Discussion Paper No. 18-059

Hotel Rankings of Online Travel Agents, Channel Pricing, and Consumer Protection

Matthias Hunold, Reinhold Kesler, and Ulrich Laitenberger

ZEW

Zentrum für Europäische Wirtschaftsforschung GmbH

Centre for European Economic Research

Discussion Paper No. 18-059

Hotel Rankings of Online Travel Agents, Channel Pricing, and Consumer Protection

Matthias Hunold, Reinhold Kesler,
and Ulrich Laitenberger

Hotel Rankings of Online Travel Agents, Channel Pricing, and Consumer Protection*

Matthias Hunold[†], Reinhold Kesler[‡] and Ulrich Laitenberger[§]

December 2018

Abstract

We investigate whether online travel agents (OTAs) assign hotels worse positions in their search results if these set lower hotel prices at other OTAs or on their own websites. We formally characterize how an OTA can use such a strategy to reduce price differentiation across distribution channels. Our empirical analysis shows that the position of a hotel in the search results of OTAs is better when the prices charged by the hotel on other channels are higher. This is consistent with the hypothesis that OTAs alter their search results to discipline hotels for aggressive prices on competing channels, thereby reducing the search quality for consumers.

Keywords: Consumer protection, free-riding, hotel booking, online travel agents, ranking, search bias.
JEL Class: D40, L42, L81

*First version: February 2017. Financial support by the State Government of Baden-Württemberg, Germany, through the research program "Strengthening Efficiency and Competitiveness in the European Knowledge Economies" (SEEK) is gratefully acknowledged. We thank three anonymous referees and an anonymous editor for extremely careful and constructive feedback. We thank the participants at the MaCCI Annual Conference 2017, the 9th ICT Conference at Télécom ParisTech 2017, the 15th ICT Conference at the ZEW 2017, the 19th ZEW Summer Workshop for Young Economists, the EARIE Conference 2017 in Maastricht, the German Economic Association Annual Meeting 2017 in Vienna, the 4th IO in the Digital Economy Workshop in Liège, seminars at the universities of Bergen, Dortmund, Düsseldorf, Freiburg, Palma, and Zürich, the Workshop on E-Commerce and Platforms at CREST, the 15th Digital Economics Summer School in Montpellier, and in particular Apostolos Filippas, Vicente Lagos, Martin Peitz, Johannes Muthers, Christoph Schottmüller, Chengsi Wang, and Julian Wright for helpful comments.

[†]Heinrich-Heine-Universität Düsseldorf, Düsseldorf Institute for Competition Economics (DICE), Universitätsstr. 1, 40225 Düsseldorf, Germany; hunold@dice.hhu.de.

[‡]Centre for European Economic Research (ZEW), MaCCI Mannheim Centre for Competition and Innovation, Address: L7,1, 68161 Mannheim, Germany; kesler@zew.de.

[§]Télécom ParisTech, 46 Rue Barrault, 75013 Paris, France, and ZEW (as above); laitenberger@enst.fr.

1 Introduction

Online travel agents (OTAs) have become important intermediaries for hotels and consumers.[1] They pool the offers of different hotels and help consumers to find a hotel room offer. A consumer searching for a hotel room at an OTA typically retrieves a list with offers from various hotels. The ranking of these offers affects which hotels consumers are likely to book (Ursu, 2018). Consumers may often believe that offers with higher match values have been ranked higher according to the consumer's own needs. More suspicious minds may argue that OTAs charge the hotels commission fees per booking, such that OTAs benefit when offers with a high likelihood of a booking *at the OTA* are made the most visible. In this article we investigate both theoretically and empirically the question of whether OTAs condition their rankings of search results on the hotel prices at other OTAs or on the hotels' websites. Our results suggest that the OTA conditions its recommended ranking on certain factors that are relevant for the OTA to maximize its profit, but arguably not to maximize the match value of consumers.

In the theory part, we first establish a reference ranking of hotel offers according to the expected match values for consumers. The labels on major OTAs' default rankings are consistent with this approach. For instance, Expedia labels its default ranking "Recommended" and Booking.com uses the terms "Our top picks" and "The best selection for business travelers."[2] We then show that an OTA which maximizes its profits may have incentives to present offers with a higher booking likelihood on the OTA website more prominently, to the detriment of the matching quality (search bias). We show that this can create a relationship between the ranking and the hotel prices on other sales channels. Through this relationship the OTA can influence the hotels in their pricing across distribution channels without directly conditioning the ranking on the price differentials across channels. This is to some extent a substitute to price parity clauses (PPCs)[3] as it can induce hotels to reduce discounts on direct channel prices relative to OTA prices. Moreover, we show that an OTA may also employ a policy whereby it conditions the ranking directly, and possibly more drastically, on whether a hotel is offering lower prices on competing channels. We consider this distinction to be relevant. For instance, in the context of narrowing down its PPCs in Europe, Booking.com committed to various na-

[1] See European Hotel Distribution Study 2018 (last accessed August 31, 2018).

[2] Labels last checked on August 31, 2018. The term for business travelers is translated from the German website version.

[3] With a price parity clause an OTA contractually obliges a hotel to not charge lower prices on (certain) competing sales channels, such as other OTAs and the direct channel of the hotel. They are also called best price clauses and (retail) most favored nation (MFN) clauses.

tional competition authorities to not directly condition the ranking on price parity across channels.[4]

We investigate empirically how changes of a hotel's prices on competing sales channels affect the hotel's position in the default rankings that appear in response to a search query via Booking.com and Expedia. We use web-scraped data of (i) Booking.com, (ii) Expedia, and (iii) the meta-search site Kayak.[5] The data comprises daily search results for 250 cities in 13 countries, mostly from Europe, between July 2016 until January 2017. Thus, we can track the listings and prices of more than 18,000 hotels for the same hotel room offers on different online channels. We use additional information to isolate the effect of a lower price on competing sales channels on the ranking position from other factors that might affect the overall ranking. These include promotional activities for specific hotel offers, as reported by the OTA, as well as information on whether hotels paid additional fees to the OTA in order to improve their rankings. We use both linear panel estimations with hotel fixed effects and rank-ordered logit regressions.

We find that for a given hotel price at an OTA, a lower price at the other OTA or on the hotel's website leads to a worse ranking position, both at Booking.com and Expedia. This relationship holds in countries both with and without PPC clauses. It is also robust to different functional forms of the price differences and significant for both downward and upward deviations. The relationship is more pronounced for substantial price differences of around 10% and above. When using specifications with lags and leads of the price differences across channels, we do not see a systematic pattern of significant correlations between price differences in the future with rankings today (which could be an indication of reverse causality). We also find that hotels price differentiate in a plausible way. For instance, hotels set the direct channel price lower if demand is high and the travel date is close to the booking date.

We conclude with a discussion of the potential risks of an OTA's ranking optimization being based on measures such as the prices of competing offers, which by themselves are unrelated to the intrinsic consumer value of the offers presented by the OTA. Regarding consumer protection, we debate whether consumers should be made more aware of how the rankings of OTAs and similar online platforms are computed.

[4] See section 3 for references and subsection 4.7 for a more detailed discussion.

[5] The website of the meta-search site Kayak looks similar to that of an OTA such as Booking.com, but – different from an OTA – presents price offers of the different OTAs and the hotel websites.

2 Related literature

Our theory relates most closely to the literature on intermediaries who can decide which product to present first or particularly prominently (Raskovich, 2007; Inderst and Ottaviani, 2012; Hunold and Muthers, 2017; Shen and Wright, 2018). Hagiu and Jullien (2011, 2014) specifically analyze biases in the rankings of search engines. In a setting where customers have heterogeneous search costs and the platform has a per-click payment scheme, Hagiu and Jullien (2011) predict distortions in the ranking in the sense that the less suitable product is displayed first to generate additional revenue from the product providers. Moreover, they also consider the situation that producers know the result of the ranking algorithm before setting prices. De Corniere and Taylor (2014) show that vertically integrated search engines distort search results, but the overall welfare effect is unclear because, for example, the integrated search engine can have a strong incentive to generate demand.

To our knowledge, to date there are no empirical research articles that analyze the relation between the position of hotels in an OTA's search results and the hotel's prices on competing sales channels. Yet, there is a literature which highlights the importance of OTA rankings for the booking choices of consumers. Chen and Yao (2016), De los Santos and Koulayev (2017), Koulayev (2014), and Ghose et al. (2012, 2014) study how rankings affect consumer online choices in the hotel market and provide estimates of the US dollar equivalent of a change in a hotel offer's indirect utility for a consumer resulting from a one-position increase in a hotel's ranking (position effect). De los Santos and Koulayev (2017) find position effects ranging from \$7.76 to \$35.15, whereas Ghose et al. (2012) find an average effect of \$6.24, and Koulayev (2014) reports a range from \$2.93 to \$18.78. Ursu (2018) exploits a random variation in the ranking of hotels at the OTA Expedia and finds lower position effects, ranging between \$0.55 to \$3.19. She finds that consumers click more often on an offer that is ranked better to get detailed information about it. However, conditional on seeing the detailed information, the ranking position does not influence the booking behavior of consumers.

Lu et al. (2015) study the relationship between the pricing of intermediaries, such as physical travel agents, and the introduction of a new online direct sales channel of a hotel chain. Using the data of hotel room transactions from 2001 to 2007, they analyze the introduction of the direct online sales channel in 2002 and find a significant reduction of the intermediaries' price premia. This result suggests that there is competition between different forms of hotel distribution channels.

