UNIVERSITY OF CAMBRIDGE
Faculty of Economics

CAMBRIDGE WORKING PAPERS IN ECONOMICS

The Impact of Qualitative Reviews on Racial Statistical Discrimination: Evidence from Airbnb

James
Morris

Abstract

This paper analyses the effect of qualitative reviews on racial statistical discrimination. Using a fine-tuned Bidirectional Encoder Representations from Transformers (BERT) language model that was developed specifically for this task, I include the effects of recent qualitative reviews on the log listing price difference between Black and White hosts on Airbnb. For properties without guest reviews, I find a 4% log listing price difference between Black and White hosts for comparable properties. Once review information becomes available, this pricing difference reduces to 1%, providing evidence against the persistence of racial listing price differences on Airbnb, and furthermore, suggesting that race is used as the primary signal of property quality only in the absence of better information. Beyond its applications within the context of Airbnb, this paper aims to explain how the early provision of detailed qualitative information can reduce the effects of statistical discrimination against minorities.

Reference Details

CWPE	2331
Published	22 March 2023
Key Words	
JEL Codes	
Website	www.econ.cam.ac.uk/cwpe

The Impact of Qualitative Reviews on Racial Statistical Discrimination: Evidence from Airbnb

James Morris

March 22, 2023

Abstract

This paper analyses the effect of qualitative reviews on racial statistical discrimination. Using a fine-tuned Bidirectional Encoder Representations from Transformers (BERT) language model that was developed specifically for this task, I include the effects of recent qualitative reviews on the log listing price difference between Black and White hosts on Airbnb. For properties without guest reviews, I find a 4% log listing price difference between Black and White hosts for comparable properties. Once review information becomes available, this pricing difference reduces to 1%, providing evidence against the persistence of racial listing price differences on Airbnb, and furthermore, suggesting that race is used as the primary signal of property quality only in the absence of better information. Beyond its applications within the context of Airbnb, this paper aims to explain how the early provision of detailed qualitative information can reduce the effects of statistical discrimination against minorities.

1. Introduction

It has long been established that ethnic minorities in Europe and the United States achieve worse economic outcomes than their ethnic majority peers (Altonji & Pierret, 2001; Lang & Manove, 2011; Bertrand & Mullainathan, 2004, 2017). Since Becker (1957), numerous studies have sought to explain the nature and source of these differences (see Lang & Lehmann (2012)).

This paper studies the extent to which the difference in prices posted by Black and White hosts on Airbnb can be explained by observable and unobservable[1] property quality. I use machine learning techniques to estimate the signal received from each property's qualitative reviews, and exploit these estimates to form a signal of current property quality. When taken alongside other objects to estimate the unobservable quality of a property, this signal provides an explanation of 75% of the previously unexplained difference.

With the growing importance of both online marketplaces and peer to peer sharing websites such as Amazon, eBay, Uber, and Airbnb, the racial discrimination literature has expanded to study these environments. Doleac & Stein (2013) find evidence of racial discrimination in response rates to online advertisements, Edelman & Luca (2014) uncover evidence of racial discrimination in responses to property listings on Airbnb, and Edelman, Luca, & Svirsky (2017) study the role of racial discrimination in guest acceptance rates on the same platform.

Airbnb is an online platform that allows individuals ("hosts") to offer customers ("guests") accommodation and experiences. The platform has a large share of the online rental market, with over 300 million nights and experiences booked worldwide during 2019(Airbnb, 2021). Prior to booking, private interactions between potential hosts and guests are prohibited. As a result, only information visible on an individual's profile can be used to make application and acceptance decisions.

In a similar manner to other online market places, Airbnb provides detailed information including accommodation facilities, personal information, photographs, and online reviews in order to reduce informational asymmetries. However, as detailed in Edelman, Luca, & Svirsky (2017), this can have the unintended consequence of facilitating discrimination against minority groups.

There is an expanding literature that studies racial differences on Airbnb outcomes Cui, Li, & Zhang (2020); Edelman & Luca (2014); Edelman, Luca, & Svirsky (2017); Laouénan &

[1]Not all components of property quality are observable at the time of booking. Unobservable components include: host attitude, property wear and tear, and noise in and around the property. Potential guests form expectations about these unobservable elements based on pictures, reviews, etc.

Rathelot (2022). The prevailing approach analyses differences in listing prices across racial groups while controlling for host and property attributes. The majority of these variables are fixed or change infrequently i.e. the number of bedrooms, location, amenities, and review average, and therefore, the analysis ignores an important dimension of the data, namely, how beliefs inferred about a property, and as an indirect consequence, prices, change based on recent experiences.

On Airbnb, qualitative reviews of recent guests are visible on the platform with the corresponding quantitative component omitted. Consequently, past econometric analysis has analysed the guest experiences using the average quantitative score, and ignored the qualitative component. This has resulted in the omission of recent performance in the modelled decision making process.

I study the importance of both quantitative and qualitative reviews on Airbnb property listing prices utilising techniques from the machine learning literature. I illustrate that a potential guest's beliefs about a property are influenced by the host's race, and that this has a material impact on listing prices where review information is unavailable. Furthermore, once review information is fully available,[2] the impact of race on price falls by 75%.

Extending the theoretical model from Laouénan & Rathelot (2022) to include the information from qualitative reviews, I study these effects in a theoretical environment where hosts set their price to maximise income from Airbnb and other work, and where guests deduce a signal of the property quality from the information provided. The qualitative review component potentially represents an important element of this signal which until now has been ignored. My model incorporates this information into the decision, ultimately, resulting in the conclusion that statistical racial discrimination is not persistent.

My paper contributes to the literature in two key ways. Firstly, I fine-tune a Bidirectional Encoder Representations from Transformers (BERT) language model, and use this to analyse the impact of recent reviews on pricing decisions. Generally, in the Airbnb literature the impacts of average customer reviews are understood, while the impact of one-off qualitative reviews are not. This addition aligns the economic analysis with the real-world decision making process, which I find to have a material impact on reducing the persistence of statistical racial discrimination.

Secondly, analysing the impact of recent reviews allows me to revisit the analysis of statistical discrimination in the Airbnb literature. Contrary to the current literature, my results show that a host's race has a material impact on listing prices only whilst review information is unavailable. Upon the release of this information, the listing price difference

[2]It is Airbnb policy that historical guests quantitative reviews only become visible to potential guests after the host has completed three guest stays.(Airbnb - Why Reviews Matter , 2019)

between the two racial groups is reduced by 75%. This suggests that a large proportion of what has previously been considered statistical racial discrimination, is in fact the effects of differences in historical pricing decisions and recent review quality.

The results from this paper motivate further research studying the impact of racial discrimination in online marketplaces. For this purpose, I am developing a model allowing for the effects of historical prices, and qualitative reviews on statistical racial discrimination, which I aim to use to study a wider number of settings.

The data used in this study is taken from Inside Airbnb, an organisation that obtains the data directly from the Airbnb website using automated web-scraping techniques. A detailed description of the data is provided in Section (4), with an example of a host's profile reproduced in Appendix (F). I study ten US and UK cities in the second and third quarters of 2022. I obtain the racial information of the individual by scraping host pictures directly from the Airbnb website and applying picture recognition technology. For the individuals that contribute to this study, namely those defined by the software as Black or White, I achieve a recall[3] of c.89%.

Building on the BERT model, I estimate the sentiment of qualitative reviews. I develop two models, the first classifies reviews into 3 categories: 1 to 3 star; 4 or 5 star; and automated response upon cancellation. Whereas, the second model divides the 4 and 5 star reviews into two separate categories. The models achieve accuracies of 95.4% and 85.9%, respectively. The estimates from the fine tuned models are then used in the analysis of racial listing price differences using a combination of non-linear least-squares and simple linear regression techniques.

2. Related Literature

A recent literature studies the impact of racial discrimination on "sharing-website" outcomes. Edelman & Luca (2014) use external experts to provide assessments of the quality of Airbnb listings and the racial group of hosts. The paper finds suggestive evidence of racial discrimination on Airbnb against Black hosts. Doleac & Stein (2013) list i-Pods for sale on online markets, where the picture of the product contains a dark or light skinned hand. Similarly, the paper finds evidence suggesting statistical racial discrimination against Black sellers. My paper extends the understanding of racial discrimination on sharing websites by

[3] Where recall is defined as:

$$Recall = \frac{tp}{tp + fn},$$

where tp is the number of true positives, and fn, the number of false negatives.

estimating the impact of qualitative reviews on Airbnb outcomes which, to the best of my knowledge, has not been done previously.

The use of customer generated reviews is prevalent on sharing sites, and as such, the racial discrimination literature has focussed on the impact of these reviews on market outcomes. Cui, Li, & Zhang (2020) finds that a positive review reduces the likelihood that White guests will be accepted over African Americans. In Laouénan & Rathelot (2022), potential guests form expectations about the quality of a property based on past reviews and the ethnicity of the host. The paper finds that statistical discrimination accounts for the entirety of the listing price differences between White and ethnic minority hosts. My work extends the model of Laouénan & Rathelot (2022) by considering the impact of recent textual reviews on outcomes, this allows me to test if beliefs about racial groups persist in the presence of more detailed information.

In the literature of racial discrimination on Airbnb, work has focussed on establishing evidence of host discrimination towards potential guests (Cui, Li, & Zhang, 2020; Kakar, et al., 2018), and the impact of discrimination against hosts on listing prices (Edelman & Luca, 2014; Edelman, Luca, & Svirsky, 2017; Laouénan & Rathelot, 2022). My paper sits within this second literature whilst extending it to align with real world decision making processes, by including the impacts of qualitative reviews on rental decisions. Furthermore, it corrects an oversight in the current literature which generally fails to recognise the impact of previous pricing decisions on guest ratings. My paper adjusts for this omission by modelling overall rating as a signal of quality and previous pricing decisions.

There is a growing literature that uses the information from text to derive economic insight. Wu (2018) applies these techniques to highlight animus against women in an academic forum, Ash, Chen, & Ornaghi (2021) analyses attitudes towards gender in the US judiciary, and Stephens-Davidowitz (2014) considers the impact of racial animus against Black political candidates. I use the BERT model, developed in Devlin et al. (2018) to estimate the signal received from qualitative reviews where the corresponding quantitative value is unknown. This allows me to consider the impact of recent reviews and experiences on outcomes which otherwise would be infeasible.

In another appeal to the methodology of different literatures, I exploit machine learning techniques, and in particular, the picture recognition software, *DeepFace*, introduced in Taigman et al. (2014), to analyse the racial identity of hosts. The use of this technological solution allows me to accurately analyse a large number of photographs without relying on manual categorisation procedures.

Finally, several recent papers have attempted to classify the type of discrimination suffered on sharing websites. Bohren, Imas, & Rosenberg (2019) analyses the presence of gender based discrimination on an online mathematics question and answer forum. Laouénan & Rathelot (2022) models racial discrimination by potential guests against hosts on Airbnb. I contribute to this literature by testing for evidence of statistical discrimination in the presence of the additional considerations outlined above.

The remainder of the paper is organized as follows. Section (3) provides an overview of the Airbnb platform, Section (4) describes the data, Section (5) describes the model, Section (6) presents my empirical strategy, and Section (7) concludes.

3. An Overview of Airbnb

Airbnb is an online platform that allows individuals ("hosts") to offer customers ("guests") accommodation and experiences. It began operating in 2007, and currently operates in more than 220 countries and regions throughout the world (Airbnb, 2021). During 2019 (the latest year available in the filed financial accounts), there were 300.6 million nights and experiences booked worldwide. Airbnb's activity is focused on a number of key areas, with 77% of the total bookings in North America (USA, Canada, and Mexico), Europe, The Middle East, and Africa (Airbnb, 2021).

When using the platform for the first time, hosts and guests must provide personal information including their name and contact details, and they should verify their identity.[4] Additionally, hosts are required to provide a large amount of information about their property, including: the property type (house, flat, etc.); location; guest capacity; amenities; and at least five photographs. A personal photograph is obligatory for hosts, but only encouraged for guests. Finally, both hosts and guests are encouraged to provide a short profile.

