Citation:

Sharma, A., Park, S., & Nicolau, J. L. (2020). Testing loss aversion and diminishing sensitivity in review sentiment. *Tourism Management*, *77*, 104020.

TESTING LOSS AVERSION AND DIMINISHING SENSITIVITY IN REVIEW SENTIMENT

ABSTRACT

This article analyzes the relationship between ratings and review sentiment by introducing, for the first time, the tenets of Prospect Theory. Specifically, we test loss aversion and diminishing sensitivity on a sample of 132,486 reviews and find that: first, negative deviations in ratings (receiving a service with worse performance than expected) bring about a higher impact on review sentiment than positive deviations of equal magnitude (receiving a service with better performance than expected), thus, confirming loss aversion; and second, regardless of whether the service received is better or worse than expected, variations in ratings closer to the reference point result in higher marginal impacts on sentiment than equivalent variations further away from the reference point, thus, proving diminishing sensitivity. These results have relevant theoretical implications related to the use of relative vs absolute measures and the cognitive bias involved, and managerial implications linked to meeting expectations and service recovery.

Keywords: Prospect theory; loss aversion; diminishing sensitivity; review sentiment; rating.

1. INTRODUCTION

While individuals might traditionally have relied on advertisers, friends and family to learn about products and services, the advent of the Internet has unquestionably transformed the ways in which consumers seek information and communicate consumption experiences. The proliferation of user-generated content fueled by the diffusion of Web 2.0 technologies has facilitated more widespread information sharing among consumers at all stages in the consumption process- before, during and after. Although such information sharing also takes place across multiple other electronic platforms such as Facebook and Twitter, online consumer reviews are undoubtedly one of the principal tools that individuals use today to learn about various products and services, and provide feedback about their personal consumption experiences with regards to these products and services.

Online customer reviews are certainly relevant in the context of hospitality and tourism- in a recent report from TripAdvisor (2019), 72% of those surveyed reported that they either always or frequently read online reviews when making decisions about where to eat and things to do, and as many as 81% always or frequently use online reviews before booking a place to stay. Given these statistics, there is little doubt that consumers place a lot of trust in online reviews when making consumptions decisions relating to hospitality and travel. In fact, a Nielsen survey suggests that as many as 70% of consumers trust online reviews as a form of advertising (Grimes, 2012).

If one were to somehow accurately exclude the reviews we knew to be deliberately falsified or fabricated, one might be tempted – certainly for the sake of convenience – to accept *all* remaining reviews as objective and trustworthy. One might even find support for such a position in the "rational actor" model of neoclassical economics under which we would deduce that reviewers – possessing limitless cognitive ability and operating under full information – would be writing reviews that are not in any way biased by any prejudices, emotions or sentiments.

If we were to adopt a more realistic approach, however, we would allow for the fact that reviewers- being human- are limited in their capacity to process information, and that the reviews they write are often skewed by many of the same biases that afflict everyday decision making and judgements. Indeed, a number of seminal studies in cognitive psychology and behavioral economics have adopted alternative models that are rigorous yet permit realistic investigations into human behavior and decision making without invoking the neoclassical doctrine of perfect rationality (for example, Simon, 1955; Thaler, 1980).

We believe that Kahneman and Tversky's (1979) prospect theory, which offers a powerful explanation of decision making under risk and uncertainty, might better uncover certain behavioral tendencies and cognitive biases that have the potential to distort the sentiment expressed in online consumer reviews. Under this theory, "prospects" – or potential outcomes – are weighed relative to a psychological reference point.

Two important properties result from this reference dependent framework of prospect theory- loss aversion and diminishing sensitivity. Whereas the former posits that losses loom larger than equivalent gains, the latter argues that for both gains and losses, as one moves further away from the reference point, changes in sensitivity occur at a decreasing rate.

In the context of a service experience one may think of this reference point as a service expectation, whereas one may think of gains and losses in terms of service performance above or below this expected level of service. Presumably, review sentiment – as reflected in sentiment scores resulting from a sentiment analysis - would reveal the extent to which the service received differs from the service expected. Accordingly, the objective of this study is to understand how loss aversion and diminishing sensitivity affect review sentiment. While this objective might have appeared to be somewhat unrealistic only a few years ago because the then limitations of text-mining programs, advances in the automated analysis of big data in the last decade have made processes such as sentiment analysis both quick and reliable. Consequently, the more specific objective of this research is to investigate whether sentiment scores (ranging from -1 to +1 obtained by using Python's VADER package to analyze a set of 132,486 TripAdvisor reviews) are susceptible to loss aversion and diminishing sensitivity- the two primary reference dependent behaviors explained by prospect theory. That certain behavioral inconsistencies may manifest themselves in aspects relating to online reviews would not in itself be a novel finding. It is long understood, for example, that negative reviews can hurt sellers more than positive reviews help (Basuroy et al., 2003). Chevalier and Mayzlin (2006) show that similar properties apply also to marginal impacts- incremental negative reviews hurt sales more than incremental positive reviews. Seemingly incongruent behavioral patterns have also been investigated in other aspects of online reviews. Tsang and Prendergast's (2009), for instance, examine whether inconsistencies in valence (positive or negative) between review text and ratings affect consumer perceptions of interestingness, trustworthiness and purchase intention. Although we readily acknowledge that the aforementioned studies provide a number of insights into the strand of literature in which the present study fits, none of the above contributions obtains measures based on sentiment analysis. Moreover, none of the studies mentioned above uses prospect theory as its explanatory framework. Interestingly, Tsang and Prendergast (2009) do briefly mention prospect theory, but develop their paper in a vastly different framework that results in objectives that are fundamentally dissimilar to our objectives. To the best of our knowledge, no prior study has examined the relationship between review rating and review sentiment under the lens of prospect theory.

