. We then counted the number of words of each review and deleted from the database empty online reviews.

Third, in order to derive information about the host's kindness, we extracted from each review the sentences where the guest commented on the host's behavior and the interaction with him during the stay. To this aim, we applied a mechanism called "tokenization". It is based on considering the review as a set of phrases, and for every phrase verifying if it is in line with the topic of interest. The criteria of conformity of the phrase with the topic of interest was based on the presence of a series of words (specifically, the words used are the personal pronouns, and the terms "host", "owner", "staff", "questions", "helpful", "help", "recommendation", "communication", "service", "friendly", "responsive") that could be associated exclusively to a person, in order to avoid more generic words, such as "wonderful" or "great", that could apply also to the apartment or to the city.

Fourth, we elaborated all the reviews associated to each host through machine learning, with the aim of identifying the level of kindness of every host. In particular, we obtain a first measure of kindness, *Score POLARITY*, by implementing a currently available tool of Sentiment Analysis, as described in Section 3.2.1. Second, we design our own ML algorithm based on neural networks to obtain a tailored measure of kindness, *Rank*, as described in Section 3.2.2. Finally, we also adopt a more intuitive, measure of kindness (*First-name*), defined as the percentage of reviews where the guest refers to the host by using her/his given name. The underlying idea is that being on first-name-terms alludes to a familiar and friendly relationship. To construct our First-name variable, we exploit a text-mining algorithm that dives into the review and examines whether the host is called by his given name, which is an information contained in the original dataset.

3.2.1 Measuring kindness by Sentiment Analysis

We elaborate a machine learning-based measure of kindness, that we call *Score POLARITY*, based on the analysis of the sentiment contained in the review through the usage of a Python package. "Score POLARITY", is based on the Polarity function of Python extracted from the Textblob package. The package receives as an input the text of the review, analyze its content by searching for words with high intensity, based on the specific dictionary of the package, and provide as output a value between -1 and 1. This value represents the level of positivity of the sentiment related to the text. Values lower than 0 represent negative reviews, whereas values increasingly positive identify reviews with a higher degree of positivity. We thus assign a score ranging between -1 and 1 to every review of every host. Since the empirical analysis is at the host level, we computed the average value of all reviews' POLARITY scores for every host and every listing. Finally, we normalized this mean score value on a scale between 1 and 10. Table 1 reports the output the Polarity algorithm for a small sample of reviews. While it exhibits a good ability to distinguish between positive and negative reviews, it is less performing when reviews are classified with higher granularity. For example, the comment "The hosts were very helpful and communicative" earns a brilliant 0.7 (in a range where 1 is the maximum), whereas a review with similar sentiment, "Ana is a great host and communicative" is associated with a quite lower sentiment of 0.6. Even more strikingly, the comment "After our initial host cancelled our trip on the day, Monica was very quick to say we could stay from that night" earns just a 0.07, barely positive.

Given the potential marginal error in interpreting the true feeling of the guest, we construct two additional indicators of the host's kindness. The first the *First-name* variable, which indicates whether the host is called by the given name in the guest's comment, and the second is the variable *Rank*. The latter indicator is based on a machine-learning algorithm that exploits neural networks to classify the host's behavior emerging from each review in one of four classes, namely negative, neutral, positive and excellent.[3] The variable Rank is obtained as a weighted average of the relative frequencies of the four classes for every listing. In the following Section we describe in details the algorithm used to classify each review into one of the four classes.

3.2.2 A neural network model to measure kindness

We now describe in more detail the design and parameterization of the neural network model used to classify reviews. The process entails a pre-processing phase, necessary to prepare the input data; a phase where the architecture of the network is designed and the model parametrized; a training phase; and finally a validation phase.

Pre-processing

A neural network takes numeric data as input, processes it and returns an output consistent with the chosen activation function. When the input is a text, it thus needs to be transformed into a numeric form to be processed by the neural network. This operation is usually executed through the approach known as bag of words, which consists of a script that reads the reviews and saves the N most frequent words. This operation produces a data structure -a "dictionary"-, which associates an index between 0 and N-1 to each of the most frequent words, in order of decreasing frequency. At this point, each text is converted in a binary vector of N elements, where each element has a 1 if the word associated to that position in the dictionary is present in the review. The crucial part of the bag of words approach is the choice of the parameter N: the higher the number of words in the dictionary, the higher the model's ability to learn, as the number of information on which to infer increases. However, there are drawbacks to a large N. First, an increase of N increases exponentially the computation cost of the model. Second, a large N exposes to the risk of overfitting. In

[3]Sentiment Analysis typically employs a classification approach based on three classes: negative, neutral and positive. However, such a classification does not fit well in our context, where the frequency of negative reviews is extremely limited, while positive reviews exhibit a significant heterogeneity, ranging from mildly positive reviews ("the host is kind") to a detailed description of the exceptional experience with the host. For these reasons, we adopt a classification based on four classes, distinguishing between positive and excellent reviews.

particular, the model becomes extremely efficient at inferring in the context of the words it has learned, but recognizes such a strong meaning to certain particular combinations of words, that it loses the ability to generalize what it has learned and therefore to infer on data external to one's own database.

Using the parts of reviews discussing the host's behaviour, we created the bag of words dictionary by means of the package of pre-processing functions included in the Keras environment of Tensorflow.[4] After several attempts, starting from N = 1000, up to extreme attempts with 100,000 words, we selected N = 20,000 as size of the dictionary. Following the implementation of the bag of words procedure, we obtain for each review a vector of 20,000 elements, where the i-th element is 1 if the review contains the i-th word of the vector, and 0 otherwise. Notably, words with lower i in the vector occur with higher frequency in the set of reviews.

Architecture of the neural network

The concept of a neural network could be illustrated as a parallelized computational structure composed of interconnected neurons that transform inputs into outputs. It is defined parallel because each layer is composed by a certain number of independent neurons. The network is obtained by combining, in different possible ways, a certain number of layers. The input layer collects input patterns. The output layer has classifications or output signals to which input patterns may map. In between there may be hidden layers, which fine-tune the input weightings until the neural network's margin of error is minimal. It is hypothesized that hidden layers extrapolate salient features in the input data that have predictive power regarding the outputs. Between two adjacent layers, multiple connection patterns are possible. They can be fully connected, with every neuron in one layer connecting to every neuron in the next layer. They can be pooling, where a group of neurons in one layer connect to a single neuron in the next layer, thereby reducing the number of neurons in that layer (Ciresan et al., 2013). Neurons with only such connections form a directed acyclic graph and are known as feedforward networks (Zell, 1994). Alternatively, networks that allow connections between neurons in the same or previous layers are known as recurrent networks (Miljanovic, 2012). The architecture of the network is also defined by the number

[4]Tensorflow is a open-source software library for machine learning developed by Google.

of layers, and the number of neurons for each layer. As size grows, the model's predictive ability increases, but also the computational cost. The last element of the neural network is the activation function, which expresses how the input of the last layer is converted into the information in the output.

For our purposes, we adopt a sequential and dense structure, i.e., each neuron of a subsequent layer is connected to all the neurons of the previous layer. This large number of connections allows us to explore a virtually infinite space of combinations. Moreover, we adopt an architecture based on six layers. Following common practice, the first layer has the size of the input data, in this case 20000 (choosing any lower number would mean performing calculations to extract data and then not include it in the analysis), while the last layer has as output the dimension of the possible alternatives of the classification. The number of neurons in the intermediate layers represents how the information contained in the vocabulary of words is processed and converged to our four lasses. For the second layer we supposed that many frequent words in the dictionary were actually of little information, such as conjunctions and articles. We thus decided to keep only one word out of 4, thus generating a second layer of 5000 neurons. We progressively reduced the number of neurons of subsequent layers, using 1000 words for the third layer, 500 for the fourth and 100 for the fifth. The sixth layer transforms the last 20 words into the required 0/1 output that classifies our review into the four classes. To convert the input of the last layer into our desired output, we use a sigmoid activation function, that is a statistical tool that, using the logistic regression carried out starting from the data of the last layer, returns the probability that the review belongs to each class. The review is then assigned to the class that displays the maximum probability. Reviews that do not mention the host or his behaviour are automatically classified as neutral.

Training and validation of the algorithm

The model has to learn the logic of the human assignment. To this aim, it has been trained on a subset of 5000 reviews, which have been manually ranked. In particular, the reviews, converted into numerical vectors by the bag of words approach, and the information on the manually associated class, were then fed to the model.fit function of the Tensorflow environment. This command performs a recursive optimization of the internal parameters of the model, with the aim to maximize the accuracy of the model. The model continues to "cycle" on the dataset, until the accuracy of the prediction is higher than in the previous iteration. The model, once optimized, settled on internal accuracy values of around 90%.

The validation of the algorithm is carried out on a different sample than the subset used during the training phase. To validate the model, we exploit the model.evalute command of the Tensorflow environment, which returned accuracy values comparable to the results of the training phase, as desirable. The validated model was deployed on the complete database. At the end of the execution, four binary variables are created, corresponding to the classes "negative", "neutral", "positive" and "excellent" review. To get a feeling about the performance of our neural network algorithm, Table 1 reports its output for a sample of reviews. As it can be seen from the Table, the Rank classification resulting from the neural network algorithm is able to identify as positive reviews also comments where the positive sentiment is somewhat more nuanced and implied. For example, the comment "After our initial host cancelled our trip on the day, Monica was very quick to say we could stay from that night" is classified as positive and the algorithm is not misled by the negative comment about the previous host. In a similar fashion, the comment "although we never met Eduardo because the check-in and check-out were self-made, he was always available by phone for any eventuality", is classified as positive despite the initial mention to a host that does not welcome guests in person. As a final step, the classified reviews are aggregated to obtain a measure of the host's kindness. To this aim, we convert each class in a (decimal) grade (10 for excellent reviews, 7.5 for positive, 5 for neutral and 0 for negative), and calculate the host's kindness (Rank) as the weighted average of the grades obtained on all the reviews received by that host. The weights are given by the frequency with which the class occurs in his set of reviews. The final output is a variable, *Rank*, which measures each host's kindness on a scale between 0 and 10. Figure 1 reports an example for the host Jordi, whose Rank is obtained as the weighted average 0*0+5*0.5+7.5*0.25+10*0.25=6.875. Table 2 shows the distribution of the relative frequencies of reviews at the host level. As expected, neutral or positive reviews are the predominant type: the "typical" host has an average of 58% of neutral reviews and a 30% of positive reviews. About 10% of reviews of each host are excellent, while negative reviews are rare and weight only 1.2% of reviews.

Finally, Figure 2 reports the distribution of the rating and kindness measures that we use in the econometric analysis. While the rating appears strongly biased upwards, the kindness measures exhibit a more balanced distribution.

— Table 1, Table 2, Figure 1 and Figure 2 around here —

3.3 The Other Variables

In this section we describe the other variables used in the empirical analysis. As outlined in the theoretical framework, the hypotheses focus on the relationship of kindness with two characteristics of the listing, the rating of its location left by the guests at the end of their stay, and the listing's demand in 2019.

Review score location rating is the average of the scores that all guests assign to the location of the listing, according to their experience. It is normalized by Insideairbnb on a scale from 1 (worst) to 10 (best) for the location category and is available and visible on the listing's website on Airbnb. The individual scores in the Airbnb rating system are highly correlated among themselves and with the overall rating score (see below), but our data show that location rating exhibits the lowest cross-correlations.

Number of reviews is the total number of reviews posted on Airbnb for each accommodation in 2019, which we use as a proxy of the listing's demand in the logarithmic form (see, e.g., Quattrone et al., 2016 and Lawani et al., 2019). Although there is no obligation to leave a review after the stay, the strong social vocation of the platform, the frequent reminders to the guest and the recent implementation of a reciprocity rule whereby the guest and the host are allowed to see each other review only if they both leave it (see Proserpio et al., 2018) reduce the possible selection bias due to reviews being non-compulsory. Not surprisingly, Airbnb estimates that at least 70% of the guests leave a review about their hosts, while (Fradkin et al., 2021) reports that at least 67% of reservations end up in a review.[5]

[5] With regard to the potential selection problem with guest reviews, we found interesting evidence provided by Fradkin and Holz (2022). Exploiting a field experiment on Airbnb, they tested the effect on the listing's market outcomes of a policy meant to incentivize guests to release their reviews. They found that the additional incentivized ratings did not affect the quantity of nights sold, i.e. the demand, suggesting that the selection problem should not be severe. Notwithstanding this, as a robustness test, we re-estimated our models using the number of reservation days and the occupancy rate, sourced by AirDNA, and found that the results are very similar. They are available on request

Another key variable in our analysis is the listing's *Average distance* from the 14 most popular places for tourists in Barcelona (mapped in Figure 3), which is designed to measure how strategic the location of accommodation is.[6] We identified Barcelona's tourist attractions by cross-referencing different tourist guidebooks and blogs (Tripadvisor, Lonely Planet, Skyscanner, Barcellona.org). Then, we exploited the latitude and longitude data of each listing to determine the distance in kilometers from each touristic attraction. Finally, for each listing we computed the average of the distance from all 14 touristic attractions. As explained in the section on the empirical strategy, this variable is the instrument that allows us to identify the kindness-generated bias, both the one to which is subject the guest when writing the review and the one to which are subject Airbnb users when choosing their accommodation on the basis of the reviews and ratings they read online.

