Diversity and Inclusion in the Sharing Economy: An Airbnb Case Study

Giovanni Quattrone,[1, 2] **Licia Capra** [3]

[1] Dept. of Computer Science, Middlesex University, United Kingdom
[2] Dept. of Computer Science, Univerity of Turin, Italy
[3] Dept. of Computer Science, University College London, United Kingdom

Abstract

The sharing economy model is a contested concept: on one hand, its proponents have praised it to be enabler of fair marketplaces, with all participants receiving equal opportunities; on the other hand, its detractors have criticised it for actually exacerbating preexisting societal inequalities. In this paper, we propose a scalable quantitative method to measure participants' *diversity* and *inclusion* in such marketplaces, with the aim to offer evidence to ground this debate. We apply the method to the case of the Airbnb hospitality service for the city of London, UK. Our findings reveal that diversity is high for gender, but not so for age and ethnicity. As for inclusion, we find strong signals of homophily both in terms of gender, age and ethnicity, thus suggesting that under-represented groups have significantly fewer opportunities to gain from this market model. Interestingly, the sentiment associated to same-group (homophilic) interactions is just as positive as that associated to heterophilic ones, even after controlling for Airbnb property's type, price and location. This suggests that increased diversity and inclusion are desirable not only for moral but also for economic and market reasons.

1 Introduction

Over the last decade, sharing economy platforms (Trenz, Frey, and Veit 2018) have been growing fast in cities worldwide, transforming the way in which goods, services, skills and spaces are shared. With no need for large upfront financial investments, and only requiring access to information technology to take part, the sharing economy model was envisaged to enable a more inclusive economic development than the traditional consumer model based on the concept of private ownership, bringing benefits to a broader spectrum of the population and ultimately uplifting society at large (Sundararajan 2016).

However, from the outset, detractors worried that platforms built around this economic model could cause more externalities than benefits, acting as large-scale online aggregators of well-known offline human dynamics and biases (Eckhardt and Bardhi 2015; Koh et al. 2019). Indeed, studies have emerged suggesting that a disproportionally high segment of participants in the Airbnb sharing economy platform are whites (Griswold 2016). Furthermore, in cities with high Airbnb presence, gentrification processes have been accelerating, contributing to the house affordability crisis of New York City and Los Angeles and exacerbating pre-existing societal divides, for example along racial lines (Wachsmuth and Weisler 2018; Lee 2016; Griffith 2017).

To help inform the debate between proponents and detractors of the sharing economy model around the specific issue of equal opportunities, in this paper we propose a scalable, easily reproducible, quantitative method to measure *diversity* and *inclusion* in sharing economy platforms. By *diversity*, we refer to the demographics of the peers that take part in a sharing economy platform (i.e., both hosts and guests in Airbnb). By *inclusion*, we mean the opportunities these peers have to prosper and benefit from this economic model. Note that achieving diversity does not automatically imply achieving inclusion too. For example bias, and more precisely what Mehrabi et al. define as *user interaction bias* (Mehrabi et al. 2021), may play against some specific demographics, despite these being equally represented within the platform (e.g., female hosts may be more frequently selected than male hosts, despite each representing 50% of the host market share in Airbnb). Being able to quantify diversity and inclusion may then inform what interventions to carry out, both from a technological (e.g., service features) and regulatory (e.g., legislation and taxation) point of view. This is important not only for moral reasons (i.e., no demographic segment should be left behind), but also for economic and market reasons (e.g., the broader the pool of people to tap into, the more economically efficient the service is). The method we propose consists of three main parts:

1. **Demographic Feature Extraction:** we leverage state-of-the-art face recognition AI tools to automatically extract demographics information (more specifically, *gender*, *age*, and *ethnicity*) from profile pictures of participants in sharing economy services. We measure how peers' representation along each of these features varies over time, as a way to study *diversity*.
2. **Preferential Attachment Estimation:** we model peer interactions as nodes and edges in a bi-partite graph, and use a statistical method based on network reshuffling to identify interactions that cannot be attributed to chance. We then measure preferential attachment along each demographic trait, as a way to capture *inclusion*; we do so while controlling for important geographic and economic

factors.

3. **Sentiment Analysis:** finally, we apply state-of-the-art NLP techniques to the reviews that peers leave to one another, and measure how sentiment varies across homophilic and heterophilic interactions, for each demographic traits.

As a case study, we then apply the method to measure diversity and inclusion of Airbnb in London, UK, analysing interactions spanning a period of 8 years (from January 2012 to December 2019).[1] Our findings reveal that:

Diversity: for both hosts and guests, gender diversity is high and consistently so over the years. Diversity is low for age and ethnicity instead, with White adults aged 30–39 prevailing among both hosts and guests. The platform composition is changing though, with young adults aged 18–29 being on the rise, as well as Black and Asian communities; older adults aged 40+ are on the decline instead.

Inclusion: homophilic stays dominate along all demographic features, even after controlling for Airbnb property type, price and location. In particular, we found that homophily is exceptionally strong along the age line for the 40+ year old group, and along the ethnicity line both for Asian and Black groups. These demographics are also the least represented within the Airbnb user base. Strong homophilic interactions of host-guest interactions within the same minority groups imply overall fewer opportunities for under represented groups. Interestingly, sentiment is not correlated with homophily, suggesting there is room to improve inclusion without negative consequences on service satisfaction.

The remainder of the paper is structured as follows: we first provide an overview of the main studies conducted in recent years of the sharing economy as a whole and Airbnb in particular (Section 2). We then describe the method we have developed in detail (Section 3), before delving into an extensive analysis of the results obtained when applying the method to the case study of Airbnb in London (Section 4). We discuss the implications of our findings, as well as their limitations (Section 5), before concluding the paper elaborating on future directions of investigation (Section 6).

2 Related Work

Even though *sharing* is an old practice, the *sharing economy* (SE) specifically is a rather recent phenomenon that has been propelled by information and communication technologies. In recent years, both the number of companies that adopt this market model, and the business value of its key players, has grown exponentially; this has been accompanied by an explosion in the number of academic studies of such phenomenon, from the point of view of different disciplines (e.g., Law, Economics, Sociology) (Hossain 2020). Scholars in the computing field have themselves been very actively studying the interplay between SE and computing (Dillahunt et al. 2017), both from a technical point of view, investigating for example the algorithms used for efficient provider/consumer matchmaking, and from a *socio-technical* one, where the interplay between the computing platform that mediates SE interactions, and the individuals accessing the SE service, are being investigated. Our work falls within this latter stream.

A predominant research question within this stream has been motivation, trying to explain why peers choose to participate (or not) in the sharing economy. Studies pertaining over 40 different sharing economy services repeatedly found motivation to span widely, from more idealistic reasons (e.g., altruism, social connection) to more instrumental ones (e.g., value) (Bellotti et al. 2015; Shih et al. 2015). Different reasons were also found to be behind the choice not to engage in these platforms: from issues of safety and (dis)trust, to issues of independence and autonomy (or lack thereof) (Dillahunt and Malone 2015; Meurer et al. 2014). Most of these studies used a *qualitative analysis approach* based on structured interviews and focus groups with a few selected participants; these methods afforded depth in the investigation of specific aspects or questions, to the expense of scale.