2. PROSPECT THEORY IN TOURISM RESEARCH

Behavioral aspects of travelers' decision making have long been of interest to tourism researchers. In this scholarship, it is generally accepted that in addition to various other drivers of choice and decision making, travelers rely extensively on past experiences (Kim, 2014; Zhang et al., 2016). In fact, Mazursky (1989) asserts that past travel experiences may be even more relevant than other external sources of information.

The importance of past experiences in molding traveler preferences underlines the applicability of prospect theory (Kahneman and Tversky, 1979) in tourism research. This is because prospect theory posits that the evaluations that people make tend to be reference dependent- that is, an individual's assessment of an outcome, product, service or experience is often made in terms of deviations from a certain point of reference. Two important and well-established principles relating to human preferences follow from this reference dependent evaluative process (Kahneman and Tversky, 1979): loss aversion and diminishing sensitivity. The former suggests that the deviations from the aforementioned reference point are valued differently depending on the direction of the deviation. More specifically, this principle asserts that individuals are more sensitive to negative deviations (losses) from the reference point than they are to positive deviations (gains). The latter – diminishing sensitivity – contends that for both positive and negative deviations, the magnitude of the distance from the reference point also determines the marginal impacts resulting from a particular deviation. For both gains and losses, deviations closer to the reference point produce higher marginal impacts than equivalent changes further away from the reference point.

The principles of loss aversion and diminishing returns explain the archetype prospect theory value function that is used to describe subjective assessments of specific outcomes. This value function tends to have a higher slope for losses than for gains (loss aversion), and, in terms of shape, is concave for gains but convex for losses (diminishing sensitivity). Prospect theory also helps explain several routinely observed human preferences that may have been viewed as irrational, or at the very least, inconsistent, under neoclassical models of human behavior. The principle of loss aversion predicts for example why a gain of $100 may not offset a loss of $100- the disutility induced by the loss exceeds in absolute terms the increase in satisfaction resulting from the $100 gain. The principle of diminishing sensitivity, on the other hand, would help explain why, for example, a $10 discount on a $20 item may appear more rewarding than a $10 discount on a $50 item, even though both cases result in equivalent savings ($10).

There are a number of studies in the tourism literature that have used the principal conclusions of prospect theory – diminishing sensitivity and loss aversion, to investigate various dimensions relating to hospitality tourism consumption. As one might expect, it is the pricing literature within the fields of hospitality and tourism where reference dependent preferences have been of some interest. The notion of `reference prices' – a standard against which consumers evaluate actual prices of a product to evaluate its attractiveness (Monroe, 1973), is invoked in a number of hospitality and tourism related studies (for example, Oh, 2003; Nicolau, 2008; Viglia et al., 2016). Reference dependent preferences have also been observed, for instance, in the context of tourism prices and tourist overspending behavior (Nguyen, 2016), tourist satisfaction scores (Kim and Canina, 2015), destination satisfaction and revisit intentions resulting from changes in destination image (Park and Nicolau, 2019), and wait times in tourism (Hernandez-Maskivker et al., 2019).

In one of the earlier applications of prospect theory in tourism research, Nicolau (2008), using a multinomial logit model, observes significance levels of reference price dependence among Spanish vacationers. The Nicolau (2008)'s study detects substantial levels of loss aversion, as manifested in Spaniards' higher levels of sensitivity to price increases relative to their reference price than the sensitivity exhibited to corresponding levels of price decreases. Asymmetric reactions to price fluctuations – consistent with the predictions of loss aversion - are also observed in Nicolau's (2011) study of destination choice using data from the Spanish Holidaying Behavior survey. Interestingly, however, the levels of loss aversion in this study are found to be moderated by certain subjective characteristics such as a person's cultural interest in the destination-individuals expressing more cultural appreciation for a destination tend to be less loss averse than those who were more culturally indifferent toward the destination.

Asymmetries explained by loss aversion have also been found in tourism demand across business cycles- Smeral (2017) argues that loss aversion may be one of the key reasons that income and price effects on tourism demand cannot be assumed to remain stable under varying macroeconomic conditions. Masiero and Qiu (2018) detect substantial levels of reference dependent decision making in destination choice. The authors observe loss aversion among long-haul leisure travelers in several relevant attributes including hospitality, attractions (cultural, natural and entertainment), services (food and dining, transportation), and travel budgets. Additionally, the authors also detect that inertia for reference levels is observed in several of these attributes (Masiero and Qiu, 2018).

Certainly, gains and/or losses representing deviations from reference points need not assume direct monetary values. For example, loss aversion and reference dependence have also been investigated in Hernandez-Maskivker et al's (2019) in the context of wait times in tourism. Using data obtained from visitors to theme parks, these authors find evidence of reference dependence in theme park visitors' willingness to pay a higher price for express passes. In this sense, waiting is perceived as a cost, and a reference points are formed by the theme park visitors based on expected wait times (Hernandez-Maskivker et al., 2019). Reference dependent evaluations are then manifested in the trade-offs between wait times and willingness to pay for express passes.

