



## BIROn - Birkbeck Institutional Research Online

Bai, S. and Bi, X. and Han, Chunjia and Yang, Mu and He, H. (2023) Exploring the over-time variation in customer concerns on sharing economy services. *Journal of Global Information Management* 31 (1), ISSN 1062-7375.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/49309/>

*Usage Guidelines:*

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>  
contact [lib-eprints@bbk.ac.uk](mailto:lib-eprints@bbk.ac.uk).

or alternatively

# Exploring the Over-Time Variation in Customer Concerns on Sharing Economy Services

Shizhen Bai, Harbin University of Commerce, China

Xinrui Bi, Harbin University of Commerce, China\*

Chunjia Han, Birkbeck, University of London, UK

Mu Yang, Birkbeck, University of London, UK

Hao He, Harbin University of Commerce, China

## ABSTRACT

The sharing economy represented by Airbnb has evolved rapidly. It is particularly important to identify and understand how consumer concerns change over time. As a result, this study employs structural topic modelling using room type and time as covariates to extract topics from 896,658 Airbnb reviews in London and to observe the variation in the prevalence of topics over time. The findings show that the topic proportion changed relatively sharply in the early years of Airbnb (2010-2013) and during the COVID-19 pandemic (2020-2022), but relatively smoothly in the middle period (2014-2019). This research also discovered that the proportion of topics on customers' special experiences has been decreasing while the proportion of topics on their overall experience has been increasing. This shift could be attributed to an increase in the number of professional hosts, which has accelerated the standardisation of the Airbnb service.

## KEYWORDS

Airbnb, London, Online Review, Overall Experience, Special Experience, Structural Text Mining, Topic Modelling, Topic Proportion

## INTRODUCTION

The sharing economy is a new socioeconomic system where connected individuals organize the distribution of excess capacity or resources sitting idle in exchange for a fee or other compensation among each other (Belk, 2014). In the sharing economy, idle resources are the first and most critical factor, which is the basis for resource owners and users to realize resource sharing (Baiké, 2022). Advances in information and communications technology (ICT), increased consumer awareness,

DOI: 10.4018/JGIM.321435

\*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

the proliferation of collaborative online communities and social commerce/sharing have combined to drive progress in the sharing economy (Wang & Zhang, 2012). As a result, the sharing economy will become the most significant factor in the social service industry. In the fields of accommodation, transportation, education services, life services, and tourism, excellent sharing economy companies continue to emerge, such as Airbnb, Uber, AAWork, VaShare, and Eatwith.

The focus area in this paper is Airbnb, the most well-known example in the shared accommodation industry. Since its inception in San Francisco in 2008, it has grown to 4 million hosts, welcoming over 1 billion guests in almost every country (Airbnb, 2022). In addition, Airbnb reported 102.1 million nights and experiences booked in the first quarter, surpassing pre-pandemic levels and exceeding expectations, and revenue was \$1.51 billion, up 70% from the previous year (Bursztynsky, 2022).

The improved development of the sharing economy, as represented by Airbnb, should be based on identifying and understanding consumer needs. However, it is important to note that these needs and concerns are not constant but change over time, making it crucial to recognize and comprehend these changes to better guide the sharing economy's development. Recently, customers are increasingly writing online reviews to share their lodging experiences (Ju et al., 2019), and these self-disclosed reviews can reflect the aspects that customers mainly care about (Hu et al., 2019). By collecting a large amount of textual Airbnb reviews, scholars have analyzed and investigated the most mentioned topics to identify the chief concerns of customers (Cheng & Jin, 2019; Gao et al., 2022; Tussyadiah & Zach, 2016), but they have ignored the effect of time on the popularity of the topics, which may also account for the different conclusions obtained from different studies. For example, the coronavirus disease 2019 (Covid-19) has caused considerable economic losses to peer-to-peer accommodation like Airbnb. In addition, some scholars have studied the impact of the pandemic on Airbnb consumer concerns (Bresciani et al., 2021; Lee & Deale, 2021; Liang et al., 2021), but these studies do not yield a continuous change in consumer concerns. So far, no research has systematically analyzed the changes in consumer concerns of Airbnb users during the entire development process.

To fill this gap, this study analyzed 896,658 Airbnb reviews from April 14, 2010 to December 9, 2021 to reveal the changes in customer concerns in sharing economy. This study used a topic modelling approach, specifically the structural topic model (STM), for data analysis. The previous study confirmed the effectiveness STM for identifying customer concerns and revealing changes in concerns over time from customer reviews (Ding et al., 2020). Take the methodological advantage of STM (Roberts et al., 2016), two covariates was incorporated in the topic model, room type and review date, which can enable the author to monitor changes in Airbnb users' concerns over time well while taking into account the impact of room type on topics.

The objectives of the present study are: (1) to identify the concerns of London Airbnb users and the importance they place on different Airbnb attributes, (2) to reveal changes in London Airbnb user concerns over time. Concerning the above-mentioned research objectives, the study contributes to the sharing economy and peer-to-peer accommodation literature by enriching and extending the research on customers' concerns about the sharing economy through a systematic analysis of customer reviews. This study also shows the feasibility and effectiveness of using STM to gain insights into a large amount of user-generated content. Sharing economy managers can make corresponding adjustments according to the results and improve user satisfaction.

## **THEORETICAL BACKGROUND**

### **Customer Reviews and User Experience of Airbnb**

Airbnb is an Internet-based point-to-point platform (Guttentag, 2015), users share their accommodation experience by leaving reviews on the platform, numerous user reviews provide a new data source for scholars to study the user experience of Airbnb (Cheng & Jin, 2019; Gao et al., 2022; Tussyadiah & Zach, 2016). However, based on the previous research findings on Airbnb user experience, these studies

have drawn similar but sometimes contradictory conclusions. For, example, many studies suggested that “Value” is the key attribute of Airbnb customer lodging experience (Bridges & Vásquez, 2018; Lyu et al., 2019), whereas Cheng and Jin (2019) concluded that “Price” is not a key influencer. This discrepancy may be attributed to the impact of time on user well-being being overlooked in the former studies.

Few studies have examined user experience shifts for Airbnb customers. The outbreak of the Covid-19 has attracted the attention of academics, some of whom have studied the impact of the pandemic on the Airbnb user experience (Bresciani et al., 2021; Lee & Deale, 2021; Liang et al., 2021). Although these studies considered the impact of time on user experience, they cannot show the continuous change in user experience. For example, Ding et al. (2020) studied the continuous changes in the top six service quality attributes of Airbnb from 2014 to 2019, but because of the time limit of the sample, it cannot completely show the changes in the user experience in the whole development process of Airbnb.

### **Techniques for Textual Data Analysis**

The existing text analysis techniques primarily include word frequency, co-occurrence, and topic modeling methods. By analyzing and comparing the frequency of representative terms or phrases, the word frequency and co-occurrence method can extract useful information and insightful findings from a large corpus (Tussyadiah & Pesonen, 2016; Ju et al., 2019; Mendon et al., 2021). However, when using word frequency-based approach, each topic is represented by a list of manually selected keywords, making it difficult to uncover the trend of specific topics because of the challenge of identifying the temporal patterns of a single keyword that indicates a certain concept (Chen et al., 2017), an identical word in different documents may indicate different topics.

On the contrary, the topic modelling method can provide both topic-related keywords and proportional distributions of different topics across each document, allowing us to monitor changing patterns of customer perceptions of specific topics more effectively by analyzing proportional changes of topics rather than single ones. Latent Dirichlet Allocation (LDA) (Blei et al., 2003) is a well-known topic model designed to automatically organize many documents based on hidden topics, measured as the word co-occurrence patterns. However, although LDA can extract hidden thematic structures from text documents, it lacks additional document-level information, making it difficult to use LDA to examine the relationship between document metadata and the content of a document model (Roberts et al., 2016).