As large amounts of readily-available data capturing the interactions of participants in SE platforms are becoming available, *quantitative analysis approaches* have started to complement qualitative ones, so to analyse SE platforms at scale (both in terms of number of participants, geographic coverage, and over time). The vast majority of such quantitative studies have been about Airbnb, largely due to data availability (indeed, studies of Airbnb have been proliferating so quickly to warrant a dedicated systematic literature review solely on such platform (Dann, Teubner, and Weinhardt 2019)). One stream of quantitative studies research has tried to investigate the *relationship between Airbnb and the city* it penetrated: for example, a study of 13 cities in Italy (Picascia, Romano, and Teobaldi 2017) found the absolute majority of Airbnb listings to be concentrated in historical centres, a trend that has been increasing steadily over the years and that was linked to the 'social desertification' of Italian historical centres. A Texas-focused study (Zervas, Proserpio, and Byers 2015) found causal evidence of the role that Airbnb played in the loss of hotel revenues. Other studies were however supportive of Airbnb in their findings: a UK-focused study (Tussyadiah, Liu, and Steinmetz 2020) tried to quantify the impact of Airbnb penetration on the local environment, and found a higher preponderance of positive effects (e.g., economic impact on the local community) than negative ones (e.g., fear of increased local crime). When zooming in London, UK, another study found strong evidence of the impact that Airbnb penetration had on increasing the value of real estate properties (Benitez-Aurioles and Tussyadiah 2020). All these works offered evidence in the emerging debates between those in favour and those against SE services, ultimately aiming to help regulating such services. As SE services keep growing, other researchers have also looked at how SE platforms had been

[1] We purposely excluded the years 2020-onward from the present study, since the COVID-19 pandemic and its associated travel restrictions have nearly halted Airbnb activity during this time period.

evolving over time, to understand what geographic, demographic and socio-economic factors are mostly associated with Airbnb adoption. Findings for several US cities of different size, wealth and population composition (Quattrone et al. 2018), and for the city of London, UK, (Quattrone et al. 2016) showed some strong similarities: areas of high Airbnb presence were usually those close to city centres and occupied by the 'talented and creative' classes; however, there were also important differences, with Airbnb penetration in London growing over time in more income deprived areas – a phenomenon that did not occur in any of the US cities analysed, thus signalling important geographic differences in the adoption of this platform.

Another stream of quantitative studies research has looked into the *relationship between Airbnb and its participants* instead. For example, user satisfaction with the SE service has been studied extensively, by means of statistical and linguistic analysis of ratings and reviews (Zervas, Proserpio, and Byers 2015; Zhu, Lin, and Cheng 2020; Santos et al. 2020; Bridges and Vásquez 2018; Alsudais and Teubner 2019). Several studies have ensued that have gone beyond sentiment analysis and looked into what participants in this hospitality service care the most about (Luo and Tang 2019; Sutherland and Kiatkawsin 2020; Joseph and Varghese 2019; Lee et al. 2020; Cheng and Jin 2019; Luo 2018; Quattrone et al. 2020). Business-oriented aspects such as property location, facilities/amenities, and communication with the hosts were consistently found to be particularly important in all studies. Interestingly, the more social-oriented aspects of the Airbnb service (i.e., interactions between hosts and guests) steadily lost importance over time (Quattrone et al. 2020), suggesting that Airbnb has been increasingly evolving into a 'platform economy' (Moazed and Johnson 2016) as opposed to a 'sharing economy' one (Trenz, Frey, and Veit 2018). Recent studies have also started to delve deeper into the patterns of interactions between Airbnb hosts and guests (i.e., who stays with whom), after several SE platforms had come under fire for episodes of discrimination: for example, in ride-sharing platforms Uber and Lyft, female passengers were found to be taken on disproportionately longer and more expensive routes, and passengers with African American-sounding names were found to be twice as likely to receive trip cancellations compared to passengers with White-sounding ones (Ge et al. 2016); in online freelance marketplaces such as TaskRabbit and Fiver, gender and race were found to have an impact on worker evaluations (Hannák et al. 2017); and Airbnb was no exception, with Airbnb hosts found to be turning down potential guests based on their racial background (Edelman and Luca 2014).

To what extent are SE platforms such as Airbnb the fair and inclusive marketplaces they aim to be, and to what extent are they becoming online aggregators of offline human biases? To answer this question, recent scholarly work has tried to characterise the interactions that prevail in such platforms. For example, a study of five Western cities within the Airbnb platform found evidence of gender, race and age homophily in Airbnb (Koh et al. 2019); these findings echoed a vast literature that found homophily – i.e., the tendency of individuals to preferentially attach/interact with similar ones – to prevail along demographics lines in many online (e.g., social media) platforms. Specifically, gender, age and race homophily was found in MySpace (Thelwall 2009); race homophily was found in Facebook (Wimmer and Lewis 2010), gender homophily in Twitter (Bamman, Eisenstein, and Schnoebelen 2014), and age homophily was documented in a study of the Facebook social graph (Ugander et al. 2011). While avoidance is universally deemed as unacceptable, some studies claim that homophily may have desirable consequences instead: for example, it can facilitate access to information (Choudhury et al. 2010), the diffusion of innovations and behaviors (Christakis and Fowler 2007), as well as the formation of social norms (Centola, Willer, and Macy 2005). However, there is also ample evidence that homophily can lead to negative consequences too, including exacerbated polarization of opinions, persistence of disagreements, and the perpetration of self segregation between groups (Golub and Jackson 2012; Centola et al. 2007). In the sharing economy context, homophilic interactions risk pushing certain demographics completely out of this business model, contrary to the SE original aim of inclusiveness and democracy; this is particularly the case if certain demographics are under-represented to begin with – as reports suggest (Griswold 2016).

To investigate the issue, we present next a computational method to study diversity and inclusion in SE platforms longitudinally and at scale.

3 Data & Method

Sharing economy services rely upon trust between participants for the 'sharing' to take place. To build such trust, most SE platforms leverage two main mechanisms: first, participants have publicly visible profiles, usually comprising the peer's *profile picture*; second, upon completion of an interaction, peers leave publicly available *reviews*, describing how satisfied they have been with the service.

Our method exploits these *publicly available data*: first, we use state-of-the-art AI face recognition tools to extract demographic features from profile pictures, then quantify diversity of hosts and guests in terms of gender, age, and ethnicity. Second, we use reviews to reconstruct an empirical bipartite graph of peer interactions, then use a statistical method based on network rewiring to quantify preferential attachment along each demographic line, and use it to study inclusion. Finally, we apply state-of-the-art NLP techniques on the content of reviews to analyse peers' opinions of their interactions, along each demographic trait. We describe the details of each step of our method next. A GitHub repository containing both the source code and the (anonymised) data used in this work have been made available for transparency and reproducibility.[2]

Face Recognition Tool Selection

Our method relies on the use of AI face recognition tools to accurately extract users' demographics (i.e., gender, age,

[2] https://tinyurl.com/3snadsra

ethnicity) from profile pictures. Several candidate tools exist, both *proprietary* and *open-source*. The former have been subject to several benchmark studies aimed at ascertaining their accuracy (Buolamwini and Gebru 2018; Raji and Buolamwini 2019). These studies have found Microsoft Azure Face API[3] and Amazon Rekognition API[4] to have significantly higher accuracy than any competitor when estimating gender and age; however, neither of them infers ethnicity from profile pictures. One of the very few proprietary tools that does so is Clarifai.[5] Its accuracy for ethnicity estimation is however not well documented in the literature. A common disadvantage of proprietary tools is their lack of transparency; furthermore, the financial cost associated to them negatively impacts reproducibility of results. An alternative to proprietary tools is offered by open-source ones: these tools afford easy reproducibility, better transparency and customisation than proprietary tools; however, their accuracy is not well documented.

To inform the choice of what tool(s) to rely on for demographics' extraction from profile pictures, we thus conducted a technology assessment study. We chose as candidate proprietary tools the Microsoft Azure Face API for gender and age estimation, and the Clarifai API for ethnicity; we favoured the Microsoft Azure Face API as proprietary tool against the Amazon Rekognition one due to lower cost and faster processing. We selected DeepFace[6] as candidate open-source tool for all demographic features under study, due its current popularity. We then performed the following investigations.

