

Figini, Paolo; Cicognani, Simona; Zirulia, Lorenzo

**Working Paper**

## Booking in the rain: Testing the impact of public information on prices

Quaderni - Working Paper DSE, No. 1137

**Provided in Cooperation with:**

University of Bologna, Department of Economics

*Suggested Citation:* Figini, Paolo; Cicognani, Simona; Zirulia, Lorenzo (2019) : Booking in the rain: Testing the impact of public information on prices, Quaderni - Working Paper DSE, No. 1137, Alma Mater Studiorum - Università di Bologna, Dipartimento di Scienze Economiche (DSE), Bologna, <https://doi.org/10.6092/unibo/amsacta/6265>

This Version is available at:

<https://hdl.handle.net/10419/213534>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-nc/3.0/>



ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna  
DEPARTMENT OF ECONOMICS

**Booking in the Rain:  
Testing the impact of public  
information on prices**

Paolo Figini  
Simona Cicognani  
Lorenzo Zirulia

*Quaderni - Working Paper DSE N°1137*



# **Booking in the Rain: Testing the impact of public information on prices**

Paolo Figini, Department of Economics and CAST – Centre for Advanced Studies on  
Tourism, University of Bologna; TREES – North-West University<sup>1</sup>

Simona Cicognani, Department of Economics, Free University of Bozen

Lorenzo Zirulia, Department of Economics, University of Bologna and ICRIOS, Bocconi  
University

## **Abstract**

The key role played by public information in market economies is paramount. Yet, investigation on how prices are affected is hindered by intricate endogeneity and parametrization issues. This paper contributes to the literature by identifying and studying a case where public information is related to weather forecasts, in a sector (tourism) where meteorological conditions exogenously impact on both suppliers' and consumers' decisions, especially in leisure destinations. We first develop a theoretical model where forecasts affect demand and hence equilibrium prices in accommodation: an improvement in the weather forecast has a positive impact on the price, the effect being stronger the higher the level of accuracy of the forecast and the level of *ex ante* uncertainty in weather conditions. We then estimate an augmented hedonic price model where dynamic characteristics (typical of advanced booking strategies) are included together with information about weather forecasts. Consistent with the theory, we find that, *ceteris paribus*, good weather forecasts are associated to higher prices, the impact being larger the higher their level of accuracy and the higher the *ex ante* level of uncertainty in weather conditions.

**Keywords:** Information uncertainty; Bayesian model; Big Data; Pricing strategy; Exogenous quality; Hotels; Weather forecast; Hedonic price.

**JEL classification:** D80, D22, L83

<sup>1</sup> Corresponding author: E.mail: paolo.figini@unibo.it. The authors acknowledge technical support from Marco Montanari and financial support from CAST, UNIRIMINI and the Fondation du Risque (Université Paris-Dauphine, ENSAE and Groupama): grant "Individuals and Risk: analysis and market response". Thanks to Anastasia Arabadzhyan for her useful comments. The usual disclaimers apply.

## Non-technical summary

The key role played by information in market economies is paramount. When information is public, all parties involved in the transaction can freely access it and evaluate it. Yet, investigation on how prices are affected by public information is difficult because of endogeneity between price and information (e.g. financial markets) and other technical issues. This paper identifies and studies the case of tourism, a sector where public information is related to weather forecasts, known to both suppliers' and consumers' and completely exogenous. The intuition is that when customers purchase through advance booking, they consult the information provided by weather forecasts before taking any decision.

The paper builds on this intuition on both theoretical and empirical sides. To do so, we first develop a theoretical model where forecasts are common knowledge and can be observed by both accommodation service providers and consumers. As the quality of the leisure experience is affected by weather conditions, weather forecasts (as a signal for expected quality) affect the willingness to pay of customers and hence equilibrium prices: an improvement in weather forecasts would increase the expected quality of the experience and drive the supply to instantaneously adapt the price, the effect being stronger the higher the level of accuracy of the forecast and the larger the degree of *ex ante* uncertainty in weather conditions.

We then estimate an augmented hedonic price model where dynamic characteristics (typical of advanced booking strategies) are included together with information about weather forecasts. The model is tested on data for the summer season of a typical sea & sun destination (Rimini, Italy) where most of the leisure activities are carried out *open air*, hence being heavily dependent on weather conditions. A web scraper collected daily prices posted by the population of Rimini hotels on Booking.com, an important booking platform, and daily weather forecasts of Ilmeteo.it, the most popular weather forecast provider for the Italian market.

Consistent with the theory, we find that, *ceteris paribus*, good weather forecasts are associated to higher prices, the impact being larger the higher the level of accuracy of the forecast (i.e., in the last 2-3 days before the due date) and the higher the *ex-ante* level of uncertainty in weather conditions (i.e., in months when rain and other meteorological negative events are more likely).

The paper contributes to a number of related streams of literature, from the one concerning the role of public information on prices to the one analysing the economic impact of weather forecasts, to the one considering the role of weather conditions in leisure activities. Moreover, our analysis also provides interesting welfare and managerial implications. Theoretically, we find that the accuracy of weather forecasts does not affect profits and consumers' surplus. As in this respect many stakeholders have recently blamed the weatherman for the low accuracy of forecasts, which would entail negative effects on the tourism market by keeping tourists away and by reducing profits, our model does not warrant this conclusion, hence warning to disentangle the impact stemming from low accuracy to the one due to bad weather.

## 1. Introduction

The understanding of prices and their dynamics is a key issue in economic analysis, as in market economies the price has the paramount role of allocating resources and conveying information (Hayek, 1945). When information is private, the price system allows the use of dispersed knowledge, with possibly different degrees of efficiency (Grossman and Stiglitz, 1980; Vives, 2014). Prices are also expected to reflect public information, when it concerns external factors affecting demand or supply conditions. At the business level, strategies of price discrimination and dynamic pricing are nowadays very popular to capture idiosyncratic variations in demand and to provide suppliers with complex but effective tools of revenue management optimization (Nocke and Peitz, 2007; Moller and Watanabe, 2010).

Within this general framework, our article aims at assessing the impact of public information on prices, identifying factors that affect the magnitude of the effect. We do this in an environment where not only information is highly relevant for the exchange, but is also completely exogenous: the one of weather forecasts and hotel prices. The quality of the stay in the destination (particularly when considering leisure activities linked to summer holidays) is expected to depend on weather conditions, which impact on demand through price variations (Scott and Lemieux, 2011; Gomez-Martin, 2005; Zirulia, 2016). Weather forecasts better inform economic agents in such markets where there is information uncertainty, as purchasing decisions need to be anticipated through advance booking. Weather forecasts are continuously produced by private and public providers in order to predict real weather conditions. A peculiarity of weather forecasts, which makes them a good terrain to investigate in terms of solving endogeneity concerns, is that they are orthogonal to both suppliers' and consumers' behaviour. Moreover, they are known to both sides of the market, which rules out issues of information asymmetry.

Our contribution is both theoretical and empirical. On the theoretical side, we develop a simple model adopting a Bayesian rational choice approach, where the value of actions stemming from individual decisions depends on realized weather conditions (Katz and Murphy, 1997; Tena and Gómez, 2011; Zirulia, 2016). The model is built to resemble the distinctive features of our setting. A firm faces consumers whose willingness to pay for a good/service depends, among other things, on external factors such as weather conditions. Consumers take their purchasing decisions before the weather state is realized, but after having observed an informative signal, i.e. the weather forecast. The signal is observed by the firm as well, which then fixes the price contingent on the weather forecast. It follows that consumers make their choice observing the weather forecast and the posted price.

Our model predicts that bad weather forecasts have a negative effect on prices, via the impact they exert on demand. In addition, the impact is expected to be larger the higher the forecast's level of accuracy and the higher the *ex ante* level of uncertainty in weather. Finally, from a normative point of view, we show that forecast accuracy has no impact on profit and consumer surplus, a result which hinges upon the symmetry of information and the perfect ability of the firm to respond to signals when setting its price.

The theoretical predictions concerning the impact of public information on prices are then brought to the data for our empirical contribution. To assess the impact of weather forecasts on hotel prices, we estimate a hedonic price model augmented with characteristics related to dynamic pricing strategies and to weather forecasts. This approach allows us to tackle the complexity of markets in which price adjustments are almost instantaneous and where the law of one price remains buried in old textbooks. We collect data through a web scraper and analyse daily prices for the population of hotels in Rimini (Italy), a typical sea & sun summer destination, highly dependent on weather conditions. More than one million observations were collected throughout the whole summer season and then merged with information regarding weather forecasts available at the time when prices were posted.

The hypotheses deriving from the model are supported by the data. We find that weather forecasts are important and independent determinants of the price: *ceteris paribus*, the worse the forecast, the lower the price. As predicted by the model, the impact is larger the higher the forecast's level of accuracy, i.e. when the forecast day is closer to departure date, and in those months in which weather is more uncertain, i.e. the higher the *ex ante* level of uncertainty in weather. Results are robust to several specifications and checks.

Our analysis concerns the impact of public information on prices, and more generally on economic outcomes and behaviour. For this reason, it should be seen as a contribution to a number of related streams of literature. First of all, there is a vast empirical literature which investigates the role of public information and media in various economic settings. In particular, a few studies have analysed the impact of public information on the stock market (Cutler et al, 1989; Mitchell and Mulherin, 1994; Engelberg and Parsons, 2011) and in future markets such as those related to crops and livestock (Dorfman and Karali, 2015). Information by media companies has been shown to affect corporate governance (Dyck et al., 2008) and political behavior (Della Vigna and Kaplan, 2007). Our novel setting presents a few advantages with respect to most of the existing literature: i) the channel through which the information impacts on the economy is easy to identify; ii) the accuracy of information can be measured precisely; iii) information is less biased than in situations where political ideology (Gentzkow and Shapiro, 2010) or conflicts of interest (Gurun and Butler, 2012) are relevant. On the

theory side, our model is related to the literature analyzing the use of information in contexts where agents have access to public and private information, while strategically interacting (Morris and Shin, 2002; Angeletos and Pavan, 2007; Myatt and Wallace, 2015; Myatt and Wallace, 2017). While private information is not relevant in our setting, the result on the impact of information accuracy on welfare in a buyer-seller scenario can be discussed in light of extant literature.

Our work also contributes to the growing literature on the impact of weather on the economy (a recent, comprehensive survey is provided by Dell et al., 2014). Weather can affect the economy in several ways. It can be an exogenous trigger of economic shocks, leading to shortage of goods and then to price increases (Heinen et al., 2018), civil conflicts (Jia, 2014), deteriorated health conditions (Cervellati et al., 2017). It can affect the psychological dimension of decision making, in choices ranging from financial investment (Saunders, 1993; Hirshleifer and Shumway, 2003) to college enrolment (Simonsohn, 2009). It can affect demand indirectly, as it is the case for energy (Mu, 2007).<sup>2</sup> Finally, it can directly influence the quality of the good or service, and therefore demand, as it is case in tourism (Shih et al., 2009; Day et al. 2013; Ridderstaat et al. 2014) or wine production (Ashenfelter, 2008).<sup>3</sup> More closely related to our paper, there is also a (mostly interdisciplinary) literature on the role of the weather forecast in the economy, which studies its overall impact (Katz and Murphy, 1997a) or in specific sectors, such as agriculture (Mjelde and Penson, 2000), energy (Considine et al., 2004) and fishery (Costello et al., 1998). To the best of our knowledge, this study is the first one looking at the impact of weather forecasts on tourism, especially on hotel prices.

The remainder of the article is organized as follows. Section 2 presents the model and derives its main predictions. Section 3 introduces the methodology and the data used in the empirical analysis. Section 4 presents the econometric results, while Section 5 concludes and discusses the general implications of our work, together with its limitations.

<sup>2</sup> An impact of weather conditions on demand is also found by Cellini and Cuccia (2019) for museum attendance. In this case, weather effects are nonlinear and differ across seasons.

