NBER WORKING PAPER SERIES

STICKY PRICE FOR DECLINING RISK?
BUSINESS STRATEGIES WITH "BEHAVIORAL" CUSTOMERS
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, Revised January 2024

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? Business Strategies with “Behavioral” Customers in the Hotel Industry
Nicola Lacetera, Claudio A. Piga, and Lorenzo Zirulia
NBER Working Paper No. 28456
February 2021, Revised January 2024
JEL No. D21,D22,D91,L11,L83

ABSTRACT

Using data from about twenty-five million hotel room postings in four countries, we document that rather than decreasing to zero as the likelihood of cancellation declines, the difference between the prices for refundable and non-refundable reservations remains positive at roughly 10% to 15% of the full price. A model where travelers have different willingness to pay and some of them overestimate the probability to cancel their trip explains these price-setting patterns more consistently than alternative interpretations. We denote these business strategies as naïveté-based price discrimination. Our data and theory therefore show that this form of apparent inertial behavior of companies regarding a major strategic variable can be an intentional managerial choice. We demonstrate, finally, that this profit-enhancing commitment to limited flexibility may also benefit customers in some cases, by expanding the reach of the market. Thus, strategies that rely on cognitive biases on the demand side may not necessarily exploit consumers.

Nicola Lacetera
University of Toronto
Institute for Management and Innovation
3359 Mississauga Road, Room KN 235
Mississauga, ON L5L 1C6
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

You could be missing out if you don't offer both Non-Refundable and Fully Flexible Rate Plans.
It's beneficial to offer both to keep your rates compelling to a wide range of customers at all times.
(Booking.com - https://partner.booking.com/en-us/solutions/advice/pricing-foundations)

1 Introduction

In sectors such as hospitality, travel, entertainment, sport events and e-commerce, consumers reserve a product or service and pay for it in advance of fruition. As such, the actual fruition, at the time of the reservation, is uncertain. In these cases, 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 over time, in this paper we provide evidence that this does not occur. We propose a theory whereby firms limit or exclude premium changes in order to leverage the systematic tendency of some consumers to overestimate the likelihood of cancelling their purchase. Under certain circumstances, the premium inertia may not only lead companies to capture more value, but may also expand the size of the market.

Our empirical context is the hotel industry. Customers can generally choose, for a given room configuration and dates, between a refundable and a non-refundable payment option. The former 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 and less valuable.

But do we observe prices for the two types of reservations actually converge to one another? To address this question, we collected data from a set of hotel establishments in nine French, thirty-two British, ten American and five Canadian cities, operating on the *Booking.com* platform—a total of about $3,500$ hotels and twenty-five million room postings. The data include information on prices for all rooms at a specific query date, the check-in day, the cancellation policy, and several characteristics of the hotels and rooms. This allows us to study a case of add-on pricing (Ellison, 2005), where the add-on is the cancellation option and all prices are observable.

As expected, prices are higher for refundable options than for the corresponding non-refundable ones. However, cancellation premia remain stable at roughly 10% to 15% of the full price, with little variation as the check-in day nears.

To account for this evidence, we propose a model of price-setting patterns as a form of consumer-side naiveté-based price discrimination, where naiveté indicates the systematic overestimation of 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 in relevant contexts—see for example Sydnor (2010). We derive that firms optimally sort consumers who are heterogeneous in both their valuation and degree of sophistication (i.e., the ability to estimate correctly the probability to cancel) by setting a price menu that includes both a refundable and a non-refundable tariff, with the cancellation premium being positive even when the cancellation probability is small.

This result is reminiscent of previous work on the airline industry by Escobari and Jindapon (2014). This study only includes consumer heterogeneity in the valuation of the service; we show that offering refundable tariffs tailored to consumers who have both high willingness to pay and biased beliefs can be profit-enhancing. Moreover, calibrations based on our model do not support standard risk aversion by travelers as a plausible explanation of the evidence, whereas assuming probability overweighting fits the data for reasonable parameter values.

In addition to providing an explanation for certain types of price rigidity, the model allows us to establish conditions under which these limited price adjustments are both profit-enhancing for companies and beneficial to consumers. This occurs when naiveté-based price discrimination includes a partial form of preference-based (second-degree) price discrimination that would not be possible otherwise. In this case, customers 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.

Existing explanations for the low responsiveness of prices to new information include menu costs (Nakamura and Steinsson, 2008) or, in the case of the labor market, the use of efficiency wages (Akerlof and Yellen, 1986). These views do not apply to the settings that we consider, for example because price adjustments are virtually costless online. Other studies considered psychological tendencies that might account for limited price adjustments, such as concerns for fairness, loss aversion, and inertial behavior of managers (Kahneman et al., 1986; Choi and Mattila, 2003; Anderson and Simester, 2010; DellaVigna and Gentzkow, 2019; Heidhues and Kőszegi, 2008, 2010; Kőszegi and Rabin, 2006, 2007). The literature in economic sociology and evolutionary economics has long considered managerial inertia as well, in the form of reliance on organizational routines or of limited individual agency with respect to structural dynamics at the industry or societal level (Barnett and Pontikes, 2008; Hannan and Freeman, 1984, 1977; Levinthal, 1997; Nelson and Winter, 1982).

Studies in strategic management, in contrast, attribute a more significant role to managerial discretion and initiative. Particularly related to our work is the contention that pricing decisions are the results of elaborate processes and deliberations, and are close to representing commitments rather than decisions that can be changed easily; as such, pricing is a key strategic lever for companies (Dutta et al., 2003; Zbaracki et al., 2004; Ritson et al., 2003). The hotels in our samples are, for the most part, within high-quality chains; they rely on sophisticated organizational architectures and are very active in managing their yield (Hollenbeck, 2017; Kosova et al., 2013; Mantovani et al., 2021). Therefore, one would expect them to not be inertial in their price setting. Our data and theory are closer to the strategic management literature, as they support the claim that the persistently positive premium is an intentional managerial choice. We find, in particular, that when the refundable and nonrefundable prices change during the period ahead of stay, their movements are often fine-tuned to not alter their difference. The evidence, moreover, appears neither consistent with systematic mispricing by managers, nor with the adoption of certain "decoy" pricing strategies in order, for example, to make only one type of tariff appealing to customers. Finally, we also test and find no support for a revenue-management explanation of non-declining premia, according to which offering a cancellation option imposes an opportunity cost on hotels if the option is exercised and the room is left unsold (Gallego and Şahin, 2010).

We therefore provide the literature with insights as to how strategic inertia and managerial agency may actually not be in contrast with each other. A major way in which managers set prices in certain industries is by leveraging potential naiveté on the side of customers (fully rational agents would not be willing to buy a more expensive cancellation option, when the uncertainty is fundamentally resolved) as well as their heterogeneous valuations. In these contexts, strategy makers may profitably commit to limited flexibility, and, in doing so, in some cases also benefit customers. Thus, strategies that rely on cognitive limits on the demand side may not necessarily exploit consumers.

Our model combines the insights from theories of price discrimination with naive consumers and of price rigidity due to non-standard preferences and beliefs. Existing research on price discrimination in the presence of time inconsistency and overconfidence of consumers (Eliaz and Spiegler, 2008; Sandroni and Squintani, 2013; Heidhues and Kőszegi, 2010, 2017) does not aim to explain price rigidity *per se*, but rather to derive optimal pricing menus and their welfare implications. By allowing consumers to be heterogeneous 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. In Ellison (2005), transparent add-on pricing when all consumers are fully rational leads to competitive price discrimination, where all the high-demand types buy both the good and the add-on; in Gabaix and Laibson (2006), add-on prices are hidden and only those consumers who are unaware buy the add-on. In our model, only naive consumers with high evaluations, but not the sophisticated ones, buy the refundable room in equilibrium. Although the add-on price is observable, some consumers buy it because they perceive its value incorrectly.

Research on pricing using posted (scraped) price data has provided useful information about the dynamics of competition in online markets, as well as their possible use to derive price indices (Alderighi et al., 2015; Cavallo et al., 2018; Cavallo, 2018; Escobari, 2012; Gorodnichenko and Talavera, 2017; Gorodnichenko et al., 2018, 2021; Williams, 2022). The reliance on posted prices may raise the concern that they do not correspond to actual prices at which firms expect to make a sale. In this paper, we show that posted refundable fares are often replaced by non-refundable fares when these remain the only option for buyers; this suggests that the refundable prices are set at a level that firms consider appropriate for their relevant markets.

In addition, unlike studies of airline markets in which the scraping of prices also leads to the identification of the unused capacity at the time the prices are posted (Alderighi et al., 2015, 2022; Escobari, 2012), the equivalent match of price and capacity data in the hotel sector is not as easily obtainable, especially for the large number of hotels that we consider. When capacity information is available, it is usually aggregated (e.g., at the monthly level) and concerns only a specific set of hotels belonging to one (or a few) chain (Kosova et al., 2013; Cho et al., 2018). The lack of capacity data may be problematic because our claim that a positive premium reflects the hotels' intention to exploit a segment of consumers may not hold if, instead, hotels keep the premium high to recover the opportunity cost of consumers that exercise the cancellation option, thus potentially leading to unsold rooms. This cost should be higher when hotels are very likely to sell out all their rooms well in advance of the check-in dates. We build on Nicolini et al. (2023) and use two possible proxy variables to identify hotels that, on certain dates, may be reaching full capacity utilization, to separate them from those that do not. The premium remain positive in both clusters, thus supporting our interpretation of the high cancellation premium being driven by characteristics of the demand side.

We describe the data in Section 2, and the key findings in Section 3. In Section 4 we present our model. Section 5 considers and rules out possible alternative explanations for the empirical findings. Section 6 provides concluding remarks. The Appendix includes additional tables and figures, as well as the analytical solution to the model and a few extensions.

2 The data

We rely on two different datasets. Both were obtained by scraping information on hotels and their room offer prices posted on a leading Online Travel Agent (OTA), *Booking.com*. The first dataset comprises European hotels located in France and the UK surveyed in the Fall and Winter of 2017. To emphasize the broader relevance of the pricing mechanism illustrated in this study, as well as its robustness and persistence over time, in 2019 we repeated and improved the data collection and built a second dataset of hotels in the USA and Canada; these additional data also allows us to consider additional interpretations of the findings.

2.1 Data Collection

European Hotels. The collection of data from France and the UK 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 are usually 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.[1]

The data cover stay dates between October 30th, 2017 and January 2nd, 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 a hotel offered; this would not be possible, for example, if we based the query on the city listing. This room-feature information is central to derive a precise value of the cancellation premium while holding all other characteristics constant. 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 a hotel is managed.

North American Hotels. Starting in July 2019, we followed an analogous process to retrieve data from hotels located in ten U.S. and five Canadian cities. In this case, we consider

[1]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 by hand identical to those made by the scraping computers. We could not find any noticeable difference.

a slightly longer arrival (check-in) period, from October 1st, 2019 until Jan. 15th, 2020.[2]

This dataset includes two variables that were not available in the European data: *i*) the expiry date of the cancellation premium, and *ii*) the maximum number of each type of room that can be booked simultaneously on the platform. The latter is particularly important as it indicates how many rooms of a particular type the hotel is offering at each point in time. Its 99th percentile value equals 10; considering the large size of most hotels (see Table 1), the variable therefore does not represent the actual number of rooms of each type that the hotel still has left to sell; moreover, because it remains stable over long period of time, its variation cannot be used as a proxy of how many reservations were made over two retrieval dates, as in the case of the airline industry (Alderighi et al., 2015, 2022).[3] However, we do observe reductions as the check-in date approaches, which we treat as an indicator that a particular type of room is getting closer to being sold-out (Nicolini et al., 2023; Piga and Melis, 2021). Furthermore, to keep the number of daily queries manageable, we restricted the sample to hotels with at least a three-star classification; we included, however, hotels not affiliated with chains.

Prices in each country are in the respective local currency. Figure 1 shows a snapshot of some room postings in New York collected in July 2022, to suggest that the data collection design continues to be relevant even in the post-pandemics period. In addition to the expiry dates of the cancellation option, the figure also shows an example of the second new variable in the North American data: the column "Select rooms" indicates that at most five Traditional King rooms could be reserved simultaneously. Tracking these values as the check-in date approaches offers the opportunity to distinguish which hotels are reaching maximum capacity occupation.

Table 1 reports the distribution of hotels by star and size. 3,511 hotels posted both the Refundable (R) and Non-Refundable (NR) prices: 881 in the UK, 1,149 in France, 1,173 in the USA and 308 in Canada. Hotels in the three and four-stars categories account for the largest proportion in all countries, whereas the share of five-stars hotels is about the same in the two European countries (8%), but larger in the US and Canada (respectively, about 12% and 18%). Hotels in the UK are larger than in France, 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 North American sample includes a larger proportion of hotels with more than 250 rooms and a rather negligible number of hotels with less than 50 rooms.

[2]The dates therefore cover a period that concluded before the onset of the COVID-19 pandemics (only a handful of cases were reported in the US before January 15, and the first reported case in Canada was at the end of January). Given this timeline, we do not expect that the pandemic affected or explains our findings.

[3]That is, after a reservation the variable should automatically decrease by one unit. The stability may be due to the use of channel-manager software that harmonizes the offer on each of the electronic sale channels with the hotel's load factor and rooms' availability.

The full sample includes also hotels always reporting either the R or NR price, but not both, and for which, therefore, no cancellation premium is available. The single-price hotels are a minority (about 10% of the full sample in all countries except Canada), are smaller than the rest, and tend to be two or three-star hotels, i.e., they tend to serve a more price-sensitive customer segment.