— Figure 3 around here —

3.3.1 Control variables

Among the many control variables we include in the econometric models, the overall Airbnb rating, *Review score rating*, plays a key role. It is a score based on a ten-level scale from 1 (worst) to 100 (best) through which the guest evaluates his overall experience.[7] The overall rating has been found by the previous literature to positively affect the listing's demand (Gunter and Önder, 2018). Another reason to control for this variable is that the host's "kindness", or personal attitude, is not separately evaluated in the Airbnb rating system and might therefore be included in the overall score given by the guest. In order to ascertain whether kindness has a role of its own, we therefore must account for the effect of the overall rating, separately.

We also add a rich set of control variables that are widely used by the literature on short-term rental platforms (see, for example, Liang et al., 2020). These variables describe

[6]Specifically, the fourteen tourist attractions are: Placa d'Espanya, Placa de Catalunya, Sagrada Familia, La Rambla, Casa Battlo, Casa Milà, Barcelonetta Beach, Boqueria, Park Guell, Historia Museum, Castell de Montjuic, Mirò Foundation, Gracia district, Music Palace.

[7]Airbnb asks the guest to evaluate the visit according to seven scores: the accuracy of the information provided by the listing's website, the apartment's cleanliness, its location, the check-in procedures, the quality of the communications with the host, the value for money and, finally, the overall rating.

the host and property attributes as well as the service and contractual terms.

The host attributes include the year when the listing first entered the platform to proxy for her skills and experience (*Host experience*), whether she owns or manages multiple listings, possibly indicating a more professional, and detached, approach to social relations (*Multiproperty*), and her *Superhost* status. Superhosts are proficient and reliable hosts with a record of very high rating scores, who are expected to provide excellent experiences for their guests (Ert and Fleischer, 2019). The Superhost status is awarded by Airbnb and it appears on the listing and profile page to help customers to identify them. We also control for variables indirectly related to the host's social behavior, such as the *Host response rate*, the share of inquiries and booking requests the host replies to (by either accepting/pre-approving or declining) within 24 hours; the *Host acceptance rate*, the percentage of the accepted reservations by the host; and a binary variable equal to one if the host identity is verified (*Verified host identity*).

Among property attributes, we control for the number of *Bathrooms*, *Bedrooms* and of guests it can *Accommodates* of the listing, as proxies for the apartment's size, and the *Number of photos* on the listing page, as a host with a nice house is expected to post more pictures thereby attracting consumers (Gunter and Önder, 2018). In addition, to further control for the pleasantness of the sojourn, we construct two variables related to the quality and luxury endowments of the house, based on a selection from the list of 191 optional amenities on the Airbnb website. *Quality amenities* is the number of functional but relatively costly comfort features - air conditioning, microwave, dishwasher, washing machine and complimentary parking space - which are available in the apartment and represent its quality endowment as well as a sunk cost for the host.Similarly, the *Luxury amenities*, indicating the presence of a garden, a swimming pool, sauna and a terrace, refer to a greater well-being or elegance and involve a larger investment for the apartment.[8]

Finally, a set of variables accounts for the terms of service conditions. *Price*, the average price of the listing in 2019, is expected to affect its demand; the binary variable *Strict*

[8]Luxury amenities have been selected through a criterion based on the frequency of their occurrence, being rare commodities. The garden is present in 5.5% of the listings, the swimming pool in 3.3%, sauna in 0.07% and the terrace in 1.1%.

cancellation denotes that a tight cancellation policy is in place,[9] while *Guest phone number* and *Security deposit* respectively indicate whether the guest must provide a phone number and an advance security deposit to book the apartment.

Table 3 reports all the variables we use in the empirical analysis, their definition and their source. Tables 4 and 5 show the descriptive statistics of this study.

— Tables 4 and 5 around here —

4 Empirical Strategy and Identification

The empirical analysis develops in three steps. First, we test if the "halo effect" (Leuthesser et al., 1995; Nicolau et al., 2020) generated by the host's kindness affects – biases - the guest's rating of the apartment's location. There are several reasons to use this variable rather than other attributes of the apartment as the instrument to identify the bias. First, distance cannot be shortened nor modified by the host to increase the apartment's attractiveness (for example, by buying a new mattress or installing the air conditioning). Second, it is objectively measured in km, as opposed to quality indices of amenities or services that are subjectively appreciated by hosts, based on their individual taste. Third, distance is rated separately by a clearly defined score within a rating system that covers six different aspects and one general evaluation of the sojourn. As such, the location's rating should reflect the valuation of the actual position of the listings rather than of other features. And indeed, the Airbnb rating system provides the guest with three items (out of six) to express his "true" opinion about his personal interchanges with the host as regards the quality of communications with the host, of check-in procedures, and of the information on the website. We reckon that a disappointed guest is more likely to vent his dissatisfaction about such personal services by using the three appropriate scores rather than the rating of location. Moreover, as already pointed out, although the cross-correlations among the specific rating scores is quite high, it is lowest for the rating of location. For all these

[9] According to the Airbnb Strict cancellation policy, guests may receive a full refund if they cancel within 48 hours of booking and at least 14 full days before the listing's local check-in time. After 48 hours, guests are only entitled to a 50% refund regardless of how far the check-in date is.

reasons, location rating may be the ideal instrument to capture the kindness-related bias. We thus attribute the guest's bias in the location's rating to the halo effect generated by the host's kind behaviour and to the guest's reluctance to give the "bad news" to a gentle but poorly positioned host.

The second step focuses on a more tangible effect of the bias, that is the impact of the host's kindness as conveyed by guests' reviews on the listing's demand by potential guests who make their choice after reading the reviews and checking the rating.

We are aware that the identification of the causal effect of kindness on the quantity sold may be difficult due to omitted variables, simultaneity and reverse causality. To address the omitted variable problem we include a large set of controls on the apartment and the host's characteristics, the website's informativeness and the terms and conditions of the contract, all sourced from InsideAirbnb website. However, as highlighted by Proserpio et al. (2018) with reference to the effect of reciprocity on pricing, unobserved improvements in the quality of listing's furnishings could influence the rating and possibly the reviews. In our analysis, though, the algorithms calculating the kindness indexes employ only sentences that describe the host's attitude, not the apartment. This should greatly reduce the potential effect of quality shocks on reviews and, in turn, our measure of kindness. In addition, we control for the quality of the apartment with two variables based on the list of amenities described in the website, and we include the Airbnb overall rating, which is expected to absorb the residual positive effect of a shock to quality. Finally, our dependent variable is the listing's demand in 2019; hence, only investments in quality that occurred in 2019 should influence the listing's demand, probably via the overall rating. Conversely, the variable of interest, kindness, is obtained from the cumulated host-specific reviews left by guests since the listing's creation, and we include the starting date as a control variable.

Another threat to identification may come from reverse causality, as hosts with inadequate or poorly located apartments facing low demand might adjust their behavior to obtain more favorable reviews. Comfortingly, the correlations (see Appendix Table A.1) between each of our three measures of kindness and quality features such as distance from focal points or valuable amenities available in the apartment are very low. Moreover, the scatterplots in Appendix Figures D.2, D.3 and D.4 confirm that host's kindness does not increase with distance. This evidence suggests that hosts do not seem to adjust their behavior to compensate the weaknesses of their apartment with a kind behaviour.

— Figure D.2, Figure D.3 and Figure D.4 around here —

Nevertheless, we further address reverse causality and simultaneity with an econometric strategy. We exploit a natural experiment provided by the extraordinary situation created by the lockdown that followed the spreading of Covid-19 in early 2020, which froze all touristic and business trips for several months in 2020 and was then lifted in May-June for a few months.[10] As renting on Airbnb portal practically stopped for several months, no more comments were added in that period. Hence, we can safely assume the new visitors that started traveling after the lockdown was removed made their renting decisions on the basis of pre-Covid, i.e. "old", reviews and old rating scores. This leaves us with a unique opportunity to test the impact of previously registered kindness on current demand, thus mitigating suspects of simultaneity or reverse causality. With an unexpected gap of nearly six months, there was nothing an unkind host could to do to rehabilitate her reputation and revamp the listing's demand. To conduct this test, moreover, we take an additional precaution: imaging that some of the listings in the market in 2019 would not be posted by risk-averse, cautious owners, we account for sample selection by estimating a two-step Heckman model. Therefore, we estimate the impact of historic kindness on current demand for the sub-sample of listings that have received at least one review in 2020, after controlling for the factors that led to the selection of the sample, i.e., the probability that they were actually posted and rented. The underlying idea for breaking up the link between contemporaneous host behavior and reviewed kindness is that, by lagging kindness nearly one semester we should be able to isolate our variable of interest from the possibility of reverse causality or simultaneity. In a way, it is as if we converted "kindness" from a fluid, adjustable characteristic of the host to a steady attribute of the listing (kindness-capital), which is expected to affect its demand. Insofar as prospective guests in summer 2020 made their choice by reading the reviews left by tourists up to before the Covid 19-related lockdown started, they allow us to

[10]Speaking to Yahoo Finance Live, Airbnb's CEO and Co-Founder Brian Chesky recently recalled that in 2020, Airbnb saw its business depleted when the coronavirus pandemic hit, losing 80 percent of its business in just eight weeks.

test whether the host's kindness affects the decision to rent the apartment, along with the other characteristics and the Airbnb ratings.

The final step of the analysis employs the host's kindness as a moderating variable to further investigate the impact of the bias induced by the host's behaviour on the listing's demand, when the listing's location is inconvenient and would per se negatively affect its demand. In this test, we estimate whether the host's kind attitude the potential guests perceive from the reviews can mitigate such negative effect and persuade them to book the house anyway. This would provide further evidence of the independent role of the kindness capital in the bundle of attributes that affect Airbnb demand.

5 Results

5.1 Does kindness bias the reviewer's rating of the apartment?

To determine if a bias exists, we estimate the relationship between the host's average kindness – something that has to do with personal interactions and empathy in the host-guest relation - and the average rating of the listing's location, - a feature that is expected to receive an objective evaluation by the guest. We control for the listing's actual average distance from 14 focal points, its demand and the average price, the number of years the apartment has been active in the platform, and a large number of additional variables. The score of location is visible in the listing's website and concurs to the overall score that is the average of the six items. As already argued, we choose "location" because it is the only intrinsic characteristic that is rated by Airbnb with a separate score and can be measured objectively, in contrast with the other items (quality of the communications with guest, of the check-in and of website information, cleanliness and "value for money" of the transaction), which are more likely flawed by subjectivity and more susceptible to be influenced the host's behaviour, or not measurable. In practice, a "bias" exists if the host, through a kinder attitude, sways the guest to assign a location rating not reflecting the actual positioning of the listing. Under the alternative, the host's kindness should not affect the location rating that a rational and objective – unbiased - reviewer assigns.

The first three columns of Table 6 report the estimates of the direct effect of our three measures of kindness on the rating of location, controlling for the average distance from the focal points (Hypothesis 1). We find, not surprisingly, that the rating of location is negatively and significantly related to the average distance. However, we also find that the location's rating is positively and significantly related to our three measures of kindness, which is less intuitive, since a rational reviewer should not be affected by the host's kindness when evaluating an objective characteristic such as the apartment's distance from the city's focal points. This evidence is consistent with our Hypothesis 1.

— Table 6 around here —

We now turn to the control variables. Knowing that the separate Airbnb ratings are highly correlated with each other and that there is a typical upward bias in all rating systems (Bridges and Vásquez, 2018; Fradkin et al., 2021), we have included a long list of control variables to capture the house and host's features that usually affect the guest's rating of other items but should be unrelated to the listing's position. Our purpose is to isolate the effect of kindness on the location's score by cleansing the confounding factors. And, indeed, we find that the location's rating is higher when the host is a Superhost, is more experienced, is not a professional operator, and his/her response rate is lower, when the quality of the apartment's endowments is higher and the house is small. All these features should not affect the rating of location.[11] The fact that they do validates our strategy to control for the implicit correlation of location rating with the other scores if we are to isolate the impact of kindness, suggesting that the rating of location, opportunely cleansed, should be a reliable measure of the guest's opinion about the listing's geographical position. Hence the evidence of a positive relationship with kindness should actually reflect the "halo effect" of the host's behavior.