In recent years, there has been a realization in the literature that travelers' perceived helpfulness of online reviews too can be based on a reference dependent evaluative process: Park and Nicolau (2015) find for example that travelers find positive reviews to be less useful than negative reviews- a behavioral preference that the authors attribute to the principle of loss aversion. There is evidence that reviewers exhibit reference dependence. Mellinas, Nicolau and Park (2019) demonstrate that reviewers' assessment of a hotel's locations are influenced by how they evaluate other attributes describing a particular hotel. Equally importantly, these authors also argue that the asymmetry observed in these ratings – with dissatisfaction resulting in more severe reviews than the corresponding praises stemming from satisfaction – is consistent with the principle of loss aversion.

2.1 The relationship between rating and sentiment

The meteoric growth of social media in recent years has facilitated unprecedented levels of information sharing among consumers. While this information sharing also encompasses multiple other electronic channels of user generated content such as Twitter and Facebook, online reviews and ratings are widely recognized today as one of the key elements in the overall consumption process. For many people, online content is even more trustworthy than information obtained from other sources such as professionals and marketers (Fotis et al., 2012; Gretzel and Yoo, 2008). Indeed, Nielsen's Global Trust in Advertising Survey suggests that online customer reviews are second only to word-of-mouth recommendations from friends and family as a trusted source of advertising (Grimes, 2012). It is hardly surprising therefore that each of the top 10 online retailers in the United States display reviews for the products they sell (Askalidis, Kim and Malthouse, 2017).

Online reviews play a central role not only in the consumption of products, but also in the consumption of services and experiences including those relating to travel and tourism. Websites such as TripAdvisor and Booking.com are frequented by travelers before, during and after the completion of trips (Liu and Park, 2015) for purposes of obtaining information from reviews, and providing feedback about completed travel experiences. Online reviews are used by travelers in selecting and evaluating travel destinations, flights, hotels, restaurants, attractions, etc.

The trust that is accorded to online reviews by prospective travelers is evidenced by the number of people who use online reviews when making travel plans. Moreover, under the rigid and perhaps unrealistic assumptions of the age-old rational actor model of economics, one would expect that reviewers would be characterized by unlimited cognitive capacity, and be able to objectively write reviews that are free of any biases or sentiments. At the same time, one must consider, however, that several studies in cognitive psychology and behavioral economics have challenged the rationality assumption, and instead offered a number of viable alternatives. If examined under the lens of some of these alternatives, one may question the merits of deciphering online reviews at their face value.

Simon's (1972) theory of bounded rationality suggests for instance that because of limitations in both the availability of information as well as in the capacity of individuals to process information, there exist bounds on rationality that affect decision-making and in turn result in suboptimal decisions. Given the inherent uncertainty defining situations involving limited or incomplete information, individuals then rely on certain strategies or heuristics in efforts to make decisions. In the context of travel behavior, online reviews and ratings could be thought of as one such heuristic mechanism that prospective travelers adopt to manage the uncertainty and information gaps that constrain travel related decision making (Park and Nicolau, 2015; Wattanacharoensil and La-ornual, 2019). One would suspect, however, that reviewers too would be impaired by many of the cognitive biases – including loss aversion and diminishing sensitivity - that affect general decision making and choice behavior (Mellinas et al., 2019; Park and Nicolau, 2015). We shall momentarily describe how loss aversion and diminishing sensitivity afflict review sentiment but it should nonetheless be obvious here that reviews distorted by cognitive biases are fundamentally different from reviews that entail deliberate fabrication and manipulation. No willful intent is required on part of the reviewer in the former case, where reviews may simply be biased by the reviewer's emotions and sentiments rather than by any calculated scheme or plan by the reviewer.

Sentiments are of course inherent in user-generated content. Sentiment analysis – sometimes referred to in the literature as sentiment mining - has in recent years been a powerful technique used in the literature (including the hospitality and tourism literature) to understand the valence of a particular review, which can range from negative to positive (Geetha, Singha, and Sinha 2017; Phillips et al., 2019). The process of sentiment analysis involves the use of computational linguistics along with natural language processing to extract subjective information (such as emotional inclination) from textual data (Salehan and Kim, 2016). The automated nature of the process makes it an efficient mechanism for processing big data like social media (Cheng, Chiang and Storey, 2012; Pang and Lee, 2008).

One would expect the previously described principle of loss aversion predicts an asymmetric effect of service performance received on rating sentiment. Recall that under the lens of prospect theory, gains are weighed differently than equivalent losses. In the context of online ratings, the principle of loss aversion would specifically suggest that receiving a service that is worse than expected would induce a larger absolute impact on review sentiment than the corresponding absolute effect on review sentiment resulting from receiving a service better than expected. In simpler terms, negative deviations from service expectations weigh more heavily than positive deviations.

These expectations are determined by a state of reference – or a reference point. This reference point "corresponds to the decision maker's current position, (but) it can also be influenced by aspirations, expectations, norms and social comparisons" (Tversky and Kahneman 1991, pp. 1046, 1047). One may therefore think of the sentiment expressed in particular online rating as the satisfaction (or dissatisfaction) resulting from the consumption of the service relative to an expectation about the service.

More generally the principle of loss aversion asserts that the disutility that individuals are subjected to from experiencing a lower than expected service performance eclipses the utility they obtain from experiencing equivalent higher than expected service performance. When investigating online consumer ratings one would thus expect that dissatisfaction relative to the reference point induces a more severe negative review sentiment as compared to the positive sentiment resulting from corresponding levels of satisfaction. When the review sentiments are plotted as a value function, we would consequently expect to observe a higher slope for losses than for gains, suggesting loss aversion. This leads to our first hypothesis:

H.1.- *Stemming from the prospect theory principle of loss aversion, we hypothesize that negative deviations in ratings from the reference point (receiving a service with worse performance than expected) bring about a higher impact on review sentiment than positive deviations (receiving a service with better performance than expected) from the reference point.*

Tversky and Kahneman's (1991) principle of diminishing sensitivity holds that the marginal impact resulting from a gain or loss varies based on distance from the previously discussed reference point. Specifically, diminishing returns suggests gains (losses) result in lower additional levels of satisfaction (dissatisfaction) as one moves away from the reference point. This property has been used across a number of areas of study in economics and psychology- consumer theory's property of diminishing marginal rates of substitution, producer theory's property of diminishing returns, and the intertemporal choice theory of discounting are all explained by the principle of diminishing returns (Hill and Neilson, 2007).