Figure 1: Examples of prices posted on a New York hotel's page

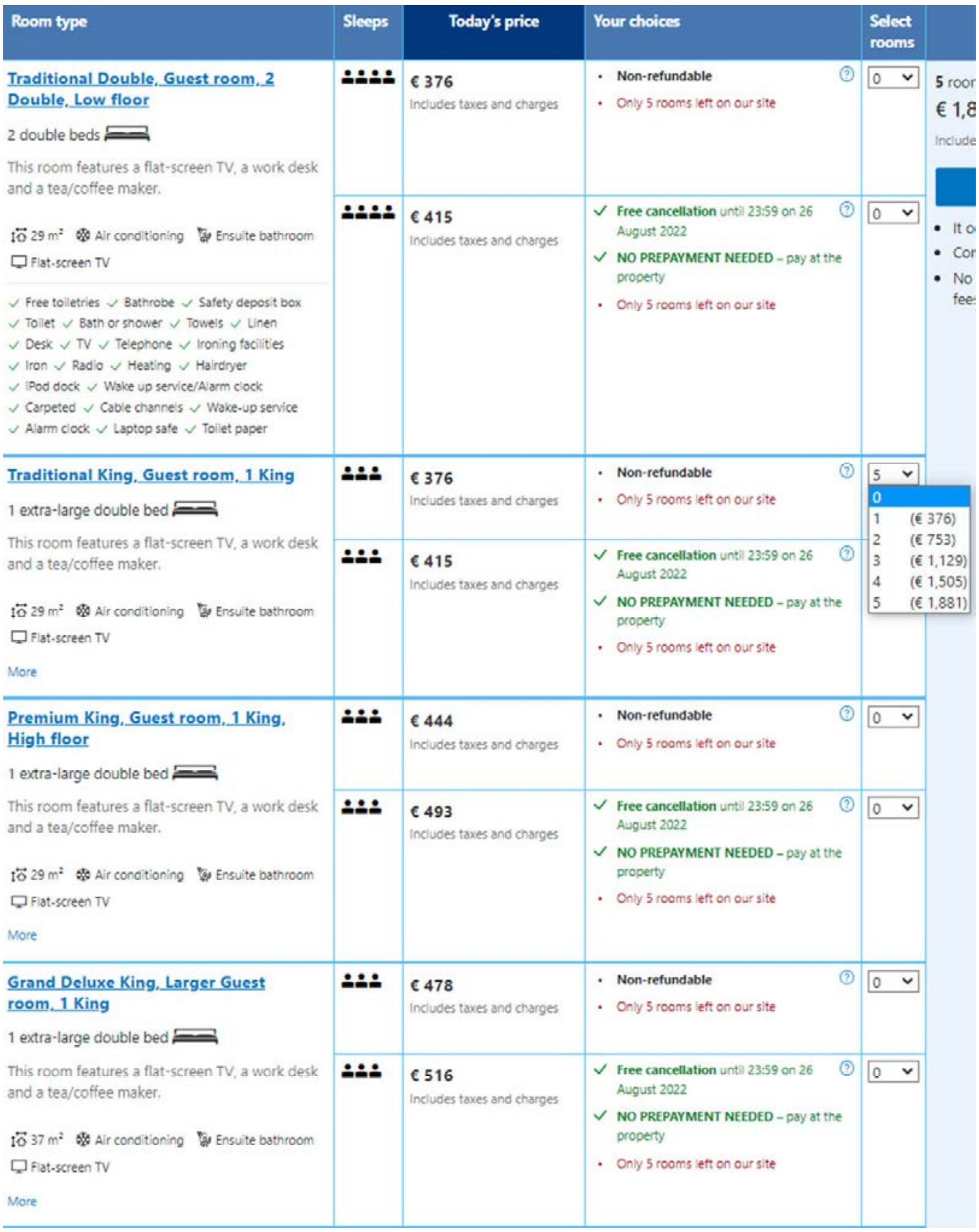


Table 1: Number of hotels using both Refundable and Non-Refundable prices, relative to full sample

	Stars:					Size (rooms)						
Sample:	2	3	4	5	All	1-49	50-99	100-49	150-99	200-49	250+	All
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
USA Both	-	540	495	138	1,173	37	129	170	111	65	225	737
USA Full	-	621	539	207	1,367	77	149	185	124	76	235	846
Canada Both	-	138	108	62	308	20	45	61	38	31	68	263
Canada Full	-	176	130	113	419	43	66	75	48	36	68	336

Notes: Cities where the hotels in our sample are located, by country. **France**: Toulouse, St.Etienne, Paris, Nice, Marseille, Lyon, Lille, Lens and Bourdeaux; **UK**: 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; **USA**: Atlanta, Boston, Chicago, Houston, Los Angeles, Minneapolis, Miami, New York, Portland, Seattle; **Canada**: Calgary, Montreal, Toronto, Vancouver, Winnipeg. Hotel capacity is not available for hotels in Minneapolis, New York and Portland.

The combined datasets comprise over twenty-five million observations, the majority of which (between 65-81%, depending on the country) report both the R and NR price. Hotels 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 the case when we retrieve a larger proportion of observations with only the NR price, in all countries (see Table A.1 in Appendix).

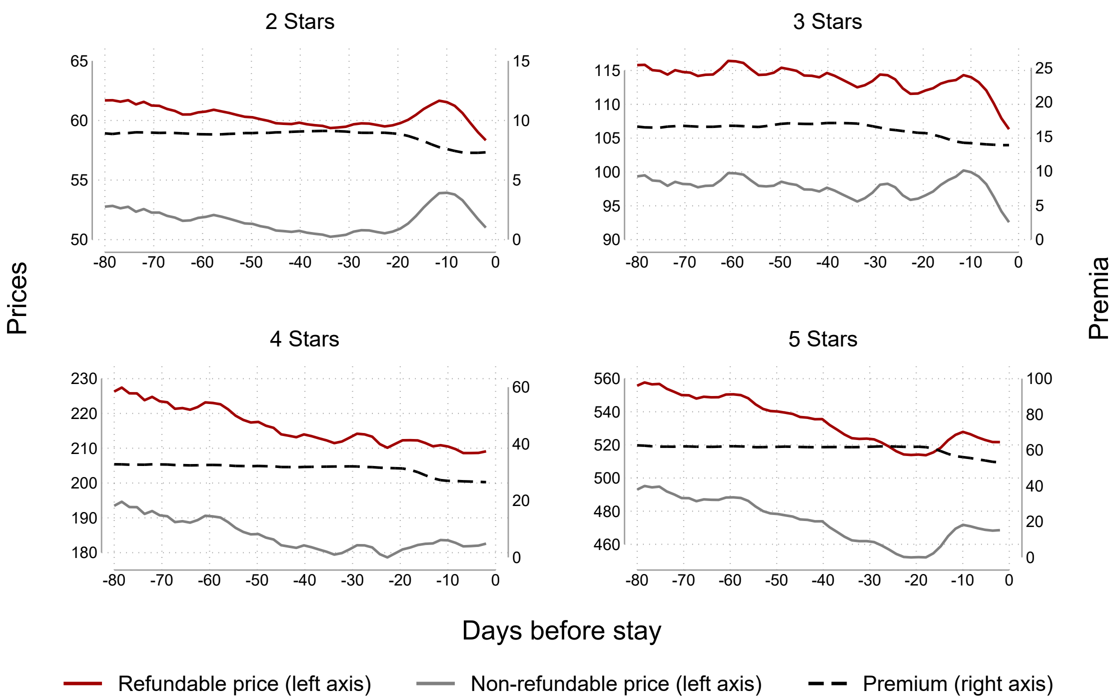
3 Empirical analysis

3.1 The movement of prices and premia: descriptive evidence

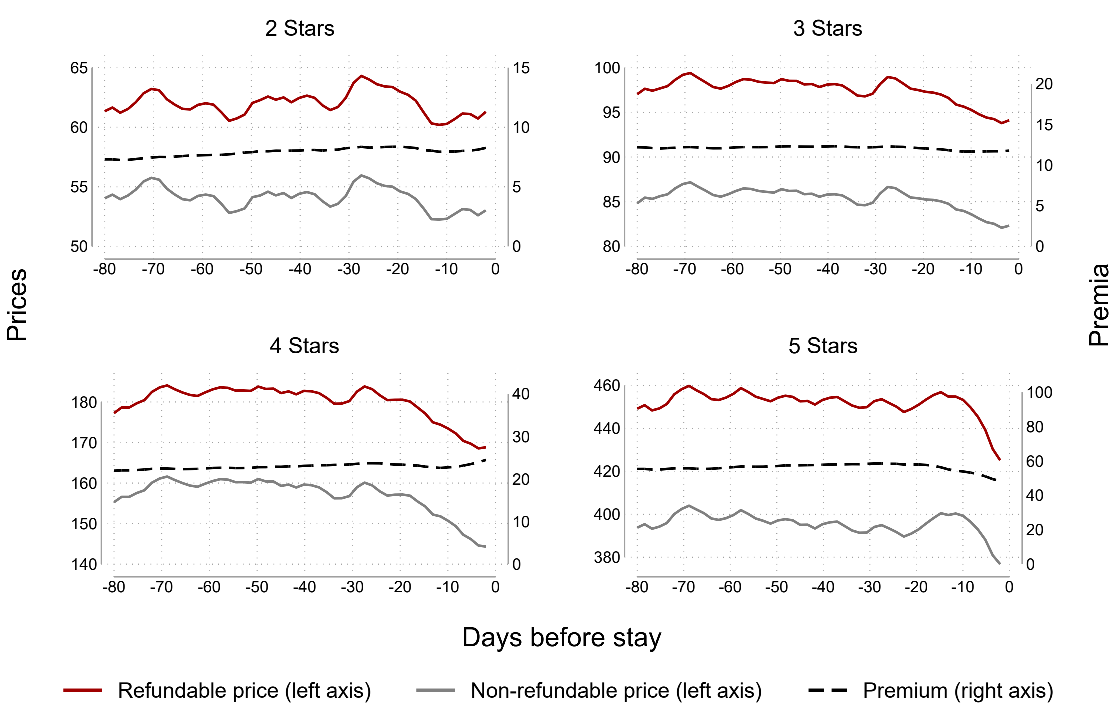
Figure 2 reports the predicted values, from a kernel-weighted local polynomial regression, of the refundable and non-refundable prices and of the cancellation premium (i.e., the difference between the two prices) on the number of days between posting of a room and the related check-in date. Premia are positive and stable for a long period up to the proximity of the fruition day. Refundable and non-refundable prices move in parallel. Changes in prices are therefore much more frequent than changes in premia, suggesting that when hotels adjust both the refundable and non-refundable prices, they most often keep the premium unchanged.

Figure 2: Smoothed values of refundable and non-refundable prices and cancellation premia, by days before check-in and stars

(a) France



(b) UK



Notes: The figures report the predicted values of kernel-weighted local polynomial regressions (degree zero, bandwidth two) of the refundable and non-refundable prices and of the cancellation premium (i.e., the difference between the two prices) on the number of days between posting of a room and the related check-in date.

This evidence is at odds with a basic prediction that cancellation premia should tend to zero because buyers' uncertainty on whether they will travel resolves as the check-in date approaches. Furthermore, because prices fluctuate more than premia, Figure 2 also suggests that the stability of premia may be the result of a deliberate decision by hotel managers.[4]

A further suggestion from Figure 2 is that the setting of both prices creates the opportunity for intertemporal arbitrage through strategic cancellation; a customer may book in advance using the refundable option and, at a later stage, she may cancel and book at the lower non-refundable price (including those offered by competing hotels). If this happened systematically, hotels would likely stop offering both prices simultaneously. Cho et al. (2018) find no evidence of strategic cancellation, even in the presence of arbitrage opportunities. Furthermore, Bachis and Piga (2011) illustrate other online arbitrage opportunities that are strongly persistent over time, possibly due to frictions on the demand side (e.g., searching and evaluating the hidden discounts on airlines' websites). In what follows, we provide further evidence of the stability of premia from regression analyses.

3.2 The persistence of refundable premia

We estimate the relationship between cancellation premia and the time between the posting and reservation date through the following linear regression model:

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

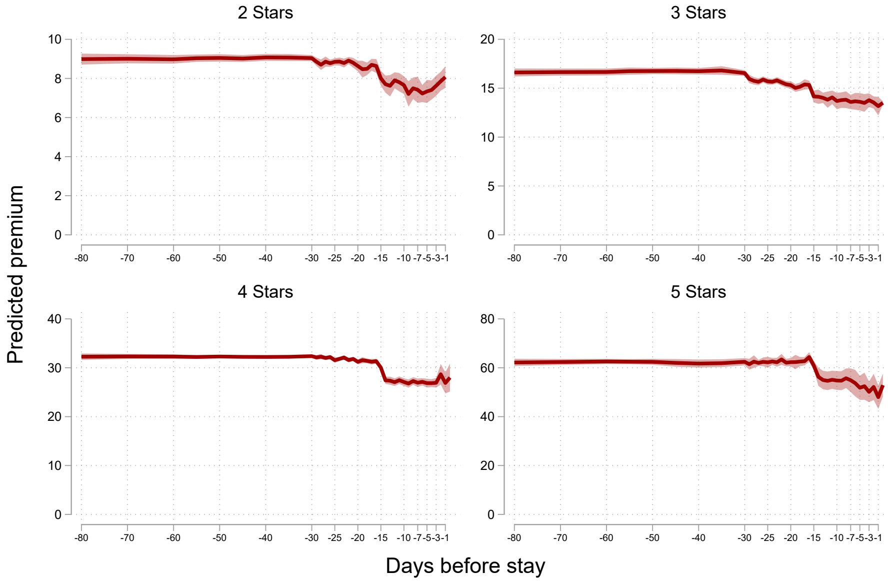
Y is the cancellation premium for a room r in a hotel h, with s stars and in country c. The indicators D_d take the value of one if the posting is d days before the date of the stay, and zero otherwise. The vector X includes indicators of the combination of hotel, check-in date, room type (double, luxury double, triple, ...), the maximum number of guests allowed in the room, and the inclusion of breakfast. Using Figure 1 as reference, one observation in our regressions would represent the "Traditional King" room in five-star hotel h in New York, allowing a maximum of three guests, with, say, check-in date 6 Nov. 2019, whose prices ($415 and $376) for, respectively, the refundable and non-refundable option (breakfast excluded) are observed on date 14 Oct. 2019 (twenty-three days before the check-in date). These combinations are akin to detailed fixed effects that allow to account for different features or "bundles" that might differentiate refundable and non-refundable options. Note, in particular, that these indicators also absorb city, hotel, and chain fixed effects. We allow the error terms ε to be correlated within hotels by estimating the standard errors at this level of aggregation.

We then estimate the parameters δ_d and β_d and plot the predicted values $\hat{Y}_{rh|s=i,d=D}$ in Figures 3 and 4, at different times before check-in, separately by number of stars and country.