Our work is also related to the recent theoretical literature on the (anti-)competitive effects of price parity clauses (PPCs) of intermediaries, such as OTAs (Edelman and Wright, 2015; Boik and Corts, 2016; Johnson, 2017; Wang and Wright, 2017; Johansen and Vergé, 2017; Ronayne and Taylor, 2018; Wals and Schinkel, 2018). To this literature we add the insight that non-contractual measures, such as ranking the search results on factors other than maximizing consumer matches, can have similar effects as PPCs.

This article also relates to the growing empirical literature on the effects of price parity clauses by online booking platforms, such as Hunold et al. (2018) and Mantovani et al. (2017). Hunold et al. (2018) study the PPCs of OTAs using meta-search price data of hotels. They find that PPCs influence the pricing and availability of hotel rooms across online sales channels. In particular, the abolition of Booking.com's narrow PPC is associated with a hotel's direct channel being the price leader more often. Moreover, they find that hotels make rooms more often available at Booking.com when it does not use the narrow PPC. These findings indicate that pricing across channels is important for hotels and also raise the question of whether non-contractual measures that aim at the OTA's prices being lowest have become more important in jurisdictions where contractual measures, in particular price parity clauses, were restricted.

3 Industry background: Online hotel booking

Distribution channels. One can book hotels through various channels. Traditionally, a large share of consumers book rooms directly at the hotel, both offline by telephone or walk-in, and online via the hotel website and e-mail. According to the European hotel association HOTREC, these channels make up 50 percent of all bookings, the latter comprising 14 and 16 percent, respectively.[6] In the last decade, OTAs became very important for hotels as a distribution channel. While the share of hotel bookings through OTAs was already 22 percent in 2013, this number steadily increased to 29 percent in 2017, comprising around 75 percent of all website-based bookings of hotel rooms. The OTA industry is highly concentrated – in 2017 Booking.com and Expedia accounted for more than 80 percent of those bookings in Europe, with Booking.com having, on average, a share of more than 60 percent.

Ranking of hotel offers at OTAs. OTAs are intermediaries that pool the offers of different hotels and help consumers to find a hotel room. In response to a search request, consumers retrieve a list of recommended offers. The label for this default list is "Our top picks" at Booking.com and "Recommended" at Expedia.[7] Despite the ability of consumers

[6] This and the following data is taken from the European Hotel Distribution Study 2018 (last accessed August 31, 2018).

[7] Labels last checked on August 31, 2018.

to refine their search through filters (sorting by price, location, etc.), there is evidence that a significant share of consumers rely on default rankings. Using a dataset from the Wharton Customer Analytics Initiative, Ursu (2018) shows that only 34 percent of consumers sort or filter hotel search results at an undisclosed OTA.[8] The rest make use of the default ranking, and as Ursu (2018) shows, high-ranked hotels receive more clicks (and hence more bookings). De los Santos and Koulayev (2017) report for a price comparison website that more than 50 percent of users rely on the default ranking. Similarly, Blake et al. (2016) report that on eBay between 67 and 85 percent of search requests have the default ranking.

Although OTAs do not reveal their default ranking algorithm, Expedia provides a general idea of the typical criteria.[9] First, the ranking depends on a price-to-value benchmark, which is based on criteria like price, ratings, location, etc. Second, it depends on how well a hotel partners with the OTA, e.g., by comparing room rates offered on the platform to those offered on other channels, or by assessing how much information the hotel provides on the platform. Third, a hotel's ranking position is affected by the commission rate it pays to the OTA. Booking.com's ranking is stated to be dependent on information provided by the hoteliers and customers, including among others conversion, availability, and pricing data.[10]

Pricing. Most OTAs today use the agency model, where hotels sign up at the OTA, provide details regarding their offers, and set retail prices on their own. Any offers that match a specific search request are listed in the corresponding search output. While signing up is usually free of charge, the OTA demands a commission for each booking on its website. Booking.com has used this model for a long time and we understand that Expedia also typically uses this model in Europe.[11] The commission is a share of the hotel price.[12] Base commission rates are around 10 to 15 percent for both Booking.com and Expedia and have remained constant in recent years. However, effective commission rates can be significantly higher.[13] Booking.com offers a "Visibility Booster" that allows hotels to pay higher commissions in exchange for better ranking positions on specific dates, and the "Preferred Partner Program," which also involves paying a higher commission for "extra

[8] See Ursu (2018), page 6 of the Online Appendix B, for details (last accessed August 31, 2018).
[9] See, for instance, a communication by Expedia that targets hoteliers (last accessed August 31, 2018).
[10] See the respective help page for hoteliers on Booking.com (last accessed August 31, 2018).
[11] See the annual reports of Booking.com and Expedia as well as the article *Expedia and TripAdvisor Separately Kill 2 Brands That Were Once Sister Companies* (last accessed August 31, 2018).
[12] See the respective help page for hoteliers on Booking.com (last accessed August 31, 2018).
[13] See Hunold et al. (2018), Online Appendix VIII, for references and details.

visibility," among other things.[14] Expedia offers a similar service called "Accelerator" through which hotels can improve their ranking position by paying a higher commission.[15] The additional commission paid for such services can be as high as 10 to 15 percent.[16]

Price parity clauses (PPCs). A major concern for OTAs is free-riding, where consumers use the platform's search and comparison service, but ultimately book through a cheaper channel, leaving the OTA without the commission.[17] As free-riding is incentivized by lower prices on other channels, many OTAs have used, and partly still use, price parity clauses (PPCs) in order to mitigate aggressive pricing. There are two main types of PPCs to distinguish. Under a wide PPC, an OTA obliges the hotel not to charge a higher price on the OTA than on any other online (and partly offline) booking channel, which in particular includes other OTAs and the hotel's own direct sales channels. Narrow PPCs prohibit the hotel from offering lower prices on its direct online sales channels than at the OTA that imposes the clause, but do not contractually restrict the hotel's room prices at other OTAs.[18] These PPCs in the contracts between OTAs and hotels differ fundamentally from price matching policies, where OTAs guarantee customers the lowest available price for an offer. Both Booking.com and Expedia had a price matching policy in place during the period of observation, with Expedia restricting this guarantee to registered users in late 2017.[19]

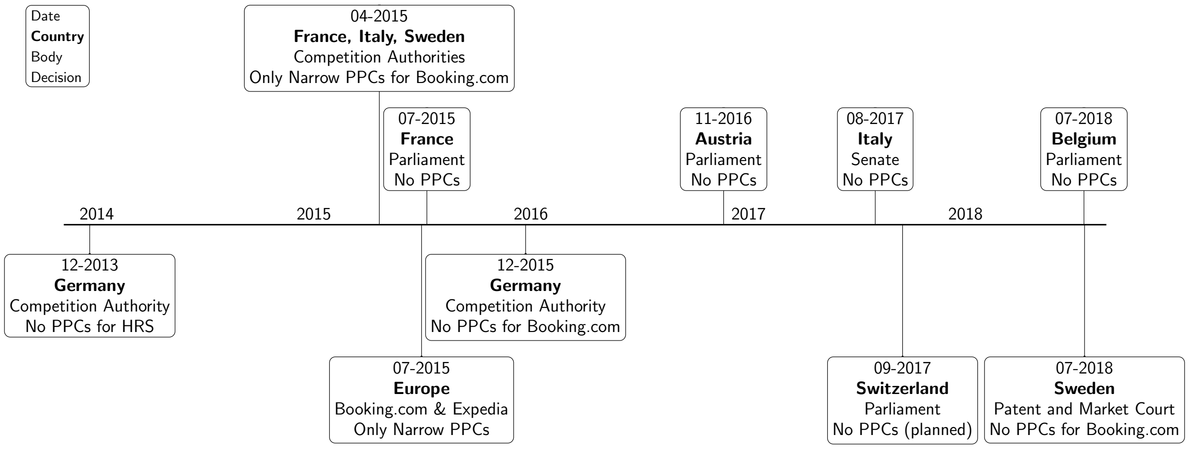
National competition law enforcement or new national laws typically target the contracts between the OTAs and the hotels in the respective jurisdiction. Various national competition authorities in Europe considered that (wide) PPCs could restrict competition between OTAs with respect to commission rates, which is the main theory of harm (see the discussion in section 2).