With regards to services, customer perceptions of service quality may be based on the gap between expected service and actual service (Parasuraman et al., 1985). While larger gaps between expected service and actual service would most certainly result in larger absolute changes in consumer perceptions of the service as reflected in review sentiment, diminishing sensitivity would suggest that as this gap increases, the resulting changes in review sentiment occur at a decreasing rate. More specifically, the principle of diminishing sensitivity would imply that when actual service received is only slightly below or above expected service, the marginal impact on review sentiment would be greater than when the service received is substantially below or above expected service. In other words, as the service received differs increasingly from the expectations as defined by the reference point, the marginal change in sentiment decreases. This is true for cases when service received fails to meet expectations, as well as when service received exceeds expectations.

This also implies that a perfectly neutral review sentiment (a sentiment score of 0) might suggest that the service received by the reviewer was exactly equal to the service expected. Accordingly, changes in service received increasingly departs from expectations, the marginal impact on the review's sentiment score falls.

Accordingly, a graphical representation of the value function for the review sentiment relating to a specific service would therefore – in addition to exhibiting a higher slope for losses than for gains – be expected to assume a convex shape for losses and a concave shape for gains. This occurs because the marginal change in sentiment increases at a decreasing rate as service received exceeds service expected, and decreases at a decreasing rate when service received is below expectations. The following therefore serves as our second hypothesis:

H.2.- *Regardless of whether the service received is better or worse than expected, variations in ratings closer to the reference point result in higher marginal impacts on sentiment than variations further away from the reference point.*

An obvious question that arises here has to do with how the reference point – in other words the expected level of service – could be measured. While the prospect theory notion of reference dependence undoubtedly provides rich theoretical insights, the testing of loss aversion and diminishing returns using real data requires the identification of actual reference points in the domain in which consumer behavior is being assessed.

In this case the domain is review sentiment, and presumably the sentiment expressed by a consumer in a review reflects the extent to which his/her service experience differs from his/her expectations as defined by the reference point. As one might expect, the identification of reference points for individual consumers is not a straightforward task when using secondary data. In an experimental setup, one may be able to manipulate reference points by varying the status quo (Hardie et al., 1993). The present study, however, uses secondary data, and experimental manipulations of the reference point are thus not an option. We must rely therefore on some alternative estimation of the service expectation which in this study describes the consumer's reference point.

Two general standards exist in the empirical prospect theory literature to identify reference points (Ref). The first of these pertains to internal memory based standards. This approach involves consumer assessments against a standard established by making use of past information. In the service context, this reference point might then be, for instance, a consumer's past experience with that particular service. A diner's perceived satisfaction during a service experience at a certain restaurant on a particular visit may therefore be assessed against his/her typical experience at that restaurant in the past. A hotel guest's perceived satisfaction during a stay may be assessed against his/her experience during previous stays at the hotel. An airline passenger's satisfaction on a particular flight could be assessed against his/her previous experiences on that route. In the diner example, we could, if the necessary data were available, use a measure that reflects the diner's median experience across multiple visits at that restaurant as a satisfactory reference point. A similar measure could be obtained for the hotel guest and air passenger mentioned above. In the context of online reviews however, there are some limitations in the estimation of reference points that may have been established using this standard. For example, we would not typically have sufficient data to discern the consumer's past experiences with that specific service. While we may have at our disposal a reviewer's review history, this history would reflect his/her consumption experiences across a multitude of non-comparable products and experiences. Indeed, it is rather unlikely that an individual reviewer would have on the same platform multiple reviews of the same restaurant with each review reflecting an independent visit to that restaurant. Similarly, it is quite unlikely that a hotel guest would leave on the same platform a different review for a particular hotel after each visit, or that an airline passenger would leave a different review for a flight experience each time he/she flies a particular route. Even if some reviewers did provide multiple reviews for the same restaurant, hotel or flight route, the number of reviewers doing so would not be sufficient to make meaningful inferences.

The second standard to estimate consumer reference point overcomes this limitation as it does not rely on the consumer's own experience with that particular service over time. This standard recognizes that reference points tend also to be established externally. The standard for comparison in the aforementioned diner example could for instance be the distribution of service experienced by *other* diners. Certainly, online reviews themselves provide a reasonably strong estimate of the service experienced by others, and therefore provide a reasonably strong estimate of the reference point. Consequently, we believe that the median service experienced by others – as reflected in the median rating of a particular service – would serve as a satisfactory reference point for purposes of this study.