To nail down the bias of the reviewers, in Columns (4)-(6) we add the interaction between kindness and average distance in order to test whether the size of the bias increases as the location of the apartment becomes more inconvenient (Hypothesis 2). This occurs if

[11] In contrast, other characteristics such as requesting the guest's profile picture, having the identity verified, quickly responding to the guest – hinting at greater eagerness to establish a friendly relationship with the host -, strict cancellation policy, the number of photos, and price do not enter significantly.

kindness mitigates the negative effect of distance, due to a biased evaluation of a physical attribute. Results show that the coefficient on the multiplicative term is positive and highly significant. Both kindness and distance remain significant. This result suggests, in line with our prediction, that the host's kind behaviour has a positive effect in soothing the judgment of the guest even on an objective characteristic such as the distance from the focal points and that the guest's reluctance to give "bad news" to the host is stronger the worse is the news, i.e., the greater is the distance. Indeed, at the bottom of the table we test the effect of kindness on the rating score at the median value of the average distance from the focal points (2.14 km) and we find that the impact is positive and significantly different from zero. Analogously, our results show that kindness mitigates the negative impact of distance as we find that, evaluated at the median value of kindness (for all three measures), the impact turns significantly positive.

Recall that none of our three measures of kindness is correlated with the listing's average distance from the focal points. Correlation with Score POLARITY is 4.8%, with Rank is 4.7% and with First-name is 0.74% (Table A1). In the appendix, Figures D.2, D.3 and D.4 confirm the absence of either a linear or nonlinear relationship. This evidence is quite relevant for our study, as it strongly advocates against the possibility that hosts behave strategically, i.e. becoming kinder to compensate for the bad location. In other words, the data suggest that the listing's owner does not adjust her/his behavior to the inconvenient location of the apartment, and that distance is a reliable instrument for the identification of the impact of kindness on the listing's demand.

5.2 Host kindness and listing demand

Table 7 reports the results of the analysis of the relationship between the log of the number of reviews in 2019 and kindness, controlling for the average price in 2019 and a large number of characteristics of the apartment (size, average distance from points of interests in the city, quality of the apartment as described by the available amenities,), the terms and conditions of the contract (average price, cancellation policy, security deposit, requests of guest's identification and the number of photos in the announcement). Moreover, we include several attributes of the host's indicative of her attitude towards the guests and her personal reading of the "sharing" component of the peer-to-peer economy (status of Superhost, acceptance and response rates, verified identity, number of photos in the website, and whether the host is an individual or a professional, i.e., part of a multi-property agency, starting date of the activity, as a proxy of experience).

— Table 7 around here —

Among the control variables, a key role is played by the listing's average Airbnb rating, which is expected to exert a strong influence on the guest's choice and the apartment's demand. Because the overall rating is likely to capture not only the physical characteristics of the apartment or a general evaluation of the stay but also some aspects of the host's attitude, it is not obvious that host's kindness might have an effect of its own, in addition to the rating. By controlling for the overall Airbnb rating, we thus test whether kindness has an independent effect on the listing's demand.

Table 7 shows the results with a quadratic specification of the relationship between the listing's demand and kindness, as measured by our three indices. At the bottom of the table, we report the tests of joint significance of the two coefficients.

Results show that the listing's demand is positively and significantly related to kindness even controlling for the positive and significant effect of the overall Airbnb rating. The evidence supports our Hypothesis 3.

Turning to the control variables, we find that the distance from the city's points of interest has a negative effect on demand, that the listing's demand is negatively related to its price and positively related to the host's experience, to the number of posted photos (a proxy for the quality of the apartment, as the host is willing to show it), and to the quality of the furnishing. Demand correlates positively also with the host's response and acceptance rates as well as with her willingness to have her identity verified. Conversely, we find a negative relationship with the request of a security deposit and with the dummy identifying apartments professionally managed by agencies or by hosts with many listings

5.3 A natural experiment exploiting Covid restrictions on tourism

We now present the results of the Heckman two-step model that estimates the effect of pre-Covid kindness on post-lockdown demand while accounting for the fact that some of the apartments were withdrawn from the market due to the pandemic. In the first step, the dependent variable is a binary variable that indicates whether the listing has received at least one review (i.e. has been rented at least once) and in the second step the outcome variable is the log of the number of reviews to the listing in 2020, when traveling restrictions due to Sars-Cov 2 were (temporarily) lifted and Barcelona reopened to tourism and business visitors. All regressors in the second step, except for kindness and average rating, and distance (which does not change), are dated at 2020 whereas, in the first step, the determinants of whether the listing will be active in 2020 are dated at 2019, since there was a hiatus from the beginning to the end of the lockdown.

Results are reported in Tables 8 and 9 respectively. In the first step, we find that the probability to participate the market when the restrictions were lifted in the summer of 2020 was positively related with the host's kindness, the overall rating score and the host's past year acceptance rate. Notably, the multi-property dummy enters with a positive sign, which suggests that business agencies were more willing to stay on the market than private hosts, probably less eager to entertain personal contacts with guests during a pandemic. As for the other control variables, most enter with the same signs as in the analysis of Table 7 .

— Table 8 and Table 9 around here —

Turning to the second step, where the dependent variable is the active listings' demand in 2020, we find that all three indicators of kindness enter significantly in the regression, confirming that when visitors have to pick an apartment for their stay, they choose based not only on the physical characteristics of the apartment, the Superhost status or the number of photos in the website, but also based on the host's gentle behavior as implied by the comments written by past guests up to the previous year. Intriguingly, those reviews are the last available information with a human/personal content before the shock of the lockdown.

Results also show that demand is negatively affected by the listing's price as well as by its distance from touristic points of attractions while, quite surprisingly, it does not seem to respond significantly to the overall rating while also the multi-property dummy is now insignificant. To sum up, this finding, exploiting the Sars-Cov-2 related shock to tourism to unravel the potentially simultaneous relationship between demand and kindness, confirms kindness as a positive and significant determinant – actually, an independent attribute in the bundle of characteristics that affect Airbnb demand.

5.4 The moderating role of kindness

In the final step of our analysis we investigate whether kindness, inferred from reviews left by previous guests can affect the current guests' choice not only directly, as shown previously in Tables 7 and 8, but also through a bias which may lead the guest to place more importance on the behavioral traits that he has learned from the reviews than on a physical characteristic like location. In Section 5.1 we showed that the host's kindness succeeds in biasing the guest's judgment when he has to rate the apartment's position, we now estimate a model that tests whether kindness can mitigate the negative effect of distance on demand by a new guest. This would indicate that the bias generated by kindness has a tangible, economic role, not only the immaterial halo effect on written comments.

We modify the model in Table 7 by adding the interactions of distance with the two polynomial kindness terms of the previous specification. As usual, we include the overall Airbnb rating to control for the effect of rating on the listing's demand. The results in Table 10 show that the kindness scores remain highly significant as well as their interactions with average distance. At face value, they suggest that the host's kind behavior mitigates the negative impact on demand of an inconvenient location of the apartment, thus confirming a tangible as well as a halo effect. At the bottom of the table, we report the tests showing that, in spite of the high cross-correlations, the interacted terms are jointly (as well as individually) significant.

— Table 10 around here —

To better illustrate our results, we calculated the elasticity of demand with respect to distance (in km) and how this elasticity changes at increasing levels of the kindness indexes.

If kindness actually succeeds in moderating the negative impact of distance, we should find that the sensitivity of demand to distance gradually decreases at higher values of kindness. In Appendix D Figures D.5, D.6 and D.7 we used the regression results in Table 10 , and plotted, for each kindness measure, the elasticity as a function of average distance for different level of kindness, i.e., at the 50th, the 90th and 95th percentile of the distributions. We find that, as expected, demand becomes more elastic, i.e. demand decreases, as listings become more and more distant from the focal points. More interestingly, we also notice that as average kindness increases, the slope flattens, suggesting that the host's kindness reduces the guest's demand sensitivity to a bad location, i.e., it moderates the negative effect of an adverse characteristic of the listing. The same evidence holds for all kindness indicators.

5.5 Extension and robustness

For completeness, this section presents two robustness tests and an extension of the empirical analysis. Results are in the Appendix.

To proxy for the listing's demand we have so far used the (log of the) number of reviews in the year. The debate on the possibility that selection bias and underestimation may affect this measure of quantity, suggests that we re-estimate our models using the number of reservation days and, alternatively, the occupancy rate, i.e., the number of booked nights divided by the sum of the available nights and booked nights as dependent variables. Focusing on hypotheses 3 and 4, which relate listing demand to host's kindness, we find that they both were not rejected when using alternative definitions, confirming the positive effect of kindness on demand (the only exception is when we test the moderating effect of kindness on the negative impact of distance using FirstName as a measure of Kindness). Results are available on request.

The second robustness test concerns the functional form of the relationship between demand and kindness, which we have so far hypothesized as quadratic to express the diminishing returns of kindness. Since the inverted U-shape form implies that the effect of kindness becomes negative after reaching the maximum value, we also experimented with a cubic form, adding a third term to the polynomial. The third term might thus capture whether the returns to kindness, after a plausible deceleration, might still remain positive. The results in Appendix C Table C.5 and Appendix Figure C.1 show that this is actually the case, as they confirm that the positive effect of kindness first increases steeply, then slows down, flattens but remains positive .

In the next piece of evidence, we extend the analysis to hosts that rent shared and single rooms instead of entire apartments. In fact, one might think that guests who actually share their living space with hosts or other guests might be more sensitive to the behavioral traits of the host. We thus repeat the full set of regressions with the sample of single/shared rooms. Results are in Appendix B Tables B.2, B3 and B.4. We find that the previous results hold for hypotheses 1 and 2 (confirming that kindness generates a biased rating of location which mitigates the effect of a poor positioning of the house) and for hypothesis 3 on the positive effect of kindness on the choice of the listing. The evidence does not hold for hypothesis 4, whereby kindness mitigates the negative effect of distance on the listing demand, suggesting that kindness might play a less relevant role in shared/single rooms, in spite of the greater "sharing" content of the accommodation. Although this result may seem surprising at first sight, one has to consider that Airbnb users who choose to go in single rooms are likely more sensitive to price than to other aspects such as the host's attitude and the apartment's furniture. And indeed, if we compare the coefficients of the control variables of single rooms and entire apartments, we find that quality is no longer significant, that luxury turns out with a negative sign (significant with Firstname) and that the verification of host identity is now insignificant. Moreover, using back on the envelop calculations of demand elasticity from log-log, we find that, on average, a price increase of 1 percent generates a demand drop of 0.18 percent in single rooms and of 0.09 (or even less) in entire apartments. The higher demand elasticity in single rooms thus suggests that their guests are less sensitive to the host's behaviour than to the accommodation's price.

6 Conclusions

The rating system provides invaluable information about the quality of transactions, and is at the foundation of most digital platforms. However, rating mechanisms are well-known for their substantial bias. Nearly 95% of Airbnb properties boasts an average rating of either 4.5 or 5 stars (the maximum); virtually none have less than a 3.5 star rating (Zervas et al., 2021). The end result is an obvious information problem, where the partial and not completely truthful disclosure of information in the rating undermines the reliability and stability of the entire mechanism.

In this paper, we shed light on the origins of the distortion by studying whether and to what extent the host's attitude biases the ratings left by guests on Airbnb platform and affects the demand for the listings. To this aim, we perform semantic analysis on reviews left on the Airbnb platform in Barcelona for entire apartments and derive two alternative measures of the host's kind attitude towards the guests, as described in the reviewers' comments. In addition, we devise a third measures of kindness that returns the degree of familiarity in the host-guest relationship, by calculating the share of the reviews in which the guest calls the host by first name.

Our results show that the kindness shown by the host to the guest during her stay can induce an upward bias on the rating left by the reviewer. We focus on the rating of the listing's location, an attribute that can be objectively measured and cannot be altered following bad reviews. By exploiting information on the position of the listing, we find that the bias on its rating is stronger, the worse is the listing location, suggesting that reviewers may find especially difficult to rate objectively when the host has been kind and a bad feature of the listing is not perceived as his fault. Hence, kindness mitigates the guest's evaluation when he has to rate the listing's location. When we turn to the market outcomes, we find that host's kindness has a positive effect on the listing demand. Prospective visitors are attracted by the kindness emerging from the reviews left by previous guests. Moreover, the host's kind behavior (as perceived by previous reviews) reduces the negative impact of a bad location on demand, with significant implications on a managerial perspective.