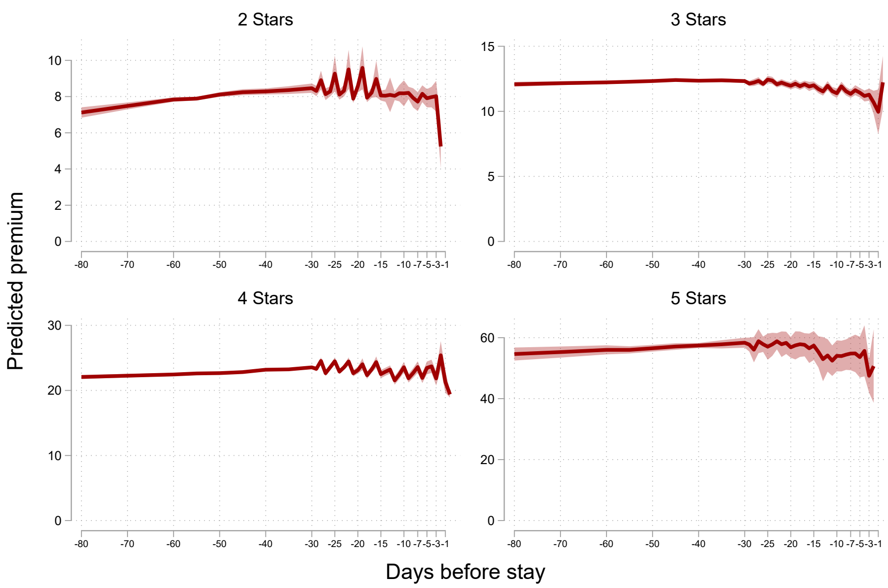
[4] We investigate the frequency of change of prices and premia in detail in Section 5.

Figure 3: Estimated cancellation premia by days before check-in and stars

(a) France



(b) UK

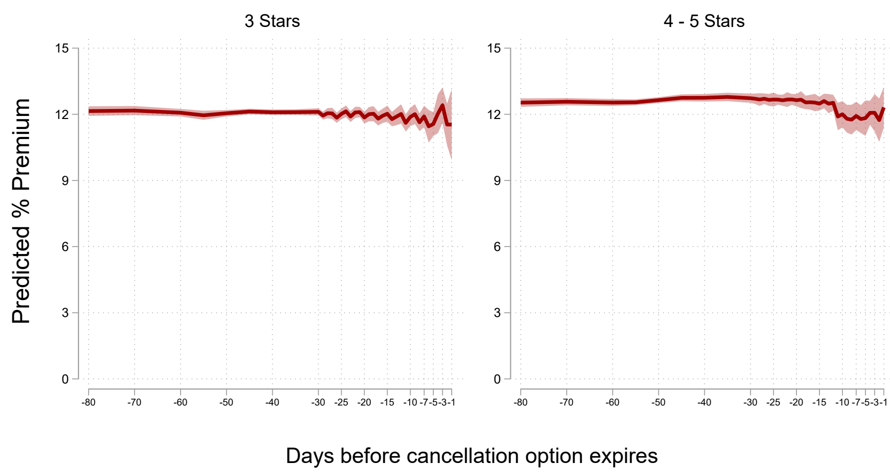


Notes: The graphs display the estimated values of the cancellation premium from model (1). The regressions are separate by stars and country. The shaded areas indicate 95% confidence intervals, with standard errors clustered at the hotel level. Fixed effects combine hotel, check-in date, room type, number 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 to 0.88 in France, and in the $0.85-0.88$ interval in the UK. Monetary values are in Euros (France) and Sterling (UK).

Figure 4: Estimated percentage cancellation premia by days before cancellation option expires and stars

(a) USA

(b) Canada



Notes: The graphs display the estimated values of the percentage 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 3 and 4-5 stars hotels are, respectively, $2,416,623$ and $3,366,508$ in USA, and $828,902$ and $905,018$ in Canada. R^2 values range in the $0.70-0.80$ interval in the in US, and between 0.74 and 0.75 in Canada.

Figure 3 refers to the European sample. The estimates are stable over a time lag between posting and check-in date from eighty days to thirty days. Subsequently, we observe a "step-wise" small reduction of the premium in the two weeks before the check-in date. In no cases, however, does the premium drop or decrease smoothly to zero in proximity of the date of stay. Estimates not shown to save space indicate that the premium expressed as a percentage of the refundable price, on average, ranges between 12% and 16%, with confidence intervals remaining far from the zero threshold.

Figure 4 shows the estimated premia from the same regression models for the North American hotels. In this case, we pooled together 4- and 5-star hotels in order to have two subsamples of similar size for each country. Because price levels (and size of the premium) depend on the quality (stars) of a hotel, in this analysis we took the natural logarithm of the premium as the outcome variable. Looking at the premium in percentage terms also provides a complementary way to analyze the data. Note also that, in this case, the x-axis denotes the number of days separating the collection date from the day of the cancellation option expiry. Nevertheless, the percentage premium in both North American countries does not fall to zero and remains just at or above the 10% level.

4 A model of naiveté-based pricing

Demand-side considerations motivate the basic prediction that cancellation premia should tend to zero as the uncertainty of travel resolves. However, we find consistent evidence of persistence of cancellation premia and lack of convergence between the refundable and the non-refundable prices. In this section, we claim that a deliberate form of naiveté-based menu pricing (Heidhues and Kőszegi, 2017), when some consumers overweight small probabilities of cancellation, can be an explanation for this evidence. Because over-weighting of small probabilities is a well-documented bias, we believe that we provide a plausible explanation for the economic phenomenon under analysis. In Section 5, we show that a few alternative explanations appear not to be supported by the data.

4.1 Model description

The model builds on Escobari and Jindapon (2014), and considers a monopolistic firm. Given the competitive nature of the hotel industry, this assumption may seem restrictive. In Section C of the Appendix, we show that the main results of the model extend to a duopoly setting with competitive price discrimination. We keep the monopoly version here for the sake of simplicity. In addition, the model described here does not consider the role of firm's capacity, by assuming that the firm can serve all consumers at each period. Section D of the Appendix provides an extension of the model that includes a capacity constraint. We briefly discuss this extension also at the end of Section 4.3 below.

The firm provides, at zero marginal cost, a service that consumers value at either v_H (high-valuation) or v_L (low-valuation), with $v_L < v_H$. We denote consumers' valuation type with θ ($= H, L$). Consumers make their purchasing decisions at period τ (i.e., the booking date), but actually enjoy the service at period t (i.e., the arrival or check-in date) with probability $\pi_\theta(\tau) \in (0,1)$, $\pi'(\tau) > 0$. With the complementary probability $1 - \pi_\theta(\tau)$, consumers do not enjoy the service and receive zero value. We assume $\pi_H(\tau) < \pi_L(\tau)$. This captures the distinction between business and leisure travelers, with the former having a higher valuation, but also a higher probability of cancellation than the latter. The cohort of consumers entering at τ for consumption at t includes $N_H(\tau, t)$ consumers of type H and $N_L(\tau, t)$ consumers of type L. Because the firm maximizes the expected short-run profits at τ, we simplify notation and omit the reference to dates from now on.[5]

A fraction β of customers has a correct perception of π_θ, with β independent of the valuation type.[6] We define these consumers *sophisticated*, and indicate their type with S. The remaining fraction $1-\beta$ holds a distorted perception of π_θ, $g(\pi)$, with $g'(\pi) > 0$. We call these consumers *naive* (index N). We assume that $\pi_\theta \geq g(\pi_\theta)$, with $\pi_\theta = g(\pi_\theta)$ for $\pi_\theta = 1$. This is justified by our focus on "high" values of π_θ (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 $u(\circ)$ (they are risk-averse).

In each period τ, 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 θ, buying a refundable tariff at price p_R, is:

$$U^R_{\theta S} = \pi_\theta u(v_\theta - p_R). \tag{2}$$

For naive consumers, the corresponding expected utility is:

$$U^R_{\theta N} = g(\pi_\theta) u(v_\theta - p_R). \tag{3}$$

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

$$U^{NR}_{\theta S} = \pi_\theta u(v_\theta - p_{NR}) + (1 - \pi_\theta) u(-p_{NR}); \tag{4}$$

finally, for naive consumers we have:

$$U^{NR}_{\theta N} = g(\pi_\theta) u(v_\theta - p_{NR}) + (1 - g(\pi_\theta)) u(-p_{NR}). \tag{5}$$

We set to zero the utility of a consumer who does not buy. Thus $c_{\theta S}$, the reservation price

[5]Karle and Möller (2020) consider a setting where firms facing loss averse consumers commit to a pricing policy that includes a price level and an advance purchase discount. However, price discrimination within each period is not considered.

[6]Our model is equivalent to Escobari and Jindapon (2014) with $\beta = 1$.

for a non-refundable tariff of a sophisticated consumer with valuation θ, is the solution to:

$$\pi_\theta u(v_\theta - c_{\theta S}) + (1 - \pi_\theta)u(-c_{\theta S}) = 0. \tag{6}$$

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

$$g(\pi_\theta)u(v_\theta - c_{\theta N}) + (1 - g(\pi_\theta))u(-c_{\theta N}) = 0. \tag{7}$$

Because $\pi_\theta \geq g(\pi_\theta)$, it follows that $c_{\theta S} \geq c_{\theta N}$, 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 type than by the level of sophistication.

4.2 The optimal tariff menu: general case

A firm can offer three types 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.

For sake of readability, the analysis of the general case, and all the corresponding proofs, are reported in the Appendix (Section B).

It turns out that we can restrict our attention to seven undominated candidate equilibria. In particular, offering the non-refundable tariff only cannot be an equilibrium, since the risk-neutral firm can always increase its profits by proposing a refundable tariff to the risk-averse consumer. In addition, if only high-valuation types are served, refundable and non-refundable tariffs cannot be conveniently used to discriminate between naive and sophisticated types. It follows that heterogeneity in valuation is a necessary condition to implement naiveté-based discrimination in our setting.

4.3 The optimal tariff menu when the cancellation probability is small

In this section we ask whether there are conditions under which the cancellation premium can be "large" even when the cancellation probability is "small", as observed in the data close to the departure date.

For this purpose, we provide the characterization of the equilibrium for $\pi_L \to 1$ and $\pi_H \to \overline{\pi} < 1$. It turns out that in this case we can restrict our attention to three undominated candidate equilibria configurations only. Configuration I exhibits naiveté-based discrimination. Naive high-valuation consumers (HN) select the refundable tariff, whereas sophisticated high-valuation consumers (HS) select the non-refundable tariff. Low types (LS and LN) select the non-refundable tariff as well. Configuration II include only preference-based discrimination: high-valuation consumers (HS and HN) select the refundable tariff, whereas low-valuation consumers (LS and LN) select the non-refundable tariff. Both in Configuration I and II, the cancellation premium is strictly positive, but is higher in the former than in the latter. In Configuration III, the firm offers only the refundable tariff, which H consumers (both sophisticated and naive) purchase.

Proposition 1 expresses the conditions that make the configuration exhibiting naiveté-based discrimination the optimal one. The proof is in the Appendix (Section B).

Proposition 1 *Define m_I as the solution to $g(\overline{\pi})u(v_H - m_I) = g(\overline{\pi})u(v_H - v_L) + (1 - g(\overline{\pi}))u(-v_L)$ and m_{II} as the solution to $\overline{\pi}u(v_H - m_{II}) = \overline{\pi}u(v_H - v_L) + (1 - \overline{\pi})u(-v_L)$, with $m_I > m_{II}$. Configuration I is the optimal tariff menu if:*

$$\overline{\pi}\left[m_I(1-\beta) - m_{II}\right] + v_L\beta > 0, \tag{8}$$

$$\overline{\pi}\left[v_H - m_I(1-\beta) - v_L\beta\right] N_H < v_L N_L. \tag{9}$$

(8) holds when Configuration I guarantees higher expected profits than Configuration II. In this case, offering a "high" refundable tariff, but to naive high type consumers only, entails higher profits than offering a "low" refundable tariff to all high-valuation customers. If (9) holds, then Configuration I guarantees higher expected profits than Configuration III. In this case, the additional profits from serving low types consumers must be higher than the loss due to the rent obtained by high type consumers.

The analysis in Section D of the Appendix shows that Configuration I may continue to be the optimal one as long as capacity is limited but is "large enough" (even if lower than $N_H + NL$). Only if capacity is "significantly small", conversely, does the firm always prefer to use it to serve the high type consumers only. Section 5 shows that the risk for our hotels to be sold-out is limited.

4.4 Risk aversion *vs* belief distortion

Proposition 1 shows that, under certain parameter configurations, firms can find optimal to keep the cancellation premium "high" even for small cancellation probability, while separating consumers based on their degree of belief distortion.

However, as long as the probability of cancellation (for high-valuation types) remains strictly positive as we assumed, we can also observe equilibria that rely on standard risk aversion to generate positive cancellation premia. In this section we show that, for reasonable parameters values, risk aversion alone is not a plausible explanation of what we observe in the data. This is the case for two reasons. First, Configuration II is the optimal menu only for a small portion of the parameter space. Second, the *magnitude* of these premia as the arrival date approaches is too small. Our results are therefore consistent with well-established "calibration theorems" (Rabin, 2000).

As for the first point, Figure 5 shows the optimal tariff menu, as function of N_H and β. 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. $v_H = 300$ and $v_L = 200$. The value of k identifies the contribution that the surplus from service consumption can have on individuals' wealth. We assume $k = 10000$; with initial wealth normalized to 1, this sets an upper bound of 3% 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). Following Wu and Gonzalez (1996), we set $\gamma = 0.71$. For the case with unbiased beliefs, $\gamma = 1$. Finally, we set $N_L = 2$ and $\overline{\pi} = 0.92$ (results are robust for different combinations of these two parameters).

What Figure 5 shows is that the purely preference-based discrimination menu associated to Configuration II is optimal only for extremely large values of β, i.e., when the fraction of naive consumers is extremely small. Bruhin et al. (2010) report that the fraction of expected utility maximizers is only 20% in their data.

