The Social Side of Airbnb: Large-scale Linguistic Analysis of Hosts and Guests Reviews

ANONYMOUS

Contrary to what their manifesto claims, sharing economy platforms have repeatedly been critiqued for being all about access as opposed to sharing, and for their users to be after utilitarian, as opposed to social value. Knowing whether an economy is about access or sharing has important implications, both in terms of how these companies market themselves, and how they are being regulated. In this paper, we perform a large-scale linguistic analysis of Airbnb hosts' and guests' reviews, spanning seven years of activity across the globe. We find strong evidence that Airbnb reviews are predominantly utilitarian, and increasingly so as time goes by; however, we also find a non-negligible number of topics that cover social values instead. Using a mixed-method approach that combines thematic analysis and machine learning techniques, we delve deeper into the social side of Airbnb reviews, and identify the sub-themes that hosts/guests discuss the most: the *people* they interact with (both in terms of *who* they are, and what their *personality* is), and *how* they interacted with one another (e.g., talking, having a meal together). Finally, we zoom in both temporally and geographically, and observe that the decline of the social side of Airbnb has been happening very smoothly over time, and in a similar way across several countries. This suggests a global shift in attitude about how the Airbnb platform is being appropriated, that subsumes both local values and platform changes.

Additional Key Words and Phrases: Sharing economy, Airbnb, linguistic analysis, thematic analysis, semi-supervised machine learning

ACM Reference Format:
Anonymous. 2018. The Social Side of Airbnb: Large-scale Linguistic Analysis of Hosts and Guests Reviews. *J. ACM* 1, 1, Article 1 (January 2018), 16 pages. https://doi.org/0000001.0000001

1 INTRODUCTION

The term "sharing economy" was originally coined to capture a class of economic arrangements whereby, rather than owning goods and services, individuals would share them with one another, under the premise that there exists substantial excess capacity in the system, and with that an opportunity to optimize resources, and increase their *value*, through sharing. The sharing economy manifesto[1] emphasizes how *social* value is as important as *financial* value, with the practice of sharing seen as a means to build stronger communities.

This vision has however received strong critiques; for example, Eckhardt and Bardhi [2] observed that sharing is a form of social exchange that takes place among people who know each other, without any profit. The moment a company, like Airbnb or Uber, acts as a financial intermediary betweens strangers, transactions become economic exchanges. In these circumstances, peers are after utilitarian value, as opposed to social value.

[1] http://www.thepeoplewhoshare.com/

Author's address: Anonymous.

0004-5411/2018/1-ART1 $15.00
https://doi.org/0000001.0000001

Being able to assess whether an economy is about access or sharing has important implications for how companies operate, and compete, in this space. For example, let us consider the case of the short-term car-ride market. Two of the major sharing economy services operating in this space are Uber and Lyft. The former markets itself as being faster and cheaper than taxi (thus emphasising utilitarian values); the latter markets itself as a friendly and community-based service (thus emphasising social values). Despite offering a near identical service, Lyft's growth has been nowhere near that of Uber, and some market research suggests that a contributing factor has been the fact that they put too much emphasis on a desire to "share" that is actually not there [9]. AirBnB, another giant of the sharing economy, has been branding itself strongly as a symbol of sharing (with its 'people, places, love and community' campaign), but market analysts have been wondering whether this was indeed a smart move, suggesting that most consumers use AirBnB because of its economic (as opposed to social) value.

To gather information about consumers' needs, market research often combines qualitative techniques (e.g., focus groups, in-depth interviews) with quantitative ones (e.g., customer surveys). These techniques require substantial financial and time investments, which limit their adoption over time and geographies. This is particularly critical for fast-evolving markets, especially if the consumers' base spans several countries – as is the case for many sharing economy platforms. However, in most of these markets, there is a continuous stream of ready-available secondary data that can be systematically analysed to extract relevant information: reviews that peers leave to one another upon completion of a service exchange.

In this paper, we propose a mixed-method technique to perform linguistic analysis of reviews in sharing economy markets, as a means to systematically gather knowledge of peers' values at a certain point in time and space. We take the case of AirBnB as an example, since the platform has been widely adopted across the world and for several years, thus affording us the execution of a comparative study in different countries and at different times. Overall, we make the following three main contributions:

(1) *Baseline Analysis* – We collect a dataset comprising 117k AirBnB guests' reviews and a further 56k hosts' reviews, covering a period of seven years of activity (from 2011 to 2017) across the world. Using the well-established Linguistic Enquiry and Word Count (LIWC) dictionary [13], we analyse these reviews and find strong evidence that business-related word categories (e.g., 'home', 'space') dominate social-related ones (e.g., 'she-he', 'humans'), and that the gap between the two increases steadily over time. LIWC is however a general-purpose dictionary and not sufficient to grasp the nuances of business vs social reviews.
(2) *Dictionary Construction* – In order to delve deeper into the topics discussed in Airbnb, we propose and use a mixed approach that involves both thematic analysis and machine learning. Using the former, we identify four main themes that AirBnB peers discuss in their reviews (i.e., property, location, business conduct and social interaction); with the latter, we then further break down these main themes into 14 distinct sub-themes, and automatically build dictionaries of the most common *n*-grams used to discuss each of them. We so end up with a custom-built dictionary that researchers can use to analyse reviews in the Airbnb sharing economy platform.
(3) *Linguistic Analysis* – Using our custom-built dictionary, we systematically study both guests' and hosts' reviews, as they vary over time. First, we confirm what already found with the LIWC dictionary: business-related themes (i.e., property, location, business conduct) dominate reviews, and they increasingly do so as time goes by. We then delve deeper into the social interaction theme, and analyse the extent to which reviews discuss with *whom* peers have interacted (e.g., husband, daughter), what their *personality* is (e.g., friendly, charming), and

how the interaction took place (e.g., talking, having a meal together). Our temporal analysis, performed both at year and at quarter granularity, reveals that social topics suffer a steady and smooth decline; furthermore, this pattern is observed both globally and when zoomed in at country level too. This suggests a global shift in attitude about how the Airbnb platform is being appropriated, that subsumes both platform changes and local values.

By understanding how peers think about these novel sharing economy markets (e.g., convenience of location *vs.* fostering a guest/host social relationship), both in different geographical contexts and at different points in time, companies can tailor and market their services so to drive successful business models (e.g., highlighting the convenience and cost-effectiveness of access, as opposed to the financial obligations of ownership, and/or the emotional obligations of sharing).

2 RELATED WORK

Sharing economy platforms like AirBnB have been studied by the academic community following three main line of enquiries: understanding the impact of such platforms on related industries, understanding motivation of both hosts and guests for participating, and analysing the ratings and reviews that peers leave to one another.

