Are environmental-related online reviews more helpful? A big data analytics approach

Article

Accepted Version

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Are Environmental-related Online Reviews more helpful?
A Big Data Analytics Approach

Purpose–Based on more than 2.7 million online reviews (ORs) collected with big data analytical techniques from Booking.com and TripAdvisor.com, this study explores if and to what extent environmental discourse embedded in ORs has an impact on electronic Word-of-Mouth (e-WOM) helpfulness across 8 major destination cities in North America and Europe.

Design/methodology/approach–This study gathered, by means of Big Data techniques, 2.7 million online reviews (ORs) hosted on Booking and TripAdvisor, and covering hospitality services in 8 different destinations cities in North America (New York City, Miami, Orlando, and Las Vegas) and Europe (Barcelona, London, Paris, and Rome) over the period 2017-2018. The ORs were analysed by means of ad hoc content analytic dictionaries to identify the *presence* and *depth* of the *environmental discourse* included in each OR. A negative binomial regression analysis was used to measure the impact of the *presence/depth* of online environmental discourse in ORs on e-WOM helpfulness.

Findings–The findings indicate that the environmental discourse presence and depth influence positively e-WOM helpfulness. More specifically those travelers who write explicitly about environmental topics in their ORs are more likely to produce ORs that are voted as helpful by other consumers.

Research implications/limitations – Implications highlight that both hotel managers and platform developers/managers should become increasingly aware of the importance that customer attach to environmental practices and initiatives and therefore engage more assiduously in environmental initiatives, if their objective is to improve online review helpfulness for other customers reading the focal reviews. Future studies might include more destinations and other operationalizations of environmental discourse.

Originality/value – This study constitutes the first attempt to capture how the *presence* and *depth* of hospitality services consumers' *environmental discourse* influence e-WOM helpfulness on multiple digital platforms, by means of a big data analysis on a large sample of online reviews across multiple countries and destinations. As such it makes a relevant contribution to the area at the intersection between big data analytics, e-WOM, and sustainable tourism research.

Keywords: Big Data analytics; e-WOM helpfulness; environmental discourse; multi-platform study; Booking; TripAdvisor.

1. Introduction

Hospitality service providers and customers are increasingly aware of the damage that excessive natural resources depletion in the guise of energy, water, and food consumption, and CO_2 emissions might bring about to the planet (Gössling and Peeters, 2015). Accordingly, a number of stakeholders including firms, consumers and policy makers in the hospitality sector are becoming increasingly aware of environmental issues (Di Pietro et al., 2013; Ettinger et al., 2018) and taking action to embrace sustainable processes (Bonilla Priego et al., 2011; Raub and Martin-Rios, 2019), offer green products (Kemper et al., 2019), and improve their Corporate Social Responsibility (CSR) reporting (Guix et al., 2018). For instance, Marriott launched and developed the "Make a Green Choice" initiative allowing guests to give up housekeeping (equating to inferior consumption of cleaning products), bartering it for loyalty points or F&B vouchers (Marriott, 2018).

From a consumer perspective, an increasing number of consumers are environmentally aware and concerned (Cho, 2015; Dolnicar et al., 2008; Liu et al., 2018) and environmental concerns have been found to influence consumers' purchase intentions of hospitality services (Di Pietro et al., 2013; Dutta et al., 2008; Gustin and Weaver, 1996; Kyung et al., 2012). Recently online reviews (ORs) have been used to capture hotel guests' perceptions of service providers' environmental practices (Brazytė et al., 2017; D'Acunto et al., 2020; Ettinger et al., 2018; Lee et al., 2016; Peiró-Signes et al., 2014; Yu et al., 2017) as part of a wider research stream related to electronic Word-of-Mouth (e-WOM). However, extant studies revolving around the relationship between environmental discourse and ORs and e-WOM features are rather fragmented and limited. Indeed, they have only focused on the relationship between some CSR dimensions and online review ratings, mostly by using either small samples of ORs related to a specific country (e.g., Yu et al., 2017) or a limited set of cities within the same continent (e.g., D'Acunto et al., 2020) and from a single platform (in most of the cases TripAdvisor). Moreover, those studies have construed implicitly online consumers' CSR discourse (e.g., D'Acunto et al., 2020) but have not critically deployed online consumers' environmental discourse – in the guise of large volumes of online reviews mentioning environmental aspects - to understand more about electronic Word-of-Mouth (e-WOM) helpfulness. This is rather surprising as consumers are increasingly environmentally aware (Cho, 2015; Dolnicar et al., 2008; Liu et al., 2018) and, since environmental features are not part of the standard hotel attributes assessed by online reviewers, one would expect that online reviewers (and online review) mentioning explicitly environmental concerns might be

particularly diagnostic and therefore helpful because they offer depth about aspects that go beyond the standard hotel service attributes.

So far, no research paper has explored if and to what extent consumers' *online environmental discourse* – namely the presence and depth of environmental-related content in online reviews proxying consumers' environmental awareness and the extent to which consumers dig in depth about environmental aspects – can affect the degree of e-WOM helpfulness and generally online review helpfulness literature has focused on a number of antecedents at the reviewer, review and hotel level that have not taken into account consumers' sensitivity to the environment. To bridge this gap, we collected more than 2.7 million online reviews (ORs) by means of sophisticated big data analytical techniques from Booking.com and TripAdvisor.com, and explored if and to what extent environmental discourse embedded in ORs has an impact on e-WOM helpfulness across 8 major destination cities in North America and Europe. Accordingly, this work makes a relevant contribution to the area at the intersection between big data analytics, e-WOM, and sustainable tourism research.

Accordingly, we make multiple contributions. First, we contribute to and extend the research stream related to e-WOM helpfulness by addressing if and how online consumers' *environmental discourse* influences e-WOM helpfulness across multiple digital platforms. Second, we propose that in the digital age consumers' environmental concerns can be captured by means of a big data analytics approach, entailing digital data streams of online consumer reviews and text analytics. Third, we contribute to the big data and analytics research stream in hospitality management literature, by suggesting that extrapolating analytics from different digital platforms (namely community-based vs. transaction-based digital platforms) while providing statistically different quantifications of environmental discourse *presence* and *depth*, offers overall consistent results when it comes to identifying the impact of environmental discourse on e-WOM helpfulness. Fourth, we enrich the body of e-WOM literature in the sustainable tourism field that so far has mainly focused on broad CSR discourses or green practices at the local or at best national level by adopting a research design that allows to seize the phenomenon across different countries and continents, thus validating the generalizability of the findings. Fifth, we contribute to the consumer behavior literature revolving around consumers' behaviors and evaluations of environmental and green aspects: rather than deploying small sample surveys asking respondents stated attitudes and behaviors (i.e., perceptions), we look at the real evaluations after consumption by means of big data analytics from a large sample of more than 2.7 million consumers' online reviews. Last, we innovatively build a joined-up body of knowledge matching the identification and examination of the drivers of e-WOM helpfulness

(e.g., Ghose and Ipeirotis, 2012), the emerging debate on consumers' perception of environmental initiatives in the tourism and hospitality industry (Ettinger et al., 2018) and the research stream at the intersection between online consumers' perceptions of environmental issues through e-WOM (e.g., Yu et al., 2017), by shedding new light on the relationship between e-WOM helpfulness and environmental discourse. From a practical point of view, by analyzing if and to what extent online consumer environmental discourse can enhance e-WOM helpfulness, we help hotel managers and platform managers shed light on one of the factors that might be used by consumers in their decision making, potentially affecting also their purchase decisions.

To make the aforementioned contributions, the paper is structured as follows. Section 2 reviews the relevant hospitality and tourism literature in the fields of big data and analytics, electronic word-of-mouth and online reviews, and sustainable tourism and puts forward the relevant research question. Section 3 illustrates the methodology adopted. In the fourth section we report and discuss the findings. Section 5 draws the main conclusions, elucidates the theoretical contributions and managerial contributions, and identifies the limitations of the study and future research directions.

2. Literature Review

2.1 Big data analytics in Hospitality and Tourism

As one of the digital technologies driving the ongoing digital revolution (Rüßmann et al., 2015), Big data (BD) are playing an increasing role in social sciences. As noted in a recent literature review (Mariani et al., 2018), the notion of BD emerged in the computer science literature almost three decades ago when scientific visualization of large amounts of data was needed (Cox and Ellsworth, 1997). Yet, the most popular definition of BD applied within the business and management domain was formulated back in 2001 by Gartner analyst Doug Laney who detected three main features of BD management, describing them as the "3Vs", encompassing: Volume (i.e., the dimension of data, today in the order of Zettabytes), Velocity (i.e., the speed of data generation, alteration and transmission), and Variety (i.e., data are generated in different guises and formats/structures) (Laney, 2001). Later, Laney's model was extended to reflect and accommodate additional characteristics such as Value (i.e., the value of the knowledge that can be extracted from data in the guise of BD analytics) and Veracity (i.e., the reliability of data), thus generating a 5Vs Big Data management framework (Fosso Wamba et al., 2015).

However, BD per se would not be sufficiently relevant for business and management scholars if they were not capable to generate BD analytics (BDA). BDA can be defined as a process to discover, collect, process, clean, analyze, use, report and interpret BD with the aim of generating business insights and enhancing business intelligence that can lead to the creation of business value (Davenport, 2014, 2017), that can ultimately translate into superior performance and competitive advantage by organizations mastering BDA capabilities (Fosso Wamba et al., 2017, 2020; Mariani and Fosso Wamba, 2020). BDA have been increasingly embraced by a number of companies and researchers seeking to generate meaningful insights in a large number of verticals in manufacturing and service industries (Davenport, 2017). The underlying techniques to generate BDA are part of data science techniques (Witten, 2016) and they include data mining techniques, including also text mining techniques allowing to capture quantitatively the unstructured part of ORs (Zhang et al., 2016), and allowing for a richer interpretation of the content of the consumption experience reviewed (Geetha et al. 2017; Xu and Li, 2016; Zhao et al., 2019). Within the service industries, hospitality is undoubtedly one of the leading sectors deploying BDA (Li et al., 2018; Mariani et al., 2018). BDA are generated from a wide variety of sources including devices in the guise of *device data* (e.g., mobile roaming data, Wifi data, GPS data, etc.), operations under the guise of *transaction data* (e.g., online booking data, web search data, etc.) and user data in the form of *user generated content (UGC) data* (e.g., online picture and text data) (Li et al., 2018). Perhaps the most popular form of data is UGC data (Li et al., 2018) that entails also online reviews (ORs) that have been extensively used to gain a better understanding of customer experience, engagement and satisfaction with hospitality services (e.g., Guo et al., 2017; Mariani and Borghi, 2018; Mariani and Predvoditeleva, 2019; Xiang et al., 2015) in online settings.