<sup>3</sup> Tourism demand can be also affected by weather in the tourists' place of origin, as shown by Agnew and Palutikof (2006) and Rosselló-Nadal et al. (2011).



## 2. The Model

The model is inspired by the literature analyzing Bayesian decision making under weather uncertainty and availability of weather forecasts (Tena and Gómez, 2011; Katz and Murphy, 1997; Zirulia, 2016).

There are two sets of agents: i) consumers and ii) a firm acting as a monopolist in the market. The monopoly structure assumes market power, which characterizes firms offering sufficiently differentiated services such as accommodation, abstracting at the same time from the complication that emerges when competitive environments are considered. Assuming market power is crucial in our analysis, since it allows the firm to react to information through price adjustments. At the same time, ignoring the effect of competition is justified by the nature of shocks, which would affect competing firms in the same way.<sup>4</sup>

Consumers have heterogeneous willingness to pay for the service sold by the firm, and willingness to pay also depends on external factors, such as weather conditions (Candela and Cellini, 1998). Weather forecasts are observed by both firms and consumers; hence, the price set by the firm is contingent on the weather forecast. Consumers make their choice observing the posted price and the weather forecast. The details of the demand and the supply sides of the model are described in Section 2.1 and 2.2, respectively.

### 2.1 Demand

$\theta_i$  denotes the willingness to pay of consumer  $i$  when the weather state is "good".  $\theta$  is uniformly distributed across consumers over the support  $[\underline{\theta}; \bar{\theta}]$ , with  $1 < \bar{\theta} < 2$ ,  $\bar{\theta} - \underline{\theta} = 1$ , and mass normalized to 1. A good weather state may correspond to sunny days for a weekend in a seaside destination, clear sky and warm temperature for an open-air concert, or a sunny day with natural snow in a ski resort. When the weather state is "bad", the willingness to pay is reduced to  $\alpha\theta$  with  $0 < \alpha \leq 1$ . The individual demand function is perfectly inelastic: the consumer either buys the service or not. Defining  $p$  the market price, the consumer net utility function is  $u = \theta - p$ , when the weather state is good, while it is  $u = \alpha\theta - p$ , when the weather state is bad. Her reservation utility is normalized to 0, so that the consumer does not buy when her utility is negative.

The purchasing decision is taken before the weather state is realized, based on probability distribution of weather conditions. Consumers have access to two pieces of information. First,

<sup>4</sup> This claim is supported by the extension of the model to a duopoly regime with vertical differentiation, which yields predictions in line with those reported in this paper. Details are available upon request.



they know the *ex ante* probability of a good state, which is  $\frac{1}{2} \leq r < 1$  (the state is bad with complementary probability).<sup>5</sup> Such a probability is common knowledge among consumers and the firm, and can be interpreted as the historical frequency of "good" and "bad" weather in that particular season. In addition, consumers access the weather forecast, a public signal of the real weather conditions; the signal is correct with a probability  $\frac{1}{2} \leq q < 1$ , and it conveys useful information to consumers, except in the extreme case of random information,  $q = 1/2$ . As a matter of notation, we will use  $g$  and  $b$  to denote the realized good and bad states, " $g$ " and " $b$ " to denote the corresponding signals, and  $p_{g^*}$  and  $p_{b^*}$  the prices conditionals to, respectively, good and bad signals. Therefore, by straightforward application of the Bayes theorem, we obtain:

$$\Pr(b | "b") = \frac{(1-r)q}{(1-r)q+r(1-q)} \quad (1)$$

$$\Pr(g | "b") = \frac{r(1-q)}{(1-r)q+r(1-q)} \quad (2)$$

$$\Pr(g | "g") = \frac{rq}{rq+(1-r)(1-q)} \quad (3)$$

$$\Pr(b | "g") = \frac{(1-r)(1-q)}{rq+(1-r)(1-q)} \quad (4)$$

Given such *ex post* probabilities on weather states, we can write the consumer's expected utility when buying the holiday, conditional on the signals received:

$$E(u | "g") = \frac{rq}{rq+(1-r)(1-q)} \theta + \frac{(1-r)(1-q)}{rq+(1-r)(1-q)} \alpha \theta - p_{g^*} \quad (5)$$

$$E(u | "b") = \frac{(1-r)q}{(1-r)q+r(1-q)} \alpha \theta + \frac{r(1-q)}{(1-r)q+r(1-q)} \theta - p_{b^*} \quad (6)$$

Conditional on " $g$ ", the consumer buys the service when the expected utility is larger than or equal to 0. This is the case if:

$$\frac{rq}{rq+(1-r)q} \theta + \frac{(1-r)(1-q)}{rq+(1-r)(1-q)} \alpha \theta - p_{g^*} \geq 0 \quad (7)$$

<sup>5</sup> Assuming that the good state is *ex ante* more likely than the bad state is theoretically inconsequential, but it simplifies the proof and it is consistent with our empirical application. In particular, the assumption is used only when we bring to the data the model implication on the role of *ex ante* uncertainty.

Condition (7) determines a threshold  $\hat{\theta}_{"g"} \equiv \frac{p[rq+(1-r)(1-q)]}{rq+\alpha(1-r)(1-q)} = p\beta_{"g"}$  such that the consumer buys the holiday if and only if her willingness to pay is larger than or equal to  $\hat{\theta}_{"g"}$ . Similarly, conditional on "b", the consumer buys the holiday when the expected utility is larger than or equal to 0. This is the case if:

$$\frac{(1-r)q}{(1-r)q+r(1-q)}\alpha\theta + \frac{r(1-q)}{(1-r)q+r(1-q)}\theta - p_{"b"} \geq 0 \quad (8)$$

Condition (8) determines a threshold  $\hat{\theta}_{"b"} \equiv \frac{p[(1-r)q+r(1-q)]}{r(1-q)+\alpha(1-r)q} = p\beta_{"b"}$  such that the consumer buys the holiday if and only if her willingness to pay is larger than or equal to  $\hat{\theta}_{"b"}$ .

It can be shown that  $\hat{\theta}_{"g"} \leq \hat{\theta}_{"b"}$ , i.e.  $\beta_{"g"} \leq \beta_{"b"}$ , where the inequality is strict if  $\alpha < 1$ . In fact,

$$\frac{rq+(1-r)(1-q)}{rq+\alpha(1-r)(1-q)} \leq \frac{(1-r)q+r(1-q)}{r(1-q)+\alpha(1-r)q} \quad (9)$$

simplifies to  $r(1-r)(1-\alpha)(1-2q) \leq 0$ , which is always satisfied. Intuitively, a good signal induces to buy even those consumers with a relatively low willingness to pay; or equivalently, demand is less responsive to price when the *ex post* likelihood of good weather increases.

Given  $\theta$  distribution, the demand functions, conditional on the signals, are derived by noting that  $\Pr(\theta \geq \hat{\theta}_{"g"}) = \bar{\theta} - \hat{\theta}_{"g"}$  and  $\Pr(\theta \geq \hat{\theta}_{"b"}) = \bar{\theta} - \hat{\theta}_{"b"}$ :

$$D_{"g"}(p_{"g"}) = \bar{\theta} - \hat{\theta}_{"g"} = \bar{\theta} - \frac{p[rq+(1-r)(1-q)]}{rq+\alpha(1-r)(1-q)} = \bar{\theta} - \beta_{"g"}p_{"g"} \quad (10)$$

$$D_{"b"}(p_{"b"}) = \bar{\theta} - \hat{\theta}_{"b"} = \bar{\theta} - \frac{p[(1-r)q+r(1-q)]}{r(1-q)+\alpha(1-r)q} = \bar{\theta} - \beta_{"b"}p_{"b"} \quad (11)$$

where  $\beta_{"g"}, \beta_{"b"} \in [1, +\infty)$  are the slopes of the linear demand functions.

## 2.2 Supply and market equilibrium

The firm operates at zero marginal and fixed costs, and chooses its prices, in order to maximize profits (or, equivalently, revenues), after having observed the signal. Conditional on good and bad signals, profits are respectively  $\Pi_{"g"} = p_{"g"}D_{"g"}(p_{"g"})$  and  $\Pi_{"b"} = p_{"b"}D_{"b"}(p_{"b"})$ . The first-order conditions for profit maximization are:<sup>6</sup>

<sup>6</sup> Second-order conditions are obviously satisfied.

$$\bar{\theta} - 2\beta_{g^n} p_{g^n} = 0 \quad (12)$$

$$\bar{\theta} - 2\beta_{b^n} p_{b^n} = 0 \quad (13)$$

from which we derive the pair of equilibrium prices,  $p_{g^n}^* = \frac{\bar{\theta}}{2\beta_{g^n}}$ ,  $p_{b^n}^* = \frac{\bar{\theta}}{2\beta_{b^n}}$ , and profits,  $\Pi_{g^n}^* = \frac{\bar{\theta}^2}{4\beta_{g^n}}$  and  $\Pi_{b^n}^* = \frac{\bar{\theta}^2}{4\beta_{b^n}}$ . In equilibrium, demand is equal to  $\bar{\theta}/2$ , irrespectively of the signal.

### 2.3 Discussion and testable implications

The aim of this section is to derive the model implications on the impact of public information on equilibrium prices, which will be then tested in Section 4. Specifically, we look at the impact of information provided by weather forecasters (good *vs* bad signal) both in absolute and percentage terms. By defining  $\Delta p^* \equiv p_{g^n}^* - p_{b^n}^*$  and  $\% \Delta p^* \equiv (p_{g^n}^* - p_{b^n}^*)/p_{b^n}^*$ , straightforward computations yield:

$$\Delta p^* = \frac{\bar{\theta}}{2} \left( \frac{1}{\beta_{g^n}} - \frac{1}{\beta_{b^n}} \right) = \frac{\bar{\theta} r(1-r)(2q-1)(1-\alpha)}{2[rq+(1-r)(1-q)][(1-r)q+r(1-q)]} \quad (14)$$

$$\% \Delta p^* = \frac{\beta_{b^n}}{\beta_{g^n}} - 1 = \frac{[rq+(1-r)(1-q)]}{rq+\alpha(1-r)(1-q)} \frac{[\alpha(1-r)q+r(1-q)]}{r(1-q)+(1-r)q} - 1 \quad (15)$$

From inspection of (14), it is immediate to see that  $\Delta p \geq 0$  (since  $q \geq \frac{1}{2}$ ), from which  $\% \Delta p^* \geq 0$  follows. In other words:

**Proposition 1** *A forecast of good weather has a positive impact on price, both in absolute and percentage terms.*

The result can be interpreted in terms of the elasticity rule, which is typical of monopoly pricing. Demand elasticities are defined as  $\frac{1}{\beta_{g^n} D_{g^n}(p_{g^n})}$  and  $\frac{1}{\beta_{b^n} D_{b^n}(p_{b^n})}$ . In equilibrium, their value is 1 (being marginal cost nil). Since demand is constant at equilibrium, prices are inversely related to the value taken by  $\beta_{g^n}$  and  $\beta_{b^n}$ : signals that increase the ex post likelihood of a good state reduce the sensitiveness of consumers to prices, and the firm reacts by increasing the price.

A second set of predictions derived from the model concerns those factors that can affect the magnitude of  $\Delta p^*$  and  $\% \Delta p^* \geq 0$ . In particular, we look at how the extent of price variation

is mediated by the precision of information, i.e. weather forecasts' accuracy, and by the *ex ante* level of uncertainty. In the model, precision is measured by  $q$ , while uncertainty is inversely related to  $r$  (as long as  $r \geq \frac{1}{2}$ , as we assumed).