Impact. Scholars from Business Studies and Economics have extensively analysed the relationship between sharing economy services and their regulated counterparts (e.g., Uber *vs.* taxis, AirBnB *vs.* hotels), to shed light onto the increasing accusations about the former being predatory and exploitative, causing severe externalities onto the latter. For example, [16, 17] focused their analysis on Texas and showed that indeed an increase in the number of AirBnB listings in this city leads to a decrease in the number of monthly hotel room accommodations. It was further found that this trend mainly impacted lower-end hotels, while leaving untouched higher-end ones; the former then reacted by decreasing their prices. Other scholars have looked more broadly at the impact that sharing economy services have had on the economy. For example, [3] found that such platforms have brought benefits to the broader tourism industry, since a reduction in accommodation prices has increased the number of tourists, and that has been driving the opening of many new job positions in related industries (e.g., restaurants). However, low-end hotels have been closing down or seen their business reduced, which has instead caused a loss of jobs. Sharing economy markets have been growing extremely fast, and some scholars have also started to look at their evolution over time, both in terms of where they grow, and in terms of who takes part in them. For example, [10] conducted a temporal analysis on the growth of Airbnb in London, UK from 2012 to 2015. They found that Airbnb properties and hotels are not located in exactly the same urban areas: while the latter are mostly present in the city center and close to tourist attractions, the former are increasingly spreading out to areas where low and middle income people reside. These results collectively point to the fact that Airbnb and hotel industries are two distinct services, although they do no shed light onto what motivates customers to gravitate around one or the other.

Motivation. A parallel stream of works has looked into motivational factors behind sharing economy platforms' uptake. For example, [14] conducted an online survey of both Airbnb customers and traditional hotels' ones. They found that the two are quite distinct in terms of their needs and motivation. They further conducted in-depth interviews with hotel executives, and found that, while top-end hotels did not fear any disruption coming from AirBnB, smaller and mid-range ones feared the competition instead, and were starting to adjust their business accordingly. [11] focused on sharing economy participants, and Airbnb hosts in particular, and delved further into their motivation to take part in these markets. By means of an online survey distributed via social media, they discovered that primary drivers of participations were convenience (e.g., speed and ease of finding accommodation) and hedonism (e.g., having fun); social influence (e.g., perception that others have of AirBnB users) was also mentioned, though to a lesser extent. [6] focused on AirBnB hosts, and used in-depth interviews to elicit their motivation. Findings for the city of Helsinki revealed that money was often the initial driver; however, over time, social factors started to gain importance. A follow-up study [8] conducted with 12 host interviewees now based in San Francisco, revealed that the financial benefits of hosting were indeed not a key factor, while social exchange and interpersonal interactions were. These findings suggest that, while hotels are business structures only, AirBnB is a platform where business and social factors come together. For platform owners to act upon these findings, they would need to gain confidence in their validity, and assess to what extent they hold across a larger sample of users, in different countries, and over time. To do so, a methodologically different approach is needed. In this paper, we propose one based on the linguistic analysis of the reviews that peers leave to one another.

Rating and Reviews. Reviews left on sharing economy platforms have already been studied by some scholars, mainly to understand the rationale behind the very high rating scores observed on these platforms. Indeed, a study conducted on AirBnB found that 95% of its properties boast an average rating of either 4.5 or 5 stars [15]. This is in stark contrast to what they found on similar platforms like TripAdvisor, where the average star rating is 3.8, and where the variance across reviews is significantly higher. Interestingly, the authors found that, once restricting the analysis to properties that are listed on both platforms, average ratings on Airbnb and TripAdvisor do get more similar, but proportionally more properties receive the highest ratings (4.5 stars and above) on Airbnb than on TripAdvisor. [5] tried to shed light onto this, by studying the actual reviews that hosts and guests leave to one another. They found that high ratings were indeed matched by reported positive experiences, and suggest that 'socially induced reciprocity', which occurs when peers interact socially with one another, causes negative information to be omitted from reviews. [12] went a step further into analysing mutual ratings, this time focusing on the CouchSurfing hospitality platform, and found evidence of power-balance mechanisms being at play.

Beyond shedding light onto rating dynamics, reviews have also started to be investigated to compare monetary *vs.* non-monetary sharing economy services. In particular, [7] collected a sample of hosts' profiles and guests' reviews in AirBnB and Couchsurfing; after manually labelling and analysing them, they found initial evidence that the primary shared asset in AirBnB is the house (i.e., its facilities, location, neighbourhood), while in Couchsurfing it is the human relationship (i.e., self-description, motivation, experience). They suggest that such findings may drive businesses to design their services differently. We concur with this suggestion, and in this paper propose a scalable method to gather quantitative evidence of such market characteristics – over large geographies and different time scales. Before presenting our method, we briefly introduce the dataset we collected and later used to instantiate the method.

3 DATASET

In conducting this study, we did not want to focus on a specific city or country; rather, we wanted to analyse Airbnb reviews at global scale. We thus started by selecting Airbnb users (both guests and hosts) at random from all over the world, and collected a pool of 203k guests' reviews, and a further 79k hosts' reviews. During an initial cleansing step, we eliminated 'malformed' reviews that did not have a year and/or a country associated with them. Furthermore, we removed reviews that were too short (less than 5 words), since these were considered useless for content analysis. At the opposite end, we removed a few outlier reviews that were too long (more than 200 words). Finally, we removed reviews written by power users (i.e., users with several hundreds of reviews), once again to get rid of outliers. We ended up with a dataset comprising 117k guests' reviews and 56k hosts' reviews, covering seven year timespan (from 1011 until 2017) from as many as 160 countries. We inferred the language used in reviews by means of the R package *textcat* and the Rosette Text Analytics API[2]. English was by far the most commonly used language. We then used the Yandex Translate API[3] to translate all remaining reviews to English. Summary statistics about the collected dataset can be found in Table 1.

