NBER WORKING PAPER SERIES

# STICKY PRICE FOR DECLINING RISK? THE CASE OF CANCELLATION PREMIA IN THE HOTEL INDUSTRY

Nicola Lacetera
Claudio A. Piga
Lorenzo Zirulia

Working Paper 28456
http://www.nber.org/papers/w28456

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2021

We are grateful Heski Bar-Isaac, Ambarish Chandra, Andrew Ching, Rahul Deb, Zheng Gong Tanjim Hossain, Yao Luo, Matt Mitchell, Devin Pope, Xianwen Shi and Justin Sydnor for their feedback. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Sticky Price for Declining Risk? The Case of Cancellation Premia in the Hotel Industry
Nicola Lacetera, Claudio A. Piga, and Lorenzo Zirulia
NBER Working Paper No. 28456
February 2021
JEL No. D21,D22,D91,L11,L83

**ABSTRACT**

Using data from about seven millions room postings by hotels in France and the UK, we document that, rather than smoothly decreasing to zero, cancellation premia remain positive at roughly 10%to 15% of the full price until two days before the stay. A model where travelers have different willingness to pay and some overestimate the probability to cancel their trip explains this price-setting mode more consistently than alternative interpretations. We denote these strategies as a form of naivet -based price discrimination. We use our model also to identify conditions under which these strategies are exploitative of certain consumers, or are welfare-enhancing instead.

Nicola Lacetera
University of Toronto
Institute for Management and Innovation
3359 Mississauga Road, Room KN 235
Mississauga, ON L5L 1C6
CANADA
and NBER
nicola.lacetera@utoronto.ca

Claudio A. Piga
University of Genova
Via Balbi, 5
16126 Genova
Italy
claudio.piga@unige.it

Lorenzo Zirulia
University of Milano
Via Festa del Perdono, 7
20122 Milano
Italy
lorenzo.zirulia@unimi.it

# 1 Introduction

According to standard economic theory, prices of goods and services reflect all the available information relevant to a particular product. Changes in relevant information should therefore lead to changes in prices.

There is evidence, however, that prices do not adjust entirely, if at all, to news. This topic has long been of interest in economics, because price dynamics determine the effectiveness of certain economic policies and the trends in various socioeconomic outcomes, such as the distribution of gains from international trade or the effect of mergers, acquisitions and other forms of consolidation or fragmentation in an industry (Bils and Klenow, 2004).

Classic explanations for the low responsiveness of prices appealed, for example, to the presence of menu costs (Nakamura and Steinsson, 2008), and to the use of efficiency wages (Akerlof and Yellen, 1986). Studies that departed from the standard rational-agent framework have considered psychological tendencies that might account for limited price adjustments, such as concerns for fairness, loss aversion, and managerial inertia (Kahneman et al., 1986; Choi and Mattila, 2003; Anderson and Simester, 2010; Heidhues and Kőszegi, 2008, 2010; Kőszegi and Rabin, 2006, 2007).

We contribute to this field of inquiry with theory and evidence of how an additional psychological mechanism may explain the limited flexibility of prices in certain markets where purchases (reservations) occur before the actual fruition of a good or service. Examples include hospitality, travel, entertainment and sport events, and e-commerce. When consumers book a good or service to enjoy at a future date, and the actual fruition is uncertain, companies may offer insurance against missed consumption at a premium: a refundable tariff. Although one may expect this premium to decline as the uncertainty resolves, companies may instead adopt pricing schemes that limit or exclude premium changes in order to leverage the systematic bias of some consumers in assessing the likelihood of cancelling their purchase.

Our empirical focus is on the hotel industry. Customers can generally choose, for a given room configuration and dates, between a refundable a non-refundable payment option. The former deal includes reimbursement in case of cancellation before the date of arrival, whereas the latter does not. The price of the refundable option is therefore higher than the non-refundable price. However, as the arrival (check-in) date approaches, the insurance component of the refundable option is less valuable.

But do we observe the prices for the refundable and non-refundable options converge to one another? To address this question, we collected information on the full set of hotel establishments in nine French and thirty-two British cities, operating on the *Booking.com* platform – a total of $2,030$ hotels and about seven millions room postings. The data include information on prices for all rooms at a specific query date, the check-in date, the cancellation policy, and several characteristics of the hotel and the specific room.

We document that refundable prices are indeed higher than the corresponding non-refundable ones. However, cancellation premia remain stable at more than 10% of the full price with little variation over time, and no convergence toward zero as the arrival day nears.

Further analyses provide evidence that the persistently positive premium is likely an intentional, profit-maximizing managerial choice. We find, in particular, that when the refundable and non-refundable prices change during the period ahead of stay, their change is often fine-tuned so as to not alter their difference. Our evidence, moreover, does not support the hypothesis of systematic mispricing by managers, nor is it consistent with the adoption of certain "decoy" pricing strategies in order, for example, to make only one type of tariff appealing to customers.

Rational customers, however, would not be willing to buy a more expensive cancellation option, when the uncertainty is fundamentally resolved. The persistence of a positive cancellation premium thus indicates potential naiveté on the side of customers, which facilitates pricing strategies of the type that we document.

To interpret our evidence, we propose a model of these price-setting patterns as a form of (consumer-side) naiveté-based price discrimination, where the naiveté consists in systematically overestimating the probability to cancel a trip. The tendency to overweight the probability of rare events is one of the implications of prospect theory (Kahneman and Tversky, 1979), and several studies demonstrated its presence and implications in relevant contexts – see for example Sydnor (2010), Barseghyan et al. (2013) and Jindal (2015). We derive that profit-maximizing firms optimally sort consumers that are heterogenous in their willingness to pay and degree of sophistication by setting a price menu with both a refundable and a non-refundable tariff.[1] This result is reminiscent of previous work on the airline industry by Escobari and Jindapon (2014); similar to hotel reservations, air travelers book their seat in advance and can choose to insure against cancellation from a menu of fares. This model only includes consumer heterogeneity in willingness to pay; we show that firms can find it profitable to offer refundable tariffs tailored to consumers who have both high willingness to pay and biased beliefs even when the (actual) cancellation probability is negligible. Calibrations based on our data do not support standard risk aversion by travelers as a plausible explanation of the evidence, whereas assuming probability over-weighting fits the model for more reasonable parameter values.

Our work combines the insights from studies of price discrimination with naive consumers

[1] Bruhin et al. (2010) provide experimental evidence on heterogeneity in probability distortion across individuals.

and results from work on price rigidities due to non-standard preferences and beliefs. Research in behavioral industrial organization analyzed price discrimination in presence of time inconsistency and overconfidence of certain consumers (Eliaz and Spiegler, 2008; Sandroni and Squintani, 2013; Heidhues and Kőszegi, 2010, 2017). The objective of these studies is not to explain price rigidities, but rather to derive optimal pricing menus and their welfare implications. In addition to providing an explanation for certain types of price rigidities, we also establish the conditions under which price discrimination in the presence of naive consumers is not exploitative (as is usually the case in the literature) but is, instead, welfare-enhancing. This occurs when naiveté-based price discrimination allows a partial form of preference-based (second-degree) price discrimination that would not be possible otherwise. In this case, consumers with a higher willingness to pay are more affected by the distortion in the perception of the cancellation risk. When price discrimination expands the market, profits increase and each consumer type is (weakly) better off. Allowing for consumers to be heterogenous both in their willingness to pay and degree of sophistication, we provide novel insights for those markets where consumers plausibly differ along both these features.

We describe the data in Section 2. Section 3 reports our key findings and potential explanations for them. In Section 4 we present our model. Section 5 provides a discussion and concluding remarks.

# 2 The data

The data collection involved four steps. First, we obtained the full list of establishments operating on the *Booking.com* platform using the identifiers of nine French and thirty-two British cities. For each establishment, we recorded the unique *url* identifier and its type (hotel, B&B, apartment, villa, inn, etc.). We focus only on establishments listed as hotels, because the other lodging types tend to be small family-run businesses, which adopt simple, unsophisticated pricing approaches (Mantovani et al., 2021).

Second, we scraped the hotels' pages to retrieve their star classification, size as measured in number of rooms, and whether they were affiliated with a chain. Because there is extensive agreement in the literature that chain membership confers a competitive advantage by improving a hotel's revenue management capabilities (Kosova and Lafontaine, 2012; Kosova et al., 2013; Hollenbeck, 2017; Mantovani et al., 2021), and that one-star hotels exhibit a very low propensity toward active pricing (Melis and Piga, 2017), we restricted the sample to only chain hotels with at least a two-stars classification.

Third, to reduce scraping problems arising from possible changes in the HTML that

*Booking.com* uses, we saved the pages on a local disk.

Fourth, we parsed the internal HTML code to create the sample that we used for the analysis. Thus, we could conveniently update the parsing program without losing information stored in the HTML file.[2]

The data cover stay dates between 30th October 2017 and 2nd January 2018, with intervals of three days to ensure that all weekdays were represented and to reduce collection time. We retrieved room prices in advance of the stay period. Starting on September 7th 2017 and continuing on a daily basis (whenever possible), we issued individual queries that specified each hotel's *url* identifier and the stay dates. Doing so allowed us to obtain the prices for all the varieties of rooms that an hotel offered; this would not be possible, for example, if we based the query on the city listing. The hotel page provides a precise relationship between the price and the offer characteristics in terms of cancellation policy, breakfast inclusion and the number of persons allowed. Figure A.1 in the Appendix shows a snapshot of some room postings. Similar examples continue to be qualitatively valid at the time of writing this paper (February 2021). This information is central to derive a precise value of the cancellation premium while holding all other characteristics constant.

Prices are in Euro and Sterling for French and British hotels, respectively. The page also includes the overall customer rating, and its division into various components: Comfort, Cleanliness, Staff quality, Facilities, which we use to proxy how well an hotel is managed.

The sample for the analysis contains 881 hotels in the UK and 1,149 in France that posted both the Refundable (R) and Non-Refundable (NR) prices. Hotels in the three and four-stars categories account for the largest proportion in both samples (respectively, 27.6% and 56.3% in France, 37.0% and 50.0% in the UK). The share of five-stars hotels is about the same in both countries (8%). Hotels in the UK are larger, with almost 69% of establishment having between fifty and two hundred rooms; the proportion is about 61% in the French sample, which includes about 32% of small hotels with less than fifty rooms.

The scraping activity retrieved also hotels that exclusively posted rooms using only either the R or NR price. We dropped these hotels from the sample, because no cancellation premium is available. The single-price hotels are a minority (about 10% of the full sample in both countries), are smaller than the rest, and tend to be two or three-star hotels, i.e., are vertically differentiated to serve a more price-sensitive customer segment.[3]

The resulting dataset includes about seven million observations, the majority of which

[2]We verified that web scraping did not engender dynamic pricing (Cavallo, 2017). First, we cleaned the cookie folder every day; second, using computers that were not used for the data collection, we issued some queries identical to those made by the scraping computers by hand. We could not find any noticeable difference.

[3]The full distribution of hotels by star and size is available in Appendix B, Table B.1.

(between 65-80%) report both the R and NR price. Hotel may only post one of the two prices for some types of rooms, or, more relevant for our analysis, they may choose to offer only one at some point before the stay date. For instance, once the cancellation option expires, which normally happens between seven and one day before the stay, one would only observe the NR price. This is indeed when we retrieve a larger proportion of observations with only the NR price, in both countries.[4]

# 3 Empirical analysis

## 3.1 The persistence of refundable premia

We estimate the relationship between room prices and the time between the posting and reservation date through the following econometric model:

$$Y_{rhd|s=i;c=j} = \sum_{d} \delta_d D_d + \sum_{r,h} \beta_d X_{rh} + \varepsilon_{rhd}. \tag{1}$$

The outcomes that we consider are the price per night, separately for refundable and non-refundable options, as well as the cancellation premium, i.e., the difference between the price of a refundable and a non-refundable option for a room $r$ in a hotel $h$, with $s$ stars and in country $c$. The binary indicators $D_d$ take the value of one if the posting is $d$ days before the date of the stay, and zero otherwise. The vector $X$ includes indicators of the combination of hotel, check-in date, room type (double, luxury double, triple, ...), the maximum number of guests allowed in the room, and the inclusion of breakfast. As an example, one observation in our regressions would represent a room in three-star *Hôtel des Princes* in France, allowing a maximum of four guests, with check-in date 6 Nov. 2017, whose prices for both a refundable and non-refundable options (breakfast excluded) are observed on date 14 Oct. 2017 (twenty-three days before the check-in date). These controls are akin to detailed fixed effects that allow to account for different features or "bundles" that might differentiate, in particular, refundable and non-refundable options. We allow the error terms $\varepsilon$ to be correlated at the hotel level; as such, we cluster the standard errors at this level of aggregation. We then estimate the vectors of parameters $\delta_d$ and $\beta_d$ and derive the predicted values $\hat{Y}_{rh|s=i,d=D}$.

Figure 1 shows the estimated premia for a refundable option at different times before check-in, separately by number of stars and country of the establishment. The estimates are stable over a time lag between posting and check-in date from eighty days to one day. In a few instances, we observe a "stepwise" reduction of the premium in the two weeks before the

[4] The full distribution of observations over the period ahead of stay is available in Appendix B, Table B.2.

check-in date. In no cases, however, does the premium drop or decrease smoothly to zero in close proximity of the date of stay. In Figure 2, we report the estimated prices for the refundable and non-refundable options. The prices, too, have little variation over different time lags. This implies that the premium changes little not only in absolute terms, but also as a percentage of price; on average, the estimated premium is about 12% to 16% of the full refundable price.[5]

### 3.2 Understanding the premium persistence

We consider a few potential explanations for why premia for refundable reservations remain positive even in close temporal proximity of the check-in date.