2.2 Electronic Word-of-Mouth, Online Reviews and e-WOM Helpfulness in Hospitality and Tourism

The advancement of digital technologies and the development and consolidation of digital platforms - including social media – have paved the way for a proliferation of online reviews (ORs) (Mariani et al., 2014; Mariani and Matarazzo, 2020). The latter ones constitute a means by which former, current and future online consumers express their perceptions and elaborate evaluations about products, services, and brands on the Internet, de facto sharing them with other Internet users (Hennig-Thurau et al., 2004). ORs have been deeply investigated in the information management (e.g., Duan et al., 2008), computer science (e.g., Hu et al., 2008) and marketing (e.g., Rosario et al., 2016) domains. In most of the aforesaid domains, ORs have

become the object of a research stream labelled first electronic word-of-mouth (e-WOM) and later online word-of-mouth (King et al., 2014). This research stream has expanded considerably over the last 15 years as e-WOM is becoming increasingly relevant vis-à-vis WOM due to its likely anonymity, rapidity of diffusion, convenience, non-existence of face-to-face contact, many-to-many reach, and communication effectiveness (Sun et al., 2006). Consequently, the antecedents and outcomes of e-WOM have been at the centre of the stage in scholarly debate and therefore examined in a number of studies as suggested for instance by recent literature reviews and meta-analytical studies in marketing (e.g., Rosario et al., 2016; You et al., 2015). In line with the vast majority of the aforementioned literature, in this paper we will use the term ORs and eWOM interchangeably.

In the hospitality industry, e-WOM has progressively become ubiquitous due to the increasing amount of ORs written and read by online consumers on Online Travel Agencies such as Booking.com and Expedia.com, online travel review websites such as TripAdvisor and CTrip, sharing economy platforms like Airbnb, and generalist social media such as Facebook (Mariani et al., 2021). Scholars in hospitality and tourism have found that ORs can elicit consumers' purchasing and booking intentions (Tsao et al., 2018) and ultimately influence sales (Ye et al., 2009) and firms' financial performance (Yang et al., 2018; Mariani and Borghi, 2020; Mariani and Visani, 2019). E-WOM from ORs is adopted as information source before purchase by as a relevant number of prospective customers of accommodation services as the quality of the focal services is unknown before consumption and problematic to assess before purchase (Filieri and McLeay, 2014; Gretzel and Yoo, 2008). For example, TripAdvisor ORs are frequently adopted by consumers to inform their purchase decisions and some of the features of ORs (such as ratings) have a positive impact on online bookings (Ye et al., 2011). Therefore, not only large volumes (i.e., Big Data) of ORs represent a critical source of business intelligence for hospitality firms (Mariani et al., 2018), but they also constitute a meaningful source to generate critical insights about consumers' perceptions, including perceptions of the usefulness of online peers' evaluations of hospitality services (discussed in section 2.2.1) and environmental perceptions (discussed in section 2.3).

2.2.1 E-WOM Helpfulness in Hospitality and Tourism

The consolidation of travel OR platforms like TripAdvisor and Online Travel Agencies like Booking.com and Expedia.com has brough about an increasing number of ORs related to hospitality services. While the abundance of ORs means more information about hospitality services, overabundance is making it difficult for consumers to process this information. The

proliferation of ORs has generated information overload for consumers (Eppler and Mengis, 2004; Furner and Zinko, 2017; Luo et al., 2013; Park et al., 2006; Park and Lee, 2008) who are challenged to find out which ORs can help them understand the quality of the service covered by the OR. To address and alleviate this problem, OR platforms in hospitality and tourism have endowed themselves with procedures letting customers like or vote the helpfulness of an OR (Singh et al., 2017), and thus reveal to other customers which ORs are the most important to gauge hospitality services' performance and quality.

OR helpfulness therefore relates to the extent to which online consumers evaluate other consumers' ORs as informative and helpful. One of the most relevant empirical found that there are reviewer- and review-level features including readability, linguistic correctness, subjectivity, and informativeness, which affect sales and perceived usefulness (Ghose and Ipeirotis, 2011). In their study, helpfulness equates to the degree of informativeness of ORs that works as a signal for other users. Thus, helpfulness of ORs in hospitality literature has been operationalized in terms of number of helpful votes (Felbermayr and Nanopoulos, 2016; Ghose and Ipeirotis, 2011) allowing online consumers to evaluate the information diagnosticity of other travellers' online reviews (e.g., Filieri, 2015; Filieri and McLeay, 2014; Kwok and Xie, 2016) due to their information cues.

In the current study, e-WOM helpfulness constitutes the capability of an online review to aid other consumers to gain a more accurate understanding of the quality and performance of a product or a service and can be operationalized as the overall number of helpful votes assigned by online readers to a given OR, indicating to what extent an OR has been found useful in the OR digital platform. A relatively large amount of studies has analyzed the antecedents of e-WOM helpfulness in both the information management (Ghose and Ipeirotis, 2006, 2007, 2011; Ghose et al., 2012; Hu et al., 2010; Hu and Chen, 2016) and marketing management (Korfiatis et al., 2012; Lee and Shin, 2014; Shin et al., 2019) field, as well as the hospitality and tourism field (e.g., Fang et al., 2016; Filieri, 2015; Liu and Park, 2015) trying to understand what makes an OR helpful.

In hospitality and tourism management studies, the most commonly examined antecedents and predictors of e-WOM helpfulness entails: 1) reviewers' features such as the reviewers' identity disclosure, experience, expertise, and reputation (e.g., Liu and Park, 2015; Fang et al., 2016; Filieri et al., 2019; Park and Nicolau, 2015); 2) reviews' features such as length, readability, ratings, extreme ratings, emotions, emotional intensity, number of days elapsed from the day the review was published (e.g., Fang et al., 2016; Kwok and Xie, 2016; Filieri et al., 2018, 2019; Lee et al., 2017; Park and Nicolau, 2015; Wang et al., 2019); 3) hotel

firms' features entailing firm size, managerial response and belonging to a chain or not (Kwok and Xie, 2016; Filieri et al., 2018, 2020).

From the literature review conducted, it is rather clear that no research has explored to what extent the presence of environmental-related content in online reviews – that we can define as consumers' *online environmental discourse* - influence the e-WOM helpfulness. This represents a critical research gap, and it is rather surprising as online consumers of hotel services are increasingly environmentally concerned and growing in number (D'Acunto et al., 2020). Moreover, one would expect that online reviewers (and online review) mentioning explicitly environmental concerns might be particularly diagnostic and therefore helpful. However, so far, no research work has analyzed if and to what extent consumers' online environmental discourse can affect the e-WOM helpfulness. This is the research gap that we intend to fill. In the following section we elaborate on the construct of environmental discourse and formulate our research questions.

2.3 Environmental discourse and e-WOM helpfulness
2.3.1 Online environmental discourse of hotel guests and e-WOM helpfulness

Hospitality and tourism firms have been found to consume a considerable amount of natural resources (Gössling and Peeters, 2015) and are, in some cases, becoming increasingly aware of environmental issues (Di Pietro et al., 2013; Ettinger et al., 2018) and taking action to embrace sustainable processes (Bonilla Priego et al., 2011; Raub and Martin-Rios, 2019), offer, promote and market green products (Dief and Font, 2010; Kemper et al., 2019), and improve their Corporate Social Responsibility (CSR) reporting (Guix et al., 2018), and competitiveness (Helen and Giebelhausen, 2012).

From a consumer perspective, an increasing number of consumers (including travelers and hotel guests) are environmentally aware (Liu et al., 2018; Pagiaslis and Krontalis, 2014) and express environmental concerns have been found to influence consumers' purchase intentions of hospitality services (Di Pietro et al., 2013; Gustin and Weaver, 1996; Kyung et al., 2012). In their analysis of U.S. hotel customers, Gustin and Weaver (1996) found a positive relationship between customers' attitudes towards U.S. hotels' pro-environmental practices and purchase intentions. Kyung et al. (2012) conduct a survey on 455 consumers in three US states and find that the level of environmental concern influences positively the willingness to pay a premium for hotels' green initiatives. Di Pietro et al. (2013) find that, despite customers

wanted to gain more knowledge about green practices, only a tiny fraction is willing to pay for environmentally friendly products and services.

Recently online consumer reviews have been used to capture hotel guests' perceptions of service providers' environmental practices (Brazytė et al., 2017; D'Acunto et al., 2020; Ettinger et al., 2018; Lee et al., 2016; Peiró-Signes et al., 2014; Yu et al., 2017). For instance, Peiró-Signes et al. (2014) find that four-star hotels endowed with an ISO 14001 environmental certification show higher online review ratings compared to those without such certification. Examining 727 green ORs related to the top ten green hotels on TripAdvisor, Yu et al. (2017) observe that consumers describe both positive and negative experiences at green hotels, and that various types of green practices affect in a different way online customer satisfaction. Brazytė et al. (2017) scrutinize a sample of 2,487 ORs of 30 Costa-Rican hotels endowed with a sustainability certification and reveal that, where sustainability indicators are explicitly used OR ratings are higher than where they are not mentioned. By employing qualitative content analysis on a sample of 1,383 ORs pertaining to 47 Austrian hotels and mentioning CSR aspects, Ettinger et al. (2018) document that a large majority of the review (more than 90%) was of a positive or neutral nature. In a study of CSR and ORs, D'Acunto et al. (2020) analyse how CSR dimensions relate to hotels' attributes such as experience, amenities, location, transactions, value, over 10 years for six European cities. They also use TripAdvisor OR ratings and CSR and find that there is a positive correlation between CSR content in online reviews and online ratings.