Figure 1 shows a timeline of policy changes regarding PPCs. In December 2013, the German competition authority prohibited the German incumbent OTA HRS from applying PPCs.[20] In April 2015, following joint investigations, competition authorities in Sweden, France, and Italy simultaneously accepted commitments from Booking.com to only impose narrow PPCs. Subsequently, in June Booking.com announced an extension of these

[14] See Booking.com's website on the Preferred Partner Programme (last accessed August 31, 2018).
[15] See Expedia's website on Visibility (last accessed August 31, 2018).
[16] See the article *First Look at Expedia Accelerator Program for Improving Hotel Placement* (last accessed August 31, 2018).
[17] See, for instance, Booking.com's arguments as discussed in the German Federal Cartel Office's decision prohibiting Booking.com's narrow price parity clauses (last accessed August 31, 2018).
[18] See Hunold (2017) for details and a comparison of different decisions with respect to PPCs in Europe.
[19] See the archived websites of Booking.com and Expedia respectively (last accessed August 31, 2018).
[20] See German Federal Cartel Office's decision prohibiting HRS to apply price parity clauses (last accessed August 31, 2018).

commitments to the rest of Europe.[21] Expedia reacted to this in July 2015 by announcing that it would also abolish their wide PPCs in Europe.[22]

Figure 1: Policy actions against price parity clauses (PPCs) of online travel agents (OTAs) – Timeline



In the following years, narrow PPCs were also prohibited in several countries. Starting with the removal of PPCs by the French parliament in July 2015, OTAs are nowadays also prohibited from using narrow PPCs in Austria, Belgium, Italy, Sweden, and soon in Switzerland.[23] In Germany, only the narrow PPCs of Booking.com and HRS are explicitly prohibited, whereas the case for Expedia is still pending. To the best of our knowledge, authorities in northern America have not taken action against the PPCs of OTAs.[24] We understand that OTAs still use wide PPCs in the US and Canada. Although the largest OTAs are not allowed to use (wide) PPCs in Europe anymore, hotels raise concerns that OTAs might implicitly enforce price parity by lowering the ranking positions of hotels which offer better rates elsewhere.[25]

[21] See the amendment by Booking.com (last accessed August 31, 2018).

[22] See the amendment by Expedia (last accessed August 31, 2018).

[23] See Hunold et al. (2018), Online Appendix I, and the respective pages of Belgium, Italy, Sweden, and Switzerland for more recent actions (last accessed August 31, 2018).

[24] In 2014, the US District Court for the Northern District of Texas dismissed an antitrust lawsuit aimed at restricting OTAs in applying PPCs. (last accessed August 31, 2018).

[25] See section 4.15, page 23 of the Report on the Monitoring Exercise carried out in the Online Hotel Booking Sector by EU Competition Authorities in 2016 (last accessed August 31, 2018).

4 Model

4.1 Best match ranking

Our reference is a ranking that maximizes consumer net match values of the offers available at the OTA. We illustrate such a ranking for the case of two symmetric, spatially differentiated hotels and one OTA. This generalizes to other forms of differentiation and more than two hotels.

Hotel L is located at the left end of a line of length one, hotel R at the right end. There is a mass one of consumers. Each consumer has an ideal location $x \in [0, 1]$, which is uniformly distributed on the unit interval. The OTA learns about the location as the consumer types in its location x in the search request. We assume that this translates into a probability of x that hotel R is the best match and, by symmetry, a complementary probability $1-x$ that hotel L is the best match. In the case of a best match, let the gross match value be V_i, $i \in \{L, R\}$, and $v_i < V_i$ otherwise (this could be zero, for instance). The net match value in the case of a best match with hotel i is thus $V_i - p_i^{\mathrm{OTA}}$ (assuming that V_i is larger than p_i^{OTA}, the price for hotel i at the OTA). The expected match value is $x \cdot \left(V_R - p_R^{\mathrm{OTA}}\right)$ for hotel R and $(1-x)\left(V_L - p_L^{\mathrm{OTA}}\right)$ for hotel L.

A best match ranking is such that for each x, the OTA shows the hotel that yields the highest expected match value on top. For symmetric valuations and prices, this means hotel R for $x \geq 0.5$ and hotel L for $x \leq 0.5$. We summarize the hotel's ranking choice with the threshold value, which we denote r. For symmetric hotels, the value $r = 0.5$ thus defines the best match ranking; a value of $r < 0.5$ means that the ranking is biased toward hotel R, and $r > 0.5$ implies a bias toward hotel L. For the purpose of this analysis we use the terms *bias* and *search bias* as descriptions of a ranking that does not maximize the match values for consumers. Hence, we do not exclude that there might be legitimate business reasons for a deviation from a best match ranking.

Lemma 1. *A best match ranking depends on the prices at the OTA, but not on the prices on other channels.*

4.2 Hotel pricing and standard OTA ranking optimization

We now analyze how an OTA can use the ranking to optimally respond to the prices charged by hotels on their different channels. Each hotel has a direct channel ("DIR") and is present at the OTA. We first consider the following game:

1. Each hotel $i \in \{\mathrm{L,R}\}$ sets prices p_i^k on the channels $k \in \{\mathrm{DIR,OTA}\}$.

2. The OTA observes these prices and optimizes its ranking with respect to r.

We refer to this as *standard profit maximization*. We solve this game of complete information for Subgame Perfect Nash Equilibria using backward induction. This means that we first solve the OTA ranking optimization for given hotel prices (subsection 4.3), and then solve the hotels' pricing game in anticipation of the OTA's ranking optimization (subsection 4.4). An OTA may also employ a policy whereby it announces it will condition the ranking directly on price differences in a certain way. We show in subsection 4.7 that it can be profitable for an OTA to commit to such a *direct conditioning* prior to hotels setting prices.

In order to keep the model simple, we conduct our analysis for a given level of commission rates that the OTAs charge the hotels per booking on the OTA website. Arguably, a major determinant of the commission rate is competition among OTAs (from which we abstract in this model). Other articles specifically study the commission rates and show that OTA competition is particularly important in this context (Boik and Corts, 2016; Johnson, 2017; Wang and Wright, 2017; Shen and Wright, 2018; Teh and Wright, 2018). Endogenizing the commission rates would make the model much more complicated and is not the core of the present analysis.[26] Our model nevertheless provides interesting insights into the relationship of pricing and ranking for any level of commission rates.

The demand of hotel i is given by

$$q_i = \underbrace{[\alpha\phi_i(r) + (1-\alpha)]}_{\text{consumers interested in hotel i}} \cdot \underbrace{\left[q_i^{\text{OTA}} + q_i^{\text{DIR}}\right]}_{\text{booking likelihood}}. \tag{1}$$

The term in the left brackets is the mass of potential consumers of hotel i. The parameter $\alpha \in [0,1]$ denotes the fraction of consumers for which the ranking is relevant.[27] For instance, α of the consumers may rely on the ranking of the OTA, whereas $1-\alpha$ of the consumers find hotels in other ways.[28] For $\alpha > 0$, the OTA can use the ranking decision r to influence whether more consumers are interested in one or the other hotel.

Assumption 1. *The function $\phi_i(r)$, $i = L, R$, with $r \in [0,1]$, satisfies the following conditions:*

[26] See, for instance, Inderst and Ottaviani (2012) who show that commission payments increase when the intermediary is less interested in customers buying the right product. Shen and Wright (2018) introduce a direct sales channel in the model and show that when the intermediary chooses the level of commission, free-riding due to lower direct selling prices tends to occur.

[27] Parameter α can also be interpreted as the degree with which a representative consumer is sensitive to the ranking.

[28] For instance, consumers might use a different ordering of offers at the OTA, such as by lowest price, or a general search engine, such as Google, or a meta-search site for hotels that collects prices from different channels, such as Kayak or Trivago.

• $\partial\phi_L/\partial r > 0 > \partial\phi_R/\partial r$ *(bias is good for one and bad for the other hotel),*
• $\frac{1}{2} = \arg\max_r \phi(r)$ *(maximum is attained at unbiased ranking of* $1/2$*),*
• $\phi_i(r) = \phi_{-i}(1-r)$ *(symmetry between hotels),*
• $\phi_i'' < 0$ *(strict concavity – to ensure well-behaved optimization problems).*

One can obtain a simple parameterization of ϕ_i from the example in subsection 4.1. For this, assume that consumers only consider buying the highest ranked hotel offer.[29] Moreover, assume that the value in the case of a bad match is sufficiently low such that a booking does not take place ($v_i = 0$, for instance). For a given threshold $r \in [0,1]$, the mass of consumers that sees hotel L as the top-ranked hotel and realizes the high match value is

$$\phi_L = \int_0^r (1-x)\,dx = \left[x - \frac{1}{2}x^2\right]_0^r = r - \frac{1}{2}r^2, \tag{2}$$

and that for hotel R is

$$\phi_R = \int_r^1 x\,dx = \left[\frac{1}{2}x^2\right]_r^1 = \frac{1}{2}\left(1-r^2\right). \tag{3}$$

The best match ranking is obtained for $r = 1/2$. If $r > 1/2$, the OTA is said to bias the ranking in favor of L (and to R for $r < 1/2$). The total mass of consumers that realize a high match value and are interested in booking a hotel, $\phi_L + \phi_R$, decreases in the OTA's bias toward one hotel, that is, $|r - \frac{1}{2}|$.

The terms $q_i^{\text{OTA}}(p_i^{\text{OTA}}, p_i^{\text{DIR}})$ and $q_i^{\text{DIR}}(p_i^{\text{DIR}}, p_i^{\text{OTA}})$ in the right brackets of equation (1) can be interpreted as the likelihood that a consumer will book hotel i through the direct channel or the OTA. The likelihood of booking on a channel decreases in the price on that channel, increases in the price of the other channel, and decreases when both prices increase.