Hotels, for one, may set prices for reasons other than conveying all the available information or the level of uncertainty. Prices for refundable reservations, for example, may be kept artificially high to establish a reference point or a "decoy" for consumers, in order to induce them to book a room at the more convenient non-refundable price. This would imply that hotel managers would think it highly unlikely that any transaction would include paying the higher refundable price. In this case, whenever refundable fares are not posted, we should observe the non-refundable fare to be lower than the refundable one that was offered in nearby dates. In our sample, this may occur if a hotel occasionally stops posting the refundable option, or after its termination date occurs, generally between seven and one day before the stay.

The histograms in Figure 3 show the distribution of the following statistic:

$$\Delta P = \left[ P_{NR}(d) - P_R(d+1) \; | P_R(d) \; \textit{not observed} \right]. \tag{2}$$

$\Delta P$ is the difference of the non-refundable price $P_{NR}$ posted $d$ days before the stay, and the refundable price $P_R$ posted the day before, conditional on the latter not being available on $d$. Both in France and in the UK, over half of the distribution mass of $\Delta P$ is at zero, i.e., the NR price converges towards the refundable level *when the latter is not offered any longer*. Because we mostly observe this during the week before the stay (not shown, but available on request), Figure 3 suggests that the refundable price does not operate as a "decoy", but it is instead a price at which the hotel expects to transact. In fact, in some cases $\Delta P > 0$, i.e., occasionally, the non-refundable price is set even above the past refundable values.

[5] Figure B.1 in the Appendix displays the estimates of parameters $\rho_d$ from the following regression equation: $\ln(price)_{rhd|s=i;c=j} = \sum_d \delta_d D_d + I(R) * \sum_d \rho_d D_d + \sum_{r,h} \beta_d X_{rh} + \varepsilon_{rhd}$, where $ln(price)$ is the natural logarithm of the price of a room, $I(R)$ is binary indicator taking value of 1 if an observation concerns the price for a refundable reservation, and zero otherwise, and all other terms and subscripts are as in Equation (1). The values of $\widehat{\rho}_d$ in Figure B.1 thus represent the cancellation premium in percentage terms.

Second, the findings may reflect systematic mispricing, thus providing consumers with arbitrage opportunities. DellaVigna and Gentzkow (2019), for example, conclude that the limited price changes within a grocery store chain for the same product in different socioeconomic areas is consistent with inertial behavior of managers.[6] Institutional features of the hotel industry, as well as evidence from our own data, allow us to rule out that the persistence of premia is due to managerial inertia. The data come from hotels that are part of established chains, and exclude one-star hotels. Previous research found that hotel chains employ sophisticated revenue-management strategies; Hollenbeck (2017), for example, discuss how information on revenue management strategies shared by chain hotels may become a source of competitive advantage; as such, it is unlikely that their pricing strategies are systematically suboptimal.

Hotels with fewer stars, moreover, are more likely to be smaller and more informally managed, but this is generally not the case for establishment with three stars or more (Melis and Piga, 2017; Mantovani et al., 2021). If there were systematic mispricing in our sample, we may expect it to occur especially among less sophisticated or more badly managed hotels, as several studies indicate that features such as market experience or the quality of education may correct certain behavioral tendencies (List, 2003; Goldfarb and Xiao, 2011; Anagol et al., 2018). Our evidence is inconsistent with this prediction. Premia, for example, stay positive for hotels of different quality as measured by stars. Within each star category, moreover, we rely on the customer ratings available in the data to evaluate management quality (Vives et al., 2018). We define a hotel as having low management quality if it scores below the median value of its star group in all the four rating measures, and high management quality otherwise. After estimating the model in equation (1) separately for hotels with high and low-rated management quality, we show the predicted premia in Figure 4. The predicted values are slightly smaller for hotels with lower management quality scores, but the patterns are similar, and the premia are stable, for all categories of hotels and in both countries.

Finally, the stability of premia may be just a direct consequence of the stability of prices. However, Figure 5 shows that this is not the case. We plot the sample proportion of observed changes, within a certain number of days before the check-in date, in both prices (dotted lines) and the premium (full lines), by stars. Changes in prices are much more frequent than changes in premia; in other words, when hotels adjust both the refundable and non-refundable prices, they most often keep the premium unchanged. This indicates that the stability of premia may be the result of a deliberate decision by hotel managers.

In the next section, we provide support to our preferred interpretation of the evidence. We

[6] Arbitrage by customers is less plausible in the context that DellaVigna and Gentzkow (2019) study than in the one we consider.

claim, in particular, that a form of naiveté-based menu pricing (Heidhues and Kőszegi, 2017), with some consumers over-weighting small probabilities of cancellation, leads to the persistence of cancellation premia and the documented lack of convergence between the refundable and the non-refundable tariffs.

# 4 A model of naiveté-based pricing

## 4.1 Model description

A monopolistic firm provides, at zero marginal cost, a service that consumers value at either $v_H$ or $v_L$, with $v_L < v_H$. Consumers make their purchasing decisions at period $\tau$ (i.e., the booking date), but actually enjoy the service at period $t$ (i.e. the arrival or check-in date) with probability $\pi(\tau) \in (0, 1)$, $\pi'(\tau) > 0$. With the complementary probability $1-\pi(\tau)$, consumers do not enjoy the service and receive zero value.

For a given $t$, cohorts of consumers entering at $\tau$ are otherwise identical. There are $N_H$ consumers of type $H$, i.e., they have a valuation $v_H$ for the service (high-valuation types), and $N_L$ consumers of type $L$ (low-valuation types). The probability $\pi(\tau)$ does not depend on consumer type; from now on, we simply express this likelihood as $\pi$.

A fraction $\beta$ of customers has a correct perception of $\pi$, with $\beta$ independent of the valuation type. We define these consumers *sophisticated*, and indicate their type with $S$. The remaining fraction $1 - \beta$ holds a distorted perception of $\pi$; we denote it as $g(\pi)$, with $g'(\pi) > 0$. We call these consumers *naive* (index $N$). We assume that $\pi > g(\pi)$; this is justified by our focus on "high" values of $\pi$ (i.e., "low" cancellation probability) and the assumption of over-weighting of small probabilities. All consumers, irrespective of their valuation and sophistication, share the same increasing and concave utility function $u(\circ)$ (consumers are risk-averse).

In each period $\tau$, the profit-maximizing firm offers a menu with a fully refundable and a fully non-refundable tariff (we exclude partial refunds).

We classify consumers in four "composite" types given by the combination of valuation and sophistication, and denote them with the indexes $HS$, $HN$, $LS$ and $LN$. The expected utility of a sophisticated consumer with valuation $\theta$ $(= H, L)$, buying a refundable tariff at price $p_R$, is:

$$U^R_{\theta S} = \pi u(v_\theta - p_R). \tag{3}$$

For naive consumers, the corresponding expected utility is:

$$U^R_{\theta N} = g(\pi)u(v_\theta - p_R). \tag{4}$$

The expected utility for a sophisticated type buying a non-refundable tariff $p_{NR}$ is:

$$U_{\theta S}^{NR} = \pi u(v_\theta - p_{NR}) + (1-\pi)u(-p_{NR}), \tag{5}$$

and for naive consumers, the corresponding expected utility is:

$$U_{\theta N}^{NR} = g(\pi)u(v_\theta - p_{NR}) + (1-g(\pi))u(-p_{NR}). \tag{6}$$

We set to zero the utility of a consumer who does not buy. Thus $c_{\theta S}$, the reservation price for a non-refundable tariff of a sophisticated consumer with valuation $\theta$, is the solution to:

$$\pi u(v_\theta - c_{\theta S}) + (1-\pi)u(-c_{\theta S}) = 0. \tag{7}$$

For naive consumers the corresponding reservation price, $c_{\theta N}$, solves:

$$g(\pi)u(v_\theta - c_{\theta N}) + (1-g(\pi))u(-c_{\theta N}) = 0. \tag{8}$$

Because $\pi > g(\pi)$, it follows that $c_{\theta S} > c_{\theta N}$, i.e., sophisticated types have a higher reservation price than naive consumers with the same valuation. Moreover, $c_{HS} > c_{LS}$ and $c_{HN} > c_{LN}$: for a given level of sophistication, high-valuation consumers have a higher reservation price. We also assume that $c_{HS} > c_{HN} > c_{LS} > c_{LN}$, i.e., reservation prices are more affected by valuation than by the level of sophistication.

## 4.2 The optimal tariff menu when cancellation probability is small

In this Section we ask whether there are conditions under which the cancellation premium can remain positive and large even when the cancellation probability is ”small”. Thus we provide the characterization of the equilibrium for $\pi \to 1$. Consistently with our previous assumption, we assume $\lim_{\pi\to 1} g(\pi) = \overline{g} < 1$, which implies $\lim_{\pi\to 1} v_H - c_{HN} > 0$ and $\lim_{\pi\to 1} v_L - c_{LN} > 0$. A general version of the model, as well as all proofs, are in Appendix C.

We can restrict our attention to four undominated candidate equilibria configurations, and classify them in two groups. In the first group, two equilibria, which we denote as $I$ and $II$, exhibit naiveté-based discrimination. In both $I$ and $II$, naive high-valuation consumers ($HN$) select the refundable tariff, whereas sophisticated high-valuation consumers ($HS$) select the non-refundable tariff. Naiveté, therefore, leads to overinsurance. The two equilibria differ because in $I$ all types (except type $HN$) select the non-refundable tariff, whereas in $II$ type $LN$ does not buy in equilibrium; in this latter case, therefore, the market is not fully covered. The non-refundable tariff in $I$ that induces $LN$ types to buy (those with the lowest reservation price for the non-refundable price) is lower than the non-refundable tariff in $II$; consequently, the refundable tariff in $I$ is lower than the refundable tariff in $II$, in order for the incentive compatibility constraint of type $HN$ to be satisfied.

Candidate equilibria in the second group do not include naiveté-based discrimination. Equilibrium $III$ corresponds to the case where all types select the refundable tariff. In equilibrium $IV$, the firm offers only the refundable tariff, which $H$ consumers (both sophisticated and naive) purchase.

Proposition 1 expresses the conditions that make each of the two configurations exhibiting naiveté-based discrimination the optimal one.

**Proposition 1** *Define $m_I$ as the solution to $g(\pi)u(v_H - m_I) = g(\pi)u(v_H - c_{LN}) + (1 - g(\pi))u(-c_{LN})$ and $m_{II}$ as the solution to $u(v_H - m_{II}) = g(\pi)u(v_H - v_L) + (1 - g(\pi))u(-v_L)$.*

*i) If $[(1-\beta)(m_{II} - m_I) + \beta(v_L - c_{LN})]\, N_H < (c_{LN} - \beta v_L)N_L$, then $p^*_{NR} = c_{LN}$ and $p^*_R = m_I$ is the optimal tariff menu if:*

$$m_I > \max\left\{\frac{v_L(N_H+N_L) - c_{LN}(\beta N_H + N_L)}{N_H(1-\beta)}; \frac{v_H N_H - c_{LN}(\beta N_H + N_L)}{N_H(1-\beta)}\right\}; \tag{9}$$

*ii) If $[(1-\beta)(m_{II} - m_I) + \beta(v_L - c_{LN})]\, N_H > (c_{LN} - \beta v_L)N_L$, then $p^*_{NR} = v_L$ and $p^*_R = m_{II}$ is the optimal tariff menu if*

$$m_{II} > \max\left\{\frac{v_L(N_H+N_L)}{N_H}; \frac{v_H N_H - \beta v_L(N_H+N_L)}{N_H(1-\beta)}\right\}. \tag{10}$$

The condition

$$[(1-\beta)(m_{II} - m_I) + \beta(v_L - c_{LN})]\, N_H < (c_{LN} - \beta v_L)N_L \tag{11}$$

holds when configuration $I$ guarantees higher expected profits than configuration $II$. The left-hand side of the inequality is always positive because $v_L > c_{LN}$ and $m_{II} > m_I$ (the latter inequality derives from the binding incentive-compatibility constraints). The right-hand side term is negative for sufficiently high values of $\beta$, and declines with $N_L$. The intuition for this result is that serving naive low-valuation types (which requires low refundable and non-refundable tariffs) is convenient only if the number of these consumers, $N_L(1-\beta)$, is sufficiently high.

Suppose now that configuration $I$ dominates $II$. To be the equilibrium, this menu must yield higher profits than $III$ and $IV$ as well. This is the case if (9) holds. If instead $II$ entails higher expected profits than $I$, the configuration is preferable to $III$ and $IV$ if (10) holds.[7]

[7] Suppose that $u(v_\theta - p) = \log(1 + \frac{v_\theta - p}{k})$, with $k = 10000, v_H = 250$ and $v_H = 200$, and assume $\overline{g} = 0.95$.

### 4.3 Welfare implications

Does the presence of naive consumers, and their consideration in a firm's pricing strategy, have welfare implications? In this section we show that naiveté-based price discrimination can be welfare-enhancing. In our setting, serving all customers is socially efficient because all valuations are higher than marginal cost. If naiveté-based discrimination increases market coverage, then social welfare increases, consistently with standard results from models of price discrimination (Schmalensee, 1981; Katz, 1983). We show, in addition, that both the firm and each type of consumers are (weakly) better off in this case. Therefore, not only can naiveté-based discrimination be welfare-maximizing (see for example Heidhues and Kőszegi (2017)); it can also be Pareto-improving.