Number of distinct reviews	172,973 reviews (117,105 guest reviews and 55,868 host reviews)
Number of distinct peers	149,089
Number of detected languages	32
Top five languages	English (80%), Italian (6%), French (6%), Spanish (4%), German (3%)
Number of analysed countries	160
Number of reviews per country	Min 1, 1st Qu. 7, Median 29, 3rd Qu. 400, Max 62k
Number of countries having more than 400 reviews (3rd Qu.)	40
Percentage of reviews produced by country (top five)	USA (36%), ITA (7%), GBR (7%), FRA (5%), CAN (4%)
Review length (number of words per review)	Min 5, 1st Qu. 22, Median 37, 3rd Qu. 62, Max 200
Time span	7 years (from the second quarter of 2011 to the second quarter of 2017)

Table 1. Main statistics of our dataset

4 BASELINE ANALYSIS

We started our investigation by performing a linguistic analysis of the 170k reviews collected above, using the Linguistic Inquiry and Word Count (LIWC) dictionary. LIWC is a popular dictionary developed over the last few decades and it includes a number of different categories of language, ranging from part-of-speech (e.g., articles or personal pronouns) to topical categories (e.g., family or money) [13]. The 2007 version of LIWC includes 64 different categories; in order to understand to what extent Airbnb reviews are business oriented vs. social oriented, we focused our attention on the following four LIWC categories:

- *Home*. Examples of words belonging to this category are 'kitchen', 'bath', 'bed'. We included this category since it captures facilities of the property being rented.
- *Space*. Examples of words belonging to this category are 'area', 'at', 'away'. We included this category since it captures information about the neighbourhood within which a property is being located.
- *Shehe*. This category includes declinations of the third-person singular pronouns 'she' or 'he'. In conversations, these have been found to indicate a focus on people. In the case of Airbnb reviews, they could indicate a more socially oriented use of the platform, with discussion about the peers involved.
- *Humans*. Examples of words belonging to this category are 'boy', 'girl, 'adult'. As for the previous category, we included this one since such words may indicate an interest of Airbnb peers to talk about the interpersonal relationships they have experienced.

Intuitively, 'home' and 'space' LIWC categories signal business-oriented discussions, while 'shehe' and 'humans' signal social-oriented discussions.

We grouped reviews by year in which they were written, keeping guests and hosts separate. We then computed the total number of words belonging to each of the four categories above, normalised by the total number of words adopted. As Figure 1 illustrates, business-oriented topics dominate Airbnb reviews written both by guests (left) and hosts (right); furthermore, the popularity of these topics remains stable over time and even slightly increases for 'home'. On the contrary,

[2]https://www.rosette.com/
[3]https://tech.yandex.com/translate/

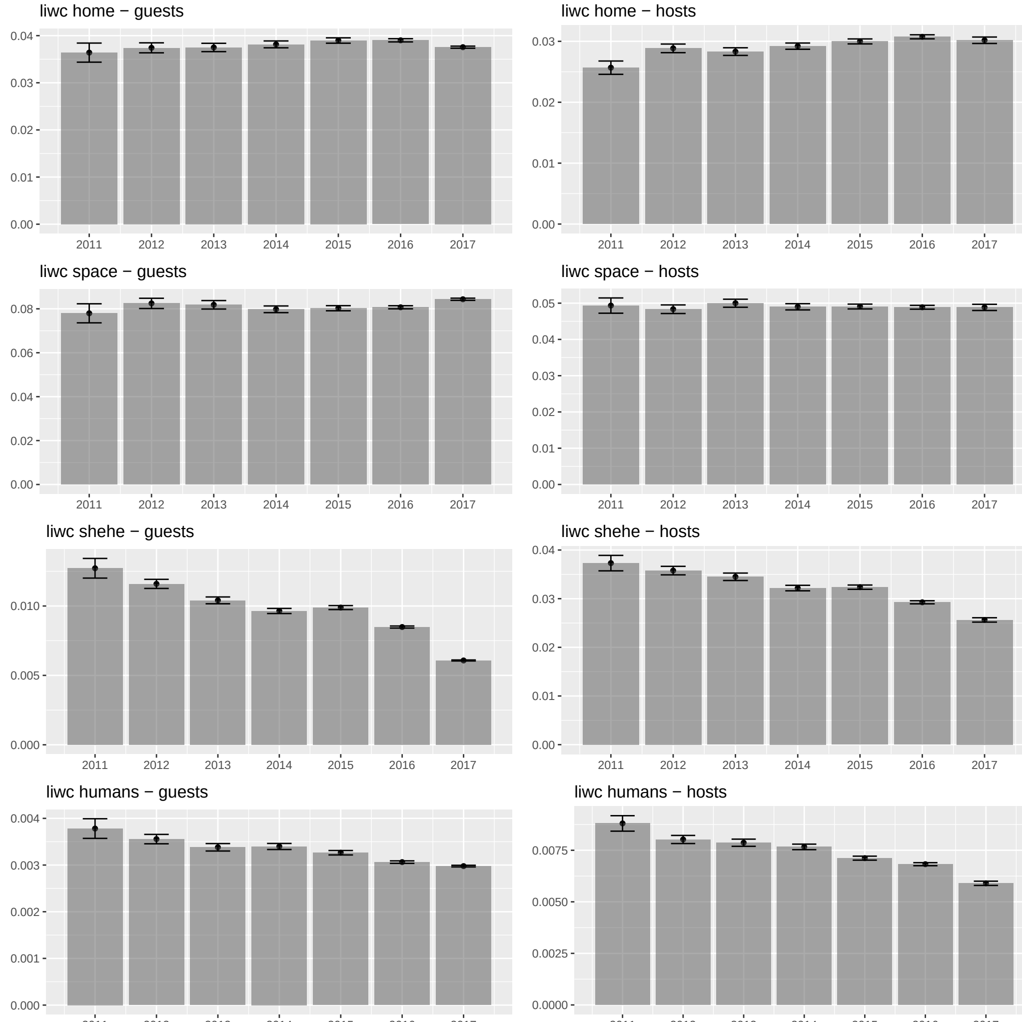

Fig. 1. LIWC adoption in Airbnb guests' and hosts' reviews. For each plot, the error bar shows the 95% confidence interval.

social-oriented topics are substantially less discussed, and the decline continues steeply as time progresses.

This is a quantitative evidence that Airbnb reviews primarily discuss the utilitarian value of this market (with emphasis on property facilities and location) as opposed to its potential social value. However, these charts also reveal that the steady decline in the discussion of social topics is not matched by an equal increase in business topics – the latter reaming mostly stable over time. One may thus wonder whether the general-purpose LIWC dictionary is an appropriate lens through which to conduct large-scale market analysis of these emerging economies, or whether a custom-built dictionary is needed instead. We developed this dictionary next.