Assumption 2. *The booking likelihood satisfies the following conditions:*
• $\partial q_i^k/\partial p_i^k < 0$ *(demand decreases in own price),*
• $\partial q_i^k/\partial p_i^{-k} > 0$, *(channels are substitutes),*
• $\partial q_i^k/\partial p_i^k + \partial q_i^k/\partial p_i^{-k} < 0$ *(demand decreases in price level),*
• $q_i^{DIR}(x,y) = q_i^{OTA}(x,y)$ *(symmetry across channels).*

To ensure that the OTA's maximization problem has an interior solution, we assume that the expected booking revenue $p_i^k q_i^k$ of a channel is a concave function of the price level,

[29] A similar modeling approach is used by Raskovich (2007); Inderst and Ottaviani (2012); Hunold and Muthers (2017); Shen and Wright (2018).

so that $\partial^2(p_i^k q_i^k)/(\partial p_i^{\mathrm{DIR}})^2 + \partial^2(p_i^k q_i^k)/(\partial p_i^{\mathrm{OTA}})^2 < 0$ for $p_i^{\mathrm{OTA}} = p_i^{\mathrm{DIR}}$.[30] For instance, this is satisfied in the case of the linear relationship $q_i^k = 1 - p_i^k + \gamma \cdot p_i^{-k}$, with $0 < \gamma < 1$.

4.3 OTA optimization (standard profit maximization)

The profit of the OTA is the commission times the bookings at the OTA:

$$\pi^{OTA} = \sum_{i \in \{L,R\}} c_i \, p_i^{\mathrm{OTA}} \cdot [\alpha \phi_i(r) + (1-\alpha)] \cdot q_i^{\mathrm{OTA}}, \tag{4}$$

where c_i is the commission rate the OTA charges hotel i. Using the symmetry assumption of ϕ_i, we can write the OTA's first-order condition with respect to r as

$$c_L \, p_L^{\mathrm{OTA}} \cdot q_L^{\mathrm{OTA}} \phi_L'(r) = c_R \, p_R^{\mathrm{OTA}} \cdot q_R^{\mathrm{OTA}} \cdot \phi_L'(1-r). \tag{5}$$

If the expected OTA revenues are equal, the optimal ranking decision is $r = \frac{1}{2}$, which maximizes the mass of interested consumers. Otherwise, the OTA optimally favors the hotel with the higher expected margin. This yields

Lemma 2. *The OTA's ranking of a hotel is better if the commission rate c_i that the hotel pays is higher (holding other things equal).*

A decrease in the hotel's direct price decreases the expected OTA profit from that hotel as $\partial q_i^{OTA}/\partial p_i^{DIR} > 0$. This yields

Proposition 1. *A hotel's rank at an OTA is worse if the hotel charges a lower price on its direct online sales channels (holding other things equal).*

An increase of the OTA price p_i^{OTA} decreases the likelihood of a booking ($\partial q_i^{\mathrm{OTA}}/\partial p_i^{\mathrm{OTA}} < 0$), but increases $c_i \, p_i^{\mathrm{OTA}}$, the revenue per booking. The relationship between the OTA price and ranking is therefore not obvious at this point, but we can still state

Lemma 3. *A hotel's rank at the OTA depends positively on the OTA price if the expected channel revenue $p_i^{OTA} q_i^{OTA}$ increases in p_i^{OTA} (holding other things equal).*

[30] Note that we abstract from direct price competition between the hotels through the booking likelihoods for a given ranking for reasons of tractability. This is also consistent with the interpretation of the booking likelihood in the previous subsection if customers only look at the top ranking.

4.4 Hotel optimization

Let r^* denote the solution to the OTA's ranking optimization characterized by (5) as a function of p_i^{OTA}, p_i^{DIR}, and c_i with $i \in \{L, R\}$. A hotel's profit is given by

$$\pi_i = \underbrace{[\alpha\phi_i(r^*) + (1-\alpha)]}_{\text{Consumers interested in booking hotel i}} \cdot \underbrace{\left[\left(p_i^{\text{DIR}} - b\right) q_i^{\text{DIR}} + \left((1-c_i)\, p_i^{\text{OTA}} - b\right) q_i^{\text{OTA}}\right]}_{\text{Hotel's expected profit per consumer } (\pi_i^C)}. \tag{6}$$

For consumers booking on the direct channel, the hotel's margin is the direct price minus variable hosting costs $b > 0$. For consumers booking on the OTA, the commission is an additional cost.[31]

In what follows, we restrict attention to demand functions that give rise to quasi-concave reduced-form hotel profits and a stable interior equilibrium. The following FOCs with respect to p_i^{DIR} and p_i^{OTA} characterize the optimal prices:

$$\frac{\partial \pi_i}{\partial p_i^k} = [\alpha\phi_i(r^*) + (1-\alpha)]\, \frac{\partial \pi_i^C}{\partial p_i^k} + \alpha\phi_i'(r^*)\frac{\partial r^*}{\partial p_i^k}\pi_i^C = 0, \tag{7}$$

with $k \in \{\text{DIR,OTA}\}$, $i \in \{\text{L, R}\}$.

4.5 Equilibrium with price parity restriction ($p_i^{\text{OTA}} = p_i^{\text{DIR}}$)

Rankings do not matter ($\alpha = 0$)

For $\alpha = 0$ the ranking does not affect a hotel's profit, which becomes

$$\pi_i = \pi_i^C = \left(p_i^{\text{DIR}} - b\right) q_i^{\text{DIR}} + \left((1-c_i)\, p_i^{\text{OTA}} - b\right) q_i^{\text{OTA}}.$$

The optimal uniform price $p_i^*(\alpha = 0) = p_i^{\text{OTA}} = p_i^{\text{DIR}}$ is defined by

$$\frac{\partial \pi_i^C}{\partial p_i^{\text{OTA}}} + \frac{\partial \pi_i^C}{\partial p_i^{\text{DIR}}} = 0. \tag{8}$$

Lemma 4. *When rankings do not matter ($\alpha = 0$) and price parity is required, the optimal uniform price is above the level that maximizes the expected channel revenue $p_i^{OTA} q_i^{OTA}$.*

Proof. See Appendix A. □

[31] We assume here for simplicity that there are no specific direct booking costs. They are most likely at least significantly lower than the OTA commission.

Rankings matter ($\alpha > 0$)

When rankings matter ($\alpha > 0$), the hotels compete for a favorable ranking of the OTA. This makes them internalize the profitability of their offers for the OTA.

Proposition 2. *When rankings matter ($\alpha > 0$), each hotel has an incentive to reduce the uniform price below the level $p_i^*(\alpha = 0)$, as this makes the ranking more favorable.*

Proof. See Appendix A. □

A hotel does not set a lower price than the price that maximizes the OTA's revenues from the hotel (a further decrease would lead to a less favorable ranking). Down to this level, the OTA benefits from price cuts – which the ranking induces. Therefore, biasing rankings is profitable for the OTA, but hurts symmetric hotels (that pay the same commission rates and the same costs) as they earn lower margins, even though in a symmetric equilibrium rankings are unbiased.

Proposition 3. *In the case of a price parity clause, biasing rankings leads to higher OTA profits, and lower profits for symmetric hotels. In equilibrium, consumers benefit through lower prices.*

To interpret the result of lower prices, one must keep in mind that we have kept the commission rate as exogenous for this analysis. Please see the discussion in subsection 4.8.

4.6 Equilibrium without price parity restriction

Case: Rankings do not matter ($\alpha = 0$)

The hotel's optimal prices p_i^{OTA} and p_i^{DIR} solve

$$\frac{\partial \pi_i^C}{\partial p_i^{\text{OTA}}} = (1 - c_i)\, q_i^{\text{OTA}} + \Big((1 - c_i)\, p_i^{\text{OTA}} - b\Big) \frac{\partial q_i^{\text{OTA}}}{\partial p_i^{\text{OTA}}} + \Big(p_i^{\text{DIR}} - b\Big) \frac{\partial q_i^{\text{DIR}}}{\partial p_i^{\text{OTA}}} = 0, \tag{9}$$

$$\frac{\partial \pi_i^C}{\partial p_i^{\text{DIR}}} = q_i^{\text{DIR}} + \Big(p_i^{\text{DIR}} - b\Big) \frac{\partial q_i^{\text{DIR}}}{\partial p_i^{\text{DIR}}} + \Big((1 - c_i)\, p_i^{\text{OTA}} - b\Big) \frac{\partial q_i^{\text{OTA}}}{\partial p_i^{\text{DIR}}} = 0. \tag{10}$$

At a zero commission rate ($c_i = 0$) the hotel sets equal prices on the direct channel and at the OTA. A positive commission rate $c_i > 0$ implies a higher OTA price for a given direct price, yielding

Lemma 5. *When rankings do not matter ($\alpha = 0$) and a hotel can price differentiate, the optimal OTA price is above the optimal direct price.*

Proof. See Appendix A. □

Case: Rankings matter ($\alpha > 0$)

When rankings matter ($\alpha > 0$), the hotel's marginal profits with respect to its two prices $p_i^{\text{OTA}}$and p_i^{DIR} are not zero at the prices defined by equations (9) and (10) of the case that rankings do not matter ($\alpha = 0$). In particular, the marginal profit $\partial\pi_i/\partial p_i^{\text{DIR}}$, as stated in equation (7), is positive because the ranking improves in the direct price: $\phi_i'(r^*) \cdot \partial r^*/\partial p_i^{\text{DIR}} > 0$ (recall Lemma 2).