When naiveté-based discrimination does not affect total welfare, it has redistributionary effects; the firm is better off, naive types are (weakly) worse off, and sophisticated types are (weakly) better off. Introducing naive consumers in the market makes sophisticated types (weakly) better off, but this does not imply that naive consumers are necessarily exploited, as in the case of Eliaz and Spiegler (2008) or Heidhues and Kőszegi (2010). Consider these two results:

- Suppose that $v_H N_H > v_L(N_H + N_L)$. If the menu choice is restricted to cases without naiveté-based discrimination (configurations $III$ and $IV$), in equilibrium for $\pi \to 1$ only high-valuation types are served, and pay a price (approximately) equal to $v_H$. In this case, the market is not fully covered, and the utility of low-valuation types is zero. With (profitable) naiveté-based discrimination, the naive high-valuation types pay $m_{II}$ or $m_I$, both lower than $v_H$. The sophisticated types are better off because $v_L < v_H$ (or $c_{LN} < v_H$). Low-valuation types are indifferent under configuration $II$, whereas the sophisticated low-valuation types enjoy a rent under configuration $I$ because $c_{LN} < v_L$. Intuitively, naiveté-based price discrimination leads to a partial form of screening that allows the firm to treat naive high types differently from low types, while serving also the latter in equilibrium (at least the sophisticated ones). This possibility hinges upon the fact that consumers with a higher willingness to pay are more affected by the distortion in the perception of the cancellation risk. Therefore our results depends crucially on preference heterogeneity.

- Now consider $v_H N_H < v_L(N_H + N_L)$. Absent naiveté-based discrimination, the market is fully covered in equilibrium, with all consumers (approximately) paying $v_L$. If naiveté-based discrimination is considered (and profitable), a NH type ends up paying $m_{II}$ (or

Configuration $II$ is the equilibrium if $\beta = 0.95, N_H = 200$ and $N_L = 10$; while it is configuration $I$ to be the equilibrium for $\beta = 0.7, N_H = 170$ and $N_L = 50$.

$m_I$), which are both higher than $v_L$: therefore, he is worse off. The HS type is indifferent under configuration $II$, while she is better off under configuration $I$. As in the previous case, low types are indifferent under configuration $II$, while the sophisticated low type obtains a rent under configuration $I$.

### 4.4 Risk aversion *vs* belief distortion

As long as the probability of cancellation remains strictly positive, standard risk aversion can generate positive cancellation premia. In this section we show, however, that for reasonable parameters values, the magnitude of these premia as the arrival date approaches is too small as compared to what we observe in the data. Our results are therefore consistent with well-established "calibration theorems" (Rabin, 2000).

Let us express the utility of consumers with the function $u(v_\theta - p) = \log(1 + \frac{v_\theta - p}{k})$, where $k$ is a positive constant. A log utility function implies a coefficient of relative risk aversion equal to 1, in line with recent empirical estimates (Chetty, 2006; Hartley et al., 2014).

We set valuations in a monetary range compatible with four-stars hotels in our sample, i.e. $v_H = 250$ and $v_L = 200$. The value of $k$ identifies the contribution that the surplus from service consumption can have on individuals's wealth. We assume $k = 10000$; with initial wealth normalized to 1, this sets an upper bound of 2.5% for this contribution. As for the choice of $g(\pi)$, we compare biased and unbiased beliefs. We define $g(\pi) = \frac{\pi^\gamma}{(\pi^\gamma + (1-\pi)^\gamma)^{\frac{1}{\gamma}}}$, where $\gamma \leq 1$, for consumers with biased beliefs (Tversky and Kahneman, 1992).[8] Following Wu and Gonzalez (1996), we set $\gamma = 0.71$. For the case with unbiased beliefs, $g(\pi) = 1$. Finally, we focus on equilibrium configuration $II$.

Detailed evidence for cancellation probabilities in the hotel industry is limited, because this information is only available from proprietary data. One of the few papers providing estimates is Falk and Vieru (2018), which looks at cancellation risk using individual bookings data from the booking system of nine hotels in a Finnish chain. The study reports a probability of cancellation that is 2.9 % for bookings made between 1 and 4 days before a stay, 4.8% for bookings between 5 and 9 days and 11% for bookings between 10 and 24 days. Cho et al. (2018) report probabilities of cancellation as a function of days prior to arrival. These range from under 0.5% (40 days prior to arrival) to a peak of 1.4% a few days before check-in. The values are comparable to Falk and Vieru (2018). We thus focus on values of $\pi$ from 0.9 to 0.97. In this interval, the absolute risk premium ranges from 53.53 to 25.28 when beliefs are biased, but,absent probability overweighting, from only 25.28 to 7.59. The

[8] $g(\pi)$ is inconsistent with the assumption $\lim_{\pi \to 1} g(\pi) = 1$. However, this in inconsequential because in this Section we consider $\pi \leq 0.97$.

estimated premium for French four-stars hotels in our sample is never lower than 25 Euros; it falls slightly below 20 pounds immediately before the check-in date in the UK. In addition, the ratio between the premium at $\pi = 0.9$ and $\pi = 0.97$ is 2.12 with biased beliefs and 3.33 without bias (risk aversion only), thus implying a lower convergence of refundable and non-refundable prices in the latter case.[9]

# 5 Conclusions

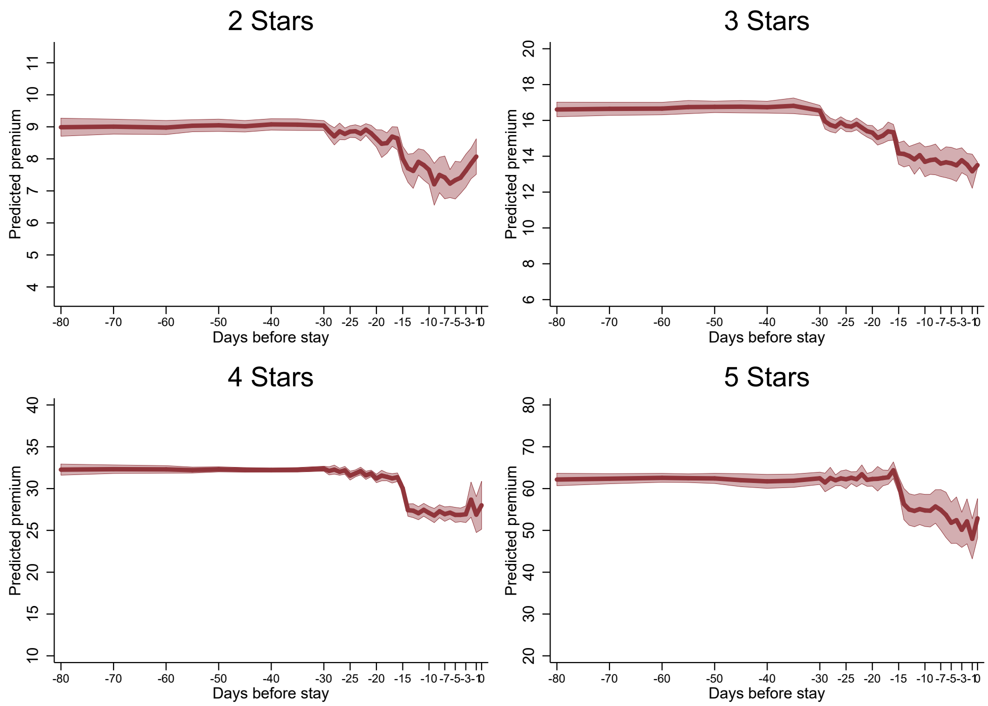
We find systematic evidence of the persistence of an insurance-like premium for the option to cancel a hotel reservation, of about 10% to 15% of the full price, even when the uncertainty about the fruition of a room is likely to be minimal, i.e., in the proximity of the check-in date. The patterns and the estimated size of the premia relate closely to a model where firms adopt pricing strategies with menus that sort customers depending both on their valuation of a good and on whether they estimate correctly or overestimate the likelihood of having to cancel. Under certain conditions, equilibrium strategies sort consumers based on their degree of sophistication and exploit naive consumers, i.e., deliver lower utility to them as compared to a situation where naiveté-based discrimination is not implemented. However, menu pricing can also benefit this type of customers, because it may expand the market and serve types of clients that might be left out with different pricing schemes that do not consider the behavioral aspects we highlight in our study.

There are several other relevant markets where there is a time lag between the purchase of a good or service and its fruition, and where contingencies can preclude fruition at the established date. Examples include travel, entertainment and sport events, as well as any market where consumers order a good rather than purchase it in person, and have the good delivered at a later date. It is likely that in all these different markets, perhaps with different prevalence, consumers differ in their attitudes toward uncertainty and the possibility of losses, and that companies take advantage of these tendencies and the heterogeneity in the population. In addition to showing how these psychological tendencies and pricing strategies play out in the hotel industry, we also provide a framework to assess under what circumstances menu pricing is exploitative, redistributive, or welfare enhancing. We believe that investigating the relevance of our approach in these other markets is a fruitful area of research.

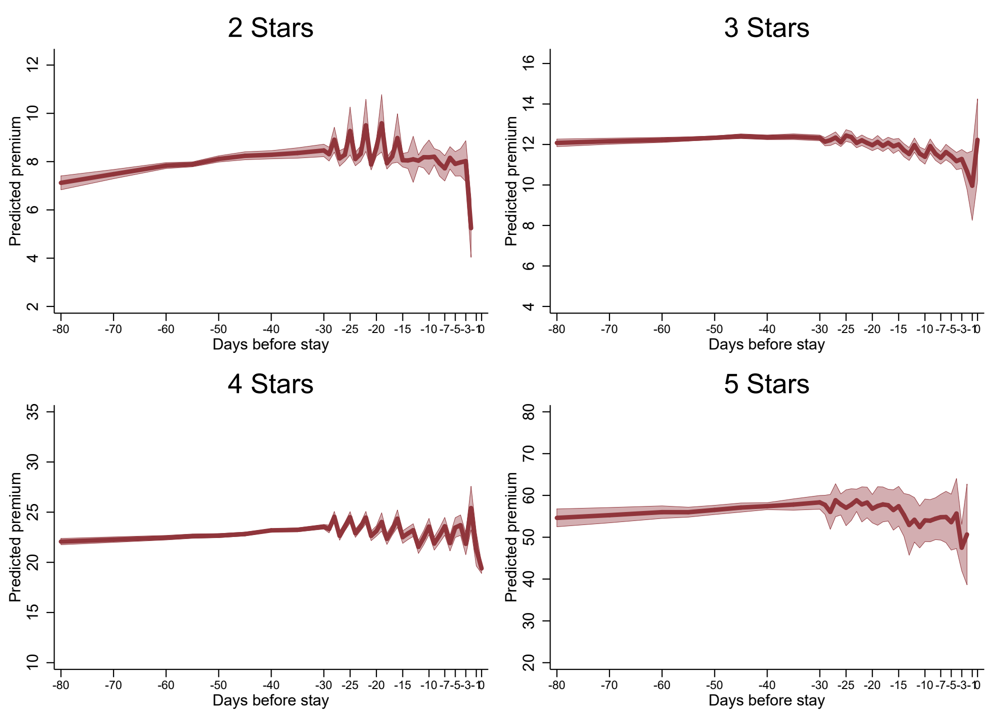
[9] In both cases, the rate of convergence is somewhat higher than the rate observed in the data. One way to deal with this issue would be to extend the model, to include the additional constraint that the refundable price at $\tau$ cannot be higher than the refundable price at any subsequent period $\tau\prime$ (otherwise, customers monitoring firm's offer would have the opportunity to cancel at $\tau^{'}$ and book the same room, most often still available, at lower price).