After registering their property, hosts choose: the dates to make the property available; the price; the cancellation policy; and any long-stay discounts. The cancellation policy ranges from 'Flexible', where any payment is fully refundable for cancellations made up to 24 hours prior to check-in, and 'Super Strict-60 Days', where 50% of the cost of the rental is refundable if cancellation is made up to 60 days prior to check-in, and non-refundable thereafter. Airbnb offers advice on several of the adjustable components of the terms of rental, including providing the mean rental price for similar properties in the area over the previous three months.

Potential guests visit the Airbnb website and can perform a search for available properties based on location, dates, and the number of guests. A list of search results is then visible that provides key details of each property including a short description, several photographs, the

[4] Hosts and guests can verify their identity via a scanned copy of a qualifying photographic identification or by confirming historical identity information.

average review score of previous stays, and the price for the selected stay.[5] Once a specific property is selected, users can see a more detailed profile including: the host's name; list of amenities; qualitative reviews; and cancellation policy.

Finally, if a property is reserved by a potential guest, the host receives a request which includes the guest's information. This request can be accepted or rejected by the host without penalty. However, if a host cancels a request after acceptance, they are subject to financial penalties, and an automated negative review will appear on their profile. Cancellation by the guest is subject to the host's cancellation policy.

The mean quantitative reviews associated with a host's Airbnb profile are the average of reviews provided by past guests in respect of seven characteristics of a guest's stay: overall; cleanliness; accuracy; communication; location; check-in; and value for money. The qualitative reviews provided by historical guests are also visible. There is no standard format for these reviews, and as a result, review length and detail vary significantly between a one-word review, and a detailed review of many aspects of an individual's stay.

4. Data

I use the data on Inside Airbnb, a free to use website that provides detailed information on Airbnb listings. This information has been used in several published academic papers, including Guo, Barnes, & Qiong (2017). The data is scraped every three months, and provides details of the Airbnb properties that are visible to users of the website at that time. For the purposes of the analysis in this paper, I use the data for the second and third quarters of 2022.

All information on the first page of a qualifying host's profile is provided,[6] in addition to the associated *Review* and *Amenities* page for each property. In order to filter out inactive listings, I discard all properties that have not received a review over the observation period.

I obtain the profile pictures of each entry directly from Airbnb using web-scraping techniques similar to those used in Koffi (2021) and Laouénan & Rathelot (2022). The race of the individuals in the profile pictures are then analysed using picture recognition technology as discussed in the subsequent section.

Generally the properties in the sample are small and on average consist of 1.6 bedrooms, accommodate 3.5 people, and have 1.4 bathrooms. 71.7% of listings are for the entire property, compared to 27.4% of private rooms, and 0.5% shared rooms. A more detailed

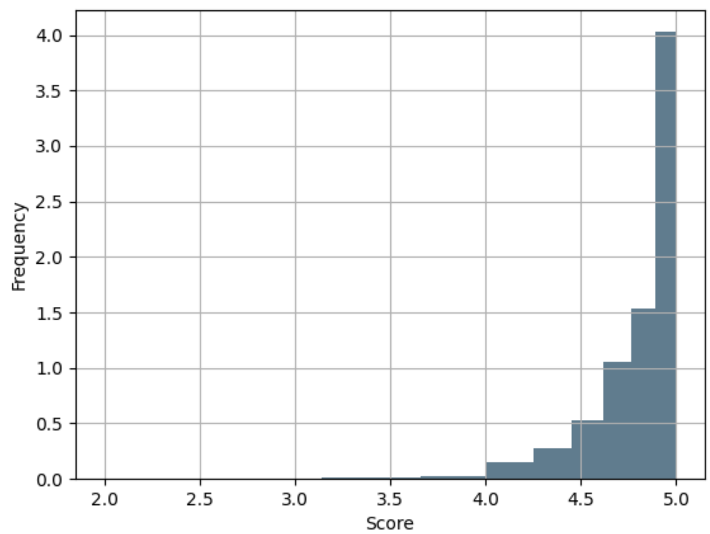
[5] The results list can be filtered based on a large number of property characteristics. Further information is provided in Appendix (F).

[6] An example can be found in Appendix (F).

description of the properties, the hosts, and the property location are provided in Tables (8), (9), and (10) of Appendix (D).

Figure (1) presents the average rating for properties within my sample. The average rating is highly skewed with a mean rating of 4.74, median rating of 4.86, and standard deviation of 0.39.[7] The high average rating reflects the fact that poorly rated properties are more likely to leave the market than higher rated properties. This can clearly be seen in Figure (5) in Appendix (E).

Figure 1: Distribution of Average Ratings on Airbnb



Notes:

(1) This figure illustrates the distribution of the average review rating in the data.

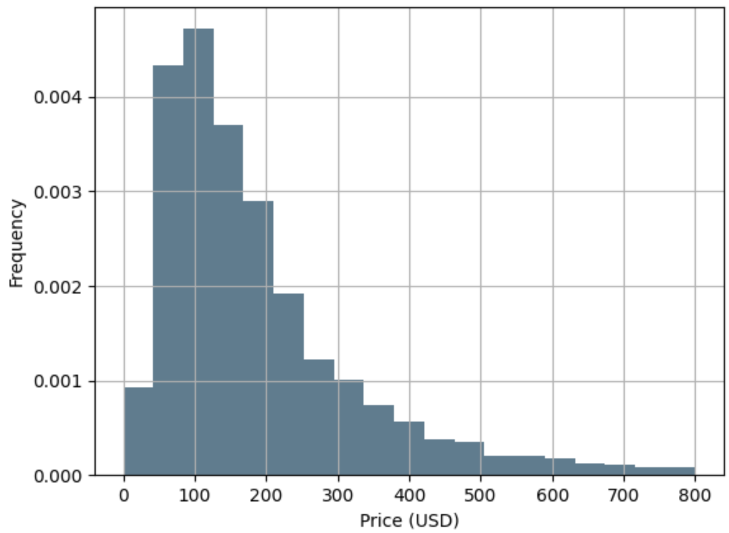
(2) The ratings are grouped into bins based on the exponential of each rating. There are 8 equally distributed bins on $[e^1, e^5]$.

Figure (2) illustrates the distribution of nightly prices on Airbnb in US Dollars (trimmed at \$800). The distribution is highly skewed. After dropping the top and bottom 2% of prices, the median price is \$150 and mean price is \$197.

Differences exist between the listing prices of properties listed by Black and White hosts. Specifically, the mean listing price for a Black host of \$144 is considerably lower than for a White host of \$218. These differences embed a number of factors including property choices, location, and available amenities. After controlling for these effects, the log listing price of a Black host remains 4% below that of a White host. A summary of this analysis can be found in Table (1).

[7]The rating scale on Airbnb is between a minimum of 1, and a maximum of 5.

Figure 2: Distribution of Prices on Airbnb



Notes:

(1) This figure provides a graphical illustration of the distribution of per night property prices in US Dollars.

(2) Properties valued at greater than $800 are not shown.

(3) Where the original price is in GBP, the price is converted to USD at the listing date's spot rate.

The results highlight a number of interesting data features that I explore in the following sections. In particular, racial differences are negatively correlated with both quantitative and qualitative review quality, and a recent qualitative review of 4 star or less has a statistically significant impact on listing price. These two observations provide motivation for the extension of the original framework provided in Laouénan & Rathelot (2022).

4.1. Analysis of Race

The race of a host is unknown to potential guests at the time of booking. As such, any beliefs about race will be formed using information obtained directly from the host's page including host name, profile pictures, and past reviews. One benefit of the platform for the study of racially sensitive outcomes is that potential guests form beliefs about the race of the host using the same dataset as the researcher.

There is a large and well-established literature dedicated to both understanding the process of inferring race, and the consequences of these racial assumptions on outcomes. Three popular techniques are: names as a signal of race (Bertrand & Mullainathan, 2004);

Table 1: Log Listing Price Differences Between White and Black Hosts

	(1)	(2)	(3)
Minority	-0.040	-0.032	-0.032
	(0.001)	(0.008)	(0.008)
Average Quantitative Reviews			
4.5 and above		-0.032	-0.016
		(0.005)	(0.005)
3.5 to 4.5		-0.039	-0.036
		(0.007)	(0.007)
Below 3.5		0.030	0.035
		(0.284)	(0.282)
Most Recent three Qualitative Reviews			
At least one review 3 star or below*			-0.004
			(0.005)
All 4 or 5 star reviews, with at least one 4 star review.			-0.012
			(0.004)
R^2	0.72	0.72	0.73
Observations	70,334	63,153	63,153

Notes:

(1) All results are from the OLS regression of log prices on an indicator for Black individuals;

(2) In addition to the listed variables, the results include: city fixed effects; neighbourhood fixed effects; the twenty most popular amenities; host response time; *superhost* status; host verification status; and the average number of reviews per month;

(3) Standard errors are clustered at a property level;

∗ The reference group are the properties with exclusively five star ratings in their three most recent qualitative reviews.

picture analysis using machine learning techniques; and, language usage (Grogger, 2008).

Despite computational costs, I use facial recognition techniques to analyse the race of hosts using profile pictures scraped from Airbnb. Specifically, I utilise the picture recognition software, the *DeepFace*, that was introduced in Taigman et al. (2014).

In order to confirm the precision and recall[8] of the Deepface software, I manually analyse the race of two hundred photographs without prior knowledge of the automated results. I

[8] Where precision and recall are defined as:

$$Precision = \frac{tp}{tp + fp}.$$

$$Recall = \frac{tp}{tp + fn},$$

where tp is the number of true positives, fp, the number of false positives, and fn, the number of false negatives.

achieve precision of 72% and recall of 71%. For the purposes of my econometric analysis, I use the subset of hosts that are classified by the software as White or Black. For this subset, I achieve a precision of 89%.

Table 2: Racial Analysis of Airbnb Hosts

Race	**Observations**	**%**
White	62,302	39.7
Black	9,368	6.0
Other	37,062	23.6
Unknown*	48,384	30.8
Total	157,116	

Notes:

(1) The table provides details of the anlaysis of host race using the *Deepface* software.

∗ A profile picture is defined as Unknown if it is: (1) a picture of the property, a corporate logo, or some other object; (2) the face is obscured; or (3) the picture is of low quality, and therefore, cannot be analysed.

4.2. Analysis of Qualitative Reviews

4.2.1 Background

As discussed in the literature section, several economics papers have utilised the information from text in their analysis (Mueller & Rauh (2018); Engle et al. (2020); Stephens-Davidowitz (2014). Despite success in other literatures, to the best of my knowledge, these techniques have not been used to explain racial differences on sharing platforms.

There are a large variety of techniques that facilitate the use of textual data for economic analysis. However, this section provides information only on the BERT language model that I use in this paper.

4.2.2 BERT Language Model

The BERT model is a state of the art language model developed in Devlin et al. (2018). The original model is:

> ... designed to pretrain deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers.(Devlin et al., 2018, pp.1)[9]

[9] A more detailed discussion of the BERT model can be found in Appendix (B).

In the spirit of the Transformer architecture, I adapt the pre-trained model to my purpose by fine-tuning it for the task of text classification using a sample of 29,096 reviews for which the review score is known. Although quantitative scores for each individual review are unavailable on Airbnb, I use the subset of properties whose number of reviews has increased by one between periods, and as such, the quantitative score associated with the review can be calculated by comparing the average review from the current and past period.

These 29,096 reviews are split into a train and test set of 21,822 and 7,274 reviews, respectively. The training set is further split into a train and validation set, and trained over 3 epochs.[10]

Table (3) provides the confusion matrix[11] for the version of the model with the three classifications: Automated review due to cancellation at short notice; Review score between 1 and 3; and Review score of 4 or over. The associated accuracy is 95.4%.

Table 3: Confusion Matrix for Text Classification with Three Categories

	Actual		
	4 or above	3 or below	Cancellation
Predictions			
4 or above	6,303	112	0
3 or below	209	255	1
Cancellation	10	0	384

Notes:

(1) The table details the predictions of my fine-tuned BERT model with three categories.

(2) I predict the value of 7,274 labelled reviews with an accuracy of 95.4%

The fine-tuned model is used to predict the value of the past three reviews for each active listing. The results of this prediction are exploited in the analysis of the racial log pricing differences as discussed in the following sections.