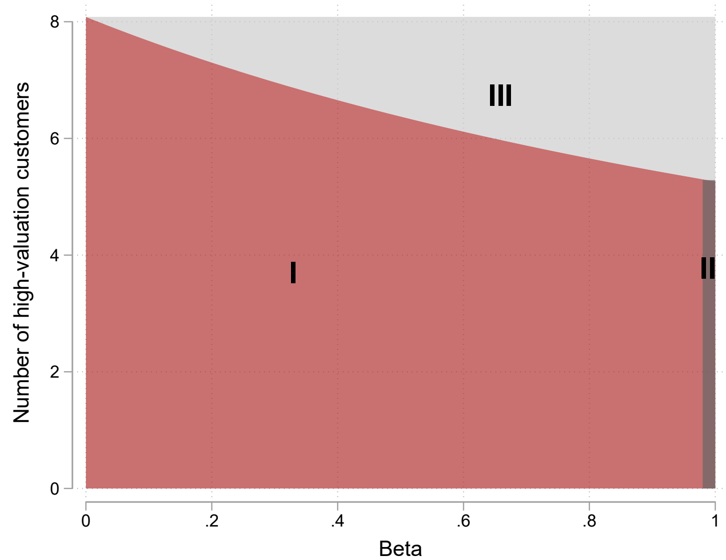
Regarding the second point, building on the estimates of cancellation probabilities provided in Falk and Vieru (2018) and Cho et al. (2018), we computed m_I and m_{II} for values of $\overline{\pi}$ from 0.92 to 0.98. [7]In this interval, the absolute risk premium ranges from 46.21 to 17.01 when beliefs are biased, but, absent probability overweighting, only from 17.73 to 4.16.

The estimated premium for French four-stars hotels in our sample is never lower than 25 Euros; it falls slightly below 20 GBP immediately before the check-in date in the UK. In addition, the ratio between the premium at $\pi = 0.92$ and $\pi = 0.98$ is 2.72 with biased beliefs and 4.26 without bias (risk aversion only), thus implying a lower convergence of refundable and non-refundable prices in the latter case.[8]

[7] Falk and Vieru (2018), using booking data from a Finnish hotel chain, report a cancellation probability of, respectively, 2.9%, 4.8% and 11.0% for bookings between 1 and 4 days, 5 and 9 days and 10 and 24 days before check-in. Cho et al. (2018) report also probabilities of cancellation as a function of days prior to arrival. These range from under 0.5% (40 days before) to a peak of 1.4% a few days before check-in.

[8] 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 τ 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 5: Parameter space for each equilibrium configuration



Notes: Region I includes the combinations of N_H and β such that naiveté-based discrimination is an equilibrium. For the parameter values in Region II, the equilibrium is one of price discrimination that is not naiveté-based. Finally, region III contains the parameter values for which, in equilibrium, the company only serves high-valuation customers. The values of the constant parameters that generate the figure are: $\pi_H = 0.92$; $\pi_L = 1$; $N_L = 2$; $k = 10000$; $v_H = 300$; $v_L = 200, \gamma = 0.71$.

4.5 Welfare implications

Suppose that, initially, the firm does not implement naiveté-based price discrimination. What would be the welfare implications of a firm's pricing strategy that considers the presence of naive consumers? The answer to this question is summarized in Lemma 1 (the proof is Section B of the Appendix).

Lemma 1 *The implementation of naiveté-based price discrimination is Pareto-improving if and only it increases market coverage, i.e., if the number of consumers served in equilibrium increases.*

The intuition behind this Lemma is as follows. In our setting, serving all customers is socially efficient because all valuations are higher than marginal cost. Then, if naiveté-based discrimination increases market coverage, social welfare increases, consistently with standard results from models of price discrimination (Schmalensee, 1981; Katz, 1983). However, what we obtain is a stronger result, in 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, so that naive consumers are not necessarily exploited, as in the case of Eliaz and Spiegler (2008) and Heidhues and Kőszegi (2010).

Intuitively, naiveté-based price discrimination leads to a partial form of screening that allows the firm to treat HN types differently from low types, while serving also the latter in equilibrium. 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 depend crucially on consumer heterogeneity.

When naiveté-based discrimination does not increase market coverage (and so does not affect total welfare), it has redistributionary consequences; the firm is better off, naive high types are worse off, sophisticated high types are better off.

5 Alternative explanations

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

Decoy pricing Hotels, for one, may set prices for reasons other than conveying all the available information. Easily observable online prices for refundable reservations, for example, may be kept artificially high to establish a reference point or a "decoy" for online consumers, to make the more convenient non-refundable price look more enticing. This would imply that hotel managers would think it highly unlikely that any transaction could be finalized under the higher refundable price. Therefore, whenever refundable prices are not posted, we should observe the non-refundable price to remain lower than the refundable one that was offered in nearby dates. In our samples, 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 6 show the distribution of the following statistic:

$$\Delta P = \left[P_{NR}(d) - P_R(d+1) \; | P_R(d) \; \textit{not observed}\right]. \tag{10}$$

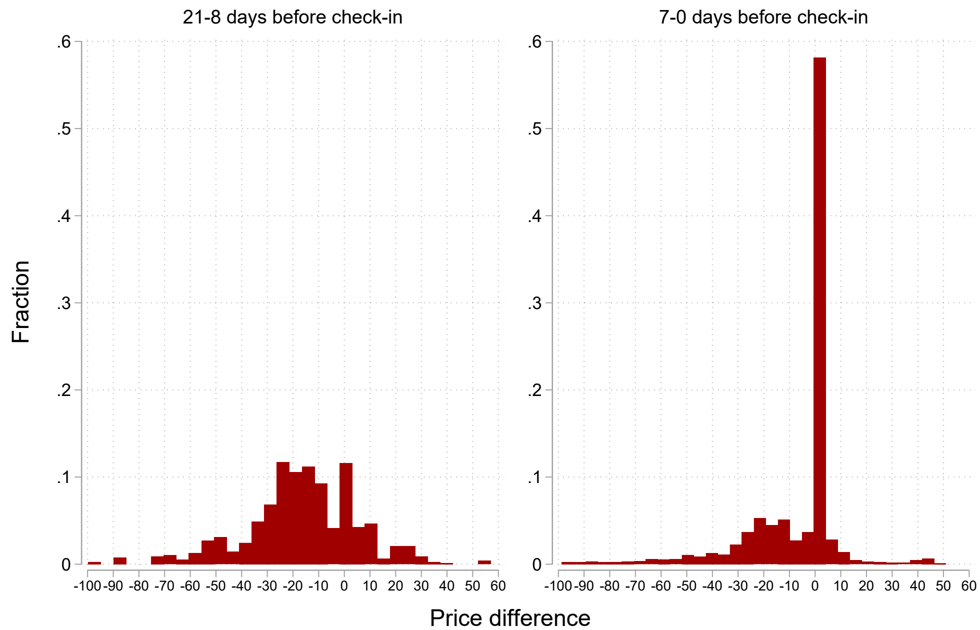
ΔP is the difference of the non-refundable price P_{NR} posted d days before the stay, and the refundable price P_R posted the day before, conditional on the latter not being available on d. Both in France and in the USA, about 60% of the distribution mass of ΔP is at zero during the last week before arrival (right panel), i.e., P_{NR} converges towards the refundable level when the latter is not offered any longer. Earlier on, however, $\Delta P < 0$, because the cancellation option is still active and the company momentarily omitted to post the refundable option.

Figure 6 therefore highlights two possible implications. First, the refundable price does not operate as a "decoy", but it is instead a price at which the hotel expects to transact. This is particularly relevant for our analysis based on posted prices. Second, the convergence between the two prices, that in Escobari and Jindapon (2014) occurs just before the cancellation option expires, in our sample may also occur, but only after. Albeit very occasionally, we even have

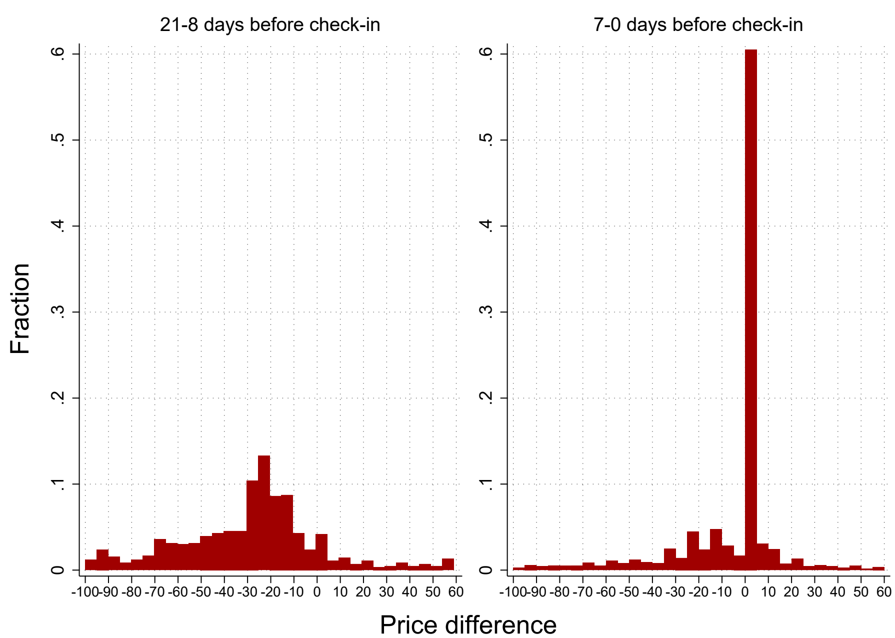
$\Delta P > 0$, i.e., the non-refundable price may be set even above the past refundable values.

Figure 6: Distribution of the difference between the non-refundable price on the posting day and the one-day lagged refundable price, when the refundable price is not available on the posting date.

(a) France



(b) USA



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

Systematic mispricing DellaVigna and Gentzkow (2019), in their analysis of price rigidity within store chains, conclude that their evidence is consistent with inertial behavior of managers. May this also be the case in our context? At least in part, by absorbing hotel-level fixed effects, our econometric specification accounts for time-invariant, firm-level differences in managerial inertia. Moreover, institutional features of the hotel industry, as well as evidence from our own data, do not support this hypothesis. First, the evidence in Figure 6 suggests that prices are not fixed at a level that is patently too high. Second, 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, discusses how information on revenue management strategies shared by chain hotels may become a source of competitive advantage; as such, it is unlikely that the pricing strategies of the hotels in our sample are systematically ill-designed. 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).

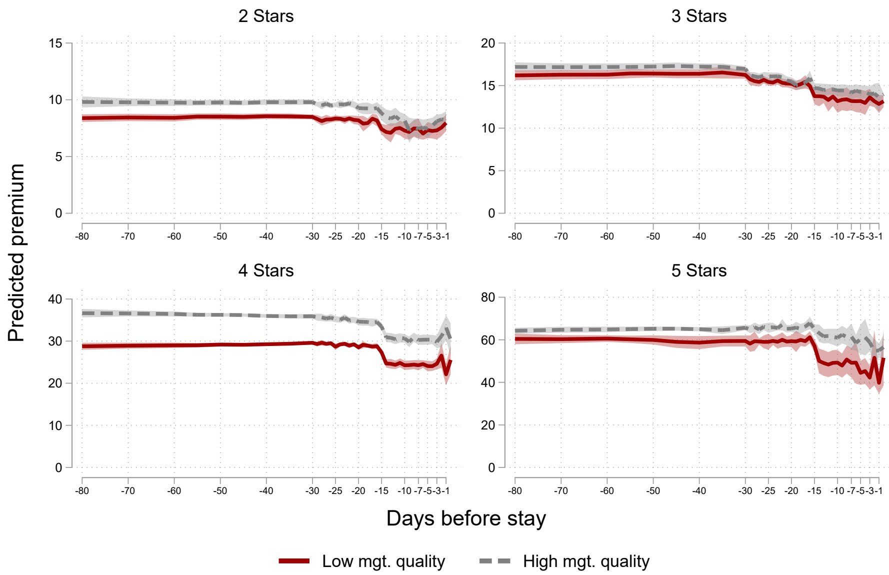
If there were systematic mispricing in our sample, we might expect it to occur especially among less sophisticated or more badly managed hotels, because features such as market experience or high skills may correct certain behavioral tendencies (List, 2003; Goldfarb and Xiao, 2011; Anagol et al., 2018). 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 (i.e., Comfort, Cleanliness, Staff quality, Facilities), and high management quality otherwise.

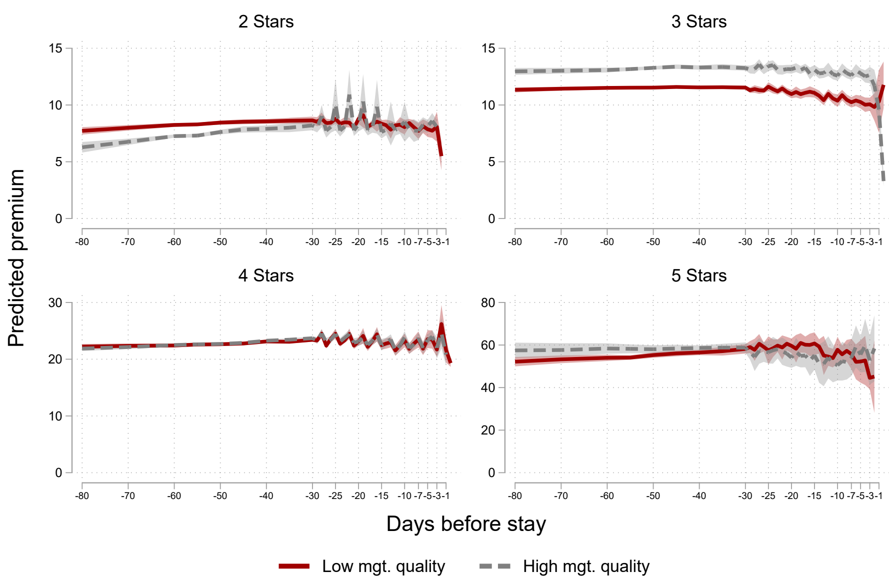
After estimating the model in equation (1) separately for hotels with high and low-rated management quality, we show the predicted premia in Figure 7 for France and the UK, and the premia in percentage terms in Figure 8 for the USA and Canada. The predicted values are generally smaller for hotels with lower management quality scores, but the patterns are similar: premia are stable and stay well above zero, for all categories of hotels and countries.