**Proposition 2** *The more accurate weather forecasts (i.e. the larger  $q$ ), the larger the impact of a good weather forecast on price, both in absolute and percentage terms.*

**Proposition 3** *The higher the ex ante level of weather uncertainty (i.e. the lower  $r$ ), the larger the impact of a good weather forecast on price, both in absolute and percentage terms.*

Proofs of Propositions 2 and 3 are in Appendix A. The interpretation of Proposition 2 is straightforward: the more accurate the signals, the higher the value of conveyed information. Once again, the effect unfolds on demand elasticity, since more reliable forecasts make demand less elastic when the signal is good (thus leading to a price increase), and more elastic when the signal is bad (thus leading to a reduction in price).

As for Proposition 3, a higher level of *ex ante* uncertainty in weather conditions makes *ex post* distribution more relying on signals, following Bayesian updating. Consequently, a forecast of good weather significantly reduces demand elasticity, and the opposite occurs for a forecast of bad weather.

#### *2.4 Public information accuracy and social welfare*

Although welfare analysis is not the main contribution of the paper, the normative implications of our model can be derived and discussed. Specifically, we investigate the relationship between the accuracy of public information and social welfare, measured by the sum of the firm profit and consumer surplus.<sup>7</sup> This relationship has been the focus of a literature initiated by the work of Morris and Shin (2002), and followed by the work of Angeletos and Pavan (2007), Myatt and Wallace (2015), Myatt and Wallace (2017), among others. Our results can be summarized in Proposition 4 (the proof is in Appendix A).

**Proposition 4** *Both expected profits and consumer surplus are independent from the level of accuracy.*

<sup>7</sup> As regards weather forecast accuracy, Bauer et al. (2015) and Alley et al. (2019) discuss and provide evidence on the great progress obtained in the last decades.

In our model, firms and consumers are in fact in a symmetric situation in terms of access to information.<sup>8</sup> Consumers observe signals, which have an impact on demand function; and the firm, observing the same signal as consumers, *instantaneously* reacts by changing the price responding to the variation in demand elasticity. If the signal is more accurate, the variation of demand is stronger, and the change in price more pronounced (see Proposition 2), while leaving the equilibrium level of demand unaffected.

This result follows the buyer-seller type of interaction that we consider and the symmetric information between the two parties with respect to the signal, together with perfectly flexible prices on the firm's side. As for the first point, in contexts where agents' choices are fully independent, the higher the accuracy of the public signal, the higher the individuals' welfare in equilibrium. On the other hand, Morris and Shin (2002) consider a setting where agents have access to both public and private signals and each agent chooses an action to maximize his own pay-off. The choice is similar to the one of the other players due to strategic complementarities in their actions, in a beauty contest-like situation. The main result of Morris and Shin (2002) is that more accurate public signals can be detrimental to social welfare, because agents may give them too much importance (thus ignoring or downplaying their own private information) in the effort of improving coordination with other agents. In our model, consumers are not interested in guessing (and mimic) what other consumers do. However, there is an agent (i.e. the firm) which is acting in a way that "counteracts" the consumers' actions, thereby cancelling any information advantage and any gain in consumer surplus.

The importance of symmetric information (and price flexibility) is testified by the different results obtained in Zirulia (2016), who considers a model where prices are not contingent on weather forecasts, but a single price is posted *ex ante* (and determined also by the accuracy of weather forecasts). In his setting, the higher accuracy of weather forecasts decreases prices, reduces profits and increases consumers' welfare, because tourists can exploit their superior information when they make their decisions. It follows that forecast accuracy matters for welfare implications only if information is asymmetric, or, more relevant in our case, if prices are "sticky".

<sup>8</sup> We observe that public information does not imply *ex post* symmetric information (even when *a priori* probabilities are identical), as long agents have different abilities in elaborating such information (Green, 2004).

### 3. Data, methodology and hypotheses formulation

#### 3.1 Data

Data have been collected through a scraper in the period June – September 2015, monitoring Rimini, an important Italian sea & sun destination hosting more than 2 million arrivals and 10 million overnight stays per year. This city has been chosen because of its relevance in the tourism sector (more than 800 hotels are in activity) and because in the summer season the city hosts leisure tourists who generally spend the day on the beach. Arguably, they are hence very sensitive to weather conditions (De Freitas, 2015).

Each day the scraper collected information on: i) weather forecasts, from the most popular – commercial - weather website/app; ii) actual weather conditions, from a public archive; iii) hotel prices and other characteristics, posted daily on Booking.com.

[Insert Tables 1, 2 and 3 here]

- I. **Weather forecasts.** The scraper collected weather forecasts published by the most popular commercial website/app in Italy: *ilmeteo.it*, an application with more than 10 million downloads in Google Play Store. Despite many other websites and apps are available, the popularity of this service makes it the most important influencer for weather forecasts in Italy, and its bulletins are regularly reposted and published by national newspapers and popular media. Table 1 compares weather forecasting apps in Italy and their rating figures, showing the typical J-shaped distribution of rating websites (Hu et al., 2009). The scraper collected daily information on forecasts for Rimini for the following 15 days, from May 13<sup>th</sup> to September 27<sup>th</sup>, 2015. Each day, forecasts of minimum and maximum temperature, as well as the literal translation of the icon summarizing the overall weather forecast for the day (sunny/ rainy/ cloudy/ etc.) were collected. Net of the missing forecasts due to unavailability of the scraper or the website, 1942 observations were collected, with the main summary statistics reported in Tables 2 and 3.
- II. **Real weather conditions.** In order to assess the accuracy of forecasts, data on actual weather conditions, as recorded by the public official archive (Arpa-ER, the Regional Agency for Environmental Protection) were also collected. Archived data included the millimeters of rain, air pressure, maximum, minimum and daily average temperature, maximum, minimum and daily average humidity, wind speed and direction. The summary statistics for these variables, for the days from May 15<sup>th</sup> to September 30<sup>th</sup>, 2015 are reported in Table 4.

III. **Hotel prices and other characteristics.** The last part of the database included the hotel prices posted daily on Booking.com, an important booking platform, for the whole population of hotels in Rimini. The scraper searched prices for one-week stays (from Saturday to Saturday) and for short weekend breaks (a two-night stay, from Friday to Sunday) in the period May 29<sup>th</sup>, 2015 (Friday) to September 26<sup>th</sup>, 2015 (Saturday) as check-in dates. A 15 days lead booking period was considered (that is, for each search, prices posted up to 15 days before the check-in date were collected) in order to match prices and weather forecasts. The whole set of available rooms and price conditions on offer were scraped. Names, labels and descriptive statistics of the variables are reported in Table 5.

[Insert Tables 4 and 5 here]

After a complex cross-check, some inconsistent observations were deleted, leaving about 730,000 observations for the statistical and econometric analysis. Further information about the construction of the dataset is in Appendix B.

### 3.2 Methodology and hypotheses formulation

A hedonic price model was estimated where, differently from standard hedonic models in which only the physical quality characteristics of the product are considered, the price was also regressed over a set of “dynamic characteristics”, i.e. characteristics that vary over time (number of available rooms, booking lead time, weather forecasts, etc.), in line with recent empirical literature on hotel dynamic pricing (Abrate et al., 2012; Melis and Piga, 2017; Falk and Vieru, 2019). We run both a linear and a log-linear version of the model, to test hypotheses in terms of both absolute and percentage variations of price.

The general specification of the linear model is the following:

$$P_{i,t,t-\tau} = a + bD_{i,t} + cR_iH_j + dC_{i,t,t-\tau} + eB_{i,t,t-\tau} + fL_{t,t-\tau} + gLOS_{i,t,t-\tau} + hADV_{i,t,t-\tau} + iA_{i,t,t-\tau} + jWF_{t,t-\tau} + \epsilon_{i,t,t-\tau} \quad (16)$$

where:

-  $P_{i,t,t-\tau}$  is the daily price (computed by dividing the posted price over the length of stay) of room  $i$  for check-in date  $t$  posted at time  $t - \tau$ , where  $0 \leq \tau \leq 15$ . In log-linear specifications of the model,  $\ln P$ , the natural logarithm of price, was alternatively used as dependent variable;



-  $D_{i,t}$  is a set of dummies built over the variable *from\_date* reported in Table 5, controlling for the different check-in dates;

-  $R_i$  is a set of dummies built over the variable *room\_newcode*, controlling for the different characteristics of the rooms, such as size, floor, view;

-  $H_j$  is a set of dummies built over the variable *hotel*, controlling for the hotels fixed effects;

-  $C_{i,t,t-\tau}$  is a dummy variable controlling for the possibility of cancelling the booking without penalties;

-  $B_{i,t,t-\tau}$  is a set of dummies built over the variable *breakfast\_code*, controlling for the type of service included (bed only, bed & breakfast, half-board, half-board with lunch, full-board);

-  $L_{t,t-\tau}$  is a numeric variable measuring the distance between the search date and the check-in date, it is then equal to  $t - \tau$  for each observation and controls for the use of dynamic pricing strategies.

-  $LOS_{i,t,t-\tau}$  is a dummy variable which indicates the "length of stay" to control for different price offers for weekends, long weekends and one-week holidays;

-  $ADV_{i,t,t-\tau}$  is a dummy indicating if the price is advertised as a special offer/discount;

-  $A_{i,t,t-\tau}$  is a set of dummies built over the variable *availability*, controlling for the supply constraint, that is, the number of rooms shown as available on the platform at the posted price (from 1 to 5). When the number of available rooms was more than 5, the variable was coded as 0.

-  $WF_{t,t-\tau}$  is the main variable of interest and is a set of dummies built over the variable *fore\_code* reported in Table 3, controlling for the weather forecast for time  $t$  posted at time  $t - \tau$ . In some of the alternative specifications of the model,  $WF_{t,t-\tau}$  was proxied by *norain*, or *sun*, binary variables built by aggregation of the different values of *fore\_code*. More specifically, *norain* (*sun*) takes the value of 1 if *fore\_code* takes the values of 1, 2 or 3 (1 or 2) as indicated in Table 3, and 0 otherwise.<sup>9</sup>

The theoretical propositions derived in Section 2 lead us to formulate the following hypotheses:

**Hypothesis 1.** *Coherently with Proposition 1, we expect that, ceteris paribus, good weather forecasts are associated with higher prices.*

<sup>9</sup> In the sensitivity analysis, alternative proxies for  $WF_{t,t-\tau}$  such as the maximum daily temperature and the average daily temperature were also tested, providing robust results. In this regard, Scott et al. (2007) find that the optimal temperature perceived by tourists, for beach environments, is 27 °C (median), five degrees more than for urban environments, and seven more than for mountain summer environments.

As regards the proxy for weather, in accordance with Scott et al. (2008), we chose the sun/rain indicator (the variable  $WF_{t,t-\tau}$  in our dataset), as it is likely to be the most relevant factor associated to beach activities in a sea & sun destination. Using as reference category the most positive forecast ( $WF_{t,t-\tau} = 1$ ), we expect negative coefficients for the other  $WF_{t,t-\tau}$  categories, with the coefficients being larger in absolute value for more negative forecasts. In the alternative specifications in which *norain* or *sun* dummy variables are used, we expect a positive coefficient associated to these variables.

**Hypothesis 2.** *Coherently with Proposition 2, we expect the impact of weather forecasts to be stronger, the more accurate the forecasts.*

In that respect, scientific evidence (Bauer et al., 2015; Alley et al., 2019) shows that forecast accuracy varies with the temporal distance between the day when the forecast is produced and the day it refers to, i.e. forecast accuracy increases the shorter the time lag of the forecast. Our data confirm this evidence, as later reported in Table 9. It follows that the impact of weather forecasts on prices is expected to be stronger the closer the forecasts to the day of the check-in.<sup>10</sup>

**Hypothesis 3.** *Coherently with Proposition 3, we expect the impact of weather forecasts to be larger when weather uncertainty is higher.*

In our case, a good proxy for weather uncertainty is the period of the year. According to historical data for Rimini, average rainfall is lower in the central part of the summer season (July and August) than in the early and late months (June and September). For instance, the average monthly rainfall for July and August in the 2011-2015 period was 49.4 mm, while it was 74.4 mm for June and September of the same years.<sup>11</sup> It follows that the impact of weather forecasts on prices is expected to be stronger at the margins than in the central part of the season.