To summarize, extant studies revolving around the relationship between environmental discourse and OR features are rather fragmented and limited. Indeed, they have only focused on the relationship between CSR dimensions - and the underlying CSR discourse – and online review ratings, mostly by using either small samples of ORs related to a specific country (e.g., Yu et al., 2017) or a limited set of cities within the same continent (e.g., D'Acunto et al., 2020) and from a single platform (in most of the cases TripAdvisor). Moreover, those studies construe implicitly online consumers' CSR discourse (e.g., D'Acunto et al., 2020) but do not critically deploy online consumers' environmental discourse to understand more about e-WOM helpfulness. This is rather surprising as consumers are increasingly environmentally aware (D'Acunto et al., 2020) and since environmental features are not part of the standard attributes assessed by online reviewers, one would expect that online reviews mentioning explicitly environmental concerns might be particularly diagnostic and therefore helpful because they offer depth about aspects that go beyond the standard hotel service attributes. In other words, while commenting on aspects that are beyond standard hotel service attributes, the consumers

are going an extra mile in their reviewing activity, and this implies that they care about their review. Caring and "going an extra mile" in reviewing activity is consistent with early theories of altruism (Carman, 1992; Nagel, 1970; Paul et al., 1993) and with marketing literature and theory that has revealed that consumers behave in an altruistic way when they generate WOM and e-WOM because they want to help other consumers thus engaging in pro-social behavior (Hennig-Thurau et al., 2004; Sundaram et al., 1998). So far, no research paper has investigated if and to what extent consumers' *online environmental discourse* – which is perhaps the highest expression of altruism and pro-social behavior in online reviewing - can affect the degree of helpfulness of online reviews. This represents a critical research gap, and it is rather surprising as online consumers of hotel services are increasingly environmentally concerned and growing in number (D'Acunto et al., 2020). Moreover, one would expect that online reviewers (and online review) mentioning explicitly environmental concerns might be particularly diagnostic and therefore helpful. However, so far, no study has examined if and to what extent consumers' online environmental discourse can affect the e-WOM helpfulness. Indeed, as illustrated in section 2.2.1, in hospitality management studies, scholars have only focused on a number of antecedents of e-WOM helpfulness such as the reviewers' identity disclosure, experience, expertise, and reputation (e.g., Liu and Park, 2015; Fang et al., 2016; Filieri et al., 2019; Park and Nicolau, 2015), reviews' features such as length, readability, ratings, extreme ratings, emotions, emotional intensity, number of days elapsed from the day the review was published (e.g., Fang et al., 2016; Kwok and Xie, 2016; Filieri et al., 2018, 2019; Lee et al., 2017; Park and Nicolau, 2015; Wang et al., 2019), and hotel firms' features entailing firm size, managerial response and belonging to a chain or not (Kwok and Xie, 2016; Filieri et al., 2018, 2020). However, no study has conceptualized or measured empirically if and how *online environmental discourse* affects e-WOM helpfulness. This is the research gap that we intend to fill. To bridge this gap, (thus offering insights to researchers working in the area at the intersection of BDA, e-WOM, and sustainable tourism research), we analyze how online consumers' *environmental discourse* about hotels influences e-WOM helpfulness, building on a large sample of more than 2.7 million ORs pertaining to hotels across 8 leading tourism destinations in America and Europe. To enhance the generalizability of the results, we take into account different types of OR platforms including transactional and community-based OR platforms that have been found to function in a different way (Gligorijevic, 2016). We conceptualize online consumers' environmental discourse as a bidimensional concept encompassing the *presence* and *depth* dimensions. The *presence* of consumers' environmental discourse relates to consumers' environmental awareness and can be captured by the presence

of words related to environmental aspects in hotels' ORs. The *depth* of consumers' environmental discourse instead relates to the extent to which consumers dig in depth about environmental aspects and concerns and can be proxied by the intensity and number of words in a hotel OR related to environmental aspects. That clarified, and based on the research gap identified, we set to address the following research question:

RQ: Do the *presence* and *depth* of environmental-related e-WOM in consumers' online reviews of hospitality services influence e-WOM helpfulness and how?

3. Methodology

3.1 Data and sample

For the purposes of this project, we collected and analyzed online review data in January 2019 using a script in the Python programming language. In particular, we deployed two leading OR websites that cover the two different types of OR platforms identified in the current literature (Gligorijevic, 2016): community-based and transaction-based platforms. More specifically, for the community-based platform type, we collected data from TripAdvisor which is considered the largest aggregator worldwide to this respect, whereas for the transaction-based platform we chose Booking.com since it integrates the largest portion of certified ORs (Revinate, 2019).

Regarding the sample, we collected OR data for the hotels listed in the aforementioned OR platforms. The hotels are located in the top destinations in terms of tourist arrivals according to the Euromonitor report 2018 (Geerts, 2018). In more detail, we included in our sample four of the top 10 tourism destinations in the Americas, namely New York City, Miami, Orlando, and Las Vegas, and four of the top 10 tourism destinations in Europe, such as London, Paris, Rome, and Barcelona. Moreover, to ensure robustness and comparability to our analyses, we collected the entire population of ORs from both platforms for the same time frame: the years 2017 and 2018. This is due to the fact that Booking.com makes available OR data on a rolling basis of two years and at the moment of the data collection task the latest available years were 2017 and 2018. From the aforementioned digital platforms, we collected not only the single ORs and their content but also reviewer-related information, such as reviewers' level of expertise and identity disclosure as well as hotel-related metrics: for instance, whether or not the hotel belongs to a chain, star rating, and overall scores in terms of reputation and popularity. Besides, the number of helpful votes received by each specific OR was collected and used in the econometric models to identify the helpfulness of a given OR.

In line with other studies concerning the utilization of text mining techniques (e.g., Bi et al., 2019), we kept in our final database only ORs written in English. As such, the final sample of ORs retrieved for TripAdvisor consists of 1,144,461 ORs, while for Booking.com it consists of 1,557,766 ORs: overall, 2,702,227 ORs were collected and analyzed. Tables 1.a and 1.b show the sample of ORs analyzed in the study by destination and platform:

Table 1.a –TripAdvisor Sample of ORs

Tourism destination	N of reviews (English)	Ratio N of reviews (English)/Total N of reviews for the destination	% on the total sample of reviews in English	N Hotels
Barcelona	68,279	52.25%	6.0%	449
Las Vegas	214,205	91.30%	18.7%	210
London	303,059	82.16%	26.5%	997
Miami	25,443	75.32%	2.2%	108
New York	213,808	80.26%	18.7%	445
Orlando	134,826	86.53%	11.8%	321
Paris	115,624	46.24%	10.1%	1,679
Rome	69,217	48.16%	6.0%	991
Total	**1,144,461**	**72.25%**	**100.0%**	**5,200**

Table 1.b –Booking.com Sample of ORs

Tourism destination	N of reviews (English)	Ratio N of reviews (English)/Total N of reviews for the destination	% on the total sample of reviews in English	N Hotels
Barcelona	136,767	34.11%	8.8%	450
Las Vegas	51,755	64.65%	3.3%	154
London	595,146	63.03%	38.2%	1,027
Miami	46,515	45.34%	3.0%	154
New York	247,471	56.07%	15.9%	418
Orlando	92,804	61.98%	6.0%	246
Paris	265,206	31.23%	17.0%	1,662
Rome	122,102	30.03%	7.8%	933
Total	**1,557,766**	**46.16%**	**100.0%**	**5,044**

3.2 Methods

To investigate our research question, we needed to select an estimation technique able to accommodate two distinctive traits related to the nature of our dependent variable (Review Helpfulness): non-negative and count measure. To this respect, the simplest and most established method to account for these characteristics is the Poisson (log-linear) regression model. Yet, it relies on an assumption that is usually found to be too restrictive in real-life scenarios (Cameron and Trivedi, 1998). Indeed, the Poisson distribution requires that the conditional mean and variance of the event under investigation are equal (Wooldridge, 2010). In our data, we are in presence of the so-called "*overdispersion*" (Cameron and Trivedi, 1998) since the variance (0.313 for Booking and 1.283 for TripAdvisor) exceeds the mean (0.07 for Booking and 0.591 for TripAdvisor) for the dependent variable. This is confirmed looking at the overall distribution of the count of helpful votes (Figure 1.a and 1.b) where the majority of ORs have not received any helpful vote in both platforms. In the OR setting this is not an unusual situation (e.g., Fang et al., 2016; Lee et al., 2017; Park and Nicolau, 2015; Yin et al., 2014). Accordingly, we reject the *equidispersion* property of the Poisson distribution, and approach the estimation problem leveraging on a less restrictive econometric model which allows for more flexibility in the variance-mean relationship. Thus, consistently with other researchers using the number of helpful votes as their dependent variable and incurring in the same overdispersion problem (Fang et al., 2016; Lee et al., 2017; Park and Nicolau, 2015; Wang et al., 2019) we deploy a Negative Binomial regression model, whose details are presented in Cameron and Trivedi (1986). In essence, the formulation of the Negative Binomial model, for a random count variable *y*, is governed by the following density function, obtained as a mixture of a Poisson and a gamma distribution:

$$f(y|\mu,\alpha) = \frac{\Gamma(\alpha^{-1} + y)}{\Gamma(\alpha^{-1})\Gamma(y + 1)} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu}\right)^{\alpha^{-1}} \left(\frac{\mu}{\mu + \alpha^{-1}}\right)^{y} , \alpha > 0$$

where Γ indicates the gamma function and μ is a deterministic function of the regressors **x**. Besides, α refers to the unknown parameter of the mixing gamma distribution, such that:

$$E[y|u,\alpha] = \mu, \quad V[y|u,\alpha] = \mu(1 + \alpha\mu)$$

At the single observation level (e.g., individual review) μ is defined as:

$$\mu_i = \exp(x_i'\boldsymbol{\beta})$$

where x_i' is the vector of characteristics related to the online review i used as regressors and β is the vector of parameters to be estimated in order to assess the effect of x_i' on the random variable under investigation, review helpfulness in this case.
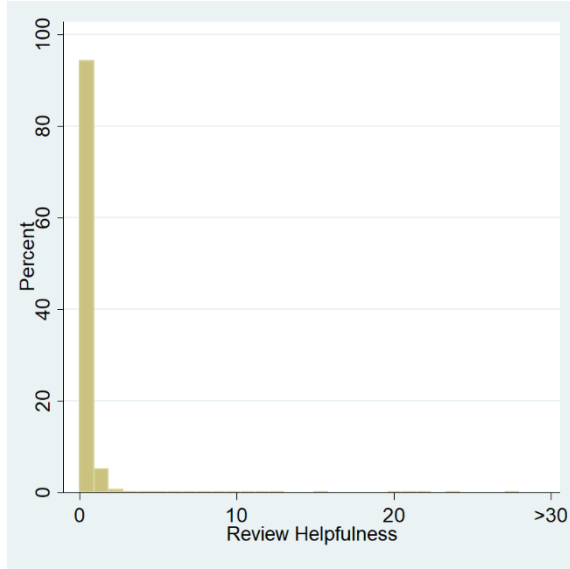