(a) Proprietary vs Open-source Tools for Gender and Age. We selected 300 Airbnb host and guest profile pictures at random, and asked three independent annotators to label them for gender (i.e., male, female) and age group (i.e., 18–29, 30–39, 40+). They got an overall Cohen's kappa coefficient equal to 0.95 for the former and 0.81 for the latter, indicating a substantial agreement among them. We then ran the Microsoft Azure Face API and the DeepFace API for gender and age on such pictures, and measured their accuracy compared to the 'ground truth' (defined as majority vote among the three annotators). With the Microsoft Azure Face API, we obtained very high accuracy: 0.81 for gender and 0.74 for age. This was expected, based on past benchmark studies of such tool. With DeepFace, we achieved significantly lower accuracy instead: 0.57 for gender and 0.32 for age. Based on these results, we deemed the accuracy of the selected open-source tool not good enough to be used as-is in our study, and thus selected the proprietary Microsoft Azure Face API as tool of choice for gender and age estimation instead.

(b) Proprietary vs Open-source Tools for Ethnicity. The literature is very sparse in terms of benchmark studies measuring the accuracy of AI face recognition tools to estimate ethnicity. We thus decided to conduct a more extensive investigation with respect to this demographic feature. First, for the same 300 Airbnb profile pictures used before, we asked three independent annotators to label them for ethnicity. Note that there is no agreement in the literature in terms of how many ethnic groups exist; for the purpose of this study, we asked annotators to select among the following six (since these are the outputs of the selected AI tools): Asian, Black, Indian, Latino, Middle Eastern, White. We obtained an overall Cohen's kappa coefficient equal to 0.74, again indicating substantial agreement among annotators. We then ran the Clarifai API and the DeepFace API for ethnicity on such pictures, and obtained an accuracy of 0.67 for the former, and 0.41 for the latter. As was the case for gender and age, the chosen proprietary tool reports significantly higher ethnicity estimation accuracy than the chosen open-source one; however, the gap is not as big as with other demographic features. Also, this study relied on the ethnicity annotations of three individuals that, while spanning different age groups and genders, they all self-identify as White; we were thus concerned that bias could have been introduced in the 'ground-truth' labeling step. To gain confidence in our technology assessment results, we then conducted a second ethnicity classification study, this time using the Chicago Face Database (Ma, Correll, and Wittenbrink 2015). This is a dataset curated by the University of Chicago, comprising images of 597 male and female models that are self-identified as Asian (18%), Black (33%), Latino (18%), and White (30%). Pictures in this benchmark dataset have been taken with individuals front-facing the camera, against a neutral background, with no clutter and in ideal lighting conditions. Although profile pictures in Airbnb are possibly much more challenging for AI face recognition tools, experiments against the Chicago Face Database can offer a valuable upper-bound in terms of estimation accuracy for ethnicity. Once again, we found the overall accuracy of the Clarifai API to be higher than that of DeepFace (0.83 versus 0.77). Based on these results, we chose to proceed using the proprietary Clarifai API tool for ethnicity estimation. Upon closer inspection, we noted that some ethnic groups where significantly easier to detect than others: with reference to the curated Chicago Face Database, accuracy for Asian/Black/White groups was $0.97/0.90/0.80$ respectively, but only 0.59 for Latino. In the reminder of this paper, we will focus our analysis and discussion on the three ethnic groups that the Clarifai tool can most confidently estimate (i.e., Asian, Black and White). For completeness, we will also report results on the other ethnic groups, combining results for Clarifai's Indian, Latino, and Middle Eastern groups into a single category labeled 'Others' and keeping in mind the degree of uncertainty arising from these demographic estimates.

Data

We chose London (UK) as case study for this work for two main reasons: London has been one of the first cities in the world with very large Airbnb penetration, enabling us to perform both a large-scale and long-running study of the platform; furthermore, it is home to one of the most ethnically diverse populations in the world, with about one third of

[3]https://azure.microsoft.com/en-us/services/cognitive-services/face/

[4]https://aws.amazon.com/rekognition/

[5]https://www.clarifai.com/models/ai-face-detection

[6]https://github.com/serengil/deepface

	Guests	*Hosts*
Number of . . .	106,736	14,012
Age distribution	18–29 (35%), 30–39 (36%), 40+ (27%)	18–29 (24%), 30–39 (46%), 40+ (29%)
Gender distribution	Female (55%), Male (45%)	Female (56%), Male (44%)
Ethnic distribution	Asian (12%), Black (3%), Others (20%), White (64%)	Asian (5%), Black (6%), Others (28%), White (61%)

Table 1: Demographics of Airbnb hosts and guests

Londoners being born overseas and over 200 languages being spoken in the capital (for National Statistics 2022), making London particularly suited for a study of diversity and inclusion. We scraped the Airbnb website for the city of London (UK) in June 2020 and specifically collected four sets of data: for each property *listing*, we collected its type (e.g., whether a full property, a private room or a shared room), its price per night, and its location (in terms of the London neighbourhood it is located in).[7] From the identifiers of London Airbnb listings, we were then able to retrieve information about London *hosts*, including their profile pictures. From the identifiers of London Airbnb listings, we were also able to retrieve all *reviews* that followed a stay; in particular, we collected the date the review was written, the text, and the identifier of the guest who left it. Finally, using these guest identifiers, we scraped *guests* profile pictures.

Our initial dataset contained 54k unique hosts, 1.3M unique guests, 86k unique listings, and 1.5M unique reviews, dating from 2009 to June 2020. For the purpose of this study, we only kept 8 years of data, from January 2012 to December 2019: we removed the first two years so to focus on the 'majority' phase of platform adoption (as opposed to the innovators/early adopters phases); we also removed the data from January 2020 onward, to discard the time period during which Airbnb usage came to almost a halt due to COVID-19.

Extracting user demographics (i.e., gender, age, ethnicity) from profile pictures has a significant cost, both financially and in terms of processing time (as discussed in the previous section, at present the most accurate state-of-art AI face recognition tools are proprietary ones, accessible via a programmatic API). We thus chose to sample our data, and in particular selected a random sample of 150k reviews (i.e., 10% of the original data). This led us to keep 17k unique hosts, 122k unique guests, and 19k unique listings.

After running the Microsoft Azure Face API and Clarifai on our 10% sample data, we further removed all guests and hosts for which demographics could not be estimated from their profile picture (e.g., some pictures do not contain a person but for example a pet; others contain multiple people). We ended up with a dataset comprising 14k Airbnb hosts and 106k Airbnb guests (writing 147k reviews). The demographics composition of such dataset is shown in Table 1.

[7]The boundaries of such neighbourhoods have been determined by Airbnb based on research with cartographers, locals and city experts. To read more: https://www.airbnb.co.uk/help/article/420/what-are-neighbourhoods.

Preferential Attachment Estimation

Our approach to studying inclusion begins with the construction of a bipartite graph, where nodes represent Airbnb hosts and guests, and an edge from guest g to host h represents a stay (i.e., guest g stayed in a property listed by host h, as inferred from a review that g left). Each graph node is associated with three labels representing the node's gender, age, and ethnicity. To investigate interactions across each demographic trait that cannot be attributed to chance, we employ a statistical network shuffling method. We begin by assuming a 'null hypothesis' that any given guest-host pairing occurs at random. We then create a large set of 'null network models' by shuffling the labels of both guest and host nodes of the empirical network, and finally comparing the properties of these null network models against those observed in the empirical one.

Our method draws inspiration from (Koh et al. 2019), which uses a random 'shuffling' step to create null-network models. However, our approach significantly improves upon it. Specifically, the method proposed in (Koh et al. 2019) shuffled the *edges* of the bipartite graph while maintaining the total number of stays of each guest and host, so to preserve the correlations between demographic characteristics and activity in Airbnb. Nevertheless, this method ignored important geographic and economic factors that could significantly influence a guest's choice of where to stay. To address this limitation, our proposed method shuffles the *node labels* of the empirical graph, while keeping the network structure intact. During the node labels' shuffling process, we then apply a further restriction to preserve crucial geographic and economic variables associated with each stay, in particular property type (full or shared), property location, and property price (divided into quartiles).

