ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

And suddenly, the rain! How surprises shape experienced utility

Paolo Figini
Veronica Leoni
Laura Vici

Quaderni - Working Paper DSE N°1185

And suddenly, the rain! How surprises shape experienced utility*

Paolo Figini[1,2], Veronica Leoni[†2,3], and Laura Vici[2,4]

[1] Department of Economics, University of Bologna, Italy
[2] Centre for Advanced Studies in Tourism, University of Bologna, Italy
[3]Department of Applied Economics, University of the Balearic Islands, Spain
[4]Department of Statistical Sciences 'Paolo Fortunati' University of Bologna, Italy

April 21, 2023

Abstract

The utility associated with a service's consumption is contingent on its intrinsic characteristics and various situational factors. One key element that influences consumer satisfaction is adherence to prior expectations. This is particularly relevant for experience goods that highly depend on external factors, such as weather. On these premises, the current study explores the role of expectations on utility by analyzing the effect of weather surprises (*i.e.*, the mismatch between forecast and realized weather) on online ratings. Results from the analysis of over 300,000 reviews posted on Booking.com indicate that weather surprises have an impact on the reported experienced utility, the effect depending on the sign of the surprise. Moreover, the consumption span moderates the surprise effect, thereby mitigating the impact of both positive and negative surprises on utility.

Keywords: weather bias; surprise effect; online reputation; experienced utility; online platforms

JEL codes: D83; D91; L81; Q54.

*Declarations of interest: none.

†Corresponding author: veronica.leoni3@unibo.it
Center for Advanced Studies in Tourism, University of Bologna. Via Angherà, 22 (47921) Rimini (Italy).
Declarations of interest: none.

Non-technical Summary

When consumers buy a product or a service, their satisfaction depends on the intrinsic quality, also compared to previous expectations, and by contingent factors. This is especially true for the so-called experience goods, in which quality can be ascertained only after consumption. In most cases, the overall experience is significantly affected by external factors that the producer does not control; consider how the weather or the climate conditions affect the activities to be carried out while on holiday, and hence its enjoyment. In the last decade, online platforms for consumers to leave feedback about specific services, like hotel accommodations or restaurant meals, have become increasingly popular. People can now report their spontaneous opinion about specific services, which can serve prospective buyers when shortlisting providers based on the overall reported level of satisfaction. Therefore, reviews might have a non-negligible effect on future buyers' behaviour and hence on the economic performance of the service providers. External factors can play a role when evaluating services, despite not being an inherent part of the service to be rated or under the service provider's control.

Given these considerations, we were interested in exploring the impact of an unexpected event on people's reported level of satisfaction. Specifically, we wanted to investigate whether and to what extent an unexpected external event, a surprise, is reflected in service evaluations and captured by online ratings. To do this, we looked at ratings left by customers who had stayed in hotels in two important Italian cities: Milan and Venice. The surprise here analyzed is the discrepancy between the expected weather (built upon the weather forecasts publicly available at the time of travelling) and actual weather conditions during the stay.

We find that an inaccurate forecast, which translates into weather surprise, significantly impacts the reported satisfaction with the accommodation service. More specifically, if the weather was better than expected (a positive surprise), people were happier with their experience, and vice versa. Importantly, and differently from what was done in previous research, what drives the evaluation is not the weather condition per se but the surprise generated by any mismatch between existing expectations (the weather forecast) and the realized weather conditions while at the destination. Our findings are of great interest since they show that online reported evaluation might be misleading, capturing factors that are external to the quality of the service itself and that attain the psychological dimension of surprisingness. Implications for service providers and consumers are discussed.

1 Introduction

People go to the cinema to watch a movie because they have high expectations, triggered, for example, by positive critics' reviews, a star-sudden cast, the movie's genre, the trailer, the director, or word-of-mouth recommendations. Consumers' movie experience is then determined by the quality of the film's intrinsic elements, with the onset of specific events sometimes playing a significant role. Unexpected plot twists or endings can negatively impact the audience's perception of the story or characters. Conversely, surprise cameos can enhance the movie's enjoyment by adding tension and suspense to the narrative, keeping the audience engaged and guessing (Ely et al., 2015). Sometimes, the quality of the experience depends on characteristics beyond the control of the producers, such as the characteristics of the theatre (the screen size, the sound system, and seating) or the behavior of the audience in the auditorium, including noise and overcrowding. As a result, the experienced utility, according to Kahneman's (1991) definition, is influenced not only by the intrinsic quality of the movie but also by unexpected circumstances surrounding its viewing and other external factors.

When it comes to purchasing experience goods, agents' decision process is much more complex as the quality of such products can be assessed only during or after consumption (Nelson, 1970). Some commodities are cheap, and their purchase does not entail a burdensome decision process; by contrast, buying other, more expensive commodities involves carefully acquiring information to form more reliable prior expectations. However, even when information is available and non-costly, the risk of receiving lower utility than expected remains and, in part, is due to the occurrence of events that were not taken into account when making decisions based on expected values and probabilities.

In the literature, different methods and measures have been used to assess consumers' experienced utility but analyzing user-generated content, such as online reviews and ratings, is gaining momentum in academic research (Agarwal et al., 2021; Helmers et al., 2019; Greiff and Paetzel, 2020; Reimers and Waldfogel, 2021; Mayzlin et al., 2014). Thanks to the steady progress of online platforms and social media, it is extremely common for subjects to offer first-hand and spontaneous feedback on their consumption experience and their degree of satisfaction with goods and services. Despite the significance of expectations in shaping the reported experienced utility, the literature has yet to explore this topic fully. Most studies in this area have primarily focused on investigating the expected quality of the goods or services to be consumed and eventually reviewed. It has been shown that subjects can be affected by social bias when writing reviews, *i.e.*, they tend to conform to the norm, embodied by prior ratings (Chevalier and Mayzlin, 2006; Cicognani et al., 2021; DeMarzo et al., 2003; Manski, 1993).

In our paper, we dig deeper into the uncharted territory of consumer behavior by delving into the under-explored yet critical role of surprises when reporting utility. More specifically, we measure how surprises, namely discrepancies between expectations and experienced consumption, affect the level of utility in the case of experience goods. The idea is that the expected effects of externalities are already incorporated in prior beliefs (*e.g.*, crowded cinemas during winter weekends), hence the reported utility derived by consuming experience goods may be affected by surprises (occurring when current beliefs are further from previous beliefs, Ely et al. (2015)).

A classic example of an experience good that aligns with the description provided above is found in the travel industry. In this study, we will focus on accommodation services, which quality relies on factors such as location, the season, and the level of the services provided. It is crucial to note that the perceived quality of such services can also be impacted by an individual's current state, emphasizing the state-dependent nature of these goods. This makes accommodation services an excellent case study for our research objectives.

In terms of the specific type of surprise here considered, we focus on the unexpected change in weather conditions compared to publicly available forecasts (already pondered in prior beliefs). Therefore, the current work aims to explore if (and to which extent) the mismatch between expectation and realization of an exogenous factor (*i.e.*, weather surprises) is reflected in customers' evaluation of the service. Conditional to the time of traveling (seasonality), subjects form their beliefs about the type of activities to carry out when on holiday or decide whether to travel depending on the weather forecasts. On these premises, surprises (*i.e.*, sudden unexpected changes in weather), can potentially impact the overall experienced utility and allegedly be captured by online ratings, despite not being an inherent part of the service to be reviewed or under the control of the service provider.

To this end, we analyze more than 300,000 online reviews for accommodation providers listed on Booking.com, the leading search engine for lodging services, in two popular Italian destinations (Milan and Venice), between September 2019 and February 2020. The period under investigation is chosen for consistency reasons, allowing us to analyze only ratings before the COVID-19 outbreak and after the reform of the reviewing system introduced by the platform in September 2019 (Leoni and Boto-García, 2023). With regard to weather surprises, we look at the potential mismatch between the weather forecasts posted days before traveling and the realised weather. The empirical strategy aims to disclose this "surprise" effect on individual ratings, conditional on the reviewers' characteristics (nationality, travel party, length of stay, and review anonymity), hotels' time-invariant features, prices, and external factors such as seasonality.

Our findings provide conclusive evidence of a statistically significant impact of surprises on experienced utility. Specifically, we observe that better-than-expected weather conditions have a positive effect on the final score, while worse-than-expected weather conditions are found to significantly decrease satisfaction levels with hotel services. Moreover, we uncover an interesting moderating role played by the consumption span, which mitigates both the positive and negative effects of weather mismatch, in line with the hedonic adaptation theory (Frederick and Loewenstein, 1999). Results are robust to several specifications and, differently from most of the existing literature on reported utility, we control for prices, hence taking into account the critical role of *value for money* when judging services.

Within this framework, the contribution of this paper lies in the empirical analysis of if and how the differential between prior belief and realized situation (surprises) may affect the quality assessment of an experience good. We investigate how consumer perception and assessment of goods quality is biased by external factors out of suppliers' action spheres, concerning the contingent state where goods are consumed. We draw some managerial and policy implications to lessen the impact of such external effects. To the best of the authors' knowledge, no prior research has explored the effect of exogenous surprises on agents' utility by investigating online user-generated content. Moreover, we contribute to analyzing if positive and negative surprises produce any asymmetric effect on consumers' utility.