From an economic point of view, our results can improve our understanding of the functioning of the system of reviews and its impact on demand. However, implications are wider. They can shed light on the mechanisms of revelation of information in online services, where a rating inflation has been recognized (Filippas et al., 2022) and guests' comments "in their own words" may play a disciplining role. Our life is increasingly pervaded by a "digital"

–and somewhat impersonal– way of doing transactions. This work can contribute to our knowledge on the implications of personal interactions within digital environments.

Table 1: Reviews and ML-based kindness variables

Text of the review	**First-name**	**Score Pol. [-1,1]**	**Rank class**
Roland was great he have us a lot of tips of Barcelona and cautioned us on things that we may not be used to m. j he let us use anything and everything in the apartment and provided a lock for our room when we would leave for safety.	1	0.8	Excellent
I am very grateful for the availability of margarita, I have rarely been lucky enough to find a host like that !	1	0.2569	Excellent
The hosts were very helpful and responsive.	0	0.2	Positive
Ana is a great host and communicative	1	0.8	Positive
Many thanks to the host.	0	0.35	Positive
After our initial host cancelled our trip on the day, Monica was very quick to say we could stay from that night!	1	0.2708	Positive
Although we never met Eduardo because the check-in and check-out were self-made, he was always available by phone for any eventuality.	1	0.4	Positive
Regrettably, the host called me at 4,30pm on the check-out day and accused me of locking the room and not leaving the place, making him unable to accommodate the next guest and wanted compensations from me. I was reachable via ¡hidden by airbnb¿ the whole day, and so if he asked me earlier, I could have explained and helped, but instead he called at 4,30pm and accused me of lying about it.	0	-0.0933	Negative
We were extremely disappointed that Eduard refused to cancel and give us a refund due to the violent street protests.	1	-0.1159	Negative
When we turned up at 12pm (earliest time stated on post to enter room) we could not get hold of host for 30 mins and we were then told the room was not ready for us. Once it was not in contact with us he did do everything he could to help but unfortunately the lack of communication prior to the stay meant that the start of our stay wasn't as imagined.	1	-0.2	Negative
Very bad service not clear announcement.	0	-0.48	Negative

Table 2: Distribution of hosts based on the *Rank* variable

Class	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Freq. Negative	4,150	0.012	0.031	0	0.333
Freq. Neutral	4,150	0.581	0.174	0	1
Freq. Positive	4,150	0.303	0.143	0	1
Freq. Excellent	4,150	0.103	0.092	0	0.833

Host	Review	Class
Jordi	The host is superb, very helpful, from showing us map and places to visit and organizing safe taxis for us.	Excellent
Jordi	Hosts are great and make check in and check out easy.	Positive
Jordi	Everything was great!	Neutral
Jordi	Check in regularly which is appreciable.	Neutral

Rank				
Host	Frequency negative	Frequency neutral	Frequency positive	Frequency excellent
Jordi	0%	50%	25%	25%

Figure 1: Example of construction of the variable *Rank*

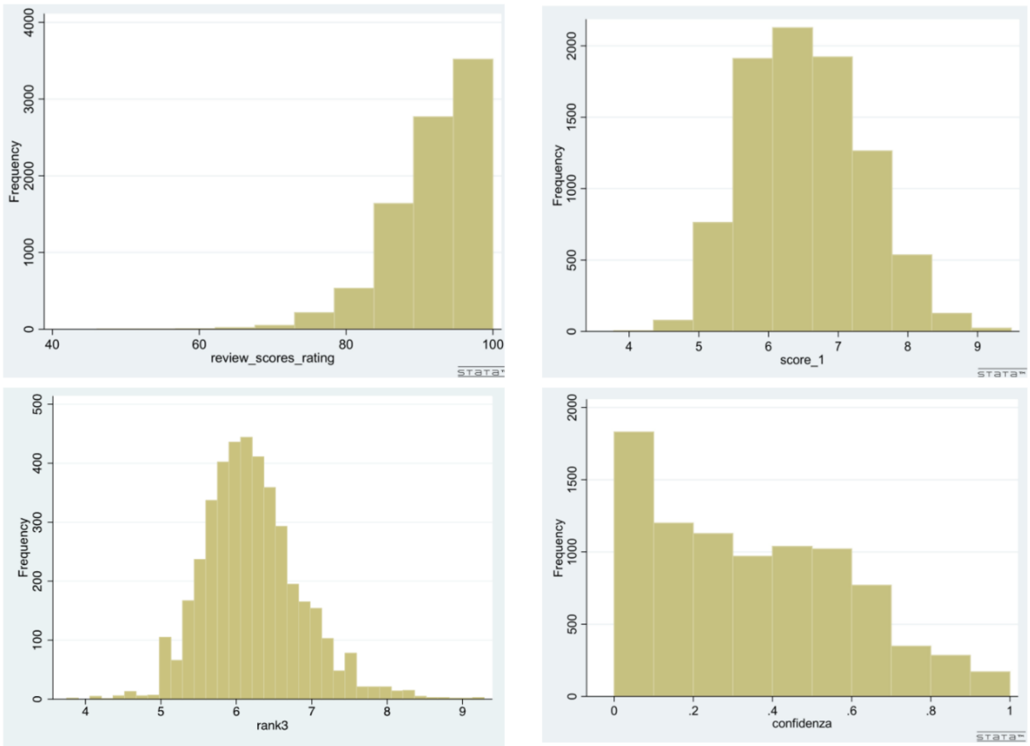

Figure 2: Distribution of ratings and kindness indicators

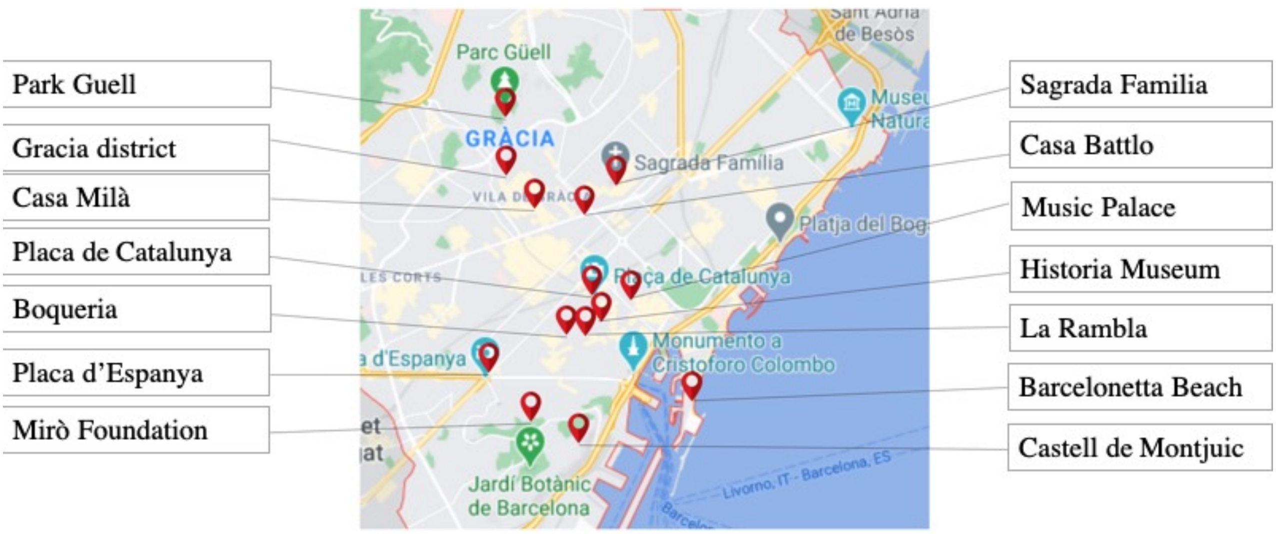

Figure 3: Map of the 14 most popular places for tourists in Barcelona

Table 3: Variables description

Variable	Description	Data source
Dependent variables		
Review score location rating	Average score of the location provided to every listing on Airbnb website	InsideAirbnb
Number of reviews	Total number of listing reviews (the demand)	InsideAirbnb
Independent variables		
Kindness (Score POLARITY)	Index of host's kindness from the machine-learning tool POLARITY	InsideAirbnb
Kindness (Rank)	Frequency-weighted Index of host's kindness from the machine-learning tool SID	
Kindness (First-name)	Percentage of reviews where the client refers directly to the host by using her/his name	InsideAirbnb
Average distance	Average distance of the listing respect to the 14 main tourist attractions in Barcelona	Google Maps
Control variables		
Price	Mean price of the listing	AirDNA
Review score rating	Average score of the listing on Airbnb website	InsideAirbnb
Host experience	The year when the host entered in Airbnb	InsideAirbnb
Host response rate	Percentage of new inquiries and reservation requests a host responds to (by either accepting/pre-approving or declining) within 24 hours	InsideAirbnb
Host acceptance rate	Percentage of accepted reservations by the host	InsideAirbnb
Verified host identity	Dummy variable equal to 1 in case the host identity is verified, 0 otherwise	InsideAirbnb
Superhost	Dummy variable equal to 1 in case the host is a superhost, 0 otherwise	InsideAirbnb
Multiproperty	Dummy variable equal to 1 in case of a multi property, 0 otherwise	InsideAirbnb
Bathrooms	Number of bathrooms of the listing	InsideAirbnb
Bedrooms	Number of bedrooms of the listing	InsideAirbnb
Number of photos	Number of photos of the apartment posted on the website	InsideAirbnb
Accommodates	Average number of accommodates in every listing	InsideAirbnb
Luxury amenities	Sum of eleven dummy variables that refer to the 11 luxury amenities, namely Beachfront, Exercise equipment, Free parking on premises, Garden, Gym, Heated floors, Jetted tab, Pool, Private pool, Sauna, Terrace	InsideAirbnb
Quality amenities	Sum of five dummy variables that refer to five quality amenities, namely the air conditioning, the washing-up machine, the washing machine, the complementary parking space, and the microwave	InsideAirbnb
Multiproperty	Dummy variable equal to 1 in case of a multi property, 0 otherwise	InsideAirbnb
Accommodates	Average number of accommodates in every listing	InsideAirbnb
Bathrooms	Number of bathrooms of the listing	InsideAirbnb
Strict cancellation	Dummy variable equal to 1 in case the cancellation is strict, 0 otherwise. According to the Airbnb Strict cancellation policy, guests may receive a full refund if they cancel within 48 hours of booking and at least 14 full days before the listing's local check-in time. After 48 hours, guests are only entitled to a 50% refund regardless of how far the check-in date is	InsideAirbnb
Guest phone number	Dummy variable equal to 1 if the guest phone is provided, 0 otherwise	InsideAirbnb
Security deposit	Dummy variable equal to 1 in case there is a security deposit, 0 otherwise	InsideAirbnb

Table 4: Summary statistics - kindness measures

	mean	sd	min	p1	p5	p10	p25	p50	p75	p90	p95	p99	max	count
Rank	6.212	0.649	3.75	5.00	5.28	5.48	5.78	6.15	6.58	7.05	7.37	8.13	9.29	4,150
POLARITY	5.844	0.503	4.06	4.95	5.12	5.26	5.49	5.79	6.15	6.49	6.70	7.29	8.45	4,150
First-name	0.276	0.254	0.00	0.00	0.00	0.00	0.07	0.21	0.43	0.65	0.81	0.98	1.00	4,150

Table 5: Summary statistics - other variables

	Mean	sd	min	max	count
Review score location rating	9.701	0.500	7.00	10.00	4,149
Number of reviews – log	2.902	0.743	1.39	5.67	4,150
Number of reviews	23.434	16.799	4.00	290.00	4,150
Average distance	2.386	0.782	1.50	7.93	4,150
Review score rating	91.280	6.097	46.00	100.00	4,150
Host experience	2016.288	2.130	2010	2020	4,150
Host response rate	0.937	0.123	0.00	1.00	4,068
Host acceptance rate	0.959	0.090	0.16	1.00	4,147
Verified host identity	0.363	0.481	0.00	1.00	4,150
Superhost	0.289	0.454	0.00	1.00	4,150
Price	139.303	109.870	16.00	1001.00	4,150
Multiproperty	0.755	0.430	0.00	1.00	4,150
Accommodates	5.211	2.161	1.00	20.00	4,150
Bathrooms	1.456	0.652	0.00	7.50	4,149
Bedrooms	2.251	1.115	0.00	9.00	4,148
Number of photos	24.303	11.667	3.00	115.00	4,143
Luxury amenities	0.146	0.430	0.00	4.00	4,081
Quality amenities	3.118	0.951	0.00	5.00	4,074
Strict cancellation	0.523	0.500	0.00	1.00	4,150
Guest phone number	0.057	0.232	0.00	1.00	4,150
Security deposit	227.966	216.505	0.00	4050.00	3,988