Theme	*Description*	*Example*
Property	Covering topics related to the property/house/room such as its cleanness, its decoration, size, furniture, or house amenities.	"the room was bright and clean"
Location	Covering topics related to the neighbourhood where the property is situated such as nearby viewpoints, restaurants, and metro stations.	"the area was full of nice restaurants"
Professional Conduct	Covering topics related to the professional conduct of guest/host such as how good their communication has been, whether the guest has left the room clean and/or she has respected the house rules, or whether the host has been flexible with check-in/check-out times.	"she left the room clean and tidy"
Social Interaction	Covering topics related to the personality of guest/host or about guest/host spending social time together.	"I enjoyed chatting with her"

Table 2. Our four main themes

5 DICTIONARY CONSTRUCTION

We built our dictionary in five steps: first, we developed a coding scheme by performing thematic analysis of a random sample of 100 Airbnb reviews (step 1); second, we validated the coding scheme by means of a crowd-sourcing study conducted on the Crowdflower[4] platform (step 2), and ended up identifying four main themes (i.e., property, location, business conduct, social interaction). Third, we conducted a second study on Crowdflower, where we asked crowd-workers to label a larger set of 1,500 reviews, using the identified themes (step 3). Using natural language processing techniques, we built a dictionary of the n-grams most representative of each such category (step 4). Finally (step 5), using hierarchical clustering techniques, we grouped together these n-grams into 14 distinct clusters, which represent a more detailed refinement of the main 4 themes manually identified at steps 1 and 2. We discuss the details of each step next.

Step 1. Developing a Coding Scheme. Using stratified sampling to cover all study years and all major countries of activity, we sampled 100 Airbnb reviews. We broke down each review into its constituting sentences, and performed a thematic analysis over these. In a way similar to [1], two independent annotators coded these resulting sentences by performing three steps: *(i)* familiarising with the data, *(ii)* generating the initial codes and searching for themes among codes, and *(iii)* defining themes. After a full round of coding, the two coders compared their results, and agreed on which themes to remove, amend, or merge. As a result, they agreed on having five main themes named 'property', 'location', 'business conduct', 'personality', 'social interaction'. Intuitively, the first three refer to business-oriented discussion topics, while the latter two refer to social-oriented ones.

Step 2. Validating the Coding Scheme. To gain confidence in the validity of the coding scheme, we asked crowd-workers to annotate sentences extracted from a new sample of 100 Airbnb reviews, using these five themes. In particular, we prepared a Crowdflower page that consisted of three sections: *(i)* a list that showed our five themes; *(ii)* for each theme, actual examples of Airbnb reviews manually labelled by us; and *(iii)* new Airbnb sentences to be labelled. We paid 0.01$ per annotation, and each Airbnb sentence was independently annotated by at least four different workers. We computed the Fleiss' kappa agreement score for the five themes [4], and two of them (i.e., 'personality' and 'social interaction') had a Fleiss' kappa score less than 0.5. We merged these

[4] Crowdflower is a crowd-sourced market of online workforce to clean, label and enrich data: https://www.crowdflower.com/.

Theme	*Frequency*
Property	35%
Location	28%
Professional Conduct	23%
Social Interaction	14%

Table 3. Frequency of our four themes

two themes into one, resulting in four main themes: 'property', 'location', 'professional conduct' and 'social interaction' (Table 2). To ascertain the effectiveness of coding with those four themes, we again asked crowd-workers to annotate a new sample of sentences extracted from yet another 100 Airbnb reviews. All four themes resulted in a Fleiss' kappa score higher than 0.5, suggesting their validity.

Step 3. Labelling Reviews. We were then ready to label a larger set of Airbnb reviews using the identified four themes. We used again Crowdflower to annotate unlabelled sentences extracted from a new set of 1,500 reviews. We gathered 22,975 distinct annotations of 4,062 sentences. We kept those sentences on which at least 75% of annotators agreed, and ended up with a set of 1,868 sentences having high agreement. Table 3 shows the frequency of occurrence of each of the four themes in these sentences. As one would expect also based on our previous LIWC analysis, the most popular theme was 'property', followed by 'location' and 'professional conduct'; 'social interaction' was the least frequent theme instead.

Step 4. Building the Dictionary. For each of the four main themes, we needed to associate representative words (n-grams). We did so in a data-driven fashion. First, for each theme c, we split the 1,868 annotated sentences into two sets: Set_c (the set of sentences labelled with theme c by at least three quarter of workers), and $Set_{\bar{c}}$ (the set of sentences labelled with theme c by no more than one quarter of workers). Second, we extracted all n-grams from Set_c and $Set_{\bar{c}}$, with $n = \{1, 2, 3\}$. For each n-gram t, we computed two measures: $tf(t,c)$ and $tf(t,\bar{c})$, respectively denoting the term frequency of t in Set_c and in $Set_{\bar{c}}$. Finally, we computed $tf_{gain}(t,c) = \frac{tf(t,c)}{tf(t,\bar{c})}$.

To each theme c, we then associated all the n-grams t such that $tf(t,c) \geq tf_{min}$, $tf(t,c) \leq tf_{max}$ and $tf_{gain}(t,c) \geq th_{gain}$, with $tf_{min} > 0$, $tf_{max} \leq 1$ and $th_{gain} > 1$. The first two thresholds, tf_{min} and tf_{max}, removed extremely unpopular and extremely popular n-grams respectively. The use of the last threshold th_{gain} enabled us to associate to a theme c only those n-grams that were *comparatively more popular* in Set_c than in $Set_{\bar{c}}$. We manually checked the outcomes obtained for different values of tf_{min}, tf_{max} and th_{gain}, in terms of accuracy and completeness; the combination of values that gave us the best results was $tf_{min} = 0.01$, $tf_{max} = 0.15$ and $th_{gain} = 3$. All n-grams having $n > 1$ were kept only if the corresponding th_{gain} was higher than all the th_{gain} values associated with the sub n-grams. Only n-grams having $n = 1$ remained. The second column of Table 4 summarises the number of n-grams that each theme contained at this point.

We then used a word embedding machine learning technique (i.e., word2vec) to further enrich our initial dictionary. We started by training the technique on the whole corpus of reviews, and mapped each n-gram into a vector having 50 dimensions. For each n-gram already present in our dictionary, we then computed a list of similar n-grams, that is, a list of n-grams having a cosine similarity higher than a threshold th_{cos}. We included these n-grams in our dictionary if they were not already present. In so doing, we enriched our dictionary with n-grams that are not frequently used in the 1,868 labelled sentences, but still widely used in the whole corpus of reviews (and similar to those previously derived from our labelled corpus). The value of th_{cos} giving us the

Theme	*Initial n-grams*	*Expanded n-grams*
Property	63	77
Location	97	109
Professional Conduct	107	119
Social Interaction	61	68

Table 4. Number of n-grams for each of our themes, in our initial and final dictionary

best accuracy/completeness trade-off was $th_{cos} = 0.7$. The third column of Table 4 shows the total number of n-grams belonging to each of the four themes after this enrichment step.