Figure 7: Estimated cancellation premia by days before check-in, stars, and management quality

(a) France



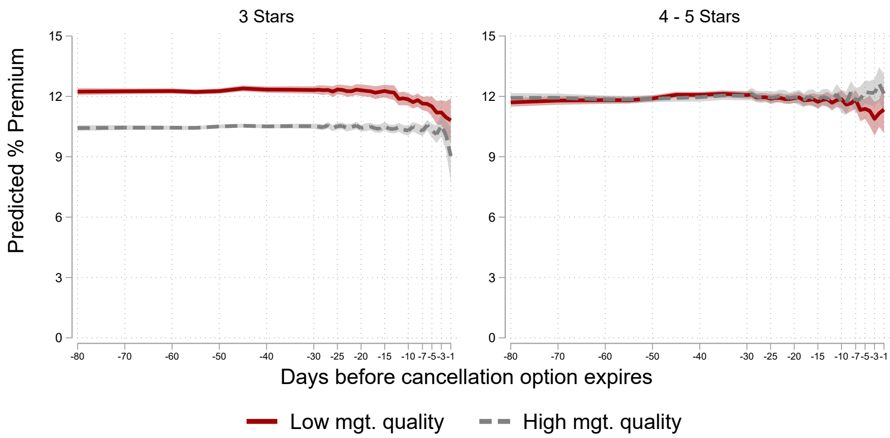
(b) UK



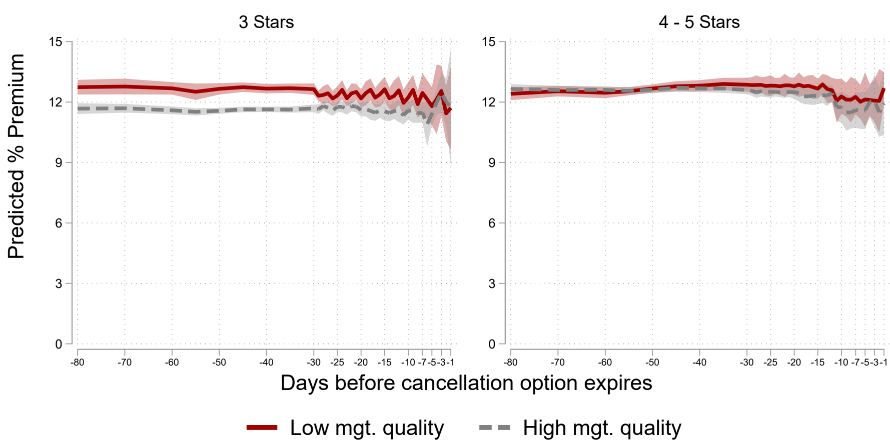
Notes: The graphs display the estimated values of the cancellation premium, from model (1). The regressions are separate by stars and country. The shaded areas indicate 95% confidence intervals, with standard errors clustered at the hotel level. Fixed effects combine hotel, check-in date, room type, 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 8: Estimated cancellation premia by days before cancellation option expires, stars, and management quality

(a) USA



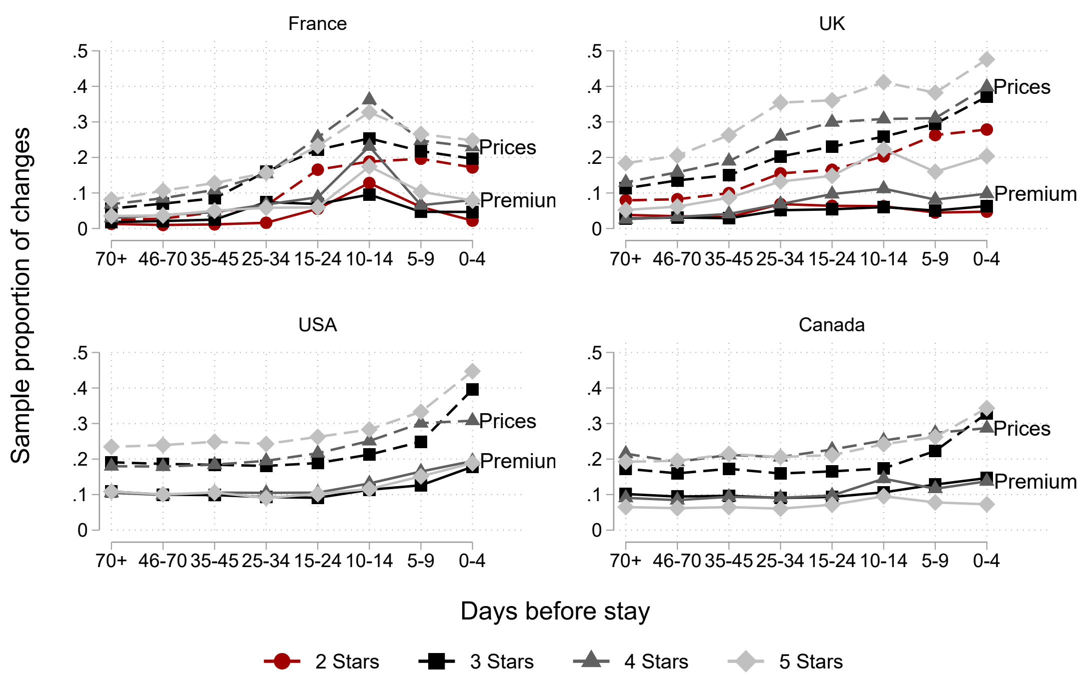
(b) Canada



Notes: The graphs display the estimated values of the percentage 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 3 and 4-5 stars hotels are, respectively, $1,229,177$ $(1,187,446)$; $1,680,358$ (1,686,150) in USA, and $462,460$ $(366,442)$; $462,776$ $(442,242)$ in Canada. R^2 values range from $0.70-0.80$ in USA, and $0.70-0.77$ in Canada.

Relatedly, managerial inertia or inability may lead to the stability of prices and, in turn, to the stability of premia. However, Figure 9 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, for both prices (dotted lines) and premia (solid lines), by stars. Changes in prices are much more frequent than changes in premia; for instance, only about 40% of four-star USA hotels change their prices within four days before check-in, but only 20% change the associated premium concurrently. That is, when hotels adjust both the refundable and non-refundable prices, they most often keep the premium unchanged. This indicates that the stability of premia in all countries may be the result of a deliberate strategy by hotel managers.

Figure 9: Frequency of price and premium changes



Notes: The graphs report the share of hotel rooms that had their prices and/or premia modified, by days before a given check-in date. We grouped days in periods of decreasing length as the check-in date approaches.

Location and time-specific effects The stability of refund premia that we document in the full datasets may derive only from certain cities, for example because they are more likely to be destinations for business travelers than others, and, in general, different location may host customers with different characteristics, including the valuation of a room and the estimates of the probability of cancellation. These difference and the makeup of travelers' features may also be seasonal. The regression estimates and predicted values that we have reported so far are from models that control also for the combination of city and check-in dates. However, the periods that the datasets cover are limited, and do not allow to fully account for time or seasonal effects. Moreover, because in our model the presence of both refundable and non-refundable options even in proximity of the arrival date depends on both individual valuation and attitudes toward losses, it is hard to attribute a specific combination of these two aspects to categories of travelers that we can identify, for example, by their destinations or days of travels. Note also that, in the model, the degree of distortion in the perception of the probability of cancellation is orthogonal to valuation.

Our data, however, allow for at least some additional partial investigations of time and location effects. Appendix Figures A.1 and A.2 display the predicted premia (in percentage value) by time before the check-in date for specific cities and limited to hotels affiliated with chains. Panels b and d of each figure report estimates further restricted to observations where the check-in date occurs on a weekend (nights of Saturday and Sunday); the clientele during business days and weekend days may again differ, for example leisure travelers are more likely to look for hotel rooms also during weekends. Although the average percentage difference between non-refundable and refundable reservations options do vary between cities and days in the week (see also Appendix Table A.4), in no case do we observe convergence of the premium toward zero.[9]

Supply-side effects Finally, supply-side factors may keep the refundable and non-refundable prices apart. From the hotel's perspective, the cost of a cancellation is primarily determined by the opportunity cost of not reselling the room, which is more likely if the cancellation occurs close to the check-in date. Again, controlling for hotel-level fixed effects addresses, at least in part, differences in these opportunity costs. The following considerations further lead us to rule out sizable supply side effects. First, the available evidence for cancellation probabilities, although limited, indicates that they are low, especially close to check-in (Falk and Vieru, 2018; Cho et al., 2018). Second, and more importantly, a late cancellation generates an opportunity cost only as long as, prior to the cancellation, the hotel must refuse a reservation request because it is sold-out, and, due to cancellations, it is subsequently left with vacant rooms on the check-in date. The analysis of the booking trajectories of the US hotels in Cho et al. (2018) indicate that, even during the busiest weekends, the hotels do not reach full capacity utilization.

Although we do not have access to direct information about capacity, and thus cannot directly test whether the lack of premia convergence toward zero is due to a higher load factor, we take two approaches to identify hotels that are more likely to have reached higher than usual occupancy rates. First, we follow Piga and Melis (2021) and assume that hotels stop posting their prices on *Booking.com* because they are sold-out. In the North American sample, we can observe whether, for a given room variant (with or without breakfast, varying number of persons, etc.), the hotel posts data until the exact date of the cancellation option. In such a case, we can hypothesize that the hotel had rooms available, the opportunity cost of offering the cancellation option is negligible and the incentive to move the premium toward

[9]The fact that percent premia do not coincide across cities and days of the week further suggest that these pricing strategies do not follow simple rule of thumbs. The statistics in Table A.4 in the Appendix, moreover, distinguish between hotels that are part of chains and hotels that are not, in the sample of North American establishments. The differences between the two groups are very small.

zero strong.[10] Table 2 indicates that the percentage of room variants that are always observed ranges from about 60% to 80% in the North American cities in our sample.

Table 2: Occupancy proxies

Room variants:	Always observed^a	Drop room variants^b	Drop available room^c	Both drops^d
Country and city:				
USA				
Atlanta	0.66	0.47	0.26	0.58
Boston	0.70	0.56	0.33	0.68
Chicago	0.73	0.48	0.30	0.60
Houston	0.70	0.48	0.21	0.57
Los Angeles	0.70	0.58	0.42	0.74
Minneapolis	0.78	0.48	0.30	0.61
Miami	0.62	0.59	0.37	0.71
New York	0.61	0.67	0.51	0.82
Portland	0.71	0.51	0.31	0.64
Seattle	0.67	0.59	0.4	0.73
CANADA				
Calgary	0.81	0.31	0.22	0.45
Montreal	0.74	0.50	0.39	0.66
Toronto	0.71	0.51	0.34	0.66
Vancouver	0.68	0.52	0.43	0.68
Winnipeg	0.76	0.46	0.29	0.60

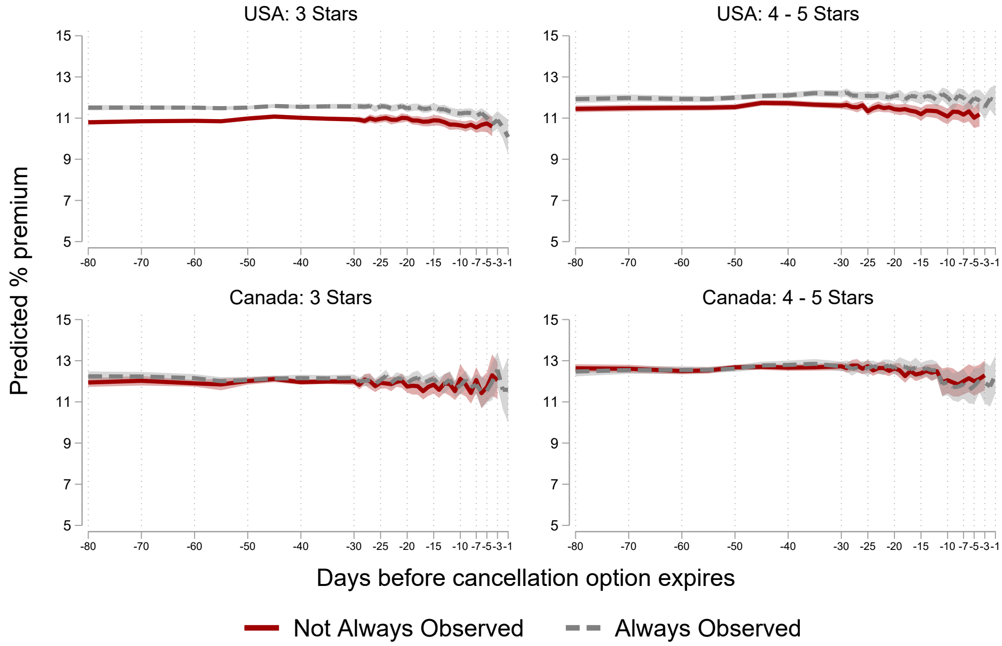
Notes: a) Room variants that are "Always Observed" until the option expiry; b) rooms whose availability drops by at least 30% between the first and the last observation, and c) rooms in hotels whose number of offered room types drops by at least 15%. All obtained holding the check-in date fixed.

Panel a of Figure 10 shows the estimated values of the percentage premia obtained from two complementary clusters of room variants that are "Always Observed" until the option expiry, or not. For both countries and all star-ratings, premia remain between 11-12%, without any noticeable difference between the clusters.

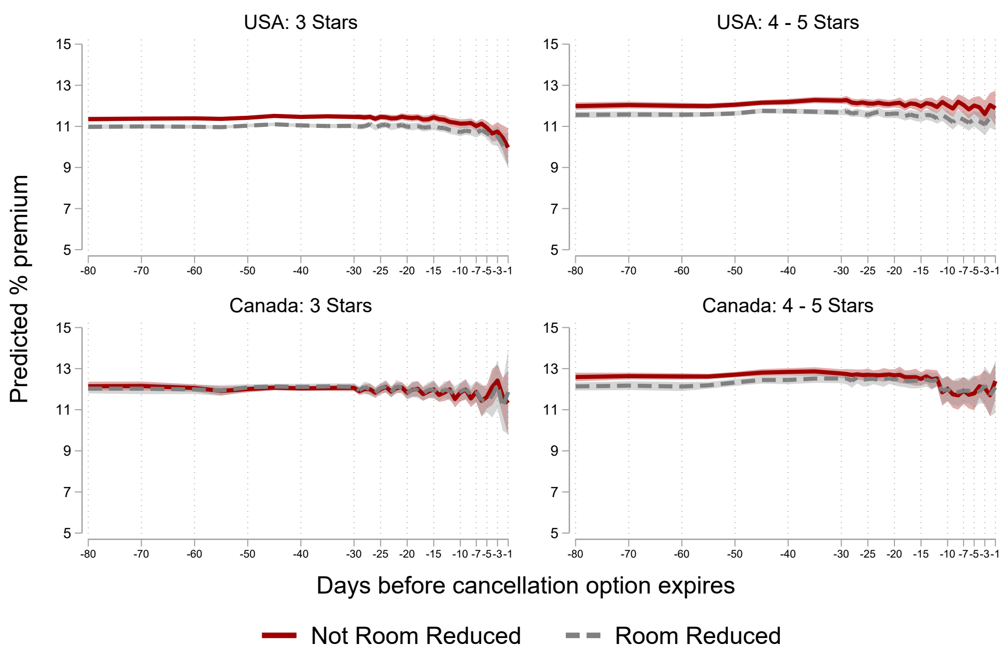
[10] If a room is cancelled after the cancellation but before the stay date, the hotel can charge at least the price for one night and resell the room. This mechanism further reduces the opportunity cost of offering the cancellation option.