<sup>10</sup> Anbarci et al. (2011) show that weather forecast accuracy is affected also by population and income at the local level. Being a study of single destination, this issue is not relevant in our case.

<sup>11</sup> Own elaboration based on official statistics provided by Arpa-ER – Regional Agency for Environmental Protection Emilia Romagna ([simc.arpae.it/dext3r](http://simc.arpae.it/dext3r)). This evidence is robust to the use of the average monthly number of rainy days, according to which there were 4.9 rainy days in July-August and 7.6 in June and September, for the same years.

## 4. Results

### 4.1 Testing Hypothesis 1: the impact of weather forecasts on hotel prices

The results of testing the first hypothesis are reported in Table 6. As recalled in Section 3.2 we regressed the daily price for room  $i$  of hotel  $j$  for the check-in date  $t$  posted at  $t - \tau$  on the weather forecast for  $t$  posted in  $t - \tau$  (after controlling for all the other variables included in equation (16)).

[Insert Table 6 here]

Model (6.1) of Table 6 reports the results for the most consistent subset of the dataset, where the daily price of one double room for party groups of two adults for a weekend break was regressed. This subsample includes about 92,000 observations. Estimations suggest that, *ceteris paribus*, bad weather was significantly associated with lower prices compared to the case of the base category, which is “hot and sunny”, the most positive forecast available in *ilmeteo.it* codification ( $WF_{t,t-\tau} = 1$ ). These findings are hence consistent with Hypothesis 1. As regression (6.1) is linear, the coefficients can be directly interpreted as the price variation between the actual weather forecast and the base “hot and sunny” weather forecast. This “discount” for sub-optimal forecasts ranges between € 3.02 (overcast, code 2) to € 7.20 (downpour, code 4) and to € 7.73 (cloudy, code 5).

All the other variables included in (6.1) are significant at the 1% level and with the expected sign: the free cancellation option ( $C$ ) was associated with a higher price (of about € 3.38 per day), which could be interpreted as the premium paid for buying the option of disattending the contract. The inclusion of breakfast in the offer was valued, on average, € 4.43 ( $B$ , code 1) while the half-board including dinner costs € 17.44 ( $B$ , code 2). Any day closer to the check-in date was associated with a higher price (of about € 1.11, see specifically the coefficient of  $L$ ), coherently with a strategy of increasing prices when the occupancy rates are high and the check-in dates approach. Special offers ( $ADV = 1$ ) do have a heavy impact on price: the average discount is estimated to be € 19.32, more than 20% of the basic price in the regression, represented by the constant (€ 73.83). Finally, the indication that there is only a limited number of available rooms on offer is relevant: the coefficients of  $A_{i,t,t-\tau}$  are positive, suggesting that a constraint from the supply side is associated with a higher price with respect to the base category, in which no signal of limitation in the supply was posted. Coherently with the simple law of demand and supply, the coefficient is the lowest when the available rooms are many ( $A = 5$ , € 2.20) and increasing as long as the available rooms decrease,

reaching the maximum when the number of available rooms signalled is  $A = 1$  (€ 7.57). This is exactly what is expected when the hotel approaches full occupancy.

As regards the other variables included in the regression but omitted from Table 6 for space limitation, the coefficients of the check-in dates dummies  $D$  are all significant and with the expected sign: prices are higher approaching the peak period of late July – mid August, signalling an active policy of seasonal pricing. Moreover, the coefficients of the interaction dummies between room type and hotel ( $R*H$ ) are also statistically significant, signalling that the overall quality of the hotel and the characteristics of the posted room are also relevant in determining the price, as expected. Overall, the dynamic characteristics of pricing strategy are quite important in explaining the posted price, improving the overall explanatory power of the regression (R-squared is around 54%) with respect to standard hedonic models.

Errors in regression (6.1) are robust but assumed to be uncorrelated across groups. As price dynamics is heavily dependent on revenue management strategies run by hotels, it is likely that errors in different time periods and for different rooms in a given hotel may be correlated: failure to control for within-cluster error correlation can generate larger t-statistics and more “optimistic” results. Model (6.2) in Table 6 tackles this issue and encompasses clustered errors at the hotel level. As expected, confidence intervals increase and some of the variables included in the model are no longer significant, namely,  $A$ ,  $C$ , and  $B$ , highlighting that hotels apply a common rule as regards these variables. In particular, the cost of breakfast, which is negligible for most of the structures, is likely to allow hotels to manipulate the inclusion of the breakfast in the posted price as a kind of promotional strategy. What is more important, however, is that the estimated coefficients for weather forecasts in Model (6.2) keep being significant.

Table 6 also reports the results of the log-linear model. In particular, regressions (6.3) and (6.4) include the logarithm of price as dependent variable, respectively with robust standard errors (6.3) and robust standard errors clustered by hotel (6.4). Overall, this specification improves the ability to explain variations in price (R-squared is now 66.6%) and results are in line with the Model (6.1): weather forecasts worse than the baseline “hot and sunny” are associated to lower prices, from 3.9% of rain ( $WF = 7$ ) to 10.3% of downpour ( $WF = 4$ ) forecasts. Given that the results of the log-linear model are very similar to the ones of the linear model, the remaining of the econometric analysis in this section alternatively considers absolute or logarithmic prices according to the immediacy of the coefficients’ interpretation.<sup>12</sup>

<sup>12</sup> All findings are available upon request and results that are not robust with the main findings are always reported in the paper.

Interpretation of individual weather forecasts, when the summary proxy is one of the 8 icons described in Table 3, can be problematic. We then simplified the analysis by clustering the 8 codes into simple binary variables, *norain* and *sun*, as explained above. In Model (6.5) of Table 6, in which *norain* was included in the model, the main result is confirmed: prices are on average 3.5% higher when forecasts exclude rain. The coefficient is very similar also in Model (6.6), in which *sun* was inserted, suggesting that the forecast of a sunny day increased the price by 3.6%.

A similar robustness test was undertaken in Models (6.7) and (6.8) where, respectively, the forecast maximum and average daily temperature were included in the model as proxies for *WF*. All the estimated coefficients for the variables included in the regression and their significance are very similar to the log-linear model in (6.5): every Celsius degree more in the forecast maximum temperature is associated to an increase of 1.76% in the average daily price, *ceteris paribus* (see Model 6.7). Similarly, every Celsius degree more in the forecast mean temperature was associated to an increase of 1.60% in the average daily price (see Model 6.8).

Finally, Models (6.9) and (6.10) in Table 6 are run on an extended dataset, where all the possible rooms (double, triple, quadruple etc.), party groups (two adults with and without two children) and length of stay (weekends and one-week holidays) are considered, with robust errors clustered by hotel. To simplify the reading, weather forecasts are proxied by the binary variable *norain* (results are however robust to the alternative inclusion of *WF* or *sun* in the regression). In Model (6.9) we only consider weekends of party groups with or without children and, consequently, a wider range of rooms, including double, triple and quadruple rooms, together with full apartments and villas (about 230,000 observations). The estimated coefficient of *norain* is very similar to the one of Model (6.5), hence reinforcing the support for Hypothesis 1. In Model (6.10) we include double rooms and parties of two adults only (as in Models 6.1 to 6.8) but both weekends and one-week holidays are considered: although the coefficient of *norain* in (6.10) is still positive and significant, its value is lower (0.0123), suggesting a weakening of the impact. This is very interesting and expected, especially if we consider that in a sea & sun destination in Southern Europe, bad weather usually lasts for one or, in the worst-case scenario, two or three days. Hence, it is unlikely that a bad weather forecast for the check-in day is going to affect the decision of travelling to the destination for a whole week.<sup>13</sup>

<sup>13</sup> In our sample, most of the hotels indeed used some form of dynamic pricing, somehow different from what found by Melis and Piga (2017). However, there were 194,956 observations (18% of the observations with a valid price) where the price of a

#### 4.2 Testing Hypotheses 2 and 3: the role of signal accuracy and ex ante uncertainty

In Table 7, we report the results for testing Hypothesis 2. Models (7.1) and (7.2) can be compared to Model (6.5) in which the logarithmic daily price of one room for two adults for a weekend stay is regressed over *norain* and the standard set of control variables outlined in equation (16). In Model (7.1) only the last three days of the booking period were considered, while in Model (7.2) we included the complementary period, i.e. from 4 up to 15 days in advance. The same comparison was carried out in (7.3) and (7.4), with the price in levels as dependent variable. This split follows common knowledge about the accuracy of forecasting models and is in line with forecasts of the public provider (the Aeronautica Militare in the Italian case), which are produced only over the previous 72 hours.

[Insert Table 7 here]

Results are consistent with Hypothesis 2: the effect of weather forecasts is stronger the shorter the time lag, as the accuracy of the information embodied in the forecasts improves; accordingly, the coefficient of *norain* estimated in (7.1) is larger than the one estimated in (7.2) and the coefficient of *norain* estimated in (7.3) is larger than the one estimated in (7.4) which becomes statistically not significant. Such results were confirmed in regressions (7.5) to (7.8), run on the extended dataset with different party groups (adults with or without children) and different room size (double, triple, quadruple, etc.): the coefficient of *norain* estimated in the loglinear regression (7.5) is larger than the one estimated in (7.6), while the coefficient estimated in the linear regression (7.7) is larger than the one estimated in (7.8), which becomes not significant.

Finally, in Models (7.9) and (7.10) we run models (7.7) and (7.8) but replacing the binary variable *norain* with the weather forecast codes described in Table 3. Although not all the forecast coefficients are significant, it is worthwhile to highlight that the strongest impact of weather forecasts in the last three days before check-in (Model (7.9)) comes from forecasts associated with consistent bad weather (cloudy, overcast), which are likely to continue for the

given room in a given hotel for the same check-in date never changed in the advance booking period. When this group of observations were excluded from the regression the coefficient of *norain* became larger (0.04496, significant at the 1% level) than the coefficient in the equivalent model (6.5), where the estimate was 0.0349, reinforcing the support of the rationale behind the model. Unsurprisingly, the coefficient of *norain* was instead insignificant (and equal to -0.0009) when the regression was run on the subsample of observations where price discrimination was not used.

whole weekend, rather than from forecasts associated with unstable bad weather (thunderstorms), which are likely to last 1-2 hours, only partially ruining a hypothetical weekend.<sup>14</sup>

In order to test Hypothesis 3, we run the hedonic model separately for June and September in Models (8.1) and (8.3): June and September are the months with higher level of weather uncertainty according to historical data for Rimini and for July and August, characterized by lower weather uncertainty (Models (8.2) and (8.4)). Results are reported in Table 8, with the daily price expressed in logarithms (Models (8.1) and (8.2)) and in levels (Models (8.3) and (8.4)). For simplicity, in Models included in Table 8 weather forecast is measured through the inclusion of the binary variable *norain*.

[Insert Table 8 here]

In line with Hypothesis 3 we find that the estimated coefficients of *norain* are higher (and significant) in the period of high uncertainty in both the linear and loglinear regressions (compare the estimated coefficient of *norain* in 8.1 and 8.3, which are higher than the one estimated respectively in 8.2 and 8.4). However, the coefficients of *norain* in the season with low weather uncertainty (July and August) are not significant. This result is arguably stemming from the fact that in the low weather uncertainty period, the frequency of *norain* is particularly low. When the extended dataset is used, including all the rooms and party groups (rather than only parties of two adults searching for one double room as in Models 8.1 to 8.4), results are indeed supportive of Hypothesis 3: in the log-linear regression the coefficient of *norain* is larger in June and September (0.055, significant at the 1%) than in July and August (0.009, this time significant at the 1%).<sup>15</sup>

#### *4.3 Additional results: impact by hotel category and the role of cancellation policy*

In Table 8 we also estimate the hedonic pricing model separately for upper-scale hotels (4- and 5-star hotels, Models 8.6 and 8.8) and low- and mid-scale hotels (1-, 2- and 3-star hotels,

<sup>14</sup> We also observe that some of the controls' coefficients differ between "even" and "odd" numbered Models: in the regressions run on prices posted in the last three days before the check-in date (odd Models), the set of A dummies is seldom significant; moreover, the positive sign of the coefficient of *L* suggests that a sort of last minute pricing is implemented; finally, the most important factor affecting the posted price is always *adv*, i.e. the price posted as a special offer.