Table 6: Kindness, "halo effect" and the rating of location (Hypotheses 1 and 2)

Dep. Var.: Review score location rating						
Kindness is:	Score POL	Rank	First name	Score POL	Rank	First name
	(1)	(2)	(3)	(4)	(5)	(6)
Kindness	0.107***	0.0681***	0.153***	-0.143***	-0.103**	-0.329***
	(0.0166)	(0.0130)	(0.0293)	(0.0522)	(0.0439)	(0.102)
Average distance	-0.238***	-0.238***	-0.237***	-0.845***	-0.681***	-0.292***
	(0.0124)	(0.0124)	(0.0124)	(0.135)	(0.124)	(0.0186)
Kindness*Average distance				0.103***	0.0708***	0.204***
				(0.0223)	(0.0191)	(0.0451)
Control variables						
Number of reviews	0.00367***	0.00377***	0.00354***	0.00370***	0.00380***	0.00354***
	(0.000644)	(0.000652)	(0.000624)	(0.000644)	(0.000652)	(0.000628)
Price	0.000149	0.000143	0.000123	0.000149	0.000147	0.000143
	(0.000106)	(0.000109)	(0.000108)	(0.000110)	(0.000111)	(0.000110)
Superhost	0.147***	0.156***	0.164***	0.145***	0.154***	0.164***
	(0.0164)	(0.0164)	(0.0154)	(0.0165)	(0.0165)	(0.0155)
Multiproperty	-0.0807***	-0.0851***	-0.0868***	-0.0767***	-0.0826***	-0.0811***
	(0.0168)	(0.0169)	(0.0168)	(0.0167)	(0.0168)	(0.0168)
Bedrooms	0.00264	0.00309	0.00136	0.00255	0.00194	0.000195
	(0.0110)	(0.0110)	(0.0111)	(0.0109)	(0.0109)	(0.0110)
Bathrooms	0.116	0.104	0.0984	0.161	0.158	0.0600
	(0.119)	(0.118)	(0.120)	(0.123)	(0.109)	(0.151)
Accommodates	-0.0148**	-0.0145**	-0.0139**	-0.0149**	-0.0142**	-0.0139**
	(0.00587)	(0.00591)	(0.00594)	(0.00582)	(0.00586)	(0.00590)
Guest phone number	0.0411	0.0378	0.0409	0.0400	0.0383	0.0399
	(0.0274)	(0.0273)	(0.0272)	(0.0271)	(0.0271)	(0.0271)
Host response rate	-0.160**	-0.137**	-0.0926	-0.128**	-0.115*	-0.0760
	(0.0621)	(0.0612)	(0.0603)	(0.0623)	(0.0615)	(0.0604)
Host Acceptance Rate	0.0276	0.00924	-0.00908	0.000138	-0.00382	-0.00991
	(0.0876)	(0.0880)	(0.0871)	(0.0874)	(0.0892)	(0.0870)
Verified host identity	-0.000129	0.00262	0.00356	0.00392	0.00572	0.00553
	(0.0157)	(0.0157)	(0.0156)	(0.0156)	(0.0155)	(0.0155)
Strict cancellation	-0.000471	-0.000861	0.00251	-0.00113	-0.000965	0.00339
	(0.0145)	(0.0145)	(0.0145)	(0.0144)	(0.0144)	(0.0144)
Security deposit	-6.30e-06	-3.94e-06	3.05e-06	-1.14e-05	-6.95e-06	-2.43e-07
	(3.28e-05)	(3.25e-05)	(3.24e-05)	(3.40e-05)	(3.32e-05)	(3.28e-05)
Luxury amenitied	-0.00254	-0.000190	-0.00300	-0.00505	-0.00328	-0.00305
	(0.0276)	(0.0277)	(0.0279)	(0.0274)	(0.0276)	(0.0278)
Quality amenities	0.0598***	0.0573***	0.0558***	0.0614***	0.0577***	0.0530**
	(0.0217)	(0.0217)	(0.0217)	(0.0216)	(0.0216)	(0.0217)
Number of photos	0.000314	0.000261	0.000469	0.000327	0.000230	0.000567
	(0.000666)	(0.000667)	(0.000661)	(0.000657)	(0.000657)	(0.000657)
Host experience	-0.00466	-0.00555	-0.00600	-0.00307	-0.00450	-0.00550
	(0.00380)	(0.00380)	(0.00379)	(0.00379)	(0.00379)	(0.00378)
Constant	18.94**	20.96***	22.24***	17.15**	19.85***	21.36***
	(7.666)	(7.672)	(7.639)	(7.638)	(7.643)	(7.630)
H0: Kind.+(Kind.*avg.dist.)*						
*median Avg.dist.= 0						
F-Statistic (p-value)				23.78(0.00)	14.84 (0.00)	24(0.00)
H0:Avg.dist.+(Kind.*avg.dist.)*						
*median Kind.= 0						
F-Statistic (p-value)				357.31	363.43(0.00)	14.24(0.00)
R-squared	0.234	0.232	0.231	0.241	0.237	0.237
Observations	3,812	3,812	3,812	3,812	3,812	3,812

Notes. Airbnb data in 2019. Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 7: Effect of Kindness on the listing's demand (Hypothesis 3)

Dep. Var.: Number of reviews (log) Kindness is:	Score POL (1)	Rank (2)	First name (3)
Kindness	5.246***	3.307***	1.280***
	(0.410)	(0.200)	(0.143)
Kindness squared	-0.441***	-0.266***	-1.325***
	(0.0342)	(0.0158)	(0.169)
Average distance	-0.0681***	-0.0649***	-0.0704***
	(0.0135)	(0.0135)	(0.0140)
Control variables			
Review score rating	0.0231***	0.0244***	0.0226***
	(0.00220)	(0.00212)	(0.00220)
Price	-0.000645***	-0.000591***	-0.000618***
	(0.000158)	(0.000155)	(0.000166)
Superhost	0.0925***	0.110***	0.0360
	(0.0303)	(0.0301)	(0.0308)
Multiproperty	-0.214***	-0.215***	-0.170***
	(0.0272)	(0.0273)	(0.0283)
Bedrooms	-0.00835	-0.00920	-0.0170
	(0.0161)	(0.0157)	(0.0163)
Bathrooms	0.420*	0.424*	0.399*
	(0.246)	(0.235)	(0.219)
Accommodates	0.0198**	0.0164**	0.0269***
	(0.00846)	(0.00823)	(0.00840)
Guest phone number	-0.0892**	-0.0953**	-0.0976**
	(0.0440)	(0.0438)	(0.0456)
Host response rate	0.196**	0.231**	0.187*
	(0.0959)	(0.0955)	(0.0983)
Host acceptance rate	1.104***	1.110***	1.313***
	(0.164)	(0.160)	(0.164)
Verified host identity	0.0583**	0.0633***	0.0540**
	(0.0229)	(0.0227)	(0.0233)
Strict cancellation	0.0244	0.0278	0.0171
	(0.0216)	(0.0214)	(0.0222)
Security deposit	-0.000532***	-0.000543***	-0.000618***
	(6.94e-05)	(6.58e-05)	(7.20e-05)
Luxury amenities	0.0175	0.0187	0.0112
	(0.0394)	(0.0387)	(0.0396)
Quality amenities	0.0679**	0.0755**	0.0756**
	(0.0318)	(0.0322)	(0.0326)
Numer of photos	0.00603***	0.00657***	0.00586***
	(0.00102)	(0.00100)	(0.00106)
Host experience	-0.0367***	-0.0374***	-0.0350***
	(0.00549)	(0.00542)	(0.00563)
Constant	57.91***	64.49***	69.59***
	(11.21)	(11.00)	(11.37)
H0: Kindness, Kindness squared= 0 (F-statistic)	83.23	145.54	41.42
(p-value)	(0.00)	(0.00)	(0.00)
R-squared	0.265	0.278	0.230
Observations	3,813	3,813	3,813

Notes. Airbnb data in 2019. Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 8: Covid-19 lockdown in early 2020 as a natural experiment (2nd stage of Heckman's sample-selection model)

Dep. Var.: Number of reviews (2020) (log) Kindness is:	Score POL (1)	Rank (2)	First name (3)
Kindness (2019)	2.388***	1.219*	0.363
	(0.838)	(0.624)	(0.318)
Kindness (2019) squared	-0.212***	-0.105**	-0.821**
	(0.0693)	(0.0494)	(0.357)
Average distance	-0.0553*	-0.0531	-0.0621*
	(0.0330)	(0.0330)	(0.0332)
Control variables			
Review score rating	0.0910	0.102	0.0815
	(0.118)	(0.118)	(0.113)
Price	-0.000714**	-0.000709**	-0.000710**
	(0.000281)	(0.000279)	(0.000281)
Superhost	0.182**	0.185**	0.175**
	(0.0724)	(0.0723)	(0.0712)
Multiproperty	-0.0312	-0.0250	-0.00344
	(0.0704)	(0.0716)	(0.0695)
Bedrooms	-0.0944**	-0.0987**	-0.0963**
	(0.0437)	(0.0449)	(0.0425)
Bathrooms	-0.309	-0.270	-0.296
	(1.145)	(1.137)	(1.124)
Accommodates	0.0528**	0.0546**	0.0535**
	(0.0223)	(0.0225)	(0.0221)
Guest phone number	-0.451***	-0.452***	-0.469***
	(0.103)	(0.104)	(0.102)
Host response rate	-0.00565**	-0.00570**	-0.00606***
	(0.00224)	(0.00222)	(0.00222)
Verified host identity	0.188***	0.188***	0.187***
	(0.0566)	(0.0563)	(0.0558)
Security deposit	-0.000111	-0.000140	-0.000160
	(0.000138)	(0.000139)	(0.000142)
Ljuxury amenities	0.0793	0.0771	0.0672
	(0.104)	(0.103)	(0.103)
Quality amenities	0.104	0.118	0.104
	(0.0968)	(0.101)	(0.0926)
Number of photos	0.00723***	0.00757***	0.00719***
	(0.00243)	(0.00243)	(0.00244)
Host experience	-0.0210	-0.0212	-0.0180
	(0.0138)	(0.0137)	(0.0138)
Constant	37.13	40.71	37.99
	(28.11)	(27.92)	(27.76)
Wald test (all var = 0): p-value	0.00	0.00	0.00
Wald test of independent equations (ρ=0): p-value	0.785	0.891	0.561
H0: Kindness, Kindness squared= 0 (F-statistic: p-value)	0.00	0.01	0.00
Observations	3,865	3,865	3,865

Notes. Airbnb data in 2020. Maximum likelihood estimates. Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 9: First-stage of Heckman's sample selection model

Dep. Var.: Number of reviews (2020) (dummy)			
Kindness is:	Score POL	Rank	First name
	(1)	(2)	(3)
Kindness (2019)	1.555**	1.172***	0.541**
	(0.710)	(0.399)	(0.263)
Kindness (2019) squared	-0.133**	-0.0988***	-0.567*
	(0.0590)	(0.0313)	(0.302)
Average distance	-0.00460	-0.00281	-0.00636
	(0.0276)	(0.0277)	(0.0276)
Control variables			
Multiproperty	0.151***	0.147***	0.174***
	(0.0532)	(0.0531)	(0.0535)
Bedrooms	-0.135***	-0.134***	-0.139***
	(0.0325)	(0.0325)	(0.0326)
Bathrooms	0.126	0.149	0.154
	(0.572)	(0.573)	(0.581)
Accommodates	0.0541***	0.0518***	0.0566***
	(0.0172)	(0.0172)	(0.0173)
Guest phone number	-0.250***	-0.251***	-0.253***
	(0.0898)	(0.0900)	(0.0899)
Security deposit	-0.000221**	-0.000211*	-0.000251**
	(0.000112)	(0.000111)	(0.000111)
Luxury amenities	0.0219	0.0194	0.0180
	(0.0801)	(0.0800)	(0.0799)
Quality amenities	0.369***	0.373***	0.372***
	(0.0626)	(0.0626)	(0.0625)
Host experience	-0.00146	-0.00262	0.000849
	(0.0105)	(0.0105)	(0.0105)
Review score rating	0.0159***	0.0167***	0.0147***
	(0.00438)	(0.00428)	(0.00425)
Price	-0.000427	-0.000423	-0.000401
	(0.000265)	(0.000268)	(0.000264)
Host acceptance rate	2.323***	2.288***	2.408***
	(0.285)	(0.285)	(0.288)
Superhost	0.126**	0.151***	0.0952*
	(0.0579)	(0.0583)	(0.0572)
Constant	-5.644	-2.298	-5.873
	(21.56)	(21.32)	(21.22)
Selected observations	2,100	2,100	2,100
Non-selected observations	1,765	1,765	1,765
Observations	3,865	3,865	3,865