3. RESEARCH DESIGN

3.1. Methodology

In order to test loss aversion and diminishing sensitivity in the context of sentiment analysis, we stem from the basic tenets proposed by Kahneman and Tversky (1979) and incorporate them into a regression model. In particular, Kahneman and Tversky's (1979) value function $v(x)$ is described in terms of gains and losses (so a reference point is needed to capture the differences between the actual value and the expected value), is steeper for losses than for gains [$v(x)<-v(-x)$, $x>0$] bringing about *loss aversion*, and has a S-shape curve (concave for gains [$v''(x)<0$, $x>0$] and convex for losses [$v''(x)>0$, $x<0$]) resulting in *diminishing sensitivity*. Therefore, the proposed model is:

$$Sent_i = \alpha + \beta \cdot Gain_i + \gamma \cdot Loss_i + \theta \cdot Gain_i^2 + \varphi \cdot Loss_i^2 + \sum_{j=1}^{J} \delta_j \cdot CV_{ij} + \varepsilon_i$$

where $Sent_i$ is the review sentiment ranging from -1 to +1, $Gain_i$ is defined as ($Actual\ Rating_i$ - $Expected\ Rating_i$) D_1, where D_1=1 if ($Actual\ Rating_i$ - $Expected\ Rating_i$)>0 and D_1=0 otherwise; $Loss_i$ is defined as ($Actual\ Rating_i$ - $Expected\ Rating_i$) D_2, where D_2=1 if ($Actual\ Rating_i$ - $Expected\ Rating_i$)<0 and D_2=0 otherwise; CV_{ij} are a set of J control variables related to the reviewer, the service and the route; and ε_i is a random term. Finally, α, β, γ, θ, φ and δ_j are coefficients to be estimated. Loss aversion will be detected if the loss parameter is higher than the gain parameter ($\gamma_i/\beta_i>1$) and diminishing sensitivity will be evidenced if the square of the gain variable has a negative and significant parameter (θ) and the square of the loss variable has a positive and significant parameter (φ). The parameters δ_j are associated with the *j-th* control variable.

3.2. Sample and variables

A sample of 157,036 airline reviews was retrieved from Tripadvisor by looking at 20 US airlines. After checking for missing values, we are left with a final sample of 132,486 observations. The dependent and independent variables used in this study are defined as follows:

Dependent variable.

In order to obtain review sentiment, we calculated sentiment scores for each review in the dataset. A sentiment analysis as part of opinon mining was applied to uncover opinions and to assess contextual polarity of online consumers within a given text (Alaei, Becken and Stantic, 2019). In order to analyze the data we used the VADER (Valence Aware Dictionary and Sentiment Reasoner) package in Python. VADER employs a lexicon and rule-based sentiment analysis. A lexicon typically refers to a lit of lexical features like words which are labelled based on semantic orientation (Hutto and Gilbert, 2014; Liu, 2010). VADER considers both intensity as well as polarity of emoption. Hutto and Gilbert (2014) show that VADER outperforms other tools/algorithms relating to sentiment analysis with regards to accuracy of text classificiation in social media data. VADER employs a dictionary to associate the lexical characteristics of the given text to intensity of emotion. Five heuristics are considered to assess how contextual elements effect the text which is being analyzed – puntuation, capitalization, degree modifiers, the impact of contrasting conjunctions like "but", and the examination of the tri-gram before a sentiment-laden lexical feature . The calculated scores from VADER show normalized lexicon ratings between -1 (extremtly negative) and +1 (extremely positive).

Independent variables

The central independent variables are "rating" which is defined as the the overal rating of the specific flight (airline and route) the reviewer used, measured on a scale from 1 to 5 and "expected rating" defined as the median of the overall rating for an airline and route. Note that, as indicated previously when discussing the types of reference points, getting proper reference points is a methodological challenge; to refine as much as possible the estimation of reference points we attempt to reflect the company and the specific product reviewed. Accordingly, for this empirical application, not only do we consider the median value of the airline but also we control for the route reviewed. We use this value as the reference point to which the reviewers compare the service they receive.

We include in the model other variables that are used as control variables: i) Reviewer's level, which refers to the extent to which the specific reviewer shares his/her experiences in TripAdvisor in general (for example, the more reviews/images a reviewer write and post in Tripadvisor, the higher level scores this reviewer obtains); ii) Review count reflects the total number of reviews written by a specific reviewer; iii) Helpful count: it is the total number of helpful votes the reviewer has received divided by the total number of reviews written; iv) Experience shows the period of time the individual has been reviewing for Tripadvisor; v) Visited cities count indicates the number of cities the reviewer has visited; vi) Photos which reflects the number of photos the reviewer has posted; vii) Distribution of ratings: within the total contributions, it shows the proportion of ratings the reviewer has classified as "Excellent", "Very good", "Average", "Poor" and "Terrible"; viii) Value for Money, which is measured on a scale from 1 to 5; ix) Domestic flight that represents the type of flight, domestic vs International; and x) Economy class, which is a variable that indicates if the reviewer flew in economy class. Table 1 shows the descriptive statistics of these variables.

4. RESULTS

Prior to estimating the model, collinearity and heteroskesdaticity are tested. Accordingly, we find, for collinearity, that all the Variance Inflation Factors are below the recommended value of 10 (Hair et al., 2006; Neter et al., 1989), and for heteroskesdaticity, that the Breusch-Pagan confirms its existence ($F=651.3$; $p<0.001$), hence the White heteroscedasticity-consistent standard errors are utilized.

Regarding the parameters of interest, in Table 2 we observe that the four key variables (*gain*, *loss*, $gain^2$ and $loss^2$) are significant. The loss parameter is significantly greater than the gain parameter (Wald test=55.78; $p<0.001$), which supports the idea that travelers react more strongly to dissatisfactions (finding a worse service than expected) than to satisfactions (finding a better service than expected), which represents evidence in favor of loss aversion supporting Hypothesis 1. In other words, receiving a service of a lower than expected quality brings about a stronger reaction in the sentiment variable than getting a service of a better than expected quality; i.e. dissatisfaction relative to the reference point induces a more severe negative review sentiment as compared to the positive sentiment resulting from corresponding levels of satisfaction.. Figure 1 shows the different slopes for the loss and gain regions.