Figure 1: Estimated refundable reservation premia by days before check-in and stars

(a) **France**



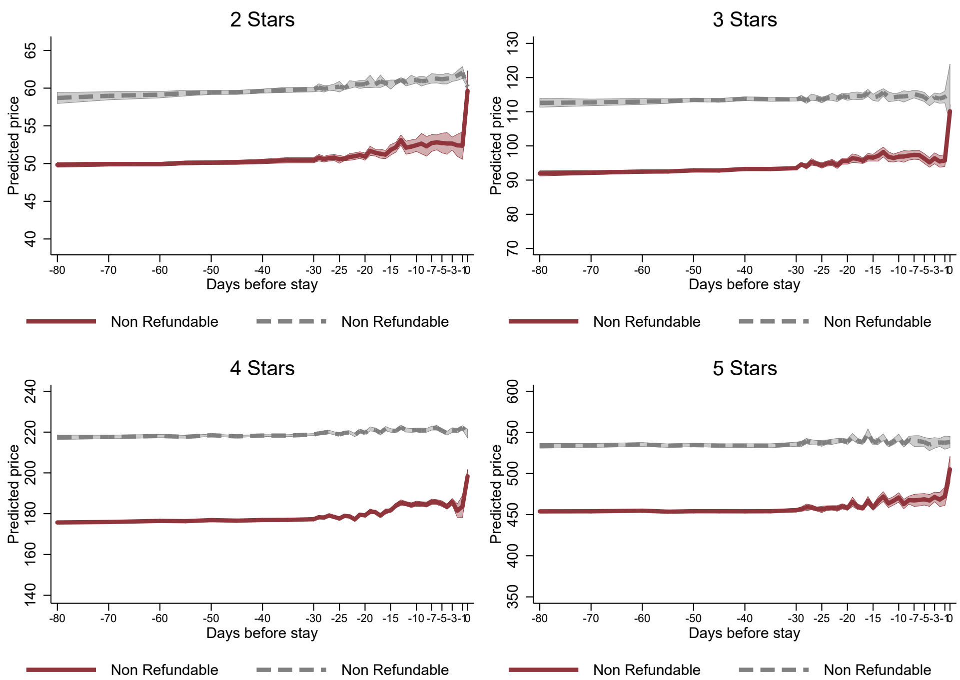

(b) **UK**



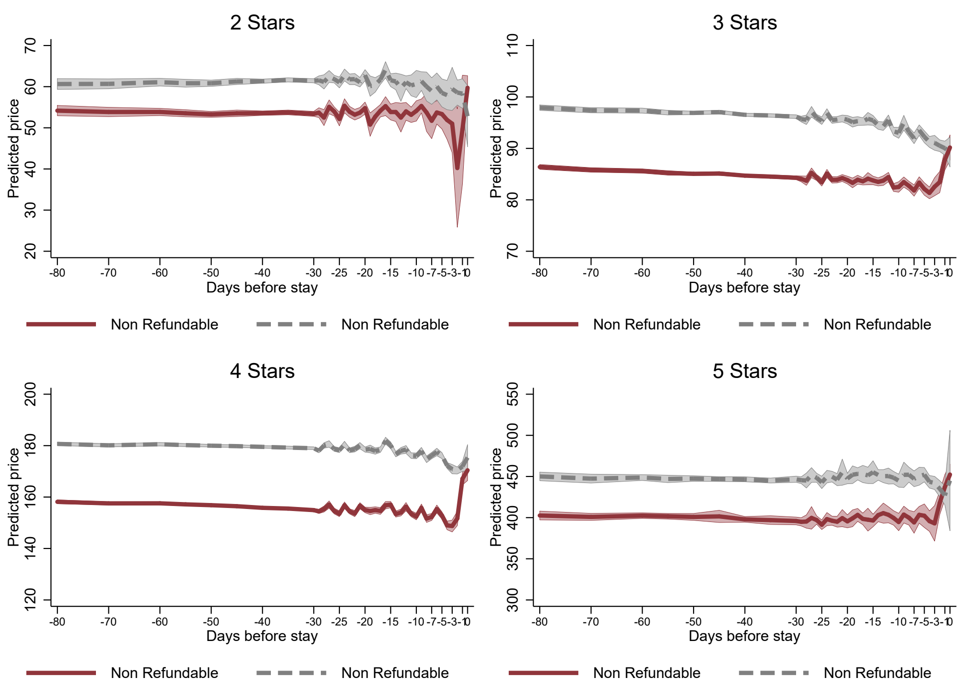

*Notes*: The graphs display the estimated values of the cancellation premium, from model (1). The regressions are separate by stars and country. The shaded areas indicate 95% confidence intervals, with standard errors clustered at the hotel level. Fixed effects combine hotel, check-in date, room type, number of allowed guests, and breakfast inclusion. The number of observations from 2, 3, 4, and 5-stars hotels are, respectively, $63,324$; $319,623$; $1,709,202$; $315,288$ in France, and $46,888$; $770,904$; $1,663,163$; $259,056$ in the UK. $R^2$ values range from $0.81-0.88$ in France, and $0.85-0.88$ in the UK. Values are in Euros (France) and Sterling (UK).

Figure 2: Estimated refundable and non-refundable reservation prices by days before check-in and stars

(a) **France**



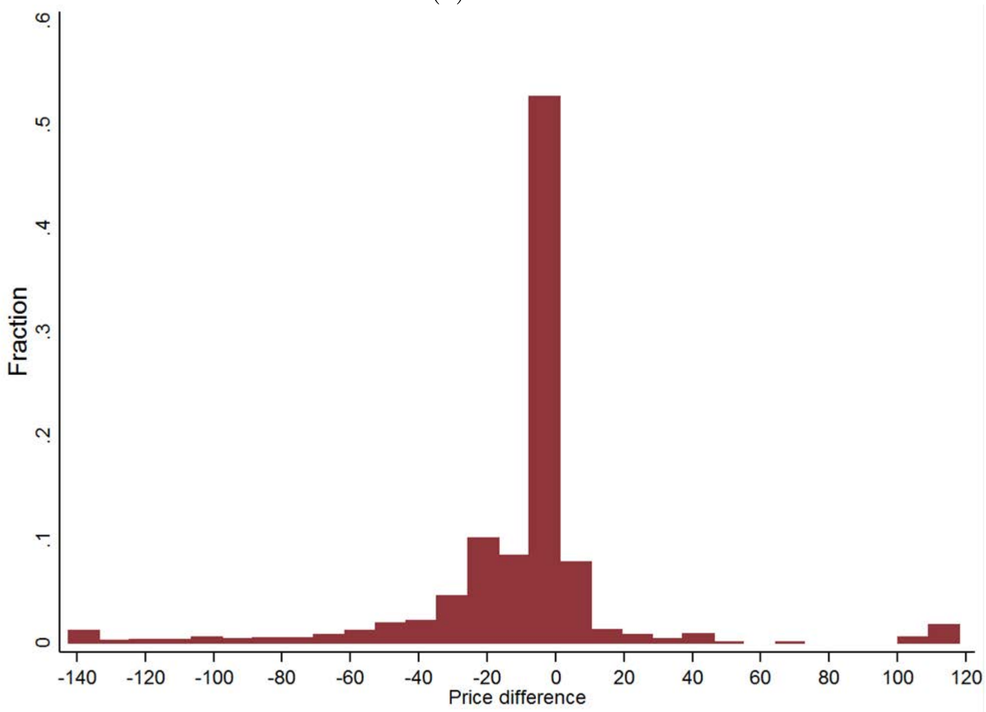

(b) **UK**



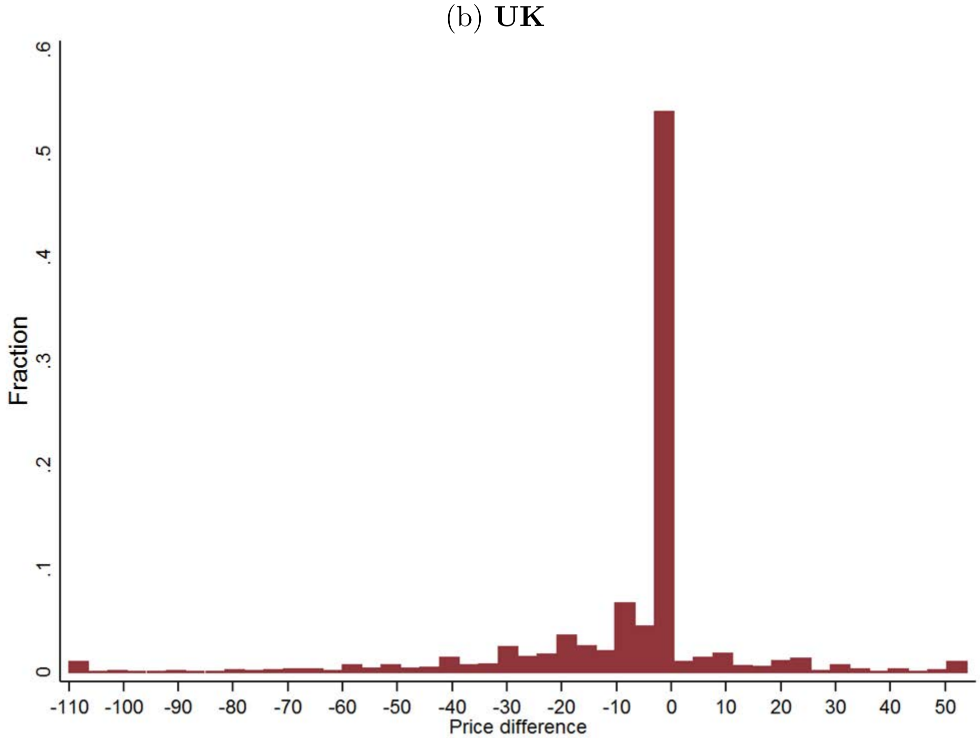

*Notes*: The graphs display the estimated values of refundable and non-refundable prices, from model (1). Regressions are separate by type of price, stars and country. The shaded areas indicate 95% confidence intervals, with standard errors clustered at the hotel level. Fixed effects combine hotel, check-in date, room type, number of allowed guests, and breakfast inclusion. The number of observations for the refundable (non-refundable) prices from 2, 3, 4, and 5-stars hotels are, respectively, $81,452$ $(102,509)$; $381,707$ (433,347); $1,920,753$ $(2,307,963)$; $357,486$ $(382,361)$ in France, and $55,951$ $(47,735)$; $875,485$ $(798,868)$; $1,987,259$ $(1,719,396)$; $311,148$ $(276,229)$ in the UK. $R^2$ values range from $0.95-0.99$ in France, and $0.91-0.97$ in the UK. Values are in Euros (France) and Sterling (UK).

Figure 3: Distribution of the difference between the non-refundable price for a room, and the one-day lagged refundable price for the same room, conditional on the refundable price not being available on the posting date.

(a) **France**



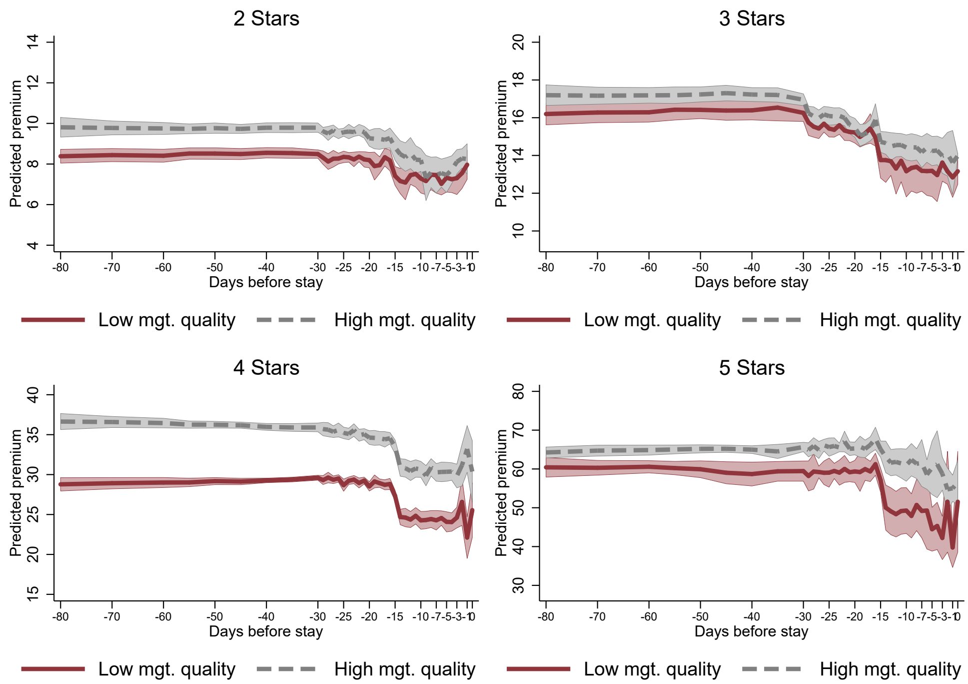

(b) **UK**



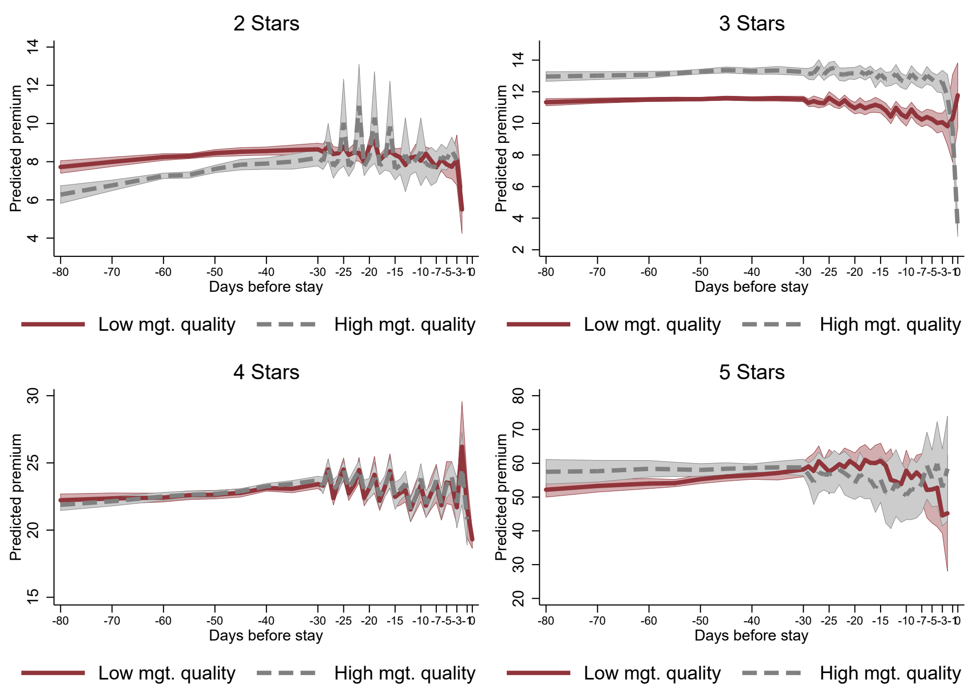

*Notes*: Values winsorized at the first and ninety-ninth percentiles. The price differences on the horizontal axis are in Euros for France, and Sterling for the UK.