Notes. Airbnb data in 2020. Maximum likelihood estimates. Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 10: Does Kindness mitigate the negative effect of distance on the listing's demand? (Hypothesis 4)

Dep. Var.: Number of reviews (the demand) Kindness is:	Score POL (1)	Rank (2)	First name (3)
Kindness	7.664***	4.815***	2.160***
	(1.079)	(0.550)	(0.404)
Kindness squared	-0.644***	-0.387***	-2.558***
	(0.0904)	(0.0431)	(0.511)
Average distance	2.919**	1.885***	-0.0391
	(1.222)	(0.668)	(0.0251)
Kindness * average distance	-1.012**	-0.632***	-0.367**
	(0.411)	(0.209)	(0.154)
Kindness squared *average distance	0.0850**	0.0505***	0.515***
	(0.0345)	(0.0163)	(0.193)
Control variables			
Review score rating	0.0230***	0.0243***	0.0225***
	(0.00220)	(0.00212)	(0.00221)
Price	-0.000656***	-0.000595***	-0.000598***
	(0.000160)	(0.000157)	(0.000167)
Superhost	0.0917***	0.110***	0.0390
	(0.0303)	(0.0301)	(0.0308)
Multiproperty	-0.214***	-0.216***	-0.172***
	(0.0272)	(0.0273)	(0.0284)
Bedrooms	-0.00844	-0.00970	-0.0178
	(0.0160)	(0.0157)	(0.0162)
Bathrooms	0.407*	0.419*	0.415*
	(0.247)	(0.241)	(0.225)
Accommodates	0.0194**	0.0163**	0.0269***
	(0.00844)	(0.00823)	(0.00842)
Guest phone number	-0.0883**	-0.0947**	-0.0955**
	(0.0441)	(0.0439)	(0.0455)
Host response rate	0.189*	0.227**	0.172*
	(0.0968)	(0.0960)	(0.0987)
Host acceptance rate	1.105***	1.116***	1.320***
	(0.165)	(0.160)	(0.166)
Verified host identity	0.0593***	0.0650***	0.0556**
	(0.0229)	(0.0227)	(0.0234)
Strict cancellation	0.0253	0.0286	0.0175
	(0.0216)	(0.0214)	(0.0222)
Security deposit	-0.000535***	-0.000541***	-0.000617***
	(7.04e-05)	(6.59e-05)	(7.31e-05)
Luxury amenities	0.0176	0.0192	0.0144
	(0.0394)	(0.0387)	(0.0397)
Quality amenities	0.0679**	0.0751**	0.0734**
	(0.0318)	(0.0322)	(0.0326)
Number of photos	0.00607***	0.00660***	0.00580***
	(0.00102)	(0.000999)	(0.00105)
Host experience	-0.0364***	-0.0372***	-0.0347***
	(0.00550)	(0.00542)	(0.00563)
Constant	50.15***	59.31***	68.85***
	(11.73)	(11.15)	(11.35)
H0: Kindness, Kindness*average distance = 0 (F-statistic)	3.04	5.08	3.59
(p-value)	(0.048)	(0.00)	(0.028)
Observations	3,813	3,813	3,813
R-squared	0.266	0.280	0.232

Notes. Airbnb data in 2019. Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

References

Alsudais, A. (2017). Quantifying the offline interactions between hosts and guests of airbnb. Twenty-third Americas Conference on Information Systems, Boston.

Archak, N., A. Ghose, and P. G. Ipeirotis (2011). Deriving the pricing power of product features by mining consumer reviews. *Management Science 57*(8), 1485–1509.

Arcidiacono, D., M. Mainieri, and I. Pais (2016). Quando la sharing economy fa innovazione sociale. il caso blablacar. Technical report, Collaboriamo.org.

Becker, C., G. T. Bradley, and K. Zantow (2012). The underlying dimensions of tipping behavior: An exploration, confirmation, and predictive model. *International Journal of Hospitality Management 31*(1), 247–256.

Bridges, J. and C. Vásquez (2018). If nearly all airbnb reviews are positive, does that make them meaningless? *Current Issues in Tourism 21*(18), 2057–2075.

Cheng, M. and X. Jin (2019). What do airbnb users care about? an analysis of online review comments. *International Journal of Hospitality Management 76*, 58–70.

Dellarocas, C. and C. A. Wood (2008). The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management Science 54*(3), 460–476.

Ert, E. and A. Fleischer (2019). The evolution of trust in airbnb: A case of home rental. *Annals of Tourism Research 75*, 279–287.

Filippas, A., J. J. Horton, and J. M. Golden (2022). Reputation inflation. *Marketing Science*, forthcoming.

Fradkin, A., E. Grewal, and D. Holtz (2021). Reciprocity and unveiling in two-sided reputation systems: Evidence from an experiment on airbnb. *Marketing Science 40*(6), 1013–1029.

Gunter, U. and I. Önder (2018). Determinants of airbnb demand in vienna and their implications for the traditional accommodation industry. *Tourism Economics 24*(3), 270–293.

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

Hu, N., J. Zhang, and P. A. Pavlou (2009, oct). Overcoming the j-shaped distribution of product reviews. *Commununications of the ACM 52*(10), 144–147.

Lawani, A., M. R. Reed, T. Mark, and Y. Zheng (2019). Reviews and price on online platforms: Evidence from sentiment analysis of airbnb reviews in boston. *Regional Science and Urban Economics 75*, 22–34.

Leuthesser, L., S. K. Chiranjeev, and R. H. Katrin (1995). Brand equity: the halo effect measure. *European Journal of Marketing 29*, 57–66.

Liang, S., M. Schuckert, R. Law, and C.-C. Chen (2020). The importance of marketer-generated content to peer-to-peer property rental platforms: Evidence from airbnb. *International Journal of Hospitality Management 84*, 102329.

Lind, M., M. Visentini, T. Mäntylä, and F. Del Missier (2017). Choice-supportive misremembering: A new taxonomy and review. *Frontiers in Psychology 8*.

Magnani, M. (2020). The economic and behavioral consequences of online user reviews. *Journal of Economic Surveys 34*(2), 263–292.

Nicolau, J. L., J. P. Mellinas, and E. Martín-Fuentes (2020). The halo effect: A longitudinal approach. *Annals of Tourism Research 83*, 102938.

Proserpio, D., W. Xu, and G. Zervas (2018). You get what you give: theory and evidence of reciprocity in the sharing economy. *Quantitative Marketing and Economics 16*, 371–407.

Quattrone, G., D. Proserpio, D. Quercia, L. Capra, and M. Musolesi (2016). Who benefits from the "sharing" economy of airbnb? *Proceedings of the 25th International Conference on World Wide Web*, 1385–1394.

Sthapit, E. and J. Jiménez-Barreto (2018). Sharing in the host–guest relationship: perspectives on the airbnb hospitality experience. *Anatolia 29*(2), 282–284.

Tillquist, K. (2008). *Capitalizing on kindness: why 21st century professionals need to be nice*. Career Press, NJ.

Zervas, G., D. Proserpio, and J. Byers (2021). A first look at online reputation on airbnb, where every stay is above average. *Marketing Letters 32*, 1–16.

A Online Appendices

Table A.1: Correlations

N.	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	Review score location rating	1																						
2	Number of reviews	0.215	1																					
3	Score POLARITY	0.187	0.114	1																				
4	Rank	0.169	0.072	0.786	1																			
5	First-name	0.173	0.169	0.444	0.450	1																		
6	Average distance	-0.350	-0.021	0.048	0.047	0.007	1																	
7	Review score rating	0.446	0.270	0.456	0.382	0.332	-0.040	1																
8	Host experience	-0.085	-0.236	-0.204	-0.181	-0.151	-0.056	-0.130	1															
9	Host response rate	0.012	0.137	0.240	0.199	0.082	0.101	0.149	-0.100	1														
10	Host acceptance rate	-0.001	0.151	-0.119	-0.129	-0.088	-0.033	-0.037	0.082	0.065	1													
11	Verified host identity	0.063	0.150	0.205	0.190	0.164	0.049	0.148	-0.334	0.086	-0.058	1												
12	Superhost	0.212	0.222	0.439	0.411	0.352	0.029	0.538	-0.134	0.175	0.019	0.160	1											
13	Price	0.050	-0.062	-0.020	-0.012	0.030	-0.155	0.087	0.010	0.050	-0.081	-0.067	0.005	1										
14	Multiproperty	-0.107	-0.174	-0.275	-0.244	-0.194	-0.110	-0.216	0.143	-0.174	0.079	-0.093	-0.250	0.118	1									
15	Bedrooms	-0.032	0.006	-0.061	-0.073	-0.050	-0.025	-0.054	-0.059	0.047	0.006	-0.019	-0.053	0.472	0.073	1								
16	Bathrooms	0.041	0.017	-0.015	-0.011	0.001	-0.075	0.008	0.017	-0.020	-0.012	0.029	0.024	0.025	-0.007	0.031	1							
17	Guest phone number	0.043	0.022	0.029	0.040	0.009	-0.043	0.045	-0.186	0.038	-0.053	0.078	0.049	-0.030	0.020	-0.020	0.009	1						
18	Security deposit	0.019	-0.156	0.086	0.079	0.034	-0.031	0.060	0.006	0.031	-0.080	0.038	0.063	0.139	0.020	0.122	0.012	-0.013	1					
19	Number of photos	0.063	0.171	0.169	0.173	0.121	0.035	0.199	-0.169	0.143	-0.022	0.146	0.180	0.220	-0.107	0.224	-0.047	0.020	0.109	1				
20	Accommodates	-0.042	0.002	-0.091	-0.108	-0.094	-0.056	-0.072	-0.043	0.032	0.012	-0.045	-0.077	0.489	0.123	0.810	0.027	-0.021	0.142	0.230	1			
21	Strict cancellation	0.040	0.015	0.034	0.030	-0.008	-0.065	0.069	-0.057	-0.117	-0.056	-0.021	0.053	-0.038	-0.072	0.022	-0.011	0.048	0.140	0.022	0.040	1		
22	Luxury amenities	-0.001	-0.006	-0.008	-0.019	0.001	-0.026	-0.026	0.008	0.002	0.008	-0.017	-0.007	0.016	0.017	0.029	-0.003	-0.010	0.001	-0.022	0.029	-0.005	1	
23	Quality amenities	0.062	0.115	0.066	0.067	0.067	0.024	0.209	-0.042	0.154	0.094	0.022	0.158	0.174	0.014	0.160	0.026	0.029	0.055	0.135	0.193	-0.019	0.005	1