Step 5. Identifying Sub-Themes. Upon performing a manual inspection of our expanded dictionary, we realised that several sub-themes could be identified within the four main ones that we manually coded at steps 1 and 2. For example, under the theme 'social interaction', we identified both n-grams that refer to *who* peers interacted with (e.g., husband, wife, daughter) as well as *how* (e.g, meals together, talking). In order to refine the top level themes and offer a more detailed taxonomic structure, we used a hierarchical divisive clustering algorithm. For each of the 4 main themes separately, we took all the n-grams associated to it and placed them in a single cluster. We then iteratively increased the number of clusters until the intra-cluster similarity of each produced cluster was higher than a given threshold th_{cls} (as before, cosine similarity was used). The value of th_{cls} that enabled us to identify a relatively small number of distinct clusters was $th_{cos} = 0.35$. After repeating the process for each of the 4 main themes, we ended up with 14 clusters: three clusters were refinements of the 'property' theme, five clusters were refinements of the 'professional conduct' theme, and a further five of the 'social interaction' theme. The 'location' theme was mapped to a single cluster, without further refinement. Table 5 provides an overview of the final dictionary we built, including our 4 manually labelled themes, 14 automatically inferred sub-themes[5] and their cluster internal similarity, number of n-grams in each sub-theme, and top 8 n-grams for each.

In terms of *property*, we observe that the three sub-themes derived from clustering directly match property description fields of Airbnb listings – that is, property type (e.g., whether a house or a flat), internal layout (e.g., kitchen, bed, cozy), and facilities (e.g., wifi, tv, fridge). In terms of *professional conduct*, distinct elements are detected: basic communication (e.g., questions, quick, responded), handling of logistics (e.g., check in, arrival), provision of advice (e.g., tips, directions), guests' care (usually found in hosts' reviews commenting about guests' behaviour – e.g., clean, tidy), and hosts' approach (orthogonal to care, usually found in guests' reviews commenting about the hosts' approach to the hospitality service – e.g., welcoming, helpful). We observe what might appear an overlap between sub-theme 'guests' care' in professional conduct, and the general 'property' theme. In practice, 'guests' care' is mostly discussed in hosts' reviews (e.g., "she left the room in a perfect state and respected the house rules"), while 'property' is discussed in guests' reviews (e.g., "'Lovely room in a bright spacious house"). In the analysis that follows, we consider these 3 themes and 9 sub-themes as indicative of reviews that 'talk business'. The category *social interaction* is now further refined in terms of 'people' (e.g., who peers interact with – e.g, husband, wife), what their 'personality' is (e.g., friendly, kind, warm), if/what they are 'sharing' (e.g., share, stories, experiences), and how – 'talking' (e.g., chat, talking, conversation) over a 'meal' (e.g., breakfast, dinner together).[6]

[5]Note that the name of the sub-theme was assigned by us after clustering.
[6]The full dictionary will be made available to the research community after publication.

Theme	*Sub-theme*	*n-grams*	*Top 8 n-grams (in a decreasing order of term frequency)*	*Internal cluster similarity*
Property	Type	17	apartment, house, home, flat, private, building, pool, property	0.48
	Interiors	43	clean, comfortable, room, bed, kitchen, spacious, small, cozy	0.42
	Facilities	17	water, hot, wifi, towels, tv, fridge, internet, facilities	0.42
Location	–	109	quiet, area, walk, located, restaurants, neighborhood, away, walking	0.40
Professional Conduct	Communication	33	communication, questions, quick, communicative, responded, answered, prompt, text	0.49
	Logistics	22	check, provided, arrival, arrived, late, offered, keys, hours	0.35
	Advice	11	information, recommendations, tips, advice, suggestions, directions, instructions, useful	0.61
	Guests' Care	18	place, apartment, clean, house, room, space, care, tidy	0.56
	Hosts' Hospitality	35	helpful, welcoming, available, accommodating, responsive, flexible, welcomed, attentive	0.37
Social Interaction	People	24	her, she, family, friends, husband, wife, daughter, son	0.44
	Personality	22	friendly, kind, warm, charming, sweet, gracious, adorable, generous	0.47
	Sharing	6	shared, share, sharing, experiences, stories, interests	0.60
	Talking	8	chat, conversation, conversations, talking, chatting, moments, chats, talks	0.55
	Meals	8	breakfast, delicious, fresh, dinner, meals, together, breakfast, cooked	0.59

Table 5. Summary statistics of the final dictionary

6 LINGUISTIC ANALYSIS

Using our newly constructed dictionary, we then went back to the Airbnb dataset we collected. To conduct a more refined linguistic analysis: first, we performed a temporal (yearly) analysis of the themes discussed across all countries, and validated results against those previously obtained with the general-purpose LIWC dictionary; we subsequently zoomed in into the social side of Airbnb reviews, and investigated what sub-themes disappear the most quickly. While doing so, we zoomed-in geographically and investigated whether this trend varies by country or whether it is a global phenomenon instead; furthermore, we zoomed-in temporally to observe whether the decline is smooth or abrupt (perhaps following major platform or legal changes).

6.1 Main Findings

We grouped Airbnb reviews by year, keeping host and guest reviews separate. For each group of reviews, we computed the fraction of n-grams in theme c:

$$adoption(c, text) = \frac{\sum_{t \in c} freq(t, text)}{\sum_{t \in v} freq(t, text)} \tag{1}$$

where $freq(t, text)$ is the number of occurrences of n-gram t in $text$, and v is the set of distinct n-grams in $text$.

Figure 2 illustrates the popularity of each theme in guests' reviews. Let us focus on business-related themes first. It is interesting to note that, while the 'home' and 'space' LIWC business categories remained mostly constant over time, using our purpose-built dictionary we can now have a much better understanding of what business aspects are actually 'growing' in popularity among Airbnb reviews over time: while 'location' ('space' in LIWC) remains mostly constant over time, we now observe a significant increase in the discussion of 'property' ('home in LIWC') and 'professional conduct' (a theme completely missing in LIWC). Conversely, mentions of 'social interactions' have been subject to significant decrease over time.