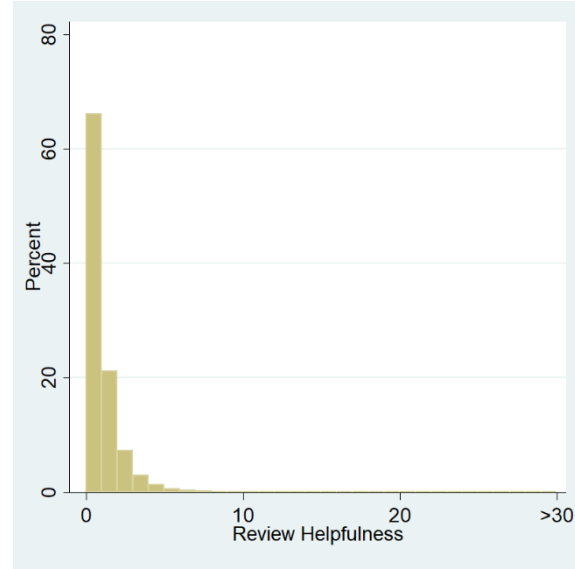**Figure 1.a** Review Helpfulness Distribution Booking.com

Figure 1.b Review Helpfulness Distribution TripAdvisor

3.3 Variables

In order to capture different nuances of the online environmental discourse in our analyses, we created two constructs and variables, namely *Environmental Presence* and *Environmental Depth*. To this aim, we leveraged on the environmental dictionary developed by Pencle and Mălăescu (2016) which includes a set of 451 environmental-related keywords. The dictionary, conceived for predicting the impact of corporate social responsibility practices in US IPOs' pricing components, has been developed through *computer-aided text analysis* (Pencle and Mălăescu, 2016). Recently, it has been successfully deployed in the hospitality and tourism setting (D'Acunto et al., 2020). *Environmental Presence* indicates the presence of environment-related aspects in an online review. It is measured as a dummy variable equal to one when the written content of an OR refers to at least a word in the environmental dictionary used in the analysis. *Environmental Depth* aims to capture the depth of environmental discourse in a given OR. To this end, it is operationalized as the ratio between the number of environment-related words and the overall amount of words in an OR.

Referring to our econometric analyses, *Review Helpfulness* has been used as a dependent variable and it has been operationalized as the overall number of helpful votes

assigned by online readers to a given OR, indicating to what extent an OR has been found useful in the OR digital platform. Moreover, in our empirical analyses, we embedded a wide range of control variables based on extant literature revolving around the antecedents of e-WOM helpfulness. To this aim, we include review-related metrics, such as the score provided by the online reviewer to judge her overall service experience (*Review Rating*). Regarding the latter, as empirically proved by Park and Nicolau (2015) and Yin et al. (2014) there exists a U-shape relationship between Review Helpfulness and the Review Rating. Thus, in our econometric model, we introduced two terms related to the review rating: a linear one (*Review Rating*) and its squared value (*Squared Review Rating*). Moreover, at the reviewer level, we control for her level of expertise (*Reviewer Expertise*) and to what extent she disclosed her identity in the digital platform (Forman et al., 2008). To this aim, we leveraged on whether or not the reviewer has provided a personalized picture (*Reviewer Personalized Image*) and if she has disclosed her country of origin (*Country Disclosure*).

Besides, we analyzed the written text provided with the OR, and added a set of text analytics measures to ensure robustness to our econometric model specification. Indeed, we control for the number of words in the OR (*Review Length*), the understandability of the text (*Review Readability*), and its polarity (*Review Polarity*). In particular, *Review Polarity,* also known as sentiment score (Ma et al., 2018), has been operationalized using a continuous variable ranging from -1 to +1 and is computed – in line with Alaei et al. (2019) – using the Valence aware dictionary for sentiment reasoning (VADER), which exploits a set of heuristics along with a specific lexicon dictionary for this particular task (Hutto and Gilbert, 2014). The aforementioned VADER-based sentiment analysis technique and related measure has been deployed because recent research (Alaei et al., 2019) has found that it outperforms other sentiment analysis classifiers used in the tourism and hospitality domain, while being consistent with them. *Review Readability* refers to the simplicity of a text for a reader's understanding. Consistently with extant literature (Korfiatis et al., 2012; Filieri et al., 2019) we operationalized readability by means of the Automated Readability Index (ARI) that has a long tradition in text mining analysis as it was developed in the sixties to estimate the number of years of formal education a person needs in the US education system to understand an (English) text on the first reading (Smith and Senter 1967). Furthermore, at the hotel level, we include the number of ORs a hotel has received over time (*Hotel Popularity*), its average rating in the OR platform (*Hotel Reputation*), its star category (*Hotel Star Rating*), if it belongs to a chain (*Chain*) and the continent where the hotel is located (*Hotel Continent*). Yet, since higher e-WOM helpfulness can be associated with ORs that have been in the platform for a longer period of

time (Wang et al., 2019), we controlled for the number of days elapsed between the data collection and the publication of the OR (Elapsed Days).

Table 2 displays the description of each of the aforementioned variables used in the econometric models and provides the references to extant studies that have used the same metric as an antecedent of e-WOM helpfulness. Besides, the descriptive statistics can be found in Tables 3.a and 3.b. Since some variables presented a skewed distribution, such as Review Length, Review Readability, and Hotel Popularity we included their logarithmic form in the econometric model.

Table 2 – Description of variables embedded in the Econometric Models

Variable	Description	Reference
Dependent Variable		
Review Helpfulness	Number of helpful votes provided to an online review	Fang et al. (2016), Lee et al. (2017), Park and Nicolau (2015)
Independent Variables		
Environmental Presence	It is a dummy variable that is equal to 1 if the review includes at least one word in the environmental dictionary developed by Pencle and Mălăescu (2016), and zero otherwise.	Own elaboration based on Mariani and Borghi (2020)
Environmental Depth	Share of environment-related words (words present in the Pencle and Mălăescu (2016) environmental dictionary) out of the total number of words in the review, multiplied by 100.	Own elaboration based on Mariani and Borghi (2020)
Control Variables		
Review Rating	Rating provided by an online reviewer to evaluate the overall service experience.	Park and Nicolau (2015), Yin et al. (2014)
Reviewer Expertise	Number of online reviews written by the online reviewer in the analysed digital platform.	Lee et al. (2017), Ghose and Ipeirotis (2011)
Reviewer Personalized Image	It is a dummy variable that is equal to 1 if the reviewer used a personalized image for its social profile in the platform, and zero otherwise.	Park and Nicolau (2015)
Country Disclosure	It is a dummy variable that is equal to 1 if the reviewer did not disclose her/his country of origin, and zero otherwise.	Filieri et al. (2019)
Review Length	It denotes the number of words included in each online review.	Fang et al. (2016), Korfiatis et al. (2012)
Review Readability	It denotes to what extent the online review is "easy to read" through the following formula: $$ARI = 4.71 \times \left(\frac{Characters}{Words}\right) + 0.5 \times \left(\frac{Characters}{Words}\right) - 21.43$$	Ghose and Ipeirotis (2011), Korfiatis et al. (2012), Park and Nicolau (2015)

Review Polarity	It ranges from -1 (extremely negative) to +1 (extremely positive) and expresses with a content and emotions. Calculated using the Valence Aware Dictionary for sentiment Reasoning (VADER) (Hutto and Gilbert, 2014).	Lee et al. (2017)	
Elapsed Days	Number of days elapsed between the data collection and the publication of the online review	Fang et al. (2016), Wang et al. (2019)	
Hotel Popularity	It represents the overall number of reviews received by the hotel in the digital platform at the moment of the data collection	Fang et al. (2016), Lee et al. (2017),	
Hotel Reputation	It denotes the average rating associated with the hotel in the digital platform at the moment of the data collection	Fang et al. (2016), Yin et al. (2014)	
Chain	It is equal to 1 if the hotel belong to a chain, 0 otherwise	Filieri et al. (2020), Gazzoli et al. (2008)	
Star Rating	It is a categorical variable that describes the hotel class category adopted to classify hotels according to their quality (from 1- to 5-stars)	Filieir et al. (2020), Silva (2015)	
Hotel Continent	It is a categorical variable that indicates the continent of the hotel. Accordingly, it embeds 2 values: America and Europe	Own elaboration	

Table 3.a - Descriptive statistics for the TripAdvisor sample

	Mean/Share	SD	Min	Max
Review Helpfulness	0.591	1.283	0.000	351.000
Log(Environmental Depth)	0.522	0.560	0.000	3.013
Environmental Presence	53.8%		0.000	1.000
Review Rating	4.120	1.164	1.000	5.000
Log (Reviewer Expertise)	2.204	1.747	0.000	9.283
Reviewer Personalized Image	27.8%		0.000	1.000
Country Disclosure	76.6%		0.000	1.000
Log (Review Length)	4.411	0.684	0.000	8.190
Log (Review Readability)	2.166	0.600	0.000	6.846
Review Polarity	0.742	0.484	-0.999	1.000
Elapsed Days	380.721	211.408	1.000	730.000
Log (Hotel Popularity)	7.927	1.268	0.693	10.527
Hotel Reputation	4.120	0.511	1.000	5.000
Chain	51.4%		0.000	1.000
Observations	1,144,461			

Table 3.b - Descriptive statistics for the Booking.com sample

	Mean/Share	SD	Min	Max
Review Helpfulness	0.070	0.313	0.000	28.000
Log(Environmental Depth)	0.352	0.687	0.000	4.615
Environmental Presence	24.8%		0.000	1.000
Review Rating	7.950	1.919	2.500	10.000
Log (Reviewer Expertise)	1.330	1.136	0.000	6.954
Reviewer Personalized Image	41.1%		0.000	1.000
Country Disclosure	99.9%		0.000	1.000
Log (Review Length)	2.956	1.157	0.000	6.649

Log (Review Readability)	1.977	0.717	0.000	7.128
Review Polarity	0.392	0.516	-0.998	0.999
Elapsed Days	351.540	206.561	1.000	730.000
Log (Hotel Popularity)	7.560	0.960	0.000	10.594
Hotel Reputation	7.950	0.919	2.500	10.000
Chain	48.6%		0.000	1.000
Observations	1,557,766			

4. Findings

Referring to the econometric analyses, Table 4 shows the empirical results of the impact of environmental discourse-related metrics (Environmental *presence* and *depth*) on the perceived helpfulness of an OR across the two OR platform analyzed (Booking.com and TripAdvisor). Assessing the appropriateness of the negative binomial regressions, the likelihood-ratio test for alpha equal to 0 is statistically significant (p<0.001) in all cases. Therefore, this implies the existence of heterogeneity in preferences at the consumer level, which support the deployment of the chosen estimation technique over the standard Poisson regression model (Cameron and Trivedi, 1998). Besides, we checked for potential multicollinearity issues calculating the variance inflation factor (VIF). Since VIF values in our models range from 1.01 to 1.85, which remain below the acceptance threshold of 10 (Hair et al., 1998), we have made sure that our results are not affected by significant multicollinearity problems.