To get statistical significant results, we generated 100 randomised network configurations for each null hypothesis to test, using the random shuffling procedure above. After each round of random shuffling, we compared the empirical labels with the shuffled labels, with the aim to ensure no correlation between them. In particular, we required a Cohen's kappa score between the empirical labels and the shuffled labels of less than 0.1 for each shuffled configuration to be accepted as part of our random set.

For each demographic feature under study (i.e., gender, age, ethnicity), we counted the number of interactions between guests and hosts for each possible combination of values (i.e., for gender, we counted the number of male-male, female-female, male-female, and female-male stays). We did so both in the empirical network and in each of the 100 corresponding randomised configuration. For each combination, we then defined its corresponding *preferential at-*

tachment (pa) value as the relative difference between the number of interactions between guests and hosts with that particular demographic feature pairing in the actual graph (denoted as $actual_{pair}$), and the number of interactions between guests and hosts with that same pairing in the randomised graph (denoted as $null_{pair}$). Specifically,

$$pa = \frac{actual_{pair} - null_{pair}}{null_{pair}}$$

A preferential attachment value of 0 indicates that the interactions between guests and hosts in the empirical network are in line with what would be expected in a randomised network configuration. In contrast, values greater or lower than 0 indicate a deviation from the expected randomised network configuration. The magnitude of the preferential attachment value (in absolute terms) reflects the degree of deviation from the randomised network configuration.

Note that, for each demographic feature under analysis, we compute 100 pa values (one for each randomised network); in our Results section, we will then show values ranging from the lower 5% percentile to the upper 95% percentile, in order to show if (and by how much) the empirical network differs from 90% of our randomised configurations.

Sentiment Analysis

Our final step is to measure how sentiment varies between homophilic and heterophilic interactions, for each demographic group. We do so in two steps.

First, we use the Valence Aware Dictionary for Sentiment Reasoning (VADER) algorithm (Hutto and Gilbert 2014) to compute the sentiment associated to each group of reviews. Since VADER has been shown to work particularly well on short text, we first break down each review into sentences and run the VADER sentiment analyser on each of these; we then compute the sentiment of a review as the mean sentiment of its sentences, and derive the sentiment of a group of interactions as the mean sentiment of its associated reviews. Note that VADER only performs sentiment analysis on English text. For London, 87% of Airbnb reviews are written in English, so for the purpose of conducting sentiment analysis we discarded the remaining 13%. We further discarded English reviews that were either too short (less than 8 words) or too long (more than 175 words – 97.5th percentile), to avoid skewing results. In total, 15% of the original set of reviews have been removed from this sentiment analysis computation.

Finally, we defined the *sentiment gain* ($sent\_gain$) for each demographic feature guest/host pairing as the relative difference between the average sentiment of stays between guests and hosts with that demographic feature pairing in the actual graph (denoted as $sent_{pair}$) and the average sentiment of all stays in the same graph (denoted as $sent_{all}$). Specifically,

$$sent\_gain = \frac{sent_{pair} - sent_{all}}{sent_{all}}$$

4 Results

Diversity

To study diversity of Airbnb hosts and guests, we calculated the share of Airbnb users that fall into each demographic group, for each year from 2012 to 2019. Diversity of Airbnb guests and hosts in London for gender is very high: as shown on the left side of Figure 1a, there is an almost perfectly-even split between male and female guests (46% vs 54% respectively in 2019). The same holds when we look at hosts (right side of Figure 1a), with only a marginally higher presence of females (53%) than males (47%). Note that, assuming the majority of London hosts are living in London (and thus captured by the London census), a slightly higher prevalence of female hosts tallies with the overall London population having a slightly higher proportion of females than males (London Datastore 2020). High diversity of Airbnb users along the gender line started from the early days of platform adoption (e.g., in 2012, there were 51% vs 49% female/male guests, and 57% vs 43% female/male hosts), and has not changed much over time.

Diversity along the age line depicts a rather different situation. Let us consider guests first: as shown on the left side of Figure 1b, in the early years of Airbnb guests travelling to London, the 30 to 39 age group dominated the scene (e.g., 44% in 2012); however, the younger segment of 18 to 29 has been steadily on the rise (from 20% in 2012 to 37% in 2019), ultimately even surpassing the 30 to 39 age group. On the other hand, an increasing proportion of 40+ travellers seems to be disengaging with the hospitality platform (from 35% in 2012, down to 25% in 2019). Let us now turn our attention to Airbnb hosts (right side of Figure 1b). In 2019, there was a disproportionately high prevalence of 30–39 year old Airbnb hosts (48%), compared to the older 40+ (30%) and especially the younger 18–29 year old (21%) segments. Unlike what we observed for guests, the dominance of the 30–39 age group for hosts has not been faltering over time (from 48% in 2012 to 44% in 2019), and although the younger segment has been increasing its presence over time (from 21% in 2012 to 30% in 2019), the older segment of the population has steadily disengaged instead (from 30% in 2012 to 25% in 2019). If we consider that, in 2019, the London population was made of approximately 18% of people aged 18 to 29, 21% of people aged 30 to 39, and 30% of people aged 40 to 65 (London Datastore 2020), our data suggests that the 30–39 segment has disproportionately high representation in the Airbnb host base; the 40+ segment has a lower-than-expected representation instead, while the youngest age group is hovering just above what one might expect (relative to the London population).

We finally turn our attention to diversity along the ethnicity line. Before we do so, recall that the AI face recognition tool we used to extract ethnicity information from profile pictures exhibited high accuracy for Asian, Black and White groups only; accuracy for Indian, Latino and Middle Eastern groups (next referred to as a combined 'Others' group) was lower instead. Despite the inaccuracies that our ethnicity estimation bears, some noteworthy patterns can still be detected with some confidence. In particular, as shown on the left side of Figure 1c, White guests currently dominate the scene, with 64% of Airbnb guests in London being White, 12% being Asian and only 4% Black; such ethnic composition of guests has remained relatively unchanged between 2012 and 2019. Ethnic diversity of hosts depicts perhaps a more encouraging picture (right side of Figure 1c): White hosts still dominate the scene, but their market share has been diminishing over the years (from 67% in 2012 to 49% in 2019), with mixed-race (Others) and Black hosts slowly but steadily growing; Asian hosts were few in the early days and remained so over time. To better interpret these numbers, we took a look again at the London population statistics of the last 20 years (London Datastore 2020), and found that the London population of White has been slowly decreasing over time, and is currently below 60%; Asian have instead been growing rapidly, and partly Black too: these ethnic groups currently represent approximately 20% and 15% of the London population respectively. If we revisit host ethnicity ratios in Airbnb (right side of Figure 1c), in light of the population statistics just mentioned, it is not surprising to see presence of White hosts at around 50% (and slightly declining over time), nor to see Black hosts accounting for approximately 12% (and steadily increasing). What is surprising though is to see that Asian, the second largest and fastest growing ethnic group in London, still represents only 6% of the Airbnb host population.