The remainder of the paper is structured as follows: the relevant streams of literature, focusing on the role played by expectations and surprise, are recalled in Section 2, while the theoretical framework is presented in Section 3. We then describe the data (Section 4) and the empirical strategy (Section 5). Finally, the results and the concluding remarks are presented in Sections 6 and 7, respectively.

2 Literature Review

Our work is related to the stream of literature investigating the role of expectations in economics, especially the one studying how individual utility is affected by a mismatch between the expected and the realized state of the world, something that can be called a *surprise*.

2.1 Surprises in Economics

Surprises are directly related to expectations and uncertainty about future outcomes. By definition, surprises emerge as a conflict with pre-existing beliefs (Baccan et al., 2015). From a psychological perspective, according to Schützwohl (1998), surprises result from schema-discrepancy, *i.e.*, an inconsistency between activated schemas and newly acquired information. When such discrepancy exceeds a certain threshold, people experiment with the feeling of surprise, which undergoes a certain degree of subjectivity because people might display different thresholds. In economics and financial studies, the surprise can be defined as the mismatch between expected and actual results, for instance, when the rate of return of investments is considered.
Economic agents are subject to a certain degree of uncertainty when making decisions; as a result, their economic behaviours inevitably involve expectations. Especially in consumption, satisfaction and enjoyment are much related to the adherence of outcomes to such priors. However, a mismatch does not compulsory imply a negative connotation. The literature discusses how the feeling of surprise is linked to an emotion that can be either positive or negative (Janakiraman et al., 2006). The "sign" of the surprise depends on each agent's consequence of the unexpected event or outcome. As Derbaix and Vanhamme (2003) discussed, an event's surprisingness amplifies the reaction. In other words, with equal events, unexpected ones will produce stronger responses.

Most of the research dealing with the effect of surprises on people's utility and enjoyment is focused on entertainment markets, such as sports games (Bryant et al., 1994; Ely et al., 2015; Peterson and Raney, 2008) or online media products (Simonov et al., 2022). Extant literature underlines the importance of suspense (*i.e.*, excitement or anxiety about what may happen in the future) and surprise dynamics in the context of entertainment. However, it is hard to empirically assess their role because beliefs and enjoyment are often not directly observable (Ely et al., 2015). In this regard, Bizzozero et al. (2016) study preferences using the size of the tennis audience (in other words, the number of people watching a tennis match) as a proxy for enjoyment. As a matter of fact, bored viewers can easily (at almost no cost) decide to switch the channel or turn the TV off. While in these specific markets' enjoyment is expressed in terms of audience, when dealing with the consumption of other services, it can be assessed via consumers' explicit evaluations in the form of online reviews (Fradkin et al., 2018). In an experimental setting, Derbaix and Vanhamme (2003) studied the effect of elicited surprises on the willingness to share information (word-of-mouth). No prior research has instead explored the effect of exogenous surprises on agents' enjoyment by investigating online user-generated content.

2.2 Exogenous shocks and individual utility

While most of the economic attention is towards the impact of an exogenous shock on the macro economy, the effect of shocks on agents' utility is a relatively unexplored topic in microeconomics literature; this fact most probably results from standard economic theory, which uses probability theory to treat expectations and uncertainty objectively. As Oswald and Powdthavee (2008) discussed, these authors build upon the hedonic adaptation theory (Frederick and Loewenstein, 1999) to assess the effect of onset disability on people's well-being. This term refers to people's reaction (and adaptation) to upcoming favorable or adverse events, which, within our context, could be defined as exogenous shocks. According to psychology literature, shocks' effect on people's utility is not permanent, *i.e.*, reactions fade over time because people tend to adapt to changes. This is what Wilson and Gilbert (2008) define as affective adaptation. Rayo and Becker (2007) equate hedonic utility to happiness and look at its volatility over time. In this regard, they consider that happiness is not constant and depends on upcoming changes, prior expectations about the future level of happiness, and peers' happiness. In a similar vein, Kettlewell et al. (2020) analyze the time path of the well-being of females affected by an exogenous shock, namely the spouse's death. Riis et al. (2005) analyze the reported level of happiness of healthy people and people with serious illnesses, finding no evidence of a significant difference. All these studies align with the theory of hedonic adaptation, showing that, after an external shock (which might be of different nature), utility tends to converge to the pre-event level. Moreover, the empirical evidence on hedonic adaptation holds also for other types of shocks, more related to individuals' disposable income. In this regard, Easterlin (1995) finds evidence that an income increase does not lead to long-run higher well-being. This is explained in terms of the hedonic treadmill: people's happiness does not increase with money since people tend to adapt to the new income level quickly.

A classic example of a purely exogenous shock is the one stemming from unexpected weather-related events. Weather affects everyday life, influencing and limiting the range of activities that can be performed. Weather conditions might influence people's choice between work and leisure (Connolly, 2008) and their overall productivity (Lee et al., 2014). In an intertemporal choice between leisure and work, bad weather could decrease the opportunity cost of working (reducing distractions resulting from good weather), hence bringing more work hours. In a similar fashion, Huysmans (2002) looks at the weather as a determinant of human sleeping, sports participation, leisure, and other recreational activities. The weather might also moderate the participants' responses to financial events (Dehaan et al., 2017).

Scaling down to the context of the current study, the weather undoubtedly has remarkable effects on travel choices and enjoyment of related activities. This relationship exists for more practical reasons; for instance, people tend not to go to the beach when raining or visit cities in the middle of a heatwave. Moreover, from a more psychological perspective, a stream of literature links the weather with mood, thinking, and judgments (Klimstra et al., 2011). In a recent paper, Brandes and Dover (2022) look at the effect of unpleasant weather (hence assuming preferences for sunny weather) on evaluating accommodation services. More specifically, they consider the weather in the origin country (which does not coincide with the destination country) and find that rain is associated with a higher review provision and lower scores. This result is most likely due to the lower opportunity cost of time when raining and the bad mood associated with bad weather.

While most of the literature has focused on the effect of weather *per se* on the enjoyment of leisure activities (Gössling et al., 2016; Jeuring, 2017; Jeuring and Peters, 2013), less is known about the effect of weather surprises on agents' utility. In other words, existing studies do not consider agents' expectations about the weather conditions when traveling. As Figini et al. (2022) discusses, weather forecasts are public information that is very relevant for travel decisions and firms' pricing strategies. Moreover, since such information flow comes at negligible costs, it is not subject to potential rational inattention (Sims, 2006). On these premises, and in line with existing studies (Derbaix and Vanhamme, 2003), it is reasonable to think that the effect of weather status on agents' utility depends on the degree of the unexpectedness of such status, *i.e.*, on the mismatch between prior information (for instance, derived from forecasts) and the realized weather. This is precisely where our work is positioned.

3 Theoretical Framework

The theoretical framework is adapted from Ely et al. (2015) where the authors model the optimal way for principals to release information to attract a greater audience, assuming that agents have a preference for surprise and/or suspense dynamics, for instance, when reading novels or watching sports matches. Consistently, we assume that visitors (the type of agents herein under investigation) hold a belief about the weather that will occur during the trip, and that will affect the type of activities to be carried out. In fact, according to the type of city visited, rainy or sunny weather might allow or prevent visitors from undertaking specific activities. Consequently, any mismatch between the expected and the realized weather while at the destination disrupts anticipated activities and is likely to affect the agents' utility and the enjoyment associated with the stay. We posit that people's expectations follow an adaptive learning process, *i.e.*, agents adjust and revise their expectations as new data get released over time (Bray and Savin, 1986). In our setting, agents adapt their expectations based on the weather forecasts published during the last days preceding traveling. It is important to highlight that weather forecasts are public information, hence available to all agents at virtually no cost.

Differently from the case discussed in Ely et al. (2015), the information released, and hence the occurrence of a state of surprise, is not ad hoc driven by the principal to maximize agents' utility. In our case, the surprise is exogenously determined by sudden changes in weather conditions, which is tantamount to stating that Mother Nature plays the role of principal. In the original model of Ely et al. (2015), agents have also a taste for suspense: more suspenseful events (sports matches, novels, films) are preferred over flat calm. By contrast, our study does not explicitly model suspense which, arguably, would negatively affect agents' utility by increasing the stress related to the uncertainty of actual weather conditions.

In its simplest form, the weather is assumed to be a binary variable, either Rainy or Sunny. Based on historical weather trends and, in case of availability, forecasts at the time of booking, customers form a belief. Beliefs get updated following a Bayesian updating mechanism, with forecasts becoming more and more reliable the closer the date to be forecasted. If the expected weather is not realized, the resulting surprise could impact utility. One caveat is in order: in this framework, it is not the weather *per se* to affect the utility, as agents decide to buy a service to be consumed in a specific period of the year and hence know the ex-ante probability of having sunny or rainy weather. In other words, when booking for specific dates, agents have already internalized the expected utility deriving from the two possible states of the world (sunny or rainy), respectively weighted by their probability, based on historical data (and common sense).