<sup>15</sup> Full results are available upon request.



Models 8.5 and 8.7). Models (8.5) to (8.8) show that the response of prices to weather forecasts is larger for upper-scale hotels than for low- and mid-scale hotels, both when the price is expressed in levels (the coefficient of *norain* is larger in Model 8.8 than in 8.7) and in logarithms (the coefficient of *norain* is larger in Model 8.6 than in 8.5). These results are robust to the use of the extended dataset too, where all the rooms and party groups are considered.

To interpret this, we observe that different hotel categories may impact in two ways on the model's rationale. First, they can impact on  $\bar{\theta}$ , which is expected to be higher for customers of upper-scale hotels than for low/mid-scale ones. If that is the case, the impact of weather forecasts should be higher for upper-scale hotels when measured in absolute terms, and independent from scale when the impact is measured in percentage terms. Second, hotel categories can interfere with parameter  $\alpha$ , which measures the utility reduction due to bad weather conditions. Theoretically, from equations (14) and (15) it is immediate to derive that  $\frac{\partial \Delta p^*}{\partial \alpha} < 0$  and  $\frac{\partial \% \Delta p^*}{\partial \alpha} < 0$ , that is, the impact of positive weather forecasts is smaller the higher  $\alpha$ . In other words, when consumers have similar willingness to pay in both states of the world ( $\alpha$  close to 1, i.e. weather conditions are relatively unimportant), weather forecasts have a limited impact on prices. Empirically, it follows that results can be accounted for by a combination of: i) different willingness to pay distribution among customers and ii)  $\alpha$  lower for customers of luxury hotels than for those of low- and mid-scale hotels (the latter condition is required to explain the result in percentage terms). *A priori*, the impact of hotel category on  $\alpha$  is ambiguous since higher quality hotels offer services that both counter-balance possible negative weather conditions (such as spa services), or rely on good weather conditions (swimming pools, open-air bars, etc).<sup>16</sup>

A last and relevant comment relates to the interaction between the posted price, the strategy of dynamic pricing, the offered discounts and the free cancellation option. The joint use of these options is the core of hotels' effort to optimize their revenue management strategy when demand is sub-optimal. Yet, these tools are not substitutes one to each other: the discount targets customers with low willingness to pay (low  $\theta$  in the model outlined in Section 2); the free cancellation option mainly targets customers with high sensibility to weather conditions (low  $\alpha$ ) and possibly high risk aversion (a factor not accounted for in the

<sup>16</sup> As regards the other coefficients, it can be interesting to notice that the full board (breakfast\_code = 4) is mainly offered by budget and mid-scale hotels, while for upper-scale hotels the coefficient is not estimated due to collinearity issues: this is an usual practice in Rimini, where 4-star hotels habitually charge separately and optionally the use of the internal restaurant, while small family-run hotels often offer an all-inclusive package to their guests.

model); the strategy of dynamic pricing follows a pre-determined pattern but can be fine-tuned to reach the hotel's target in terms of occupancy rates. How do these factors interact with weather forecasts?

A few considerations might help answer this question. First, the coefficient of  $ADV$  is always negative and significant, with the discount being marginally larger in the last days before the check-in. Second, the coefficient of  $C$  is never significantly associated to price in the last three days before the check-in (see Models 7.1, 7.3, 7.5, 7.7). Third, in the same Models the coefficient linked to weather forecasts,  $norain$ , is always significant. Fourth, the coefficient of  $C$  is positive and significant in the remaining Models (see Models 7.2, 7.4, 7.6, 7.8) run on the subset of prices posted between 4 and 15 days before the check-in date.

The joint reading of this evidence suggests that in the last 2-3 days before the check-in date the cancellation option is likely not to be effective to tackle lower demand associated with bad but reliable weather forecasts (since forecasts are reliable, weather is likely to be bad: customers will then cancel last minute the reservation and hotel occupancy will suffer). On the contrary, hotels would respond to lower demand associated to a bad weather forecast by decreasing the price posted online, either directly or through the promotion of a special offer, discounting the lower quality of the tourists' experience in case of bad weather.

This interpretation is supported by the frequency of the cancellation option. Overall, 46% of the prices in our sample composed by double rooms available for a weekend has the free cancellation option. This share substantially changes along the lead booking period: in the last three days of advance booking, we report only 5.6% of the prices associated with free cancellation (1197 observations), while from 4 to 7 days we report 48.5% of the sample (13009 observations). Finally, from 8 to 15, we report 62.7% of the sample (30969 observations).

Moreover, when we run a model where  $L$  is transformed in a dummy and then interacted with  $C$ , we find that the coefficient of  $norain$  is still significant (0.044,  $t=6.90$ , higher than the corresponding Model 6.5), prices increase from 15 up to 8 days before the check-in, then decrease around 6-7 days in advance, and then increase again up to the check-in date. The interaction terms show that the free cancellation option costs about 25% more than the refundable price two weeks in advance: it is roughly around 20% up to 5 days in advance, decreases to 13-14% four days before and to less than 10% one day before, with only few free cancellation options available.

We interpret these results by suggesting that the free cancellation option is used to increase bookings during the lead time. This "insurance premium" costs about 15-25% to the customer, it tends to be used to a lower extent by hotels in the last few days of the booking period and,

when it appears to convince undecided customers, it is used as a kind of last-resort discount. This interpretation is confirmed by a further robustness check. If the Model is run only on prices posted with the option of free cancellation ( $C = 1$ ), the coefficient of *norain* is 0.0166 and significant at the 5% level only. Alternatively, if the Model is run only on prices posted without the option of free cancellation ( $C = 0$ ), the coefficient of *norain* is 0.0418 and significant at the 1% level (we recall that the corresponding coefficient in the baseline model run in (6.5) is 0.0349). This corroborates the conclusion that in the last days before the check-in, in case of negative weather forecasts, the policy of lowering the price is more effective than the free cancellation option.

## 5. Conclusions and Discussion

The key role played by public information in market economies is self-evident. Alas, empirical investigation on how it affects prices is hindered by the coexistence of private information, ambiguity on how information might impact the economic variables of interest, and endogeneity issues. This article contributes to the ongoing debate on the role of information in price setting by analysing the impact of weather forecasts (as example of public information, predominant on any source of private information) on prices in tourism, a market where weather conditions (which are clearly exogenous) directly affect demand. To do so, we first develop a theoretical model where forecasts are common knowledge and can be observed by both accommodation service providers and consumers. As the quality of the leisure experience is affected by weather conditions, weather forecasts (as a signal for expected quality) affect the willingness to pay of customers and hence equilibrium prices: an improvement in weather forecasts would increase the expected quality of the experience and drive the supply to instantaneously adapt the price, the effect being stronger the higher the level of accuracy of the forecast and the larger the degree of *ex ante* uncertainty in weather conditions.

Predictions of the theoretical model are then estimated through an augmented hedonic model where dynamic characteristics (typical of advanced booking strategies) are included together with information about weather forecasts and other control variables. The model is tested on data for the summer season of a typical sea & sun destination (Rimini, Italy) where most of the leisure activities are carried out *open air*, hence being heavily dependent on weather conditions. Results robustly support the theory.

Our analysis also provides some welfare implications. Theoretically, the accuracy of weather forecasts is not supposed to affect total welfare, neither profits nor consumer surplus (Proposition 4). In this respect, many stakeholders have recently blamed the weatherman for the low accuracy of forecasts, which would entail negative effects on the tourism market by keeping tourists away and by reducing profits (Scott and Lemieux, 2011). Our model does not warrant this conclusion, based on the effect on price.

However, in order to empirically assess the welfare implication of forecast (in)accuracy, and to disentangle the loss for the hotel sector due to the weather forecast inaccuracy from the loss due to bad weather, we should know the actual number of rooms sold at each price in each day, something that is private information of the hotel. Then, this information should be merged with the corresponding weather forecasts and with real weather conditions. The simple observation of posted prices does not offer any insight on the number of transactions: possibly, when forecasts are bad, hotels do not sell rooms, while they might sell them the

next day, at a different price, should the forecasts improve. Similarly, the correctness of the forecast in signaling real weather conditions depends on its accuracy: if the forecast is inaccurate the transaction would occur at a “wrong” price. In this respect, we should also consider the reverse scenario, when the forecast is wrong but positive (sunny forecast, matched with actual rain). In those cases, the hotel is “gaining” by quoting a price higher than the optimal, since it is not discounting the rain.

A complete analysis of forecast accuracy is beyond the scope of this work; however, data for the case under investigation show a level of accuracy around 80-85% and provide evidence of a “wet-bias” effect (i.e. weather forecasters usually report a higher probability of rain than the “correct” one stemming from meteorological models), consistent with previous literature (Silver, 2012). Table 9 shows that the match between forecast of rain and actual rainy weather is lower than the match between forecasted sunny and actual sunny weather. By multiplying the coefficient of *norain* estimated in the baseline models of Table 6 with the number of rooms sold at the “wrong” price when the forecast did not match the real weather, we could compute the net loss due to weather forecast inaccuracy. Our estimates suggest that this figure is likely to be negligible. Access to companies’ revenue management data would allow to shed light on this issue.

[Insert Table 9 here]

Our work has a number of limitations, which also constitute the main avenues for future research on the topic. For instance, the theoretical approach could be extended to a more general scenario where the two signals of good and bad weather have different degrees of accuracy, in line with the wet-bias evidence. Also, including a behavioural dimension seems a promising direction, since there is evidence that individuals have problems in interpreting weather forecast probability in the correct way (Gigerenzer et al., 2005; Juanchich and Sirota, 2016).

On the empirical side, the first issue concerns replicability, as our results are based on one single destination and one single summer. More in general, different destinations might respond to weather conditions in different ways. We expect that the relevance of weather would differ according to the tourism type: cultural tourism is probably less sensitive to weather than leisure tourism (in terms of model set-up,  $\alpha$  would be higher for cultural tourism). Therefore, the price elasticity to weather forecasts would be lower in cultural destinations. Moreover, the computation of the sector’s net loss in terms of profits and welfare due to bad weather and bad forecasts needs access to microdata of hotels, something that is

worthy to be investigated given the important policy implications at the destination level. Finally, in an era of climate alterations, weather conditions change, thus implying more variability within the season and higher frequency of extreme events (Dell et al., 2014). These changes are likely to affect both the accuracy of weather forecasts and the willingness to pay of customers for more risky holiday breaks. Hence, a future extension of the empirical analysis will have to monitor changes in consumer behavior and innovation in the implementation of pricing strategies by hotels.

## Appendix A – Proofs

### Proof of Proposition 2

The impact of  $q$  on prices is inversely related to the impact of  $q$  on  $\beta_{g''}$  and  $\beta_{b''}$ . Straightforward derivations yield:

$$\frac{d\beta_{g''}}{dq} = \frac{(2r-1)[rq+\alpha(1-r)(1-q)]-[r-\alpha(1-r)][rq+(1-r)(1-q)]}{[rq+\alpha(1-r)(1-q)]^2} \quad (A1)$$

The denominator in (A1) is positive. The numerator is 0 for  $\alpha = 1$  and it is increasing in  $\alpha$ . Therefore  $\frac{d\beta_{g''}}{dq} \leq 0$ , i.e.  $p_{g''}^*$  is increasing in  $q$ .