Lemma 6. *Starting from the prices that are optimal when rankings do not matter, the hotel has an incentive to raise the direct price in order to improve its OTA ranking when rankings matter.*

As the initial OTA price is above the direct price, this means that the ranking induces the hotel to reduce the discount for the direct channel. Similarly, the hotel has an incentive to reduce the OTA price if an increase in this price reduces the expected OTA revenue per consumer ($c_i p_i^{\text{OTA}} q_i^{\text{OTA}}$).[32]

The practice of including the probability of booking into the ranking algorithm tends to reduce the price dispersion between the OTA and the direct channel, similar to a parity clause.[33] In other words, such non-contractual measures can act as a substitute to the price parity clause – which policymakers in various countries prohibited (see section 3). Biasing rankings increases the OTA profit, both when a price parity is applied and when it is not applied, as the hotels internalize the OTA profit and adjust their prices accordingly. It is interesting to note that biasing rankings tends to be more profitable for the OTA when no price parity is used, compared to when it is used.[34] The intuition is that a price parity clause and biasing rankings (when used as a non-contractual disciplinary measure) are substitutes to some extent. Both reduce the markup of the OTA price over the direct price (or more generally a price on a channel with lower distribution costs).

4.7 Ranking with direct conditioning on other channel prices

We previously studied the case of standard profit maximization where the OTA ranks hotels for given hotel prices with the objective to maximize the product of commission and booking likelihood for a given website user. In this case, the ranking depends on prices on other channels only indirectly through the booking likelihood. The OTA could, how-

[32] This is clearly the case with uniform prices, see Lemma 4.

[33] One can show that this is the case for large parameter ranges when using a linear specification for q_i and the parameterizations in (2) and (3) for ϕ_i.

[34] See footnote 33.

ever, also react more directly and possibly more drastically to lower prices on competing channels.

To illustrate this channel in the framework of our model, assume that the OTA publicly commits itself before hotels set prices to maximally worsen the ranking for the hotel with the higher discount on the direct channel.[35] Formally, this means

$$r = \begin{cases} 0 & \text{if } \left(p_L^{\mathrm{OTA}} - p_L^{\mathrm{DIR}}\right) > \max\left\{0,\ \left(p_R^{\mathrm{OTA}} - p_R^{\mathrm{DIR}}\right)\right\}, \\ \frac{1}{2} & \text{if } \left(p_L^{\mathrm{OTA}} - p_L^{\mathrm{DIR}}\right) = \left(p_R^{\mathrm{OTA}} - p_R^{\mathrm{DIR}}\right), \\ 1 & \text{if } \left(p_R^{\mathrm{OTA}} - p_R^{\mathrm{DIR}}\right) > \max\left\{0,\ \left(p_L^{\mathrm{OTA}} - p_L^{\mathrm{DIR}}\right)\right\}. \end{cases} \tag{11}$$

Although this ranking policy might not raise the OTA's profit in the short-term, it can discipline hotels and thereby lead to much less aggressive prices elsewhere. This can ultimately raise the OTA's profits even more than the standard profit maximization, which takes the hotel prices as given.

To illustrate matters further, assume that $\alpha = 1$ and $\phi_L(0) = \phi_R(1) = 0$. This yields a hotel profit of zero if one hotel discounts its direct price more than the other hotel (see equation (6)).

Proposition 4. *Assume that $\alpha = 1$ and $\phi_L(0) = \phi_R(1) = 0$, and that the OTA commits itself to the drastic ranking policy described in* (11) *before hotels set prices. As a consequence, no hotel will discount its direct price in equilibrium. This has the same effect as a price parity clause.*

Proof. For $\left(p_L^{\mathrm{OTA}} - p_L^{\mathrm{DIR}}\right) > \max\left\{0,\ \left(p_R^{\mathrm{OTA}} - p_R^{\mathrm{DIR}}\right)\right\}$, hotel L makes zero profits. As a consequence, it must be profit increasing for L to reduce the discount to the level $p_R^{\mathrm{OTA}} - p_R^{\mathrm{DIR}}$, as this yields a positive profit. An even better adjustment is to reduce the discount slightly further, as then the ranking changes from $r = \frac{1}{2}$ to $r = 1$, which implies a discrete increase of ϕ. However, now R will react in the same way until $p_i^{\mathrm{DIR}} \geq p_i^{\mathrm{OTA}}$ for $i \in \{L, R\}$. □

This extreme case is chosen to make the optimization simple and illustrative. Apart from a strict enforcement of a parity clause, we conjecture that a more realistic parameterization is much less extreme. However, the rationale that a policy to react rather drastically to lower prices elsewhere can be attractive for an OTA also applies to less extreme cases. In the symmetric equilibrium, the OTA does not need to bias its ranking and therefore the

[35] In reality, such a commitment does not need to be a public announcement. It may be sufficient that the OTA acts in this way and hotels learn about the behavior over time.

ranking still has value for consumers. In the case of symmetric hotels, only in states of disequilibrium – which may be short term and might only affect a few search results – does the ranking need to be biased.

It is noteworthy that a distinction between *standard profit maximization* and a *direct conditioning* of the ranking on prices on competing channels finds support in the commitments of Booking.com toward various European competition authorities in 2015. These state that "Booking.com's ranking algorithm will not take into account directly whether an accommodation refuses to enter into or does not comply with Price Parity [...]."[36] This indicates that for competition authorities such a policy was a cause for concern. At the same time, an "indirect" conditioning on the prices of other channels (as it occurs in the case of the profit maximization previously modeled) is not explicitly ruled out.

Moreover, under the headline "Understand and help impact your visibility," Expedia states in a text addressed to hoteliers that the hotel's visibility depends on "past behavior" and, among other things, "a measure of how well you partner with Expedia Group, that takes into account [...] your room and rate competitiveness on Expedia Group site."[37] This suggests to us a possibly more direct conditioning of the ranking on prices at competing channels than just through the booking likelihood as modeled above. We have not spotted such a statement on the Booking.com website. In view of the above commitments, this is not surprising.

In general, it appears to be sufficient that an OTA employs a certain ranking policy and hotels learn about the behavior over time. For instance, a report by various European competition authorities states that "many hotels mentioned measures taken by OTAs to 'penalize' unwanted behavior by hotels – such as price and/or availability differentiation – without relying on parity clauses."[38]

4.8 Ranking bias and consumer welfare

We have focused on symmetric hotels and hotel-symmetric equilibria. This typical modeling approach simplifies the analysis, particularly of the different price levels, and is common in the literature.[39] A consequence of the symmetry among hotels is that their prices on each channel are the same so that the OTA is better off not biasing the ranking

[36] See section 4 of the Booking.com commitments to the Swedish competition authority of April 2015 (last accessed August 31, 2018).

[37] See Expedia's website Visibility (last accessed August 31, 2018).

[38] The report is based on a survey among hoteliers. See section 8 for references and more details.

[39] Other articles that deal with consumer steering and also consider symmetric equilibria (in which typically no biasing takes place) include Raskovich (2007); Inderst and Ottaviani (2012); Hunold and Muthers (2017); Shen and Wright (2018).

in equilibrium. Nevertheless, biasing the ranking in off-equilibrium states is crucial for the pricing incentives of the hotels and thus the resulting prices.

Beyond the theoretical workhorse of symmetric players and a symmetric equilibrium, there are many reasons why one might empirically observe biased rankings. First of all, the hotels might well be heterogeneous. For instance, the direct channel may be more important for the one hotel than for the other hotel. One can easily incorporate this in the above model and obtain biased rankings in equilibrium. Similarly, one hotel might be active at one OTA only, while another hotel is active at two OTAs. This can lead to different optimal prices across channels in equilibrium, even when hotels take the influence of their pricing on their ranking at an OTA into account.

Moreover, an economic system is not necessarily in equilibrium at a particular point in time. Instead, there may be ongoing adjustment processes. This can particularly be the case in markets where many firms are active and there are frequent demand changes, as is often plausible for hotel markets.

In the present model, consumer welfare depends both on the price levels and on the ranking. Starting from a symmetric outcome, several comparative statics are immediate:

- For a given ranking (r), consumers benefit from a price decrease by any hotel on any channel.
- Given symmetric hotel valuations and prices ($p_L^k = p_R^k$, $k = \text{DIR}, \text{OTA}$), a reduction in ranking bias (that is moving closer to $r = 1/2$) increases consumer welfare as consumers realize more optimal matches.
- Given symmetric hotel quality and prices at the OTA, but a lower direct price of hotel L ($p_L^{\text{DIR}} < p_R^{\text{DIR}}$), the corresponding ranking that is optimal for the OTA tends to reduce consumer welfare compared to the best match ranking because it favors hotel R (see Proposition 1).

One may ask whether biased rankings can also benefit consumers. Within our model framework, the symmetric equilibrium with price parity indeed features lower consumer prices if, off-equilibrium, the OTA were to bias the ranking in favor of a hotel with lower prices (see Proposition 3). However, at the same time the prices increase in the commission rates – which we have kept exogenous. The main result of various academic studies with competing OTAs is that PPCs can restrict OTA competition and lead to excessive commission rates and hotel prices – even above the monopoly level (see section 2). This has also been the main theory of harm of various competition authorities and has led to the prohibition of PPCs in various countries (see section 3). We have shown that the ranking can have effects similar to PPCs in reducing the incentives of hotels to charge lower prices on other channels. We conjecture that in a richer model with competing OTAs, this can then also lead to higher commission rates and prices (if there is a drastic reaction equivalent to a PPC, as shown in subsection 4.7, the insights of the literature on PPCs should apply). Notwithstanding, one may argue that a profit-maximizing ranking involving a "bias" may be necessary for the OTA to be profitable. We discuss this issue further in the concluding section 8.