Appendix B Single rooms

Table B.2: Hypothesis 1 and 2: Single rooms

Dep. Var.: Rating Score Location Kindness is:	Score POL (1)	Rank (2)	First name (3)	Score POL (4)	Rank (5)	First name (6)
Kindness	0.0940***	0.0560***	0.181***	-0.00762	-0.0424	-0.0943
	(0.0217)	(0.0163)	(0.0433)	(0.0580)	(0.0423)	(0.120)
Average distance	-0.187***	-0.188***	-0.185***	-0.453***	-0.465***	-0.231***
	(0.0113)	(0.0113)	(0.0112)	(0.154)	(0.122)	(0.0256)
Kindness * average distance				0.0414*	0.0404**	0.116**
				(0.0238)	(0.0175)	(0.0518)
Number of reviews	0.00184***	0.00189***	0.00173***	0.00185***	0.00189***	0.00173***
	(0.000453)	(0.000464)	(0.000445)	(0.000453)	(0.000463)	(0.000443)
Price	0.000834**	0.000773**	0.000698**	0.000803**	0.000746**	0.000709**
	(0.000334)	(0.000331)	(0.000319)	(0.000334)	(0.000331)	(0.000318)
Superhost	0.154***	0.165***	0.167***	0.154***	0.163***	0.168***
	(0.0194)	(0.0193)	(0.0180)	(0.0193)	(0.0192)	(0.0179)
Multiproperty	-0.0208	-0.0259	-0.0224	-0.0224	-0.0268	-0.0226
	(0.0183)	(0.0183)	(0.0183)	(0.0183)	(0.0183)	(0.0183)
Bedrooms	0.0776***	0.0803***	0.0756***	0.0764***	0.0780***	0.0726***
	(0.0255)	(0.0263)	(0.0258)	(0.0257)	(0.0266)	(0.0260)
Bathrooms	-0.0655	-0.0578	-0.0791	-0.0712	-0.0706	-0.0746
	(0.110)	(0.112)	(0.111)	(0.108)	(0.111)	(0.108)
Accommodates	-0.0752***	-0.0765***	-0.0741***	-0.0752***	-0.0761***	-0.0727***
	(0.0112)	(0.0111)	(0.0112)	(0.0112)	(0.0112)	(0.0113)
Guest phone number	0.0159	0.0142	0.0197	0.0140	0.0115	0.0111
	(0.0403)	(0.0407)	(0.0409)	(0.0400)	(0.0406)	(0.0408)
Host response rate	0.0362	0.0374	0.0253	0.0400	0.0484	0.0336
	(0.0661)	(0.0663)	(0.0657)	(0.0659)	(0.0662)	(0.0656)
Host acceptance rate	-0.165**	-0.190***	-0.184***	-0.169**	-0.194***	-0.188***
	(0.0696)	(0.0695)	(0.0694)	(0.0695)	(0.0689)	(0.0693)
Verified host identity	-0.0355*	-0.0345*	-0.0377**	-0.0360*	-0.0339*	-0.0378**
	(0.0187)	(0.0188)	(0.0188)	(0.0187)	(0.0187)	(0.0187)
Strict cancellation	-0.00576	-0.00565	-0.00200	-0.00574	-0.00718	-0.00465
	(0.0175)	(0.0175)	(0.0175)	(0.0175)	(0.0175)	(0.0174)
Security deposit	6.54e-05**	7.01e-05**	7.53e-05**	6.60e-05**	7.51e-05**	7.79e-05**
	(3.20e-05)	(3.26e-05)	(3.29e-05)	(3.18e-05)	(3.34e-05)	(3.35e-05)
Luxury amenities	0.0199	0.0206	0.0219	0.0162	0.0194	0.0182
	(0.0313)	(0.0313)	(0.0309)	(0.0314)	(0.0315)	(0.0307)
Quality amenities	0.0146	0.0151	0.0180	0.0153	0.0151	0.0176
	(0.0179)	(0.0179)	(0.0178)	(0.0179)	(0.0179)	(0.0178)
Number of photos	-0.000173	-0.000338	-0.000276	-0.000119	-0.000239	-0.000308
	(0.000861)	(0.000861)	(0.000859)	(0.000862)	(0.000860)	(0.000857)
Constant	9.753***	9.992***	10.32***	10.41***	10.68***	10.42***
	(0.198)	(0.180)	(0.134)	(0.399)	(0.321)	(0.140)
Observations	2,111	2,111	2,111	2,111	2,111	2,111
R-squared	0.259	0.255	0.257	0.262	0.259	0.260

Notes. Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table B.3: Hypothesis 3: Single rooms

Kindness is:	Score POL (1)	Rank (2)	First name (3)
Kindness	4.394***	2.422***	1.647***
	(0.637)	(0.315)	(0.205)
Kindness squared	-0.348***	-0.186***	-2.022***
	(0.0498)	(0.0226)	(0.225)
Average distance	-0.124***	-0.121***	-0.130***
	(0.0139)	(0.0140)	(0.0140)
Review scores rating	0.0146***	0.0178***	0.0154***
	(0.00343)	(0.00337)	(0.00332)
Price	-0.00260***	-0.00266***	-0.00236***
	(0.000475)	(0.000462)	(0.000474)
Superhost	0.0940**	0.121***	0.0876**
	(0.0368)	(0.0365)	(0.0366)
Multiproperty	-0.0654**	-0.0862***	-0.0676**
	(0.0301)	(0.0301)	(0.0303)
Bedrooms	-0.110***	-0.120***	-0.117***
	(0.0406)	(0.0424)	(0.0417)
Bathrooms	0.235	0.174	0.0925
	(0.192)	(0.200)	(0.168)
Accommodates	0.0757***	0.0688***	0.0729***
	(0.0216)	(0.0214)	(0.0214)
Guest phone number	-0.0925	-0.107	-0.0695
	(0.0777)	(0.0779)	(0.0782)
Host response rate	0.235**	0.223**	0.245**
	(0.105)	(0.111)	(0.111)
Host acceptance rate	1.714***	1.691***	1.809***
	(0.127)	(0.125)	(0.123)
Verified host identity	-0.0156	-0.0176	-0.0239
	(0.0293)	(0.0295)	(0.0300)
Strict cancellation	-0.0305	-0.0210	-0.0308
	(0.0290)	(0.0292)	(0.0296)
Security deposit	-5.97e-05	-6.05e-05	-7.25e-05
	(0.000118)	(0.000118)	(0.000123)
Luxury amenities	-0.0613	-0.0632	-0.0946*
	(0.0533)	(0.0533)	(0.0562)
Quality amenities	0.0404	0.0310	0.0288
	(0.0296)	(0.0298)	(0.0298)
Number of photos	0.00407***	0.00526***	0.00497***
	(0.00134)	(0.00136)	(0.00137)
Constant	-13.81***	-8.030***	-0.282
	(2.034)	(1.079)	(0.376)
Observations	2,112	2,112	2,112
R-squared	0.221	0.216	0.203

Notes. Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table B.4: Hypothesis 4: Single rooms

Kindness is:	Score POL (1)	Rank (2)	First name (3)
Kindness	5.648***	2.015**	1.662***
	(1.419)	(0.808)	(0.509)
Kindness squared	-0.448***	-0.155***	-2.058***
	(0.111)	(0.0592)	(0.585)
Average distance	1.370	0.174	-0.131***
	(1.439)	(1.019)	(0.0384)
Kindness * average distance	-0.480	0.174	-0.00627
	(0.449)	(0.301)	(0.185)
Kindness squared * average distance	0.0383	-0.0131	0.0152
	(0.0348)	(0.0221)	(0.216)
review scores rating	0.0148***	0.0176***	0.0154***
	(0.00342)	(0.00341)	(0.00332)
Price	-0.00261***	-0.00264***	-0.00235***
	(0.000476)	(0.000462)	(0.000477)
Superhost	0.0933**	0.122***	0.0876**
	(0.0367)	(0.0365)	(0.0366)
Multiproperty	-0.0678**	-0.0858***	-0.0676**
	(0.0301)	(0.0301)	(0.0304)
Bedrooms	-0.113***	-0.120***	-0.117***
	(0.0400)	(0.0426)	(0.0418)
Bathrooms	0.234	0.176	0.0923
	(0.194)	(0.199)	(0.168)
Accommodates	0.0772***	0.0684***	0.0729***
	(0.0217)	(0.0214)	(0.0215)
Guest phone number	-0.0924	-0.106	-0.0699
	(0.0780)	(0.0779)	(0.0784)
Host response rate	0.234**	0.221**	0.245**
	(0.105)	(0.111)	(0.112)
Host acceptance rate	1.708***	1.691***	1.808***
	(0.126)	(0.126)	(0.123)
Verified host identity	-0.0151	-0.0181	-0.0238
	(0.0294)	(0.0296)	(0.0300)
Strict cancellation	-0.0296	-0.0207	-0.0309
	(0.0290)	(0.0292)	(0.0297)
security deposit	-5.74e-05	-6.15e-05	-7.24e-05
	(0.000117)	(0.000118)	(0.000123)
Luxury amenities	-0.0615	-0.0633	-0.0947*
	(0.0534)	(0.0534)	(0.0563)
Quality amenities	0.0410	0.0318	0.0288
	(0.0296)	(0.0298)	(0.0298)
Number of photos	0.00409***	0.00526***	0.00497***
	(0.00134)	(0.00137)	(0.00137)
Constant	-17.71***	-6.678**	-0.280
	(4.524)	(2.787)	(0.389)
Observations	2,112	2,112	2,112
R-squared	0.221	0.216	0.203

Notes. Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Appendix C Cubic demand specification

Table C.5: Cubic demand specification

Dep. Var.: Number of reviews						
Kindness is:	Score POL	Rank	First name	Score POL	Rank	First name
	(1)	(2)	(3)	(4)	(5)	(6)
Kindness	20.93***	8.236***	2.528***	53.29***	8.010*	4.409***
	(4.107)	(1.355)	(0.275)	(10.24)	(4.589)	(0.857)
Kindness squared	-2.998***	-1.040***	-5.426***	-8.109***	-0.897	-9.872***
	(0.663)	(0.211)	(0.826)	(1.669)	(0.708)	(2.561)
Kindness cube	0.138***	0.0400***	3.145***	0.404***	0.0267	5.577***
	(0.0354)	(0.0108)	(0.641)	(0.0901)	(0.0361)	(1.980)
Average distance	-0.0660***	-0.0640***	-0.0714***	28.19***	0.369	-0.0230
	(0.0135)	(0.0135)	(0.0139)	(8.781)	(3.851)	(0.0278)
Avg dist. * Kindness				-13.54***	0.0666	-0.783**
				(4.313)	(1.790)	(0.329)
Avg dist. *Kindness sq.				2.140***	-0.0556	1.849*
				(0.702)	(0.275)	(0.967)
Avg dist. *Kindness cube				-0.111***	0.00532	-1.009
				(0.0378)	(0.0139)	(0.740)
Review scores rating	0.0223***	0.0241***	0.0228***	0.0223***	0.0240***	0.0227***
	(0.00219)	(0.00211)	(0.00219)	(0.00219)	(0.00211)	(0.00220)
Price	-0.000631***	-0.000593***	-0.000643***	-0.000655***	-0.000596***	-0.000634***
	(0.000154)	(0.000156)	(0.000169)	(0.000157)	(0.000158)	(0.000171)
Superhost	0.116***	0.126***	0.0562*	0.115***	0.126***	0.0585*
	(0.0308)	(0.0304)	(0.0309)	(0.0307)	(0.0304)	(0.0309)
Other Control var.	Yes	Yes	Yes	Yes	Yes	Yes
F-Stat (p-value)						
Joint sign.						
All kindness terms	75.2(0.00)	113.3(0.00)	36.7(0.00)	25.6(0.00)	27.8(0.00)	3.40(0.02)
Joint sign.						
Square and cubic terms	108.06(0.00)	160.2 (0.00)	49.8(00)	4.55(0.00)	3.53(0.01)	13.34(0.00)
Observations	3,813	3,813	3,813			
R-squared	0.270	0.281	0.236			

Notes. Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

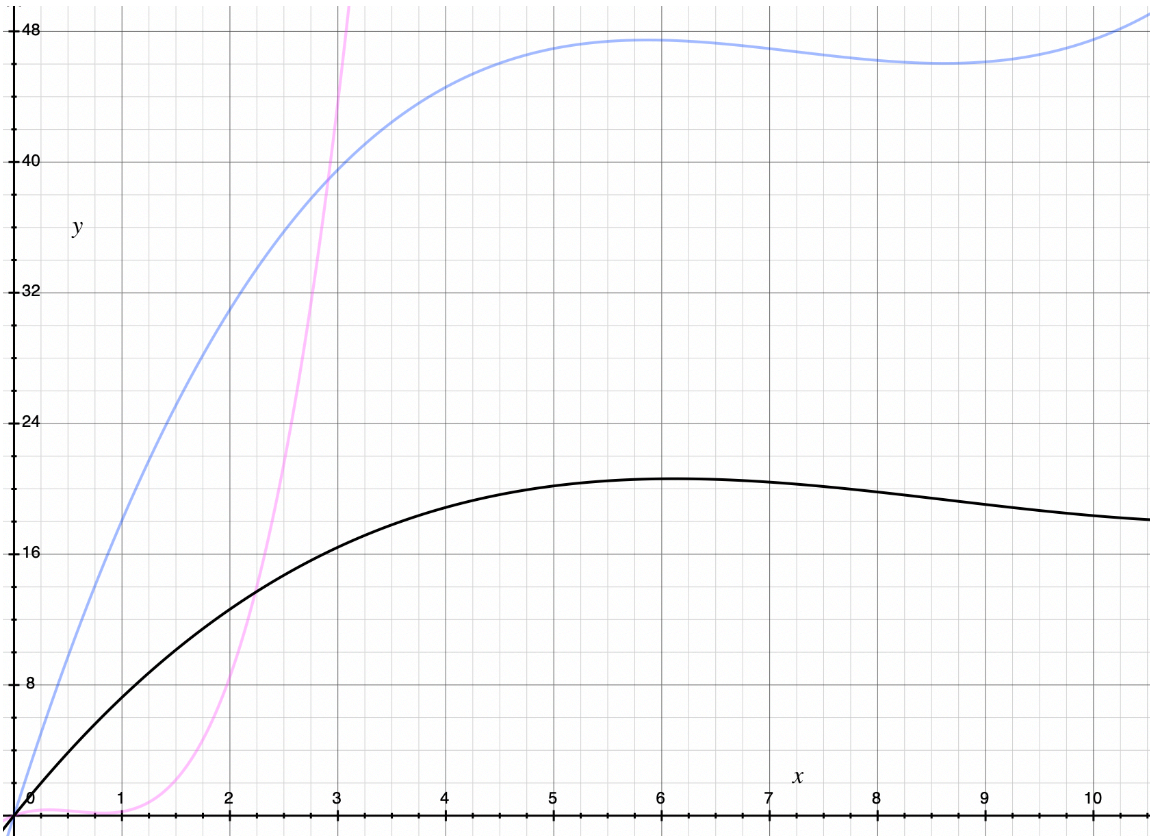

Figure C.1: Number of reviews (on the vertical axis) in function of kindness (on the horizontal axis). Kindness is: Rank (black), Score POL (blue), First Name (pink). Note: coefficients from Table C.5, columns (1), (2) and (3).