In terms of the parameter estimates, both focal independent variables, the *presence* and *depth* of environmental discourse in ORs have a positive and statistically significant impact on the number of helpful votes received by an OR (p<0.001, except for model (3) where p<0.1) across the digital platforms considered (community- vs. transaction-based platform). These findings suggest that ORs embedding the assessment of environment-related aspects have a higher probability to receive helpfulness votes. Furthermore, in light of the positive and significant effect of the *depth* of the environmental discourse, ORs delving deeper into the evaluation of environment-related features appear to be more helpful to the eyes of online readers. These novel results imply a profound reconsideration of the importance of environmental discourse in online review platforms since environmental discourse clearly plays a role in the evaluation of the helpfulness of ORs, with the latter being increasingly deployed by OR platforms to rank ORs and by online readers in their decision-making processes (Gottschalk et al., 2017). It is noteworthy to notice that the effects of the two main independent variables are robust regardless of the wide range of controls embedded in the econometric models.

Regarding the control variables, in line with Park and Nicolau (2015) and Yin et al. (2014) we found a U-shaped relation between review valence and e-WOM helpfulness. Indeed, the linear review valence term (i.e., *Review Rating*) has a negative and significant impact (p<0.001) whereas the quadratic valence term (i.e., *Squared Review Rating*) has a positive and significant impact (p<0.001) implying that extreme ratings are usually assessed as more helpful. To our knowledge, this is the first study that has found such a U-shaped relationship by leveraging data of the transaction platform Booking.com.

As far as the text analytics metrics are concerned, *Review Length* influences positively (p<0.001) e-WOM helpfulness in all the four models considered in line with extant literature (Fang et al., 2016; Mudambi and Schuff, 2010) as online readers tend to find more useful reviews covering in-depth a wide range of aspects of the tourist experience. Conversely, *Review polarity* affects negatively (p<0.001) e-WOM helpfulness in all the four models, confirming a negativity bias inherent in the evaluation of e-WOM helpfulness, and the fact that online readers perceive, as helpful, less extremely positive ORs in terms of polarity (Lee et al., 2017). Furthermore, a higher *Review Readability* score is associated with a higher number of helpful votes (p<0.001) in line with Ghose and Ipeirotis (2011) and Wang et al. (2019). Moreover, the time elapsed after the submission of the OR has a positive effect (p<0.001) on its perceived helpfulness (Wang et al., 2019), meaning that ORs that have spent more time in the digital platform have a higher probability of being associated with a higher number of helpful votes. At the reviewer-level, *expertise* has a positive and significant effect (p<0.001), consistently with extant research (Lee et al., 2017; Park and Nicolau, 2015), whereas identity disclosure (*Reviewer personalized image* and *Country Disclosure*) have discordant effects across the digital platform considered. Indeed, on Booking, having a personalized image and disclosing the country of origin increases the chances of an OR to receive a higher amount of helpful votes (p<0.001), while on TripAdvisor displaying a customize profile image significantly impact e-WOM helpfulness, but showing the country of origin has a negative effect on the number of helpful votes provided by online readers (p<0.001). This could shed light on the fact that online readers use different cognitive heuristics based on the type of digital platform consulted (Xiang et al., 2017). However, the explanation of these differences goes beyond the scope of the current manuscript. Finally, at the hotel-level, in line with Filieri et al. (2020), belonging to a chain negatively influences e-WOM helpfulness (p<0.001) since travelers, in the presence of an established brand, are less likely to conduct activities with the aim of reducing their uncertainty, such as the evaluation of a wide range of online reviews.

Table 4 – Econometric models: environmental discourse impact on e-WOM helpfulness

	Booking (1) Presence	TripAdvisor (2) Presence	Booking (3) Depth	TripAdvisor (4) Depth
Environmental Presence	0.0729***	0.0280***		
	(0.00810)	(0.00357)		
Environmental Depth			0.0108*	0.0329***
			(0.00522)	(0.00307)
Review Rating	-0.451***	-1.158***	-0.451***	-1.158***
	(0.0110)	(0.00767)	(0.0110)	(0.00767)
Squared Review Rating	0.0289***	0.151***	0.0289***	0.151***
	(0.000782)	(0.00112)	(0.000782)	(0.00112)
Log (Reviewer Expertise)	0.0745***	0.0520***	0.0749***	0.0519***
	(0.00305)	(0.00111)	(0.00305)	(0.00111)
Reviewer Personalized Image	0.175***	0.0338***	0.175***	0.0338***
	(0.00694)	(0.00386)	(0.00694)	(0.00386)
Country Disclosure	0.714***	-0.0565***	0.712***	-0.0564***
	(0.141)	(0.00407)	(0.141)	(0.00407)
Log (Review Length)	0.367***	0.539***	0.378***	0.545***
	(0.00358)	(0.00259)	(0.00335)	(0.00242)
Log (Review Readability)	0.0834***	0.0174***	0.0858***	0.0171***
	(0.00523)	(0.00279)	(0.00523)	(0.00279)
Review Polarity	-0.0823***	-0.0573***	-0.0805***	-0.0577***
	(0.00772)	(0.00411)	(0.00772)	(0.00411)
Elapsed Days	0.00116***	0.000393***	0.00116***	0.000393***
	(0.0000166)	(0.00000777)	(0.0000166)	(0.00000777)
Chain	-0.163***	-0.0759***	-0.163***	-0.0761***
	(0.00752)	(0.00327)	(0.00752)	(0.00327)
Further Hotel Controls				
Star Rating	YES	YES	YES	YES
Hotel Popularity	YES	YES	YES	YES
Hotel Reputation	YES	YES	YES	YES
Hotel Continent	YES	YES	YES	YES
Constant	-1.527***	-2.540***	-1.512***	-2.566***
	(0.153)	(0.0256)	(0.153)	(0.0254)
Observations	1,557,766	1,144,461	1,557,766	1,144,461
Pseudo R^2	0.042	0.057	0.042	0.057
AIC	755,903.0	2,246,256.4	755,979.4	2,246,202.8
LR Chi2	32,747.4***	136,417.1***	32,671.0***	136,470.7***
Log Likelihood	-377,930.4	-1,123,107.1	-377,968.6	-1,123,080.4
LR test of $\alpha = 0$	23,799.6***	164,264.8***	23,813.4***	164,183.2***

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.01$, *** $p < 0.001$

To further ensure the reliability of our results we estimated the econometric models excluding ORs having more than 30 helpful votes in the TripAdvisor sample, since they can be perceived as possible outliers (Wang et al., 2019). The findings of these further models are consistent with those presented in Table 4 and enhance the robustness of our results.

5. Discussion and Conclusions

5.1 Conclusions

This work has contributed to expand our knowledge of e-WOM helpfulness by identifying a further antecedent of e-WOM helpfulness, namely online consumers' *environmental discourse* beyond past "non digital" conceptualizations (Peeters and Gössling, 2008) and in line with the seminal study conceptualizing environmental discourse presence and depth (Mariani and Borghi, 2020). Accordingly, we extended the research stream related to e-WOM helpfulness in general (e.g., Ghose and Ipeirotis, 2006, 2011, 2012) and more specifically in the hospitality context (e.g., Liu and Park, 2015; Fang et al., 2016; Filieri et al., 2018, 2019; Lee et al., 2017; Wang et al., 2019). Moreover, we contributed to the big data and analytics research stream in hospitality management literature (Li et al., 2018; Mariani et al., 2018), by suggesting that extrapolating big data analytics from different digital platforms (namely community vs. transaction-based digital platforms) while providing statistically different quantifications of online consumer environmental discourse *presence* and *depth*, offers overall consistent results when it comes to identifying the impact of environmental discourse *presence* and *depth* on e-WOM helpfulness. As such, this study has made relevant contributions to the area at the intersection between e-WOM, big data analytics and sustainable tourism research.

5.2 Theoretical implications

This paper makes several contributions to the area at the intersection of Big Data Analytics, e-WOM and sustainable tourism research in hospitality. Firstly, this study is virtually the first one addressing if and how online consumers' *environmental discourse* influences eWOM helpfulness on multiple digital platforms. This represents a relevant contribution to and extension of the research stream related to e-WOM helpfulness in general (e.g., Ghose and Ipeirotis, 2006, 2011, 2012) and more specifically in the hospitality context (e.g., Liu and Park, 2015; Fang et al., 2016; Filieri et al., 2018, 2019; Lee et al., 2017; Wang et al., 2019). We innovatively find that online consumers' *environmental discourse* influences positively eWOM helpfulness: the underlying mechanism bringing online users to engage with environmental discourse is underpinned by altruism and pro-social behavior theories (Bierhoff, 2002; Carman, 1992; Nagel, 1970; Paul et al., 1993). Accordingly, our study is novel as it is the first in both the wider marketing area and in tourism and hospitality research, to develop a theoretical linkage between environmental discourse, measured through text analytics, and eWOM helpfulness. Secondly, we propose that in the digital age consumers' environmental concerns can be captured by means of digital data streams (Pigni et al., 2016) of online consumer reviews and the related text analytics: without

a big data analytics approach we would have not been able to collect such a vast amount of data and to operationalize and measure two different nuances of online consumers' *environmental discourse* in line with previous literature (Mariani and Borghi, 2020)*: environmental discourse presence* and *environmental discourse depth.* Thirdly, we contribute to the big data and analytics research stream in hospitality management literature (Li et al., 2018; Mariani et al., 2018), by suggesting that extrapolating analytics from different digital platforms (namely community-based vs. transaction-based digital platforms) while providing statistically different quantifications of environmental discourse *presence* and *depth*, offers overall consistent results when it comes to identifying the impact of environmental discourse *presence* and *depth* on e-WOM helpfulness. Fourthly, we enrich the body of e-WOM literature in the sustainable tourism field that so far has mainly focused on broad CSR discourses or green practices at the local or at best national level (e.g., Brazytė et al., 2017; Lee et al., 2016; Peiró-Signes et al., 2014; Yu et al., 2017) by adopting a research design that allows to seize the phenomenon under scrutiny across different countries and continents, this validating the generalizability of the findings. Fifthly, we contribute to the consumer behavior literature revolving around consumers' behaviors and evaluations of environmental and green aspects (e.g., Dolnicar et al., 2008; Thøgersen et al., 2010): rather than deploying small sample surveys asking respondents stated attitudes and behaviors (i.e., perceptions), we look at the real evaluations after consumption by means of big data analytics from a large sample of more than 2.7 million consumers' online reviews. This is certainly in line with the recommended use of big data research in management studies (George et al., 2014). Last, we innovatively build a joined-up body of knowledge matching the identification and examination of the drivers of e-WOM helpfulness (Ghose and Ipeirotis, 2006, 2011, 2012), the emerging debate on consumers' perception of environmental initiatives in the tourism and hospitality industry (Ettinger et al., 2018) and the research stream at the intersection between online consumers' perceptions of environmental issues through e-WOM (e.g., Brazytė et al., 2017; Lee et al., 2016; Yu et al., 2017), by shedding new light on the relationship between e-WOM helpfulness and environmental discourse.