5 Empirical strategy

5.1 Predictions

To review and compare the different theoretical predictions, let us write the ranking of hotel i at an OTA in response to a search request s as the following function:

$$Ranking_{i,s}^{\text{OTA}} = f\left(v_i,\, p_i^{\text{OTA}}, q_i^{\text{OTA}}, p_i^{\text{COM}}, c_i\right), \tag{12}$$

and let a higher value mean that the ranking is better (closer to the top of the search results).

Best match ranking. The ranking should be a positive function of the hotel's gross match value v_i for the average consumer ("hotel quality"), which should increase in the number of stars, the user rating, free breakfast, and so on. Given the gross value v_i, the price p_i^{OTA} should negatively affect the ranking as, other things equal, a higher price should mean a lower net match value for the average consumer. Moreover, the booking likelihood q_i^{OTA} (conditional on the true match value and price) should not influence the ranking. Likewise, the commission rate c_i and the price p_i^{COM} for the hotel offer on other channels should not matter for the best match ranking of the OTA's hotel offers.

Standard OTA profit maximization. The booking likelihood should also have a positive influence now as the OTA generates income on a per-booking basis. For the same reason, the commission rate should have a positive influence. Whether the hotel price at this OTA enters negatively or positively is not as obvious. On the one hand, a higher price means a lower net match value for the consumer, which reduces the booking likelihood. However, conditional on observing the booking likelihood (which is possibly a function of the prices), the hotel price at the OTA should have no (direct) negative effect. On the other hand, the commission is the product of the commission rate and the OTA price, which speaks for a positive influence of the OTA price on the ranking.

When the OTA maximizes its profit for given prices, one could argue that the gross match value v_i does not matter anymore. Moreover, conditional on the booking likelihood, the hotel's prices on competing channels should have no influence. However, it is clear that the match value should positively affect the booking likelihood, whereas lower prices on competing channels should decrease the booking likelihood.

Direct conditioning of the ranking on prices of other channels. In this case, the OTA directly conditions the ranking on price differentials in order to discipline the hotel. Different from the standard profit maximization, for a given booking likelihood, higher hotel prices on competing channels should now have an effect and lead to a worse ranking position.

5.2 Methodology

Our focus is to investigate whether an OTA's ranking of hotels depends on how the hotel's prices at other OTAs or on the hotel's direct online sales channel relate to the hotel's price at this OTA. For this we estimate equation 12 using two complementary approaches: the non-linear, rank-ordered logit and the linear ordinary least squares (OLS) with hotel fixed effects.

Rank-ordered logit. In the rank-ordered logit model, we assume that the OTA ranking can be thought of as the result from assigning every hotel a score based on a linear equation of the form

$$Score_{s,i} = \beta' X_{s,i} + \varepsilon_{s,i}, \tag{13}$$

where s is the set of offers that are responsive to the user's search request and i is the hotel. The vector of explanatory variables $X_{s,i}$ includes observable variables as in equation 12 that influence the score. For a technical description of the estimator, see Beggs et al. (1981).[40] In short, the estimator estimates the parameters of the scoring function by comparing selected to non-selected alternatives like a standard logit. Additionally, it accounts for the ranking of the alternatives by comparing better to worse ranking

[40] It is also known as the Placket-Luce model (Marden, 1995), the exploded logit model (Punj and Staelin, 1978), and the choice-based method of conjoint analysis (Hair et al., 2010).

positions. The best-ranked hotel should have a higher score than the second-best ranked hotel, and the latter a higher score than the third-best ranked hotel. Under the assumption that the error terms $\varepsilon_{s,i}$ are independently distributed and follow an extreme value type I distribution, the probability that a hotel is ranked first can be written in the multinomial logit form

$$\pi_{i,s} = Pr\{Score_{s,1} > max(Score_{s,2}, ..., Score_{s,n})\} = \frac{exp(Score_{s,i})}{\sum_{j=1}^{n} exp(Score_{s,j})}. \tag{14}$$

The probability of observing a specific ranking can be written as the product of the terms in expression 14. This is consistent with sequential decision-making where the algorithm first chooses the best-ranked hotel, and then the best-ranked hotel among the rest, and so on. The rank-ordered logit can therefore also be thought of as a conditional logit for each subset of rankings. We estimate the parameters with maximum likelihood.[41] A higher coefficient implies a higher score, that is, a better ranking position, and thus a better visibility for consumers. As in all non-linear models, without further assumptions the estimated coefficients are not identified in their scale, but only their direction.

Linear fixed effect model. As including hotel fixed effects is computationally too burdensome in the rank-ordered logit model, we additionally estimate linear models of the form

$$Ranking_{i,s} = \beta' X_{i,s} + \gamma' Z_s + \xi_i + \varepsilon_{i,s} \tag{15}$$

where Z_s are additional time-varying controls that do not vary for the hotels of a search result (for instance, controls for local seasonality and booking-month fixed effects) and ξ_i are hotel fixed effects. The latter allow us to remove all unobserved time-constant heterogeneity between hotels and focus on whether variation in prices for a given hotel changes that hotel's rank. Note that a better ranking because of a higher score corresponds to a lower number. Therefore, we invert the dependent variable in the regressions by subtracting it from the number of entries in the search result to ease comparison.

Mapping from predictions to observed variables. Our dependent variable in both approaches is the observed ranking position of each hotel offer in an OTA's search results. The set of explanatory variables consists of three parts. First, we control for the gross

[41] To decrease the computational burden, we treat rankings below the 10th position as incomplete. This ensures that the estimates are more precise for the most relevant top positions. However, the results are qualitatively the same for other thresholds (e.g., the 5th or 20th position).

match value of the hotel by the average rating of the hotel on the OTA as a measure of consumer satisfaction, further variables accounting for quality (such as an included breakfast and payment terms) and time-constant hotel characteristics, such as the hotel size (number of rooms), the type of the hotel (hotel chain vs. independent hotel), the proximity to public transport, the number of stars and, of course, the price at the OTA. Second, the OTA might rank hotel offers higher if they are more profitable. Besides the price at the OTA, we therefore include measures indicating whether the hotel pays a higher commission rate (i.e., elevated fees in order to improve the ranking) and information about the number of past bookings on the OTA website itself (as reported by the OTA) to proxy for the past conversion rate. Third, to see whether the OTA directly conditions rankings on prices of other channels (for a given booking likelihood), we include the availability of the same offer on other channels, as well as the respective relative price difference. We discuss later to what extent we fully capture the booking likelihood with the other variables (including past bookings).

5.3 Identification

The key challenge is to identify whether there is a causal link between the price charged by a hotel on a competing sales channel and the rank of that hotel at a particular OTA. There are several issues to address.

Heterogeneity between hotels. If for an OTA the absolute commission income is higher – other things equal – it has an incentive to rank the hotel better. A high commission income can be either due to a higher commission rate paid by the hotel or a higher price at a given rate. If for some reason hotels with a lower absolute commission typically discount the direct channel price more, we could get a spurious negative correlation between a good ranking position and the OTA's price markup relative to the direct channel.

We deal with this unobserved heterogeneity across hotels in several ways. One is that we include indicators of the commission rate. These are promotional activities of hotels as reported by the OTA (which are available in the case of Expedia) and information on whether hotels paid elevated fees in order to improve their rankings (as reflected in the Preferred Partner Program of Booking.com). We also include the price level at the OTA as a control variable, as this has a direct effect on the absolute commission income of the OTA. Finally, in the linear regressions we remove time-constant unobserved heterogeneity between hotels by including hotel fixed effects.

Changes in commission rates. In addition to unobserved time-constant heterogeneity across hotels, there might also be unobserved changes in the commission rate that applies to the hotel. These changes might affect both the pricing across channels and the ranking of the hotel at an OTA. In particular, as a higher commission rate implies a higher distribution cost for the hotel when selling through the OTA, the hotel might also increase the price it charges at the OTA. Indirectly, it might also affect the price the hotel charges on its website. An obvious reason for this is a price parity clause.

We control for changes in the price charged by the hotel at the OTA under investigation, as this price should be mainly affected by an increase of the distribution costs at this OTA. Our main approach is that we regress the hotel's rank at an OTA on the (absolute) OTA price and the relative difference of the price on other channels and that OTA's price. Relative differences allow the comparison between hotels with different absolute average price levels and decrease the potential spurious correlation between the price and the deviation. As robustness checks, we also consider alternative specifications, in particular the absolute difference of prices across channels or indicators for the level of deviations.

In addition, economic reasoning helps us to deal with this potential identification issue. An identification problem arises only if, in response to an increase of the commission at an OTA, the hotel increases the prices on competing – and now relatively cheaper channels – more than at the OTA (and analogously for a cut). In this case there would be a spurious positive correlation between a better ranking position and a higher (= less aggressive) price on competing sales channels. However, in response to an increase of the commission rate at an OTA a larger price increase on competing channels than at the OTA does not seem likely. In particular, this does not occur in our theoretical model presented above, and

i. if there is a parity such that all prices, and changes therein, are equal, and

ii. if for some reason the hotel always sets the direct price equal to a fraction of the OTA price (as long as it is lower).[42]

Regarding (i), note that if in response to an increase of the hotel's marginal costs at the OTA all prices increase uniformly, we can control for this with the OTA's price. In particular, the difference between the OTA price and competing prices would be unchanged.