Figure 3 illustrates the popularity of themes in hosts' reviews instead. Note that, as expected, hosts do not discuss their property or its location in their reviews, so only two themes are relevant. Note the steep and steady rise of hosts' reviews focusing on professional conduct (of their guests) – a theme that we could not study using LIWC. The discussion of social interactions is declining

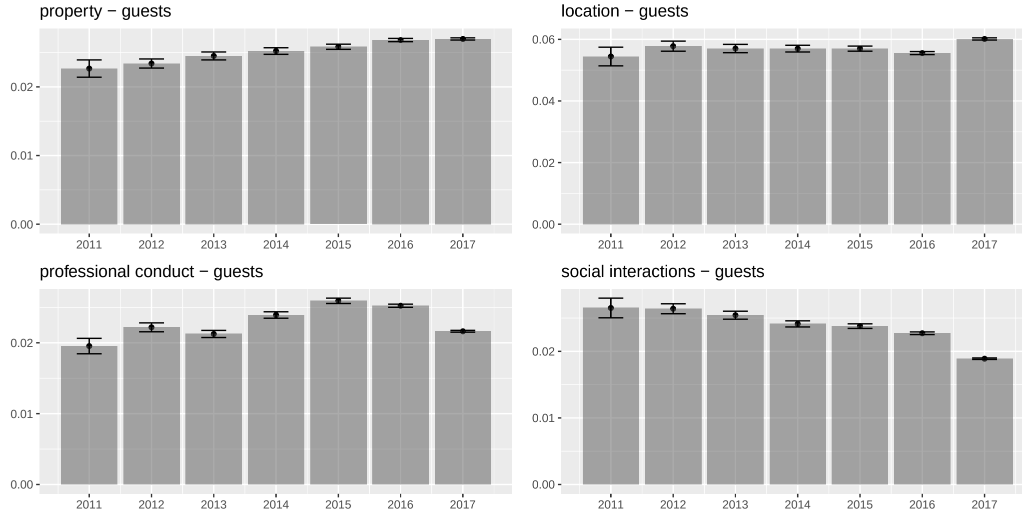

Fig. 2. Adoption of the four main themes for Airbnb guests. The error bar is showing the 95% confidence interval.

over time instead, although not as quickly as in guests' reviews. Using the more fine-grained (five) sub-themes for 'social interactions' identified in our custom-built dictionary, we zoom in into this aspect next.

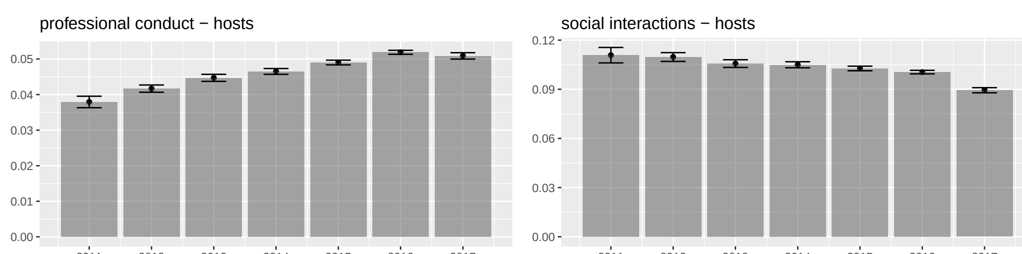

Fig. 3. Adoption of the 'professional conduct' and 'social interaction' main themes for Airbnb hosts . The error bar is showing the 95% confidence interval.

Figure 4 shows the popularity of each sub-theme over the years, for guests (left) and hosts (right) separately. Overall, we observe that the sub-themes 'people' and 'personality' are the most discussed, while 'sharing', 'talking' and 'meals' are (and have always been) much less popular discussion topics in reviews. Let us now focus on changes over time: we observe that both guests and hosts discuss who they interacted with much less as time passes; the decline is observable for the personality sub-theme too, although to a lesser extent. Discussions about sharing and talking vary over time, but overall they are in decline. It is interesting to observe a different trend between guests and hosts when it comes to discussing sharing a meal: while hosts have significantly reduced their discussion of the topic, guests seem to pay increasing attention to it instead.

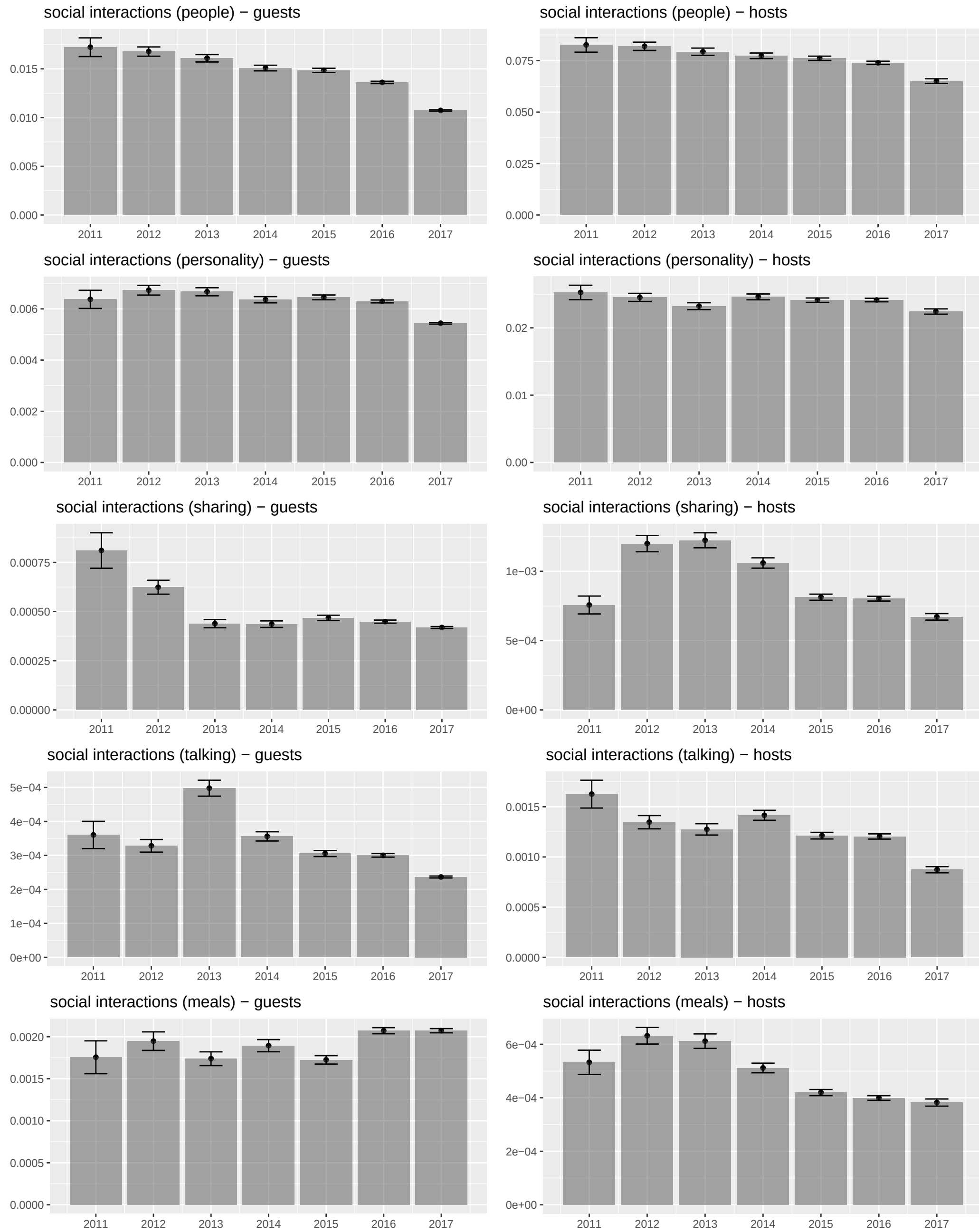

Fig. 4. Adoption of the five sub-themes of 'social interaction' for Airbnb guests and hosts. The error bar is showing the 95% confidence interval.