Appendix D Figures

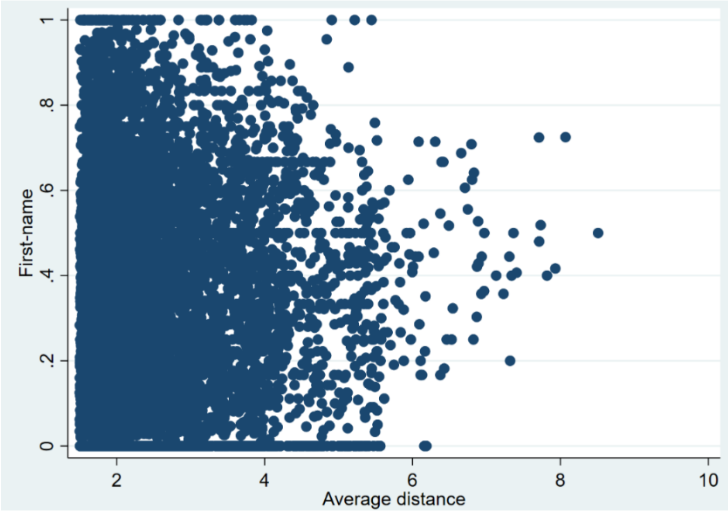

Figure D.2: Scatterplot of First-name according to average distance

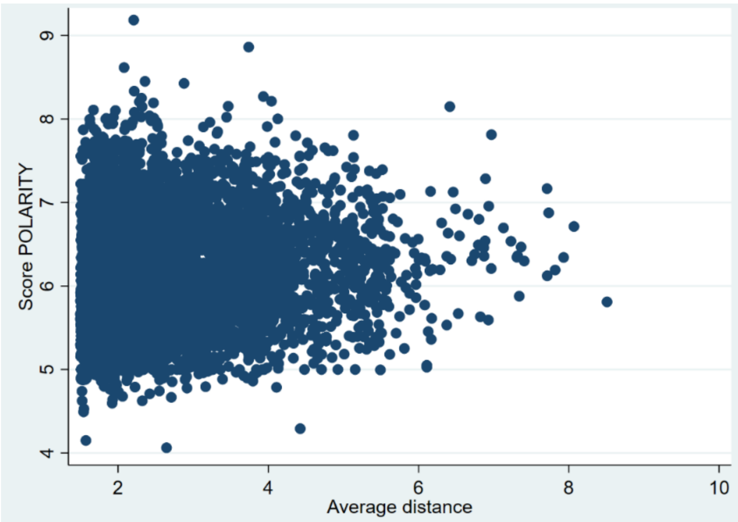

Figure D.3: Scatterplot of Score Polarity according to average distance

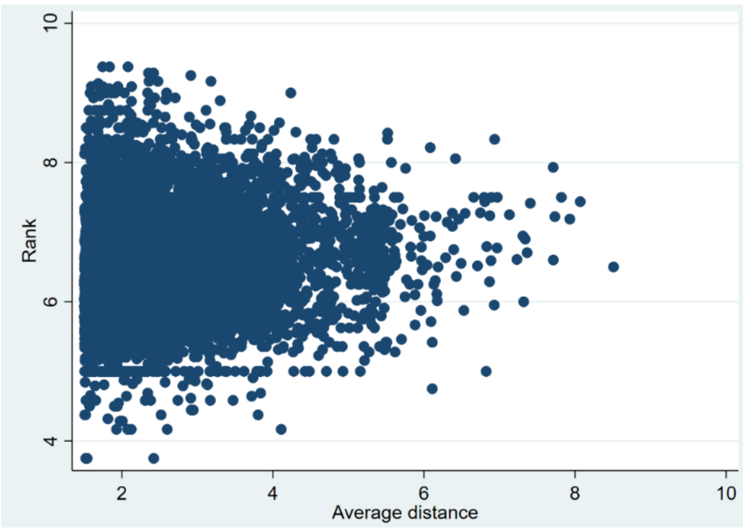

Figure D.4: Scatterplot of Rank according to average distance

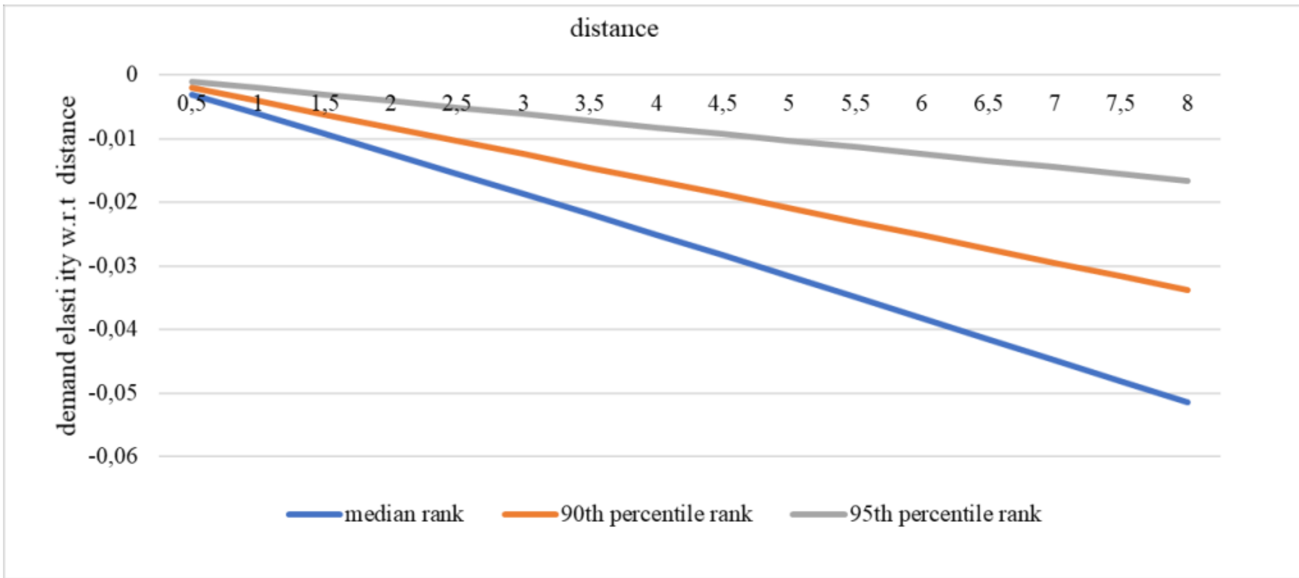

Figure D.5: Marginal effect of average distance on listing demand moderated by kindness - Rank. Note: Elasticity of demand (on the vertical axis) in function of the average distance (on the horizontal axis), for different levels of Rank.

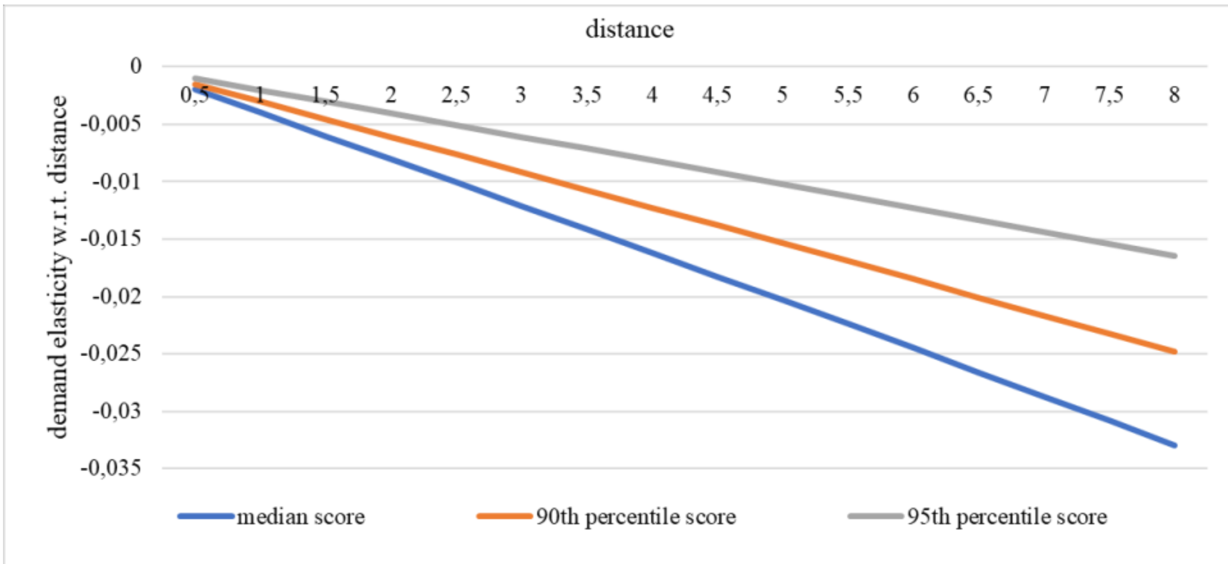

Figure D.6: Marginal effect of average distance on listing demand moderated by kindness - Score Polarity. Note: Elasticity of demand (on the vertical axis) in function of the average distance (on the horizontal axis), for different levels of Score Polarity.

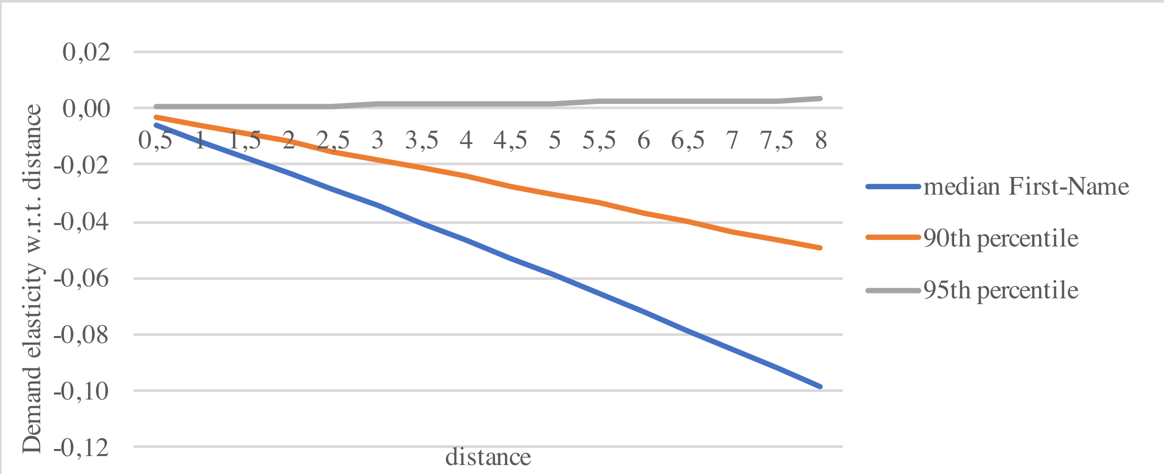

Figure D.7: Marginal effect of average distance on listing demand moderated by kindness - First-Name. Note: Elasticity of demand (on the vertical axis) in function of the average distance (on the horizontal axis), for different levels of First-Name.