As an extended model of LDA, structural topic modeling (STM) can allow researchers to incorporate arbitrary document information which appears as covariates to estimate the per-document topic distributions (topic prevalence) and per-topic word distributions (topic content) (Roberts et al., 2016). STM has been applied to analyzing consumer comments, and scholars can set different covariates according to their own research purposes (Gao et al., 2022; Hu et al., 2019). For example, (Gao et al., 2022) set accommodation type, property rating, price, and comment sentiment as covariates, and used STM to analyze Airbnb consumer reviews to understand users’ preferences. The effectiveness of adding the time covariate to STM to observe the change of topic prevalence over time has been confirmed in previous studies (Ding et al., 2020; Yang & Han, 2021; Bai, He, Han, Yang, Yu, Bi, Gupta, Fan & Panigrahi, 2023). However, very few studies use STM and add the time covariate to analyze the change of the proportion of topics formed by Airbnb users’ reviews. Only (Ding et al., 2020) used STM and set the nationality and review date of Airbnb users as covariates to analyze the changing trend of topic prevalence.

## **DATA AND METHODS**

### **Data Collection**

The London Airbnb data for this study were acquired from the Inside Airbnb website (InsideAirbnb, 2022). London was chosen in this study because it has the highest number of Airbnb listings in

European cities in 2020 (Lock, 2020) and it is one of the top ten travel cities in the U.K. (Singh, 2019), the large number of accommodation listings and user reviews makes the dataset a good representative sample. The Airbnb dataset contains 1,043,004 customer reviews from April 14, 2010, to December 9, 2021. London has 66,641 listings and 44,695 hosts.

We show the review frequency in different periods in Figure 1. In this study, we classified hosts with two or more rental properties as professional hosts, while we classified hosts with only one as nonprofessional (Li et al., 2015). Table 1 shows that most Airbnb hosts in London are nonprofessional hosts, with only 16.09% being professional hosts. However, the difference in the number of listings is not very large, implying that many professional hosts own over two listings, which can be demonstrated in Figure 2.

### Text Preparation

The London Airbnb dataset includes the reviews' text content and metadata, such as listing id, host id, review creation time, reviewer id, reviewer name, room type, and price. Each review's text is considered a document, and the overall reviews are a whole corpus, a collection of individual documents. A compilation of typical common natural language processing techniques, detailed in Berliner et al. (2018), was used to pre-process the documents ensuring compatibility with unsupervised text mining techniques applied later. We processed the dataset with the following four steps. First, we removed non-English reviews. Although, there are multiple languages in the comments (such as English, French, Russian, and Chinese), our study focuses on a single language, English, to ensure consistency throughout. Second, we converted capital letters into lower case to get a more unified form and reduce the vocabulary size. Third, we removed numbers, stop words, punctuation marks, and

**Table 1.**  
 Summary of the number of Airbnb hosts and listings of the two host types in London

Host Type	Host Number	Host Proportion	Listing Number	Listing Proportion
Nonprofessional host	37505	83.91%	37505	56.28%
Professional host	7190	16.09%	29136	43.72%

**Figure 1.**  
 Yearly review frequency

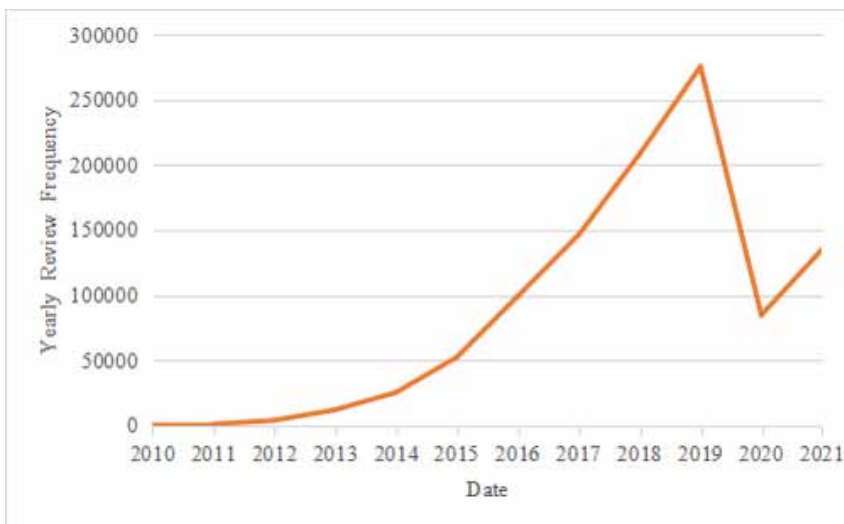
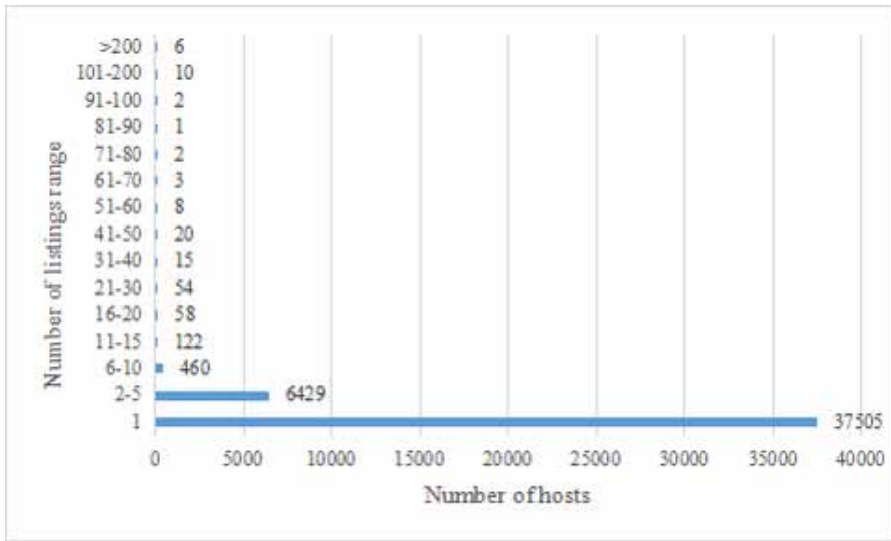


Figure 2. Distribution of the number of Airbnb hosts with different ranges of number of listings in London



extra white space. Fourth, text tokenization was undertaken and we removed words that appeared in under 30 documents to exclude insignificant words. Finally, our corpus comprised 896,658 processed review documents and 8294 unique words.

### Model Set-Up

Conventional document topic modeling uses the classical Latent Dirichlet Allocation (LDA) (Blei et al., 2003) topic model. Although the academic community has generally accepted this model, it is still inadequate for this study to explore the time-varying concerns of customers about sharing economy services. The STM (Structural Topic Model) (Roberts et al., 2014) is used in this work, and its covariate-based topic modeling characteristics are closer to the research objectives, allowing the trend of topic popularity over time to be observed with the help of covariates. The STM is an unsupervised machine learning method for identifying pattern features of topic distribution in documents. The model is based on the earlier Latent Dirichlet Distribution (LDA) topic model and incorporates metadata (covariates) that can explore the role of document metadata (covariates) concerning the document topic distribution. In this study, metadata refers to the information associated with the user reviews, e.g., posts listing id, review creation time, room type, and price. We show the basic schematic diagram in Figure 3.

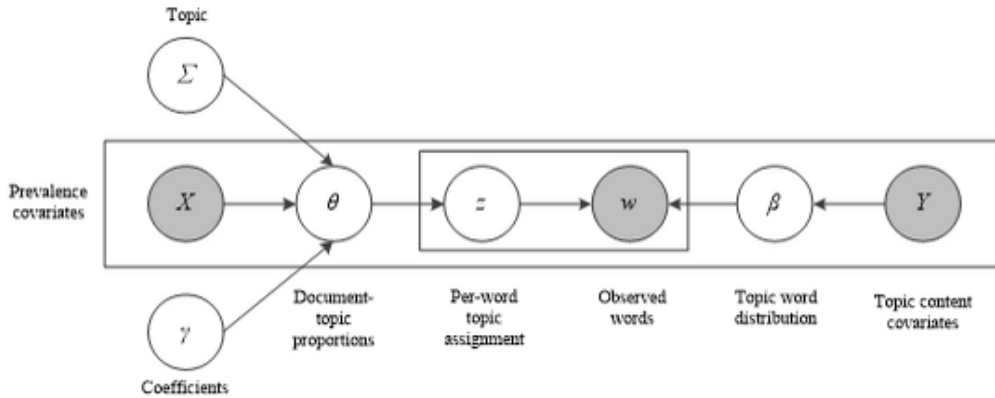
The STM is a hierarchical model in which a document  $d$ 's prevalence of each topic (denoted by  $\theta_d$ ) is drawn from a logistic-normal distribution whose mean is a function of document covariates  $X_d$ . For example, here, document  $d$  represents a review in our study:

$$\theta_d \sim \text{LogisticNormal}(X_d, \Sigma)$$

Then, given the topic-prevalence vector, one specific topic  $z_{d,n}$  is associated with the position which needs to be filled through the following process where  $n$  is the index of each word in document  $d$ :

$$z_{d,n} \sim \text{Multinomial}(\theta_d)$$

Figure 3.  
 A graphical illustration of STM using plate notation



Next, the words of each document  $w_{d,n}$  are assigned to the topics:

$$w_{d,n} \sim \text{Multinomial}(\beta_{d,z})$$

where  $\beta_{d,z}$  is the probability of choosing vocabulary word  $w$  to fill a position in document  $d$  given the topic assignment variable  $z$ . We used the *stm* package (Roberts et al., 2019) in the R programming language to set up the model for our analysis, inputting review text content as documents and unique words as vocabulary words.