5.3 Practical implications

In this work, we have attempted to explain the role played by online consumer *environmental discourse* in consumers' assessment of e-WOM helpfulness. By analyzing if and to what extent online consumer environmental discourse can enhance e-WOM helpfulness, we help hotel managers and platform managers shed light on one of the factors that might be used by

consumers in their decision making, potentially affecting also their purchase decisions. A number of implications have been developed, with reference to hotel managers, platform developers and platform managers.

As far as hotel managers are concerned, first we suggest that they should become increasingly aware of the importance that customer attach to environmental practices and initiatives and therefore engage more assiduously in environmental initiatives, if their objective is to improve online review helpfulness for other customers reading the reviews written on their hotel. Second, those hotels that are doing something for the environment, either in the form of green initiatives (Kyung et al., 2012) or under the guise of environmental certifications (Peiró-Signes et al., 2014), should make this more explicit both online (on their websites and company's social media) to appeal consumers before their reservations, and also offline during the hotel service encounter. Indeed, eliciting comments regarding environmental aspects would increase e-WOM helpfulness, thus aiding other consumers in their decision-making processes. However, helpful reviews might be highly positive or highly negative (Park and Nicolau, 2015) and therefore communicating with too much emphasis environmental initiatives might be a double hedged sword, as green consumers that have read a positive helpful review about a hotel engaging with environmental issues might be rather severe during the hotel service experience and after it (while leaving a review). Accordingly, hotel managers should make sure to bridge the expected-perceived hotel service gap (Parasuraman et al., 1985), without overpromising or overstating their environmental commitment and initiatives. Third and related to the previous point, more emphasis should be given by managers to educational and edutainment initiatives revolving around their green and environmental practices and the likely impact of their activities. Therefore, an increased educational attention might be relevant for managers to communicate even more effectively their environmental commitment: this might translate not only in helpful e-WOM but also in higher levels of satisfaction with the perceived service (Yu et al., 2017). Fourth, in terms of response strategies, hotel managers that have received helpful online reviews mentioning environmental aspects and impact should explore if they are negatively valenced (i.e., if they display negative ratings) and in that case respond to those reviewers explaining how they intend to tackle the issue or, if unclear, ask the reviewer to articulate what the issue was (as the evaluations are related to attributes that are beyond the standard hotel attributes). This approach might be positively appreciated by other online consumers and potentially reduce negative opinions and disengagement with the hotel. Last, hotel mangers might explicitly ask the hotel guests if they have specific environmental preferences or what are their views about environmental aspects upon check-in and during the

stay, to accommodate their communication strategies and render visible the company's efforts (if any) to engage with environmental practices.

As far as platform developers and managers are concerned, first they could consider developing an environmental tag – for instance in the form of a *green leaf* – that can be associated with the focal ORs to signal if they cover environmental-related aspects. This might be relevant to capture travelers' attention, especially for the segment of prospective hotel customers that care about the environment (i.e., green consumers) and are sensitive to green hotel practices. Secondly, and based on a combination of text analytics (including those computed in this paper), platform developers might introduce a filter – a green filter - allowing customers to rank ORs based on their environmental content. This might enable prospective hotel guests in their information processing before purchase and ultimately increase potential "green" hotel guests' reservation intentions. Third, and related to the previous point, we suggest that platform managers might juxtapose the new "green filter" to the extant "helpfulness filter" to give more possibilities to OR readers to choose how to filter reviews and cope with information overload (Eppler and Mengis, 2004). Fourth, and more generally, platform managers might develop an additional service attribute to juxtapose to the traditional standard attributes (i.e., location, service, staff, etc.) a novel attribute to be named "eco-friendliness". This might help both hotel managers and platform managers to generate insights (beyond those stemming from the text analytics) about customers' perception of environmental and green practices. This is likely to translate in lower bounce rates and increased reservations for the green segments. Last, while the aforementioned recommendations apply to both online travel agencies and community travel review platforms, we would like to stress that they are particularly relevant for OTAs whose business models rely entirely on reservations and need therefore to engage effectively with customers.

5.4 Limitations and Future Research

This study displays a few limitations. First, while we considered a fair number of leading destinations in two continents, future research might extend the analysis to other relevant destinations in Asia (including for instance Hong Kong and Bangkok). Second, consistently with a large body of literature (e.g., Bi et al., 2019) we retained only reviews written in English (which represent the majority in the selection analyzed) as most of the dictionaries used to generate text analytics work effectively with the English language. While we are confident that the results are not significantly affected by this choice, future research might consider dictionaries in other languages (Mariani, Borghi and Okumus, 2020) when and if they will be

extensively validated in relevant scholarly work. Third, future research might consider additional operationalization to be juxtaposed to those that we elaborated (presence and depth) and enrich the model specification with a variable able to capture the depth in a qualitative fashion. Fourth, scholars might consider expanding our work to other OR platforms, namely sharing economy platforms such as Airbnb that apparently are rather hybrid as they mix some transactional characteristics with social and community-based characteristics.

REFERENCES

Alaei, A.R., Becken, S., and Stantic, B. (2019), "Sentiment Analysis in Tourism: Capitalizing on Big Data". *Journal of Travel Research*, Vol. 58 No. 2, pp. 175-191.

Bi, J.W., Liu, Y., Fan, Z.P., and Zhang, J. (2019), "Wisdom of crowds: Conducting importance-performance analysis (IPA) through online reviews", *Tourism Management*, Vol. 70, pp. 460-478.

Bierhoff, H.W. (2002), *Prosocial behaviour*, Psychology Press., New York.

Bonilla Priego, M.J., Najera, J.J. and Font, X. (2011), "Environmental management decision-making in certified hotels", *Journal of Sustainable Tourism*, Vol. 19 No. 3, pp. 361-381.

Brazytė, K., Weber, F., and Schaffner, D. (2017), "Sustainability management of hotels: how do customers respond in online reviews?", *Journal of Quality Assurance in Hospitality and Tourism*, Vol. 18 No. 3, pp. 282-307.

Cameron, A.C., and Trivedi, P.K. (1986), "Econometric models based on count data. Comparisons and applications of some estimators and tests", *Journal of applied econometrics*, Vol. 1 No. 1, pp. 29-53.

Cameron, A.C., and Trivedi, P.K. (1998), *Regression analysis of count data*, Cambridge University Press, Cambridge.

Carman, J.M. (1992), "Theories of Altruism and Behavior Modification Campaigns", *Journal of Macromarketing*, Vol. 12(Spring), pp. 5–18.

Chan, R.Y.K. (2001), "Determinants of Chinese consumers' green purchase behavior." *Psychology & Marketing*, Vol. 18 No. 4, pp. 389-413.

Cho, Y.N. (2015), "Different Shades of Green Consciousness: The Interplay of Sustainability Labeling and Environmental Impact on Product Evaluations", *Journal of Business Ethics*, Vol. 128 No. 1, pp. 73-82.

Cox, M. and Ellsworth, D. (1997), "Managing Big Data for scientific visualization", ACM Siggraph, *MRJ/NASA Ames Research Center*, Vol. 5, pp. 1-17.

Davenport, T.H. (2017), "How analytics have changed in the last 10 years", *Harvard Business Review*, June 2017.

Davenport, T.H. (2014), "How strategists use "big data" to support internal business decisions, discovery and production", *Strategy & Leadership*, Vol. 42 No. 4, pp. 45-50.

D'Acunto D., Tuan A., Dalli, D., Viglia, G., and Okumus, F. (2020), "Do consumers care about CSR in their online reviews? An empirical analysis*", International Journal of Hospitality Management*, Vol. 85, 102342.

Di Pietro, R.B., Cao, Y., and Partlow, C. (2013), "Green practices in upscale foodservice operations", *International Journal of Contemporary Hospitality Management,* Vol. 25 No. 5, pp. 779-796.

Dief, M.E., and Font, X. (2010), "The determinants of hotels' marketing managers' green marketing behavior", *Journal of Sustainable Tourism*, Vol. 18 No. 2, pp. 157-174.

Dolnicar, S., Knezevic Cvelbar, L., and Grün, B. (2017), "Do pro-environmental appeals trigger pro-environmental behavior in hotel guests?", *Journal of Travel Research*, Vol. 56 No. 8, pp. 988-997.

Dolnicar, S. Crouch, G.I. and Long, P. (2008), "Environment-friendly tourists: What do we really know about them?", *Journal of Sustainable Tourism*, Vol. 16 No. 2, pp. 197-210.

Duan, W., Gu, B., and Whinston, A.B. (2008), "Do online reviews matter?—An empirical investigation of panel data", *Decision support systems*, Vol. 45 No. 4, pp. 1007-1016.

Dutta, K., Umashankar, V., Choi, G., and Parsa, H.G., (2008), "A comparative study of consumers' green practice orientation in India and the United States: a study from the restaurant industry", *Journal of Foodservice Business Research*, Vol. 11 No. 3, pp. 269-285.

Eppler, M.J., and Mengis, J. (2004), "The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines", *The information society,* Vol. 20 No. 5, pp. 325-344.

Ettinger, A., Grabner-Kraeuter, S., and Terlutter, R. (2018), "Online CSR communication in the hotel industry: Evidence from small hotels", *International Journal of Hospitality Management*, Vol. 68, pp. 94-104.

Fang, B., Ye, Q., Kucukusta, D., and Law, R. (2016), "Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics", *Tourism Management*, Vol. 52, pp. 498-506.

Fang, Y.-H. (2014) "Beyond the credibility of electronic word of mouth: Exploring eWOM adoption on social networking sites from affective and curiosity perspectives", *International Journal of Electronic Commerce*, Vol. 18 No. 3, pp. 67-102

Felbermayr, A. and Nanopoulos (2016), "The Role of Emotions for the Perceived Usefulness in Online Customer Reviews", *Journal of Interactive Marketing*, Vol. 36, pp. 60-76.

Filieri, R. (2015), "What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in eWOM", *Journal of Business Research, Vol.* 68 No. 6, pp. 1261-1270.