As for the quadratic terms, both are significant, negative for the gain parameter and positive for the loss parameter, resulting in a concave line for gains and convex curve for losses (see Figure 2), in line with the diminishing sensitivity property of prospect theory. This also supports Hypothesis 2 that, regardless of whether the service received is better or worse than expected, variations in ratings closer to the reference point result in higher marginal impacts on sentiment than equivalent variations further away from the reference point. Certainly, larger gaps between expected service and actual service bring about larger absolute changes in consumer perceptions of the service as reflected in review sentiment, and with the diminishing sensitivity property we find that as this gap increases, the changes in review sentiment manifest themselves at a decreasing rate.

Most of the control variables show significant effects: the level that the reviewer has attained in Tripadvisor has a significant effect, in fact, the higher the level the greater the impact on sentiment. The number of reviews posted, the proportion of times the reviewer has described the service as excellent or very good, and value for money have positive effects on sentiment.

The period of time (experience) the individual has been reviewing for Tripadvisor, the proportion of times the reviewer has described the service as average, poor or terrible, the domestic character of the flight and the economy-type seat present negative impacts.

Finally, the number of times these reviews have been helpful, the number of cities visited and the number of posted photos do not seem to have any effects on sentiment.

5. CONCLUSIONS

This article analyzes the relationship between ratings and sentiment by introducing the tenets of prospect Theory. Specifically, we test loss aversion and diminishing sensitivity on a sample of 132,486 airlines reviews and find first that negative deviations in ratings (receiving a service with worse performance than expected) bring about a higher impact on sentiment than positive deviations (receiving a service with better performance than expected), thus confirming loss aversion.. Second, regardless of whether the service received is better or worse than expected, variations in ratings closer to the reference point result in higher marginal impacts on sentiment than variations further away from the reference point, thus proving diminishing sensitivity.

Other variables also have a positive impact on sentiment: the reviewer's level attained in Tripadvisor, number of reviews posted, proportion of times the reviewer has described the service as excellent or very good, and value for money. Other variables that have a negative effect are individual's experience reviewing for Tripadvisor, proportion of times the reviewer has described the service as average, poor or terrible, the domestic character of the flight and the economy-type seat.

The results obtained have relevant theoretical and managerial implications. Regarding the theoretical implications, it is important to note the following ones: 1) The use of prospect theory provides further insights in sentiment analyses. While some previous studies have looked at the effects of reviews on some outcomes (such as purchase intention or sales ranks), none of them have used sentiment measures and tested prospect theory in a tourism context. The fact that prospect theory has been confirmed in the relationship between sentiment and reviews implies that, in order to have a comprehensive view of the effect of reviews on sentiment measures, the principles of prospect theory should be considered; otherwise, relevant knowledge may be omitted. 2) The analysis of the effects of reviews on sentiment should include relative measures rather than just absolute metrics. For research to identify potential asymmetries, the studies should use reference points; while using absolute values can give an indication about the influence of reviews on sentiment, we have shown that *loss aversion* and *diminishing sensitivity* exist, and both properties provide richer information about this influence and better reflect the way people make their assessments and provide their rating values. 3) The effect of reviews on sentiment is not free of cognitive bias. Beyond the fact that reviewers have limited cognitive capacity (as expected), the main consequence is that they cannot objectively write reviews that are bias-free. Thus, the significant parameters found in this study regarding loss aversion (*gain* and *loss*) and diminishing sensitivity ($gain^2$ and $loss^2$) prove that these are cognitive biases that should be taken into account when analyzing the valence of ratings and effect of reviews. This consideration should take place either to explicitly include the relevant variables (reference values) to control for this cognitive bias or to recognize that spurious estimates can be obtained if these benchmark values are not included.

Concerning the managerial implications, the confirmation of loss aversion and diminishing sensitivity properties has critical implications for decision-makers.

First, when the service received fails to meet expectations, loss aversion implies that the negative impact on sentiment is greater that the positive effect of an increase in ratings of the same amount. For example, a reduction in the overall ratings from an average value of, say, 4 to an actual value of 3 can be perceived as a reduction in the quality of service which has an effect on review sentiment. If managers try to solve this situation and decide to implement strategies to increase the ratings back to 4, the increase from 3 to 4 will not bring about the same size of variation (with a different sign) in sentiment as the aforementioned reduction. The variation in ratings from 3 to 4 will cause a lower positive effect on sentiment than the negative effect derived from the variation from 4 to 3. Consequently, in practical terms, the "efforts" to get the rating back to the previous upper levels necessarily have to be greater than the “inattention” that caused the reduction in the rating. Note that, even though our study uses an overall rating to test prospect theory, if the rating values of individual attributes were available, the analysis can be easily extrapolated so that the specific attributes with lower-than-expected quality can be detected. Accordingly, measures to change this low-quality level could be implemented and, in line with the results obtained and the aforementioned suggestion, the efforts to increase this quality should be more intense than the inattention that led to its reduction.

