

How Does Customer Feedback Vary Across Leading Airline Review Platforms?

Adeim Suvanbekova^{1,2}, Christian González-Martel³, Teresa Aguiar-Quintana^{4,2}

¹Universidad de Las Palmas de Gran Canaria,

²TIDES Institute,

³Department of Quantitative Methods in Economics, Universidad de Las Palmas de Gran Canaria,

⁴Department of Economics and Business Management, Universidad de Las Palmas de Gran Canaria,

Corresponding author: Teresa Aguiar-Quintana,

Abstract

This study investigates score heterogeneity in online consumer reviews (OCRs) for airlines across TripAdvisor and Skytrax. The research addresses the variability in review scores, focusing on two airlines: Ryanair and Lufthansa. Previous studies have largely ignored cross-platform score variability, and this study fills that gap by applying a Bayesian hierarchical model. This model is suitable for its ability to incorporate prior knowledge and handle complex data structures. The dataset spans from January 1, 2023, to December 31, 2023. Results indicate no significant differences in review scores between platforms or airlines, suggesting consistent customer satisfaction. However, a significant interaction effect between TripAdvisor and Ryanair indicates lower scores for this combination. These findings highlight the importance of platform-airline interactions in customer feedback analysis. The study's implications underscore the need for tailored strategies in managing OCRs and validate the use of Bayesian methodologies in hospitality research. Future research should expand the dataset for broader validation.

Plain Language Summary

This study looks at how customer reviews for airlines differ between TripAdvisor and Skytrax. We analyzed reviews for Ryanair and Lufthansa using a special statistical method to see if there are big differences in ratings. We found that, on average, ratings are similar across both platforms and airlines. However, Ryanair reviews on TripAdvisor tend to be lower. These findings can help airlines understand and improve their online reputation.

1 Introduction

As technology continues to advance industries globally and digitalization transforms operations, the hospitality sector has undergone significant changes. This evolution, driven by Web 2.0 technologies, has given rise to online review platforms featuring user-generated content (UGC), providing valuable insights for prospective customers and fundamentally reshaping customer engagement with businesses. Thus, online consumer reviews (OCRs) have begun to significantly influence consumer decision-making processes (Mendes-Filho et al., 2018), shaping brand image (Lam et al., 2020), and enhancing electronic Word-of-Mouth (eWOM) (Muda & Hamzah, 2021; Nusairat et al., 2021). Furthermore, within the academic field, scholars are increasingly recognizing the advantages of applying UGC due to its accessibility and capacity for in-depth analysis. Consequently, there has been a notable growth in scholarly research incorporating UGC.

Given the unique characteristics of airline services, both in-flight and on-ground, effective customer communication and feedback are indicators of the operational efficiency and overall customer experience. Existing literature on airline services heavily relies on surveys as the primary data source to explore perceptions of service quality and customer satisfaction. The recent increasing integration of online reviews into scholarly research is attributed to accessibility, diverse perspectives, and rich data content, thereby overcoming the sampling limitations associated with traditional surveys. The incorporation of UGC within the airline industry has been widely utilized for various research aims, such as assessing passengers' emotions following flight delays (Song et al., 2020), investigating the impact of airline alliance strategy on service quality (Migdadi, 2022), and predicting airline recommendations. (Jain et al., 2022; Siering et al., 2018). However, despite the growing literature utilizing customer-shared experiences, there remains a gap in understanding the variability in online customer ratings within the airline industry. To address this gap, this study proposes a Bayesian approach for measuring and managing score heterogeneity across major online review platforms.

(Methodology/Need to check) Recently, (Martel-Escobar et al., 2023) noted a lack of research on the application of Bayesian techniques in publications within the tourism and hospitality sectors, particularly in studies focused on eWOM. While Bayesian approaches have found extensive adoption across various fields such as finance, healthcare, and marketing over the past decades, their application in hospitality research remains limited. Despite numerous advantages, including the incorporation of prior data, flexibility, and the ability to handle small sample sizes, the integration of Bayesian methods in hospitality studies remains constrained. According to Bianchi & Heo (2021) Bayesian statistics provide a more nuanced analysis compared to classical methods, which is particularly beneficial for the hospitality sector. Through the analysis of data obtained from primary OCR platforms for airlines, namely TripAdvisor and Skytrax, the study presents a novel Bayesian approach to understanding score heterogeneity among the principal review portals for airlines.