We set the prevalence function as follows:

$$\text{prevalence} \sim \text{roomtype} + s(\text{reviewscreatedtime})$$

where  $s$  is the smooth function of time and room type is one of the topical prevalence covariates. According to the room information mentioned in the dataset, Airbnb rooms are divided into four types: “Entire home”, “Hotel room”, “Private room”, and “Shared room” (Table 2).

A challenging aspect of the topic modeling process is determining the number of topics. The *SearchK* function in the *stm* package measures the number of topics by combining different metrics, including Held-Out, Semantic Coherence, Residuals, and Lower Bound. The basic idea of Held-Out is to extract some words from a set of documents, and the training model uses document-level latent

Table 2.  
 Summary of the number of listings and Airbnb user reviews for four different room types in London

Room Type	Listing Number	Listing Proportions	Frequency	Frequency Proportions
Entire home	37472	56.23%	498436	47.79%
Hotel room	257	0.39%	4649	0.45%
Private room	28402	42.62%	533690	51.17%
Shared room	510	0.77%	6229	0.60%

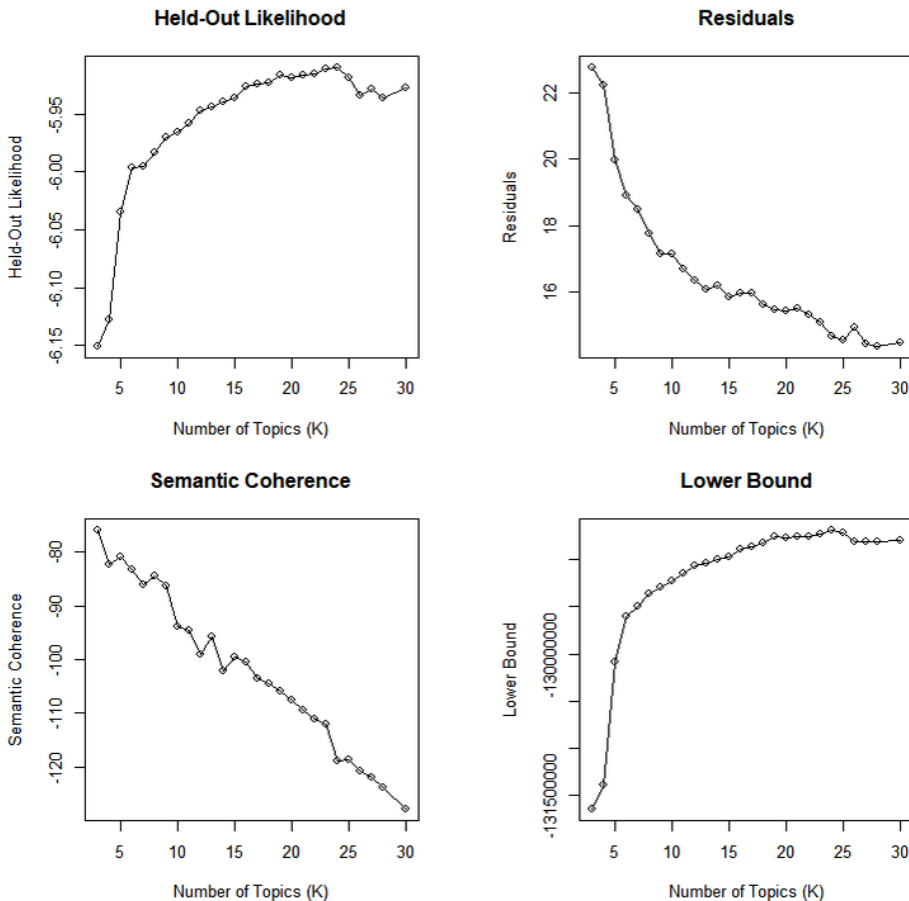
variables to evaluate the probability of retention. Semantic coherence is maximized when the most likely words to appear in a specific topic frequently occur together. In the STM data generation process, residuals are calculated as a test for over-dispersion of the polynomial variance; if the residuals are over-dispersed, the number of topics set is low, and more topics are required to absorb some additional variance. Finally, the Lower Bound indicates that convergence can be checked by approximating the variation. The higher the likelihood of document retention and semantic consistency, the lower the residuals and bounds, and the better the model performance.

We can see from Figure 4, after  $K=25$ , the change of held-out likelihood, residuals, and lower bound starts gentle and almost remains horizontal. In addition, when  $K=25$ , the values for the likelihood of document retention and semantic consistency are relatively high, and the values for residuals and bounds are relatively low, so  $K = 25$  is selected as an optimal number of topics.

### Topics

This model identifies 25 topics that best characterize the content of 896,658 reviews (Table 3). Each topic represents an underlying word distribution where every word in our documents is given a probability of assignment to that topic, as described in the previous section. The model outputs a list of top words for each topic, which are the words that have the highest probability of appearing in

Figure 4.  
Diagnostic values by number of topics





the topic but the least probability in other topics. Manually labeling the topic is a common practice of topic modeling (Gao et al., 2022). Thus, each topic’s top words and representative reviews were used in the topic labeling process. In addition, the naming of each topic also referred to previous Airbnb studies (Ding et al., 2020; Han & Yang, 2021).

By examining the representative reviews of 25 topics, we can find that most of the reviews contained in Topic 13 Room Condition for sleep and Topic 1 Issues are negative reviews from users, whereas most reviews in the other topics are positive. This result is consistent with the findings of

**Table 3.**  
**Topic summary and labeling**

Topic No.	Topic Label	Topic Proportions (%)	Top Words
	<i>Topic category: General experience</i>	61.71	
Topic 8	Good feeling	12.20	great, locat, place, clean, communic, stay, excel
Topic 2	Stay experience	8.84	stay, realli, nice, definit, recommend, love, enjoy
Topic 5	Home feeling	7.12	love, home, welcom, hous, beauti, feel, made
Topic 16	Comfortable feeling	5.06	comfort, well, quiet, spacious, transport, conveni, close
Topic 21	Future visit intention	4.56	time, london, will, visit, next, book, alway
Topic 12	Thankful experience	3.68	thank, much, experi, best, hospit, first, one
Topic 10	Perfect experience	3.25	perfect, need, everyth, work, anyth, weekend, smooth
Topic 11	Recommending feeling	3.20	recommend, high, quick, respons, question, respond, answer
Topic 24	Easy access	2.95	get, easi, even, around, find, check, better
Topic 7	Host review	2.79	travel, airbnb, want, live, person, one, like
Topic 15	Apartment service	2.32	apart, gave, inform, check, tip, lot, met
Topic 23	Satisfaction feeling	2.28	make, way, sure, provid, offer, breakfast, went
Topic 6	Honest advertisement	1.86	look, just, like, exact, pictur, describ, photo
Topic 25	London experience	1.60	london, also, area, night, stay, lot, coupl
	<i>Topic category: Location</i>	15.57	
Topic 17	Transportation	7.57	walk, station, tube, minut, close, london, away
Topic 4	Flat location	3.72	flat, restaur, street, shop, market, bar, area
Topic 14	Decent location	2.82	citi, near, nice, hous, can, london, far
Topic 20	Parking	1.46	park, ideal, free, base, secur, bonus, handi
	<i>Topic category: Facilities</i>	12.92	
Topic 9	Room facilities	4.18	room, bed, bathroom, comfi, clean, kitchen, privat
Topic 13	Room condition for sleep	3.36	bit, floor, bedroom, littl, night, window, build
Topic 1	Issues	2.96	issu, didn, work, one, wasn, check, howev
Topic 18	Kitchen accessories	2.42	kitchen, etc, use, two, studio, cook, small
	<i>Topic category: Service</i>	6.73	
Topic 3	Host attitude	4.76	host, help, friend, accommod, kind, extrem, autom
Topic 19	Booking and cancellation	1.97	day, arriv, late, reserv, cancel, post, earli
	<i>Topic category: Value</i>	3.07	
Topic 22	Good value	3.07	good, valu, clean, price, locat, overall, money

Note: All data in the Top words column in the table is output by R programming software. Each word may appear incomplete because the English text requires part of speech restoration to obtain a stem, which does not affect our understanding of its meaning.



reviews when making comments, and their comments could affect them. The Covid-19 pandemic may cause a change in the proportion of topics from 2020 to 2022.