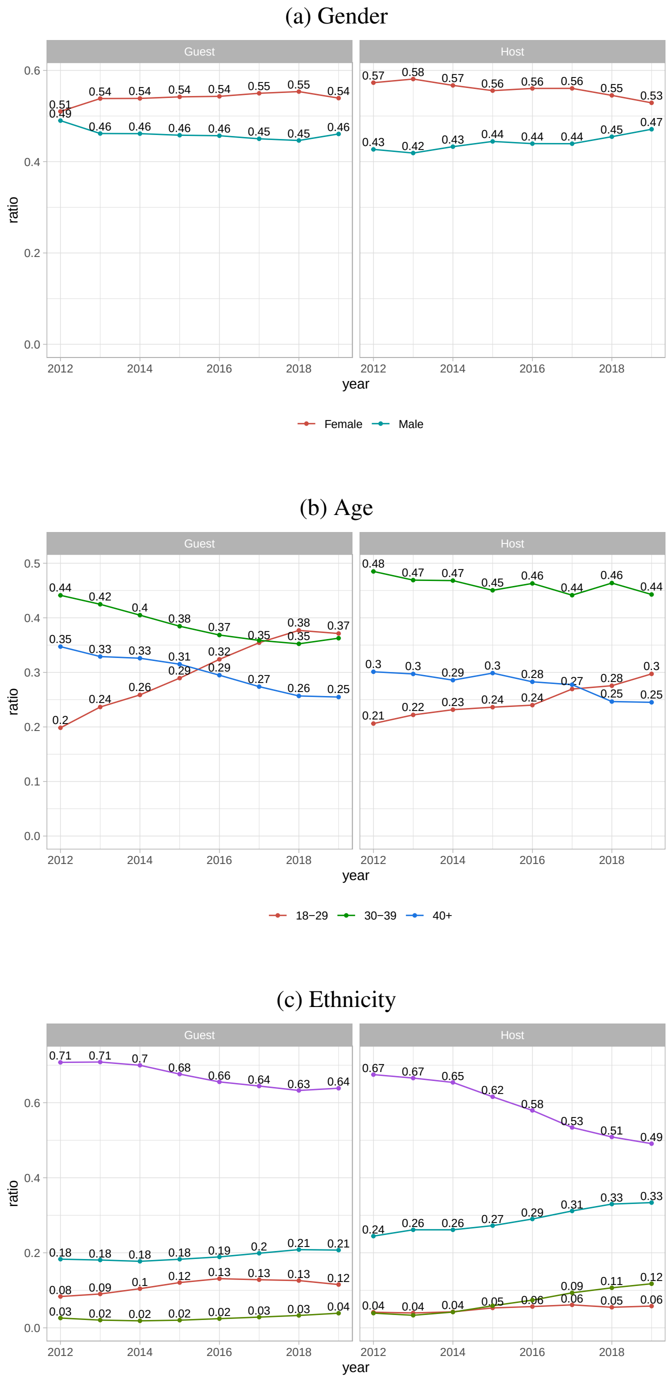

Figure 1: Airbnb users' demographics from January 2012 to December 2019

Inclusion

Having studied the composition of the Airbnb user base for both guests and hosts, along the main demographic features (i.e., gender, age and ethnicity), we now turn our attention to *inclusion*, that is, the actual interactions that these peers have. We infer inclusion of different Airbnb demographics using the *preferential attachment* formulation presented in Section 3. For conciseness, we only present results for the aggregate time period from 2012 to 2019; we also analysed inclusion year-by-year, but observed no significant temporal variation, especially from 2015 onwards.

To begin with, we analyse host/guest interactions focusing on gender. Previous studies (Koh et al. 2019) had revealed a slight prevalence of gender homophilic interactions in Airbnb; our results confirm and expand this (Figure 2a), showing gender homophily is present even after controlling for various exogenous factors (i.e., property type, price and location). More precisely, male guests are preferentially attached to male hosts (median $pa = 0.05$), and female guests are preferentially attached to female hosts (median $pa = 0.04$). This is despite the number of 'shared room' properties in Airbnb London being less than 1% – indeed, nearly 40% are private rooms and 60% entire homes/apartments (the latter suggesting that hosts and guests do not actually spend much time together). Because females and males are almost equally represented in Airbnb both in terms of hosts and guests (i.e., high gender diversity), inclusion is not hampered by a prevalence of homophilic stays.

We now consider host/guest stays focusing on age groups. As it emerges from Figure 2b, there is a greater issue in terms of inclusion for the elder age group of 40+. Indeed, we previously observed this is the least represented group in terms of age, both among hosts and guests; on top of that, we now observe they mostly engage in homophilic stays too. Specifically, 40+ guests preferentially stay with hosts aged

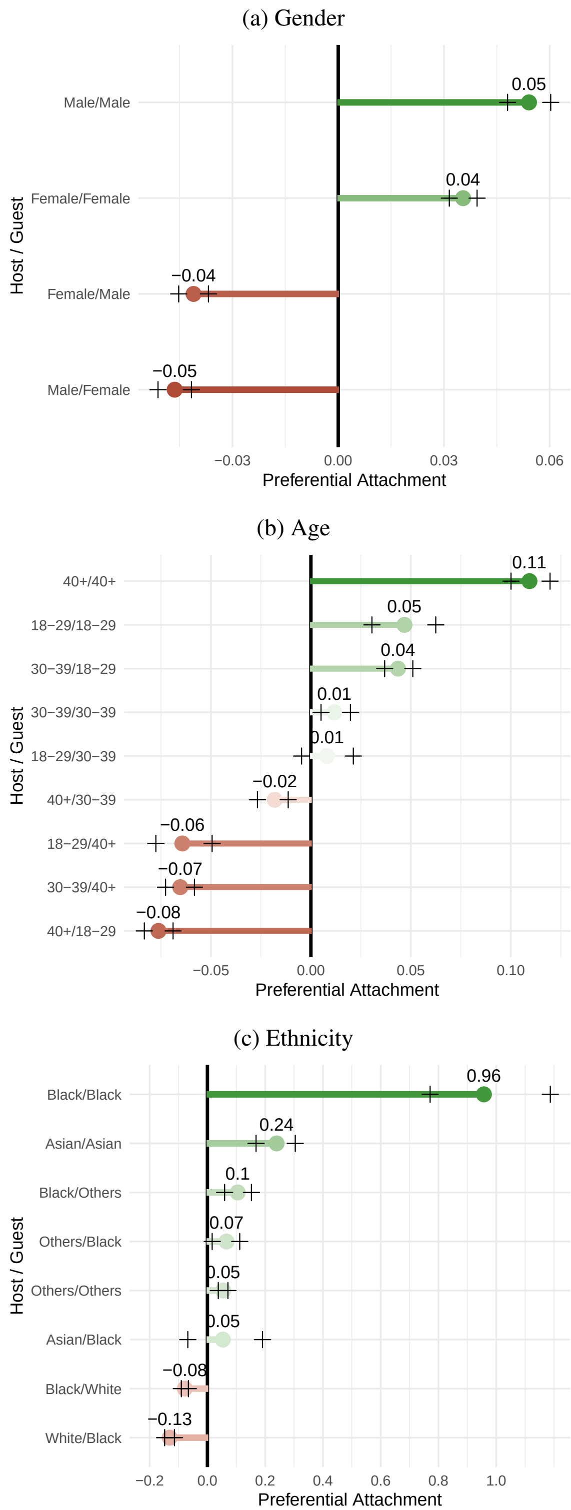

Figure 2: Preferential attachment between hosts and guests by gender, age and ethnicity for the aggregate period January 2012 – December 2019. Median values are shown as dots, bottom 5% and top 95% percentiles are shown with a $+$. To enhance readability, Figure 2c displays the top eight preferential attachment pairings (in absolute values) only.

40+ too (median $pa = 0.11$, being it the highest preferential attachment we achieve for all age groups), as opposed to -0.06 for 18–29 and -0.07 for 30–39 hosts. Likewise, when acting as *hosts*, they preferentially receive 40+ year old guests, as opposed to 18–29 (median $pa = -0.08$, being it the lowest preferential attachment we achieve for all age groups). This may reduce the opportunities that 40+ year old Airbnb users have, both as guests (i.e., fewer properties to choose from) and hosts (i.e., attracting fewer paying guests). The two other age groups of 18–29 and 30–39 still preferentially stay with the same age group, however their preferential attachment is significantly lower than the one obtained for 40+ hosts/guests ($pa = 0.05$ for 18–29 hosts/guests and $pa = 0.01$ for 30–39 hosts/guests).