Second, when the service received exceeds expectations, loss aversion means that while the potential above-the-standard point is beneficial to the customers (positive sentiment), this excellence should be maintained at those levels; otherwise, if a reduction in service performance is observed, the negative effect on sentiment will be drastic compared to the initial increase. If the expected rating is 3 and the customer’s rating is 4, there will be a positive impact on sentiment; however, for the next consumption occasion the new expected rating will be 4, so if the actual rating for this future consumption occasion is 3, then the impact on sentiment will be more negative than the aforementioned positive impact. Let us suppose an airline is implementing a new speedy boarding system so that passengers get on the plane in a more efficient and quick way, reducing the boarding time by a certain number of minutes. While this will be perceived as an increase in the quality of the service, the airline should regard the new boarding times as long-term values so that the necessary arrangements need to be organized to maintain these times. Instead of looking at this action as a short-term tactic to entice customers, a more strategic long-term view is required because the new boarding times are the new reference values the passengers will set in their mind and will remember. Although the time reduction brings about an increment in satisfaction, the dissatisfaction caused by a subsequent time increase (even if occasional) will be higher than that increment in satisfaction.

Third, because of the diminishing sensitivity found, when a firm is enjoying a very good reputation materialized by high rating values, it seems to be more protected as a “cushion” seems to exist. Remember that if the service received is better than expected, variations in ratings further away from the reference point result in lower marginal effects than equivalent variations closer to the reference point. In practical terms, it means that the negative effect of a not-so-good experience on sentiment will be lower if the rating changes from 4 to 3 than if the rating shifts from 2 to 1. Consequently, it is obvious that when facing a service failure, high- and low-rated companies must try their best to implement service recovery strategies; nevertheless, while high-rated companies may have some leeway (assuming it is not anything major), low-rated companies must work harder to solve the issue and better compensate its customers because the negative influence of a bad experience will reduce further the sentiment measure of these low-rated companies than that of high-rated firms.

Regarding limitations, if customers look at reviews, the publicly available average ratings are most likely to determine the customer's reference points; however, if reviewers remember the values with which they rated the service in the last consumption occasion, these values could form these reviewers' reference points. In other words, while in this study we rely on external reference points (i.e. published average values of ratings), it could be interesting to see, as future research, whether internal memory-based reference points (e.g. last value with which the reviewer rated the service) offer similar results.

Also as future avenue for research stands out the effect of ratings of specific items on review sentiment. We have looked at the overall rating, thus it could be interesting to see whether the ratings of individual items exert an effect and which ones are more determinant. Finally, even though the dataset is large, cross-validation is still needed so that different platforms and different industries would help reinforce the results obtained.

REFERENCES

Alaei, A. R., Becken, S., & Stantic, B. (2019). Sentiment analysis in tourism: capitalizing on big data. Journal of Travel Research, 58(2), 175-191.

Askalidis, G., Kim, S. J., & Malthouse, E. C. (2017). Understanding and overcoming biases in online review systems. Decision Support Systems, 97, 23-30.

Basuroy, S., Chatterjee, S., & Ravid, S. A. (2003). How critical are critical reviews? The box office effects of film critics, star power, and budgets. Journal of Marketing, 67(4), 103-117.

Chen, Hsinchun, Roger HL Chiang, and Veda C. Storey. (2012). "Business intelligence and analytics: From big data to big impact." MIS Quarterly 36, 4.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. Journal of Marketing Research, 43(3), 345-354.

Fotis, J. N., Buhalis, D., & Rossides, N. (2012). Social media use and impact during the holiday travel planning process (pp. 13-24). Springer-Verlag.Geetha, M., Singha, P., & Sinha, S. (2017). Relationship between customer sentiment and online customer ratings for hotels-An empirical analysis. Tourism Management, 61, 43-54.

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

Grimes, M. (2012). “Global Consumers’ Trust in ‘Earned’ Advertising Grows in Importance,”Nielsen, April 10, https://www.nielsen.com/us/en/press-releases/2012/nielsen-global-consumers-trust-in-earned-advertising-grows/

Hardie, B. G., Johnson, E. J., & Fader, P. S. (1993). Modeling loss aversion and reference dependence effects on brand choice. Marketing Science, 12(4), 378-394.

Hernandez-Maskivker, G. M., Nicolau, J. L., Ryan, G., & Valverde, M. (2019). A reference-dependent approach to WTP for priority. Tourism Management, 71, 165-172.

Hill, S. A., & Neilson, W. (2007). Inequality aversion and diminishing sensitivity. *Journal of Economic Psychology*, 28(2), 143-153.

Hutto, C.J. & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment

analysis of social media text . ICWSM proceedings. Palo Alto, CA: Association for the Advancement of Artificial Intelligence (AAAI).

Kim, J. H. (2014). The antecedents of memorable tourism experiences: The development of a scale to measure the destination attributes associated with memorable experiences. Tourism Management, 44, 34-45.

Kim, J. Y., & Canina, L. (2015). An analysis of smart tourism system satisfaction scores: The role of priced versus average quality. Computers in Human Behavior, 50, 610–617.

Liu, B. (2010). Sentiment analysis and subjectivity. Handbook of natural language processing, 2(2010), 627-666.

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

Mazursky, D. (1989). Past experience and future tourism decisions. Annals of Tourism Research, 16(3), 333-344.

Masiero, L., & Qiu, R. T. (2018). Modeling reference experience in destination choice. *Annals of Tourism Research*, 72, 58-74.

Mellinas, J. P., Nicolau, J. L., & Park, S. (2019). Inconsistent behavior in online consumer reviews: The effects of hotel attribute ratings on location. *Tourism Management*, 71, 421-427.

Mintel. (2013). Holiday planning and booking process – UK. Mintel Report, November.

Nicolau, J. L. (2008). Testing reference dependence, loss aversion and diminishing sensitivity in Spanish tourism. *Investigaciones Económicas*, 32(2), 231-255.

Nicolau, J.L. (2011) "Differentiated price loss aversion in destination choice: The effect of tourists' cultural interest", *Tourism Management*, 32, 5, 1186-1195.