Figure 4: Estimated refundable reservation premia by days before check-in, stars, and management quality

(a) **France**



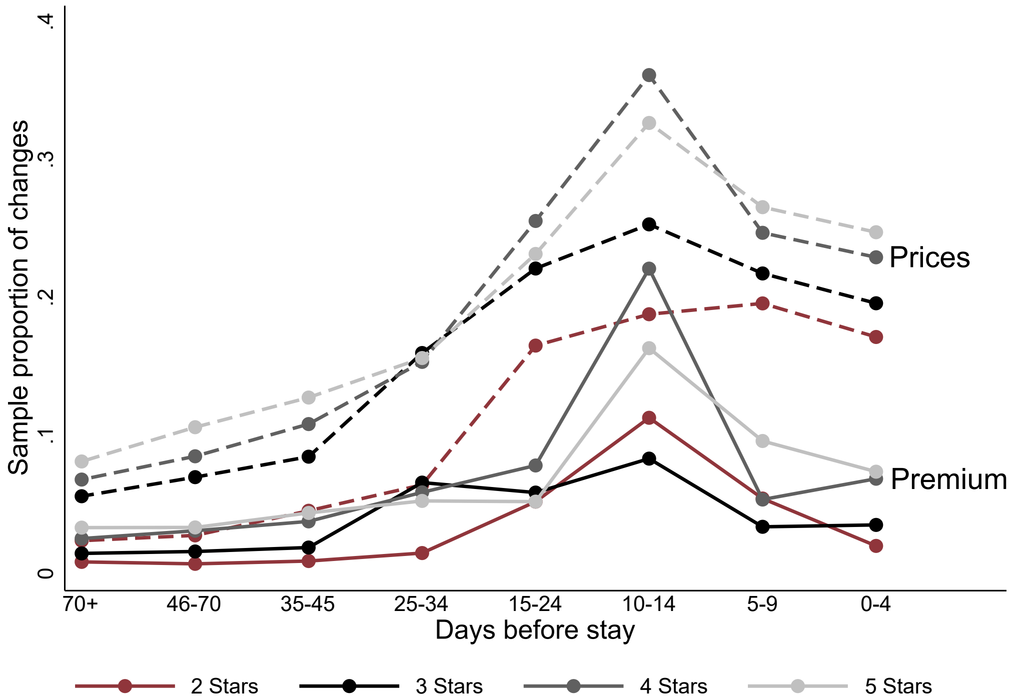

(b) **UK**



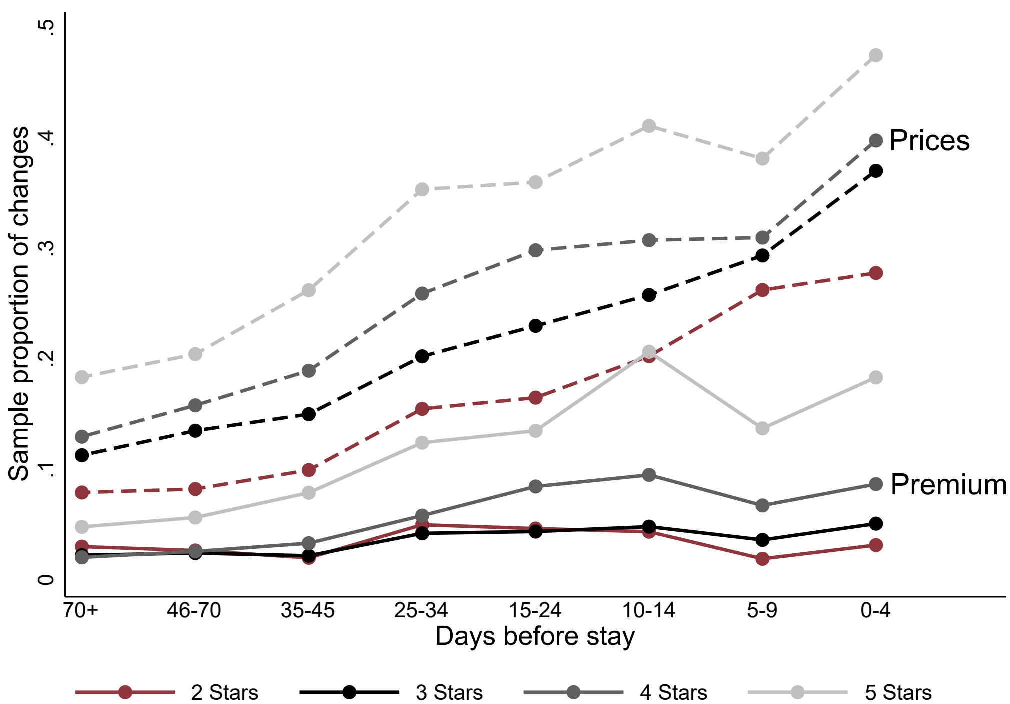

*Notes*: The graphs display the estimated values of the cancellation premium, from model (1). The regressions are separate by stars and country. The shaded areas indicate 95% confidence intervals, with standard errors clustered at the hotel level. Fixed effects combine hotel, check-in date, room type, number of allowed guests, and breakfast inclusion. The number of observations from High (Low) quality 2, 3, 4, and 5-stars hotels are, respectively, 26,709 (36,615); 132,866 (187,757); 752,873 (959,329); 149,519 (165,769) in France, and 19,287 (27,601); 352,753 (418,151); 733,006 (925,157); 119,514 (139,542) in the UK. $R^2$ values range from $0.79 - 0.92$ in France, and $0.80 - 0.90$ in the UK. Values are in Euros (France) and Sterling (UK).

Figure 5: Fraction of rooms experiencing changes in prices and cancellation premia at different lags before stay.

(a) **France**




(b) **UK**



# References


Akerlof, G. A. and Yellen, J. L., editors (1986). *Efficiency wage models of the labor market.* Cambridge University Press., Cambridge , UK.

Anagol, S., Balasubramaniam, V., and Ramadorai, T. (2018). Endowment Effects in the Field: Evidence from India's IPO Lotteries. *Review of Economic Studies*, 85(4):1971–2004.

Anderson, E. T. and Simester, D. I. (2010). Price stickiness and customer antagonism. *Quarterly Journal of Economics*, 125(2):729–765.

Barseghyan, L., Molinari, F., O'Donoghue, T., and Teitelbaum, J. C. (2013). The nature of risk preferences: Evidence from insurance choices. *American Economic Review*, 103(6):2499–2529.

Bils, M. and Klenow, P. J. (2004). Some Evidence on the Importance of Sticky Prices. *Journal of Political Economy*, 112(5):947–985.

Bruhin, A., Fehr-Duda, H., and Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, 78(4):1375–1412.

Cavallo, A. (2017). Are online and offline prices similar? Evidence from large multi-channel retailers. *American Economic Review*, 107(1):283–303.

Chetty, R. (2006). A new method of estimating risk aversion. *American Economic Review*, 96(5):1821–1834.

Cho, S., Lee, G., Rust, J., and Yu, M. (2018). Optimal Dynamic Hotel Pricing. mimeo.

Choi, S. and Mattila, A. S. (2003). Hotel revenue management and its impact on customers' perception of fairness. *Journal of Revenue and Pricing Management*, 2(4):303–314.

DellaVigna, S. and Gentzkow, M. (2019). Uniform Pricing in U.S. Retail Chains. *The Quarterly Journal of Economics*, 134(4):2011–2084.

Eliaz, K. and Spiegler, R. (2008). Consumer optimism and price discrimination. *Theoretical Economics*, 3(4):459–497.

Escobari, D. and Jindapon, P. (2014). Price Discrimination through Refund Contracts in Airlines. *International Journal of Industrial Organization*, 34(3):1–8.

Falk, M. and Vieru, M. (2018). Modelling the cancellation behaviour of hotel guests. *International Journal of Contemporary Hospitality Management*, 30(10):3110–3116.

Goldfarb, A. and Xiao, M. (2011). Who thinks about the competition? Managerial ability and strategic entry in US local telephone markets. *American Economic Review*, 101(7):3130–3161.

Hartley, R., Lanot, G., and Walker, I. (2014). Who really wants to be a millionaire? Estimates of risk aversion from gameshow data. *Journal of Applied Econometrics*, 29(6):861–879.

Heidhues, P. and Kőszegi, B. (2008). Competition and price variation when consumers are loss averse. *American Economic Review*, 98(4):1245–1268.

Heidhues, P. and Kőszegi, B. (2010). Exploiting naivetè about self-control in the credit market. *American Economic Review*, 100(5):2279–2303.

Heidhues, P. and Kőszegi, B. (2017). Naivetè-based discrimination. *Quarterly Journal of Economics*, 132(2):1019–1054.

Hollenbeck, B. (2017). The economic advantages of chain organization. *Rand Journal of Economics*, 48(4):1103–1135.

Jindal, P. (2015). Risk preferences and demand drivers of extended warranties. *Marketing Science*, 34(1):39–58.

Kahneman, D., Knetsch, J. L., and Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *American Economic Review*, 74:728–741.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: an Analysis of Decision under Risk. *Econometrica*, 47:263–291.

Katz, M. L. (1983). Non-uniform pricing, output and welfare under monopoly. *Review of Economic Studies*, 50(1):37–56.

Kőszegi, B. and Rabin, M. (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics*, 121(4):1133–1166.

Kőszegi, B. and Rabin, M. (2007). Reference-dependent risk attitudes. *American Economic Review*, 97(4):1047–1073.

Kosova, R. and Lafontaine, F. (2012). Much ado about chains: A research agenda. *International Journal of Industrial Organization*, 30(3):303–308.

Kosova, R., Lafontaine, F., and Perrigot, R. (2013). Organizational form and performance: evidence from the hotel industry. *Review of Economics and Statistics*, 95(4):1303–1323.

List, J. (2003). Does Market Experience Eliminate Market Anomalies. *Quarterly Journal of Economics*, 118(1):41–71.

Mantovani, A., Piga, C. A., and Reggiani, C. (2021). Online Platform Price Parity Clauses: Evidence from the EU Booking.com Case. *European Economic Review*, forthcoming, available at SSRN: https://ssrn.com/abstract=3381299.

Melis, G. and Piga, C. A. (2017). Are all hotel prices created dynamic? An empirical assessment. *International Journal of Hospitality Management*, 67:163–173.

Nakamura, E. and Steinsson, J. (2008). Five Facts about Prices: A REevaluation of Menu Costs Models. *Quarterly Journal of Economics*, 123(4):1415–1464.

Rabin, M. (2000). Risk Aversion and Expected-Utility Theory: A Calibration Theorem. *Econometrica*, 68(5):1281–1292.

Sandroni, A. and Squintani, F. (2013). Overconfidence and asymmetric information: The case of insurance. *Journal of Economic Behavior and Organization*, 93:149–165.

Schmalensee, R. (1981). Output and welfare implications of monopolistic third-degree price discrimination. *American Economic Review*, 71(1):242–247.

Sydnor, J. (2010). (Over) insuring modest risks. *American Economic Journal: Applied Economics*, 2(4):177–199.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323.

Vives, A., Jacob, M., and Payeras, M. (2018). Revenue management and price optimization techniques in the hotel sector: A critical literature review. *Tourism Management*, 24(6):720–752.

Wu, G. and Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science*, 42(12):1676–1690.

# Online Appendix - not for publication

# A Examples of posted data

Figure A.1: Examples of prices posted on a London hotel's page

| Room type | Sleeps | Today's price | Your choices |
|---|---|---|---|
| Standard Twin Room<br>Only 7 rooms left on our site!<br>2 single beds<br>183 feet² Air conditioning<br>Private bathroom Flat-screen TV<br>Free WiFi<br>More<br>Prices are per room<br>**Included:** 20 % VAT, Breakfast | 2 | £104<br>includes taxes and charges | Good breakfast included<br>• Non-refundable |
| | 2 | £119<br>includes taxes and charges | Good breakfast included<br>FREE cancellation before 14:00 on 5 March 2020<br>NO PREPAYMENT NEEDED - pay at the property |
| Triple Room<br>Only 7 rooms left on our site!<br>1 double bed and 1 sofa bed<br>205 feet² Air conditioning<br>Private bathroom Flat-screen TV<br>Free WiFi<br>More<br>Prices are per room<br>**Included:** 20 % VAT, Breakfast | 2 | £129<br>includes taxes and charges | Good breakfast included<br>• Non-refundable |
| | 2 | £144<br>includes taxes and charges | Good breakfast included<br>FREE cancellation before 14:00 on 5 March 2020<br>NO PREPAYMENT NEEDED - pay at the property |
| Single Room<br>Only 7 rooms left on our site!<br>1 single bed<br>Air conditioning Private bathroom<br>Flat-screen TV Free WiFi<br>More<br>Prices are per room<br>**Included:** 20 % VAT, Breakfast | 1 | £79<br>includes taxes and charges<br>Great value | Good breakfast included<br>• Non-refundable |
| | 1 | £94<br>includes taxes and charges | Good breakfast included<br>FREE cancellation before 14:00 on 5 March 2020<br>NO PREPAYMENT NEEDED - pay at the property |

# B Sample Statistics

Table B.1: Number of hotels using both Refundable and Non-Refundable prices, relative to full sample

| | Star | | | | | Size | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2* | 3* | 4* | 5* | Total | 1-49 | 50-59 | 100-149 | 150-199 | 200-249 | 250 plus | Total |
| UK Both | 44 | 326 | 440 | 71 | 881 | 83 | 229 | 232 | 138 | 82 | 108 | 872 |
| UK Full | 50 | 370 | 467 | 87 | 974 | 127 | 256 | 246 | 143 | 85 | 117 | 974 |
| France Both | 87 | 317 | 647 | 98 | 1,149 | 368 | 412 | 201 | 84 | 28 | 56 | 1,149 |
| France Full | 129 | 361 | 694 | 112 | 1,296 | 426 | 482 | 213 | 87 | 31 | 57 | 1,296 |

French Cities: Toulouse, St.Etienne, Paris, Nice, Marseille, Lyon, Lille, Lens and Bourdeaux; British Cities: Aberdeen, Bath, Belfast, Birmingham, Blackpool, Brighton, Bristol, Carnarfon, Cambridge, Canterbury, Cardiff, Chester, Dundee, Glasgow, Inverness, Leeds, Liverpool, London, Manchester, Newcastle, Newquay, Norwich, Nottingham, Oxford, Scarborough, Sheffield, Skegness, Swansea, Torquay and York.