Figure 10: Controlling for hotel occupancy

(a) Occupancy proxy: Observed until cancellation option expires



(b) Occupancy proxy: Room reduction



Notes: The graphs display the estimated values of the percentage cancellation premium for hotel rooms in the US and Canada. The regressions are separate by country and by different proxies for occupancy rates: a) room variants that are "Always Observed" until the option expiry, and those that are not; b) rooms whose availability drops by at least 30% between the first and the last observation, and room types drops by at least 15%, both holding the check-in date fixed;

Second, we construct and combine two possible indicators based on rooms availability to control for potentially high occupancy. One is based on the idea that a higher load factor should be reflected by fewer rooms available for each variant. This information is drawn from the hotel's page (see Section 2.1). A first indicator thus flags all the observations of a given variant if the number of available rooms falls by at least 30% between the first and the last observation, holding the check-in date fixed. Table 2 above shows that the cities with the largest proportion of observed drops are Vancouver in Canada (43%) and New York in the USA (51%). The second measure derives from the prediction that the number of listed room types should decline, as some sell out and thus disappear from the platform. The second proxy that we use is an indicator for whether the hotel's number of room types reduces by at least 15%, holding the check-in date fixed. Between 31% (Calgary) and 67% of room offers are observed in hotels where the indicator is active. In Panel b of Figure 10 at least one of the two indicators is active in the "Room Reduced" cluster, as summarized in the fourth column of Table 2. The combined indicators thus provide a proxy for the hotels with the higher opportunity cost of the cancellation option. However, no significant difference is found across countries and star-rating categories, as well as no declining trend in the "Not Room Reduced" cluster.

6 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 overestimate the likelihood of having to cancel. Reducing the response of prices to information may actually represent an intentional, rather than inertial and suboptimal, managerial conduct that leverage demand heterogeneity in valuation and propensity to cognitive biases of customers. 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.

A direction for future research in the context of the hotel industry would be to extend observations to longer periods within a year, to assess potential seasonal differences in these pricing practices. In this study, moreover, we relied on posted prices and made inferences about capacity constraints from the data at our disposal. Although this allowed us to provide reliable insights on the stability of refundability premia, integrating the analysis with available information on actual transactions and hotel capacity would expand the range of questions that one could address. Because this information is generally not available at the level of the individual hotel for a large number of establishments, a complementary approach would be to focus on smaller samples and collect more detailed data analyses or case studies.

Finally, there are several other 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 future research.

References

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

Alderighi, M., A. A. Gaggero, and C. A. Piga (2022). Hidden prices with fixed inventory: Evidence from the airline industry. *Transportation Research Part B: Methodological 157*, 42–61.

Alderighi, M., M. Nicolini, and C. A. Piga (2015). Combined effects of load factors and booking time on fares: Insights from the yield management of a low-cost airline. *Review of Economics and Statistics 97*, 900–915.

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

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

Bachis, E. and C. A. Piga (2011). Low-Cost Airlines and online Price Dispersion. *International Journal of Industrial Organization 26(6)*, 655–667.

Barnett, W. and E. Pontikes (2008). The red queen, success bias, and organizational inertia. *Management Science 54*(7), 1237–1251.

Bruhin, A., H. Fehr-Duda, and T. Epper (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.

Cavallo, A. (2018). Scraped Data and Sticky Prices. *Review of Economics and Statistics 100(1)*, 105–109.

Cavallo, A., E. Diewert, R. C. Feenstra, R. Inklaar, and M. P. Timmer (2018). Using Online Prices for Measuring Real Consumption Across Countries. *AEA Papers and Proceedings 108*, 483–487.

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

Cho, S., G. Lee, J. Rust, and M. Yu (2018). Optimal Dynamic Hotel Pricing. mimeo, available at https://cowles.yale.edu/3a/hpabridged-optimal-dynamic-hotel-pricing.pdf.

Choi, S. and A. S. Mattila (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 M. Gentzkow (2019, 06). Uniform Pricing in U.S. Retail Chains. *The Quarterly Journal of Economics 134*(4), 2011–2084.

Dutta, S., M. Zbaracki, and M. Bergen (2003). Pricing process as a capability: a Resource-Based perspective. *Strategic Management Journal 24*, 615–630.

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

Ellison, G. (2005). A model of add-on pricing. *Quarterly Journal of Economics 120(2)*, 585–637.

Escobari, D. (2012, December). Dynamic pricing, advance sales, and aggregate demand learning in airlines. *Journal of Industrial Economics 60*(4), 697–724.

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

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

Gabaix, X. and D. Laibson (2006, 05). Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets. *The Quarterly Journal of Economics 121*(2), 505–540.

Gallego, G. and O. Şahin (2010). Revenue management with partially refundable fares. *Operations Research 58*(4), 818–833.

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

Gorodnichenko, Y., S. Sheremirov, and O. Talavera (2018). Price Setting in Online Markets: does IT Click? *Journal of the European Economic Association 16*, 1764–1811.

Gorodnichenko, Y. and O. Talavera (2017). Price setting in online markets: Basic facts, international comparison, and cross-border integration. *American Economic Review 107(1)*, 249–282.

Gorodnichenko, Y., O. Talavera, and N. Vu (2021). Quality and Price Setting of High-Tech Goods. *Economic Modelling 98*, 69–85.

Hannan, M. and J. Freeman (1977). Structural inertia and organizational change. *American Sociological Review 82*(5), 929–964.

Hannan, M. and J. Freeman (1984). The population ecology of organizations. *American Journal of Sociology 49*, 149–164.

Hartley, R., G. Lanot, and I. Walker (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 B. Kőszegi (2008). Competition and price variation when consumers are loss averse. *American Economic Review 98(4)*, 1245–1268.

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

Heidhues, P. and B. Kőszegi (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.

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

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

Karle, H. and M. Möller (2020). Selling in advance to loss averse consumers. *International Economic Review 61(1)*, 441–468.

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

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

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

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

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

Levinthal, D. (1997). Adaptation on rugged landscapes. *Management Science 43*(7), 934–950.

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

Mantovani, A., C. A. Piga, and C. Reggiani (2021). Online platform price parity clauses: Evidence from the EU Booking.com case. *European Economic Review 131*, 103625.

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

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

Nelson, R. R. and S. G. Winter (1982). *An evolutionary theory of economic change*. Princeton University Press.

Nicolini, M., C. A. Piga, and A. Pozzi (2023). From Uniform to Bespoke Prices: Hotel Pricing during EURO 2016. *Quantitative Marketing and Economics 21*, 333–355.

Piga, C. A. and G. Melis (2021). Identifying and measuring the impact of cultural events on hotels' performance. *International Journal of Contemporary Hospitality Management 33*(4), 1194–1209.

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

Ritson, M., M. Bergen, S. Dutta, D. Levy, and M. J. Zbaracki (2003). Shattering the Myth of Costless Price Changes. *European Management Journal 21(6)*, 663–669.

Sandroni, A. and F. Squintani (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.

Stole, L. A. (1995). Nonlinear pricing and oligopoly. *Journal of Economics and Management Strategy 4(4)*, 529–562.

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

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

Valletti, T. (1999). A model of competition in mobile communications. *Information Economics and Policy 11(1)*, 61–72.

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

Williams, K. R. (2022). The Welfare Effects of Dynamic Pricing: Evidence From Airline Markets. *Econometrica 90(2)*, 831–858.

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

Zbaracki, M., M. Ritson, D. Levy, S. Dutta, and M. Bergen (2004). Managerial and customer costs of price adjustment: Direct evidence from industrial markets. *Review of Economics and Statistics 86(2)*, 514–533.

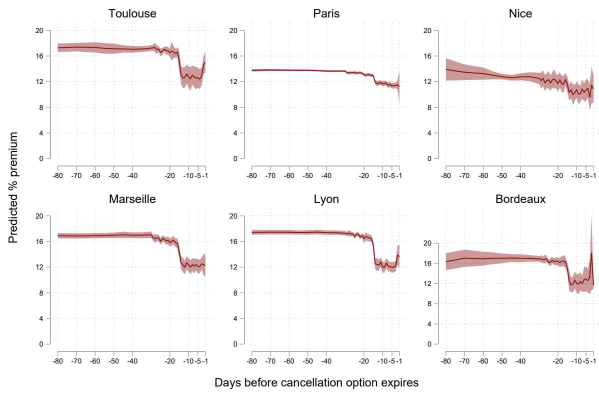
Sticky price for declining risk? Business strategies with "behavioral" customers in the hotel industry

Appendix

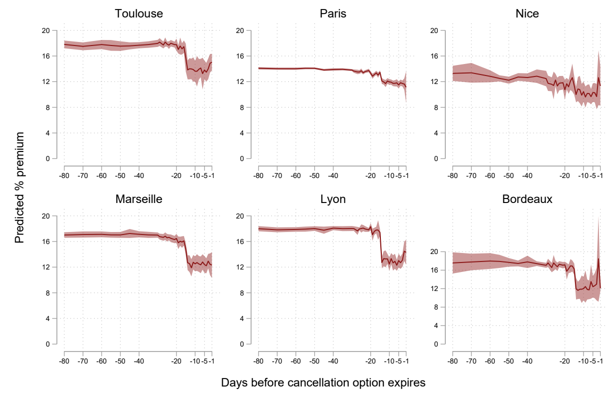
A Sample Statistics

Figure A.1: Predicted Percentage Cancellation Premia, by city and country: Europe

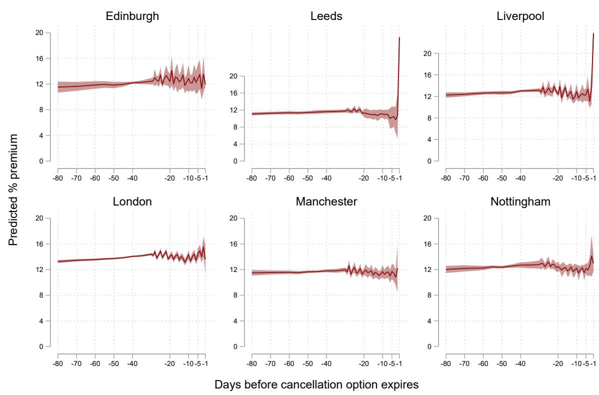
(a) France



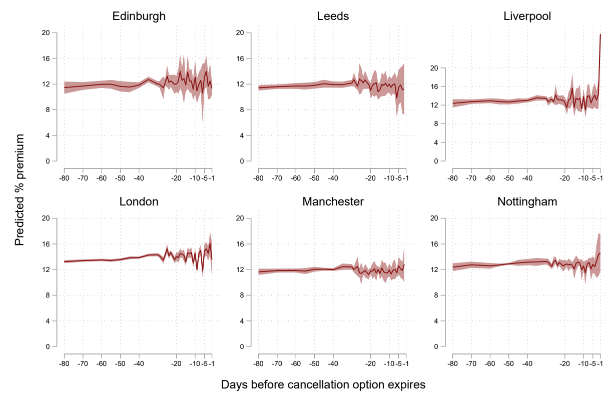
(b) France - Weekends



(c) UK



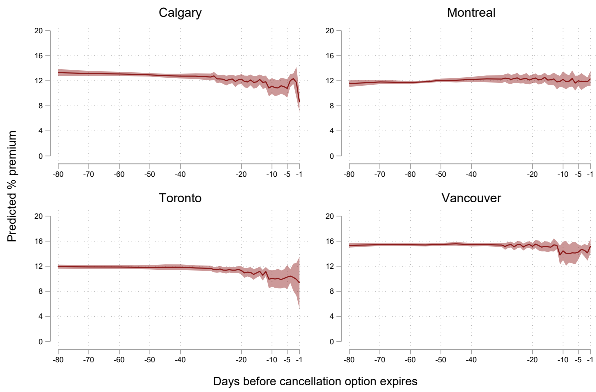
(d) UK - Weekends



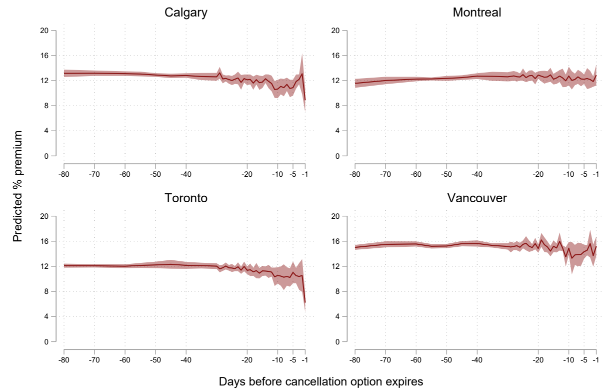
Notes: Four and Five-star hotels combined. All hotels affiliated with a chain.

Figure A.2: Predicted Percentage Cancellation Premia, by city and country: North America

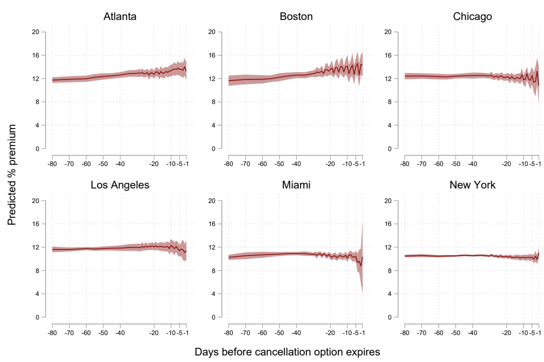
(a) Canada



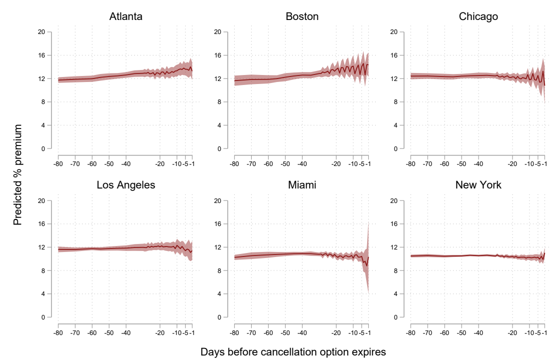
(b) Canada - Weekend



(c) USA



(d) USA - Weekend



Notes: Four and Five star hotels combined. All hotels affiliated with a chain.