Given the theoretical context and the available data, four specific points in time are relevant for the purpose of our study:
i. the time of purchase (which corresponds to the early booking of the service);
ii. the time of consumption (which corresponds to the delivery of the purchased service);
iii. any other intermediate point (when the agent acquires the latest available information before the consumption);
iv. the review time (when the online rating is provided).

If consumption happens at time t, the review is written at time t+k (with k¿0). Going backward, purchasing happens at time t-j (with j¿0). In between t and t-j (t-x, with x¡j), the agent receives updates about the weather forecast of time t. As the consumption event approaches, signals become increasingly more reliable. On the consumption's eve (t-x), the agent acquires a trustworthy signal that serves as the basis for expectation. To ensure an easier comprehension, this timeline is graphically represented in Figure 1.

Beliefs are updated in a finite time window, ranging between the purchasing and the

Figure 1: The consumption timeline

intermediate point. Theoretically, updating beliefs could occur regularly and constantly within the time frame (even very close to t). For sake of simplicity, we postulate that the final update, which is relevant for building expectations, occurs on the eve of departure through a check of the weather forecast.[1]

As regards the weather signal, let us denote with Ω the finite space of states ω. $\omega \in \Omega = S, R$, where S denotes a sunny day and R represents a rainy day. At the time of purchasing the service, the agent forms a belief μ, with μ^{ω} being the probability of the state ω. μ_0 is the prior, *i.e.*, the ex-ante likelihood of ω to occur before considering any new (posterior) information. We expect that for values of j big enough (early bookings), the agent does not avail of reliable forecasts so that the expectations are made on past trends (for example, the historical weather trend for the same period of the year and common sense). The agent receives a signal w (weather forecast), which affects the prior μ_0 and generates a stochastic path of Markov-martingale beliefs, *i.e.*, a sequence where each μ_{t+1} depends onμ_t. θ is the accuracy of each w, which increases as we approach the consumption date. Agents update their belief about the weather and hence signals ($W = \sum_{i=j}^{x} w_i$) affect expectations about the activities that can be performed [2].

Given the binary state of signals (S, R), at the time of consumption, depending on the realized ω, and the signal w, we have the four scenarios represented in Table 1.

		Realised (ω)	
		S	R
Forecasted (w)	S	No surprise	Surprise
	R	Surprise	No Surprise

Table 1: The surprise matrix

The surprise can assume two values: 0 if w is correct (or, alternatively, $w = \omega$); 1 if

[1]It is common for individuals to consult the weather forecast prior to embarking on a trip, as it enables them to be better equipped for any weather contingencies that may arise, such as rain or extreme temperatures. In a straightforward manner, this information is utilized to make informed decisions on what attire and personal belongings to pack for the trip.

[2]For simplicity, we hold the expected quality of the accommodation service constant. Agents choose the service provider based on a set of characteristics and reviews from previous guests. However, one might expect that online reviews (which are other types of signals) posted between purchase and consumption could also generate a path of evolving expectations about the hotel quality and hence affect future enjoyment. In line with existing literature, we expect that for verified platforms, such as Booking.com, ratings tend to be quite stable over time (Figini et al., 2020). Moreover, consumers tend to read reviews and gather information only before booking accommodation, to ensure they make an informed decision

w is incorrect (or, alternatively, $w \neq \omega$). *ceteris paribus*, we expect that surprises (=1) will affect agents' experienced utility. Moreover, according to the good-news bad-news theory (Eil and Rao, 2011), the effects S→R and R→S may not be symmetrical. These effects' magnitude could be different for positive (good news) and negative (bad news) information.

Before discussing the empirical setting, one crucial point should be mentioned. The concept of utility used in the current work is not Decision Utility, used in standard economic literature to explain agents' choice, but is meant in the sense of Experienced Utility (Kahneman and Thaler, 1991). As discussed by these authors, Experienced Utility is measurable and has a neutral point on the boundary between desirable and undesirable. This is indeed more fitting in our study, which recalls utility as the evaluation of an episode. More specifically, among the different types of experienced utility, we link to the specific concept of remembered utility, which consists of a "measure of hedonic and affective experience, inferred from a subject's retrospective reports of the total pleasure or displeasure associated with past outcomes" (Kahneman and Thaler, 2006).

Finally, coherently with the theoretical framework proposed by Ely et al. (2015), the weather is not considered instrumental information in determining the decision to purchase accommodation services. Instead, weather (better said, sudden weather changes) is perceived as non-instrumental information that can potentially influence the perceived and reported utility of the service. In other words, weather surprises may affect the enjoyment of the stay, but it does not directly impact the decision to book the accommodation, which was made prior to the arrival.

4 Data and Methods

4.1 Data

We investigate how people react to surprises (sudden changes in weather conditions) in the context of accommodation stays, a typical experience good. Specifically, we use individual ratings as a proxy for experienced utility. Ratings and other characteristics of the stay are retrieved from Booking.com, a popular hotel reservation platform, while weather forecasts are collected from different online websites, including private providers (such as AccuWeather) and local public providers (Arpa Lombardia and Arpa Veneto, the regional environmental protection agencies that, according to the Italian law, are required to produce and publish weather forecasts at least up to 72 hours). We study two Italian cities, Venice and Milan, covering the period between September 2019 and February 2020. Italy is one of the top destinations in the world, offering a wide range of diverse activities, and the choice of these two cities allows us to explore contexts mainly characterized by culture and leisure (Venice) and business (Milan).

Concerning the temporal frame, the study period allows for a fair comparison of experienced utilities since all scores have been generated under the same review system. In fact, Booking.com changed its rating algorithm starting in September 2019. In the new system, reviewers are asked to provide a numerical rating (from 1 to 10) of the overall experience. Unlike the previous system, which requested reviewers to only evaluate six specific hotel items through smileys (while the overall score was generated as a simple average), the current individual score is explicitly rated by the customer and could better reflect the influence of exogenous factors that are beyond the control of the hotel (for more details on the differences between the old and the new systems, we refer to Leoni and Boto-García (2023)). Moreover, we stopped the data collection in February 2020 to avoid any confounding effects generated by the COVID-19 pandemic and subsequent lockdown measures.

4.1.1 Data collection and descriptive statistics

Accommodation The accommodation data were collected through an ad-hoc build web crawler that mimicked online users' surfing and recorded reviews (and prices) for any room offered on the platform by the population of 3- to 5-star hotels in the two cities. Over the six months, 341,039 observations were collected and included in the dataset, each consisting of an individual evaluation of an accommodation service of a specific provider and for a specific travel date.

For each observation, we avail of data about: (i) the hotel and its characteristics; (ii) the date of the review; (iii) the month of the stay; (iv) the consumers' profile: name, country of origin, travel party and length of stay at the hotel; (v) the pre-existing score, namely the average score assigned by previous costumers up to the week before the stay; (v) the price. This last variable represents a precious piece of information, considering that empirical studies investigating online customer ratings tend to overlook the price, hence not taking into account value-for-money judgments and the role of price in building consumers' expectations. In this regard, it is important to highlight that we do not have information on the exact price associated with each stay, which also depends on the booking time, *i.e.*, how in advance the booking was completed. Therefore, for each date, we built a daily price computed as the simple weekly average of the different prices posted by the hotel for any room type for bookings with different advances[3]. This procedure helps average out differences stemming from different booking windows and room types.

Table 2 includes the descriptive statistics of the relevant variables. The individual score, the dependent variable in our study, proxies the experienced utility with the stay and is a variable ranging between 1 and 10. The mean score is 8.43, and the variable is characterized by highly left skewness, with around 26% of observations scoring 10/10. 67% of reviews are for providers in Venice (125 structures), while the remaining 33% is for Milan (47 structures). This imbalance in the geographic distribution of observations is attributable to the higher number of stays in Venice compared to Milan (ISTAT, 2020)[4]. Moreover, the type of consumers staying overnight in the two cities is quite different. As a matter of fact, Venice hosts a high share of leisure visitors, while Milan is mainly known

[3]We avail of the trajectory of prices set for a specific room, in different points in time from date -15 days up to the actual date of traveling.

[4]Official Statistics on Tourism Flow, Istat. Website: https://www.istat.it/it/files//2020/12/C19.pdf. Accessed on December, 13th 2022.

Label	Description	Mean/Percentage	St. Dev.	Min	Max
`Score`	Individual score (outcome variable)	8.430	1.643	1	10
`LengthOfStay`	Length of the stay (in days)	1.817	0.771	1	3
`Couple`	=1 if travel party=couple	0.456			
`Family`	=1 if travel party=family	0.260			
`Group`	=1 if travel party=group	0.087			
`SoloTraveler`	=1 if travel party=solo traveler	0.125			
`Domestic`	=1if Italian	0.260			
`Anonymous`	=1 if anonymous reviewer	0.135			
`LnPrice`	average price (in ln)	5.361	0.641	2.944	8.422
`PriorRatings`	Average of pre-existing score	8.482	0.624	4.450	10

Table 2: Summary Statistics of Booking.com variables.

for business travelers.

On average, each provider received approximately 3240 reviews in the period under investigation, with a high degree of heterogeneity across providers (min=20; max= 8262; St. dev. = 1926.43). 26% of reviewers are Italians; 45% traveled as a couple, 26% as a family, 12% are solo travelers, and groups represent the remaining 13%. 14% of scores are from anonymous customers. However, as per the platform policy, only real guests can rate the service, which ensures high reliability in this metric. The average number of nights spent by consumers in the accommodation structure is 1.8[5]. The average price per night is €204, displaying a high heterogeneity across dates and providers (min= €50.40; max= €4491.76).