Table B.2: Number of observations by type of prices that hotels offer during the booking period

| | UK | | | | FRANCE | | | |
|---|---|---|---|---|---|---|---|---|
| Days from stay | Both R-NR | Only R | Only NR | Total | Both R-NR | Only R | Only NR | Total |
| 0-1 | 1,008 | 14,893 | 37,725 | 53,626 | 6,387 | 25,608 | 4,530 | 36,525 |
| 2 | 10,751 | 28,395 | 9,234 | 48,380 | 11,706 | 26,192 | 2,275 | 40,173 |
| 3-6 | 102,411 | 59,432 | 5,779 | 167,622 | 144,910 | 42,360 | 8,984 | 196,254 |
| 7-9 | 93,521 | 26,718 | 3,004 | 123,243 | 119,142 | 21,882 | 7,096 | 148,120 |
| 10-13 | 117,757 | 31,329 | 3,512 | 152,598 | 144,860 | 24,175 | 8,782 | 177,817 |
| 14-20 | 199,371 | 44,305 | 4,537 | 248,213 | 262,027 | 33,525 | 12,578 | 308,130 |
| 21-29 | 253,968 | 40,598 | 4,792 | 299,358 | 312,682 | 30,645 | 23,948 | 367,275 |
| 30-39 | 304,843 | 40,245 | 5,354 | 350,442 | 300,520 | 24,475 | 52,194 | 377,189 |
| 40-49 | 300,670 | 36,197 | 5,015 | 341,882 | 253,915 | 21,834 | 73,913 | 349,662 |
| 50-59 | 346,296 | 38,789 | 5,358 | 390,443 | 261,833 | 22,138 | 111,183 | 395,154 |
| 60-69 | 284,994 | 31,825 | 4,641 | 321,460 | 198,539 | 17624 | 108,964 | 325,127 |
| 70 plus | 724,421 | 97106 | 13266 | 834793 | 390,916 | 43,503 | 404,296 | 838,715 |
| Total N | 2740011 | 489,832 | 102,217 | 3332060 | 2407437 | 333,961 | 818,743 | 3560141 |

Table B.3: Estimates from model (1). Dependent Variable: Cancellation Premium. France

| days before check-in | 2-star | 3-star | 4-star | 5-star |
|---|---|---|---|---|
| 100 + | 9.161*** | 16.22*** | 31.49*** | 62.04*** |
| 80-99 | -0.0762 | 0.189 | 0.567** | -0.325 |
| 70-79 | -0.172 | 0.396 | 0.779* | 0.128 |
| 65-69 | -0.155 | 0.431 | 0.829* | 0.323 |
| 60-64 | -0.183 | 0.444 | 0.806 | 0.536 |
| 55-59 | -0.127 | 0.528 | 0.725 | 0.450 |
| 50-54 | -0.113 | 0.539 | 0.812 | 0.397 |
| 45-49 | -0.143 | 0.550 | 0.750 | -0.0275 |
| 40-44 | -0.0845 | 0.524 | 0.729 | -0.311 |
| 35-39 | -0.0926 | 0.599 | 0.757 | -0.156 |
| 30-34 | -0.121 | 0.330 | 0.889 | 0.421 |
| 29 | -0.302 | -0.268 | 0.629 | -0.582 |
| 28 | -0.458 | -0.464 | 0.779 | 0.493 |
| 27 | -0.302 | -0.559 | 0.522 | -0.0475 |
| 26 | -0.383 | -0.329 | 0.695 | 0.400 |
| 25 | -0.309 | -0.518 | 0.0617 | 0.194 |
| 24 | -0.298 | -0.559 | 0.328 | 0.547 |
| 23 | -0.372 | -0.410 | 0.609 | 0.213 |
| 22 | -0.247 | -0.619 | 0.0943 | 1.381 |
| 21 | -0.362 | -0.814* | 0.323 | 0.0474 |
| 20 | -0.525 | -0.896* | -0.278 | 0.258 |
| 19 | -0.688 | -1.188*** | 0.0683 | 0.304 |
| 18 | -0.669 | -1.065** | -0.0794 | 0.501 |
| 17 | -0.465 | -0.829* | -0.264 | 0.655 |
| 16 | -0.523 | -0.876* | -0.130 | 2.351 |
| 15 | -1.152** | -2.059*** | -1.448 | -1.109 |
| 14 | -1.454*** | -2.088*** | -4.063*** | -5.736* |
| 13 | -1.537*** | -2.199*** | -4.131*** | -7.042** |
| 12 | -1.256*** | -2.390*** | -4.432*** | -7.378** |
| 11 | -1.354*** | -2.153*** | -4.027*** | -6.947** |
| 10 | -1.501*** | -2.525*** | -4.397*** | -7.260** |
| 9 | -1.954*** | -2.431*** | -4.714*** | -7.320** |
| 8 | -1.664*** | -2.394*** | -4.202*** | -6.315** |
| 7 | -1.741*** | -2.622*** | -4.541*** | -7.122** |
| 6 | -1.933*** | -2.553*** | -4.364*** | -8.329*** |
| 5 | -1.825*** | -2.606*** | -4.633*** | -10.21*** |
| 4 | -1.751*** | -2.722*** | -4.635*** | -9.582** |
| 3 | -1.535*** | -2.439*** | -4.559*** | -11.90*** |
| 2 | -1.305*** | -2.678*** | -2.813* | -9.880*** |
| 1 | -1.089** | -3.052*** | -4.602*** | -14.05*** |
| 0 | - | -2.709*** | -3.487** | -9.147*** |
| N | 63,324 | 319,623 | 1,709,202 | 315,288 |
| $R^2$ | 0.868 | 0.812 | 0.869 | 0.879 |

*Notes*: The table reports the regression estimates used to plot Figure 1a, where the outcome variable is the cancellation price. 100+ denotes the base category, so the estimate represents the predicted value of the premium for time lags longer than 100 days. For earlier days before check-in, the figures are the estimated differences in premium with respect to the base category. Regressions are separate by stars. Robust standard errors clustered at the hotel level. Fixed effects combine hotel, check-in date, room type, number of allowed guests, and breakfast inclusion. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively. Values are in Euros.

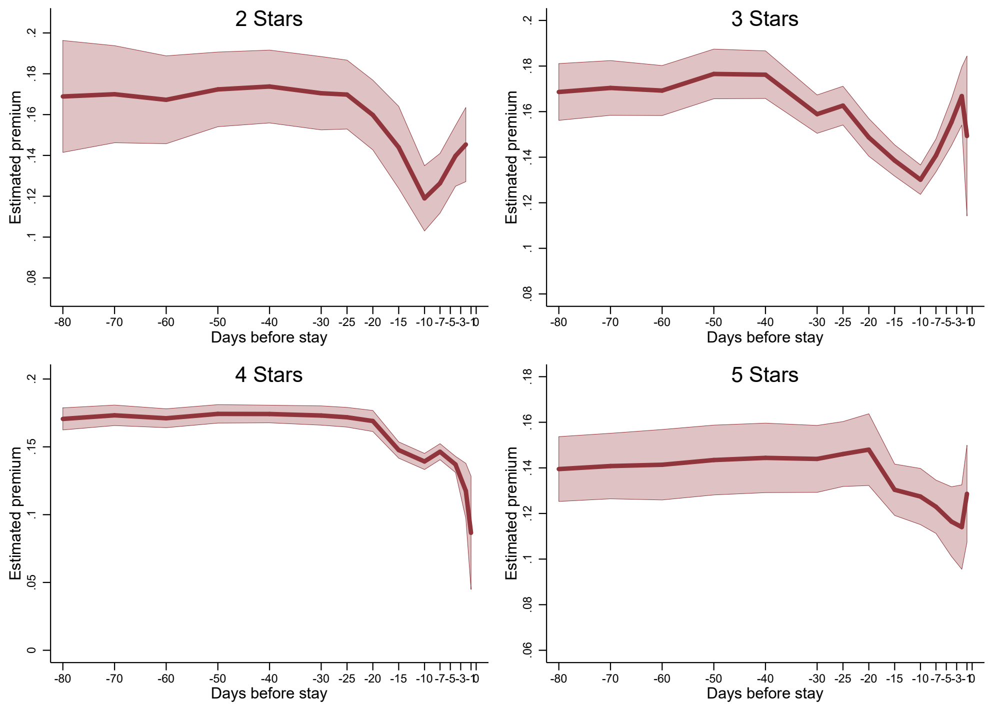
Table B.4: Estimates from model (1). Dependent Variable: Cancellation Premium.

| Days before check-in | 2-star | 3-star | 4-star | 5-star |
|---|---|---|---|---|
| 100 + | 6.304*** | 12.45*** | 21.47*** | 54.82*** |
| 80-99 | 0.320*** | -0.253** | 0.264** | 0.0266 |
| 70-79 | 0.817*** | -0.373*** | 0.599*** | -0.188 |
| 65-69 | 1.179*** | -0.291* | 0.794*** | 0.457 |
| 60-64 | 1.533*** | -0.227 | 0.986*** | 1.178 |
| 55-59 | 1.584*** | -0.179 | 1.141*** | 1.172 |
| 50-54 | 1.808*** | -0.122 | 1.194*** | 1.715 |
| 45-49 | 1.938*** | -0.0454 | 1.352*** | 2.281 |
| 40-44 | 1.982*** | -0.0990 | 1.719*** | 2.613 |
| 35-39 | 2.055*** | -0.0689 | 1.771*** | 2.997 |
| 30-34 | 2.160*** | -0.128 | 2.077*** | 3.523 |
| 29 | 2.007*** | -0.315 | 1.851*** | 2.915 |
| 28 | 2.604*** | -0.240 | 3.069*** | 1.241 |
| 27 | 1.828*** | -0.111 | 1.177*** | 4.031 |
| 26 | 1.982*** | -0.343 | 2.095*** | 3.018 |
| 25 | 2.965*** | -0.0114 | 3.023*** | 2.229 |
| 24 | 1.806*** | -0.0854 | 1.432*** | 2.990 |
| 23 | 2.022*** | -0.377 | 2.084*** | 4.007 |
| 22 | 3.198*** | -0.245 | 2.989*** | 3.028 |
| 21 | 1.579*** | -0.376 | 1.168** | 3.483 |
| 20 | 2.267*** | -0.487* | 1.652*** | 2.014 |
| 19 | 3.283*** | -0.328 | 2.553*** | 2.647 |
| 18 | 1.633*** | -0.522** | 0.874* | 3.026 |
| 17 | 1.911*** | -0.368 | 1.757*** | 2.850 |
| 16 | 2.674*** | -0.543* | 2.887*** | 1.698 |
| 15 | 1.754*** | -0.450* | 1.042* | 2.595 |
| 14 | 1.745*** | -0.735** | 1.412** | 0.527 |
| 13 | 1.791*** | -0.923*** | 1.740*** | -1.877 |
| 12 | 1.737*** | -0.484* | 0.0720 | -0.657 |
| 11 | 1.880*** | -0.901*** | 0.989 | -2.371 |
| 10 | 1.875*** | -1.060*** | 2.097*** | -0.781 |
| 9 | 1.901*** | -0.532* | 0.406 | -0.862 |
| 8 | 1.625*** | -0.907*** | 1.206** | -0.388 |
| 7 | 1.419*** | -1.112*** | 2.123*** | -0.0276 |
| 6 | 1.852*** | -0.828*** | 0.453 | 0.0110 |
| 5 | 1.604*** | -1.020*** | 1.998*** | -1.202 |
| 4 | 1.668*** | -1.271*** | 2.236*** | 0.845 |
| 3 | 1.718*** | -1.171*** | 0.376 | -7.314** |
| 2 | -1.060 | -1.773*** | 3.937*** | -4.153 |
| 1 | - | -2.488*** | -0.162 | - |
| 0 | - | -0.223 | -2.073*** | - |
| N | 46,888 | 770,904 | 1,663,163 | 259,056 |
| $R^2$ | 0.848 | 0.858 | 0.860 | 0.879 |