We finally turn our attention to host/guest stays focusing on ethnic groups (Figure 2c). Our diversity analysis already signalled significant under-representation of Asian and Black, both as hosts and guests. If we look at actual stays, an even more alarming reality emerges: the median pa value among Asian guests/hosts is equal to 0.24, and almost the entirety of stays involving Black hosts or guests is indeed homophilic – pa reaching the highest value of 0.96 for Black guests staying into properties rented by Black hosts, as opposed to $pa = -0.13$ for Black guests staying in properties rented by White hosts. The combined Indian, Latino, and Middle Eastern ethnic group (labelled 'Others') does not appear to be disadvantaged in terms of inclusion; however, this is the group for which our facial recognition tool achieved the lowest accuracy, so our results may have a greater degree of uncertainty than those shown for the other ethnic groups.

Overall, our inclusion analysis has revealed that, for Airbnb in London UK, three demographic groups are mostly engaged in homophilic interactions: those aged 40+ (median $pa = 0.11$), those of Black ethnicity (median $pa = 0.96$), and those of Asian ethnicity (median $pa = 0.24$). These demographics are also those with lowest representation within Airbnb, as previously evidenced by our results on diversity. The combination of these two findings suggests that these demographic groups are at risk of being significantly excluded from the sharing economy business model. To counter this, Airbnb could consider implementing platform interventions (e.g., a recommender system) to favour more inclusive (heterophilic) interactions; however, before doing so, one first needs to understand whether satisfaction with the hospitality service is comparable between heterophilic and homophilic stays, or whether perhaps satisfaction is significantly higher for the latter (as one might expect based on studies of interactions on social media – e.g., (Block and Grund 2014)). We explore this next.

Sentiment

Airbnb does not make available the individual ratings that guests leave after a stay, only an overall average rating per host. We thus built a proxy for individual user satisfaction with their own stay by focusing on the reviews written after a stay, and by measuring the sentiment associated to them.

We calculated an average sentiment score across all Airbnb stays of $+0.93$ (over a range of $[-1, +1]$), corroborating previous studies that found sentiment in Airbnb reviews to be positively skewed (Zhu, Cheng, and Wong 2019; Martinez et al. 2017; Zervas, Proserpio, and Byers 2015; Zhu, Lin, and Cheng 2020; Santos et al. 2020; Bridges and Vásquez 2018; Alsudais and Teubner 2019). At an aggregate level, Airbnb guests thus appear to be very satisfied with their experience of the hospitality service; here we go a step further, dissecting sentiment as it varies between homophilic and heterophilic stays for each demographic group under study.

The most striking observation we found is that homophilic interactions are no more satisfactory than heterophilic ones; indeed, *sentiment gain* remains very close to 0 across all interaction groups (up to 0.02 for gender, up to 0.03 for age, up to 0.02 for ethnicity). Furthermore, we found no statistical correlation between preferential attachment and sentiment gain (Spearman correlation $= -0.09$, p-value > 0.6), suggesting that the review sentiment score is not associated with the likelihood of having guests and hosts preferentially connected.

5 Discussion

Implications

The presence of homophily has been found in several studies of human networks (McPherson, Smith-Lovin, and Cook 2001); our results illustrate strong presence of homophily in Airbnb too, both along gender, age and ethnicity lines. This is despite the fact that, in our case study of London, the majority of Airbnb properties are entire houses / flats, meaning that hosts and guests most likely complete a short-term rental without ever meeting or bonding with one another, thus raising the question as to why there is such strong presence of homophily in the platform. Indeed, as our study further evidenced, homophily is not associated to sentiment, even after controlling for property type, price and location. One may thus wonder whether this tendency of Airbnb hosts and guests to complete transactions with demographically similar others is nothing but *an online manifestation of our offline biases*. If so, there are both moral, economic, and market reasons for considering interventions to reduce this phenomenon.

From a moral point of view, interventions should be considered to achieve social justice, removing systemic barriers that may lead some segments of the population (e.g., older adults) to become increasingly marginalised and to miss opportunities to contribute to society. From an economic point of view, diversity and inclusion have been found to be associated to economic efficiency (Burns 2012), thus advocating for organisations to tap into diverse pools. And from a market perspective, businesses must be able to reflect the diversity of their market base, if they are to keep growing their market share; for both Airbnb UK and US, this means for example attracting the non-White and older population, since these are the fastest growing demographics in these countries (for National Statistics 2022; Vespa et al. 2018).

In practical terms, what role can technology play to increase diversity and inclusion in sharing economy platforms such as Airbnb? Some technological interventions have already taken place, mostly aimed at addressing potential *hosts' biases*: for example, the 'instant booking' option was introduced a few years back to curb discrimination (Tam 2016); when hosts list their property using this option, guests can complete a reservation request instantly, without needing an explicit host's approval. In an attempt to further reduce racial discrimination in cases the instant booking option is not used, a trial has just started to show hosts the names of potential guests only after a booking has been confirmed (Rylah 2022).

Similar interventions could be considered to tackle *guests' biases* too: for example, not showing names and profile pictures of hosts until a reservation request has been made, specifically when searching for entire homes/apartments. In these cases, new recommender system functionalities could also be introduced to more proactively foster inclusion: for example, at the moment the onus to browse through available accommodations to find a suitable one is entirely on the guest; however, a recommender system could suggest properties the guest might enjoy among those matching the search criteria (e.g., price, date and location), while also promoting inclusion; based on our findings, increased inclusion would not come at the expense of satisfaction. Gamification elements could be considered as well to promote inclusive stays, such as collecting badges and rewards. Although the above interventions might be suitable when considering entire homes/apartments, the situation becomes more complicated when considering rooms in shared properties; in these cases, issues of personal safety may come into play (for example, female guests may only be willing to share a property with female hosts). When dealing with shared properties, masking names and profile pictures could be a barrier to trust building and discourage interactions altogether.

The above technological interventions might help with inclusion, but not with diversity. In order to design technological interventions aimed at increasing the diversity of Airbnb peers in the first instance, one needs to understand what barriers are currently preventing under-represented groups to take part in sharing economy services such as Airbnb. Indepth qualitative studies with past Airbnb users who have disengaged with the platform might shed some light into current barriers to diversity and inclusion; broader market research studies are also needed to understand motivation and fears from specific demographic groups.

Limitations

The findings of the present study should take into consideration a number of limitations.

First, our method relies on AI tools to infer demographic features from profile pictures; as such, our findings can only be as robust as these tools are. Based on our technology assessment study, findings related to diversity and inclusion with respect to gender and age can be stated with higher confidence than those concerning ethnicity (especially beyond the Asian, Black and White ethnic groups). As highlighted by another recent study (Cavazos et al. 2021), further research is necessary to enhance the accuracy of ethnicity estimation in AI face recognition tools before findings derived from such tools can be relied upon. While gender classification accuracy is currently high, at present all tools can only process gender as a simple binary concept. It is also important to note that while face recognition algorithm accuracy will continue to improve, we assume that profile pictures are true representations of Airbnb peers, which we believe is a plausible assumption since user pictures are a fundamental element of the platform's trust and reputation system.