Weather Data on weather conditions were collected from AccuWeather website, owned by an American media company providing commercial weather forecasting services worldwide. For a robustness check, we have also complemented data with the official data on weather conditions recorded by the Regional Environmental Protection Agency of the two involved regions (Arpa Veneto for Venice and Arpa Lombardia for Milan).

As reported in Table 3, there was unexpected rainy weather (*i.e.*, the forecast announced sunny weather while it actually rained) in around 6% of the days under investigation. If we also consider the cases in which the forecast announced cloudy weather, but it actually rained (UnexpectedRainExtended), the percentage increased to 16%. Analogously, 0.3% of days unexpected sunny weather (*i.e.*, the forecast announced rainy weather while it was actually sunny) was experienced, increasing to 3.1% if we also consider forecasted cloudy weather, which turned out to be sunny[6]. This set of dummy variables has been created starting from a discrete variable generated using data from AccuWeather. This variable describes the daily weather according to the phrase and icon that the provider presents on its website and app. It has been consequently categorized into three levels: sunny, cloudy, and rainy. The use of weather icons, as opposed to more

[5]Data have been restricted only to short stays (1-3 days). The reasons behind this choice are discussed later in this Section.

[6]As shown in Appendix A, Table A2, we found no cases of unexpected sunny weather for Venice.

Label	Description	(%)
`UnexpectedRain`	Forecast: sunny → rainy	0.063
`UnexpectedRainExtended`	Forecast: sunny or cloudy → rainy	0.164
`UnexpectedSun`	Forecast: rainy → sunny	0.003
`UnexpectedSunExtended`	Forecast: rainy or cloudy → sunny	0.031
`Sun`	Sunny weather	0.353
`Rain`	Rainy weather	0.266
`Cloudy`	Cloudy weather	0.381

Table 3: Summary Statistics weather data

detailed information about humidity, air pressure, solar radiation, etc., makes weather forecasts easier to be interpreted by final users and hence further reduces the cost of information.

For robustness checks, we have also created the dummy `sunniericon` which captures a situation where the estimated weather has turned out to be sunnier than the forecasted one. We also have generated two other continuous variables representing the difference between the actual and forecasted hours of sun and the difference between the actual and forecasted rain probability. For space reasons, we leave the descriptive statistics and further description of these variables in Appendix A, which also contains a breakdown of weather variables per city (Table A1-A2).

4.2 Sample construction

We built a panel dataset combining Booking.com and weather conditions data and proceeded with micro econometric analysis to test the effect of surprises on the score of each individual review i left at time t+k by a consumer who stayed at time t in the accommodation h. Booking.com includes the exact date of the review but only the month and the year of the rated stay. Building upon existing literature on the evaluation of accommodations services (Brandes and Dover, 2022), we assume that the reviews are written, on average, two days after the stay (hence, k=2) (*e.g.*, if the review date is October 10^{th}, we assume that the customer ended his/her stay on October 8^{th}). Hence, we match any review with the weather of the stay (*i.e.*, October 8^{th}), and with the weather forecast of four days before the stay (*i.e.*, October 4^{th}) (hence, x=4)[7]. To reduce the potential error in the assignment of a specific stay date, we have limited our analysis only to short stays (length of stays which are less than or equal to 3 days) and only to those reviews where the month of the review and those of the stay coincides.[8] By limiting the sample to short stays, and based on the evidence that the review is written two days after the stay, we assume that the weather forecast taken into consideration is 2 days before the

[7]Most of the hospitality services are bought with more than 4 days of advance. However, we understand that for $j \leq 4$, the model would consider only three points in time.

[8]We acknowledge that for reviews written in the first two days of the month, the reported stay month may differ from the month of the review. For example, if a guest checks out from a hotel on October $31st$, the review will likely be written in November. We have taken great care to handle such cases appropriately in our analysis.

stay for those staying 2 nights, while it is of 3 (1) days before the stay for those staying 1 (3) nights. On these premises, the specific choice of the last weather check (x=4) considers two main aspects: (i) the reliability of weather forecasts (95% accuracy about 72 hours before t[9]); (ii) given the restricted sample (we only consider short stays), weather checks are made, on average, two days before departure. We match each stay date t with the corresponding weather conditions of the city where the accommodation is located (Venice or Milan). The panel dataset is unbalanced because a specific provider might have received no review or more than one single review in a specific t.

4.3 Empirical Strategy

To explore the effect of surprises on experienced utility, we estimate the following utility model:

$$\mathrm{U}_{ijht} = \beta' \boldsymbol{X_i} + \mu' \boldsymbol{Z_{ht}} + \tau Surprise_{jt} + \alpha_h + \alpha_m + \epsilon_{ijht} \tag{1}$$

Where U_{ijht} is the (experienced) utility of the individual consumer i who consumed the accommodation service of the provider h, located in city j, at time t. The utility is proxied by the individual review score. $\boldsymbol{X_i}$ is a vector of consumer characteristics (nationality, anonymity, length of stay, and travel party). $\boldsymbol{Z_{ht}}$ is a vector of time-varying provider characteristics (price (in logarithm) and the stock of previous reviews). τ is the main parameter to be estimated, capturing the effect of the surprise on the outcome. α_h are provider-fixed effects, which address issues of time-unvarying unobserved heterogeneity across different providers. α_m are monthly fixed effects, controlling for seasonal variations and hence for prior beliefs, and ϵ_{ihjt} is the stochastic error term.

Based on the hedonic adaptation theory, we expect that the effect weakens for longer stays (the effect gets diluted/spread over time). Therefore, we provide an alternative specification including interaction terms between Surprise and the stay duration (expressed in days). More formally:

$$\mathrm{U}_{ijht} = \beta' \boldsymbol{X_i} + \mu' \boldsymbol{Z_{ht}} + \tau Surprise_{jt} + \phi Surprise_{jt} \times LenghtOfStay_i + \alpha_h + \alpha_m + \epsilon_{ijht} \tag{2}$$

We remind (see Subsection 4.1) that only short stays (≤3 days) are considered. The models in 1 and 2 are estimated via Ordinary Least Squares with a robust variance estimator [10].

5 Results

5.1 Main Findings

Table 4 shows the results of Model 1 on the pooled dataset. The main finding is that the surprise, measured as the mismatch between expected and realized weather, systematically affects the customers' experienced utility, proxied by their rating score. More

[9] Reliability of weather forecast, from the Aeronautica Militare webpage: https://www.meteoam.it/it/attendibilita-previsioni. Accessed on January, 27^{th} 2023.

[10] In the main estimations, we use Huber-White heteroskedasticity-robust standard errors.

specifically, `unexpectedrain` is associated with a significant decrease in the overall utility (-0.142***, Column 1) when the rain comes as a surprise during the stay, *ceteris paribus*. The magnitude of such an effect is weaker when the mismatch between expected and realized weather is less dramatic. In fact, when also considering cases of forecasted cloudy (`unexpectedrainextended`), the surprise's impact is almost halved (-0.066***, Column 2). Although the drop's magnitude might appear negligible, we recall that the dependent variable ranges on a 1 to 10 scale, with a relatively low variation around the mean (st.dev. =1.643). Surprises with the opposite sign (sunshine when forecasts announced rainy weather) translate into higher experienced utility. As shown in Column 3, `unexpectedsun` is associated with higher scores (0.298***). Like the case of rain, also for the sun, a weaker surprise, which also includes cases of forecasted cloudy (`unexpectedsunextended`), produces a sensibly lower (yet still positive) effect (0.128***, Column 4). These results also confirm the asymmetry in the reactions to different types of surprise, in line with studies on asymmetric responses to financial information. However, we have no supporting evidence of any negativity bias (Schwager and Rothermund, 2013). Extant literature, theoretically grounded in the good news/bad news theory, consistently finds that negative information has a much greater impact on individuals' attitudes than positive information does (Eil and Rao, 2011; Nguyen and Claus, 2013; Soroka, 2006). By contrast, our estimates suggest that the increase in reported utility when having unexpected good weather is higher than the drop caused by unexpected unfavorable weather.

In Table 4 we also explore the heterogeneity in the surprise effect for the two cities, as modeled by introducing interactions between the surprise variables and the city dummy. There is no significant difference in the effect of `unexpectedrain` for the two cities, as the coefficient of the interaction variable is statistically insignificant. By contrast, the effect of `unexpectedrainextended` varies sensibly across the two subsets: while for Venice, rain negatively affects ratings (-0.149***, Column 6), even when cloudy weather was expected, the effect for Milan is actually positive. By contrast, the effect of an `unexpectedsunextended` is positive for both cities, but the magnitude is much lower for Venice (-0.196***, Column 7)[11].