The airline industry has experienced notable transformations following the introduction of Low-Cost Carriers (LCCs), which have impacted passengers' perceptions and shaped distinct expectations according to airline's business model. The evaluation of services provided by LCCs differ compared to Full-Service Carriers (FSCs) due to disparities in airfare pricing and customer anticipations. This distinction is underscored by the customer satisfaction coefficient, revealing the heightened importance of different attributes (Byun & Lee, 2016). Thus, to conduct a comprehensive analysis, both LCCs and FSCs will be examined.

Given research presents the methodology and findings of Bayesian analysis, based on data collected from leading online review platforms for both LCCs and FSCs. Through in-depth approach, the study not only contributes to the existing literature on UGC and airline service quality but also brings attention to rating heterogeneity, offering a comprehensive understanding of score variations. These insights are valuable for theoretical scholars utilizing UGC for their analysis, as the majority of studies tend to concentrate on a single UGC platform, while also holding practical implications for stakeholders seeking to enhance service quality, foster eWOM, and manage brand image effectively within the highly competitive industry.

2 Literature Review

2.1 The importance of online consumer reviews in tourism.

Online reviews have captured considerable attention of both academia and industry due to their significant impact on consumer behavior and purchasing decisions. Thus, UGC has become increasingly essential in the tourism sector, particularly considering the intangible nature of hospitality and the growing reliance on reviews for travel planning. Notably, such content is considered more credible than official sites and tourist destination pages (Bilos et al., 2022), influences travelers' behavior more than marketing-generated-content (Tsiakali, 2018), and positively influences tourists' intentions to book a hotel and visit a destination (Ukpabi & Karjaluoto, 2018). Research by Cox et al. (2009) in the accommodation sector reveals that the majority of consumers prioritize the views of other travelers over the hotel's self-description available online. Furthermore, UGC has more credibility and trustworthiness compared to professionally generated content, and this trustworthiness remains consistent regardless of whether participants encounter positive or negative information (Cheong & Morrison, 2008).

According to Muritala et al. (2020), the research on online reviews and electronic eWOM in tourism and hospitality is reaching a mature stage, characterized by a rapidly growing volume of studies built on a solid base of seminal research and the application of the latest analytical techniques, including artificial neural networks and deep learning. Academics have extensively studied UGC in relation to travel planning behaviour (Bilos et al., 2022; Cox et al., 2009; Mendes-Filho et al., 2018; Tsiakali, 2018), corporate reputation management (Baka, 2016; Nicoli & Pa-

padopoulou, 2017), motivations for content contribution (Bronner & Hoog, 2011; Daugherty et al., 2008; Kaur & Singh, 2021; Yoo & Gretzel, 2009), credibility perception (Cheong & Morrison, 2008; Filieri et al., 2021; Lo & Yao, 2019), helpfulness determinants (Filieri et al., 2021; Hu & Yang, 2020; Huang et al., 2015; Kwon et al., 2021; Lo & Yao, 2019), predicting recommendations (Jain et al., 2022; Siering et al., 2018), and service recovery (Etemad-Sajadi & Bohrer, 2019; Xu et al., 2019).

The increasing adoption of OCRs in scholarly research is attributed to numerous advantages. Firstly, it provides direct access to extensive datasets with organically generated and diverse data. In certain instances, OCRs enable researchers to gather socio-demographic details of content contributors, supporting cross-cultural studies and analyses of age or gender disparities (Dike et al., 2024; Punel et al., 2019; Stamolampros et al., 2019; Sudhakar & Gunasekar, 2020). **(DO I NEED TO PUT ARTICLES?)**. Additionally, given the unique nature of airline services, reviews often contain valuable information on cabin class, airline type, and routes for evaluating moderating effect and comparative analysis, along with additional ratings on certain service attributes for extended examination (Lucini et al., 2020; Sezgen et al., 2019; Tahanisaz & shokuhyar, 2020). Furthermore, reviews allow researchers to track changes in customer perceptions, preferences, and behaviors in response to significant external factors such as the COVID-19 pandemic (Hassan & Salem, 2022; Nilashi et al., 2022; Sulu et al., 2022) and economic crises. However, study by W. Lu & Stepchenkova (2015) underscores that the anonymity of UGC creators, sample representativeness and the generalizability of results poses challenges to research validity, thereby questioning the reliability of findings.