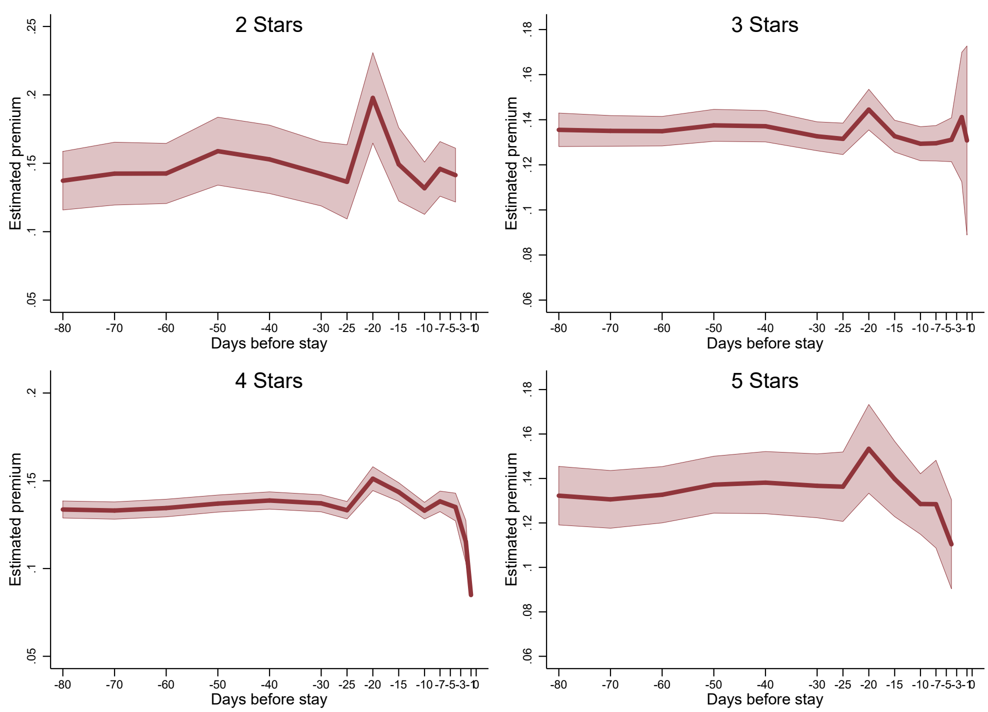
*Notes*: The table reports the regression estimates used to plot Figure 1b, where the outcome variable is the cancellation price. 100+ denotes the base category, so the estimate represents the predicted value of the premium for time lags longer than 100 days. For earlier days before check-in, the figures are the estimated differences in premium with respect to the base category. Regressions are separate by stars. Robust standard errors clustered at the hotel level. Fixed effects combine hotel, check-in date, room type, number of allowed guests, and breakfast inclusion. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively. Values are in Sterling.

Figure B.1: Estimated percentage refundable reservation premia by days before check-in and stars

(a) France




(b) UK




*Notes*: The regressions are separate by stars and country. The shaded areas indicate 95% confidence intervals, with standard errors clustered at the hotel level. Fixed effects combine hotel, check-in date, room type, number of allowed guests, and breakfast inclusion. The number of observations from 2, 3, 4, and 5-stars hotels are, respectively, $126,648$; $639,246$; $3,418,404$; $630,576$ in France, and $93,776$; $1,541,808$; $3,326,326$; $518,112$ in the UK. $R^2$ values range from $0.96-0.98$ in France, and $0.93-0.97$ in the UK.

# C Solving the model

In this Appendix, we fully characterize the solution of the model, which leads to Proposition 1 in the paper.

## C.1 The optimal tariff menu: the general case

In general, the firm can offer three typologies of menus: i) a menu in which both the refundable and the non-refundable tariffs are chosen in equilibrium by at least a type $\varpi \in \{LS, LN, HS, HN\}$; ii) a menu such that only the refundable tariff is chosen; or iii) a menu with only the non-refundable tariff being selected. For each consumer type to buy in equilibrium, their participation ($PC$) and incentive compatibility ($IC$) constraints must hold.

The following lemmas simplify the firm's profit maximization problem:

**Lemma 1** *Suppose that $PC_\omega$ is satisfied for a type $\omega$ selecting $p_{NR}$. Then $IC_{\omega'}$ implies $PC_{\omega'}$ for all types $\omega$ displaying $c_{\omega'} > c_\omega$.*

***Proof*** *If $U_\omega^{NR} \geq 0$, then $c_\omega \geq p_{NR}$. It follows that $U_{\omega'}^{NR} \geq 0$ because $c_{\omega'} > c_\omega \geq p_{NR}$. If type $\omega'$ selects $p_R$, then $U_{\omega'}^{R} \geq U_{\omega'}^{NR} \geq 0$.* ■

**Lemma 2** *Suppose that type $\theta S$ ($\theta = L, H$) selects $p_R$. Then type $\theta N$ selects $p_R$ as well.*

**Proof** *The proof is by contradiction. Suppose that $U_{\theta S}^{R} \geq U_{\theta S}^{NR}$ and $U_{\theta N}^{NR} \geq U_{\theta N}^{R}$. The two conditions can be re-written as:*

$$\pi \left[u(v_\theta - p_{NR}) - u(v_\theta - p_R) - u(-p_{NR})\right] \leq -u(-p_{NR})$$

$$g(\pi) \left[u(v_L - p_{NR}) - u(v_\theta - p_R) - u(-p_{NR})\right] \geq -u(-p_{NR}).$$

*Because $\pi > g(\pi)$, the two inequalities are incompatible.* ■

**Lemma 3** *Suppose that type $LS$ ($LN$) selects $p_R$. Then type $HS$ ($HN$) selects $p_R$ as well.*

***Proof*** *Let us consider the case of sophisticated types first. The proof is by contradiction. Suppose that $U_{LS}^{R} \geq U_{LS}^{NR}$ and $U_{HS}^{NR} \geq U_{HS}^{R}$. The two conditions can be re-written as*

$$\pi \left[u(v_\theta - p_{NR}) - u(v_L - p_R)\right] \leq -(1-\pi)u(-p_{NR})$$

$$\pi \left[u(v_H - p_{NR}) - u(v_H - p_R)\right] \geq -(1-\pi)u(-p_{NR}).$$

*Due to concavity of* $u(\circ)$, $u(v_H - p_{NR}) - u(v_H - p_R) < u(v_L - p_{NR}) - u(v_L - p_R)$. *Therefore, the two inequalities are incompatible. The proof for naive types is obtained substituting* $g(\pi)$ *to* $\pi$. ■

Lemma 1 shows that the reservation price for non-refundable tariff identifies an ordering over type exclusion when such a tariff is part of the equilibrium menu: if type $\omega$ is served in equilibrium, then all types $\omega'$ for which $c_{\omega'} > c_\omega$ must also be served. Lemmas 2 and 3 derive from the fact the attractiveness of refundable tariffs is higher for those customers who have more to gain to get insured, because they have a higher (subjective) probability of not enjoying the service (Lemma 2) or they have a higher valuation for it (Lemma 3).

In the first group of equilibrium configurations, where each tariff is chosen by at least one type, Lemmas 1-3 implies that there are six configurations to consider. For each configuration, we can determine the candidate equilibrium tariffs by setting $p^*_{NR}$ equal to the lowest reservation price for types selecting such a tariff, and $p^*_R$ such that the incentive compatibility constraints of types selecting in equilibrium such a tariff hold, with at least one of them with an equality sign.

We summarize the candidate equilibrium tariffs in the following propositions:

**Proposition 2** *Suppose* $HS$, $HL$ *and* $LN$ *select the refundable tariff and* $LS$ *selects the non-refundable tariff (configuration 1). Then* $p^*_{NR} = c_{LS}$ *and* $p^*_R = v_L$. *The firm's expected profit is* $v_L\pi(N_H + N_L(1-\beta)) + c_{LS}\beta N_L$.

***Proof*** *The eight constraints for expected profit maximization problem are:*

$$\pi u(v_H - p_R) \geq \pi u(v_H - p_{NR}) + (1-\pi)u(-p_{NR}) \qquad (IC_{HS})$$

$$g(\pi)u(v_H - p_R) \geq g(\pi)u(v_H - p_{NR}) + (1-g(\pi))u(-p_{NR}) \qquad (IC_{HN})$$

$$g(\pi)u(v_L - p_R) \geq g(\pi)u(v_L - p_{NR}) + (1-g(\pi))u(-p_{NR}) \qquad (IC_{LN})$$

$$\pi u(v_L - p_{NR}) + (1-\pi)u(-p_{NR}) \geq \pi u(v_L - p_R) \qquad (IC_{LS})$$

$$\pi u(v_H - p_R) \geq 0 \qquad (PC_{HS})$$

$$g(\pi)u(v_H - p_R) \geq 0 \qquad (PC_{HN})$$

$$g(\pi)u(v_L - p_R) \geq 0 \qquad (PC_{LN})$$

$$\pi u(v_L - p_{NR}) + (1-\pi)u(-p_{NR}) \geq 0 \qquad (PC_{LS})$$

$PC_{LN}$ *implies* $PC_{HS}$, $PC_{HN}$ *and* $PC_{LS}$ *as from Lemma 1, whereas* $IC_{HS}$ *implies* $IC_{HN}$ *following 2.* $PC_{LN}$ *is binding when* $p_R = v_L$. *We ignore* $IC_{HS}$ *for now.* $p_R = v_L$ *implies that* $IC_{LS}$ *is satisfied if* $p^*_{NR} \leq c_{LS}$. *If this constraint is binding, then* $IC_{LN}$ *holds because*

*$c_{LS} \geq c_{LN}$. Finally, if we substitute $p_R = v_L$ and $p_{NR} = c_{LS}$ into $IC_{HS}$ we obtain*

$$\pi u(v_H - v_L) \geq \pi u(v_H - c_{LS}) + (1-\pi)u(-c_{LS}),$$

*which always holds because $c_{LS} < v_L$.* ■

**Proposition 3** *Suppose $HS$ and $HN$ select the refundable tariff and $LS$ and $LN$ select the non-refundable tariff (configuration 2). Then $p^*_{NR} = c_{LN}$ and $p^*_R = m_2$ where $m_2$ is the solution of $\pi u(v_H - m_2) = \pi u(v_H - c_{LN}) + (1-\pi)u(-c_{LN})$. The firm's expected profit is $m_2 \pi N_H + c_{LN} N_L$.*

***Proof*** *The eight constraints of the expected profit maximization program are:*

$$\begin{aligned}
&\pi u(v_H - p_R) \geq \pi u(v_H - p_{NR}) + (1-\pi)u(-p_{NR}) && (IC_{HS})\\
&g(\pi)u(v_H - p_R) \geq g(\pi)u(v_H - p_{NR}) + (1-g(\pi))u(-p_{NR}) && (IC_{HN})\\
&g(\pi)u(v_L - p_{NR}) + (1-g(\pi))u(-p_{NR}) \geq g(\pi)u(v_L - p_R) && (IC_{LN})\\
&\pi u(v_L - p_{NR}) + (1-\pi)u(-p_{NR}) \geq \pi u(v_L - p_R) && (IC_{LS})\\
&\pi u(v_H - p_R) \geq 0 && (PC_{HS})\\
&g(\pi)u(v_H - p_R) \geq 0 && (PC_{HN})\\
&g(\pi)u(v_L - p_{NR}) + (1-g(\pi))u(-p_{NR} \geq 0 && (PC_{LN})\\
&\pi u(v_L - p_{NR}) + (1-\pi)u(-p_{NR}) \geq 0 && (PC_{LS})
\end{aligned}$$

*$PC_{LN}$ implies $PC_{HS}$, $PC_{HN}$ and $PC_{LS}$ following Lemma 1; $IC_{HS}$ implies $IC_{HN}$ as per Lemma 2. $PC_{LN}$ is binding for $p_{NR} = c_{LN}$. For $p_R = m_2$, $IC_{HS}$ is binding. Because*

$$\pi u(v_H - v_L) > \pi u(v_H - c_{LN}) + (1-\pi)u(-c_{LN}),$$

*it follows that $m_2 > v_L$, which implies that $IC_{LS}$ and $IC_{LN}$ hold.* ■

**Proposition 4** *Suppose $HS$ and $HN$ select the refundable tariff, $LS$ selects the non-refundable tariff and $LN$ does not buy (configuration 3). Then $p_{NR} = c_{LS}$ and $p_R = m_3$ where $m_3$ is the solution of $u(v_H - m_3) = \pi u(v_H - c_{LS}) + (1-g(\pi))u(-c_{LS})$. The firm's expected profit is $m_3 \pi N_H + c_{LS} \beta N_L$.*

***Proof*** *These are the six constraints for the expected profit maximization problem in this case:*

$$\pi u(v_H - p_R) \geq \pi u(v_H - p_{NR}) + (1-\pi)u(-p_{NR}) \qquad (IC_{HS})$$
$$g(\pi)u(v_H - p_R) \geq g(\pi)u(v_H - p_{NR}) + (1-g(\pi))u(-p_{NR}) \qquad (IC_{HN})$$
$$\pi u(v_L - p_{NR}) + (1-\pi)u(-p_{NR}) \geq \pi u(v_L - p_R) \qquad (IC_{LS})$$
$$\pi u(v_H - p_R) \geq 0 \qquad (PC_{HS})$$
$$g(\pi)u(v_H - p_R) \geq 0 \qquad (PC_{HN})$$
$$\pi u(v_L - p_{NR}) + (1-\pi)u(-p_{NR}) \geq 0 \qquad (PC_{LS})$$