The key novel point of our work relates to the analysis of the surprise effect rather than the effect of the weather *per se*, which was largely analysed in related literature (Brandes and Dover, 2022; Lee et al., 2014; Radic and Lück, 2018; Štumpf et al., 2021). Interestingly, by running a specification where only the actual weather (not the mismatch with prior expectations) is considered (Table 5, Column 1), we notice that the coefficients of `sunny` and `rainy` are both negative and significant (respectively -0.052 and -0.054, Column 1, reference category: cloudy). Arguably, the "strange" coefficient for `sunny` might signal an omitted variable bias stemming, in line with the theoretical framework herein reported, from the inclusion of the weather status alone, not considering the agents' expectations about the weather (the available forecasts).

In Table 5, we test a slightly different specification of Model 1 to further explore

[11] As previously mentioned in Section 4.3, an interaction term for Venice and the variable `unexpectedsun` was not included in the model as we did not have any instances of this scenario within the specified study period.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LengthOfStay	-0.0252***	-0.0249***	-0.0250***	-0.0248***	-0.0252***	-0.0245***	-0.0248***
	[-6.85]	[-6.77]	[-6.80]	[-6.75]	[-6.85]	[-6.67]	[-6.77]
Domestic	-0.0424***	-0.0430***	-0.0426***	-0.0419***	-0.0424***	-0.0438***	-0.0420***
	[-6.42]	[-6.51]	[-6.46]	[-6.35]	[-6.42]	[-6.64]	[-6.35]
Couple	0.0171**	0.0174**	0.0162*	0.0182**	0.0171**	0.0170**	0.0181**
	[2.07]	[2.10]	[1.95]	[2.20]	[2.07]	[2.06]	[2.19]
Family	0.0465***	0.0453***	0.0457***	0.0471***	0.0465***	0.0445***	0.0476***
	[5.18]	[5.05]	[5.09]	[5.25]	[5.18]	[4.97]	[5.30]
Single	0.00126	0.00104	-0.00195	0.000349	0.00126	0.000877	-0.000923
	[0.12]	[0.10]	[-0.18]	[0.03]	[0.12]	[0.08]	[-0.09]
Anonymous	-0.203***	-0.202***	-0.202***	-0.201***	-0.203***	-0.202***	-0.202***
	[-25.73]	[-25.50]	[-25.60]	[-25.43]	[-25.73]	[-25.53]	[-25.46]
LnPrice	-0.279***	-0.288***	-0.284***	-0.282***	-0.279***	-0.284***	-0.286***
	[-17.15]	[-17.62]	[-17.44]	[-17.28]	[-17.06]	[-17.41]	[-17.53]
PriorRatings	-2.037***	-2.048***	-2.040***	-2.035***	-2.037***	-2.053***	-2.032***
	[-27.35]	[-27.52]	[-27.42]	[-27.37]	[-27.35]	[-27.55]	[-27.34]
UnexpectedRain	-0.142***				-0.139***		
	[-10.84]				[-6.63]		
UnexpectedRainExtended		-0.0660***				0.0325***	
		[-8.42]				[2.64]	
UnexpectedSun			0.298***				
			[9.00]				
UnexpectedSunExtended				0.128***			0.270***
				[8.78]			[10.76]
UnexpectedRain#Venice					-0.00424		
					[-0.16]		
UnexpectedRainExtendes#Venice						-0.149***	
						[-9.49]	
UnexpectedSunExtended#Venice							-0.196***
							[-6.37]
Constant	27.25***	27.38***	27.29***	27.23***	27.25***	27.41***	27.23***
	[42.92]	[43.16]	[43.04]	[42.97]	[42.91]	[43.15]	[42.98]
Provider FE	YES	YES	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES	YES	YES
N	320685	320685	320685	320685	320685	320685	320685

Table 4: Regression estimates of the models in 1
.

this matter. Conditional of being rainy, we look in Column 2 at the effect of a stronger (`sunny to rainy`) and a weaker surprise (`cloudy to rainy`), with no surprise (`rainy to rainy`) as a reference category[12]. Similarly, conditional on being sunny, Column 3 studies the effect of a strong (`rainy to sunny`) and weak (`cloudy to sunny`) surprise. Results show that, compared to a rainy day with no surprise (*i.e.*, the rain was correctly forecasted), there is a significant and negative effect on the rating score of the unexpected rain when the forecast was of a sunny day (-0.129***, Column 2). Interestingly, the effect is much weaker (statistically nil) when the unexpected rainy weather follows a forecast of cloudy weather only (`cloudy to rainy`, Column 2). In the same fashion, there is a strong increase in the reported rating scores when agents experience sunny weather during their stay after a wrong forecast of rain (0.522***, Column 3). Again, such an effect is weaker when the forecast is of cloudy weather only (0.192***, Column 3).

In Column 4, we then look at the effect of different types of surprises. While we consider the reference category as "no surprise" (when the forecast equals the realized weather), in line with the specific literature finding that tourists prefer good weather conditions (Jeuring and Peters, 2013), we define a `positive surprise` the case when the weather is sunny after a forecast of rain or cloudy weather, while a `negative surprise` in the opposite case.[13] Moreover, we control for the actual weather, considering cloudy as the reference category. This further check confirms the importance of expectations. In fact, while the effect of the actual weather is ambiguous (the coefficient of rainy is not significant, and the coefficient of sunny is negative), this specification highlights once more that what drives the utility (measured by the score) is the surprisingness of an event: *ceteris paribus*, compared to a no surprise scenario, unexpected sun leads to a higher level of reported utility (0.165***, Column 4) while unexpected rain lowers utility (-0.091***, Column 4). Again, the effect is asymmetric, with the positive surprise effect being approximately 80% higher than the impact of the negative surprise.

A potential source of heterogeneity is the consumption span: it is likely that the length of the episode dilutes any surprise effect. Consistently, we expect that the surprise effect diminishes with the length of stay. In Table 5, Columns 5 and 6, we estimate Model 2 by interacting the surprise variables with the number of days spent in the city (`unexpectedrain#lenghtofstay`; `unexpectedsun#lenghtofstay`). Evidence supports the theoretical construct: the effect of surprise keeps being significant and with the expected sign, coherently with the above discussion (-0.371*** for `unexpectedrain`, Column 5; 0.511*** for `unexpectedsun`, Column 6), the effect weakens for longer stays (0.125***, Column 5; -0.132**, Column 6). Hence, in line with the hedonic adaptation theory (Frederick and Loewenstein, 1999), staying longer means having more time to adapt to the circumstances and the surprise gets more diluted, spreading over a longer period.

[12] The number of observations used in the estimation of the model in Columns 2 and 3 is less than that of Column 1 because the observations were limited to rainy days (Column 2) or sunny days (Column 3) only.

[13] We define these surprises as "positive" and "negative" because we assume that, for leisure purposes, sunny weather is preferred to rain. However, it is mostly a matter of labeling, and this definition could vary according to the type of activities to be carried out when on holiday.

	(1)	(2)	(3)	(4)	(5)	(6)
`Unexpectedrain`					-0.371***	
					[-11.80]	
`Unexpectedsun`						0.511***
						[5.18]
`Sunny`	-0.0519***			-0.0716***		
	[-8.28]			[-11.02]		
`Rainy`	-0.0544***			0.00901		
	[-7.27]			[0.80]		
`Sunnytorainy`		-0.129***				
		[-7.07]				
`Cloudytorainy`		0.0102				
		[0.72]				
`Rainytosunny`			0.522***			
			[11.18]			
`Cloudytosunny`			0.192***			
			[11.33]			
`Unexpectedrain#LenghtOfStay`					0.125***	
					[8.49]	
`Unexpectedsun#LenghtOfStay`						-0.132**
						[-2.36]
`Positive Surprise`				0.165***		
				[10.90]		
`Negative Surprise`				-0.0908***		
				[-7.45]		
`Constant`	27.50***	24.06***	28.28***	27.35***	27.23***	27.29***
	[43.38]	[19.71]	[26.47]	[43.15]	[42.85]	[43.04]
N	323026	79957	119312	323026	320685	320685
Consumer Characteristics	YES	YES	YES	YES	YES	YES
Provider FE	YES	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES	YES
adj. R2	0.146	0.173	0.159	0.147	0.147	0.146

Table 5: . Extensions on the surprise effect model in 2

.

5.2 Further results

Although the main goal of the work is to isolate the effect of surprise from other confounding elements that could also affect the rating score, this sub-section offers a quick overview of the estimated coefficients of the other covariates. Estimated coefficients in Tables 4 and 5, despite some marginal differences in their magnitude across specifications (pooled and with interaction models), remain consistent. As expected, price (`lnprice`) negatively impacts scores. Higher prices entail higher expectations about the overall quality, resulting, *ceteris paribus*, in lower scores. Interestingly, `anonymous` reviews are associated with lower scores, a result that aligns with deindividuation theories. As Deng et al. (2021) explain, anonymity enables reviewers to give worse ratings because of a lower self-awareness and social presence[14]. `Domestic` (*i.e.*, Italian guests) tend to provide lower scores compared to foreigners, perhaps due to a better familiarity with the reference quality that can be expected from the service (Cordell, 1997). Moreover, `couples` and `families` are associated with higher scores compared to people traveling in groups.