Second, our findings are based on the analysis a 10% sample of the original data. This was due to financial restrictions derived from the use of proprietary AI face recognition tools (a choice that was deemed necessary to avoid compromising accuracy). The sample we were left with was still remarkable (i.e., 14k Airbnb hosts, 106k Airbnb guests, and 147k reviews). However, we acknowledge that data sampling may lead to certain subgroups within the population being either over- or under-represented in the sample, casting doubts in the validity of the findings. To reduce susceptibility to sampling bias, we chose our sample *randomly*. The only profiles that were *intentionally* excluded were those whose picture did not contain a person. Future research may look into whether some demographic groups are more likely than others to mask their true identity (e.g., favouring the picture of a landscape or a pet instead). We chose to exclude these profiles from the present study since our work aimed to unveil patterns of interactions based on the *user demographics* as revealed in the profile pictures chosen by hosts/guests on the Airbnb platform. Future studies may also look into interactions dynamics when profile pictures do not represent a human being instead (though it is worth noting these were relatively uncommon in our data, representing approximately 14% of the sample). Finally, as open-source AI tools continue to improve, we envisage it will be possible to conduct this type of studies with no need for data sampling at all.

Third, we measure satisfaction as the sentiment expressed in reviews. In doing so, we consider a single 'sentiment scale' across all users. A recent study of TripAdvisor shows that the way we write reviews vary depending on our demographics: in particular, self-identified females are much less likely to write negative reviews than male ones (Proserpio, Troncoso, and Valsesia 2021). We therefore need to be careful when inferring the satisfaction with Airbnb stays from user reviews as higher sentiment score does not always imply higher user satisfaction. In order to gain a more comprehensive understanding of the subject matter, future studies could delve deeper into their definition of sentiment scale, as well as explore the topics covered in Airbnb reviews.

Fourth, the findings of this study only pertain the city of London, UK, and cannot be generalised any further. This is because Airbnb, and more broadly the sharing economy, are fundamentally a *urban* phenomenon, and local geographic factors ranging from urban structure to population density patterns to social structure might significantly impact our findings, thus calling for the present study to be repeated in other geographic contexts.

Last but not least, while we focused on *demographic* diversity and inclusion of Airbnb users, there are other social and economic factors that can influence a user's decision to join the platform. For example, individuals with lower incomes may face barriers to accessing the platform, both as hosts (e.g., if they do not have spare rooms to offer) and as guests (e.g., if they do not have sufficient funds to travel). We partly considered geographic and economic factors within our models, by controlling for property price, location, and type. Further studies may look into other aspects of diversity and inclusion, to offer a more comprehensive socio-economic view of the platform dynamics.

Ethical Considerations

This work used publicly accessible data from the Airbnb company website. Data was anonymised and aggregated, to avoid traceability to any individual. Since the study pertains personal characteristics (i.e., gender, age and ethnicity), we first obtained study approval from the University Research Ethics Committee (Ethics Application 6725/003).

6 Conclusion

In this paper, we have proposed a quantitative method to measure diversity and inclusion longitudinally and at scale in sharing economy platforms. The method leverages face recognition AI tools to measure diversity, and network shuffling to measure inclusion. We have applied this method to conduct a longitudinal study that aimed to unveil diversity, inclusion, and overall sentiment of Airbnb hosts and guests for the specific case of London, UK. Unlike conventional census and survey-based methods, that rely on human-based data collection techniques, our approach utilises software-based methods that can be easily replicated to study other cities, as well as different sharing economy platforms (where it is common for users to have profile pictures, and for ratings and reviews to be left upon completion of an interaction). The method can also be repeatedly applied over time, so to measure diversity and inclusion as they vary, possibly as a result of platform interventions.

The method, when applied to the specific case of Airbnb in London, UK, detected high diversity for gender, but low diversity for age and ethnicity; while the latter is now steadily improving, the former is not. Considering that the elder segment of the population is the one that is predicted to increase the most over the next years, investigating why this demographic group is so under represented is an important direction of research, so that interventions can be enacted if systemic barriers are found to be responsible. Our findings also revealed issues of inclusion, as a consequence of a preponderance of homophilic interactions over heterophilic ones, despite these not being linked to higher sentiment. These findings can not be extrapolated beyond London; in order to inform platform interventions, future studies need to be conducted in different locales.

Future work also needs to look at the impact that COVID-19 has had on individuals and their behaviour towards sharing. We chose to exclude data from January 2020 onward from the present study since Airbnb came almost to a stop as the world battled through the pandemic, with travel bans imposed widely. Now that restrictions have been lifted, an interesting question is to understand to what extent our behaviours have bounced back to pre-pandemic, and to what extent our sharing attitudes have changed instead, possibly in a different way for different demographics.

Acknowledgments

We are grateful to Aniket Dixit for his diligent efforts in exploring AI tools for face recognition which have significantly improved the depth and quality of our research.

References

Alsudais, A.; and Teubner, T. 2019. Large-Scale Sentiment Analysis on Airbnb Reviews from 15 Cities. In *Twenty-fifth Americas Conference on Information Systems*.

Bamman, D.; Eisenstein, J.; and Schnoebelen, T. 2014. Gender identity and lexical variation in social media. *Journal of Sociolinguistics*, 18(2): 135–160.

Bellotti, V.; Ambard, A.; Turner, D.; Gossmann, C.; Demkova, K.; and Carroll, J. M. 2015. A Muddle of Models of Motivation for Using Peer-to-Peer Economy Systems. In *Proc. of the 33rd Annual ACM CHI Conference*, 1085–1094.

Benitez-Aurioles, B.; and Tussyadiah, I. 2020. What Airbnb does to the housing market. *Annals of Tourism Research*.

Block, P.; and Grund, T. 2014. Multidimensional homophily in friendship networks. *Network Science*, 2(2): 189–212.

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

Buolamwini, J.; and Gebru, T. 2018. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency*, 77–91. PMLR.

Burns, C. 2012. The Costly Business of Discrimination. Https://tinyurl.com/2p8bmy2f.

Cavazos, J. G.; Phillips, P. J.; Castillo, C. D.; and O'Toole, A. J. 2021. Accuracy Comparison Across Face Recognition Algorithms: Where Are We on Measuring Race Bias? *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 3(1): 101–111.

Centola, D.; Gonzalez-Avella, J. C.; Eguiluz, V. M.; and Miguel, M. S. 2007. Homophily, Cultural Drift, and the Co-Evolution of Cultural Groups. *Journal of Conflict Resolution*, 51(6): 905–929.

Centola, D.; Willer, R.; and Macy, M. 2005. The Emperor's Dilemma: A Computational Model of Self-Enforcing Norms. *American Journal of Sociology*, 110(4): 1009–1040.

Cheng, M.; and Jin, X. 2019. What do Airbnb users care about? An analysis of online review comments. *International Journal of Hospitality Management*, 76: 58–70.

Choudhury, M. D.; Sundaram, H.; John, A.; Seligmann, D. D.; and Kelliher, A. 2010. "Birds of a Feather": Does User Homophily Impact Information Diffusion in Social Media? *CoRR*, abs/1006.1702.

Christakis, N. A.; and Fowler, J. H. 2007. The Spread of Obesity in a Large Social Network over 32 Years. *New England Journal of Medicine*, 357(4): 370–379.

Dann, D.; Teubner, T.; and Weinhardt, C. 2019. Poster child and guinea pig – insights from a structured literature review on Airbnb. *Intl. Journal of Contemporary Hospitality Management*.

Dillahunt, T. R.; and Malone, A. R. 2015. The Promise of the Sharing Economy among Disadvantaged Communities. In *Proc. of SIGCHI Conference*, 2285–2294.

Dillahunt, T. R.; Wang, X.; Wheeler, E.; Cheng, H. F.; Hecht, B.; and Zhu, H. 2017. The Sharing Economy in Computing: A Systematic Literature Review. *Proc. of ACM CSCW*.

Eckhardt, G. M.; and Bardhi, F. 2015. The sharing economy isn't about sharing at all. *Harvard business review*.

Edelman, B. G.; and Luca, M. 2014. Digital discrimination: The case of airbnb.com. Technical report, Harvard Business School Working Paper.

for National Statistics, O. 2022. National population projections: 2020-based interim.