5. Model and Link to the Theory

5.1. Background

As discussed in Section (2), prior research has provided empirical evidence for the existence of a racial listing price gap. To date, the impacts of recent reviews, and historical pricing differences have not been considered. To this end, this section builds on the model of

[10] In machine learning an epoch is defined as one pass of the entire training data through the algorithm.

[11] A confusion matrix is a tableau representation of a classification exercise.

Laouénan & Rathelot (2022) with a view to deepen the understanding of ethnic listing price differences. The model in this paper diverges from their model in that the triple of recent qualitative reviews, quantitative reviews, and value for money provide a signal of property quality, as opposed to the signal being derived entirely from the quantitative review.

5.2. Overview

There are two sides of the market, one of hosts, and another of potential guests. Hosts set their rental price, p_t, which they receive if their property is successfully rented. A guest who rents the property pays p_t directly to the host.[12] The rental market is sufficiently large such that neither a host's nor a potential guest's rental decision impact prices.

A potential guest chooses an Airbnb property based on the information available on the website. She requests to stay in a property if that choice maximises her expected utility between staying in properties available on Airbnb, and her outside option, potentially renting a property from an alternative source, or not making the trip.

For potential guests, property quality is only partially observable at the booking stage, and fully observable after the completion of the stay. The partially observable component of quality includes: the number of bedrooms; bathrooms; and location. Whereas initially unobservable quality includes: host attitude; exact location; and the noise and smell of the property.

When searching for a property, potential guests form beliefs about the unobservable component of quality which determines their willingness to pay. The preferences and beliefs of White and Black guests are the same.

Finally, a host's pricing decision is myopic, and she does not consider the impact of current price on future ratings, nor the impact of future ratings on future prices.[13]

5.3. Host Pricing Decision

Following (Laouénan & Rathelot, 2022, pp. 117), in each time period, hosts have one unit of time which they split between paid work for $1 - L$ hours, and on their Airbnb property for L hours. Rental activity has decreasing returns to scale, and L hours spent on Airbnb results in $L^{\tilde{\alpha}}$ nights supplied, where $\tilde{\alpha} \in (0, 1)$.

A host maximises their current period income by solving the following problem:

[12] On Airbnb hosts receive rental income net of fees. However, for the purposes of the model, these Airbnb fees are ignored.

[13] In my data, each property has a mean availability of 9 days in the next 30 days, and 22 days in the next 60. Given the penalties for cancellation, this suggests that hosts do not update the price of their properties for each new review.

$$\max_{p_t} \left\{ p_t D_t(p_t, Q, m) + (1 - D_t^{\frac{1}{\tilde{\alpha}}}(p_t, Q, m))W \right\}, \tag{1}$$

where p_t is current period price, W is the income from paid work, and $D_t(p_t, Q, m)$ is the demand for a property defined as:

$$D_t(p_t, Q, m) = \frac{f(t)^{\eta} Q^{\beta}}{{p_t}^{\kappa} \Gamma^{m}}, \tag{2}$$

where Q is the sum of observable and unobservable quality, m is the host's race, $f(t)$ is a local seasonality function, and β, η, κ, and Γ are constants.

Solving equation (1) using a first order approach, taking the natural logarithm, and substituting in the value for expected unobservable quality, yields:

$$p_t = p_0 - \lambda\gamma m + \lambda\alpha w + \lambda\beta Q + \lambda\eta g(t),$$

where $p_t = \log P_t$, $w = \log W$, $\gamma = \log \Gamma$, $\alpha = \frac{\tilde{\alpha}}{1-\tilde{\alpha}}$, $\lambda = \kappa + \alpha^{-1}$, $g(t) = \log f(t)$, and $p_0 = \lambda\alpha \log \frac{\tilde{\alpha}(\kappa-1)}{\kappa}$.

5.4. Listing signal of quality

First, I mirror the approach in Bohren, Imas, & Rosenberg (2019), whereby, quality, Q, is the sum of two orthogonal components, observable quality, q, and unobservable quality, ν,

$$Q = q + \nu.$$

Each host has an observable racial identity $m \in \{1, 0\}$, where 0 indicates White and 1 Black. The distribution of unobservable quality differs across races, and can be defined as $\nu \sim \mathcal{N}(\overline{\nu}_m, \sigma^2)$.

When viewing a listing, prospective guests extract a signal from the review information provided by past guests. This information includes the subset of qualitative reviews read by the prospective guest,[14] and the average quantitative review in respect of seven criteria:

[14] Several properties have hundreds of qualitative reviews. It is unrealistic that prospective guests read each review. As such, I assume prospective guests read a subset of these reviews.

overall; cleanliness; accuracy; communication; location; check-in; and value for money.

From the description of quality, the quantitative reviews of: cleanliness; accuracy; communication; location; and check-in, may be considered as components of unobservable quality. Whereas value for money, although desirable from a guest's perspective, may more reasonably be considered a function of historical pricing decisions.

Although Airbnb guests provide review scores for each component individually, I assume that overall rating is a function of the other six components. For this reason, I treat the tuple of average overall rating and average value for money, $(\overline{r}_{it}, \overline{v}_{it})$, as a signal of unobservable quality.

In line with the literature (Bohren, Imas, & Rosenberg, 2019; Koffi, 2021; Laouénan & Rathelot, 2022), I assume that the signals from both qualitative and quantitative reviews are drawn from normal distributions centered at the true value of ν.

Given the distributional assumption on the signals, as is shown in Appendix (A), the expected quality of a property with K reviews, and where N of the qualitative review components are read by a prospective guest is given by:

$$\mathbb{E}(\nu|s, m, K, N) = \frac{K\, s(\overline{r}) + \frac{NK}{K-N\rho^2}\frac{\sigma_r^2}{\sigma_{\tilde{r}}^2}\, s(\overline{\tilde{r}}|\overline{r}) + \frac{\sigma_r^2}{\sigma_\nu^2}\overline{\nu}_m}{K + \frac{NK}{K-N\rho^2}\frac{\sigma_r^2}{\sigma_{\tilde{r}}^2} + \frac{\sigma_r^2}{\sigma_\nu^2}}, \tag{3}$$

where $s(\overline{r})$ is the signal from one quantitative review of value $\overline{r}$, and $s(\overline{\tilde{r}}|\overline{r})$ is the conditional signal from a qualitative review of value $\overline{\tilde{r}}$. σ_r^2, $\sigma_{\tilde{r}}^2$, and σ_ν^2 are the variance of a quantitative review, qualitative review, and the average variance of a Black individual's quality. Finally, ρ is the correlation coefficient of qualitative and quantitative reviews.

5.5. The Host's Price Setting Equation

If the host is aware of the consumer's estimation of quality calculation, she will account for this in her price maximisation problem. As such, the final equation becomes:

$$p_t = p_0 - \lambda\gamma m + \lambda\alpha w + \lambda\beta q + \lambda\eta g(t) + \lambda\beta\mathbb{E}[\nu|\overline{r}_t, \overline{v}_t, \tilde{r}_t, K], \tag{4}$$

Where the expected unobservable quality is defined as per equation (3).

6. Listing Price Differences - Empirical Analysis

Previous sections provided motivation and evidence of an ethnic listing price gap, and how this changes over time. This section provides the empirical specifications and analysis that test this theory.

6.1. Benchmark Model

Let p_{it} be the log rental price of property i at time t, $\overline{r}_{it}$ the average overall rating, K_{it} the number of reviews, N_{ib} the number of qualitative reviews rated 3 star or less in the most recent N reviews, and N_{ig} the corresponding number of 5 star reviews. In the baseline analysis $N = 3$. The coefficient of interest, β_B, measures the log price difference between Black and White hosts for comparable properties with zero reviews.

To ensure the comparability of properties, a large number of property, and host characteristics, z^1_{it}, z^2_i, respectively, are included as control variables.[15] The explanation for the inclusion of the majority of these variables is self-evident. For example, why the property size, location, and available amenities may influence the property price. However, there are a small number of variables that warrant further discussion. In particular, both reviews per month, and a binary variable that signifies multi-property listings is included to account for professional users of Airbnb to whom occupation rates may be more important than headline prices.

The log listing price, p_{it}, is given by the following model:

$$\begin{aligned} p_{it} = \beta_1 + \beta_2 z^1_i + \beta_3 z^2_{it} + f_b(K_{it}, N_{it})\, \beta_B\, \mathbb{1}\{m_i = 1\} + f_1(K_{it}, N_{it}) \sum_{\overline{r}} \beta_{\overline{r}}\, \mathbb{1}\{\overline{r}_{it} = \overline{r}\} \\ + f_2(K_{it}, N_{it}) \Big(\beta_{N_b}\, \mathbb{1}\{N_{ib} \geq 1\} + \beta_{N_a} \big(1 - \mathbb{1}\{N_{ib} \geq 1\}\big) + \beta_{N_e}\, \mathbb{1}\{N_{ig} = N\} \Big) + \varepsilon_{it}. \end{aligned} \tag{5}$$

The functions f_b, f_1, and f_2 are functions of the underlying parameters as per equation (8) and are defined as follows:

$$\begin{aligned} f_b(K, N) &= \frac{\sigma^2_r}{\sigma^2_\nu} f(K, N). \\ f_1(K, N) &= K\, f(K, N); \\ f_2(K, N) &= \frac{NK}{K - N\rho^2} \frac{\sigma^2_r}{\sigma^2_{\tilde{r}}} f(K, N). \end{aligned}$$

[15] A full list of control variables can be found in Tables (9) and (10) in Appendix (D)

where:

$$f(K,N) = \frac{1}{K + \frac{NK}{K-N\rho^2}\frac{\sigma_r^2}{\sigma_{\tilde{r}}^2} + \frac{\sigma_r^2}{\sigma_\nu^2}};$$

I solve equation (5) iteratively, firstly by estimating the underlying structural parameters by non-linear least squares, and substituting these estimates into equation (5) which is then estimated by OLS, with errors clustered at the property level. These steps are repeated until convergence. A summary of the results of this estimation process can be found in Tables (4) and (5).

6.2. Interpretation

Specification (1) is the original specification per (Laouénan & Rathelot, 2022, pp.121) and provides a useful benchmark against which to judge the impact of the additional elements of my specification.

Specification (2) embeds information from the latest three qualitative reviews into the signal of unobservable quality. In particular, it quantifies having one or more reviews below 4 star, the impact of having a mixture of 4 and 5 star reviews, and the impact on price of having exclusively 5 star reviews in the most recent three reviews. This modification captures several important features of the setting that are otherwise ignored, namely: agents are interested in both recent and average performance; they rely on the information contained in qualitative reviews to make decisions; and, where ratings are highly concentrated, information from additional sources is highly relevant.

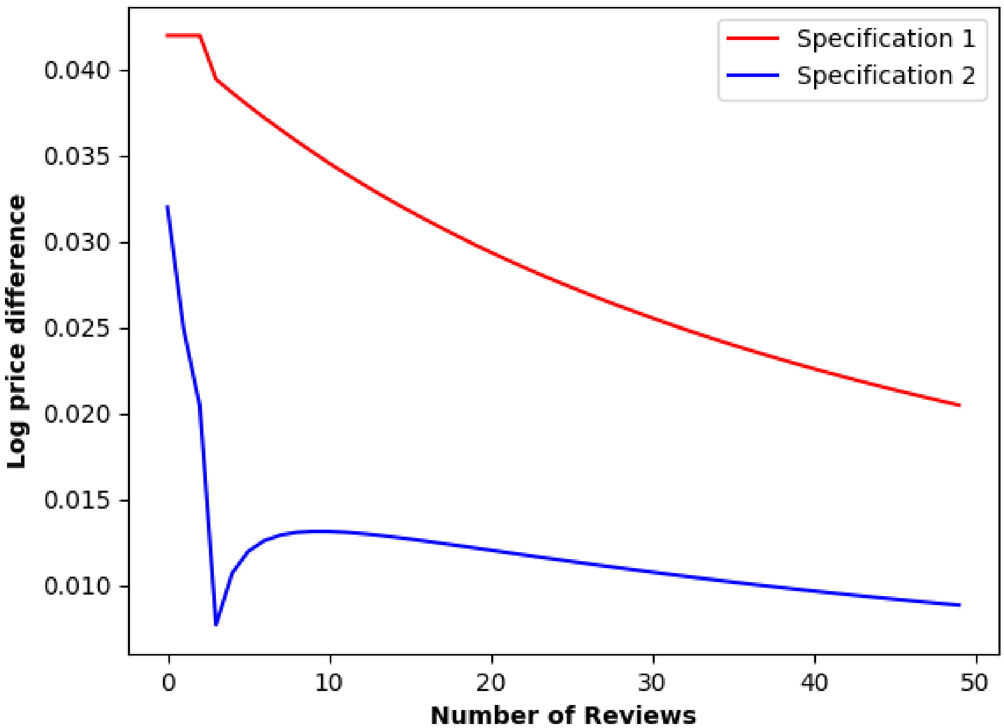
Upon the inclusion of the qualitative review information, the coefficient of interest, β_B, shrinks from -0.042 to -0.032, and its significance reduces.[16]

In specification (1) more than 50% of the initial ethnic price gap remains until a property has 46 reviews. In contrast, in the model with qualitative reviews, less than 50% of the initial ethnic price gap remains after three reviews.[17] The effects on the log price difference as the number of reviews changes can clearly be seen in Figure (3) below.