Finally, longer stays (`LenghtOfStay`) are associated with lower scores: it is reasonable to think that spending more time in the accommodation, people might be more analytical about the services evaluations (Kim and Han, 2022; Boto-García and Leoni, 2023) and that longer stays decrease the comfort sensation with hotel facilities. Another interesting result concerns the effect of `priorratings`. As discussed in the theoretical framework, the nature of the analysed product as an experience good (*i.e.*, a good whose quality can be ascertained only upon consumption) makes other agents' opinions an important source of information when choosing a service provider. Our results suggest that with everything else being equal, a higher average rating translates into lower individual scores, perhaps due to the internalization of higher expectations about the quality. We highlight that the quality of the service is not fully contained in price since both variables appear to have a strong significant effect in all different model specifications.

5.3 Robustness checks

We performed a battery of robustness checks to test the sensitivity of our findings. First, we re-estimated Model 1 with bootstrapped standard errors to consider the variance of our estimators. The significance of the coefficients remains unaltered (Appendix A, Table A4). Second, we estimated the model with a two-way fixed effects specification, including both provider and country of residence fixed effects. This choice was made to consider individual-specific effects for people coming from the same country, which could be correlated with experienced utility. Again, the results are similar to the ones of the main model (Appendix A, Table A5). Third, we have included alternative variables to identify the surprise effect, namely the difference between the forecasted and actual hours of sunshine, the mismatch between predicted and actual rain probability, and a dummy variable taking value 1 when the weather is sunnier than expected (Appendix A, Table A6). Finally, we have repeated the tests run in Table 5 on the effect of actual rain and

[14] A further check included the gender (obviously excluding anonymous agents) and we found that, *ceteris paribus*, women tend to report lower scores than men. Results are available upon request to the authors.

sun on scores using Arpa data as an alternative public source of information. Again, results are in line with Table 5 and are available upon request to the authors.

The period under investigation includes an extreme High-Water event that took place in Venice on November 12^{th}, 2019. To avoid inconsistent estimates, we performed a separate analysis for Venice, excluding the week of the event (Appendix A, Table A7) and another model which includes a control for the event (Appendix A, Table A8). Results are robust also to these alternative specifications.

Lastly, given the potential difference between week vs weekend visitors (mostly concerning the business or leisure nature of the visit), Table A9 in the Appendix runs the model both on the full sample (Column 1) and on the restricted weekend subset (Column 2). Results are in line with the main results except for the impact of the negative surprise, which is not significantly different from a correctly forecasted rain. In general, we can conclude that our findings are robust to the inclusion of alternative measures, variables, and model specifications.

6 Concluding remarks

The utility experienced by subjects after a consumption act depends not only on the intrinsic quality of the service but also on adherence to prior expectations. This is especially relevant for experience goods, whose quality can be ascertained only upon consumption, and that might also be affected by other contextual factors ((Brandes and Dover, 2022)). We argue that the hospitality industry is a relevant case to test the existence of a "surprise effect", which was theoretically proposed by Ely et al. (2015) for entertainment and sports activities. We adapt and modify their approach in two respects: one, given the nature of travel as good of exogenous quality (Candela and Cellini, 1997), we assume that the surprise effect can also stem from elements not under the control of the supplier, such as unexpected weather conditions; two, given that visitors generally have strong preferences for sunny weather conditions, we assume that the surprise effect can be perceived either as positive or negative according to the sign of the difference between forecasted and realized weather.

We test the model using the individual ratings posted on Booking.com by customers who stayed in accommodation services in Milan and Venice (two important destinations differing in their customer mix) between September 2019 and February 2020. The main findings support the theoretical model. Specifically: (i) The reported utility is affected by external aspects, namely weather conditions; (ii) Differently from existing literature, and coherently with Ely et al. (2015), we find that the driver is not the situational factor **per se**, but the surprise: the difference between expectations (which are captured by public weather forecasts available in the last days before the stay) and the realized weather; (iii) *ceteris paribus*, a positive surprise (the realized weather is sunnier than expected) is associated to higher experienced utility, while a negative surprise (the realized weather is rainier than expected) is associated to lower experienced utility; (iv) In contrast to good-news bad-news theory, we find no evidence of negativity bias, with positive surprises exerting stronger effects than negative ones; (iv) The effect is moderated by the length of stay (the longer the consumption span, the weaker the surprise effect), and is slightly different between Milan and Venice, with attitudes towards cloudy weather varying in the two cities.

To the best of the authors' knowledge, ours is a novel contribution to both the economics literature on the effect of surprises and the one on the online reported utility. Our research builds on the ground-breaking work of Ely et al. (2015), which opened a new path for exploring suspense and surprise dynamics, and we offer an empirical application in a different context. Moreover, as per the literature on the effect of weather on recreational activities, our work investigates another facet of the relationship between weather-related information and agents' behaviour. In fact, while the existing works mostly focus on the instrumental value of weather information, and hence its role in predicting behaviours or influencing market dynamics, we explicitly include expectations and look at the non-instrumental power of weather dynamics to drive agents' utility.

As for more practical takeaways, indeed, service providers do not have the power to command the weather, but certainly, more care is needed from their side to investigate how external effects can affect the utility and how to provide customers with alternative ways of "enjoying" the stay, even last minute, if the weather turns out to be not as expected. Moreover, considering the climate crisis, the change in weather conditions, their predictability, and the increase in the frequency and intensity of extreme events will impact business operations and hence the customers' evaluation. Finally, we stress the relevance of weather forecast providers in the travel industry, with non-neutral implications on consumers' enjoyment and, in turn, future behaviours. While, on the one hand, providers are hardly working to have more precise, reliable, and easily available weather forecasts, on the other hand, it is well known that they apply a rain bias (Silver, 2012), given the asymmetric effect that a wrong forecast has on customers. Interestingly, such rain bias has two opposite effects on hotels: one, it might discourage people from booking a stay, given that they expect worse-than-optimal weather conditions, with an impact on revenues and prices (Figini et al., 2022); two, the rain bias might trigger better rating scores, given the positive surprise effect found in this work.

We believe there is room for further research in this line of investigation. We indicate three main avenues: (i) the extension of our methodology to different areas, characterised by different external events or different tastes for weather events; (ii) the analysis of longer episodes, potentially characterised by multiple surprises. This would allow exploring the extent to which the timing of the surprise (*e.g.*, at the start or at the end of an episode) plays a role (Kahneman and Thaler, 2006) in shaping utility. (iii) This analysis should also expand to other contexts where experienced utility might be affected by exogenous surprises (*e.g.*, sports events), hence adding further empirical support to Ely et al. (2015).

References

Agarwal, S., Wang, L., and Yang, Y. (2021). Impact of transboundary air pollution on service quality and consumer satisfaction. *Journal of Economic Behavior & Organization*, 192:357–380.

Baccan, D., Sbruzzi, E., and Macedo, L. (2015). On the cognitive surprise in risk management: An analysis of the value-at-risk (VaR) historical. In *Progress in Artificial Intelligence*, pages 402–413. Springer International Publishing.

Bizzozero, P., Flepp, R., and Franck, E. (2016). The importance of suspense and surprise in entertainment demand: Evidence from wimbledon. *Journal of Economic Behavior & Organization*, 130:47–63.

Boto-García, D. and Leoni, V. (2023). Noisy signals: do ratings' volatility depend on the length of the consumption span?

Brandes, L. and Dover, Y. (2022). Offline context affects online reviews: The effect of post-consumption weather. *Journal of Consumer Research*, 49(4):595–615.

Bray, M. M. and Savin, N. E. (1986). Rational expectations equilibria, learning, and model specification. *Econometrica*, 54(5):1129.

Bryant, J., Rockwell, S. C., and Owens, J. W. (1994). "Buzzer Bearters" and "Barn Burners": The effects on enjoyment of watching the game go "down the wire ". *Journal of Sport and Social Issues*, 18(4):326–339.

Candela, G. and Cellini, R. (1997). Countries' size, consumers' preferences and specialization in tourism: a note. *Rivista Internazionale di Scienze Economiche e Commerciali*, 44:451–458.

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

Cicognani, S., Figini, P., and Magnani, M. (2021). Social influence bias in ratings: A field experiment in the hospitality sector. *Tourism Economics*, 28(8):2197–2218.

Connolly, M. (2008). Here comes the rain again: Weather and the intertemporal substitution of leisure. *Journal of Labor Economics*, 26(1):73–100.

Cordell, V. V. (1997). Consumer knowledge measures as predictors in product evaluation. *Psychology and Marketing*, 14(3):241–260.

Dehaan, E., Madsen, J., and d. Piotroski, J. (2017). Do weather-induced moods affect the processing of earnings news? *Journal of Accounting Research*, 55(3):509–550.

DeMarzo, P. M., Vayanos, D., and Zwiebel, J. (2003). Persuasion bias, social influence, and unidimensional opinions. *The Quarterly Journal of Economics*, 118(3):909–968.

Deng, L., Sun, W., Xu, D., and Ye, Q. (2021). Impact of anonymity on consumers' online reviews. *Psychology & Marketing*, 38(12):2259–2270.

Derbaix, C. and Vanhamme, J. (2003). Inducing word-of-mouth by eliciting surprise – a pilot investigation. *Journal of Economic Psychology*, 24(1):99–116.

Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior & Organization*, 27(1):35–47.