Similarly,

$$\frac{d\beta_{b''}}{dq} = \frac{-(2r-1)[r(1-q)+\alpha(1-r)q]+[r-\alpha(1-r)][r(1-q)+(1-r)q]}{[r(1-q)+\alpha(1-r)q]^2} \quad (A2)$$

The denominator in (A2) is positive. The numerator is 0 for  $\alpha = 1$  and it is decreasing in  $\alpha$ . Therefore  $\frac{d\beta_{b''}}{dq} \geq 0$ , i.e.  $p_{b''}^*$  is decreasing in  $q$ . It follows that  $\Delta p^*$  and  $\% \Delta p^*$  are both increasing in  $q$ .

### Proof of Proposition 3

By deriving  $\Delta p^*$  with respect to  $r$ , after some manipulations one obtains:

$$\frac{d\Delta p^*}{dr} = \frac{\bar{\theta} (2q-1)(1-\alpha)}{2} \frac{(1-2r)[rq+\alpha(1-r)(1-q)][r(1-q)+\alpha(1-r)q]-r(1-r)(2r-1)(2q-1)}{[rq+\alpha(1-r)(1-q)]^2[r(1-q)+\alpha(1-r)q]^2} < 0 \quad (A3)$$

which is negative since  $r \geq \frac{1}{2}$  and  $q \geq \frac{1}{2}$ .

Similarly, deriving  $\% \Delta p^*$  with respect to  $r$ , yields, after lengthy manipulations:

$$\frac{d\% \Delta p^*}{dr} = - \frac{(2q-1)^2(1-\alpha)q(1-q)[r^2-\alpha(1-r)^2]}{[rq+(1-r)(1-q)]^2[r(1-q)+\alpha(1-r)q]^2} < 0 \quad (A4)$$

which is negative since  $r \geq \frac{1}{2}$ .

### Proof of Proposition 4

Firm's expected profit is given by:



$$E\Pi = \Pr("g") \frac{\bar{\theta}}{4\beta_{"g"}} + \Pr("b") \frac{\bar{\theta}}{4\beta_{"b"}} = \frac{\bar{\theta}}{4} [r + (1 - r)\alpha] \quad (\text{A5})$$

which is independent from  $q$ .

As for expected consumer surplus, this is given by:

$$\begin{aligned} ECS = \Pr("g") \int_{\frac{\bar{\theta}}{2}}^{\bar{\theta}} [\Pr(g| "g") \theta + \Pr(b| "g") \alpha \theta - p_{"g"}] d\theta \\ + \Pr("b") \int_{\frac{\bar{\theta}}{2}}^{\bar{\theta}} [\Pr(g| "b") \theta + \Pr(b| "b") \alpha \theta - p_{"b"}] d\theta \end{aligned} \quad (\text{A6})$$

Solving the integrals yields:

$$ECS = \frac{3}{8} \bar{\theta}^2 \left(1 - \frac{\bar{\theta}}{2}\right) [r + \alpha(1 - r)]$$

which is also independent from  $q$ .

## **Appendix B – Technical issues related to the construction of the data set.**

**Party Groups.** Since Rimini is a popular destination for both couples and families, searches for party groups of two adults only, and for two adults with two children were collected each day for each type of stay (weekend and one-week holiday).

**Coding of Rooms.** The scraper searched for all the types of rooms on offer. This was necessary to avoid bias in the estimation due to the alternative availability or unavailability of certain rooms along the lead booking period and along the season. Consequently, in order to track the correct price changes over time, we had to control for the different types of rooms that each hotel was offering. Each room was hence manually recoded according to two dimensions: size and category. In the final dataset, the room code was composed by two digits. The first digit indicates the size of the room: 1\* indicates a single room, 2\* a double room, 3\* a triple room, 4\* a quadruple room, 5\* a twin double room; 6\* an apartment; 7\* a suite, 8\* a beach package, 9\* a bed in a dormitory. The second digit indicates the category: \*0 means that no further information about the type of room was provided, \*1 indicates that the room was economy / basic / budget, \*2 that the room was standard / classic / business / comfort, \*3 that the room was standard / classic / business / comfort but with view / balcony / access spa / access beach, \*4 indicates that the room was superior / deluxe / executive / premium / king, \*5 that the room was superior / deluxe / executive / premium / king but with view / balcony / access spa / access beach. For example, a superior double room with balcony was coded as 25, the same code of a room defined deluxe double room with sea view. This coding allowed us to exploit as much information as possible by normalizing for the type of room, and to compare prices offered by different hotels, which often name rooms differently.

**Data cleaning.** After data cleaning and checking (some observations were deleted because of errors in the scraping or due to missing relevant information), the dataset of almost 1.1 million observations and 23 variables related to 874 hotels was reduced to about 833,000 observations and 17 relevant variables. In fact, some data were lost because of the frequent changes in the structure of the booking platform, which implied the re-programming of the scraper, an action that in some cases took more than one day.

## References

- Alley, R. B., Emanuel, K. A., and Zhang, F. (2019). Advances in weather prediction. *Science*, 363(6425), 342-344.
- Agnew, M. D., and Palutikof, J. P. (2006). Impacts of short-term climate variability in the UK on demand for domestic and international tourism. *Climate Research*, 31(1), 109-120.
- Anbarci, N., Boyd III, J., Floehr, E., Lee, J., and Song, J. J. (2011). Population and income sensitivity of private and public weather forecasting. *Regional Science and Urban Economics*, 41(2), 124-133.
- Anderson, R. W., Johnson, N. D., and Koyama, M. (2016). Jewish persecutions and weather shocks: 1100–1800. *The Economic Journal*, 127(602), 924-958.
- Angeletos, G. M., and Pavan, A. (2007). Efficient use of information and social value of information. *Econometrica*, 75(4), 1103-1142.
- Ashenfelter, O. (2008). Predicting the quality and prices of Bordeaux wine. *The Economic Journal*, 118(529), F174-F184.
- Bauer, P., Thorpe, A., and Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, 525(7567), 47.
- Candela, G., Cellini, R. (1998). "I mercati dei beni di qualità esogena" [Market for exogenous quality goods], *Politica economica*, 14(2):217-244.
- Cellini, R., & Cuccia, T. (2019). Weather conditions and museum attendance: a case-study from Sicily. *Climatic Change*, 1-17.
- Cervellati, M., Sunde, U., and Valmori, S. (2017). Pathogens, weather shocks and civil conflicts. *The Economic Journal*, 127(607), 2581-2616.
- Considine, T.J., Jablonowski C., Posner B., et al. (2004). The value of hurricane forecasts to oil and gas producers in the Gulf of Mexico. *Journal of Applied Meteorology*, 43(9): 1270–1281.
- Costello, C.J., Adams, R.M. and Polasky, S. (1998). The value of El Nino forecasts in the management of salmon: a stochastic dynamic assessment. *American Journal of Agricultural Economics*, 80(4): 765–777.
- Cutler, D. M., Poterba, J. M., and Summers, L. H. (1989). What moves the stock market. *Journal of Portfolio Management*, 15(3), 4-11.
- Day, J., Chin, N., Sydnor, S., and Cherkauer, K. (2013). Weather, climate, and tourism performance: A quantitative analysis. *Tourism Management Perspectives*, 5, 51-56.
- Dell, M., Jones, B. F., and Olken, B. A. (2014). What do we learn from the weather? The new climate-economy literature. *Journal of Economic Literature*, 52(3), 740-98.

- Della Vigna, S., and Kaplan, E. (2007). The Fox News effect: Media bias and voting. *The Quarterly Journal of Economics*, 122(3), 1187-1234.
- De Freitas, C. R. (2015). Weather and place-based human behavior: recreational preferences and sensitivity. *International Journal of Biometeorology*, 59(1), 55-63.
- Dyck, A., Volchkova, N., and Zingales, L. (2008). The corporate governance role of the media: Evidence from Russia. *The Journal of Finance*, 63(3), 1093-1135.
- Dorfman, J. H., and Karali, B. (2015). A Nonparametric Search for Information Effects from USDA Reports. *Journal of Agricultural and Resource Economics*, 124-143.
- Engelberg, J. E., and Parsons, C. A. (2011). The causal impact of media in financial markets. *The Journal of Finance*, 66(1), 67-97.
- Falk, M., and Vieru, M. (2019). Myth of early booking gains. *Journal of Revenue and Pricing Management*, 18(1), 52-64.
- Gentzkow, M., and Shapiro, J. M. (2010). What drives media slant? Evidence from US daily newspapers. *Econometrica*, 78(1), 35-71.
- Gigerenzer, G., Hertwig, R., Van Den Broek, E., Fasolo, B., and Katsikopoulos, K. V. (2005). "A 30% chance of rain tomorrow": How does the public understand probabilistic weather forecasts?. *Risk Analysis: An International Journal*, 25(3), 623-629.
- Gomez-Martin, B. (2005). Weather, climate and tourism: a geographical perspective. *Annals of Tourism Research*, 32(3): 571–591.
- Green, T. C. (2004). Economic news and the impact of trading on bond prices. *The Journal of Finance*, 59(3), 1201-1233.
- Grossman, S. J., and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, 70(3), 393-408.
- Gurun, U. G., and Butler, A. W. (2012). Don't believe the hype: Local media slant, local advertising, and firm value. *The Journal of Finance*, 67(2), 561-598.
- Hayek, F. A. (1945). The use of knowledge in society. *American Economic Review*, 35(4), 519-530.
- Heinen, A., Khadan, J., and Strobl, E. (2018). The Price Impact of Extreme Weather in Developing Countries. *The Economic Journal*, 129(619), 1327-1342.
- Hirshleifer, D., and Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *The Journal of Finance*, 58(3), 1009-1032.
- Hu, N., Zhang, J., and Pavlou, P. A. (2009). Overcoming the J-shaped distribution of product reviews. *Communications of the ACM*, 52(10), 144–147.
- Juanchich, M., and Sirota, M. (2016). How to improve people's interpretation of probabilities of precipitation. *Journal of Risk Research*, 19(3), 388-404.

- Katz, R.W. and Murphy, A.H. (eds) (1997). *Economic Value of Weather and Climate Forecasts*. Cambridge: Cambridge University Press.
- Melis, G., and Piga, C. A. (2017). Are all online hotel prices created dynamic? An empirical assessment. *International Journal of Hospitality Management*, 67, 163-173.
- Mitchell, M. L., and Mulherin, J. H. (1994). The impact of public information on the stock market. *The Journal of Finance*, 49(3), 923-950.
- Mjelde, J.W. and Penson, J.B. (2000). Dynamic aspects of the impact of the use of perfect climate forecasts in the Corn Belt region. *Journal of Applied Meteorology*, 39(1): 67–79.
- Myatt, D. P., and Wallace, C. (2015). Cournot competition and the social value of information. *Journal of Economic Theory*, 158, 466-506.
- Myatt, D. P., and Wallace, C. (2017). Information Use and Acquisition In Price-Setting Oligopolies. *The Economic Journal*, 128(609), 845-886.
- Möller, M., and Watanabe, M. (2010). Advance purchase discounts versus clearance sales. *The Economic Journal*, 120(547), 1125-1148.
- Morris, S., and Shin, H. S. (2002). Social value of public information. *American Economic Review*, 92(5), 1521-1534.
- Mu, X. (2007). Weather, storage, and natural gas price dynamics: Fundamentals and volatility. *Energy Economics*, 29(1), 46-63.
- Nelson, R.R. and Winter, S.G. (1964). A case study in the economics of information and coordination: the weather forecasting system. *Quarterly Journal of Economics*, 78(3): 420–441.
- Nocke, V., and Peitz, M. (2007). A theory of clearance sales. *The Economic Journal*, 117(522), 964-990.
- Ridderstaat, J., Oduber, M., Croes, R., Nijkamp, P., and Martens, P. (2014). Impacts of seasonal patterns of climate on recurrent fluctuations in tourism demand: Evidence from Aruba. *Tourism Management*, 41, 245-256.
- Rosselló-Nadal, J., Riera-Font, A., and Cárdenas, V. (2011). The impact of weather variability on British outbound flows. *Climatic change*, 105(1-2), 281-292.
- Saunders, E. M. (1993). Stock prices and Wall Street weather. *The American Economic Review*, 83(5), 1337-1345.
- Scott, D., & Lemieux, C. (2011). Weather and climate information for tourism. *Procedia Environmental Sciences*, 1, 146–183.
- Scott, D., Jones, B., and Konopek, J. (2007). Implications of climate and environmental change for nature-based tourism in the Canadian Rocky Mountains: A case study of Waterton Lakes National Park. *Tourism Management*, 28, 570–579.