Filieri, R. and McLeay, F. (2014), "E-WOM and accommodation: An analysis of the factors that influence travelers' adoption of information from online reviews", *Journal of Travel Research*, Vol. 53 No. 1, pp. 44-57.

Filieri, R., Raguseo, E., and Vitari, C. (2018), "When are extreme ratings more helpful? Empirical evidence on the moderating effects of review characteristics and product type", *Computers in Human Behavior*, Vol. 88, pp. 134-142.

Filieri, R., Raguseo, E., and Vitari, C. (2019), "What moderates the influence of extremely negative ratings? The role of review and reviewer characteristics", *International Journal of Hospitality Management*, Vol. 77, pp. 333-341.

Filieri, R., Vitari, C., and Raguseo, E. (2020), "Extremely Negative Ratings and Online Consumer Review Helpfulness: The Moderating Role of Product Quality Signals", *Journal of Travel Research*, forthcoming.

Font, X., and Lynes J. (2018), "Corporate social responsibility in tourism and hospitality", *Journal of Sustainable Tourism*, Vol. 26 No. 7, pp. 1027-1042.

Forman, C., Ghose, A., and Wiesenfeld, B. (2008), "Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets", *Information Systems Research*, Vol. 19 No. 3, pp. 291–313.

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., and Gnanzou, D. (2015), "How big data can make big impact: Findings from a systematic review and a longitudinal case study", *International Journal of Production Economics*, Vol. 165, pp. 234-246.

Fosso Wamba, S., Gunasekaran, A., Akter, S., Ren, S.J.F.,Dubey, R., Childe, S.J. (2017), "Big data analytics and firm performance: Effects of dynamic capabilities", *Journal of Business Research*, Vol. 70, pp. 356–365.

Fosso Wamba, S., Dubey, R., Gunasekaran, A., and Akter, S. (2020), "The performance effects of big data analytics and supply chain ambidexterity: The moderating effect of environmental dynamism", *International Journal of Production Economics*, 107498.

Furner, C.P., and Zinko, R.A. (2017), "The influence of information overload on the development of trust and purchase intention based on online product reviews in a mobile vs. web environment: an empirical investigation", *Electronic Markets*, Vol. 27 No. 3, pp. 211–224

Gazzoli, G., Kim, W.G., and Palakurthi, R. (2008), "Online distribution strategies and competition: are the global hotel companies getting it right?", *International Journal of Contemporary Hospitality Management*, Vol. 20 No. 4, pp. 375-387.

Geerts, W. (2018), "Top 100 City Destinations 2018, Euromonitor International", https://go.euromonitor.com/white-paper-travel-2018-100-cities (accessed on 29 May, 2020).

Geetha, M., Singha, P., and Sinha, S., (2017), "Relationship between customer sentiment and online customer ratings for hotels-An empirical analysis." *Tourism Management*, Vol. 61, pp. 43-54.

George, G., Haas, M.R., and Pentland, A. (2014), "Big data and management", *Academy of Management Journal*, Vol. 57, pp. 321–325

Getz, D., and Page, S. (2016), "Progress and prospects for event tourism research", *Tourism Management*, Vol. 52, pp. 593-631.

Ghose, A., and Ipeirotis, P.G. (2006), "Designing ranking systems for consumer reviews: The impact of review subjectivity on product sales and review quality", *Proceedings of the 16th annual workshop on information technology and systems*, pp. 303–310.

Ghose, A., and Ipeirotis, P.G. (2007), "Designing novel review ranking systems: predicting the usefulness and impact of reviews, *Proceedings of the ninth international conference on Electronic commerce*. ACM.

Ghose, A., and Ipeirotis, P.G. (2011), "Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics", *IEEE Transactions on Knowledge and Data Engineering, Vol.* 23 No. 10, pp. 1498-1512.

Ghose, A., Ipeirotis, P.G., Li, B. (2012), "Designing ranking systems for hotels on travel search engines by mining user-generated and crowdsourced content", *Marketing Science, Vol.* 31 No. 3, pp. 493-520.

Gligorijevic, B. (2016), "Review platforms in destinations and hospitality". In: R. Egger, I. Gula and D. Walcher, eds. *Open tourism*. Springer Berlin Heidelberg, pp. 215-228.

Gössling, S., and Peeters, P. (2015), "Assessing tourism's global environmental impact 1900 2050", *Journal of Sustainable Tourism,* Vol. 23 No. 5, pp. 639-659.

Gössling, S., and Peeters, P. (2007), ""It does not harm the environment!" An analysis of industry discourses on tourism, air travel and the environment", *Journal of Sustainable Tourism,* Vol. 15, No. 4, pp. 402-417.

Gottschalk, S.A., and Mafael, A. (2017), "Cutting through the online review jungle— investigating selective eWOM processing", *Journal of Interactive Marketing*, Vol. 37, pp. 89-104.

Gretzel, U., and Yoo, K.H. (2008), "Use and impact of online travel reviews", *Information and communication technologies in tourism* 2008, pp. 35-46.

Guix, M., Bonilla Priego, M.J. and Font, X. (2018), "The process of sustainability reporting in international hotel groups: An analysis of stakeholder inclusiveness, materiality and responsiveness". *Journal of Sustainable Tourism*, Vol. 26, pp. 1-22.

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

Gustin, M.E., and Weaver, P.A. (1996), "Are hotels prepared for the environmental consumer?", *Journal of Hospitality and Tourism Research,* Vol. 20 No. 2, pp. 1-14.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., and Tatham, R.L. (1998), *Multivariate data analysis*, Upper Saddle River, NJ: Prentice hall.

Helen Chun, H., and Giebelhausen, M. (2012), "Reversing the green backlash in services: credible competitors help large companies go green", *Journal of Service Management*, Vol. 23 No. 3, pp.400-415.

Hennig-Thurau, T., Gwinner, K.P., Walsh, G., and Gremler, D.D. (2004), "Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet?", *Journal of Interactive Marketing, Vol. 18* No. 1, pp. 38-52.

Hong, H., Xu, D., Wang, G. A., and Fan, W. (2017). Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decision Support Systems, Vol. 102,* pp. 1–11.

Hu, N., Liu, L., and Zhang, J.J. (2008), "Do online reviews affect product sales? The role of reviewer characteristics and temporal effects", *Information Technology and management*, Vol. 9 No. 3, pp. 201-214.

Hu, W., Gong, Z., and Guo, J. (2010), "Mining product features from online reviews". In *Proceedings - IEEE International Conference on E-Business Engineering, ICEBE 2010*, pp. 24-29.

Hu, Y. H. and Chen, K. (2016), "Predicting hotel review helpfulness: The impact of review visibility, and interaction between hotel stars and review ratings", *International Journal of Information Management, Vol. 36* No. 6, pp. 929-944.

Hutto, C.J., and Gilbert, E. (2014), "Vader: A parsimonious rule-based model for sentiment analysis of social media text", *In Eighth international AAAI conference on weblogs and social media*.

Kemper, J.A., Hall, C.M., and Ballantine, P.W. (2019), "Marketing and sustainability: Business as usual or changing worldviews?" *Sustainability* (Switzerland), Vol. 11 No. 3, 780.

King, R.A., Racherla, P., and Bush, V.D. (2014), "What we know and don't know about online word-of-mouth: A review and synthesis of the literature", *Journal of Interactive Marketing, Vol. 28* No.3, pp. 167-183

Korfiatis, N., García-Bariocanal, E., and Sánchez-Alonso, S. (2012), "Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content", *Electronic Commerce Research and Applications*, Vol. 11 No. 3, pp. 205-217.

Kyung, H.K., Stein, L., Heo, C.Y., and Lee, S. (2012), "Consumers' willingness to pay for green initiatives of the hotel industry", *International Journal of Hospitality Management*, Vol. 31 No. 2, pp. 564-572.

Kwok, L., and Xie, K.L. (2016), "Factors contributing to the helpfulness of online hotel reviews: Does manager response play a role?" *International Journal of Contemporary Hospitality Management, Vol. 28* No. 10, pp. 2156-2177.

Laney, D. (2001), "3D Data Management: Controlling Data Volume, Velocity and Variety", META Group Research Note, 6. https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf (accessed 4th June, 2020).

Lahuerta-Otero, E., and Cordero-Gutiérrez, R. (2016), "Looking for the perfect tweet. The use of data mining techniques to find influencers on Twitter", *Computers in Human Behavior*, Vol. 64, pp. 575-583.

Lee, H., Jai, T.M., and Li, X. (2016), "Guests' perceptions of green hotel practices and management responses on TripAdvisor", *Journal of Hospitality and Tourism Technology*, Vol. 7 No. 2, pp. 182-199.

Lee, E. J., and Shin, S. Y. (2014), "When do consumers buy online product reviews? Effects of review quality, product type, and reviewer's photo", *Computers in Human Behavior*, Vol. 31, pp. 356–366.

Lee, M., Jeong, M., and Lee, J. (2017). Roles of negative emotions in customers' perceived helpfulness of hotel reviews on a user-generated review website. *International Journal of Contemporary Hospitality Management, Vol.* 29 No. 2, pp. 762-783.

Li J., Xu L., Tang L., Wang S., and Li L. (2018), "Big data in tourism research: A literature review", *Tourism Management*, Vol. 68, pp. 301-323.

Litvin, S. W., Goldsmith, R. E., and Pan, B. (2008), "Electronic word-of-mouth in hospitality and tourism management", *Tourism Management*, Vol. 29 No. 3, pp. 458-468.

Liu, Z., and Park, S. (2015), "What makes a useful online review? Implication for travel product websites", *Tourism Management, Vol. 47*, pp. 140-151.

Loria, S., Keen, P., Honnibal, M., Yankovsky, R., Karesh, D., and Dempsey, E. (2014), "Textblob: simplified text processing". Secondary TextBlob: Simplified Text Processing.

Luo, C., Lan, Y., Wang, C., and Ma, L. (2013), "The effect of information consistency and information aggregation on eWOM readers' perception of information overload", *Paper presented at the Proceedings - Pacific Asia Conference on Information Systems, PACIS*.

Ma, E., Cheng, M., and Hsiao, A. (2018), "Sentiment analysis—a review and agenda for future research in hospitality contexts", *International Journal of Contemporary Hospitality Management*, Vol. 30 No. 11, pp. 3287-3308.

Mariani, M.M., Baggio, R., Buhalis, D., and Longhi, C. (Eds.). (2014). *Tourism management, marketing, and development: volume I: the importance of networks and ICTs* (Vol. 1). Springer, New York.

Mariani, M.M., Baggio, R., Fuchs, M., and Höpken, W. (2018), "Business Intelligence and Big Data in Hospitality and Tourism: A Systematic Literature Review",

International Journal of Contemporary Hospitality Management, Vol. 30 No. 12, pp. 3514-3554.