Regarding (ii), consider the case that the original OTA price was 100, and the direct price is set at 90 percent of that. If the hotel now sets the OTA price to 110, the direct price

[42] For instance, hoteliers indicated in our phone interviews that they would set the direct price equal to the OTA price minus half the OTA commission rate.

would be 99, such that the difference would increase from 10 to 11. As a consequence, the absolute markup of an OTA price over a direct price would have increased. This would yield a positive correlation between the OTA markup and attractive ranking, which is contrary to our hypothesis. As a consequence, if we nevertheless find results in line with our hypothesis, this should not be driven by the above described mechanism.

Reverse causality. We argue that there is a causal link between the price charged by a hotel on a competing sales channel and the rank of that hotel at a particular OTA. Another concern might be that the link is reversed – that means a hotel facing a bad ranking position might decide to offer a lower price only at competing sales channels. Although it is far from clear that this would be the optimal response of a profit-maximizing hotel, we nevertheless address this concern econometrically by analyzing the timing of the deviations of hotels and the changes in the subsequent ranking to show that it is indeed the OTA reacting to a change in behavior and not the reverse.

6 Data

6.1 Data collection

For our empirical analysis we primarily need data on 1) hotel *rankings* of OTAs, and 2) *prices* that hotels post on different channels. As controls, we need data on the characteristics of hotels that can explain their attractiveness to consumers, as well as data on other determinants of the profitability of a hotel offer for the OTA (such as promotions and indicators of different commission rates). We collected hotel characteristics and the default rankings of hotels directly from the major OTAs Booking.com and Expedia (denoted by "Our top picks" and "Recommended," respectively). We obtained the prices of hotels on channels other than Booking.com and Expedia from Kayak.[43] Kayak is a travel meta-search engine that collects information from various online channels such as the OTAs Booking.com, Expedia, and the hotels' direct online channel. We understand that Kayak derives revenues from advertising placements on its websites, as well as from mobile apps and from sending referrals to travel service providers, such as hotels and OTAs.[44] Kayak displays information on availability and prices and uses interfaces with the OTAs that allow this information to be retrieved from them while hotels can provide

[43] We use the German edition of the respective websites as price discrimination between different countries is not very frequent in the hotel industry, as discussed in section 3.

[44] See page 2 of Booking Holdings Annual Report 2017 (last accessed August 31, 2018).

information on their direct channel prices by making use of their own booking engine or a third-party booking engine provider.[45]

From the three websites Booking.com, Expedia, and Kayak, we collected data between mid-July 2016 until the end of January 2017. The data collection took place by performing search requests, consisting of a travel destination, the travel dates, the number of travelers and the number of rooms, e.g., two persons looking for one room in Rome for an overnight stay in two weeks from today. The prices[46] are for overnight stays for two persons in one room on the same day, the 7th, and 14th day ahead. We collected search results and hotel prices for 250 travel destinations. These are from 13 different countries and constitute the countries' largest cities as well as popular tourist destinations.[47] As, according to previous studies (Hannak et al., 2014), user-specific rankings and prices seem less relevant on hotel OTA websites, and to decrease the burden for the platforms, we focused on collecting data for an anonymous "average" user.[48] In response to a search request, all three websites list results that are then collected and combined into one database. For Booking.com and Expedia the "default" rankings of the recommended search results are retrieved, without any search refinements.[49] We merge hotels' data across platforms using the textual similarity of the hotel name (after filtering for city and country) and by retrieving the first Google hit when searching for a hotel for each platform.[50] Hotels which we did not find by these techniques we looked up manually. We also manually checked every matched entry in the end for consistency. In theory, if all hotels always posted their room offers on all three websites, one would observe the same set of results. Of course, this is not the case, and in practice there are additional differences because of the different search criteria of each website (i.e., within-municipality or perimeter-around-geographic center). Technical problems during data collection (due to server downtimes or access changes) also led to the non-availability of specific channel prices in some instances. When this happened, we excluded from our database the search results from all three websites.

45 Booking engines are providers that offer the services necessary to connect the hotel to Kayak, such as Fastbooking, Travelclick, and Derbysoft.

46 Prices on all platforms always include the value added tax (VAT) and can include tourism taxes, depending on whether this is mandatory under national or local legislation.

47 See Appendix B for a detailed list.

48 We ran an identical search query on each of the three platforms, with the script starting at the same time from servers located in Germany, which are randomly chosen from a pool of over 100 different IP addresses. We used the same browser profile for every search request, with cookies and browsing history being cleared after each request and a new browser session being created every time.

49 If a hotel had more than one entry in the list shown in response to a search request, we just kept the entry of the hotel with the best rank. We excluded hotels that the OTA presented as being booked out (and hence without a price).

50 To find hotel A in city B on platform www.ota.com, we ran search requests on Google such as "hotel A B site:www.ota.com" and took the first result as the hotel's profile on this OTA.

We amend our dataset with two measures for destination-specific demand. First, to measure the non-availability of rooms or hotels at a given travel destination for Booking.com, we use the information provided at the top of the search results ("City x is y% unavailable for your dates on our site").[51] For Expedia, this information was not available such that we took the average of the non-availability measures from Booking.com and Kayak. Second, we retrieved time series data from Google Trends for our observation period to approximate the tourism demand for hotels in particular destinations. The data comprises the aggregated search volume of specific queries on Google over time.[52] Furthermore, to have a control variable for the booking likelihood (conversion rate), we include the number of past bookings as reported in the search results by the platforms.[53]

6.2 Descriptive statistics

We create separate datasets for Booking.com and Expedia. Each dataset includes the search results of the OTA, as well as the hotels' prices on other channels, which we took from both Kayak and the other OTA website directly.

Table 1: Characteristics of hotels on Booking.com

	Mean	Std. Dev.	Min.	Max.
Hotel Stars	3.22	0.79	1.00	5.00
(Average) User Rating	8.14	0.72	3.30	10.00
(Average) Number of User Ratings	555.13	713.65	3.00	12350.87
Number of Rooms	61.08	71.55	1.00	1590.00
Chain Hotel	0.27	0.44	0.00	1.00
Public Transportation	0.68	0.47	0.00	1.00
Preferred Partner (at least once)	0.42	0.49	0.00	1.00
Observations	22984			

Hotels in the dataset. The dataset of Booking.com contains almost 23,000 hotels (Table 1). The average hotel in the dataset has around three out of five stars and received an average of 8.14 points from 10 in roughly 555 ratings from past hotel consumers. There is also information on how many rooms a hotel has (around 61 on average), whether the

51 Anecdotally, we heard from hoteliers that they themselves use this measure as a demand predictor.

52 Similar data has been used as a predictor of actual tourism data in other studies. The collection and validation of this data is explained in Online Appendix III of Hunold et al. (2018).

53 As in many jurisdictions it is illegal to base ads on false information (see section 5 of the Act Against Unfair Competition for Germany), we are confident that these website data, although addressing consumers, are informative of the actual number of bookings (last accessed August 31, 2018).

hotel is part of a chain (27 percent), and whether the hotel is located close to public transportation (68 percent).[54] Forty-two percent registered for the Preferred Partner Program in the observation period or before, which essentially means that the hotel paid an extra commission to the platform in order to get a better ranking.[55]

Table 2: Characteristics of hotels on Expedia

	Mean	Std. Dev.	Min.	Max.
Hotel Stars	3.13	0.90	0.00	5.00
(Average) User Rating	3.98	0.51	1.20	5.00
(Average) Number of User Ratings	194.68	341.26	1.00	7682.22
Number of Rooms	68.03	81.90	0.00	2300.00
Chain Hotel	0.31	0.46	0.00	1.00
Sponsored Ad (at least once)	0.05	0.22	0.00	1.00
Observations	18977			

The composition of hotels in the Expedia dataset differs only slightly (Table 2) from that of Booking.com. The average number of stars is also about three, the average rating is close to four out of five points, which is comparable to the Booking.com average of eight out of 10 points. Hotels on Expedia have fewer ratings on average (about 195 versus 555) than hotels on Booking.com, more rooms (68 versus 61), and are more often part of a chain (31 versus 27 percent). Five percent of the hotels on Expedia paid at least once for a sponsored advertisement.

Ranked hotel offers. The unit of observation is a hotel room offer on a certain search date for a specific travel period. Table 3 summarizes the search results of Booking.com. In total, the Booking.com dataset consists of more than two million observations. The number of listed hotel offers in response to a search query varies across travel destinations. The average ranking of a search result is at 108 and the maximum number of ranked entries is 1,599. The average price of a hotel posted on Booking.com is 130.77 EUR. We excluded prices outside the range of 15 EUR to 2,000 EUR.[56] For offers on Booking.com, we also have information on whether the breakfast is included (it is in 46 percent of offers) and whether a cancellation is free of charge (31 percent). On average, Booking.com reports

[54] We retrieved the information about the number of rooms directly from the websites of Booking.com and Expedia. If it was not available, we excluded those observations amounting to less than 1 percent. The information about chain membership was collected from Booking.com. For hotels on Expedia that we did not find on Booking.com, we manually identified the remaining chains.

[55] See page 20 of the Report on the Monitoring Exercise carried out in the Online Hotel Booking Sector by EU Competition Authorities in 2016 (last accessed August 31, 2018).

[56] One might still wonder whether these extreme cases emerge from data errors. We checked the offers of the respective accommodations manually and found that this was not the case.

that a hotel has been booked 13 times in the last three days. The share of sold out rooms is, on average, 46 percent and the average Google Trends search volume measure is at 72.85 out of 100, which is the maximum.