- Scott, D., Jones, B., and Konopek, J. (2008). Exploring the impact of climate-induced environmental changes on future visitation to Canada's Rocky Mountain National Parks. *Tourism Review International*, 12, 43-56.
- Shih, C., Nicholls, S., and Holecek, D. F. (2009). Impact of weather on downhill ski lift ticket sales. *Journal of Travel Research*, 47(3), 359-372.
- Silver, N. (2012). The Weatherman is not a moron. *New York Times*, September the 7<sup>th</sup>, 2017.
- Simonsohn, U. (2009). Weather to go to college. *The Economic Journal*, 120(543), 270-280. 2012.
- Tena, E. C., and Gómez, S. Q. (2011). Economic value of weather forecasting: the role of risk aversion. *Top*, 19(1), 130-149.
- Vives, X. (2014). On the possibility of informationally efficient markets. *Journal of the European Economic Association*, 12(5), 1200-1239.
- Jia, R. (2014). Weather shocks, sweet potatoes and peasant revolts in historical China. *The Economic Journal*, 124(575), 92-118.
- Zirulia, L. (2016). "Should I stay or should I go?": Weather forecasts and the economics of 'short breaks'. *Tourism Economics*, 22 (4),837-846.

## Tables and Figures

**Table 1. Popularity and ratings of the most important weather forecast apps used in Italy**

Name	IImeteo.it	Meteo.it	3Bmeteo	MeteoAM
<b>Nr. of downloads (Play Store)</b>	> 10 million	> 5 million	> 1 million	> 0.5 million
<b>Average rating</b>	4.3	4.0	4.3	4.1
<b>Number of ratings</b>	460187	101475	141262	5074
<b>Share of 1/5 votes</b>	3.61%	8.91%	2.27%	6.90%
<b>Share of 2/5 votes</b>	2.34%	4.72%	1.64%	5.66%
<b>Share of 3/5 votes</b>	7.93%	12.56%	6.87%	11.17%
<b>Share of 4/5 votes</b>	33.63%	32.13%	39.53%	27.36%
<b>Share of 5/5 votes</b>	52,59%	41.68%	49.70%	48.92%

Notes: data were updated on April 10<sup>th</sup>, 2018.

**Table 2. Summary statistics for IIMeteo weather forecasts**

Variable	Obs	Mean	Std. Dev.	Min	Max
<b>Dump time</b>	1942			13/05/2015	27/09/2015
<b>Forecast date</b>	1942			15/05/2015	30/09/2015
<b>Min temp.</b>	1942	21.03	3.28	11	27
<b>Average temp.</b>	1942	24.94	2.68	13.5	31.5
<b>Max temp.</b>	1942	27.72	3.64	15	37
<b>Forecast code*</b>	1942	5.76	3.06	1	8
<b>Search lag time**</b>	1942	6.91	4.32	0	15

\* Codes relative to the summary weather conditions, as described in Table 3.

\*\* Difference (in days) between the day when the forecast was produced and the forecast day.

**Table 3. Weather forecasts produced by IIMeteo for Rimini, Summer 2015**

<b>Forecast icon (and code)</b>	<b>Frequency</b>	<b>Share</b>	<b>Cumulative share</b>
Hot and sunny ( <i>fore_code</i> = 1)	279	14.37	14.37
Blue sky ( <i>fore_code</i> = 2)	185	9.47	23.84
Overcast / unstable ( <i>fore_code</i> = 3)	544	28.01	51.91
Downpour / clear up ( <i>fore_code</i> = 4)	397	20.44	72.35
Cloudy ( <i>fore_code</i> = 5)	119	6.12	78.48
Misty rain ( <i>fore_code</i> = 6)	14	0.72	79.20
Rain ( <i>fore_code</i> = 7)	284	14.62	93.82
Shower / thunderstorm ( <i>fore_code</i> = 8)	120	6.18	100.00
<b>Total</b>	<b>1942</b>	<b>100.00</b>	<b>100.00</b>

**Table 4. Summary statistics for real weather, Rimini Summer 2015**

<b>Variable</b>	<b>Obs</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Min</b>	<b>Max</b>
Daily rain (mm)	138	2.53	7.99	0	48.6
Daily pressure (bar)	138	1013.22	4.18	1004.35	1023.63
Daily max temp. (C°)	138	27.03	4.21	15	36.3
Daily aver. temp. (C°)	138	23.42	3.82	13.18	30.53
Daily min temp. (C°)	138	19.26	3.67	10.9	26.1
Daily max hum. (%)	138	79.11	9.90	52	99
Daily mean hum. (%)	138	61.26	9.51	36.67	90.96
Daily min hum. (%)	138	41.35	10.55	19	76
Daily wind direction*	138	4.14	2.05	1	8
Daily wind speed (m/s)	138	2.17	.58	1.4	5.08

\* Coded clockwise from 1 (northern) to 8 (north-western)



**Table 5. Price and booking variables**

<b>Variable name</b>	<b>Description</b>	<b>Nr. Observations</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Min</b>	<b>Max</b>
price	Total price of booking	833543	486.56	512.49	0	14000
P	Daily price of booking	833543	103.96	87.04	0	2800
hotel_code	Hotel code	996919	460.87	243.29	1	879
ADV	Dummy: is it a special offer?	996920	.256	.436	0	1
score	Hotel's rating	890217	7.79	.923	3.5	9.8
stars	Category of the hotel	870935	2.97	.680	1	5
availability	Nr. of available rooms on the platform	784451	1.76	1.07	1	5
adults	Nr. of adults to accommodate	990160	2	0	2	2
Kids	Nr. of children to accommodate	990160	.276	.690	0	2
from_date	Date of check in	998562			29/05/2015	25/09/2015
to_date	Date of check out	996920			31/05/2015	27/09/2015
LOS	Length of stay	996920	4.65	2.48	2	7
C	Dummy: is there free cancellation?	996920	.354	.478	0	1
score_code*	Encode of the overall Booking evaluation	890217	6.39	1.17	1	8
breakfast_code**	Type of board	996920	1.03	.211	0	4
search_lagtime	Diff. between check in and search dates	996920	6.81	5.40	0	15
room_newcode***	Type of room	747025	32.09	18.97	10	90

\* Categorical variable coding the overall evaluation of the hotel in Booking.com

\*\* Categorical variable coding the type of board provided in the offer: 1 = breakfast not included; 2 = bed & breakfast; 3 = half-board; 4 = full-board.

\*\*\* Categorical variable coding the type of room in terms of capacity (e.g. double room, triple room, etc.) and in terms of quality (e.g. Superior room, room with sea view, etc.).

	(6.1)	(6.2)	(6.3)	(6.4)	(6.5)	(6.6)	(6.7)	(6.8)	(6.9)	(6.10)
	<i>P</i>	<i>P</i>	<i>lnP</i>	<i>lnP</i>	<i>lnP</i>	<i>lnP</i>	<i>lnP</i>	<i>lnP</i>	<i>lnP</i>	<i>lnP</i>
R*H	YES	YES	YES	YES	YES	YES	YES	YES		
H									YES	YES
R									YES	YES
D	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
C	3.384***[5.21]	3.384 [1.29]	0.0344***[9.88]	0.0344**[2.53]	0.0304**[2.24]	0.0326**[2.41]	0.0284**[2.10]	0.0338**[2.49]	0.0227***[10.21]	0.0267**[2.23]
1.B	4.426***[4.09]	4.426 [0.58]	0.0797***[8.11]	0.0797*[1.66]	0.0757 [1.58]	0.0759 [1.59]	0.0763 [1.61]	0.0804*[1.69]	0.0389***[6.07]	0.0262 [0.57]
2.B	17.44***[9.42]	17.44**[1.99]	0.293***[19.32]	0.293***[4.18]	0.290***[4.11]	0.290***[4.10]	0.293***[4.24]	0.297***[4.24]	0.281***[25.23]	0.183***[3.04]
3.B	16.78***[9.71]	16.78*[1.72]	0.257***[15.50]	0.257***[3.17]	0.253***[3.13]	0.254***[3.13]	0.250***[3.07]	0.253***[3.17]	0.322***[23.01]	0.260***[2.72]
4.B	23.09***[6.92]	23.09**[2.20]	0.398***[10.99]	0.398***[7.09]	0.380***[6.80]	0.371***[6.63]	0.389***[7.07]	0.385***[6.95]	0.200***[5.86]	0.178***[3.58]
L	-1.109***[-15.49]	-1.109***[-4.06]	-0.00663***[-16.39]	-0.00663***[-5.46]	-0.00701***[-5.98]	-0.00739***[-6.31]	-0.00541***[-4.59]	-0.00836***[-7.18]	-0.00574***[-23.57]	-0.00362***[-3.95]
ADV	-19.32***[-39.59]	-19.32***[-6.64]	-0.195***[-52.56]	-0.195***[-11.02]	-0.190***[-10.93]	-0.190***[-10.91]	-0.185***[-10.58]	-0.187***[-10.72]	-0.199***[-86.66]	-0.187***[-12.74]
1.A	7.573***[9.79]	7.573*[1.76]	0.0940***[15.99]	0.0940***[2.84]	0.0929***[2.81]	0.0930***[2.82]	0.0941***[2.85]	0.0935***[2.84]	0.128***[27.17]	0.0694**[2.25]
2.A	6.522***[8.68]	6.522 [1.53]	0.0887***[15.21]	0.0887***[2.76]	0.0883***[2.74]	0.0880***[2.74]	0.0883***[2.75]	0.0887***[2.76]	0.109***[23.27]	0.0779**[2.56]
3.A	2.842***[3.66]	2.842 [0.70]	0.0485***[7.87]	0.0485 [1.51]	0.0476 [1.49]	0.0479 [1.50]	0.0476 [1.49]	0.0477 [1.50]	0.0885***[18.28]	0.0578**[1.98]
4.A	2.891***[3.54]	2.891 [0.67]	0.0531***[7.60]	0.0531 [1.49]	0.0522 [1.45]	0.0512 [1.44]	0.0529 [1.48]	0.0521 [1.46]	0.0921***[18.03]	0.0362 [1.00]
5.A	2.201***[2.69]	2.201 [0.54]	0.0502***[7.21]	0.0502 [1.55]	0.0502 [1.55]	0.0496 [1.53]	0.0502 [1.55]	0.0501 [1.55]	0.0515***[9.50]	0.0446 [1.34]
2.WF	-4.080***[-4.81]	-4.080**[-2.49]	-0.0463***[-9.12]	-0.0463***[-4.94]						
3.WF	-3.022***[-3.89]	-3.022**[-2.34]	-0.0566***[-12.29]	-0.0566***[-6.65]						
4.WF	-7.196***[-7.97]	-7.196***[-4.14]	-0.109***[-18.74]	-0.109***[-10.31]						
5.WF	-7.725***[-4.61]	-7.725***[-2.64]	-0.0785***[-10.90]	-0.0785***[-5.79]						
7.WF	-3.395***[-3.32]	-3.395**[-2.18]	-0.0393***[-6.54]	-0.0393***[-4.22]						
8.WF	-6.823***[-3.74]	-6.823**[-2.22]	-0.0449***[-4.92]	-0.0449***[-3.00]						
norain					0.0349***[6.74]				0.0370***[19.27]	0.0124***[3.06]
sun						0.0362***[6.45]				
t_max							0.0176***[8.97]			
t_mean								0.0160***[6.32]		
kids									0.0881***[27.01]	
LOS										-0.0452 [-1.48]

_cons	73.83***[15.03]	73.83***[9.12]	3.962***[79.20]	3.962***[63.29]	3.907***[63.29]	3.910***[63.98]	3.516***[44.51]	3.625***[45.95]	3.555***[137.06]	4.097***[36.60]
<i>N</i>	92137	92137	92137	92137	92137	92137	92137	92137	230092	187505
R2	539	539	657	657	656	656	658	656	639	659
adj. R2	536	536	655	655	653	653	655	654	638	658

Notes: t-statistics in brackets; \* p < 0.10; \*\* p < 0.05; \*\*\* p < 0.01. Robust errors in all regressions. In models (6.2) and (6.4) to (6.10) errors are clustered by hotel. All models except (6.9) include only observations related to a double room for two adults; Model (6.9) includes any combination of rooms for a party group of two adults and two kids. All models except (6.10) include only observations related to a weekend stay of two nights (from Friday to Sunday); Model (6.10) includes all observations related to both weekends and full weeks (from Saturday to Saturday).