Eil, D. and Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2):114–138.

Ely, J., Frankel, A., and Kamenica, E. (2015). Suspense and surprise. *Journal of Political Economy*, 123(1):215–260.

Figini, P., Cicognani, S., and Zirulia, L. (2022). Booking in the rain. testing the impact of public information on prices. *Italian Economic Journal*, pages Online first: https://doi.org/10.1007%2Fs40797–022–00199–y.

Figini, P., Vici, L., and Viglia, G. (2020). A comparison of hotel ratings between verified and non-verified online review platforms. *International Journal of Culture, Tourism and Hospitality Research*, 14(2):157–171.

Fradkin, A., Grewal, E., and Holtz, D. (2018). The determinants of online review informativeness: Evidence from field experiments on airbnb. *SSRN Electronic Journal*.

Frederick, S. and Loewenstein, G. (1999). 16 hedonic adaptation. *D. Kahneman ED, N. Schwarz., Editors. Well-Being the Foundations of Hedonic Psychology*, pages 302–329.

Greiff, M. and Paetzel, F. (2020). Information about average evaluations spurs cooperation: An experiment on noisy reputation systems. *Journal of Economic Behavior & Organization*, 180:334–356.

Gössling, S., Abegg, B., and Steiger, R. (2016). "it was raining all the time!": Ex post tourist weather perceptions. *Atmosphere*, 7(1):10.

Helmers, C., Krishnan, P., and Patnam, M. (2019). Attention and saliency on the internet: Evidence from an online recommendation system. *Journal of Economic Behavior & Organization*, 161:216–242.

Huysmans, F. (2002). Influence of the weather on time use.

Janakiraman, N., Meyer, R. J., and Morales, A. C. (2006). Spillover effects: How consumers respond to unexpected changes in price and quality. *Journal of Consumer Research*, 33(3):361–369.

Jeuring, J. H. and Peters, K. B. (2013). The influence of the weather on tourist experiences. *Journal of Vacation Marketing*, 19(3):209–219.

Jeuring, J. H. G. (2017). Weather perceptions, holiday satisfaction and perceived attractiveness of domestic vacationing in the netherlands. *Tourism Management*, 61:70–81.

Kahneman, D. and Thaler, R. (1991). Economic analysis and the psychology of utility: Applications to compensation policy. *The American Economic Review*, 81(2):341–346.

Kahneman, D. and Thaler, R. H. (2006). Anomalies: Utility maximization and experienced utility. *Journal of economic perspectives*, 20(1):221–234.

Kettlewell, N., Morris, R. W., Ho, N., Cobb-Clark, D. A., Cripps, S., and Glozier, N. (2020). The differential impact of major life events on cognitive and affective wellbeing. *SSM - Population Health*, 10:100533.

Kim, J. M. and Han, J. (2022). Impact of the length of stay at hotels on online reviews. *International Journal of Contemporary Hospitality Management*, 34(4):1249–1269.

Klimstra, T. A., Frijns, T., Keijsers, L., Denissen, J. J. A., Raaijmakers, Q. A. W., van Aken, M. A. G., Koot, H. M., van Lier, P. A. C., and Meeus, W. H. J. (2011). Come rain or come shine: Individual differences in how weather affects mood. *Emotion*, 11(6):1495–1499.

Lee, J. J., Gino, F., and Staats, B. R. (2014). Rainmakers: Why bad weather means good productivity. *Journal of Applied Psychology*, 99(3):504–513.

Leoni, V. and Boto-García, D. (2023). Apparent and actual hotel scores under booking.com new reviewing system. *International Journal of Hospitality Management*, 111:103493, Online first: https://doi.org/10.1016/j.ijhm.2023.103493.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531.

Mayzlin, D., Dover, Y., and Chevalier, J. (2014). Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review*, 104(8):2421–2455.

Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, 78(2):311–329.

Nguyen, V. H. and Claus, E. (2013). Good news, bad news, consumer sentiment and consumption behavior. *Journal of Economic Psychology*, 39:426–438.

Oswald, A. J. and Powdthavee, N. (2008). Does happiness adapt? a longitudinal study of disability with implications for economists and judges. *Journal of Public Economics*, 92(5-6):1061–1077.

Peterson, E. M. and Raney, A. A. (2008). Reconceptualizing and reexamining suspense as a predictor of mediated sports enjoyment. *Journal of Broadcasting &amp Electronic Media*, 52(4):544–562.

Radic, A. and Lück, M. (2018). (don't) blame it on the weather: the impact of weather on cruise tourism service quality perception and satisfaction. *Australian Journal of Maritime &amp Ocean Affairs*, 10(4):205–221.

Rayo, L. and Becker, G. S. (2007). Evolutionary efficiency and happiness. *Journal of Political Economy*, 115(2):302–337.

Reimers, I. and Waldfogel, J. (2021). Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings. *American Economic Review*, 111(6):1944–71.

Riis, J., Loewenstein, G., Baron, J., Jepson, C., Fagerlin, A., and Ubel, P. A. (2005). Ignorance of hedonic adaptation to hemodialysis: A study using ecological momentary assessment. *Journal of Experimental Psychology: General*, 134(1):3–9.

Schwager, S. and Rothermund, K. (2013). Motivation and affective processing biases in risky decision making: A counter-regulation account. *Journal of Economic Psychology*, 38:111–126.

Schützwohl, A. (1998). Surprise and schema strength. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(5):1182–1199.

Silver, N. (2012). *The signal and the noise: Why so many predictions fail-but some don't*. Penguin.

Simonov, A., Ursu, R. M., and Zheng, C. (2022). Suspense and surprise in media product design: Evidence from twitch. *Journal of Marketing Research*, 60(1):1–24.

Sims, C. A. (2006). Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, 96(2):158–163.

Soroka, S. N. (2006). Good news and bad news: Asymmetric responses to economic information. *The Journal of Politics*, 68(2):372–385.

Štumpf, P., Vojtko, V., McGrath, R., Rašovská, I., Ryglová, K., and Šácha, J. (2021). Destination satisfaction comparison excluding the weather effect. *Current Issues in Tourism*, 25(15):2404–2421.

Wilson, T. D. and Gilbert, D. T. (2008). Explaining away: A model of affective adaptation. *Perspectives on Psychological Science*, 3(5):370–386.

And suddenly, the rain! How surprises *surprises* shape experienced utility

Appendix A

Label	Mean/Percentage	St. Dev	Min	Max
Unexpected rainy	.055	.229	0	1
Unexpected rainy extended	.157	.366	0	1
Unexpected sunny	.008	.089	0	1
Unexpected sunny extended	.031	.175	0	1
Hours of sun	5.086	2.810	.800	11.400
Hours of rain	1.920	3.630	0	16
Sunny	.379	.487	0	1
Rainy	.271	.446	0	1
Cloudy	.349	.478	0	1

Table A1. Summary Statistics weather data - Milan subset

	Mean/ Percentage	St. Dev.	Min	Max
Unexpected rainy	.069	.255		
Unexpected rainy extended	.170	.377		
Unexpected sunny	0	0		
Unexpected sunny extended	.031	.175		
Hours of sun	5.167	2.800	.800	10.900
Hours of rain	1.432	2.904	0	20
Sunny	.331	.472		
Rainy	.263	.441		
Cloudy	.406	.493		

Table A2. Summary Statistics weather data - Venice subset

	Mean/ Percentage	St. Dev.	Min	Max
Sunnier icon	.154	.362	0	1
Diff. hours of sun	.228	2.020	-7.100	7.100
Diff. rain probability	-.113	19.486	-81	57

Table A3. Summary Statistics of additional weather variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LengthOfStay	-0.0252***	-0.0249***	-0.0250***	-0.0224***	-0.0230***	-0.0323***	-0.0246***
	[-7.07]	[-6.73]	[-6.77]	[-5.93]	[-6.13]	[-8.37]	[-7.17]
Domestic	-0.0424***	-0.0430***	-0.0426***	-0.0443***	-0.0363***	-0.0431***	-0.0426***
	[-6.27]	[-6.74]	[-6.09]	[-6.17]	[-5.24]	[-6.77]	[-6.59]
Couple	0.0171*	0.0174**	0.0162**	0.0215**	0.0160*	0.0177**	0.0163**
	[1.91]	[2.13]	[2.02]	[2.55]	[1.82]	[2.02]	[2.06]
Family	0.0465***	0.0453***	0.0457***	0.0596***	0.0455***	0.0457***	0.0457***
	[4.74]	[5.29]	[5.65]	[6.63]	[5.47]	[4.61]	[4.99]
Single	0.00126	0.00104	-0.00195	0.0159	0.00614	0.00107	-0.00187
	[0.11]	[0.10]	[-0.19]	[1.55]	[0.61]	[0.09]	[-0.19]
Anonymous	-0.203***	-0.202***	-0.202***	-0.202***	-0.199***	-0.203***	-0.202***
	[-26.43]	[-23.95]	[-23.47]	[-25.31]	[-25.13]	[-25.20]	[-25.97]
Lnprice	-0.279***	-0.288***	-0.284***	-0.270***	-0.266***	-0.282***	-0.284***
	[-17.48]	[-15.44]	[-14.32]	[-16.60]	[-17.84]	[-17.01]	[-18.66]
Prior_Ratings	-2.037***	-2.048***	-2.040***	-2.055***	-2.073***	-2.032***	-2.040***
	[-30.17]	[-24.58]	[-25.58]	[-24.86]	[-29.81]	[-27.77]	[-26.65]
Unexpectedrain	-0.142***					-0.371***	
	[-10.31]					[-12.44]	
Unexpectedrain_extended		-0.0660***					
		[-8.24]					
Unexpectedsun			0.298***				0.511***
			[9.28]				[5.17]
Unexpectedsun_ extended							
Sunny					-0.0519***		
					[-7.47]		
Rainy					-0.0544***		
					[-7.43]		
Unexpectedrain#C.Los						0.125***	
						[9.09]	
Unexpectedsun#C.Los							-0.132**
							[-2.53]
Constant	27.25***	27.38***	27.29***	27.34***	27.50***	27.23***	27.29***
	[47.31]	[39.06]	[40.28]	[38.97]	[46.31]	[43.72]	[41.32]
N	320685	320685	320685	328577	323026	320685	320685
R2	0.147	0.147	0.147	0.146	0.147	0.147	0.147
adj. R2	0.146	0.146	0.146	0.146	0.146	0.147	0.146