Ge, Y.; Knittel, C. R.; MacKenzie, D.; and Zoepf, S. 2016. Racial and gender discrimination in transportation network companies. Technical report, National Bureau of Economic Research.

Golub, B.; and Jackson, M. O. 2012. How Homophily Affects the Speed of Learning and Best Response Dynamics. *The Quarterly Journal of Economics*, 127(3): 1287–1338.

Griffith, M. W. 2017. Airbnb as a Racial Gentrification Tool? Brookly Deep.

Griswold, A. 2016. The dirty secret of Airbnb is that it's really, really white. Https://tinyurl.com/mr3r2vfk.

Hannák, A.; Wagner, C.; Garcia, D.; Mislove, A.; Strohmaier, M.; and Wilson, C. 2017. Bias in Online Freelance Marketplaces: Evidence from TaskRabbit and Fiverr. In *CSCW*, 1914–1933.

Hossain, M. 2020. Sharing economy: A comprehensive literature review. *International Journal of Hospitality Management*, 87.

Hutto, C.; and Gilbert, E. 2014. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proc. of ICWSM*, 8(1): 216–225.

Joseph, G.; and Varghese, V. K. 2019. Analyzing Airbnb Customer Experience Feedback Using Text Mining. *Big Data and Innovation in Tourism, Travel, and Hospitality*.

Koh, V.; Li, W.; Livan, G.; and Capra, L. 2019. Offline biases in online platforms: a study of diversity and homophily in Airbnb. *EPJ Data Science*, 8(1).

Lee, C. K. H.; Tse, Y. K.; Zhang, M.; and Ma, J. 2020. Analysing online reviews to investigate customer behaviour in the sharing economy: the case of Airbnb. *Information Technology & People*, 33(3): 945–961.

Lee, D. 2016. How Airbnb short-term rentals exacerbate Los Angeles's affordable housing crisis: Analysis and policy recommendations. *Harvard Law & Policy Review*, 10: 229.

London Datastore. 2020. Population Statistics and Analysis at the Greater London Authority. Https://data.london.gov.uk/demography.

Luo, Y. 2018. *What Airbnb Reviews can Tell us? An Advanced Latent Aspect Rating Analysis Approach*. Ph.D. thesis, Iowa State University.

Luo, Y.; and Tang, R. L. 2019. Understanding hidden dimensions in textual reviews on Airbnb: An application of modified latent aspect rating analysis (LARA). *International Journal of Hospitality Management*, 80: 144–154.

Ma, D. S.; Correll, J.; and Wittenbrink, B. 2015. The Chicago Face Database. *Behavior research methods*, 47(4): 1122–1135.

Martinez, R. D.; Carrington, A.; Kuo, T.; Tarhuni, L.; and Abdel-Motaal, N. A. Z. 2017. The Impact of an AirBnb Host's Listing Description 'Sentiment' and Length On Occupancy Rates. arXiv:1711.09196.

McPherson, M.; Smith-Lovin, L.; and Cook, J. M. 2001. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27: 415—-444.

Mehrabi, N.; Morstatter, F.; Saxena, N.; Lerman, K.; and Galstyan, A. 2021. A Survey on Bias and Fairness in Machine Learning. *ACM Computing Surveys*, 54(6).

Meurer, J.; Stein, M.; Randall, D.; Rohde, M.; and Wulf, V. 2014. Social Dependency and Mobile Autonomy: Supporting Older Adults' Mobility with Ridesharing. In *Proc. of CHI*, 1923–1932.

Moazed, A.; and Johnson, N. L. 2016. *Modern Monopolies: What It Takes to Dominate the 21st Century Economy*. St. Martin's Press.

Picascia, S.; Romano, A.; and Teobaldi, M. 2017. The airification of cities. Making sense of the impact of peer to peer short term letting on urban functions and economy. In *Annual Congress of the Associationof European Schools of Planning*. Lisbon, Portugal.

Proserpio, D.; Troncoso, I.; and Valsesia, F. 2021. Does Gender Matter? The Effect of Management Responses on Reviewing Behavior. *Marketing Science*, 40(6): 1199–1213.

Quattrone, G.; Greatorex, A.; Quercia, D.; Capra, L.; and Musolesi, M. 2018. Analyzing and predicting the spatial penetration of Airbnb in U.S. cities. *EPJ Data Science*, 7.

Quattrone, G.; Nocera, A.; Capra, L.; and Quercia, D. 2020. Social Interactions or Business Transactions? What Customer Reviews Disclose about Airbnb Marketplace. In *Proc. of WWW*, 1526–1536.

Quattrone, G.; Proserpio, D.; Quercia, D.; Capra, L.; and Musolesi, M. 2016. Who Benefits from the "Sharing" Economy of Airbnb? In *Proc. of WWW*, 1385–1394.

Raji, I. D.; and Buolamwini, J. 2019. Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products. In *Proc. of AAAI/ACM Conference on AI, Ethics, and Society*, 429–435.

Rylah, J. B. 2022. Airbnb's new experiment to reduce racismn. Https://thehustle.co/01072022-airbnb-racial-bias/.

Santos, G.; Mota, V. F. S.; Benevenuto, F.; and Silva, T. H. 2020. Neutrality may matter: sentiment analysis in reviews of Airbnb, Booking, and Couchsurfing in Brazil and USA. *Social Network Analysis and Mining*, 10(1): 45.

Shih, P. C.; Bellotti, V.; Han, K.; and Carroll, J. M. 2015. Unequal Time for Unequal Value: Implications of Differing Motivations for Participation in Timebanking. In *Proc. of CHI*, 1075–1084.

Sundararajan, A. 2016. *The Sharing Economy: The End of Employment and the Rise of Crowd-Based Capitalism*. MIT Press.

Sutherland, I.; and Kiatkawsin, K. 2020. Determinants of Guest Experience in Airbnb: A Topic Modeling Approach Using LDA. *Sustainability*, 12(8).

Tam, D. 2016. Airbnb thinks increased instant bookings will curb discrimination. Https://tinyurl.com/3nsbxxsn.

Thelwall, M. 2009. Homophily in myspace. *Journal of the Association for Information Science and Technology*, 60(2): 219–231.

Trenz, M.; Frey, A.; and Veit, D. 2018. Disentangling the facets of sharing. *Internet Research*, 888–925.

Tussyadiah, I.; Liu, A.; and Steinmetz, J. L. 2020. Impact of Perceived Peer to Peer Accommodation Development on Community Residents' Well-being. *Current Issues in Tourism*.

Ugander, J.; Karrer, B.; Backstrom, L.; and Marlow, C. 2011. The anatomy of the Facebook social graph. *arXiv preprint arXiv:1111.4503*.

Vespa, J.; Armstrong, D. M.; Medina, L.; et al. 2018. *Demographic turning points for the United States: Population projections for 2020 to 2060*. US Department of Commerce, Economics and Statistics Administration.

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

Wimmer, A.; and Lewis, K. 2010. Beyond and below racial homophily: ERG models of a friendship network documented on Facebook. *American J. of Sociology*, 116(2).

Zervas, G.; Proserpio, D.; and Byers, J. W. 2015. The Impact of the Sharing Economy on the Hotel Industry: Evidence from Airbnb's Entry Into the Texas Market. In *Proc. of ACM Conference on Economics and Computation*, 637.

Zhu, L.; Cheng, M.; and Wong, I. A. 2019. Determinants of peer-to-peer rental rating scores: the case of Airbnb. *International Journal of Contemporary Hospitality Management*, 31(9): 3702–3721.

Zhu, L.; Lin, Y.; and Cheng, M. 2020. Sentiment and guest satisfaction with peer-to-peer accommodation: When are online ratings more trustworthy? *International Journal of Hospitality Management*, 86: 102369.