Acknowledging the importance of online reviews, scholars have begun investigating their credibility and willingness to use. Research conducted by Lo & Yao (2019) revealed higher perceived credibility scores among reviews that are negative, written by experts and with a consistent rating. Consumers seek the opinions of others to minimize risks and gather prepurchase information, making these opinions more valuable than advertising. According to Bilos et al. (2022) respondents rely on UGC more than their friends' recommendations and promotional content. For analyzing the intention to use online consumer reviews, scholars rely on the Technology Acceptance Model (TAM) (Davis, 1989) as a predominant framework in hospitality literature, which identifies perceived usefulness and ease of use as key factors. In addition, Changchit et al. (2020) identified perceived credibility and importance of online reviews, along with computer self-efficacy, as additional factors influencing users' intentions. According to Mendes-Filho et al. (2018), there is a heightened intention to use UGC in travel planning if the content provides perceived empowerment and usefulness. The study by Assaker (2020) found that gender and age significantly moderate travellers' intention to use UGC, with perceived usefulness being more influential for male travellers and perceived ease of use being more impactful for female and older travellers.

Online travel review authors are primarily motivated by a desire to help service providers, concern for other consumers, and the need for enjoyment and positive self-enhancement (Yoo & Gretzel, 2009). Kaur & Singh (2021) analysing motivations behind review submission, defined altruism, reciprocity, egocentric behavior, knowledge self-efficacy, and collectivism as a key antecedents. Identifying profile of vacationers contributing to review sites, Bronner & Hoog (2011) reveal them to typically be under 55 years, from high and lower-middle-income groups, and often couples with or without children. The authors highlighted the importance of distinguishing between self-directed and other-directed motivations, which influence the type of review site and the manner of expression. Self-directed motivations favor marketer-generated sites, commenting on fewer aspects, predominantly negative reviews, and contributing to less accessible platforms. Vacationers with other-directed motivations prefer consumer-generated sites, engaging in broader commenting, posting mainly positive

reviews, using both text and ratings, and contributing to sites accessible to other vacationers. According to Wilson et al. (2012), national attitudes and motivations significantly influence the type of UGC posted. In their study, authors found that British and Swiss users tend to share photos and videos on social media to relive vacation experiences, while Spanish users prefer using text and comments to make recommendations to other travelers.

With the widespread availability of online reviews and the consequent information overload, scholars have begun to study reviews in relation to their perceived helpfulness. However, despite extensive research into the determinants of review helpfulness within tourism and hospitality, findings remain mixed. A recent study by Hu & Yang (2020) identified review length as the most influential factor, followed by reviewer expertise and age, while the impact of review valence and readability was found to be insignificant. Conversely, Filieri et al. (2018) contended that long reviews are not necessarily perceived as helpful. Instead, consumers favor fact-based, objective, and logically structured reviews especially in the context of high involvement purchase decisions. According to Kwok & Xie (2016), the helpfulness of online hotel reviews is negatively affected by review length and rating, whereas positively impacted by manager response and reviewer experience. Huang et al. (2015) determined the relationship between word count and helpfulness remained significant for reviews containing 144 words or fewer. However, once the word count exceeded, the relationship lost statistical significance. Analysing extremely negative ratings, Filieri et al. (2019), found that these ratings are often considered helpful and diagnostic, as they provide valuable insights to justify the extreme evaluation, thereby significantly contributing to consumer decision-making.