Table A.1: Number of observations by combination of prices retrieved before the stay date

Country:	FRANCE				UK			
Price:	NR and R	Only R	Only NR	All	NR and R	Only R	Only NR	All
Days from stay:								
0-2	18,093	51,800	6,805	76,698	11,759	43,288	46,959	102,006
3-4	66,515	23,414	5,006	94,935	44,936	32,838	3,814	81,588
5-6	78,395	18,946	3,978	101,319	57,475	26,594	1,965	86,034
7-9	119,142	21,882	7,096	148,120	93,521	26,718	3,004	123,243
10-13	144,860	24,175	8,782	177,817	117,757	31,329	3,512	152,598
14-20	262,027	33,525	12,578	308,130	199,371	44,305	4,537	248,213
21-29	312,682	30,645	23,948	367,275	253,968	40,598	4,792	299,358
30-39	300,520	24,475	52,194	377,189	304,843	40,245	5,354	350,442
40-49	253,915	21,834	73,913	349,662	300,670	36,197	5,015	341,882
50-59	261,833	22,138	111,183	395,154	346,296	38,789	5,358	390,443
60+	589,455	61,127	513,260	1,163,842	1,009,415	128,931	17,907	1,156,253
Total	2,407,437	333,961	818,743	3,560,141	2,740,011	489,832	102,217	3,332,060

Country:	CANADA				USA			
Price:	NR and R	Only R	Only NR	All	NR and R	Only R	Only NR	All
Days from stay:								
0-2	8,971	29,450	44,529	82,950	8,894	67,804	129,459	206,157
3-4	28,279	50,688	4,291	83,258	77,404	170,598	14,510	262,512
5-6	31,711	47,852	3,094	82,657	113,793	154,126	6,724	274,643
7-9	84,129	73,450	5,008	162,587	270,510	261,886	11,224	543,620
10-13	116,807	101,622	6,440	224,869	392,225	363,713	14,310	770,248
14-20	204,568	158,541	10,422	373,531	695,410	555,439	22,209	1,273,058
21-29	269,265	203,056	14,025	486,346	889,878	687,479	27,158	1,604,515
30-39	285,084	205,746	15,471	506,301	900,157	685,149	26,671	1,611,977
40-49	250,942	178,212	13,807	442,961	773,777	587,711	23,125	1,384,613
50-59	226,444	160,609	12,850	399,903	706,031	527,535	21,026	1,254,592
60+	1,062,271	760,036	62,183	1,884,490	3,518,388	2,516,279	107,167	6,141,834
Total	2,568,471	1,969,262	192,120	4,729,853	8,346,467	6,577,719	403,583	15,327,769

Table A.2: Mean price of rooms with cancellation, by stars and country

Country:	FRANCE			UK		
Stars:	3	4	5	3	4	5
Days from stay:						
0-2	96.5	186.4	485.3	87.2	154.9	425.3
3-4	104.5	205.7	524.1	90.4	169.0	453.1
5-6	109.9	212.3	530.1	91.7	172.7	458.3
7-9	111.9	212.3	530.4	92.8	173.8	458.1
10-13	111.0	212.9	534.4	93.3	174.2	451.9
14-20	112.1	214.3	519.0	95.9	181.1	452.4
21-29	112.5	214.8	520.5	97.5	184.3	450.5
30-39	113.0	214.7	523.3	96.9	181.9	448.8
40-49	114.9	217.4	532.4	98.0	183.9	452.3
50-59	115.8	223.2	541.6	98.6	185.3	454.0
60+	119.4	234.8	557.3	97.3	178.1	445.4
Country:	CANADA			USA		
Stars:	3	4	5	3	4	5
Days from stay:						
0-2	174.1	257.7	565.1	158.8	284.0	751.5
3-4	182.0	299.0	529.9	185.0	315.2	731.6
5-6	186.2	309.5	534.9	195.8	328.8	742.0
7-9	184.6	301.4	529.3	195.5	323.4	754.8
10-13	184.5	301.5	529.1	197.8	325.2	774.8
14-20	188.0	307.0	530.6	203.3	333.4	789.8
21-29	189.7	307.2	533.2	205.1	331.9	781.4
30-39	192.1	312.4	540.6	212.0	340.4	790.8
40-49	190.8	309.9	535.9	211.6	335.4	786.6
50-59	190.0	308.1	536.0	214.5	336.9	783.9
60+	196.1	315.2	554.6	217.5	333.1	762.6

Notes: Two-star hotels in France and UK not included to save space, but available on request. Prices expressed in local currency.

Table A.3: Mean percentage cancellation premia by stars, days of the week and country

Country:	FRANCE						UK					
Stars:	3		4		5		3		4		5	
Weekend:	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Days from stay:												
0-2	14.0	15.0	11.5	11.4	10.7	11.3	12.4	12.4	15.6	16.5	12.6	11.9
3-4	13.1	14.2	13.1	13.3	11.1	11.6	12.6	12.6	13.8	14.4	12.2	11.8
5-6	12.9	13.2	13.2	13.4	11.2	11.5	12.3	12.3	13.4	13.8	12.1	11.7
7-9	12.5	13.4	13.0	13.4	11.8	12.2	12.1	12.9	13.2	13.6	12.6	12.0
10-13	12.8	13.4	13.2	13.4	11.9	12.0	12.2	12.7	13.0	13.6	12.5	12.4
14-20	13.8	14.4	14.9	15.6	13.3	13.4	12.3	12.6	12.9	13.2	12.8	12.8
21-29	14.5	15.2	15.5	16.5	13.3	13.7	12.5	12.8	13.1	13.2	12.9	13.1
30-39	15.6	16.6	15.5	16.6	13.2	13.6	12.6	13.2	12.9	13.5	12.8	13.1
40-49	15.6	16.1	15.6	16.2	13.0	13.4	12.5	12.8	12.7	12.9	12.7	12.5
50-59	15.0	16.0	15.2	16.2	12.9	13.3	12.4	12.9	12.4	12.6	12.4	12.2
60 plus	14.3	14.7	14.8	15.2	12.6	13.1	12.5	12.8	12.4	12.4	12.5	12.1
Country:	CANADA						USA					
Stars:	3		4		5		3		4		5	
Weekend:	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Days from stay:												
0-2	12.9	13.4	13.2	14.7	15.9	16.0	13.0	14.1	12.9	12.8	15.0	16.7
3-4	12.6	13.2	11.7	12.2	14.4	14.0	12.5	12.8	12.3	13.3	13.8	13.9
5-6	12.2	12.8	11.1	11.9	14.3	14.0	12.0	11.9	11.6	12.4	13.8	13.0
7-9	11.8	11.9	11.4	11.6	14.2	14.1	11.2	11.4	11.8	12.5	13.1	13.0
10-13	11.9	12.2	11.5	11.6	13.9	13.8	11.1	11.3	11.6	12.5	13.4	13.6
14-20	12.0	12.3	12.0	12.3	13.8	13.9	11.2	11.5	11.2	12.0	13.6	13.7
21-29	11.9	12.1	12.0	12.4	13.9	14.1	11.2	11.4	11.3	11.9	13.7	13.9
30-39	12.0	12.2	12.1	12.4	14.1	13.9	11.2	11.6	11.4	12.2	13.7	14.2
40-49	11.9	12.3	12.1	12.6	14.0	14.2	11.3	11.6	11.4	12.0	13.6	14.1
50-59	11.8	12.1	12.1	12.3	14.0	13.9	11.2	11.5	11.2	11.9	13.5	13.9
60 plus	11.8	12.2	12.0	12.0	13.9	14.0	11.0	11.3	10.9	11.6	13.9	14.2

Notes: Two-star hotels in France and UK not included to save space, but available on request.

Table A.4: Mean percentage cancellation premia by stars, days of the week and city

Country:	FRANCE		**Country:**	CANADA			
Chain:	Yes		**Chain:**	No		Yes	
Weekend:	No	Yes	**Weekend:**	No	Yes	No	Yes
City:			**City:**				
Toulouse	16.4	17.6	Calgary	11.9	12.0	11.9	12.1
St.Etienne	16.6	17.2	Montreal	11.6	11.5	12.0	12.3
Paris	13.0	13.4	Toronto	12.2	12.2	11.7	12.1
Nice	12.6	12.3	Vancouver	14.7	14.7	15.1	15.1
Marseille	15.7	16.0	Winnipeg	13.0	13.9	11.2	11.7
Lyon	15.6	16.5					
Lille	16.9	17.6					
Lens	16.0	15.4					
Bordeaux	15.5	16.0					

Country:	UK		**Country:**	US			
Chain:	Yes		**Chain:**	No		Yes	
Weekend:	No	Yes	**Weekend:**	No	Yes	No	Yes
City:			**City:**				
Birmingham	11.2	11.7	Atlanta	13.9	13.5	11.4	11.6
Cambridge	11.9	12.3	Boston	10.7	10.7	11.5	11.8
Cardiff	12.6	12.6	Chicago	13.2	13.1	12.3	12.5
Edinburgh	12.3	12.4	Houston	12.4	12.2	12.3	12.8
Leeds	12.0	12.5	Los Angeles	12.4	12.4	11.5	12.5
Liverpool	12.9	13.2	Minneapolis	10.7	10.7	12.0	12.4
London	13.6	13.7	Miami	13.6	13.9	10.5	10.4
Manchester	11.9	12.0	New York	12.1	12.2	10.2	11.3
Nottingham	12.1	12.6	Portland	12.6	12.3	12.6	13.4
Oxford	12.6	12.3	Seattle	13.0	13.1	10.4	11.1
Sheffield	12.7	12.9					

Notes: Hotels in France and UK are all chain affiliated. Statistics for some UK cities not included to save space, but available on request.

B Solving the model

In this section, we report all the results and the corresponding proofs that leads to Proposition 1.

B.1 The optimal tariff menu: the general case

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

Lemma B.1 *Suppose that PC_ω is satisfied for a type ω selecting p_{NR}. Then $IC_{\omega'}$ implies $PC_{\omega'}$ for all types ω 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 ω' selects p_R, then $U_{\omega'}^R \geq U_{\omega'}^{NR} \geq 0$.* ■

Lemma B.2 *Suppose that type θS ($\theta = L, H$) selects p_R. Then type θN selects p_R as well, as long as $\pi_\theta < 1$.*

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:*

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

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

Because $\pi_\theta > g(\pi_\theta)$, the two inequalities are incompatible. ■

Lemma B.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*

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

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

Due to the 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 $g(\pi_\theta)$ to π_θ. ■

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

In the first group of equilibrium configurations, where each tariff is chosen by at least one type, Lemmas B.1-B.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 also observe that configurations thatdiscriminate between naive and sophisticated consumers of type θ require $\pi_\theta < 1$.

We summarize the candidate equilibrium tariffs in the following propositions:

Proposition B.1 *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 $v_L(\pi_H N_H + \pi_L N_L(1-\beta)) + c_{LS}\beta N_L$.*

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

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

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

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

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

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

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

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

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

PC_{LN} implies PC_{HS}, PC_{HN} and PC_{LS} as from Lemma B.1, whereas IC_{HS} implies IC_{HN} following B.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

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

which always holds because $c_{LS} \leq v_L$. ■

Proposition B.2 *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 $\pi_H u(v_H - m_2) = \pi_H u(v_H - c_{LN}) + (1-\pi_H)u(-c_{LN})$. The firm's expected profit is $m_2\pi_H N_H + c_{LN}N_L$.*

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

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

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

$$\pi_H u(v_H - v_L) \geq \pi_H u(v_H - c_{LN}) + (1 - \pi_H)u(-c_{LN}),$$

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

Proposition B.3 *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 $\pi_H u(v_H - m_3) = \pi_H u(v_H - c_{LS}) + (1 - g(\pi_H))u(-c_{LS})$. The firm's expected profit is $m_3\pi_H N_H + c_{LS}\beta N_L$.*

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

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

PC_{LS} implies PC_{HS} and PC_{HN} (Lemma B.1), and IC_{HS} implies IC_{HN} as per Lemma B.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:

$$\pi_H u(v_H - v_L) \geq \pi_H u(v_H - c_{LN}) + (1 - \pi_H)u(-c_{LN}),$$

it follows that $m_3 \geq v_L$, 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 $c_{LS} > c_{LN}$) as long as $\pi_L < 1$.* ■

Proposition B.4 *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 $g(\pi_H)u(v_H - m_4) = g(\pi_H)u(v_H - c_{LN}) + (1 - g(\pi_H))u(-c_{LN})$. The expected profit is $m_4\pi_H(1-\beta)N_H + c_{LN}(\beta N_H + N_L)$.*
Proof *The eight constraints for expected profit maximization are the following;*

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

PC_{LN} implies that all the other participation constraints hold, as per Lemma B.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

$$g(\pi_H)\left[u(v_H - c_{LN}) - u(v_H - m_4) - u(-c_{LN})\right] = -u(-c_{LN}),$$

which for $\pi_H < 1$ implies:

$$\pi_H\left[u(v_H - c_{LN}) - u(v_H - m_4) - u(-c_{LN})\right] < -u(-c_{LN}),$$

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 B.5 *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 $g(\pi_H)u(v_H - m_5) = g(\pi_H)u(v_H - c_{LS}) + (1 - g(\pi_H))u(-c_{LS})$. The firm's expected profit is $m_5\pi_H(N_H(1-\beta)) + c_{LS}\beta(N_H + N_L)$.*

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

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

From Lemma B.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

$$g(\pi_H)\left[u(v_H - c_{LS}) - u(v_H - m_5) - u(-c_{LS})\right] = -u(-c_{LS}),$$

which for $\pi_H < 1$ implies:

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

from which IC_{HS} follows. Moreover because the following inequality holds:

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

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 $c_{LS} > c_{LN}$ when $\pi_L < 1$. ■

Proposition B.6 *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 $v_H\pi_H(1-\beta)N_H + c_{HS}\beta N_H$.*

Proof *The four constraints for expected profit maximization are the following:*

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

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 $U_\theta^{S,R} \geq 0$ implies $U_\theta^{N,R} \geq 0$ and vice versa). It follows that the candidate equilibrium tariffs are determined by fixing the refundable tariff equal to the lowest valuation among the customers served by the firm. This leads to the following Propositions.