## General Experience

The first and largest topic category is “General experience”, accounting for 61.71% of all reviews, which indicates that Airbnb users attach great importance to the sense of accommodation experience. The 14 topics in “General experience” describe the user’s experience with Airbnb. We give examples of the top six topics in this category in terms of the proportion of topics below:

*[Topic 8 Good feeling] Great location fab place to stay. Excellent host, great communication and easy check in process.*

*[Topic 2 Stay experience] Have a great stay at Michelle’s place, really nice room, nice people. I would definitely stay here again and would highly recommend to anyone who wants a nice clean place to stay.*

*[Topic 5 Home feeling] We had a lovely time at Claire’s home and were made to feel very welcome. Breakfast, a particular highlight with loganberries from the garden and homemade marmalade. Such a treat. Thank you Claire.*

*[Topic 16 Comfortable feeling] This flat is well situated for public transport links in a quiet residential area and is well equipped and comfortable.*

*[Topic 21 Future visit intention] Booked for my parents for a week, they were very happy and will be returning on their next visit to London.*

*[Topic 12 Thankful experience] By far one of my greatest AirBnB experience to date, Viven is such a caring person. If i could give stars out or i would have, as she totally deserves it and more. Thanks for a wonderful stay.*

The reviews and top words corresponding to the topics revealed that the topics in “General experience” can be divided into two types of experience: the overall experience (Topic 8, Topic 2, Topic 16, Topic 21, Topic 10, Topic 11, Topic 6, Topic 25) influenced by the four attributes of Location, Facilities, Service and Value, and the special experience (Topic 5, Topic 12, Topic 24, Topic 7, Topic 15, Topic 23) influenced primarily by the services provided by the host.

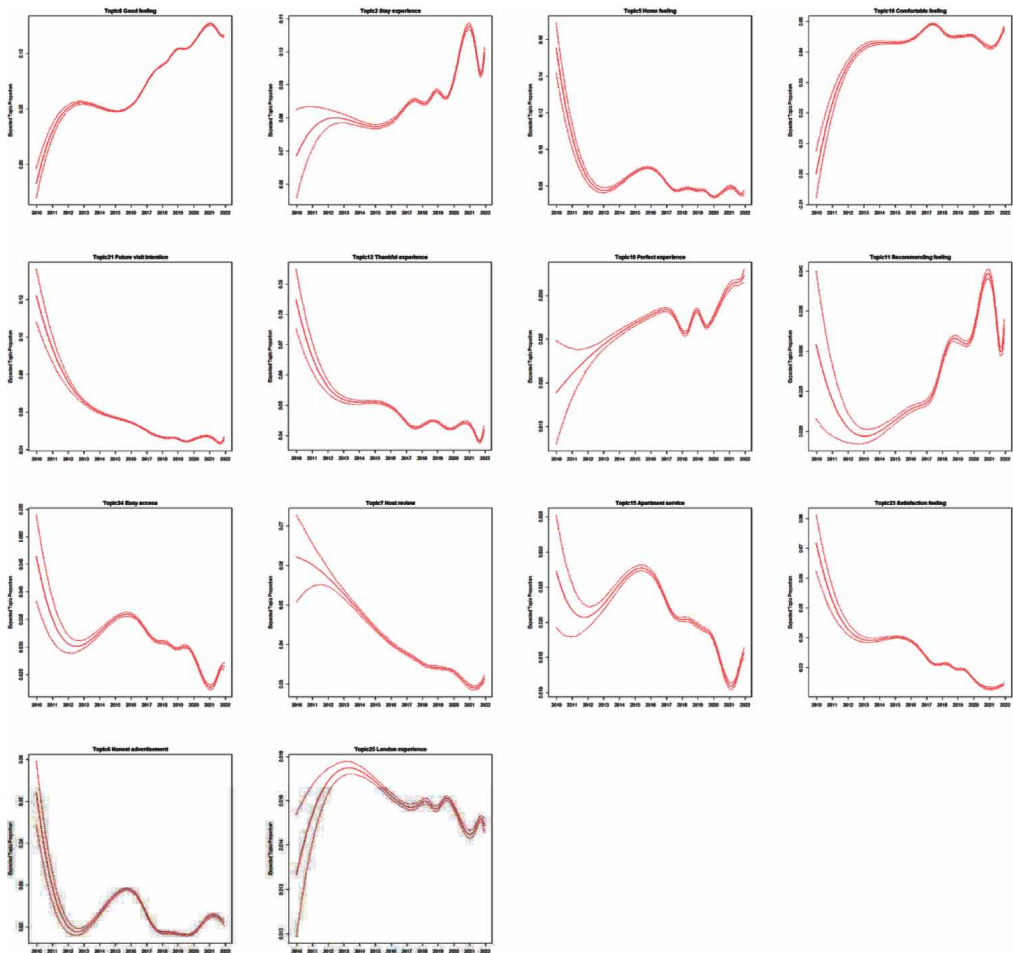
After assessing the proportion of topics, we chose the two highest subjects from each experience type as representatives for further analysis. Ultimately, Topic 8 Good Feeling (12.20%) and Topic 2 Stay Experience (8.84%) were selected as representatives of the overall experience, and Topic 5 Home Feeling (7.12%), and Topic 12 Thankful Experience (3.68%) were selected as representatives of the special experience.

Analyzing the deviation in the prevalence of four topics through the years (Figure 6), it is evident that, at the beginning of 2010, Topic 5 Home Feeling and Topic 12 Thankful Experience were comparatively high (for example, Topic 5 exceeded 14%), while Topic 8 Good Feeling and Topic 2 Stay Experience were relatively low. Following 2010, the proportion of both Topics 5 and 7 deteriorated, with an overall downward trend, while Topics 8 and 2 increased, with an overall upward trend.

In the early days of Airbnb, one of its most appealing features was the ability to offer an authentic tourist-host encounter that could not be replicated in conventional hotels (Tussyadiah, 2016); as a result, users focused on describing the unique experience provided by differentiated services. Meanwhile, because the hosts are primarily individuals with little professional training and experience who use Airbnb as a side business, the role of the hosts in the Airbnb context ranges from being extremely useful to not helpful at all. As a result, the Airbnb experience can be more polarized for the guest than hotel operations and services. As a result, there were relatively few comments on the overall good experience at that time.

In response to the varying quality of host service, Airbnb implemented several measures to improve the quality of service, including the global launch of the Airbnb Plus business, the establishment of requirements for standardized services to hosts, and the recruitment of more hotel professionals to this platform. More and more hosts manage multiple listings on Airbnb (Table 1), and they no longer use

Figure 6. Topic prevalence over time of general experience



Airbnb as a sideline (or ‘gig’) job but have become professional hosts. As can be seen in Figure 2, there are already 32 hosts operating over 50 listings in London, with the maximum number of hosts owning listings reaching 845, which is the size of a standard hotel, implying that the hosts in charge are no longer individuals, but rather more professional operating companies. These measures made Airbnb’s services more professional and standardized, effectively raising customer satisfaction and increasing comments about the overall positive experience. However, the perception of personalized experience provided by standardized services diminished, which caused a decline in the proportion of Topic 5 and Topic 12.