[16]It should be noted that the ratio of review variance to racial variance is approximately 3 times larger in my calculation compared to the original formulation in Laouénan & Rathelot (2022). This is likely an artefact of the different definitions of race between the two papers.

[17]Full review information becomes available after three reviews.

Figure 3: Log Price Difference between White and Black Hosts over Time



Notes:

(1) This figure provides a graphical illustration of the log price gap between Black and White hosts under Specifications (1) and (2) as the number of reviews increases.

The impact of having a mix of 4 and 5 star reviews compared to exclusively 5 star is of a similar magnitude as the penalty against Black hosts. Figure (4) in Appendix (E) provides details of how the probability of having no reviews lower than 4 stars, and the probability of exclusively 5 star reviews increases with the number of reviews. This is a result of poor quality hosts invariably exiting the market, such that most remaining active hosts have high recent review scores.[18] Against this backdrop the negative coefficient on properties whose recent reviews are not exclusively 5 stars becomes clearer.

[18]Figure (4) in Appendix (E) provides a graphical illustration of this feature.

Table 4: Log Listing Price Differences

	(1)	(2)
Racial Minority	-0.042***	-0.032**
	(0.01)	(0.01)
Average Review Score		
Rated above 4.5	-0.131***	-0.139***
	(0.01)	(0.01)
Rated 3.5 to 4.5	-0.235***	-0.260***
	(0.02)	(0.02)
Rated below 3.5	-0.019	0.060
	(0.19)	(0.42)
Latest Qualitative Reviews		
One or more reviews below 4 star		-0.001
		(0.01)
All 4 or 5 star reviews		-0.035***
		(0.01)
All recent reviews 5 star		-0.008
		(0.01)
Observations	70,334	63,153

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes:

(1) Results are from the OLS regression of log prices on a binary variable for Black individuals.

(2) Specification (1) is an estimation of the model per Laouénan & Rathelot (2022). Specification (2) is an estimation of equation (5).

(3) In addition to the listed variables, the results include: city fixed effects; neighbourhood fixed effects; the twenty most popular amenities; host response time; *superhost* status; host verification status; and the average number of reviews per month.

(4) Standard Errors are clustered at a listing property level.

6.3. Robustness Checks

6.3.1 Value for Money

In the above model agents are assumed to be myopic decision makers, and in particular, they do not consider the impact of price on ratings. If this assumption is incorrect, historical prices may have a material impact on the analysis.

Given the hypothesis that the listing price for an equivalent property is lower for a Black host than a White one, if Black and White hosts have observationally equivalent properties (including average review scores), the Black host should be the host of a less desirable property due to the impact of historical pricing on ratings.

Table 5: Structural Parameter Estimates for Specification (1) and (2)

	(1)	(2)
$\frac{\sigma_r^2}{\sigma_\nu^2}$	46.604***	29.037**
	(2.84)	(11.16)
$\frac{\sigma_r^2}{\sigma_{\bar{r}}^2}$		9.853**
		(3.94)
Observations	70,333	63,151

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes:

(1) This table provides details of the structural parameter estimates for the regression of log prices under Specifications (1) and (2).

In light of the above discussion, the following empirical specification includes the effects of past prices on current log prices by incorporating an individual's Value for Money score into the signal of property quality.

Both the dependent variable and independent variables in equation (6) are the same as equation (5) with the exception that the average Value for Money score, $\overline{v}_{it}$, is included in the signal of unobservable quality as follows:

$$p_{it} = \beta_1 + \beta_2 z_i^1 + \beta_3 z_{it}^2 + f_b(K_{it}, N_{it})\beta_B \mathbb{1}\{m_i = 1\} + f_1(K_{it}, N_{it}) \sum_{(\overline{r},\overline{v})} \beta_{\overline{r},\overline{v}} \mathbb{1}\{(\overline{r}_{it}, \overline{v}_{it}) = (\overline{r}, \overline{v})\}$$
$$+ f_2(K_{it}, N_{it}) \Big(\beta_{N_b} \mathbb{1}\{N_{ib} \geq 1\} + \beta_{N_a} \left(1 - \mathbb{1}\{N_{ib} \geq 1\}\right) + \beta_{N_g} \mathbb{1}\{N_{ig} = N\} \Big) + \varepsilon_{it}. \tag{6}$$

From Table (11), the estimates of β_B remain unchanged from the baseline specification. In contrast, there are some immaterial changes to the underlying structural parameters. However, given that the conclusion of the analysis is materially unchanged, it would appear that the pricing differences discussed above are not a major feature of the data.

6.3.2 Heterogeneous impact of qualitative reviews

In the benchmark model, it is assumed that the impact of recent qualitative reviews is homogenous across hosts which is likely to be an oversimplification. For example, an exceptionally good review should impact the desirability of a property with a low average rating more than a high rated property.

In order to test this additional feature, I estimate the following equation:

$$
\begin{aligned}
p_{it} = \beta_1 + \beta_2 z_i^1 + \beta_3 z_{it}^2 + f_b(K_{it}, N_{it})\, \beta_B\, \mathbb{1}\{m_i = 1\} + f_1(K_{it}, N_{it}) \sum_{\overline{r}} \beta_{\overline{r}}\, \mathbb{1}\{\overline{r}_{it} = \overline{r}\} \\
+ f_2(K_{it}, N_{it}) \sum_{\overline{r}} \Big(\beta_{N_{b,\overline{r}}}\, \mathbb{1}\{N_{ib} \geq 1,\, \overline{r}_{it} = \overline{r}\} + \beta_{N_a} \big(1 - \mathbb{1}\{N_{ib} \geq 1,\, \overline{r}_{it} = \overline{r}\}\big) \\
+ \beta_{N_{g,\overline{r}}}\, \mathbb{1}\{N_{ig} = N,\, \overline{r}_{it} = \overline{r}\} \Big) + \varepsilon_{it}.
\end{aligned} \tag{7}
$$

Where all parameters are defined as per equation (5). I solve equation (7) iteratively in an analogous manner to the benchmark case. The results of this estimation are presented in Tables (6) and (7).

The above specification offers additional granularity on the impacts of qualitative reviews. In particular, a property rated below 3.5 stars with three recent 5 star ratings will be priced at a 24% premium compared to other properties rated below 3.5 stars.

In a similar manner to specification (2), the price of a property rated above 4.5 stars with at least one recent four star review is 6% lower than other properties rated above 4.5 stars. Additionally, for properties rated 4.5 stars or above, there is a negative coefficient on having one or more recent review of 3 stars or lower. However, this coefficient is statistically insignificant. These observations reflect the reality that in the group of properties that are rated 4.5 stars or above, a recent rating below 5 star is both undesirable and unexpected. However, given the low probability of having an average rating above 4.5 stars and a recent review below 4 star, the coefficient on this specific group is insignificant.

Allowing for the impacts of qualitative reviews to be heterogeneous on prices results in the ethnic listing price gap being more persistent than under specification (2). However, this gap remains considerably less persistent than in specification (1). In fact, it takes 70% fewer reviews for half the initial listing price gap to disappear. Illustrative of this is that for all properties with at least three reviews, the racial listing price gap is below 2%, an amount which is only reached after fourteen reviews in the analysis per Laouénan & Rathelot (2022). Figure (7) in Appendix (E) provides a graphical illustration of the racial listing price gap under the three specifications.

Table 6: The Heterogeneous Impact of Qualitative Reviews on Log Listing Price Differences

	(2a)
Racial Minority	-0.039***
	(0.01)
Average Review Score	
Rated above 4.5	-0.139***
	(0.01)
Rated 3.5 to 4.5	-0.272***
	(0.02)
Rated below 3.5	-0.374
	(0.55)
Latest Qualitative Reviews per review rating	
Rated above 4.5 stars	
One or more reviews below 4 star	-0.038
	(0.03)
All reviews 4 or 5 star	-0.067***
	(0.02)
All recent reviews 5 star	-0.015
	(0.01)
Rated 3.5 to 4.5	
One or more reviews below 4 star	0.045
	(0.03)
All reviews 4 or 5 star	-0.006
	(0.03)
All recent reviews 5 star	-0.011
	(0.02)
Rated below 3.5	
One or more reviews below 4 star	0.105
	(0.08)
All reviews 4 or 5 star	-0.056
	(0.21)
All recent reviews 5 star	0.275***
	(0.14)
Observations	63,153

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes:

(1) Results are from the OLS regression of log prices on a binary variable for Black individuals.

(2) In addition to the listed variables, the results include: city fixed effects; neighbourhood fixed effects; the twenty most popular amenities; host response time; *superhost* status; host verification status; and the average number of reviews per month.

(3) Standard Errors are clustered at a listing property level.

Table 7: Structural Estimates of Parameters for the Heterogeneous Impact of Qualitative Reviews on Log Listing Prices

	(2a)
$\frac{\sigma_r^2}{\sigma_\nu^2}$	40.773***
	(5.359)
$\frac{\sigma_r^2}{\sigma_{\tilde{r}}^2}$	5.803***
	(1.891)
Observations	63,151

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes:

(1) This table provides details of the structural parameter estimates for the regression of log prices with heterogeneous effects of qualitative review.

6.4. Threats to Identification

In order to interpret the results from estimating equations (5) and (7) as causal, I need to establish that neither measurement error nor omitted variable bias are responsible for the effects. The following section provides evidence against what I consider to be the biggest threats to this interpretation.

6.4.1 Qualitative Reviews

Given the nature of my data, a risk of measurement error comes from my calculation of qualitative review quality. As such, I test my results for sensitivity in respect of this calculation.

To this end, I construct a set of 50 placebo experiments where I randomly classify the latest qualitative reviews of a property. Firstly, I randomly assign reviews of 4 stars or more to different properties, with probability equal to the unconditional probability from my original analysis. Subsequently, I randomly assign properties with all 5 star reviews where the probability conditional on all reviews being greater than 4 star is equal to the probability in my original analysis.

In lieu of recalculating the structural parameters for each of the 50 experiments, I estimate equation (5) using the original structural estimates and the simulated qualitative review quality.

In 50% of the experiments, no coefficient estimates are significant at any conventional level, and only 2% of the experiments have coefficient estimates that are significant at the 1% level.

Of the experiments with significant coefficients, they are invariably of limited economic impact. The largest coefficient is approximately 70% of the original, with the majority being significantly smaller. Finally, the significant coefficients are equally distributed across the 3 groups: at least one review less than 4 star; all reviews 4 star or above, with at least one 4 star review; and all recent 5 star reviews. Details of the first ten simulations can be found in Tables (14) and (15) in Appendix (E).

As a result, the simulations provide support that both the BERT model, and the associated regression analysis, explain an important feature of the data.

6.4.2 Difference in Property Usage

If White and Black hosts use their properties in different ways, for example, one group were predominantly professional hosts, and the other group used Airbnb sporadically to supplement their income, I would expect different residency rates across racial groups. By regressing reviews per month on the host and property characteristics, I find that race is insignificant at all conventional levels. Table (13) in Appendix (D) provides a summary of this regression. Therefore, I argue that my results are not a feature of different property usage.