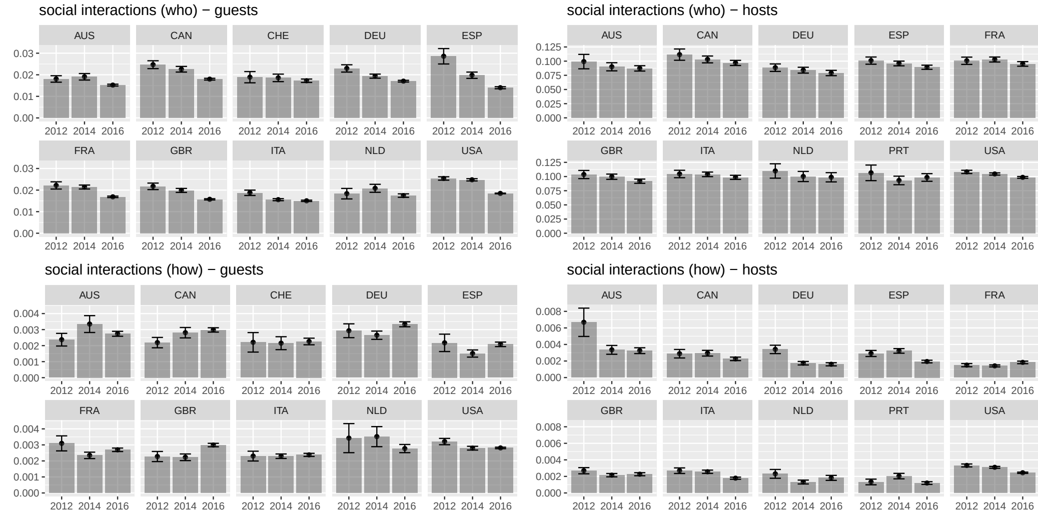

Fig. 5. Adoption of the 'social interaction' theme for Airbnb guests and hosts, separated in sub-themes 'who' (comprising 'people' and 'personality') and 'how' (comprising 'sharing', 'talking' and 'meals'), for the ten countries producing the highest numbers of reviews. Error bars show the 95% confidence interval.

6.2 Spatial Analysis

One may wonder if previous results are swayed by certain countries dominating our sample of reviews, or whether they are reflected at country level. In other words, is the popularity of business-related themes dominating over social-related ones everywhere in the world, or are there countries where a different trend emerges? Similarly, is the temporal decline in popularity of social-related themes a common trend in each country? Being able to answer these questions is particularly important for companies like Airbnb, who can decide to market their business differently in different parts of the world, based on a detailed understanding of what peers are after in different regions.

We repeated the analysis conducted before, but this time we grouped reviews by country. In order to have a statistical sample of reviews in each country, we considered the biennials 2012/13, 2014/15, 2016/17 as unit of time. Also, we combine sub-themes 'people' and 'personality' together, as well as 'sharing', 'talking' and 'meals'. We refer to the former as 'who' and to the latter as 'how'. Figure 5 shows results in the ten countries producing the highest numbers of reviews.

We observe that the 'who' sub-theme dominates over the sub-theme 'how' (by at least one order of magnitude) in all countries analysed, for both hosts' and guests' reviews. This result suggests that trend previously visualised at global level also stands at country level, at least for countries producing a high number of Airbnb reviews. In terms of temporal trend, we can further observe that the 'who' sub-theme decreases in most countries, for both guests' and hosts' reviews, though at a different rate (e.g., if we focus on guests' reviews, the decline in Spain is much steeper than in Italy, for example). The sub-theme 'how' decreases in hosts' reviews, but not so in guests' reviews – this again confirms at individual country level the trend previously detected at global level.

6.3 Temporal Analysis

As a final investigation, we zoom in temporally, rather than spatially, and repeat the previous analysis while grouping reviews by quarters (rather than annually). We do so to observe whether the decline in socially oriented reviews happened smoothly over time (thus suggesting a genuine

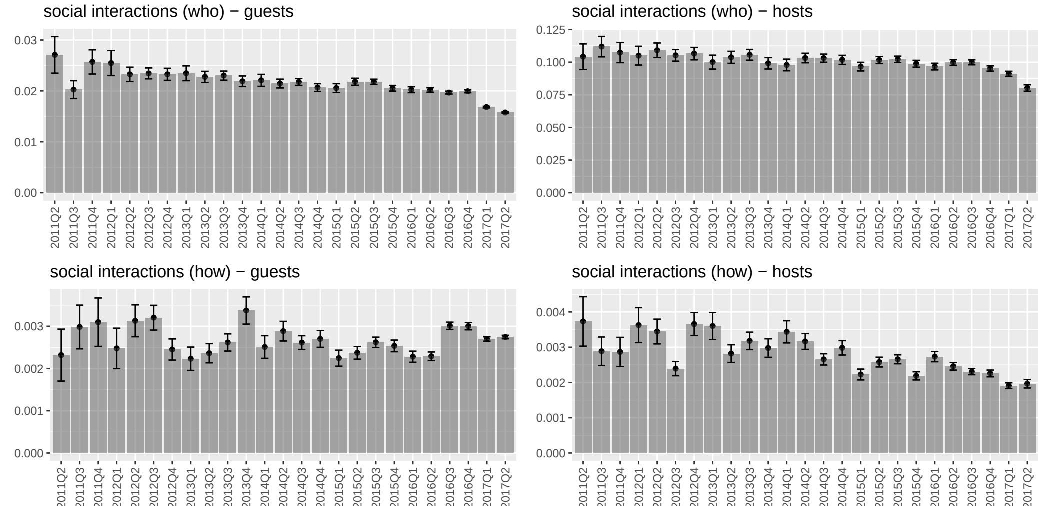

Fig. 6. Adoption of the 'social interaction' theme for Airbnb guests and hosts, separated in sub-themes 'who' (comprising 'people' and 'personality') and 'how' (comprising 'sharing', 'talking' and 'meals'), analysed at quarters' intervals. Error bars show the 95% confidence interval.

and gradual shift of peers' interests), or whether we can identify distinct points in time after which the decline has been more (or less) dramatic. These points might coincide, for example, with significant platform changes (e.g., in Airbnb, the introduction of blind reviews, and more recently of the instant booking facility), which may have resulted in peers' behaviour changes (even though our analysis cannot reveal any causality between the two).