### Location

“Location” is the second-largest topic category, containing four topics and accounting for 15.57% of all reviews. The topics included in “Location” describe the accommodation’s convenience. As the largest topic in “Location”, Topic 17 Transportation discussed the convenience of accommodation location for transportation (e.g. bus stop, train, underground station). Topic 4 Flat Location and Topic 14 Decent Location are all related to the convenience of the accommodation to major tourist attractions (e.g., Big Ben, London Eye) or points of interests (e.g. shop, cafe), where Topic 4 Flat Location focuses more on the environment around the flat, Topic 14 focuses more on the convenience

to the city center. Finally, topic 20 Parking mainly focuses on the accessibility of parking positions. We present an example of each of the four topics as follows:

*[Topic 17 Transportation] minutes bus ride from Heathrow terminal which cost you just. Close to king street bus station and famous Southall Gurudwara. Tesco express shops food joints are just minutes walking distance.*

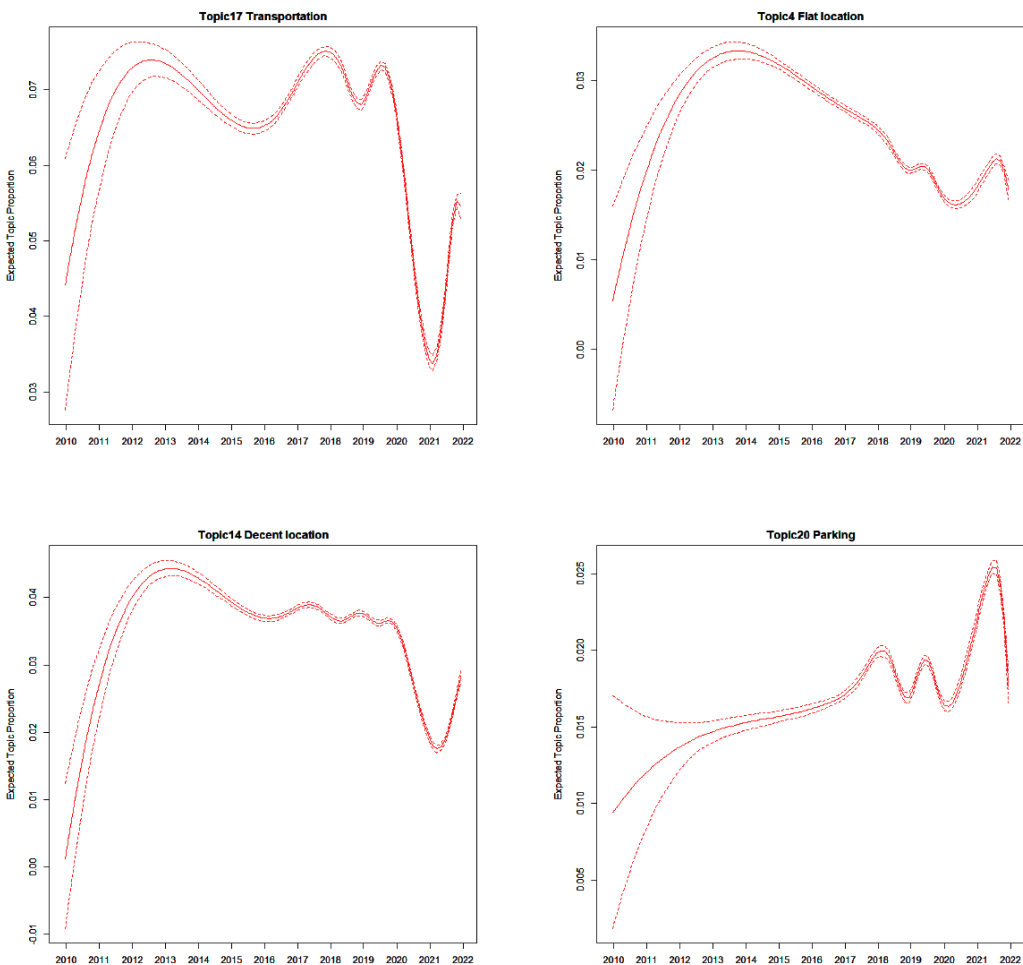
*[Topic 4 Flat location] Lovely and roomy flat and in a fantastic location, right in the west end and close to lots of restaurants and shopping areas.*

*[Topic 14 Decent location] The house is very near to the station and to all bus stops. There are a lots of shops and cafe and you can reach the center very easily. The house is clean and the landlady is very easygoing and available. Nice place*

*[Topic 20 Parking] Venue was handy to local shops. Ideal for an overnight stay and free parking a bonus.*

Observing the variation in the prevalence of four topics over time (Figure 7), it can be found that: From 2010 to 2013, the proportion of the four topics showed the same dramatic upward trend, and after

Figure 7. Topic prevalence over time of location



2013, the topic proportion changed and became relatively flat; From 2020 to 2022, the proportion of four topics again fluctuated more sharply, Topic 17, Topic 4 and Topic 14 showed roughly the same trend, the proportion of topics showed a downward trend from 2020 to 2021 and reached a minimal point at the beginning of 2021, after which the proportion of topics rose. In contrast to the trend of the previous three topics, Topic 20 shows an upward trend from 2020 to 2021 and reached a maximum point at the beginning of 2021, after which the topic proportion declined.

From 2010 to 2013, the increasing popularity of topics relating to Airbnb shows the growing appreciation of the convenient locations. This upsurge likely results from the numerous listings that entered the platform over this period.

The COVID-19 pandemic seems to impact the change in theme proportion from 2020 to 2022. The findings show less discussion of location convenience during the severe pandemic. One plausible explanation is that many popular sites in the city center were closed during the pandemic, making accommodation addresses near the city center less appealing to users. Another likely reason is that users actively chose the accommodation address to avoid the crowd to ensure their safety. This reason aligns with the finding of Liang et al. (2021), who discovered that tourists sought natural and less crowded destinations in the suburbs during the pandemic, resulting in a notable decentralized form.

Another interesting finding is that, while users discussed transportation convenience less, they discussed parking convenience more. Yet another plausible explanation is that fewer foreign visitors resulted in a decrease in long-distance travel. The mobility restrictions forced many tourists to embrace hyper-local approaches to travel and enjoy natural amenities (Dušek & Sagapova, 2021; Liang et al., 2021). Therefore, there was less discussion about the transportation modes used in long-distance travel, such as trains, planes, and subways. The potential reason is that to avoid physical contact and crowding, domestic tourists may have preferred the safer mode of travel by car. Therefore, users talked more about parking.

## Facilities

The third-largest topic category “Facilities” contains four topics, accounting for 12.92% of all reviews, including facilities of the room, the kitchen, and the night-time environment. Topic 9 Room Facilities and Topic 18 Kitchen Accessories are concerned with the room and kitchen facility configurations, respectively. Topic 13 Room condition for sleep focuses on the negative reviews of Airbnb users on the sleeping environment, mainly because noise interferes with people’s sleep. Finally, the facility issues users encountered during their stay dominate Topic 1. We present an example of each of the four topics as follows:

*[Topic 9 Room facilities] a clean comfortable en suite room with a shared living dining room, fridge to share with the guest in the other room.*