2.2 The research on OCR applied to airlines industry

Despite the growing adoption of reviews in scholarly research, studies specifically addressing online reviews related to airlines remain limited. Analyzing service quality within tourism and hospitality industry using UGC, W. Lu & Stepchenkova (2015) discovered that the majority of studies focused on the lodging setting, with relatively fewer investigations conducted in the airline, online travel agent, and restaurant settings. Similarly, an analysis of online customer evaluations in tourism and hospitality by Schuckert et al. (2015) revealed that over half of the reviewed articles concentrated on hotels. Existing research of airline service through online customer-shared experiences are primarily focused on perceptions of service quality (Brochado et al., 2023; Korfiatis et al., 2019; Rasool & Pathania, 2021), customer satisfaction (Ban & Kim, 2019; Chatterjee et al., 2023; Dike et al., 2024; Lucini et al., 2020), and behavioural intentions (Siering et al., 2018; Wang et al., 2023).

Consumer feedback platforms have undergone a significant expansion, gaining increasing popularity with the advent of Web 2.0. Online reviews assessing airline services within the scholarly literature primarily rely on data sourced from Skytrax and TripAdvisor platforms, with some studies also incorporating UGC sourced from social media, particularly Twitter. (DO I NEED TO ADD TWITTER ARTICLES). A study by L. Lu et al. (2023) underscores the importance of incorporating content from multiple review channels to gain a more accurate understanding of service quality, recognizing the potential disparities across platforms. The study on US airlines revealed that Skytrax reviewers prioritize boarding and baggage experiences, whereas Twitter users emphasize customer service interactions. Bilos et al. (2022) providing a comparative analysis with Airbnb, found that TripAdvisor reviews are perceived as more valuable for making purchasing decisions. Through conducting a comparative analysis between results sourced from Twitter and the Department of Transportation's official service quality report for US airlines, (Tian et al., 2020) **SEARCH REFERENCE** established the reliability of service quality metrics derived from social media.

Cross-cultural investigations in the field have uncovered differences in the evaluation of provided services, influenced by the cultural backgrounds of reviewers. Punel et al. (2019) revealed the influence of passengers' country of residence to travel experience, perception, and the evaluation of airline services. Notably, North American passengers prioritize cost over in-flight services and complain more about their national airlines, while East and Southeast Asian passengers are satisfied with Asian airlines and value in-flight services. Middle Easterners are critical of non-local airlines, while Oceania travelers expect high service standards and South Americans have lower expectations for in-flight amenities with an emphasis on price. According to Dike and Davis (2023) **SEARCH REFERENCE** passengers from individualistic cultural backgrounds tend to perceive service quality as lower compared to customers from collectivist cultural backgrounds. Stamolampros et al. (2019) found that airline passengers' satisfaction with service quality varies depending on their cultural values, impacting both overall satisfaction and specific aspects such as staff interactions, food quality, seat comfort, and cleanliness. Additionally, the study discovered that cultural characteristics significantly shape both numerical ratings and textual content in reviews.

The existing literature examining the airline industry underscores distinct disparities between Low-Cost Carriers (LCCs) and Full-Service Carriers (FSCs), emphasizing the significant impact of airline category on passenger expectations and perceptions of service quality. Consequently, each type of carrier should strategically prioritize specific features to enhance overall service delivery. According to Lim & Lee (2020), FSCs should focus on enhancing tangible features, while LCCs should concentrate on reliability. Furthermore, variations in recommendation drivers exist across different airline business models. (Siering et al., 2018) revealed that LCCs should focus on improving augmented service aspects like amenities to enhance recommendations, while FSCs should prioritize core service aspects such as in-flight amenities and customer service quality.

2.2.1 OCR studies based on data collected in Tripadvisor.

Tripadvisor, the world's largest travel guidance platform, features over 1 billion reviews and is accessible in 43 markets and 22 languages (Tripadvisor, 2024) **SEARCH REFERENCE**. In addition to its general community guidelines, the platform has established specific posting guidelines for reviews, which are displayed on an organization's listing page and are associated with a bubble rating ranging from 1 to 5. Beyond overall ratings, contributors evaluate additional service attributes, including legroom, seat comfort, in-flight entertainment, onboard experience, customer service, value for money, cleanliness, check-in and boarding processes, as well as food and beverage quality.