6.4.3 Pricing Conventions

If Black hosts initially underprice their properties and over time learn their true property value, the pricing history of Black hosts would be similar to those under the model of racial discrimination. In this setting, if Black hosts were systematically under-pricing their properties, the market would respond in the form of higher ratings and / or increased demand for Black properties. Figure (5) illustrates that ratings for Black hosts are consistently below those for White hosts, and Table (13) in Appendix (D) illustrates that residency rates are not significantly different between Black and White hosts. When taken alongside Yu & Margolin (2022) that finds demand is lower for Black properties, the features we see in my data are unlikely to result from Black hosts underpricing their properties.

6.4.4 Market exit rates

One explanation for the results in this paper is that low quality Black hosts leave the market quicker than low quality White hosts, and this is most pronounced at the point full review information becomes available. Although Figure (5) in Appendix (E) illustrates that a marginally greater proportion of Black hosts leave the market than their White peers, Figure (6) in Appendix (E) provides evidence against these hosts being lower quality than their

White peers. As such, the results being driven by a disproportionate number of low quality Blacks leaving the market is not supported by the data.

6.4.5 Difference in Guest Acceptance Rates

If Black hosts underpriced their properties in order to increase demand, and ultimately, be more selective with respect to guest quality or race, review quality would be endogenous. As can be seen in Figure (9), acceptance rates for Black hosts are lower than their White peers. Furthermore, over 60% of hosts have acceptance rates above 95%, suggesting guest selection is not a large feature of my data.

6.4.6 Racial Differences in Review

Throughout this paper I have assumed that reviews are free from racial bias. However, it is possible that this an oversimplification. To better understand if race influences the review process, I analyse how review content differs across racial groups. To this end, I use the Latent Dirichlet Allocation (LDA) model that was originally proposed in Blei, Ng, & Jordan (2003).[19]

In an analogous way to Guo, Barnes, & Qiong (2017), I train an LDA model to analyse the key dimensions in which reviews differ. As can be seen from Figure (8), the mean topic distribution is remarkably similar for White and Black hosts.[20] This suggests that any racial bias that exists within reviews does not manifest in differences in the content of these reviews.

Tables (16) and (17) in Appendix (D) illustrate how little variation in the topic distribution of a review can be explained by the host's race. This further illustrates that review content does not vary significantly with race, and as such, I find no evidence of systematic racial bias in the qualitative reviews in this dataset.

7. Conclusion

Racial minorities continue to achieve worse economic outcomes in online market places. Several papers have made efforts to quantify the difference on Airbnb (Edelman & Luca (2014); Edelman, Luca, & Svirsky (2017)), and Laouénan & Rathelot (2022) recently categorised the differences as deriving from statistical discrimination. However, few if any studies have sought to use the information from qualitative reviews to supplement this analysis.

[19] A detailed discussion can be found in Appendix (4.2).

[20] Table (18) in Appendix (D) provides details of the Topics that differ the most across racial groups.

After adding the effects of qualitative reviews, I find that differences in outcomes are considerably less persistent than those found in the wider literature, albeit still derived from statistical discrimination. In my model, the listing price gap shrinks by over 70% after full review information is released. Whereas under Laouénan & Rathelot (2022)'s model, it takes fourteen reviews for the listing price gap to shrink by 50%.[21] To put these numbers into perspective, in my model, 91% of Black hosts suffer racial statistical discrimination that is less than 50% of the initial amount for unreviewed properties, whereas, the comparable number for the Laouénan & Rathelot (2022) model is 37% of Black hosts.

The results have strong implications for policy on Airbnb, and other online marketplaces. Firstly, given the impact of qualitative reviews and the variable nature of the length of reviews, users should be incentivised to leave detailed qualitative reviews. Secondly, consideration should be given to the necessity of obscuring reviews until a minimum number is reached. Potentially this policy is in place to limit the impact of fake reviews. However, in its absence, high quality hosts are likely to achieve better market outcomes regardless of race.

Beyond its implications in the online marketplace, this paper aims to explain the mechanisms through which statistical racial discrimination acts on outcomes. In particular, what are the best forms of feedback mechanism to reduce the impact of initial discrimination, and what are the most efficient ways to evaluate quality where the individual is unknown? This leads to some important questions which I aim to explore in subsequent work such as the potential impact of referral letters in recruitment decisions.

Finally, this paper proposes that more prominence is given to qualitative reviews in online marketplaces in order to reduce the impact of racial discrimination in online market outcomes.

[21] This is a conservative estimate using the result calculated in Laouénan & Rathelot (2022). Using the estimates from my paper, it would take forty-six reviews.

Appendix

A. Expectation of Unobserved Quality

This section follows the model described in (Laouénan & Rathelot, 2022, pp.117) and (Koffi, 2021, pp.xi), with a view to extend the signalling model.

As discussed in Section (5.4), the quality of each property, Q, is the sum of two orthogonal components: observable quality, q; and unobservable quality, ν,

$$Q = q + \nu.$$

Each host has an observable racial identity $m \in \{1, 0\}$, where 0 signifies White, and 1 Black. The distribution of unobservable quality differs across races, and can be defined as $\nu \sim \mathcal{N}(\overline{\nu}_m, \sigma_\nu^2)$.

Each property listing is defined by a set of features, namely, the observable quality components, and the set of reviews. The set of reviews can be written as $r = \{\overline{r}, \tilde{r}_1, .., \tilde{r}_K\}$, where $\overline{r}$ is the average review score, $\tilde{r}_i$ is the qualitative review for stay i, and K is the total number of historical guests.

In general, it is infeasible that a potential guest reads all the qualitative reviews that are visible on a host's profile. Instead, each host will read a subset of the available reviews. Although the number of reviews that each perspective guest reads will differ, for the purpose of this analysis, I assume that this number is a fixed $N \in \mathbb{N}$.

Bohren, Imas, & Rosenberg (2019); Koffi (2021); and Laouénan & Rathelot (2022) model an environment where a signal is elicited from each feature. These signals are assumed to be i.i.d. normal around the true value of unobservable quality. Given the deviation from their models, I extend the framework such that signals are independent over time, but, the signal pair, $\big(S(r_i), S(\tilde{r}_i)\big)$, is jointly normal as follows:

$$\begin{pmatrix} S(r_i) \\ S(\tilde{r}_i) \end{pmatrix} \sim N\left(\begin{pmatrix} \nu \\ \nu \end{pmatrix}, \begin{pmatrix} \sigma_r^2 & \rho\sigma_{\tilde{r}}\sigma_r \\ \rho\sigma_{\tilde{r}}\sigma_r & \sigma_{\tilde{r}}^2 \end{pmatrix}\right),$$

where ν, $\sigma_{\tilde{r}}$, σ_r are defined as above, and ρ is the correlation coefficient between the two signals.

Given the above distributional assumptions, the average quantitative and qualitative signals, $S(\overline{r})$ and $S(\overline{\tilde{r}})$, are distributed as follows:

$$\begin{pmatrix} S(\overline{r}) \\ S(\overline{\tilde{r}}) \end{pmatrix} \sim N\left(\begin{pmatrix} \nu \\ \nu \end{pmatrix}, \begin{pmatrix} \frac{\sigma_r^2}{K} & \tilde{\rho}\,\frac{\sigma_r}{\sqrt{K}}\frac{\sigma_{\tilde{r}}}{\sqrt{N}} \\ \tilde{\rho}\,\frac{\sigma_r}{\sqrt{K}}\frac{\sigma_{\tilde{r}}}{\sqrt{N}} & \frac{\sigma_{\tilde{r}}^2}{N} \end{pmatrix}\right),$$

where $\tilde{\rho}$ is the correlation coefficient between the two average signals.

The remainder of this section provides the theoretical basis that supports the calculation of expected unobservable quality from these signals.

The value of $\tilde{\rho}$ varies with K and N, as such, I calculate this in terms of estimable parameters.

$$\begin{aligned} Cov\big(S(\overline{r}),\, S(\overline{\tilde{r}})\big) &= Cov\Bigg(\frac{1}{K}\sum_{i=1}^{K} S(r_i),\, \frac{1}{N}\sum_{i=1}^{N} S(\tilde{r}_i)\Bigg); \\ &= \frac{1}{KN}\sum_{i=1}^{K}\sum_{j=1}^{N} Cov\big(S(r_i), S(\tilde{r}_j)\big); \\ &= \frac{1}{KN} N Cov\big(S(r), S(\tilde{r})\big); \\ &= \frac{1}{K}\,\rho\,\sigma_{\tilde{r}}\,\sigma_r. \end{aligned}$$

Finally, given:

$$\tilde{\rho} = \frac{Cov\big(S(\overline{r}),\, S(\overline{\tilde{r}})\big)}{\frac{\sigma_r}{\sqrt{k}}\,\frac{\tilde{\sigma}_r}{\sqrt{N}}}.$$

Substituting in the value of the covariance of average signals, we have:

$$\tilde{\rho} = \frac{\sqrt{N}}{\sqrt{K}}\,\rho.$$

In line with Laouénan & Rathelot (2022), I take the weighted average of the signals. However, given the non-zero value of $\tilde{\rho}$, there is an extra step to calculate the variance of the signal of qualitative reviews conditional on the qualitative review average. Appealing to known results, this is simply:

$$\sigma^2_{\overline{\tilde{r}}|\overline{r}} = (1-\tilde{\rho}^2)\,\frac{\sigma^2_{\tilde{r}}}{N}.$$

Finally, taking the weighted average of these signals, where the weights are the inverse of their variance, provides the following expression for the expectation of unobservable quality:

$$\mathbb{E}(\nu|s,m,K,N) = \frac{K\,s(\overline{r}) + \frac{NK}{K-N\rho^2}\frac{\sigma_r^2}{\sigma_{\bar{r}}^2}\,s(\bar{\bar{r}}|\overline{r}) + \frac{\sigma_r^2}{\sigma_\nu^2}\overline{\nu}_m}{K + \frac{NK}{K-N\rho^2}\frac{\sigma_r^2}{\sigma_{\bar{r}}^2} + \frac{\sigma_r^2}{\sigma_\nu^2}}, \tag{8}$$

B. BERT Model

B.1. BERT Overview

The BERT (Bidirectional Encoder Representations from Transformers) model is a state of the art language model developed in Devlin et al. (2018). It was pre-trained on the tasks of Masked Language Modelling, and Next Sentence Prediction, on a corpus that totalled 3.3 billion words. The following section is based on Devlin et al. (2018). However, it has purposefully been altered to be understood by a wider audience. Any mistakes in the course of this simplification are my own, and are not attributed to the original article.

The model takes an input which is a series of symbols, transforms them into a different representation, and then passes them to any subsequent levels. Once these inputs have passed through the series of encoders, they are then decoded into an output series of symbols.

The model architecture is bidirectional which is described in (Devlin et al., 2018, pp.1) as "...conditioning on both the right and left context of each word". This allows the model to consider the deeper context of the inputs, in contrast to simple counts using Bag of Words, N-Grams or other techniques. This approach means that the same word may result in several different encodings based on its context in each specific occurrence.

The model sits within the wider group of models that use a system of self-Attention. This is a form of weighting mechanism that places greater emphasis on some encodings and less on others. This is applied at each level of the model, with the previously generated symbols being used as additional input to generate the next layer.

Finally, the model can be fine-tuned for various tasks including sentiment classification. Using labelled data, the underlying parameters of the model are updated subject to a loss condition, in order to maximise performance on the given task. It is in exactly this manner that I use the model.

B.2. BERT Output

The following are examples of reviews that were classified by the fine-tuned model. These reviews have been chosen at random, and are not intended to be representative of model performance.

Examples of 1 to 3 Rated Reviews

"The property is undergoing some repairs we were told. However the listing does not come close to match the current state nor the amenities the current listing shows in pictures (i.e. closed community pool). The overall condition is pretty poor and if you plan to spend any amount of time at the property - other than say simply a crash pad - I would have pause."

"Place is farther away from town then advertised 20 - 30 mins, sheets were dirty and overall the place was messy. Bathtub was backed up and there was no hand soap except for an expired one in the shared bathroom. Trying to up charge us $125 a week after we checked out. I use Airbnb all the time and have great reviews so this seems like a scam."

"Needed to be cleaner. The white bathroom rug was black with footprints - not sure if it had ever been washed. Mold on the ceiling in the bathroom that looks like it had been building up for years. Fridge didn't have a working ice or water machine - kid a dirty bag of ice in the freezer. Felt like the apartment hadn't been cleaned in months."