Mariani, M. and Borghi, M. (2020), "Environmental discourse in hotel online reviews: a big data analysis"*, Journal of Sustainable Tourism.* doi: 10.1080/09669582.2020.1858303.

Mariani, M. M. and Borghi, M. (2020), "Online review helpfulness and firms' financial performance: an empirical study in a service industry", International Journal of Electronic Commerce, Vol. 24 No. 4, pp. 421-449.

Mariani, M. M., Borghi, M. and Okumus, F. (2020), "Unravelling the effects of cultural differences in the online appraisal of hospitality and tourism services", International Journal of Hospitality Management, Vol. 90, 102606.

Mariani, M. M. and Fosso Wamba, S. (2020), "Exploring how consumer goods companies innovate in the digital age: The role of big data analytics companies", Journal of Business Research, Vol. 121, pp. 338-352.

Mariani, M. M. and Visani, F. (2019). "Embedding eWOM into efficiency DEA modelling: An application to the hospitality sector", *International Journal of Hospitality Management*, Vol. 80, pp. 1-12.

Mariani, M.M., Ek Styven, M. and Nataraajan, R. (2021), "Social comparison orientation and frequency: A study on international travel bloggers", *Journal of Business Research*, Vol. 123, pp. 232-240.

Mariani, M. and Predvoditeleva, M. (2019), "How do online reviewers' cultural traits and perceived experience influence hotel online ratings?", International Journal of Contemporary Hospitality Management, Vol. 31 No. 12, pp. 4543-4573.

Mariani, M. M. and Matarazzo, M. (2020), "Does cultural distance affect online review ratings? Measuring international customers' satisfaction with services leveraging digital platforms and big data", Journal of Management and Governance, pp. 1-22.

Mariani, M.M., Borghi, M., (2018), "Effects of the Booking.com rating system: Bringing hotel class into the picture", *Tourism Management*, Vol. 66, pp. 47-52.

Marriot (2018), "What is 'Make a Green Choice'?" https://help.marriott.com/s/article/Article-22164#:~:text='Make%20a%20Green%20Choice'%20is,Rate%20Plan%20or%20Booking%20Method (accessed on 29 May, 2020).

Mudambi, S.M., and Schuff, D. (2010), "Research note: What makes a helpful online review? A study of customer reviews on Amazon.com", *MIS quarterly*, pp. 185-200.

Naderi, I., and Van Steenburg, E. (2018), "Me first, then the environment: young Millennials as green consumers", *Young Consumers*, Vol. 19 No.3, pp. 280-295.

Nagel, T. (1970), *The Possibility of Altruism*, Clarendon, Oxford, England.

Pagiaslis, A., and Krontalis, A.K. (2014) "Green Consumption Behavior Antecedents: Environmental Concern, Knowledge, and Beliefs", *Psychology & Marketing*, Vol. 31 No.5, pp. 335-348.

Parasuraman, A., Zeithaml, V.A., and Berry, L.L. (1985). A conceptual model of service quality and its implications for future research. *Journal of marketing*, Vol. 49 No. 4, pp. 41-50.

Park, D.-H., and Lee, J. (2008), "eWOM overload and its effect on consumer behavioral intention depending on consumer involvement", *Electronic Commerce Research and Applications*, Vol. 7 No. 4, pp. 386-398.

Park, D.-H., Lee, J., and Han, I. (2006), "Information overload and its consequences in the context of online consumer reviews", *Proceedings of the 10th Pacific Asia Conference on Information Systems: ICT and Innovation Economy*, pp. 197–209

Park, S., and Nicolau, J.L. (2015), "Asymmetric effects of online consumer reviews", *Annals of Tourism Research*, Vol. 50, pp. 67-83.

Paul, E.F., Miller, F.D., Jr., and Paul, J. (1993), *Altruism*, Cambridge University Press, Cambridge, UK.

Pedersen, E.R., and Neergaard, P. (2006) "Caveat Emptor - Let the buyer beware! Environmental labelling and the limitations of 'Green' consumerism", *Business Strategy and the Environment*, Vol. 15 No. 1, pp. 15-29.

Peeters, P. and Gössling, S. (2008), "Environmental discourses in the aviation industry: The reproduction of mobility", *Tourism and Mobilities: Local Global Connections*. CABI, pp. 187-203.

Peiró-Signes, A., Segarra-Oña, M.D.V., Verma, R., Mondéjar-Jiménez, J., and Vargas-Vargas, M., (2014), "The impact of environmental certification on hotel guest ratings." *Cornell Hospitality Quarterly*, Vol. 55 No. 1, pp. 40-51.

Pencle, N., Mălăescu, I. (2016), "What's in the words? Development and validation of a multidimensional dictionary for CSR and application using prospectuses", *Journal of Emerging Technologis in Accounting*, Vol. 13 No. 2, pp. 109-127.

Pigni, F., Piccoli, G., and Watson, R. (2016), "Digital Data Streams: Creating value from the real-time flow of big data", *California Management Review*, Vol. 58 No. 3, pp. 5-25.

Plutchik, R. (1994), *The psychology and biology of emotion*, Harper Collins College Publishers, New York, NY.

Raub, S. and Martin-Rios, C. (2019), ""Think sustainable, act local" – a stakeholder-filter-model for translating SDGs into sustainability initiatives with local impact", *International Journal of Contemporary Hospitality Management*, Vol. 31 No.6, pp. 2428-2447

Revinate (2019). Global Hotel Reputation Benchmark Report 2019. https://www.revinate.com/benchmark/reputation/ (accessed on 29 May, 2020).

Rosario, A.B., Sotgiu, F., De Valck, K., and Bijmolt, T.H.A. (2016), "The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors." *Journal of Marketing Research,* Vol. 53 No. 3, pp. 297-318.

Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., and Harnisch, M. (2015), "Industry 4.0: the future of productivity and growth in manufacturing industries", *Boston Consulting Group*, 9, 2015, Retrieved from https://www.bcg.com/publications/2015/engineered_products_project_business_industry_4_future_productivity_growth_manufacturing_industries.aspx

Shin, S., Chung, N., Xiang, Z., and Koo, C. (2019), "Assessing the impact of textual content concreteness on helpfulness in online travel reviews", *Journal of Travel Research*, Vol. 58 No. 4, pp. 579-593.

Silva, R. (2015), "Multimarket contact, differentiation, and prices of chain hotels", *Tourism Management*, Vol. 48, pp. 305-315.

Singh, J.P., Irani, S., Rana, N.P., Dwivedi, Y.K., Saumya, S., and Roy, P.K. (2017), "Predicting the "helpfulness" of online consumer reviews", *Journal of Business Research, Vol.* 70, pp. 346-355.

Smith, E.A., and Senter, R.J. (1967), "Automated Readability Index", *Aerospace Medical Research Laboratories*, Wright-Patterson Airforce Base, OH.

Sun, T., Youn, S., Wu, G., and Kuntaraporn, M. (2006), "Online word-of-mouth (or mouse): An exploration of its antecedents and consequences" *Journal of Computer-Mediated Communication, Vol.* 11 No. 4, 1104-1127.

Sundaram, D. S., Mitra, K., and Webster, C. (1998), "Word-of-mouth communications: A motivational analysis", *ACR North American Advances*.

Thøgersen, J., Jørgensen, A.-K., and Sandager, S. (2012), "Consumer Decision Making Regarding a "Green" Everyday Product", *Psychology & Marketing*, Vol. 29 No.4, pp. 187-197.

Thøgersen, J., Haugaard, P., and Olesen, A. (2010), "Consumer responses to ecolabels", *European Journal of Marketing*, Vol. 44 No. 11/12, pp. 1787-1810.

Tsao, W.C., Hsieh, M.T., Shih, L.W., and Lin, T.M. (2015), "Compliance with eWOM: The influence of hotel reviews on booking intention from the perspective of consumer conformity", *International Journal of Hospitality Management*, Vol. 46, pp. 99-111.

Wang, X., Tang, L.R., and Kim, E. (2019), "More than words: Do emotional content and linguistic style matching matter on restaurant review helpfulness?", *International Journal of Hospitality Management*, Vol. 77, pp. 438-447.

Witten, I. H., Frank, E., Hall, M. A., and Pal, C. J. (2016), "Data Mining: Practical machine learning tools and techniques" (4th ed.). Cambridge, MA, USA: Morgan Kaufmann.

Wooldridge, J.M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Xiang, Z., Du, Q., Ma, Y., and Fan, W. (2017), "A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism", *Tourism Management*, Vol. 58, pp. 51-65.

Xiang, Z., Schwartz, Z., Gerdes Jr, J. H., and Uysal, M. (2015), "What can big data and text analytics tell us about hotel guest experience and satisfaction?", *International Journal of Hospitality Management*, Vol. 44, pp. 120-130.

Xu, X. and Li, Y. (2016), "The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: a text mining approach", *International Journal of Hospitality Management,* Vol. 55, pp. 57–69.

Yang, Y., Park, S., and Hu, X. (2018), "Electronic word of mouth and hotel performance: a meta-analysis", *Tourism Management, Vol.* 67, pp. 248-260.

Ye, Q., Law, R., Gu, B., and Chen, W. (2011), "The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings", *Computers in Human behavior*, Vol. 27 No. 2, pp. 634-639.

Ye, Q., Law, R., and Gu, B. (2009), "The impact of online user reviews on hotel room sales". *International Journal of Hospitality Management, Vol.* 28 No. 1, pp. 180-182.

Yin, D., Bond, S. D., and Zhang, H. (2014), "Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews", *MIS quarterly*, Vol. 38 No. 2, pp. 539-560.

You, Y., Vadakkepatt, G.G., and Joshi, A.M. (2015), "A Meta-analysis of Electronic Word-of-mouth Elasticity", *Journal of Marketing*, Vol. 79 No. 2, pp. 19-39.

Yu, Y., Li, X., and Jai, T.M. (2017), "The impact of green experience on customer satisfaction: evidence from TripAdvisor", *International Journal of Contemporary Hospitality Management*, Vol. 29 No. 5, pp. 1340-1361.

Zhang, X., Yu, Y., Li, H., and Lin, Z. (2016), "Sentimental interplay between structured and unstructured user-generated contents: an empirical study on online hotel reviews", *Online Information Review,* Vol. 40 No. 1, pp. 119–145.

Zhao, Y., Xu, X., and Wang, M. (2019), "Predicting overall customer satisfaction: big data evidence from hotel online textual reviews", *International Journal of Hospitality Management*, Vol. 76, pp. 111-121.