Nguyen, Q. (2016). Linking loss aversion and present bias with overspending behavior of tourists: Insights from a lab-in-the-field experiment. *Tourism Management*, 54, 152-159.

Oh, H. (2003). Price fairness and its asymmetric effects on overall price, quality, and value judgments: the case of an upscale hotel. Tourism management, 24(4), 387-399.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1-135.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. Journal of marketing, 49(4), 41-50.

Park, S., & Nicolau, J. L. (2015). Asymmetric effects of online consumer reviews. *Annals of Tourism Research*, 50, 67–83.

Park, S., & Nicolau, J. L. (2019). Image effect on customer-centric measures of performance. *Annals of Tourism Research*, 76, 226-238.

Phillips, P., Antonio, N., de Almeida, A., & Nunes, L. (2019). The Influence of Geographic and Psychic Distance on Online Hotel Ratings. *Journal of Travel Research*, https://doi.org/10.1177/0047287519858400

Salehan, M., & Kim, D. J. (2016). Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. *Decision Support Systems*, 81, 30-40.

Simon, H. A. (1955). A behavioural model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99-118.

Smeral, E. (2017). Tourism forecasting performance considering the instability of demand elasticities. *Journal of Travel Research*, 56(7), 913-926.

Thaler, R. (1980). Toward a positive theory of consumer choice. Journal of Economic Behavior & Organization, 1(1), 39-60.

TripAdvisor (2019), Online Reviews Remain a Trusted Source of Information When Booking Trips, Reveals New Research, available at: https://www.tripadvisor.com/pdfs/OnlineTravelReviewReport.pdf, (accessed September 29th, 2019)

Tsang, A. S., & Prendergast, G. (2009). Is a "star" worth a thousand words? The interplay between product-review texts and rating valences. European Journal of Marketing, 43(11/12), 1269-1280.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106(4), 1039-1061.

Viglia, G., Mauri, A., & Carricano, M. (2016). The exploration of hotel reference prices under dynamic pricing scenarios and different forms of competition. *International Journal of Hospitality Management*, 52, 46-55.

Wattanacharoensil, W., & La-ornual, D. (2019). A systematic review of cognitive biases in tourist decisions. Tourism Management, 75, 353-369.

Zhang, H., Yang, Y., Zheng, C., & Zhang, J. (2016). Too dark to revisit? The role of past experiences and intrapersonal constraints. *Tourism Management*, 54, 452-464.

Table 1. Descriptive statistics

Variable	Mean/Proportion	Std. Error
Overall rating	3.73	1.31
Expected rating	3.88	1.04
Gain	0.25	0.54
Loss	-0.44	0.84
Level0	7.79%	-
Level1	6.5%	-
Level2	10.54%	-
Level3	17.8%	-
Level4	15.31%	-
Level5	15.47%	-
Level6	26.59%	-
Review count	102.2	274.9
Helpful count	40.07	165.2
Experience	5.50	3.44
Visited cities count	90.3	209.7
Photos	154.3	1890
Excellent	46.06%	-
Very good	25.39%	-
Average	9.82%	-
Poor	3.51%	-
Terrible	2.73%	-
Value for money	3.61	1.30
Domestic flight	69.15%	-
Economy class	85.14%	-

Table 2. Effect of loss aversion and diminishing sensitivity on sentiment

Variable	Coefficient	Std. Error
Gain	0.071[a]	0.005
Loss	0.136a	0.006
$Gain^2$	-0.021[a]	0.002
$Loss^2$	0.009[a]	0.002
Level2	0.037[a]	0.010
Level3	0.040[a]	0.010
Level4	0.042[a]	0.010
Level5	0.049[a]	0.010
Level6	0.062[a]	0.010
Review count	2E-05[a]	6E-06
Helpful count	0.001	0.004
Experience	-0.002[a]	4E-04
Visited cities count	1E-05	9E-06
Photos	1E-06	9E-07
Excellent	0.031[a]	0.010
Very good	0.179[a]	0.012
Average	-0.082[a]	0.021
Poor	-0.300[a]	0.033
Terrible	-0.257[a]	0.031
Value for money	0.211[a]	0.002
Domestic flight	-0.017[a]	0.003
Economy class	-0.085[a]	0.004
Constant	-0.279[a]	0.009
R-squared	0.3126	
Adjusted R-squared	0.3125	
F-statistic	2739.6[a]	

[a]=$p<0.001$

Figure 1. Loss aversion in sentiment analysis.

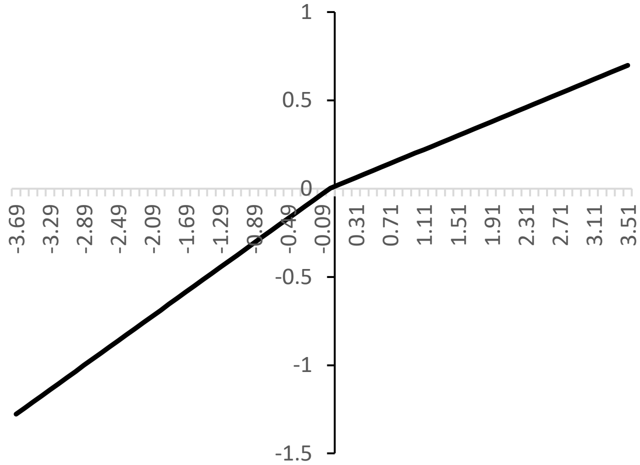

Figure 2. Loss aversion and diminishing sensitivity in sentiment analysis.