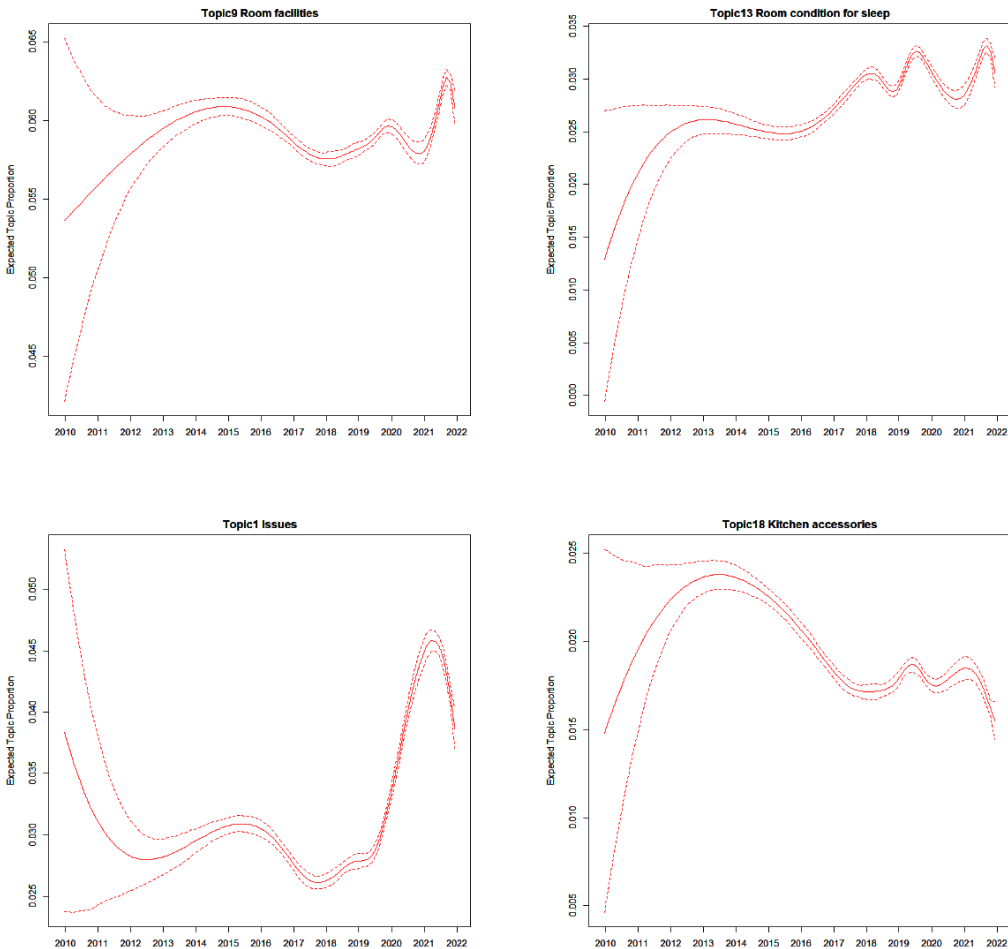
*[Topic 13 Room condition for sleep] Summer is a little hot, requiring open windows as in other places. Road noise can be an issue. Perhaps ear plugs.*

*[Topic 1 Issues] There were safety hazards in this house. The house was extremely dirty when we moved in there were repeated safety incidents. The bathtub in the upstairs bathroom had a bad leak which could have led to a dangerous slip. It was repaired but the problem recurred the next time I showered there.*

*[Topic 18 Kitchen accessories] All amenities including crockery, fork, spoon, pot, cups, electric kettle toaster, microwave oven and iron are available.*

The changing trends of the proportion of Topic 9 Room Facilities and Topic 18 Kitchen Accessories are roughly the same; both showed a dramatic upward trend from 2010 to 2013 (Figure 8), possibly because the hosts understood the user needs through reviews and improved the room facilities and kitchen. By observing the change in the topic proportion of the two topics from 2020

Figure 8. Topic prevalence over time of Facilities



to 2022, we can find that the topic proportions did not show a sharp fluctuation, which shows that Covid-19 had little impact on the two topics.

Topic 13 Room Condition for sleep showed a significant upward trend from 2010 to 2013, indicating an increase of users' complaints about the sleeping environment. The reason may be that the accommodation location was biased towards the city center in the early stage, so more traffic noise affected users' sleep. Since then, the proportion of Topic 13 has remained relatively stable, including the epidemic period from 2020 to 2022. Topic 1 Issues had a dramatic upward trend from 2010 to 2013, which could be attributed to the hosts' timely and effective resolution of the problems raised by the users. Unlike the other three topics, the proportion of Topic 1 fluctuated significantly between 2020 and 2022. The proportion of Topic 1 rose from 2020, reached the maximum point in 2021, and then declined, indicating the increased frequency of users reporting problems during the pandemic. Because of the closure of tourist attractions and travel restrictions during the pandemic, they moved many activities indoors, forcing people to spend more time in their rooms. At this time, it was important for users that the facilities function properly. If the facilities failed, it would harm the users' accommodation experience. In addition, during the severe pandemic period, hosts may not repair the facilities in time, which led to an increase in the proportion of negative reviews in user reviews.

## Service

The fourth topic category “Services” contains two topics, accounting for 6.73% of all reviews, mainly for services provided by hosts to users. Topic 3 Host Attitude is the users’ evaluation of the hosts’ attitude. Topic 19 Booking and Cancellation is the automatic posting of the system when the user cancels the reservation. We give examples of the two topics below:

*[Topic 3 Host attitude] excellent location with a friendly and respectful host whose conscientious and helpful.*

*[Topic 19 Booking and cancellation] The reservation was canceled days before arrival. This is an automated posting.*

Analyzing the fluctuations in the frequency of Topic 3 Host Attitude, it can be noted that unlike most other topics, the Topic 3 remained relatively flat from 2010 to 2013 (Figure 9), implying that evaluations of hosts did not change significantly during Airbnb’s initial stages, likely because Airbnb hosts have consistently provided superior service attitudes and have garnered users’ praise. The proportion of the topic has been growing since 2020 throughout the Covid-19 pandemic, except for a slight decrease after 2021, signifying a rise in the rate of positive user feedback about hosts’ attitudes during the pandemic. This is probably because of the attentive service of Airbnb hosts during the outbreak, which made users appreciate the friendly and helpful attitude of the hosts. The proportion of Topic 19 Booking and cancellations decreased significantly from 2010 to 2013, the proportion of order cancellation decreased, which may be related to the improvement of the overall experience of Airbnb users from 2010 to 2013 (Figure 6). Except for a slight increase after 2021, the proportion of Topic 19 has decreased since 2020 throughout the Covid-19 pandemic. One possible explanation is that travelers had considered the impact of the pandemic before booking, and the pandemic did not reduce the positive experience, but increased it, as evident by the spike in the ratio of Topic 8 and Topic 2 during this timeframe (Figure 6).

## Value

As the last topic category, “Value” contains only one topic (Topic 22 Good Value). “Value” is a relatively small category compared to other categories, accounting for 3.07% of all reviews. Topic 22 Good value is the user’s evaluation of Airbnb’s price. We present an example of Topic 22 in the following:

*[Topic 22 Good value] Very good location and good price compared to other options friendly staff.*