The platform has been extensively utilized as a data source by various investigators worldwide, contributing to diverse research findings. In particular, online reviews serve as an efficient tool for examining customer satisfaction. Sudhakar & Gunasekar (2020) analyzed Indian airlines using TripAdvisor feedbacks to reveal varying factors influencing satisfaction across nationalities and airline types, with Indian passengers prioritizing value for money and foreign passengers prioritizing service attributes. Sezgen et al. (2019) examined 45 international airlines to define the key drivers of passenger satisfaction and dissatisfaction across full-service and low-cost carriers, as well as economy and premium cabins. Despite variations in satisfaction drivers, fundamental factors such as friendly staff, value, and service quality remain universally important. In the US airline market, Park et al. (2020) underscores the need for sustainable strategies beyond price competition, highlighting the asymmetric impact of service quality attributes on customer satisfaction and their critical importance in airline operations. According to Korfiatis et al. (2019), integrating unstructured data from online reviews into service quality assessment enhances understanding and predictability of customer satisfaction. Examining

Tripadvisor reviews, the study revealed temporal variations in dimensions, thereby emphasizing the importance of monitoring and adapting to changing customer preferences over time.

Additionally, Tripadvisor reviews have been used to identify key themes by Brochado et al. (2019) and Sulu et al. (2022) during the pandemic. The study by Brochado et al. (2019) was aimed to identify the main themes in traveler reviews and their relationship with value-for-money ratings by examining airlines affiliated with the largest alliances, SkyTeam and Star Alliance, across France, the United States, Indonesia, Canada, Taiwan, and Germany. Whereas, Sulu's (2022) research focused on online airline reviews during the COVID-19 outbreak, aiming to define principal themes and identify disparities among passengers of various nationalities. The analysis revealed that European passengers exhibited the most polarized opinions, ranging from highly critical to highly satisfied, while American and Australian passengers predominantly expressed positive reviews.

2.2.2 OCR studies based on data collected in Skytrax.

Skytrax is the international air transport rating organisation with Air Travel review forum (www.airlinequality.com) that has become a leading review site for airline, airport and associated air travel traveller reviews **SEARCH REFERENCE** (Skytrax, 2024). Reviews submitted to the platform undergo a multi-stage checking process to ensure compliance with the Editorial Policy and requires reviewers to submit feedback that follows a structured evaluation framework. The platform enables passengers to share detailed opinions on their experiences across four categories: airline, seat, airport, and lounge. Passengers can rate airlines on a scale from 1 to 10 and provide textual comment, along with additional details such as country of residence, date of flight, route, cabin class, aircraft type, type of travel, and likelihood of recommending the airline. Moreover, contributors assess supplementary service attributes including value for money, ground service, seat comfort, cabin staff service, food & beverages, in-flight entertainment, and cabin Wi-Fi & connectivity.

Various scholars have used Skytrax reviews to define key factors influencing customer experience and satisfaction within the airline sector. Ban & Kim (2019) identified six key clusters, including seat comfort, staff service, entertainment, ground service, value for money, and airline brand, with all factors except entertainment significantly impacting customer satisfaction and recommendation. Shadiyar et al. (2020) **SEARCH REFERENCE** explored the utilization of big data analytics and online customer reviews to enhance customer satisfaction in airlines, identifying factors such as seat comfort, staff behavior, food and beverage quality, ground service, and particularly value for money as main. The study by Lucini et al. (2020) was aimed to measure customer satisfaction by analyzing Skytrax reviews to provide airlines with guidelines to improve competitiveness. The study revealed that the type of cabin significantly impacts customer satisfaction dimensions, prioritizing customer service for first-class passengers, comfort for premium economy passengers, and efficient baggage handling and minimal waiting times for economy class travellers, whereas the type of passenger had minimal impact on satisfaction dimensions. Dike et al. (2024) analyzed Skytrax feedbacks to explore how customer culture, travel purpose, and seat type shape customer expectations and satisfaction. The study highlighted that customer service quality, delays, and baggage handling significantly impact perceived service performance, with empathy and reliability having the strongest influence on perceived satisfaction, while cultural backgrounds affect levels of customer satisfaction. Chatterjee et al. (2023) conducted a ten-year analysis of Skytrax reviews, identifying core attributes such as value for money, seating comfort, and cabin staff behaviour as more strongly related to customer satisfaction and recommendation than augmented aspects like food and beverages and in-flight entertainment. Authors found that core service attributes such as responsiveness