Figure 6 shows the adoption of socially-oriented sub-themes 'who' (i.e., 'people' and 'personality' combined) and 'how' (i.e,. 'sharing', 'talking,' and 'meals' combined), for Airbnb guests and hosts, from the second quarter of 2011 to the second quarter of 2017. The 'who' sub-theme has a mostly smooth decline for both guests' and hosts' reviews, although a couple of discontinuity points can be identified for guests (a neat one in Q3 and Q4 of 2011, and a smaller one in Q4 of 2016). In the future, it will be worth investigating what platform (or regulatory) changes might have taken place at these times. Conversely, the sub-theme 'how' shows high variability over time, as well as different overall trends between guests and hosts; this insight may be useful when it comes to decide what marketing strategy to adopt (e.g., a focus on 'how' the sharing takes place in Airbnb may be unlikely to please both sides).

7 DISCUSSION

With a newly developed dictionary at hand, we were able to study to what extent Airbnb reviews tended to equally focus on social matters and business matters, as previous work suggested [6, 8]. We did so by analysing reviews written during seven years and across the whole world. As one might expect based on conventional wisdom, for Airbnb guests and hosts, business considerations are more salient than social ones. But, as one might not have expected, the social-business gap has not remained constant over the years: it has significantly increased instead, and it has done so everywhere. By drilling down on five distinct sub-themes related to 'social interactions', we have also been able to identify 'people' as the social aspect that is mostly discussed in reviews, but that is also subject to the steepest decline over time. Interestingly, sharing 'meals' is the only social aspect that is being increasingly discussed over time – but that is only so in guests' reviews, suggesting a potential divergence in what matters to hosts and guests in Airbnb. These global patterns have been confirmed also when zooming in at country level, and they overall offer the first quantitative cross-country evidence on to what extent the sharing economy is really about sharing.

However, some words of caution are in order. First, we do not know whether certain socio-demographic groups are more or less prone to write reviews than others and, therefore, whether they tend to be over- or under- represented. Second, we collected 282k random Airbnb reviews; however, since Airbnb is mostly adopted in Western countries, our dataset and findings may not be representative for non-Western countries. Third, in our analysis, any non-English review had been automatically translated into English before being analysed. Although we manually checked a number of translations and we were satisfied by the accuracy provided by the automated translation tool, we acknowledge that some original subtle meaning may have been lost in translation. However, we believe that the results showed in this work are not much altered by the translation procedure because our linguistic analysis is mainly based on word count and not on the analysis of sentence structure.

Despite those limitations, our method (comprising dictionary construction and linguistic analysis) is readily applicable to other platforms such as Couchsurfing and TaskRabbit. Aside from offering valuable knowledge to marketing experts, the method also offers a variety of practical implications, ranging from globally/locally tracking the "sense of community" within these platforms, to developing novel platforms features (e.g., ranking properties and listings based on hosts' sociability) informed by knowledge extracted from reviews.

REFERENCES

[1] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[2] Giana M Eckhardt and Fleura Bardhi. 2015. The sharing economy isnâĂŹt about sharing at all. *Harvard business review* 28, 01 (2015), 2015.

[3] Bin Fang, Qiang Ye, and Rob Law. 2016. Effect of sharing economy on tourism industry employment. *Annals of Tourism Research* 57 (2016), 264–267.

[4] Joseph L Fleiss. 1971. Measuring nominal scale agreement among many raters. *Psychological bulletin* 76, 5 (1971), 378.

[5] Andrey Fradkin, Elena Grewal, Dave Holtz, and Matthew Pearson. 2015. Bias and reciprocity in online reviews: Evidence from field experiments on airbnb. In *Proc. of EC '15*. ACM, 641–641.

[6] Tapio Ikkala and Airi Lampinen. 2015. Monetizing network hospitality: Hospitality and sociability in the context of Airbnb. In *Proc. of CSCW '15*. ACM, 1033–1044.

[7] Jiwon Jung, Susik Yoon, SeungHyun Kim, Sangkeun Park, Kun-Pyo Lee, and Uichin Lee. 2016. Social or Financial Goals?: Comparative Analysis of User Behaviors in Couchsurfing and Airbnb. In *Proc. of CHI '16*. ACM, 2857–2863.

[8] Airi Lampinen and Coye Cheshire. 2016. Hosting via Airbnb: Motivations and financial assurances in monetized network hospitality. In *Proc. of CHI '16*. ACM, 1669–1680.

[9] Dean Millson. 2016. *Uber vs Lyft: Brand positioning, customer expectation and internal culture.* Technical Report. Branding and Strategy.

[10] Giovanni Quattrone, Davide Proserpio, Daniele Quercia, Licia Capra, and Mirco Musolesi. 2016. Who Benefits from the Sharing Economy of Airbnb?. In *Proc. of WWW '16*. 1385–1394.

[11] Sampo Satama et al. 2014. Consumer adoption of access-based consumption services-Case AirBnB. (2014).

[12] Bogdan State, Bruno D Abrahao, and Karen Cook. 2016. Power Imbalance and Rating Systems.. In *Proc. of ICWSM '16*. 368–377.

[13] Yla R Tausczik and James W Pennebaker. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology* 29, 1 (2010), 24–54.

[14] Arup Varma, Nenad Jukic, Almir Pestek, Clifford J Shultz, and Svetlozar Nestorov. 2016. Airbnb: Exciting innovation or passing fad? *Tourism Management Perspectives* 20 (2016), 228–237.

[15] Georgios Zervas, Davide Proserpio, and John Byers. 2015. A first look at online reputation on Airbnb, where every stay is above average. (2015).

[16] Georgios Zervas, Davide Proserpio, and John Byers. 2016. The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. *Boston U. School of Management Research Paper* 2013-16 (2016).

[17] Georgios Zervas, Davide Proserpio, and John W Byers. 2015. The impact of the sharing economy on the hotel industry: Evidence from airbnb's entry into the texas market. In *Proc. of EC '15*. ACM, 637–637.

Received April 2018; revised ???; accepted ???