**Table 6. Impact of weather forecasts on prices**

	(7.1)	(7.2)	(7.3)	(7.4)	(7.5)	(7.6)	(7.7)	(7.8)	(7.9)	(7.10)
	<i>lnP</i>	<i>lnP</i>	<i>P</i>	<i>P</i>	<i>lnP</i>	<i>lnP</i>	<i>P</i>	<i>P</i>	<i>P</i>	<i>P</i>
R*H	YES	YES	YES	YES						
H					YES	YES	YES	YES	YES	YES
R					YES	YES	YES	YES	YES	YES
D	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
C	-0.0313 [-1.06]	0.0361*[1.85]	-3.492 [-0.72]	3.218 [0.95]	-0.000969 [-0.12]	0.0279***[9.81]	-1.182 [-0.95]	3.080***[6.17]	-1.017 [-0.82]	3.435***[6.96]
1.B	-0.0300 [-0.44]	0.102**[2.05]	-8.176 [-0.94]	7.151 [0.91]	0.0304**[2.27]	0.0414***[5.61]	19.30***[5.92]	8.160***[6.87]	19.38***[5.94]	8.471***[7.13]
2.B	0.170**[2.21]	0.309***[4.10]	-6.528 [-0.75]	20.54**[2.22]	0.200***[7.93]	0.292***[22.96]	19.79***[5.13]	39.91***[12.57]	19.50***[5.05]	39.97***[12.58]
3.B	0.316***[3.96]	0.231***[3.25]	6.065 [0.53]	16.90*[1.77]	0.381***[13.88]	0.300***[17.96]	33.92***[7.05]	34.21***[16.23]	33.86***[7.07]	34.89***[16.57]
4.B	0.0347 [0.43]	0.569***[9.07]	1.033 [0.10]	36.61***[3.03]	0.159***[3.70]	0.226***[5.28]	42.39***[5.93]	20.63***[4.86]	40.92***[5.79]	22.75***[5.33]
L	0.0146***[2.87]	-0.0102***[-8.64]	0.258 [0.33]	-1.398***[-5.01]	0.0154***[8.20]	-0.00852***[-26.90]	0.502 [1.45]	-1.123***[-18.40]	1.165***[2.70]	-1.047***[-16.80]
ADV	-0.202***[-7.48]	-0.192***[-10.94]	-20.64***[-5.15]	-19.13***[-6.81]	-0.216***[-45.21]	-0.199***[-75.50]	-23.12***[-28.72]	-19.54***[-57.08]	-23.03***[-27.84]	-20.41***[-56.19]
1.A	0.117**[2.49]	0.0774**[2.18]	3.858 [0.53]	6.008 [1.36]	0.0984***[9.67]	0.125***[23.42]	4.085***[2.83]	12.57***[18.69]	4.142***[2.87]	12.61***[18.77]
2.A	0.0686 [1.56]	0.0872**[2.49]	-1.521 [-0.24]	7.088 [1.55]	0.0612***[6.11]	0.114***[21.56]	-2.076 [-1.48]	9.837***[14.87]	-2.014 [-1.43]	9.855***[14.91]
3.A	0.0343 [0.73]	0.0400 [1.15]	-5.134 [-0.70]	2.312 [0.54]	0.0506***[4.70]	0.0887***[16.30]	-1.051 [-0.64]	7.472***[10.94]	-1.023 [-0.63]	7.557***[11.07]
4.A	0.103*[1.78]	0.0333 [0.91]	3.575 [0.51]	0.922 [0.19]	0.0739***[6.65]	0.0884***[15.31]	1.775 [1.18]	6.886***[10.27]	1.801 [1.20]	6.925***[10.35]
5.A	0.0124 [0.25]	0.0482 [1.42]	-2.215 [-0.32]	1.893 [0.42]	-0.00635 [-0.53]	0.0557***[9.09]	-7.213***[-4.32]	2.708***[3.73]	-7.255***[-4.35]	2.734***[3.77]
norain	0.0391**[2.36]	0.0182***[3.86]	14.15***[2.81]	-0.149 [-0.22]	0.0350***[5.39]	0.0217***[10.10]	11.90***[6.49]	-0.188 [-0.49]		
2.WF									2.107 [1.34]	-5.114***[-6.51]
3.WF									0.285 [0.21]	-5.251***[-7.86]
4.WF									-9.935***[-4.93]	-7.240***[-9.53]
5.WF									-14.87***[-5.36]	-1.255 [-1.05]
7.WF										-1.443*[-1.72]
8.WF									-7.639***[-3.26]	-7.005*** [-3.47]
kids					0.106***[13.98]	0.0716***[20.27]	20.25***[9.69]	11.43***[16.56]	20.24***[9.68]	11.14***[16.19]
_cons	4.615***[55.86]	3.918***[62.35]	106.4***[9.76]	68.45***[8.70]	4.049***[127.44]	3.660***[71.94]	30.42***[5.92]	29.96***[6.99]	40.46***[8.30]	33.76***[7.83]
<i>N</i>	19793	72344	19793	72344	49599	180493	49599	180493	49599	180493

R2	736	652	617	541	716	635	613	540	613	541
adj. R2	728	649	606	538	712	633	607	539	608	539

Notes: t-statistics in brackets;

\* p < 0.10; \*\* p < 0.05; \*\*\* p < 0.01. In all specifications, errors are clustered by hotel. Models (7.1) to (7.4) only include observations related to a double room for two adults; models (7.5) to (7.10) include observations related to any combination of rooms for a party group of two adults and two kids. All models only include observations related to a weekend stay of two nights (from Friday to Sunday). Models (7.1), (7.3), (7.5), (7.7) and (7.9) only include observations for L < 4 (i.e. prices posted in the last three days of advance booking before the check-in date). Models (7.2), (7.4), (7.6), (7.8) and (7.10) only include observations for L > 3 (i.e. prices posted from 15 to 4 days in advance before the check-in date).

**Table 7. Impact of accuracy of weather forecasts on prices**

	(8.1)	(8.2)	(8.3)	(8.4)	(8.5)	(8.6)	(8.7)	(8.8)
	<i>lnP</i>	<i>lnP</i>	<i>P</i>	<i>P</i>	<i>lnP</i>	<i>lnP</i>	<i>P</i>	<i>P</i>
R*H	YES	YES	YES	YES	YES	YES	YES	YES
D	YES	YES	YES	YES	YES	YES	YES	YES
C	0.040***[2.83]	0.001 [0.06]	4.848**[2.58]	-1.521 [-0.29]	0.026**[2.06]	0.061 [1.43]	0.768 [0.40]	14.09 [1.47]
1.B	0.088**[2.12]	0.093 [1.07]	6.311 [1.29]	6.703 [0.44]	0.128***[2.96]	-0.116 [-0.99]	12.48***[2.60]	-25.02 [-1.23]
2.B	0.452***[6.41]	0.242**[2.10]	36.57***[5.67]	18.75 [1.12]	0.347***[4.60]	0.057 [0.36]	29.45***[4.11]	-23.85 [-1.07]
3.B	0.256***[2.93]	0.268***[3.11]	17.22**[2.08]	20.67 [1.44]	0.295***[4.08]		23.10***[2.98]	
4.B		0.0917 [1.01]		7.120 [0.43]	0.413***[8.30]		30.74***[4.30]	
L	-0.008***[-5.56]	-0.006***[-3.54]	-0.631***[-3.43]	-1.627***[-3.20]	-0.007***[-5.28]	-0.009***[-3.42]	-0.839***[-3.83]	-1.769**[-2.53]
ADV	-0.137***[-7.61]	-0.277***[-8.31]	-16.61***[-8.17]	-25.56***[-5.84]	-0.182***[-9.08]	-0.203***[-6.54]	-17.27***[-5.91]	-22.86***[-4.11]
1.A	0.113**[2.58]	0.065**[2.03]	8.351*[1.87]	12.82**[2.16]	0.107**[2.29]	0.080*[1.83]	8.761 [1.52]	8.422 [1.09]
2.A	0.109**[2.50]	0.061*[1.77]	7.482*[1.68]	12.24**[1.97]	0.106**[2.34]	0.077*[1.68]	7.266 [1.28]	7.620 [0.97]
3.A	0.091**[2.05]	0.011 [0.36]	9.204*[1.89]	1.233 [0.23]	0.067 [1.44]	0.031 [0.77]	5.450 [0.98]	-0.186 [-0.03]
4.A	0.048 [0.98]	0.061 [1.46]	3.114 [0.73]	4.312 [0.54]	0.088*[1.81]	-0.025 [-0.56]	7.459 [1.35]	-7.240 [-0.97]
5.A	0.053 [1.20]	0.028 [0.83]	3.339 [0.77]	4.768 [0.86]	0.078*[1.76]	0.002 [0.04]	6.800 [1.46]	-5.450 [-0.69]
norain	0.057***[8.69]	0.003 [0.44]	3.641***[5.58]	2.343 [1.35]	0.033***[5.62]	0.041***[4.14]	3.108***[3.01]	3.739*[1.82]
_cons	3.907***[60.09]	4.783***[57.00]	61.59***[9.67]	128.8***[8.85]	3.870***[48.83]	3.649***[31.99]	62.96***[6.44]	39.25**[2.42]
<i>N</i>	48999	43138	48999	43138	69917	22220	69917	22220
R2	589	671	537	603	616	658	509	557
adj. R2	584	667	532	598	614	656	505	555

Notes: t-statistics in brackets;

\*  $p < 0.10$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ . In all specifications, errors are clustered by hotel. All models only include observations related to a double room for two adults and related to a weekend stay of two nights (from Friday to Sunday). Models (8.1) and (8.3) only include observations of the months of June and September; Models (8.2) and (8.4) only include observations of the months of July and August. Models (8.5) and (8.7) only include observations of the 1, 2 and 3 star hotels; Models (8.6) and (8.8) only include observations of the 4 and 4 star hotels.

**Table 8. Impact of weather uncertainty on prices**

**Table 9. Correct and wrong forecasts**

Forecast	15-day period	Last 3-day period
Rain	<b>Rain=73.64%</b> <i>Sun=26.36%</i>	<b>Rain=84.91%</b> <i>Sun=15.09%</i>
Sun	<b>Sun=84.75%</b> <i>Rain=15.25%</i>	<b>Sun=86.23%</b> <i>Rain=13.77%</i>

Notes: correct forecasts in bold, wrong forecasts in italic.



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>