Examples of 4 Rated Reviews

"Older home but lots of character and love."

"The location is great, such an easy walk to everything in little Italy as well as other parts of the city. The place was up 4 flights of very narrow and steep stairs which was a bit difficult with luggage. The only sink is in the bathroom so limit your usage of dishes and cups."

"Small apartment in a great location"

Examples of 5 Rated Reviews

"The unit is in a great location, and it's spacious with tons of natural light. Pool is right outside your door and a great way to relax. Betsy was extremely communicative and helpful as well. Overall a great stay!"

"Araceli was a great host and the house was awesome!"

"This house was magnificent! Lovely location, very clean, wonderful hosts...everything one could ask for in a vacation rental."

C. Latent Dirichlet Allocation Model

The LDA is a topic model in which each document of the corpus, i.e., each review in a collection of reviews, is a mixture over a group of K topics, where each topic is itself a distribution over the global vocabulary of words. Each review is modelled as a random draw of words, drawn independently by first drawing a topic from the topic distribution, and subsequently drawing a word from the specific distribution defined for each topic.

For example, when writing a review an individual writes about the pertinent parts of their stay from K potential topics. In the case of Airbnb this could be 'value for money', 'cleanliness', 'host responsiveness', and 'property characteristics' with the proportion of a specific review taken up by each category being 20%, 25%, 30%, and 25%, respectively.

A more technical description drawn from Blei, Ng, & Jordan (2003) is that each document ω in a corpus of documents $\mathcal{D}$ follows the generative process:

1. Choose the length of the document, N, which follows a Poisson process (ζ);
2. Choose the distribution over topics, θ, that follows a $Dir(\alpha)$, Dirichlet distribution on the $K-1$ simplex with hyper-parameter α;
3. For each of the N words, w_n:
    a) Choose a topic z_n which is drawn from the Multinomial distribution with parameter θ;
    b) Choose a word w_n from $p(w_n|z_n,\beta)$, which is a Multinomial distribution over the vocabulary of words, conditioned on z_n and β, where β is a hyper-parameter.

As shown in detail in Blei, Ng, & Jordan (2003), the probability of a Corpus is given by:

$$p(\mathcal{D}|\alpha,\beta) = \prod_{d=1}^{M} \int p(\theta_d|\alpha) \Bigg(\prod_{n=1}^{N_d} \sum_{z_{d_n}} p(z_{d_n}|\theta_d) p(w_{d_n}|z_{d_n},\beta) \Bigg) d\theta_d$$

Where θ_d is the mixing distribution for document d, z_{d_n} is the topic of word n of document d, and w_{d_n} is word n of document d.

C.1. Data Cleaning Steps

The following steps were taken to prepare the data for analysis in the LDA model. These steps were not necessary for the analysis by the BERT model.

The first step of the data cleaning process is to tokenize the textual reviews, collecting the words and n-grams (a series of n words that frequently occur together) into a list/vector. Words that are frequently used to negate the meaning of the word that follows i.e., 'no' and 'cannot' are removed and 'not' is added to the word that follows.

Next the data is 'stemmed' whereby words with the same root are replaced by that root i.e., "connected", "connecting" and "connects" is replaced with "connect." The stemming tool used is the Porter Stemmer that was originally proposed in Porter (1980).

Next 'stop' words, those words that are included in sentences as dictated by grammatical convention yet provide a limited amount of additional information i.e., 'the', 'of', 'a', are removed from the vector of terms. Finally, both common and uncommon words are removed using the Term-Frequency Inverse Document Frequency ('TF-IDF') method.

D. Tables

Table 8: Description of Observations by City and Date

City	Date	Observations
Austin	08/06/2022	8,759
Austin	12/09/2022	11,854
Boston	20/03/2022	1,655
Boston	15/09/2022	4,018
Chicago	17/03/2022	3,577
Chicago	14/09/2022	5,786
Dallas	17/05/2022	2,924
Dallas	14/09/2022	5,219
London	09/06/2022	20,713
London	12/09/2022	33,146
Los Angeles	06/06/2022	18,798
Los Angeles	10/09/2022	27,592
Manchester	24/03/2022	2,036
Manchester	20/09/2022	3,604
New York	04/06/2022	13,759
New York	07/09/2022	19,994
San Francisco	03/06/2022	3,190
San Francisco	07/09/2022	3,990

Notes:

(1) The table provides details of the number of observations in each city per scrape date.

Table 9: Summary of Property Details

Property Type	(%)
Entire Property	71.7
Private Room	27.8
Shared Room	0.5
Instantly Bookable	34.8
Long stays allowed	84.2
Facilities	(Mean)
Number of Bedrooms	1.6
Accommodates	3.5
Number of Bathrooms*	1.4
Number of Reviews	(%)
Less than 5	18.6
5 to 10	11.3
11 to 20	14.6
20 to 50	22.0
Greater than 50	33.6
Amenities**	(%)
Smoke Alarm	96.3
Essentials	92.1
Wifi	90.4
Kitchen	88.3
Hangers	85.4
Hot water	84.0
Hairdryer	83.6
Carbon Monoxide Detector	82.5
Dishes and Silverware	81.7
Refrigerator	75.7
Cooking Basics	75.2
Heating	73.3
Microwave	70.6
Shampoo	69.5
Coffee Maker	69.0
Bedlinen	68.7
Fire Extinguisher	60.6
First-aid Kit	53.0

Notes:

(1) The table provides details of the properties in my data.

∗ This includes both shared and private bathrooms.

∗∗ These are the most common phrases listed in the amenities section of each listing.

Table 10: Summary of Host Details

Host Race	(%)
White	39.7
Black	6.0
Other	23.6
Unknown*	30.8
Host Joined	(%)
2008	0.03
2009	0.3
2010	0.9
2011	3.0
2012	5.7
2013	7.1
2014	10.2
2015	12.9
2016	13.1
2017	9.6
2018	8.4
2019	8.9
2020	5.5
2021	7.7
2022	6.3
Characteristics	(%)
Super Host	45.8
Owner of Multiple Airbnb Properties	62.6
Verified Identity	88.0

Notes:

(1) The table provides details of the host characteristics in my data.

(2) The host race was calculated as described in Section (4).

∗ The race of a host is defined as Unknown if the profile picture is: (1) a picture of the property, a corporate logo, or some other object; (2) the face is obscured; or (3) the picture is of low quality, and cannot be analysed.

Table 11: Log Listing Price Differences and Race under Different Specifications

	(1)	(2)	(3)	(4)
Racial Minority	-0.042***	-0.032**	-0.042***	-0.032**
	(0.01)	(0.01)	(0.01)	(0.01)
Average Review Score				
Rated above 4.5	-0.131***	-0.139***	-	-
	(0.01)	(0.01)	-	-
Rated 3.5 to 4.5	-0.235***	-0.260***	-	-
	(0.02)	(0.02)	-	-
Rated below 3.5	-0.019	0.060	-	-
	(0.19)	(0.42)	-	-
Review Score and Value for money				
Above 4.5 Review and Above 4.5 Value	-	-	-0.123***	-0.134***
	-	-	(0.01)	(0.01)
Above 4.5 Review and 3.5 to 4.5 Value	-	-	-0.092***	-0.094***
	-	-	(0.02)	(0.02)
Above 4.5 Review and below 3.5 Value	-	-	1.525	2.275
	-	-	(1.01)	(1.45)
3.5 to 4.5 Review and Above 4.5 Value	-	-	-0.237***	-0.260***
	-	-	(0.02)	(0.03)
3.5 to 4.5 Review and 3.5 to 4.5 Value	-	-	-0.207***	-0.234***
	-	-	(0.02)	(0.02)
3.5 to 4.5 Review and below 3.5 Value	-	-	-0.347	-0.291
	-	-	(0.21)	(0.40)
Latest Qualitative Reviews				
One or more reviews below 4 star	-	-0.001	-	0.00
	-	(0.01)	-	(0.02)
All reviews 4 or 5 star	-	-0.035***	-	-0.038***
	-	(0.01)	-	(0.01)
All recent reviews 5 star	-	-0.008	-	-0.008
	-	(0.01)	-	(0.01)
Observations	70,334	63,153	70,334	63,153

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes:

(1) The table shows the results of the OLS regression of log listing price on the control variables: property characteristics as listed in Table (9); host characteristics as listed in Table (10); review score; value for money score; and both city, and neighbourhood fixed effects.

(2) Result (1) is the analysis of the specification per Laouénan & Rathelot (2022), and is included for comparative purposes only.

(3) Standard errors are clustered at the property level.

Table 12: Structural Parameter Estimates for Specifications (1) to (4)

	(1)	(2)	(3)	(4)
$\frac{\sigma_r^2}{\sigma_\nu^2}$	46.604***	29.037**	40.870***	28.256**
	(2.84)	(11.16)	(2.60)	(9.36)
$\frac{\sigma_r^2}{\sigma_{\tilde{r}}^2}$		9.853**		8.008**
		(3.94)		(3.30)
Observations	70,333	63,151	70,333	63,151

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes:

(1) The table provides details of the structural parameter estimates under Specifications (1) to (4).

Table 13: Reviews per month and Race

Racial Minority	-0.01
	(0.03)
Average Review Score	
Rated 3.5 to 4.5	-0.08***
	(0.01)
Rated below 3.5	-0.40***
	(0.04)
Host Characteristics	
Host Identity Verified	0.09***
	(0.018)
Host is a *superhost*	0.10***
	(0.01)
Host has multiple properties	-0.01**
	(0.01)
Property Characteristics	
Private Room	-0.04
	(0.04)
Shared Room	-0.29***
	(0.10)
Observations	70,334
R-Squared	0.229

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes:

(1) The table provides details of the OLS regression of reviews per month on the set of control variables listed in Table (11).

(2) Standard errors are clustered at the property level.

Table 14: The Value of Qualitative reviews - Simulations 1-5

	Simulation Number				
Qualitative Reviews	(1)	(2)	(3)	(4)	(5)
One or more below 4 star	-0.008	-0.006	-0.020**	-0.006	-0.012
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
All reviews 4 or 5 star	-0.015	-0.001	-0.003	-0.012	-0.017*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
All reviews 5 star	-0.010	-0.014*	-0.011	-0.011	-0.008
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)

Notes:

(1) The table shows the results of the OLS regression of log listing price on the simulated values for the last three qualitative reviews.

(2) The control variables included are: property characteristics as listed in Table (9); host characteristics as listed in Table (10); review score; value for money score; and both city, and neighbourhood fixed effects.

(3) Standard errors are clustered at the property level.

Table 15: The Value of Qualitative reviews - Simulations 6-10

	Simulation Number				
Qualitative Reviews	(6)	(7)	(8)	(9)	(10)
One or more below 4 star	-0.013	-0.012	-0.006	-0.013	0.004
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
All reviews 4 or 5 star	-0.005	-0.010	-0.021**	-0.002	-0.005
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
All reviews 5 star	-0.012	-0.010	-0.009	-0.012	-0.015*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)

Notes:

(1) The table shows the results of the OLS regression of log listing price on the simulated values for the last three qualitative reviews.

(2) The control variables included are: property characteristics as listed in Table (9); host characteristics as listed in Table (10); review score; value for money score; and both city, and neighbourhood fixed effects.

(3) Standard errors are clustered at the property level.

Table 16: The Importance of Race in Topic Prevalence (Part 1)

	Topic Number				
	(0)	(3)	(7)	(8)	(10)
Log Price	0.004***	-0.001***	-0.001***	-0.002***	-0.001**
	(0.0012)	(0.0003)	(0.0002)	(0.0005)	(0.0004)
Average Review Score					
Rated 3.5 to 4.5	0.000	0.001**	0.001***	-0.006	-0.012
	(0.0018)	(0.0004)	(0.0003)	(0.01)	(0.01)
Rated below 3.5	0.008	-0.003*	0.003***	0.001	0.002***
	(0.0070)	(0.0017)	(0.0010)	(0.0007)	(0.0006)
Host Characteristics					
Racial Minority	-0.005***	0.002***	0.001***	0.001**	0.001*
	(0.0017)	(0.0004)	(0.0002)	(0.0007)	(0.0005)
Superhost	-0.001	-0.000	-0.001***	-0.001	-0.002***
	(0.0011)	(0.0003)	(0.0002)	(0.0004)	(0.0004)
Property Characteristics					
Private Room	-0.020***	0.001	0.002***	0.003***	0.001
	(0.0031)	(0.0008)	(0.0004)	(0.0012)	(0.0010)
Shared Room	-0.005	0.004*	0.004***	0.009***	-0.001
	(0.0085)	(0.0021)	(0.0012)	(0.0034)	(0.0028)
Observations	47,175	47,175	47,175	47,175	47,175
R-Squared	0.043	0.039	0.054	0.033	0.040

Notes:

(1) The table shows the results of an OLS regression of the topic percentage per review on race and a number of control variables.