*$PC_{LS}$ implies $PC_{HS}$ and $PC_{HN}$ (Lemma 1), and $IC_{HS}$ implies $IC_{HN}$ as per Lemma 2. $PC_{LS}$ is binding for $p_{NR} = c_{LS}$. For $p_R = m_3$, $IC_{HS}$ is binding. Moreover, because the following inequality holds:*

$$\pi u(v_H - v_L) > \pi u(v_H - c_{LN}) + (1-\pi)u(-c_{LN}),$$

*it follows that $m_3 > v_L$ , and therefore $IC_{LS}$ is verified. Finally, we observe that LN types would obtain a negative expected utility both from the refundable tariff (since $m_3 > v_L$) and the non-refundable tariff (since $c_{LS} > c_{LN}$).* ■

**Proposition 5** *Suppose HN selects the refundable tariff and HS, LS and LN select the non-refundable tariff (configuration 4). Then, $p^*_{NR} = c_{LN}$ and $p^*_R = m_4$ where $m_4$ is the solution of $g(\pi)u(v_H - m_4) = g(\pi)u(v_H - c_{LN}) + (1-g(\pi))u(-c_{LN})$. The expected profit is $m_4\pi(1-\beta)N_H + c_{LN}(\beta N_H + N_L)$.*

***Proof*** *The eight constraints for expected profit maximization are the following;*

$$\pi u(v_H - p_{NR}) + (1-\pi)u(-p_{NR}) \geq \pi u(v_H - p_R) \qquad (IC_{HS})$$
$$g(\pi)u(v_H - p_R) \geq g(\pi)u(v_H - p_{NR}) + (1-g(\pi))u(-p_{NR}) \qquad (IC_{HN})$$
$$\pi u(v_L - p_{NR}) + (1-\pi)u(-p_{NR}) \geq \pi u(v_L - p_R) \qquad (IC_{LS})$$
$$g(\pi)u(v_L - p_{NR}) + (1-g(\pi))u(-p_{NR}) \geq g(\pi)u(v_L - p_R) \qquad (IC_{LN})$$
$$\pi u(v_H - p_R) \geq 0 \qquad (PC_{HS})$$
$$g(\pi)u(v_H - p_{NR}) + (1-g(\pi))u(-p_{NR}) \geq 0 \qquad (PC_{HN})$$
$$\pi u(v_L - p_{NR}) + (1-\pi)u(-p_{NR}) \geq 0 \qquad (PC_{LS})$$
$$g(\pi)u(v_L - p_{NR}) + (1-g(\pi))u(-p_{NR}) \geq 0 \qquad (PC_{LN})$$

*$PC_{LN}$ implies that all the other participation constraints hold, as per Lemma 1. $PC_{LN}$ is*

*binding for $p_{NR} = c_{LN}$. For $p_R = m_4$, $IC_{HL}$ is binding. $IC_{HL}$ can be re-written as*

$$g(\pi)\left[u(v_H - c_{LN}) - u(v_H - m_f) - u(-c_{LN})\right] = -u(-c_{LN}),$$

*which implies:*

$$\pi\left[u(v_H - c_{LN}) - u(v_H - m_f) - u(-c_{LN})\right] < -u(-c_{LN}),$$

*from which $IC_{HS}$ follows. Because*

$$g(\pi)u(v_H - v_L) > \pi u(v_H - c_{LN}) + (1-\pi)u(-c_{LN}),$$

*then $m_4 > v_L$. It follows that $IC_{LS}$ and $IC_{LN}$ are verified, too.* ■

**Proposition 6** *Suppose $HN$ selects the refundable tariff, $HS$ and $LS$ select the non-refundable tariff and $LN$ does not buy (configuration 5). Then, $p^*_{NR} = c_{LS}$ and $p^*_R = m_5$ where $m_5$ is the solution of $u(v_H - m_h) = g(\pi)u(v_H - c_{LS}) + (1 - g(\pi))u(-c_{LS})$.Firm expected profit is $m_5\pi(N_H(1-\beta)) + c_{LS}\beta(N_H + N_L)$.*

***Proof*** *The six constraints for expected profit maximization are:*

$$\begin{aligned}
&\pi u(v_H - p_{NR}) + (1-\pi)u(-p_{NR}) \geq \pi u(v_H - p_R) && (IC_{HS})\\
&g(\pi)u(v_H - p_R) \geq g(\pi)u(v_H - p_{NR}) + (1-g(\pi))u(-p_{NR}) && (IC_{HN})\\
&\pi u(v_L - p_{NR}) + (1-\pi)u(-p_{NR}) \geq \pi u(v_L - p_R) && (IC_{LS})\\
&\pi u(v_H - p_{NR}) + (1-\pi)u(-p_{NR}) \geq 0 && (PC_{HS})\\
&g(\pi)u(v_H - p_R) \geq 0 && (PC_{HN})\\
&\pi u(v_L - p_{NR}) + (1-\pi)u(-p_{NR}) \geq 0 && (PC_{LS})
\end{aligned}$$

*From Lemma 1, $PC_{LS}$ implies $PC_{HS}$. $PC_{LS}$ is binding for $p_{NR} = c_{LS}$. For $p_R = m_5$, $IC_{HL}$ is binding. We can rewrite $IC_{HL}$ as*

$$g(\pi)\left[u(v_H - c_{LS}) - u(v_H - m_h) - u(-c_{LS})\right] = -u(-c_{LS}),$$

*which implies:*

$$\pi\left[u(v_H - c_{LN}) - u(v_H - m_h) - u(-c_{LN})\right] > -u(-c_L),$$

*from which $IC_{HS}$ follows. Because the following inequality holds:*

$$g(\pi)u(v_H - v_L) > \pi u(v_H - c_{LN}) + (1-\pi)u(-c_{LN}),$$

*then $m_5 > v_L$, so $IC_{LS}$ is verified. Finally, note that $LN$ would derive a negative expected utility both from the refundable tariff ($m_5 > v_L$) and from the non-refundable tariff because $c_{LS} > c_{LN}$.* ■

**Proposition 7** *Suppose $HN$ selects a refundable tariff, $HS$ selects a non-refundable tariff and $LS$ and $LN$ do not buy (configuration 6). Then $p^*_{NR} = c_{HS}$ and $p^*_R = v_H$, and the firm's expected profit is $v_H\pi(1-\beta)N_H + c_{HS}\beta N_H$.*
***Proof*** *The four constraints for expected profit maximization are the following:*

$$\begin{array}{ll}
\pi u(v_H - p_{NR}) + (1-\pi)u(-p_{NR}) \geq \pi u(v_H - p_R) & (IC_{HS}) \\
g(\pi)u(v_H - p_R) \geq g(\pi)u(v_H - p_{NR}) + (1-g(\pi))u(-p_{NR}) & (IC_{HS}) \\
\pi u(v_H - p_R) \geq 0 & (PC_{HS}) \\
g(\pi)u(v_H - p_{NR}) + (1-g(\pi))u(-p_{NR}) \geq 0 & (PC_{HN})
\end{array}$$

*Suppose that both participation constraints are binding. Then $IC_{HS}$ is also binding and $IC_{HN}$ is satisfied. Finally, note that both $LN$ and $LS$ would derive a negative expected utility both from the refundable tariff (because $v_H > v_L$) and the non-refundable tariff because $c_{HS} > c_{LS} > c_{LN}$.* ■

In the second group of configurations, customers choose only the refundable tariff. In this case, the participation constraints do not depend on the degree of sophistication (because $U^{S,R}_\theta \geq 0$ implies $U^{N,R}_\theta \geq 0$ and vice versa). It follows that the candidate equilibrium tariffs are determined by fixing to refundable tariff equal to the lowest valuation among the customers served by the firm. This leads to the following Propositions.

**Proposition 8** *Suppose $HS$, $HL$, $LS$ and $LN$ select the refundable tariff (configuration 7). Then, $p^*_{NR} = p > c_{HS}$ and $p^*_R = v_L$. The firm's expected profit is $\pi v_L(N_H + N_L)$.*

**Proposition 9** *Suppose $HS$ and $HN$ select the refundable tariff and $LS$ and $LN$ do not buy (configuration 8). Then, $p^*_{NR} = p > c_{LS}$ and $p^*_R = v_H$. The firm's expected profit is $\pi v_H N_H$.*

Finally, in the third group of configurations, only the non-refundable tariff is chosen. In this case, given Lemma 1 above, the candidate equilibrium tariffs are determined by fixing to non-refundable tariff equal to the lowest reservation price among the customers served by the firm. Therefore, we can derive the following Propositions:

**Proposition 10** *Suppose $HS$, $HL$, $LS$ and $LN$ select the non-refundable tariff (configuration 9). Then $p^*_{NR} = c_{LN}$ and $p^*_R = p > v^H$.Firm expected profit is $c_{LS}(N_H + N_L)$.*

**Proposition 11** *Suppose $HS, HL$ and $LS$ select the non-refundable tariff (configuration 10). Then $p^*_{NR} = c_{LS}$ and $p^*_R = p > v^H$. Firm expected profit is$c_{LS}(N_H + \beta N_L)$.*

**Proposition 12** *Suppose $HS$, $HN$ and $LS$ select the non-refundable tariff and $LN$ does not buy (configuration 11). Then $p^*_{NR} = c_{HN}$ and $p^*_R = p > v^H$. Firm expected profit is $c_{HN}N_H$.*

**Proposition 13** *Suppose $HS$ selects the non-refundable tariff and $HL$, $LS$ and $LN$ do not buy. (configuration 12). Then $p^*_{NR} = c_{HS}$ and $p^*_R = p > v^H$. Firm expected profit is $c_{HS}\beta N_H$.*

Based on expected profit comparison, the following Proposition shows that we can restrict our attention to seven candidate equilibria.

**Proposition 14** *The candidate equilibrium 7 always guarantees higher expected profit than candidate equilibria 9 and 10. The candidate equilibrium 8 always guarantees higher expected profit than candidate equilibria 6, 11 and 12.*

***Proof*** *8 dominates 6 because $v_H > c_{HS}$. 7 dominates 9 and 10 and 8 dominates 11 and 12 because $\pi v_\theta > c_{\theta S}$. (7) can be re-written as $\pi = \frac{-u(-c_{\theta S})}{u(v_\theta - c_{\theta S}) - u(-c_{\theta S})}$. Multiplying both sides by $\frac{v_\theta}{c_\theta}$ we obtain $\frac{\pi v_\theta}{c_\theta} = \frac{-u(-c_{\theta S})/c_\theta}{u(v_\theta - c_{\theta S}) - u(-c_{\theta S})/v_\theta}$. The right term is greater than 1 because $u''(\circ) < 0$. $\pi v_\theta > c_\theta$ follows.* ■

## C.2 The optimal tariff menu: the case of $\pi \to 1$

In this Section we provide the characterization of the equilibrium for $\pi \to 1$. The expected profits are summarized in table below.

| **Candidate equilibrium** | **Expected profit** |
|---|---|
| 1 | $v_L(N_H + N_L)$ |
| 2 | $c_{LN}(N_H + N_L)$ |
| 3 | $v_L(N_H + \beta N_L)$ |
| 4 | $m_4(1-\beta)N_H + c_{LN}(\beta N_H + N_L)$ |
| 5 | $m_5(N_H(1-\beta)) + v_L\beta(N_H + N_L)$ |
| 7 | $v_L(N_H + N_L)$ |
| 8 | $v_H N_H$ |

The following Proposition shows that in this case we can further restrict our attention to five candidate equilibria only.

**Proposition 15** *The candidate equilibria* 1 *and* 7 *always guarantee higher expected profits than candidate equilibria* 2 *and* 3 *.*

***Proof*** 1 *and* 7 *dominate* 2 *because* $v_L > c_{LN}$*, and* 3 *because* $\beta < 1$*.* ■

The five undominated equilibria can be divided in two groups. Candidate equilibria in the first group do not implement naiveté-based discrimination. Candidate equilibrium 7 corresponds instead to the case where all types select the refundable tariff. Candidate equilibrium 8 corresponds to the case where the firm offers only the refundable tariff, to be selected by type $H$ consumers (both sophisticated and naive). Candidate equilibria 7 and 8 correspond to configurations $III$ and $IV$ in the paper, respectively.

In the second group, candidate equilibria 1, 4 and 5 do exhibit naiveté-based discrimination. In 1, naive low types are pooled with high types (both sophisticated and naive), and select the refundable tariff, whereas the sophisticated low type selects the non-refundable tariff. However, the cancellation premium $p_R^* - p_{NR}^*$ converges to 0 for $\pi \to 1$, because both $p_R^*$ and $p_{NR}^*$ converge to $v_L$, as in Escobari and Jindapon (2014). In this case, candidate equilibria 1 and 7 converge to the same menu in which all consumers end up paying $v_L$ with probability 1. For that reason, without any substantial loss in our discussion, we exclude 1 as possible global solution in the remaining analysis.

In both 4 and 5, naive type $H$ consumers select the refundable tariff, whereas sophisticated type $H$ consumers select the non-refundable tariff. In other words, naiveté leads to overinsurance, because $HN$ is willing to get insured, and pay a cancellation (i.e. insurance) premium, while the sophisticated type with the same valuation is not. The two candidate equilibria differ because in 4, all types (except type $HN$) select the non-refundable tariff (so that the market is covered), whereas in 5 type $LN$ does not buy in equilibrium (so that the market is not covered). The non-refundable tariff in 4 is lower than the non-refundable tariff in 5 to induce $LN$ (who is the type with the lowest reservation price for the non-refundable tariff) to buy; consequently, the refundable tariff in 4 is lower than the refundable tariff in 5, in order for the incentive compatibility constraint of type $HN$ to hold. Candidate equilibria 4 and 5 correspond to configurations $I$ and $II$ in the paper, respectively. The conditions that make each of these two configurations the optimal one are described in the main text.