Table A4. Regression estimates with bootstrapped standard error.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Length Of Stay	-0.0176***	-0.0174***	-0.0175***	-0.0174***	-0.0155***	-0.0244***	-0.0171***
	[-4.72]	[-4.66]	[-4.70]	[-4.66]	[-4.16]	[-6.37]	[-4.58]
Couple	0.0126	0.0125	0.0110	0.0131	0.00792	0.0130	0.0111
	[1.50]	[1.50]	[1.32]	[1.56]	[0.95]	[1.56]	[1.33]
Family	0.0304***	0.0289***	0.0289***	0.0303***	0.0277***	0.0295***	0.0289***
	[3.35]	[3.18]	[3.19]	[3.35]	[3.06]	[3.25]	[3.19]
Single	0.0218**	0.0213**	0.0179*	0.0202*	0.0238**	0.0215**	0.0179*
	[2.02]	[1.98]	[1.65]	[1.87]	[2.21]	[2.00]	[1.66]
Anonymous	-0.176***	-0.175***	-0.175***	-0.174***	-0.173***	-0.176***	-0.175***
	[-22.25]	[-22.02]	[-22.12]	[-21.99]	[-21.92]	[-22.18]	[-22.12]
Lnprice	-0.255***	-0.264***	-0.260***	-0.257***	-0.238***	-0.257***	-0.260***
	[-15.61]	[-16.10]	[-15.90]	[-15.73]	[-14.47]	[-15.73]	[-15.90]
Prior_Ratings	-2.046***	-2.056***	-2.048***	-2.044***	-2.085***	-2.041***	-2.048***
	[-27.42]	[-27.58]	[-27.48]	[-27.45]	[-27.99]	[-27.33]	[-27.48]
Unexpectedrain	-0.146***					-0.365***	
	[-10.97]					[-11.48]	
Unexpectedrain_ extended		-0.0700***					
		[-8.91]					
Unexpectedsun			0.315***				0.564***
			[9.10]				[5.45]
Unexpectedsun_ extended				0.119***			
				[8.17]			
Sunny					-0.0641***		
					[-10.28]		
Rainy					-0.0620***		
					[-8.33]		
Unexpectedrain#Lenghtofstay						0.120***	
						[8.02]	
Unexpectedsun#Lenghtofstay							-0.154***
							[-2.59]
Constant	27.17***	27.30***	27.20***	27.15***	27.44***	27.15***	27.20***
	[42.71]	[42.95]	[42.83]	[42.77]	[43.21]	[42.65]	[42.83]
N	320297	320297	320297	320297	322638	320297	320297
R2	0.167	0.167	0.167	0.167	0.167	0.167	0.167
ADJ. R2	0.166	0.166	0.166	0.166	0.166	0.166	0.166

Table A5. Regression estimates with double fixed effects(provider fixed effects and country fixed effects).

	(1)	(2)	(3)
LengthOfStay	-0.0203***	-0.0224***	-0.0223***
	[-5.31]	[-6.15]	[-6.12]
Domestic	-0.0179***	-0.0442***	-0.0440***
	[-2.63]	[-6.76]	[-6.73]
Couple	0.0325***	0.0214***	0.0226***
	[3.79]	[2.62]	[2.77]
Family	0.0472***	0.0594***	0.0605***
	[5.07]	[6.71]	[6.83]
Single	0.00287	0.0158	0.0169
	[0.26]	[1.51]	[1.61]
Anonymous	-0.190***	-0.202***	-0.202***
	[-23.59]	[-25.89]	[-25.87]
Lnprice	-0.261***	-0.272***	-0.274***
	[-15.12]	[-16.97]	[-17.12]
Prior_Ratings	-2.127***	-2.056***	-2.056***
	[-27.27]	[-27.82]	[-27.88]
Sunnier_Icon	0.0337***		
	[4.54]		
Diff_Hoursofsun		0.0237*	
		[1.83]	
Diff_Precipitation			-0.00975***
			[-6.83]
Constant	27.88***	27.34***	27.36***
	[41.95]	[43.41]	[43.51]
N	299752	328577	328577
R2	0.144	0.146	0.146
adj. R2	0.144	0.146	0.146

*t statistics in brackets * p <0.10, ** p <0.05, *** p <0.01*

Table A6. Regression estimates with alternative weather variables.

	(1)	(2)	(3)	(4)	(5)
Unexpectedrain	-0.174***				
	[-10.45]				
Unexpectedrain_ext		-0.133***			
		[-12.89]			
Unexpectedsun_ext			0.0934***		
			[5.25]		
Sunny				-0.0396***	-0.0576***
				[-5.24]	[-7.32]
Rainy				-0.107***	0.0144
				[-10.90]	[0.84]
Surprise (+)					0.118***
					[6.35]
Surprise (-)					-0.158***
					[-8.51]
Constant	25.07***	25.30***	25.10***	25.02***	24.90***
	[35.65]	[35.97]	[35.72]	[35.68]	[35.52]
N	206451	206451	206451	207727	207727
CONSUMER FIXED EFFECTS	YES	YES	YES	YES	YES
MONTHLY FIXED EFECTS	YES	YES	YES	YES	YES
R2	0.159	0.159	0.159	0.160	0.161
adj. R2	0.159	0.159	0.158	0.160	0.160

*t statistics in brackets * p <0.10, ** p <0.05, *** p <0.01*

Table A7. Estimates on Venice subset, excluding the week of High-Water episode (10^{th} November 2019)

	(1)	(2)	(3)	(4)	(5)
HW	-0.0261	-0.00757	-0.0149	0.0298	-0.00411
	[-1.25]	[-0.36]	[-0.72]	[1.40]	[-0.19]
Unexpectedrain	-0.174***				
	[-10.45]				
Unexpectedrain_ext		-0.123***			
		[-12.29]			
Unexpectedsun_ext			0.0885***		
			[4.98]		
Sunny				-0.0374***	-0.0535***
				[-4.94]	[-6.80]
Rainy				-0.107***	-0.0118
				[-10.86]	[-0.73]
Surprise (+)					0.111***
					[5.98]
Surprise (-)					-0.123***
					[-7.20]
Constant	24.91***	25.12***	24.94***	24.88***	24.78***
	[35.23]	[35.53]	[35.31]	[35.29]	[35.15]
N	215669	215669	215669	216945	216945
TOURIST FIXED EFFECTS	YES	YES	YES	YES	YES
MONTHLY FIXED EFFECTS	YES	YES	YES	YES	YES
R2	0.158	0.158	0.158	0.159	0.159
adj. R2	0.157	0.158	0.157	0.158	0.159

*t statistics in brackets * p <0.10, ** p <0.05, *** p <0.01*

Table A8. Estimates on Venice subset, controlling for the week of High-Water episode (10^{th} November 2019)

	(1)	(2)
LenghtOfStay	-0.0224***	-0.0630***
	[-6.11]	[-10.38]
Domestic	-0.0364***	-0.0686***
	[-5.53]	[-6.38]
Couple	0.0184**	-0.0900***
	[2.22]	[-6.82]
Family	0.0471***	-0.0447***
	[5.27]	[-3.08]
Single	0.00752	-0.135***
	[0.71]	[-7.62]
Anonymous	-0.197***	-0.200***
	[-24.95]	[-16.05]
Surprise (+)	0.165***	0.147***
	[10.90]	[3.85]
Surprise (-)	-0.0908***	-0.000335
	[-7.45]	[-0.02]
Sunny	-0.0716***	-0.0495***
	[-11.02]	[-4.66]
Rainy	0.00901	0.00990
	[0.80]	[0.57]
Constant	27.35***	30.08***
	[43.15]	[26.18]
N	323026	127048
R2	0.147	0.159
adj. R2	0.147	0.158

Table A9. Estimates on the pooled sample (Column 1) and on weekends (Column2)

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Strada Maggiore 45
40125 Bologna - Italy
Tel. +39 051 2092604
Fax +39 051 2092664
http://www.dse.unibo.it