(2) The control variables included are: property characteristics as listed in Table (9); host characteristics as listed in Table (10); review score; value for money score; and both city, and neighbourhood fixed effects.

(3) The results are limited to those where racial minority is significant at the 10% level.

Table 17: The Importance of Race in Topic Prevalence (Part 2)

	Topic Number			
	(12)	(13)	(17)	(18)
Log Price	0.002***	0.002*	-0.001***	-0.001***
	(0.0007)	(0.0009)	(0.0003)	(0.0002)
Average Review Score				
Rated 3.5 to 4.5	-0.003***	0.000	0.001***	0.002***
	(0.0011)	(0.0014)	(0.0005)	(0.0004)
Rated below 3.5	-0.017***	-0.018***	-0.003	-0.002
	(0.0041)	(0.0054)	(0.0019)	(0.0014)
Host Characteristics				
Racial Minority	0.006***	0.002*	0.002***	0.001**
	(0.0010)	(0.0013)	(0.0004)	(0.0003)
Superhost	0.001*	-0.002***	-0.001***	-0.001***
	(0.0006)	(0.0008)	(0.0003)	(0.0002)
Property Characteristics				
Private Room	0.008***	0.009***	-0.002**	-0.001
	(0.0019)	(0.0024)	(0.0008)	(0.0006)
Shared Room	0.005	0.006	0.001	0.001
	(0.0050)	(0.0066)	(0.0023)	(0.0017)
Observations	47,175	47,175	47,175	47,175
R-Squared	0.059	0.051	0.034	0.034

Notes:

(1) The table shows the results of an OLS regression of the topic percentage per review on race and a number of control variables.

(2) The control variables included are: property characteristics as listed in Table (9); host characteristics as listed in Table (10); review score; value for money score; and both city, and neighbourhood fixed effects.

(3) The results are limited to those where racial minority is significant at the 10% level.

Table 18: Word List of Topics That Differ Most Across Racial Groups

Topic Number				
(0)	(3)	(12)	(13)	(17)
day	en	transport	easi	great
bite	alway	public	check	locat
room	prompt	brilliant	spot	recommend
get	di	link	super	would
apart	per	ideal	park	highli
night	avail	transit	great	love
good	con	flat	cute	clean
overal	answer	local	close	apart
one	el	close	access	commun
bathroom	pleasur	access	locat	amaz
work	para	situat	clean	excel
nice	het	connect	commun	wonder
bed	understand	well	conveni	definit
nois	al	good	nice	beauti
issue	lo	support	apart	accommod

Notes:

(1) The table provides the fifteen most popular words for five topics from the LDA model. The topics listed are those that differ the most substantially across racial groups after applying controls.

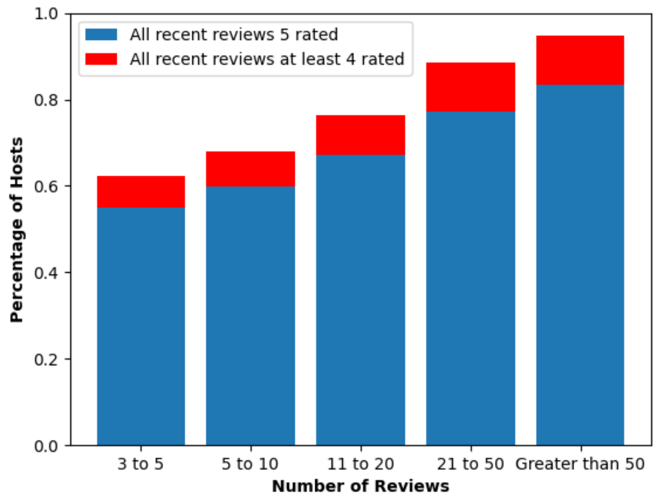
(2) The topics are a list of the words that occur together most frequently, and do not necessarily carry any collective meaning. As discussed in Appendix C.1, the 'words' in the table are the root of the words after applying the Porter Stemmer.

(3) Topic (3) appears to be predominantly non-English words. Although I removed exclusively non-English reviews, several bi-lingual reviews remained, and I conclude that they are marginally more likely to occur for Black hosts.

(4) Topics (12), (13) and (17) appear to relate to: transport links; location and check-in; and recommendation, respectively. These topics again occur with marginally more propensity for Black hosts. Finally, Topic (0) appears to be a topic with very limited interpretable meaning. This topic occurs more frequently in reviews of White hosts.

E. Figures

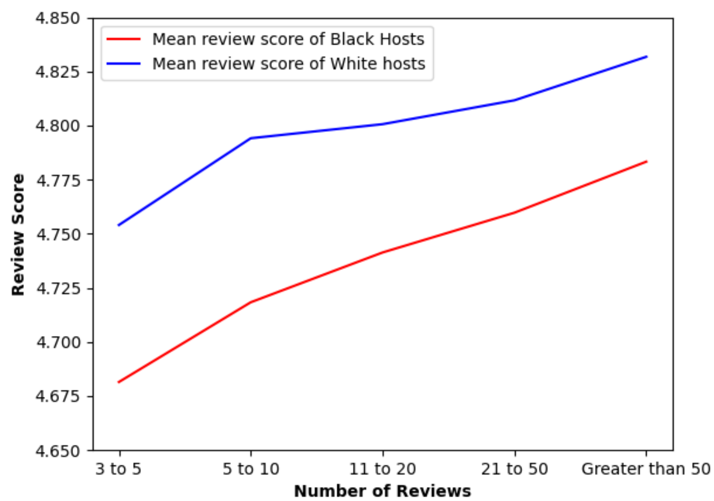
Figure 4: Percentage of Hosts with High Quality reviews per Number of Reviews



Notes:

(1) This figure provides a graphical illustration of the proportion of hosts with high quality reviews as the number of reviews increases.

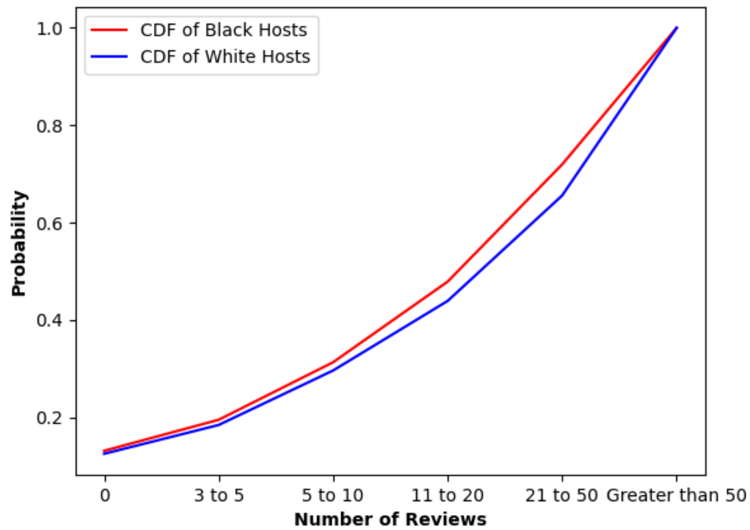
Figure 5: Average Review Scores of Black and White Hosts by Number of Reviews



Notes:

(1) This figure provides a graphical illustration of the difference in average review scores between White and Black hosts as the number of reviews increases.

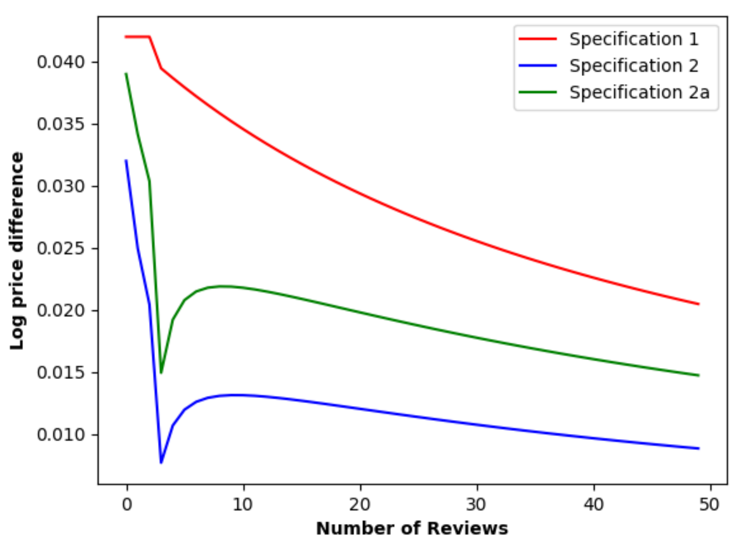
Figure 6: Cumulative Frequency of the Number of Reviews by Black and White Hosts



Notes:

(1) This figure provides a graphical illustration of the difference in the number of reviews between White and Black hosts.

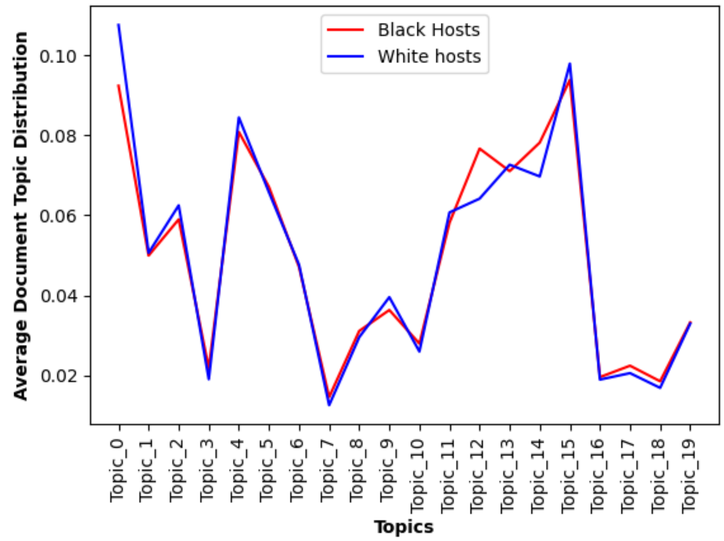
Figure 7: Log Price Difference between White and Black Hosts over Time



Notes:

(1) This figure provides a graphical illustration of the difference in log prices between White and Black hosts as the number of reviews increases.

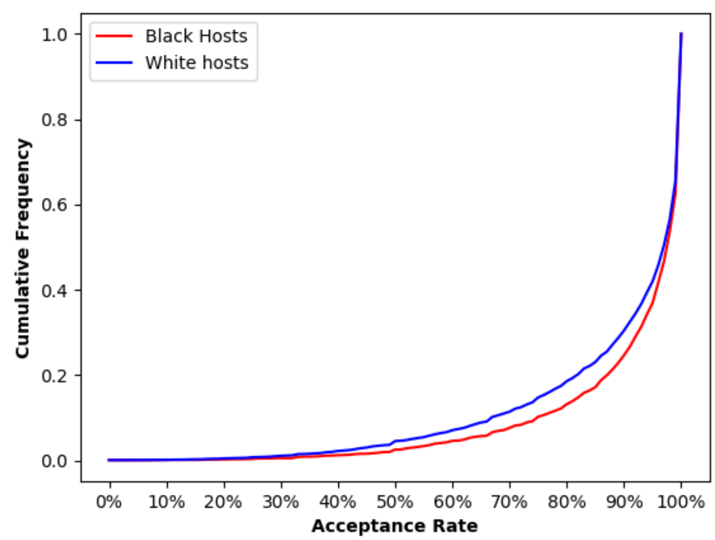
Figure 8: The Relative Importance of Topics by Host Race



Notes:

(1) This figure provides a graphical illustration of the difference in the topic distribution between White and Black hosts. The results presented are for a 20 topic LDA model trained on 169,443 reviews.