Figure 9. Topic prevalence over time of service

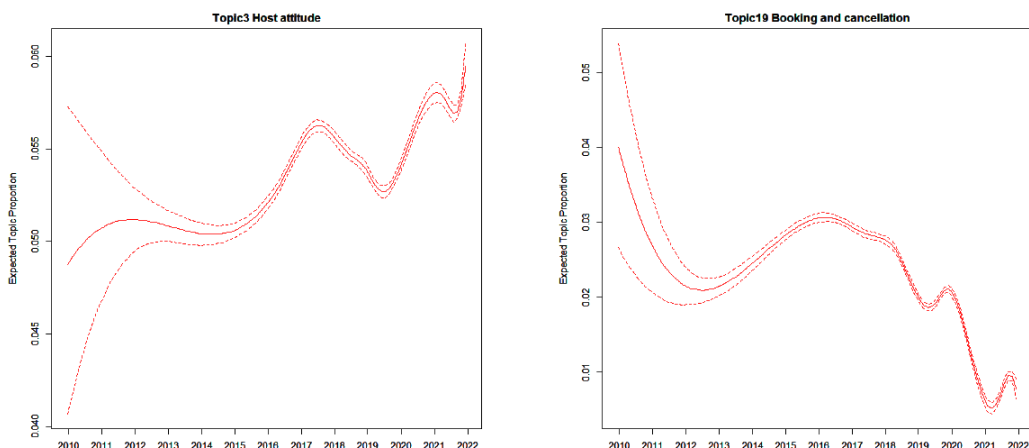
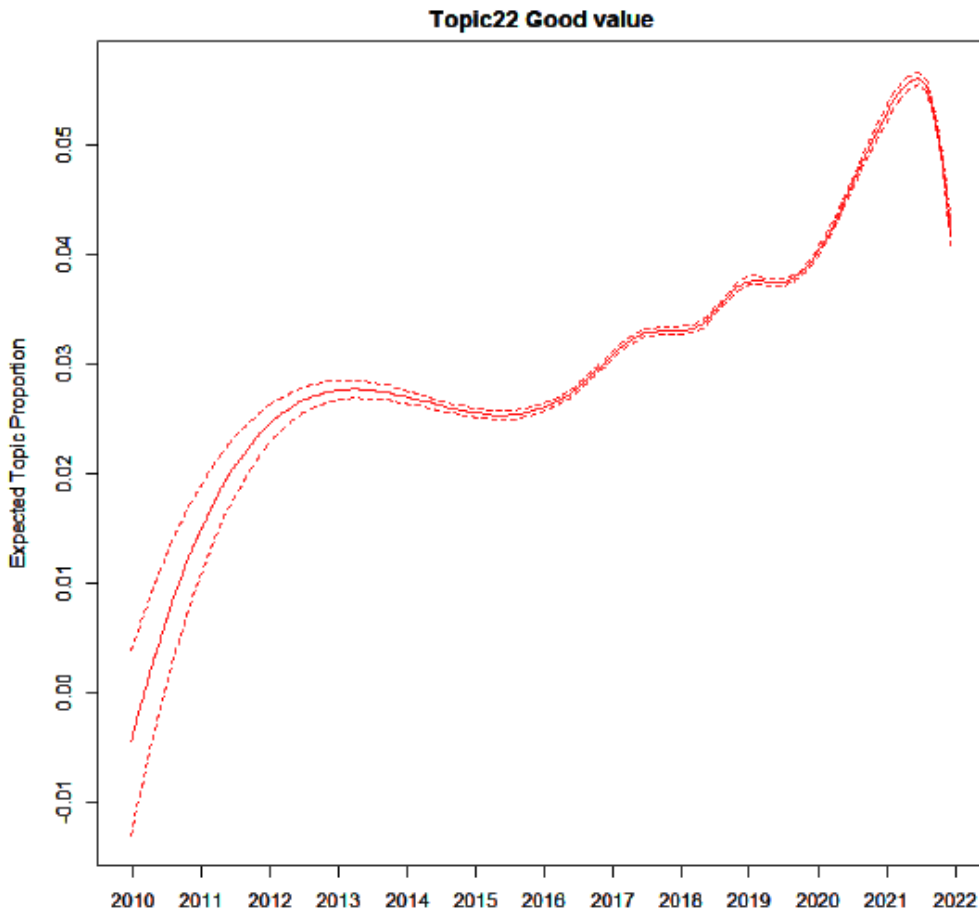




Figure 10. Topic prevalence over time of value



Regarding the change in their prevalence over time, Topic 22 showed an overall upward trend. The rapid increase in the proportion of Topic 22 between 2010 and 2013 (Figure 10) suggests that Airbnb gained early acceptance from users through price, thanks to the more competitive price compared to the hotel industry (Guttentag, 2015). During the pandemic, the discussion about the good value of Airbnb increased until the proportion decreased in the later stage of the pandemic, which could be attributed to the price adjustments made by hosts during the pandemic (David, 2022).

## THEORETICAL AND MANAGERIAL CONTRIBUTIONS

### Theoretical Implications

First, it demonstrates the feasibility and effectiveness of using an unsupervised machine learning approach fused with structural topic modeling to gain insights from a large amount of user-generated content in peer-to-peer accommodation research.

Second, this study enriches existing research on the concerns of sharing economy users. Using network analysis, the 25 topics formed by Airbnb users' reviews were divided into five categories: "General experience", "Location", "Facilities", "Service" and "Value," and determined the strength of user interest based on the proportions of topic categories. In addition, the content of the topics

included in each category were also analyzed from a micro perspective and found a clear positive bias in the reviews from most of the other topics, except for users' complaints about sleep environment and facility issues.

Third, this study extends the research on the change of user concerns in the sharing economy. According to the variation in the prevalence of topics over time, for most topics, the proportion changed relatively sharply in the early years of Airbnb (2010-2013) and during the pandemic (2020-2022), and relatively smoothly in the middle period (2014-2019). Combined with the product life cycle theory (Vernon, 1966), we have tried to develop a preliminary division of the development process of Airbnb as a service product. 2010-2013 is the introduction period of this new form of accommodation. In 2014, Airbnb entered a growth period. Until the outbreak of COVID-19 at the beginning of 2020, the undisturbed growth period of Airbnb suddenly ended. The impact of the pandemic on users' concerns results from complex external factors, so our research has important and enlightening significance for future research on the impact of such crises.

### **Practical Implications**

First, this study provides solutions for peer-to-peer accommodation practitioners to transform user reviews into useful business insights.

Second, we found that customers' discussions on their overall experience have been increasing while the proportion of the topics on customers' special experience has been reducing during the same time. We could attribute this shift to an increase in the number of professional hosts, who have sped up the standardization of Airbnb service, improving the customers' overall experience. This appears to be a departure from the original purpose of the sharing economy and may lead to Airbnb becoming less competitive than the traditional hospitality industry. It is, therefore, a question for practitioners to consider how to balance the development of professionalism in Airbnb with a unique experience.

Third, this study reveals the impact of the Covid-19 pandemic on users' concerns. Based on the change in user concerns in the "Location", it can be concluded that people are more likely to choose accommodation away from the crowds during the pandemic and are more likely to choose to travel by car. Furthermore, the proportion of reviews about facility issues has increased, possibly due to hosts being unable to fix faults promptly during the epidemic. Therefore, practitioners should pay attention to the changes in user focus during the epidemic. Although the change in customer concerns has recovered in late 2021 (Figure 6), the pandemic may have some long-term effects on the travel and tourism habits of people worldwide and the accommodation industry, so practitioners should adapt accordingly.

### **LIMITATIONS AND FURTHER RESEARCH**

First, the sample for this study is London Airbnb user review data, as policies in different countries significantly affect the development of Airbnb (Liang et al., 2021), so the findings of this study may have some geographical limitations. Future research could address applicability by collecting and analyzing Airbnb user reviews from other countries (e.g., China) for comparative analysis.

Second, due to time constraints, this study can only analyze changes in user concerns from the beginning of 2010 to the end of 2021, interspersed with the impact of the Covid-19 pandemic (an external factor) which affected the natural changes in Airbnb user concerns. Finally, future research could collect and analyze data from Airbnb user reviews following the outbreak to explore whether the impact of the pandemic will have a long-term effect on customer concerns and whether it will cause a trend change in the sharing economy.

Third, this study did not conduct sentiment analysis of Airbnb users' reviewers as previous studies have done, but rather analyzed examples related to the topics to infer user sentiment roughly, which may be subjective. Future research could add sentiment analysis to improve the persuasiveness of the findings.

## **CONCLUSION**

This study explores the over-time variation in customer concerns about sharing economy services. The London Airbnb dataset includes 896,658 user-generated reviews from April 14, 2010, to December 9, 2021. In terms of methodology, this study employs structural topic modeling with document covariates, with the two covariates of room type and time as a priori inputs allowing the model to cluster better than the LDA model and to obtain trends in the proportion of topics. Analyzing the research results, we discovered the topic proportion changed relatively sharply in the early years of Airbnb (2010-2013) and during the Covid-19 pandemic (2020-2022) and relatively smoothly in the middle period (2014-2019). This research also finds that customers' discussions on their overall experience increased while the proportion of the topics on customers' special experience reduced. The increase in professional hosts who have sped up the standardization of Airbnb service might trigger this change. These findings can assist sharing economy managers in making timely adjustments to consumer needs and improving consumer satisfaction, contributing to the sharing economy's continued development.