Proposition B.7 *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 $v_L(\pi_H N_H + \pi_L N_L)$.*

Proposition B.8 *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 $\pi_H v_H N_H$.*

Finally, in the third group of configurations, only the non-refundable tariff is chosen. In this case, given Lemma B.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 B.9 *Suppose HS, HL, LS and LN select the non-refundable tariff (configuration 9). Then $p^*_{NR} = c_{LN}$ and $p^*_R = p > v^H$.The firm's expected profit is $c_{LS}(N_H + N_L)$.*

Proposition B.10 *Suppose HS, HL and LS select the non-refundable tariff and LS does not buy (configuration 10). Then $p^*_{NR} = c_{LS}$ and $p^*_R = p > v^H$. The expected profit is $c_{LS}(N_H + \beta N_L)$.*

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

Proposition B.12 *Suppose HS selects the non-refundable tariff and HL, LS and LN do not buy. (configuration 12). Then $p^*_{NR} = c_{HS}$ and $p^*_R = p > v^H$. The firm's 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 B.13 *The candidate equilibrium 7 always guarantees higher expected profits than candidate equilibria 9 and 10. The candidate equilibrium 8 always guarantees higher expected profits than candidate equilibria 6, 11 and 12.*

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

B.2 The optimal tariff menu: the case of $\pi_L \to 1$ and $\pi_L \to \overline{\pi}$

In this Section we provide the characterization of the equilibrium for $\pi_L \to 1$ and $\pi_H \to \overline{\pi}$. We observe that $\lim_{\pi_L \to 1} c_{LN} = \lim_{\pi_L \to 1} c_{LS} = v_L$. which implies $m_2 = m_3$ and $m_4 = m_5$. The expected profits are summarized Table B.1 below.

Table B.1: Expected profits for each equilibrium configuration

Equilibrium candidate:	**Expected profits:**
1	$v_L \overline{\pi} N_H + v_L N_L$
2	$m_2 \overline{\pi} N_H + v_L N_L$
3	$m_3 \overline{\pi} N_H + v_L \beta N_L$
4	$m_4 \overline{\pi}(1-\beta) N_H + v_L(\beta N_H + N_L)$
5	$m_5 \overline{\pi}(N_H(1-\beta)) + v_L \beta(N_H + N_L)$
7	$v_L \overline{\pi} N_H + v_L N_L$
8	$\overline{\pi} v_H N_H$

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

Proposition B.14 *The candidate equilibria* 2 *always guarantee higher expected profits than candidate equilibria* 1, 7 *and* 3*. The candidate equilibrium* 4 *always guarantees a higher expected profit than* 5*.*

Proof 2 *dominates* 1 *and* 7 *because* $m_2 > v_L$*, and* 3 *because* $\beta < 1$*.* 4 *dominates* 5 *because* $\beta < 1$. ■

Candidate equilibria 4, 2 and 8 correspond to Configuration I, II and III respectively. The condition that makes Configuration I the optimal menu based on expected profit comparison is described in the main text.

B.3 Proof of Lemma 1

Suppose that $\overline{\pi} v_H N_H > \overline{\pi}\ m_{II} N_H + v_L N_L$. This condition implies that configuration III yields higher profits than Configuration II for $\pi_L \to 1$ and $\pi_H \to \overline{\pi}$. Therefore, if the menu choice is restricted to cases without naiveté-based discrimination (Configuration II and III), in the optimal menu only high-valuation types are served, and pay a price equal to v_H. In this case, the market is not fully covered, and the utility of low-valuation types is zero. With naiveté-based discrimination, i.e. if Configuration I is feasible and profitable, the naive high-valuation types pay m_I, lower than v_H, and the same is for the sophisticated types, since $v_L < v_H$. Low-valuation types are indifferent under Configuration I and III.

Now consider $\overline{\pi} v_H N_H < \overline{\pi}\ m_{II} N_H + v_L N_L$. Absent naiveté-based discrimination, the market is fully covered in equilibrium. If naiveté-based discrimination is considered (and profitable), a NH type ends up paying m_I, which is higher than m_{II}: therefore, he is worse off. The HS type is better off, since $v_L < m_I$. As in the previous case, low types are indifferent.

C A model with competition

We sketch a tractable model of competitive price discrimination, following Stole (1995) and Valletti (1999). We assume that consumers are located along a line of unit length, with point x being a consumer's position, while two firms are located at the extremes of the line (points 0 and 1). The distribution of consumers' "augmented" type (ω) does not depend on x. If θ is the consumer willingness to pay and x her position, then the utility function $u(\circ)$ when she buys from the firm located in 0 is given by $v_\theta - p - tx$, while it is $v_\theta - p - t(1-x)$ if she buys from the firm located in 1. As in Stole (1995) and Valletti (1999), we assume that the "horizontal" parameter x is observable, so that firms can offer refundable and non-refundable tariffs contingent on x. In equilibrium, each firm serves the half of the market which is the closest to it, while the other firm fixes the lowest possible tariffs (in our setting, both equal to zero) at these locations. From now on we focus on the case of small cancellation probability ($\pi_L \to 1$ and $\pi_H \to \overline{\pi} < 1$) and Configuration I, corresponding to naiveté-based discrimination. Since firms are symmetric, we consider the firm located at the left extreme (point 0) only. In addition to meeting the constraints for the monopoly case, the firm must guarantee each consumer a utility which is at least the utility provided by other firm, so that both the refundable and non-refundable price must be lower than $t(1-2x)$. We define these conditions as the competition constraints.

Ignoring competition constraints, Configuration I now entails $p^*_{NR} = v_L - tx$ and $p^*_R = m_I$ where m_I is the solution to $g(\pi_H)u(v_H - m_I) = g(\pi_H)u(v_H - v_L) + (1 - g(\pi_H))u(-v_L + tx)$. If competition is mild, i.e. if x is low and/or t is high, competition constraints are not binding, so that the monopoly case is unaffected. If competition is intense, i.e. if x is high and/or t is low, both constraints are binding. It follows that the cancellation premium is zero. For an intermediate level of competition intensity, the competition constraint is binding for the high type only. It follows that $p^*_{NR} = v_L - tx$ and $p^*_R = t(1-2x)$. This requires $v_L - tx < t(1-2x)$, i.e. $v_L < t(1-x)$. In this case, the cancellation premium is given by $p^*_R - p^*_{NR} = t(1-x) - v_L > 0$.

D A model with a capacity constraint

In this section, we outline an extension of the model that includes a capacity constraint. We assume that the firm can serve up to K customers; K therefore represents the available capacity at the beginning of the period. Consistently with the set-up of the basic model, the firm is myopic and maximizes expected profits in each period; for this reason, we omit the time index for capacity just as we did for the other variables. In addition, we focus on the case of $\pi_L \to 1$ and $\pi_H \to \overline{\pi}$, and further restrict the analysis to $\beta \geq 1/2$. By assuming that the fraction of sophisticated consumers is "large", we consider the area of the parameter space for which the conditions for the existence of naiveté-based discrimination are less likely to be met. *A fortiori*, our conclusions should hold for lower values of β. Finally, we assume that the firm adopts a quantity-based revenue management approach, by allocating capacity to each tariff. Because there is no uncertainty in demand, the firm serves fully the group of customers purchasing the tariff associated with higher expected profits; they then use the residual capacity to sell the other tariff. For $K \geq N_H + N_L$, the analysis is just the same as the one presented in Section 4 above. Here we consider $K < N_H + N_L$.

To solve the model, we first observe that, in each of the twelve configurations considered as candidate equilibria in Appendix Section B, the tariffs that maximize expected profits are unrelated to the number of consumers of each type served; in contrast, expected profits do, of course, depend on the number of consumers per type. To determine the equilibrium, we therefore proceed as follows:

- For each configuration, we determine i) the condition on K that makes a configuration *admissible*, i.e., such that the firm can serve (at least partially) the consumer types that are supposed to be served under that configuration; and ii) the corresponding expected profits, as a function of K.

- For each value of K, we compare the expected profits of all the admissible configurations to identify the optimal menu.

For each configuration, the condition for admissibility and the corresponding expected profits are as follows:

(Configuration 1) In this case, consumer types HS, HL and LN select the refundable tariff, whereas LS customers select the non-refundable tariff. The condition for this configuration to be admissible is $K > N_H + N_L(1-\beta)$. The firm's expected profit is $v_L K \pi^*_{R1} + v_L \beta N_L$, where $\pi^*_R = \overline{\pi}\frac{N_H}{N_H+N_L(1-\beta)} + \frac{N_L(1-\beta)}{N_H+N_L(1-\beta)}$ is the average cancellation probability for the consumers choosing the refundable tariff.

(Configuration 2) In this case, HS and HN select the refundable tariff and LS and LN choose the non-refundable tariff. The condition for admissibility of this configuration is $K > N_H$. The expected profit is $m_2 \overline{\pi} N_H + v_L(K - N_H)$. Note that with $\pi_L \to 1$ and $\pi_H \to \overline{\pi}$, configuration 3 is equivalent to configuration 2.

(Configuration 4) HN customers select the refundable tariff while HS, LS and LN opt for the non-refundable tariff. The condition for this configuration to be admissible is $K > N_H(1-\beta)$. The expected profit is $m_4\overline{\pi}(1-\beta)N_H + v_L(K-(1-\beta)N_H)$. In the case of $\pi_L \to 1$ and $\pi_H \to \overline{\pi}$ on which we focus in this extension,configuration 5 is equivalent to configuration 4.

(Configuration 6) Customers of type HN select a refundable tariff, HS selects a non-refundable tariff, and LS and LN do not buy. This case requires $K > N_H(1-\beta)$ to be admissible. The expected profit is $v_H\overline{\pi}(1-\beta)N_H + c_{HS}\min\{\beta N_H; K-(1-\beta)N_H)\}$.

(Configuration 7) In this case, all consumer types select the refundable tariff. This requires $K > N_H$. The firm's profit is $v_L K\pi^*_{R2}$, with $\pi^*_{R2} = \overline{\pi}\frac{N_H}{N_H+N_L} + \frac{N_L}{N_H+N_L}$.

(Configuration 8) HS and HN types select the refundable tariff and LS and LN types do not buy. This configuration is always admissible. The expected profit is $\overline{\pi}v_H\min\{K; N_H\}$.

(Configuration 9) In this case, all four types choose the non-refundable tariff. This configuration requires $K > N_H + N_L\beta$ t be admissible. The expected profit is $v_L K$.

(Configuration 10) HS, HL and LS types select the non-refundable tariff, whereas LN types do not buy. The admissibility condition is $K > N_H$, and the expected profit is $v_L\min\{K; N_H + \beta\ N_L\}$.

(Configuration 11) In this case, HS and HN customers select the non-refundable tariff, whereas LS and LN do not buy. The condition for this configuration to be admissible is $K > \beta N_H$. The expected profit is $c_{HN}\min\{K; N_H\}$.

(Configuration 12) HS customers buy the non-refundable tariff while HL, LS and LN do not buy. This configuration is always admissible, and will lead to an expected profit of $c_{HS}\min\{K; \beta N_H\}$.

Expected profit comparisons for each value of K lead to the following results.

1. If $K \leq (1-\beta)N_H$, the only admissible configurations are (8) and (12). Configuration (8) always guarantees higher expected profits than (12) because $\overline{\pi}v_H > c_{HS}$.

2. If $(1-\beta)N_H < K \leq \beta N_H$, the admissible configurations are (4), (6), (8) and (12), with configuration (8) dominating the others because the condition for (8) yielding higher expected profits than (6) is $K > (1-\beta)N_H$, and the condition for (8) to be superior to (4) is $K > \frac{\overline{\pi}m_4 - v_L}{\overline{\pi}v_H - v_L}(1-\beta)N_H$. This condition always holds because $\frac{\overline{\pi}m_4 - v_L}{\overline{\pi}v_H - v_L} < 1$ is

implied by $\overline{\pi} v_H > \overline{\pi} m_4 > v_L$ (the second inequality is demonstrated in Escobari and Jindapon (2014)).

3. If $\beta N_H < K \leq N_H$, the admissible configurations are (4), (6), (8), (11) and (12), with configuration (8) being again the dominating one. This occurs because (8) is preferred to (11) under $\overline{\pi} v_H > c_{HN}$.

4. If $N_H < K \leq N_H + (1-\beta)N_L$, the admissible configurations are (2), (4), (6), (7), (8), (10), (11) and (12). Configuration (8) dominates (10) because $\overline{\pi} v_H > v_L$. In turn, (10) yields higher expected profits than (7) because $\pi^*_{R2} < 1$. For configuration (4) to be the one yielding the highest expected profits, the requirement is that $K > N_H(1-\beta) + N_H \frac{\overline{\pi}}{v_L}(v_H - m_4(1-\beta)$ and $\overline{\pi}\left[m_4(1-\beta) - m_2\right] + v_L\beta > 0. If \mathrm{N}_H + (1-\beta) < K \leq N_H + \beta N_L$, the admissible configurations are (1), (2), (4), (6), (7), (8), (10), (11) and (12). (1) is dominated by (10) because $\pi^*_{R1} < 1$. The conditions for Configuration (4) to the one with the highest expected profits are the same as before.

6. If $\beta N_L + N_H < K \leq N_H + N_L$, all configurations are admissible. Configuration (9) is dominated by (2) since $\overline{\pi} m_2 > v_L$. The conditions for configuration (4) to be the one with the highest expected profit are the same as before.

The following Proposition describes the condition under which configuration 4, the one that exhibits naiveté-based discrimination, is the optimal one.

Proposition D.1 *If $K < N_H$, Configuration (4) is never the optimal configuration. If $K > N_H$, configuration (4) is optimal if $K > N_H(1-\beta) + N_H \frac{\overline{\pi}}{v_L}(v_H - m_4(1-\beta))$ and $\overline{\pi}\left[m_4(1-\beta) - m_2\right] + v_L\beta > 0$.*