Figure 9: Host Acceptance Rate by Host Race



Notes:

(1) This figure provides a graphical illustration of the difference in host acceptance rates by host race.

F. Sample Airbnb Host Profile

TWO BEDROOMS DELUXE

★ 4.54 · 96 reviews · New York, United States

Share Save

Entire rental unit hosted by Sergo

6 guests · 2 bedrooms · 1 bed · 1 bathroom

Great communication

90% of recent guests rated Sergo 5-star in communication.

aircover

Every booking includes free protection from Host cancellations, listing inaccuracies, and other issues like trouble checking in.

Learn more

The apartment is located in a 4 family Townhouse. We

pagespages

are in the center of all shopping areas and major restaurants. The apartment is a 5 minutes walking distance to groceries including a whole food market or the metro station. The neighborhood is very friendly and you will certainly enjoy your stay in Harlem USA. We are...

Show more >

Where you'll sleep

Bedroom 1
1 queen bed

Bedroom 2
1 queen bed

What this place offers

- Kitchen
- Wifi
- Pets allowed
- TV
- Washing machine
- Dryer
- Air conditioning
- Hair dryer
- Long-term stays allowed

£241 night
★ 4.54 · 96 reviews

CHECK-IN 21/06/20...
CHECKOUT 24/06/20...
GUESTS 2 guests

Reserve

You won't be charged yet

£241 x 3 nights	£723
Cleaning fee	£41
Service fee	£151
Total	**£915**

Report this listing

Show all 15 amenities

3 nights in New York

21 Jun 2022 - 24 Jun 2022

June 2022

Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue
		~~1~~	~~2~~	~~3~~	~~4~~	~~5~~		
~~6~~	~~7~~	~~8~~	~~9~~	~~10~~	~~11~~	~~12~~	~~4~~	5
~~13~~	~~14~~	~~15~~	~~16~~	~~17~~	~~18~~	~~19~~	11	12
~~20~~	21	22	23	24	25	~~26~~	18	19
~~27~~	~~28~~	~~29~~	~~30~~				25	26

★ 4.54 · 96 reviews

Cleanliness	4.5
Accuracy	4.5
Communication	4.7
Location	4.8

Check-in	4.7
Value	4.5

Alanna
November 2019

Sergo's house is beautiful, clean, and he is a great & responsive host. Recommend!

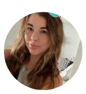
Satova
June 2022

ideal location, walking distance to the metro, starbucks, gym, grocery store, and restaurants. apartment is exactly as depicted in listing. spacious and has a nice view from an upper level apartment. the host is the best super helpful, nice disposition, and communicative. i could go on...

Show more >

Gerson
June 2022

Sergo was lovely and the communication was very good.
Everything in the apartment looked like the photos, very spacious and only 2 minutes walk to the subway.



Madison
May 2022

Really beautiful building, with exposed brick, spacious living area and kind host. Highly reccomend

Donovan
May 2022

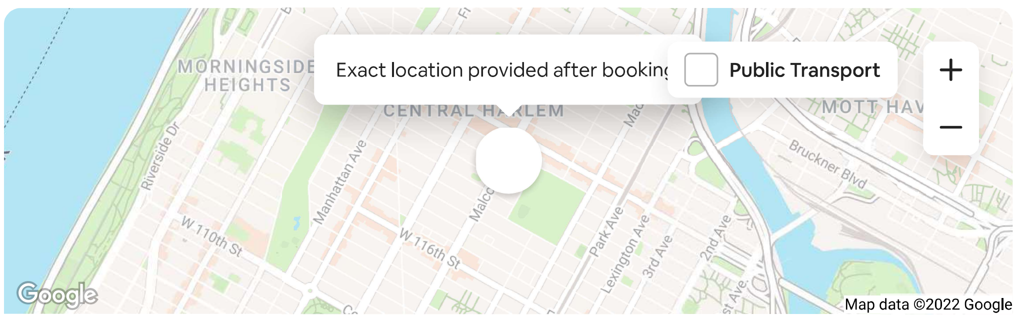
The place was amazing and the view of the street below was very clear. Bring k-cups for the coffee machine and avoid using the hot water boiler as it didn't quite work.

Match
May 2022

Nice place to stay.

Show all 96 reviews

Where you'll be



New York, United States

We have a Whole Foods Market 2 blocks away and you can walk to Central Park in 10 minutes. We are a Landmark neighborhood.

Show more >

Hosted by Sergo

Joined in March 2019 · Individual Host

★ 124 Reviews

Hello,
My name is Sergo Jeune and I have been a host in Harlem, New York City since May 2016. I have great reviews from guests and I take a lot of proud in hosting the guests. I have lived and worked in the City since 1982.
Thanks.

Thanks,
Sergo

During your stay

As your host, I am available on the property most of the times. I welcome your late arrivals and I will safeguards your luggage when you have late flight departures.

Response rate: 100%

Response time: within an hour

Contact host

To protect your payment, never transfer money or communicate outside of the Airbnb website or app.

Things to know

House rules

- Check-in: After 15:00
- No smoking
- No parties or events
- Pets are allowed

Health & safety

- Airbnb's COVID-19 safety practices apply
- Carbon monoxide alarm
- Smoke alarm

Show more >

Cancellation policy

Cancel before check-in on 21 Jun for a partial refund.

Review the Host's full cancellation policy which applies even if you cancel for illness or disruptions caused by COVID-19.

Show more >

G. Bibliography

Airbnb Form 10-k. (2021). *Retrieved from* `https://investors.airbnb.com/financials/default.aspx#sec` *(Accessed: 17 July 2022).*

Airbnb - Why Reviews Matter. (2019). *Retrieved from* `https://www.airbnb.co.uk/resources/hosting-homes/a/why-reviews-matter-41?_set_bev_on_new_domain=1658768178_NmRkM2ZjMWRlZmEx#:~:text==Your%20overall%20star%20rating%20appears%20on%20your%20listing,three%20reservations%20that%20total%20at%20least%20100%20nights.` *(Accessed: 18 January 2023).*

Altonji, J., & Pierret, C. (2001). Employer Learning and Statistical Discrimination. *Quarterly Journal of Economics: Vol. 116(1), 313-50.*

Ash, E., Chen, D., & Ornaghi, A. (2021). Gender Attitudes in the Judiciary: Evidence from U.S. Circuit Courts. *American Economic Journal: Applied Economics (Revise & Resubmit).*

Becker, G. (1957). The Economics of Discrimination. *Chicago: University of Chicago Press.*

Bendick, M. (2007). Situation Testing for Employment Discrimination in the United States of America. *Horizons Stratégiques: Vol. 3(5), 17-39.*

Bertrand, M., & Duflo, E. (2017). Field Experiments on Discrimination. *Handbook of Economic Field Experiments, Vol. 1, 309–93.*

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review: Vol. 94(4), 991-1013.*

Blei, D., Ng, A., & Jordan, M. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research, Vol. 3, 993-1022.*

Blei, D., & McAuliffe, J. (2007). Supervised Topic Models. *Advances in neural information processing systems, Vol. 20.*

Bohren, A., Imas, A., & Rosenberg, M. (2019). The Dynamics of Discrimination: Theory and Evidence. *American Economics Review: Vol. 109(10), 3395-3456.*

Cui, R., Li, J., & Zhang, D.(2020). Reducing Discrimination with Reviews in the Sharing Economy: Evidence from Field Experiments on Airbnb. *Management Science, Vol. 66(3), 1074-91.*

Devlin, J., Chang, M.W., Lee, K. & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805.*

Doleac, J., & Stein, L. (2013). The Visible Hand: Race and Online Market Outcomes. *Economic Journal, Vol. 123(572), F469-92.*

Edelman, B., Luca, M., & Svirsky, D. (2017). Racial Discrimination in the Sharing Economy : Evidence from a Field Experiment *American Economic Journal: Applied Economics, Vol. 9(2), 1-22.*

Edelman, B., & Luca, M. (2014). The Case of Airbnb.com *Harvard Business School Working Paper, No. 14-054.*

Engle, R., Giglio, S., Kelly, B., Lee, H., & Stroebel, J. (2020). Hedging Climate Change News.*The Review of Financial Studies, Vol.33 (3), 1184-1216.*

Fryer, G., & Levitt, S. (2004). The Causes and Consequences of Distinctly black Names. *Quarterly Journal of Economics, Vol. 119(3), 767-805.*

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as Data. *Journal of Economic Literature 2019, Vol. 57(3), 535–574.*

Gensim - Creating LDA Topic Model. *Available at:* `https://www.tutorialspoint.com/gensim/gensim_creating_lda_topic_model.htm` *(Accessed: 11 July 2022).*

Griffiths, T., & Steyvers, M. (2004). Finding Scientific Topics. *National Academy of Sciences Proceedings of the National Academy of Sciences - PNAS, Vol.101 (14), p.5228.*

Grogger, J. (2008). Speech Patterns and Racial Wage Inequality. *University of Chicago Harris School of Public Policy Studies Working Paper 0813.*

Guo, Y., Barnes, S., & Qiong, J. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism management, Vol.59, 467-483.*

Hawramani, I. (2015). The Arabic Baby Name Book; More than 5,000 Names for Boys and Girls. Independent publisher.

Kakar, V., Voelz, J., Wu, J., & Franco, J. (2021). The Visible Host: Does race guide Airbnb rental rates in San Francisco? *Journal of Housing Economics, Vol. 40, 25-40.*

Koffi, M. (2021). Innovative Ideas and Gender Inequality. *University of Waterloo, Canadian Labour Economics Forum (CLEF) Working Paper Series, No. 35..*

Lang, K., & Lehmann, J.K. (2012). Racial Discrimination in the Labor Market: Theory and Empirics. *Journal of Economic Literature, Vol. 50(40), 959-1006.*

Lang, K., & Manove, M. (2011). Education and Labor Market Discrimination. *American Economic Review, Vol. 101(4), 1467-96.*

Laouénan, M., & Rathelot, R. (2022). Can information reduce Ethnic Discrimination? Evidence from Airbnb. *American Economic Journal: Applied Economics, 14(1), 107-132.*

Luca, M. (2016). Reviews, Reputation, and Revenue: The Case of Yelp.com. *Harvard Business School Working Paper, No. 12-016.*

Moniz, A., & de Jong, F. (2014). Predicting the Impact of Central Bank Communications on Financial Market Investors' Interest Rate Expectations. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol.8798, 144-155.*

Mueller, H., & Rauh, C. (2018). Reading Between the Lines: Prediction of Political Violence Using Newspaper Text. *American Political Science Review, Vol. 112(2), 358–375.*

Murphy, K. (2012). Machine Learning: A Probabilistic Perspective. *Cambridge: MIT Press.*

Nosek, B., Smyth, F., Hansen, J., Devos, T., Lindner, N., Ranganath, K., Smith, C., Olson, K., Chugh, D., Greenwald, A. & Banaji, M. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology, Vol. 18, 36 – 88.*

Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment Classification using Machine Learning Techniques. *Proceedings of EMNLP 2002.*

Pang, B., & Lee, L. (2005). Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. *Proceedings of ACL 2005.*

Porter, M.F. (1980). An Algorithm for Suffix Stripping. *Program, Vol. 14(3), 130-7.*

Stephens-Davidowitz, S. (2014). The cost of racial animus on a black candidate: Evidence using Google search data. *Journal of Public Economics, Vol. 118, 26–40.*

Taddy, M. (2013). Measuring Political Sentiment on Twitter: Factor Optimal Design for Multinomial Inverse Regression. *Technometrics, Vol. 55(4), 415-425.*

Taigman, Y., Yang, M., Ranzato, M., & Wolf, L. (2014). Deepface: Closing the Gap to Human Level Performance in Face Verification. *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1701-1708*

Tirunillai, S. & Tellis. G. (2014). Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation. *Journal of Marketing Research, Vol. 51, 463-479.*

Wu, A. (2018). Gendered Language on the Economics Job Market Rumors Forum. *AEA Papers and Proceedings, Vol. 108, 175-179.*

Yu, C. & Margolin, D.(2022). Sharing inequalities: Racial discrimination in review acquisition on Airbnb.*New Media & Society : New Media & Society, 2022. Web.*