## REFERENCES

- Airbnb. (2022). *About us*. Retrieved June 7, 2022, from <https://www.airbnb.com.au/about/about-us>
- Bai, S., He, H., Han, C., Yang, M., Yu, D., Bi, X., Gupta, B. B., Fan, W., & Panigrahi, P. K. (2023). Exploring thematic influences on theme park visitors' satisfaction: An empirical study on Disneyland China. *Journal of Consumer Behaviour*, 1–17. 10.1002/cb.2157
- Baike. (2022). *Sharing economy*. Retrieved August 18, 2022, from <https://baike.baidu.com/item/sharing-economy/965155>
- Belk, R. (2014). You are what you can access: Sharing and collaborative consumption online. *Journal of Business Research*, 67(8), 1595–1600. doi:10.1016/j.jbusres.2013.10.001
- Berliner, D., Bagozzi, B. E., & Palmer-Rubin, B. (2018). What information do citizens want? Evidence from one million information requests in Mexico. *World Development*, 109, 222–235. doi:10.1016/j.worlddev.2018.04.016
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Bresciani, S., Ferraris, A., Santoro, G., Premazzi, K., Quaglia, R., Yahiaoui, D., & Viglia, G. (2021). The seven lives of Airbnb. The role of accommodation types. *Annals of Tourism Research*, 88, 103170. doi:10.1016/j.annals.2021.103170
- Bridges, J., & Vásquez, C. (2018). If nearly all Airbnb reviews are positive, does that make them meaningless? *Current Issues in Tourism*, 21(18), 2065–2083. doi:10.1080/13683500.2016.1267113
- Bugalski, Ł. (2020). The Undisrupted Growth of the Airbnb Phenomenon between 2014–2020. The touristification of European cities before the COVID-19 Outbreak. *Sustainability*, 12(23), 9841. doi:10.3390/su12239841
- Bursztynsky, J. (2022). *Airbnb beats estimates with 70% revenue growth as travel rebounds*. Retrieved June 7, 2022, from <https://www.cnbc.com/2022/05/03/airbnb-abnb-earnings-q1-2022.html>
- Chen, H., Zhang, G., Zhu, D., & Lu, J. (2017). Topic-based technological forecasting based on patent data: A case study of Australian patents from 2000 to 2014. *Technological Forecasting and Social Change*, 119, 39–52. 10.1016/j.techfore.2017.03.009
- Cheng, M., & Jin, X. (2019). What do Airbnb users care about? An analysis of online review comments. *International Journal of Hospitality Management*, 76, 58–70. doi:10.1016/j.ijhm.2018.04.004
- Chiesi, A. M. (2001). Network analysis. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 10499–10502). Elsevier. doi:10.1016/B0-08-043076-7/04211-X
- David, B. G. (2022). Heterogeneous price adjustments among Airbnb hosts amid COVID-19: Evidence from Barcelona. *International Journal of Hospitality Management*, 102, 103169. doi:10.1016/j.ijhm.2022.103169 PMID:35095167
- Ding, K., Choo, W. C., Ng, K. Y., & Ng, S. I. (2020). Employing structural topic modelling to explore perceived service quality attributes in Airbnb accommodation. *International Journal of Hospitality Management*, 91, 102676. doi:10.1016/j.ijhm.2020.102676
- Dušek, R., & Sagapova, N. (2021). Effect of the COVID-19 global pandemic on tourists' preferences and marketing mix of accommodation facilities – Case study from Czech Republic. *SHS Web of Conferences*, 92, 1009. doi:10.1051/shsconf/20219201009
- Gao, B., Zhu, M., Liu, S., & Jiang, M. (2022). Different voices between Airbnb and hotel customers: An integrated analysis of online reviews using structural topic model. *Journal of Hospitality and Tourism Management*, 51, 119–131. doi:10.1016/j.jhtm.2022.03.004
- Guttentag, D. (2015). Airbnb: Disruptive innovation and the rise of an informal tourism accommodation sector. *Current Issues in Tourism*, 18(12), 1192–1217. doi:10.1080/13683500.2013.827159
- Han, C., & Yang, M. (2021). Revealing Airbnb user concerns on different room types. *Annals of Tourism Research*, 89, 103081. doi:10.1016/j.annals.2020.103081

Hu, N., Zhang, T., Gao, B., & Bose, I. (2019). What do hotel customers complain about? Text analysis using structural topic model. *Tourism Management*, 72, 417–426. doi:10.1016/j.tourman.2019.01.002

InsideAirbnb. (2022). *London, England, United Kingdom*. Retrived June 20, 2022, from <http://insideairbnb.com/get-the-data>

Ju, Y., Back, K., Choi, Y., & Lee, J. (2019). Exploring Airbnb service quality attributes and their asymmetric effects on customer satisfaction. *International Journal of Hospitality Management*, 77, 342–352. doi:10.1016/j.ijhm.2018.07.014

Lee, S. H., & Deale, C. (2021). Consumers' perceptions of risks associated with the use of Airbnb before and during the COVID-19 pandemic. *International Hospitality Review*, 35(2), 225–239. doi:10.1108/IHR-09-2020-0056

LiJ.MorenoA.ZhangD. J. (2015). Agent behavior in the sharing economy: Evidence from Airbnb (Ross School of Business Paper No. 1298). SSRN. 10.2139/ssrn.2708279

Liang, S., Leng, H., Yuan, Q., & Yuan, C. (2021). Impact of the COVID-19 pandemic: Insights from vacation rentals in twelve mega cities. *Sustainable Cities and Society*, 74, 103121. doi:10.1016/j.scs.2021.103121 PMID:34540564

Lock, S. (2020). *Number of Airbnb listings in selected European cities as of December 2020*. Retrieved April 18, 2022, from <https://www.statista.com/statistics/815145/airbnb-listings-in-europe-by-city/>

Lyu, J., Li, M., & Law, R. (2019). Experiencing P2P accommodations: Anecdotes from Chinese customers. *International Journal of Hospitality Management*, 77, 323–332. doi:10.1016/j.ijhm.2018.07.012

Mendon, S., Dutta, P., Behl, A., & Lessmann, S. (2021). A hybrid approach of machine learning and lexicons to sentiment analysis: Enhanced insights from twitter data of natural disasters. *Information Systems Frontiers*, 23(5), 1145–1168. doi:10.1007/s10796-021-10107-x

Roberts, M. E., Stewart, B. M., & Airoidi, E. M. (2016). A model of text for experimentation in the social sciences. *Journal of the American Statistical Association*, 111(515), 988–1003. doi:10.1080/01621459.2016.1141684

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). stm: An R package for structural topic models. *Journal of Statistical Software*, 91(2). Advance online publication. doi:10.18637/jss.v091.i02

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58(4), 1064–1082. doi:10.1111/ajps.12103

Singh, S. (2019). *10 Best Cities in UK to Visit | Major Cities in UK*. Retrieved April 18, 2022, from <https://www.adequatetravel.com/blog/list-of-best-cities-in-united-kingdom-to-visit/>

Tussyadiah, I. P. (2016). Factors of satisfaction and intention to use peer-to-peer accommodation. *International Journal of Hospitality Management*, 55, 70–80. doi:10.1016/j.ijhm.2016.03.005

Tussyadiah, I. P., & Pesonen, J. (2016). Impacts of peer-to-peer accommodation use on travel patterns. *Journal of Travel Research*, 55(8), 1022–1040. doi:10.1177/0047287515608505

Tussyadiah, I. P., & Zach, F. (2016). Identifying salient attributes of peer-to-peer accommodation experience. *Journal of Travel & Tourism Marketing*, 34(5), 636–652. doi:10.1080/10548408.2016.1209153

Vernon, R. (1966). International investment and international trade in the product cycle. *The Quarterly Journal of Economics*, 80(2), 190–207. doi:10.2307/1880689

Wang, C., & Zhang, P. (2012). The evolution of social commerce: The people, management, technology, and information dimensions. *Communications of the Association for Information Systems*, 31. Advance online publication. doi:10.17705/1CAIS.03105

Yang, M., & Han, C. (2021). Revealing industry challenge and business response to Covid-19: A text mining approach. *International Journal of Contemporary Hospitality Management*, 33(4), 1230–1248. Advance online publication. doi:10.1108/IJCHM-08-2020-0920

*Shizhen Bai is a professor in the School of Management at the Harbin University of commerce, Harbin. His research interests are supply chain finance, marketing and brand.*

*Xinrui Bi is a master in the School of Management at the Harbin University of Commerce, Harbin. Her research interests are open innovation and user behaviour analysis.*

*Chunjia Han is Senior Lecturer in Business Planning and Innovation with over 10 years of experience of working with user innovation, open innovation, and crowd sourcing, both as a practitioner and an academic.*

*Mu Yang is Senior Lecturer in Business Strategy and Innovation with expertise in data analytics, digital innovation, and user behaviour analysis.*

*Hao He is a master in the in the School of Management at the Harbin University of Commerce, Harbin. His research interests are complex group decision-making, tourism management and consumer satisfaction.*