Table 3: Descriptive statistics for the search results of Booking.com

	Mean	Std. Dev.	Min.	Max.
Ranking Position	108.09	162.89	1.00	1599.00
Price at Booking	130.77	96.70	15.00	2000.00
Preferred Partner	0.40	0.49	0.00	1.00
Hotel Stars	3.30	0.82	1.00	5.00
User Rating at Booking	8.10	0.71	3.30	10.00
Number of User Ratings	789.27	893.56	3.00	14757.00
Number of Rooms	76.35	90.23	1.00	1590.00
Chain Hotel	0.32	0.47	0.00	1.00
Breakfast	0.46	0.50	0.00	1.00
Free Cancellation	0.31	0.46	0.00	1.00
Public Transportation	0.88	0.32	0.00	1.00
# Bookings Last 3 Days	13.26	22.38	0.00	1266.00
% Hotels Sold Out	0.46	0.29	0.00	1.00
Google Trends	72.85	18.02	4.00	100.00
Kayak Dummy	0.67	0.47	0.00	1.00
Direct Dummy	0.16	0.37	0.00	1.00
Direct Lower	0.04	0.20	0.00	1.00
(Booking-Direct)/Direct	0.00	0.04	-0.50	0.99
Expedia Dummy	0.70	0.46	0.00	1.00
Expedia Lower	0.11	0.31	0.00	1.00
(Booking-Expedia)/Expedia	0.00	0.08	-0.50	1.00
Observations	2307513			

We also report whether we can find a specific hotel room offer also at Kayak and, in particular, whether Kayak lists a direct channel offer. We show whether and how the respective direct channel price differs from the price posted on Booking.com. We also find the hotel offer on Kayak the same day for 67 percent of the Booking.com observations. For 16 percent of all observations, we note that the hotel offers a direct price on Kayak.

Before going into more detail about the deviations, we first conclude this subsection by describing the properties of the Expedia dataset, which differ only slightly (Table 4) from those of the Booking.com dataset. The average posted prices on Expedia are slightly lower (121 EUR vs. 131 EUR). Our understanding of a sponsored advertisement on Expedia is that hotels pay elevated commissions to be additionally shown in designated positions in the search results, while they still appear in the generic search results. In our dataset, we include only the generic search results, but measure whether, in the same search result, the hotel was additionally shown in the form of a sponsored advertisement. This is the case for only 1 percent of the observations. As an additional characteristic, we record whether the payment can be made on arrival ("Pay Later"), which is the case for 22 percent of the offers. Hotels listed on Expedia are more often found on Kayak according to our data (compared with hotels listed on Booking.com). The remaining statistics with respect to the direct channel are comparable with those of Booking.com. For every offer on Expedia, we identify a corresponding direct channel listing on Kayak for 18 percent and a listing on Booking.com for 63 percent of observations.

Table 4: Descriptive statistics for the search results of Expedia

	Mean	Std. Dev.	Min.	Max.
Ranking Position	133.61	245.18	1.00	2092.00
Price at Expedia	121.20	89.43	16.00	2000.00
Sponsored Ad	0.01	0.12	0.00	1.00
Hotel Stars	3.26	0.86	0.00	5.00
User Rating at Expedia	3.95	0.49	1.10	5.00
Number of User Ratings	240.00	424.27	1.00	10246.00
Number of Rooms	82.98	90.35	0.00	2300.00
Chain Hotel	0.39	0.49	0.00	1.00
Pay Later	0.22	0.42	0.00	1.00
# Bookings Last 2 Days	1.00	3.59	0.00	276.00
% Hotels Sold Out	0.61	0.15	0.00	0.99
Google Trends	72.58	17.07	4.00	100.00
Kayak Dummy	0.68	0.47	0.00	1.00
Direct Dummy	0.18	0.38	0.00	1.00
Direct Lower	0.05	0.21	0.00	1.00
(Expedia-Direct)/Direct	0.00	0.04	-0.50	1.00
Booking Dummy	0.63	0.48	0.00	1.00
Booking Lower	0.13	0.34	0.00	1.00
(Expedia-Booking)/Booking	0.00	0.08	-0.50	1.00
Observations	2457164			

6.3 Price differences across channels

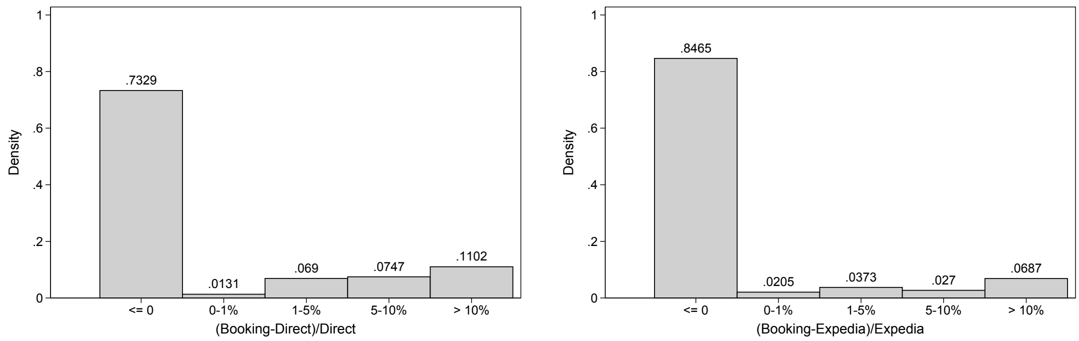
In our analysis we focus on the differences in the prices of the two major OTAs, Booking.com and Expedia, as well as the hotel website. Given the significance of these channels, price differences among these should be of primary interest for the OTAs and hotels.

Figure 2 depicts the differences between the Booking.com price and the direct price (left panel) as well as the Expedia price (right panel).[57] In about 27 percent of the observations, the direct price is below the Booking.com price (cf. Figure 2). For these cases the average

[57] For our analyses, we have ignored relative differences of less than -50 percent and more than 100 percent as they might stem from the misclassifications of hotel offers. However, these cases only account for a negligible fraction of the dataset.

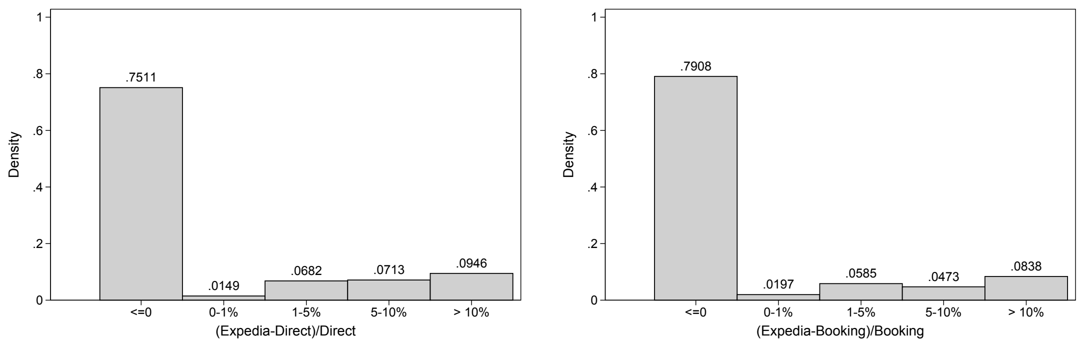
difference amounts to about 12 percent. In about 15 percent of the observations the Expedia price is below the Booking.com price, and the average difference in these cases is 12 percent. Correspondingly, negative deviations are less frequent between OTAs.

Figure 2: Extent of downward deviations of the direct channel on Kayak and Expedia toward Booking.com



We observe a similar pattern in the deviations toward Expedia in Figure 3. While there is a substantial share of negative deviations of the direct channel toward Expedia, negative deviations on Booking.com toward Expedia are less frequent.

Figure 3: Extent of downward deviations of the direct channel on Kayak and Booking.com toward Expedia



7 Estimation results

7.1 Main regressions

We now present our regression of the hotels' ranking position on the price differences across channels and the control variables. A positive coefficient indicates that the explanatory variable affects the ranking position positively, both in the rank-ordered logit and the linear estimation with fixed effects.

We estimate four different specifications for each dataset. The baseline specification (1) includes only the hotel characteristics as well as the price and booking history at Booking.com. Specification (2) additionally contains an indicator of whether we could identify the hotel offer on the direct channel ("Channel Availability") and – if this is the case – also the difference in the Booking.com price and the direct price.[58] Specification (3) analogously contains the information on availability at Expedia and the respective price difference. Specification (4) combines the information added in specifications (2) and (3).

Table 5: Determinants of Booking.com ranking – Rank-ordered logit results

	(1)	(2)	(3)	(4)
Price at Booking/100	0.10***	0.11***	0.11***	0.11***
Preferred Partner	1.58***	1.56***	1.56***	1.56***
Hotel Stars	0.22***	0.19***	0.19***	0.19***
User Rating at Booking	0.12***	0.12***	0.12***	0.12***
Number of User Ratings/100	0.02***	0.02***	0.02***	0.02***
Number of Rooms/100	0.08***	0.08***	0.08***	0.08***
Chain Hotel	0.19***	0.12***	0.13***	0.12***
Breakfast	0.10***	0.13***	0.13***	0.13***
Free Cancellation	0.07***	0.07***	0.07***	0.07***
Public Transportation	0.33***	0.29***	0.29***	0.29***
# Bookings Last 3 Days	0.01***	0.01***	0.01***	0.01***
(Booking-Direct)/Direct		-0.44***		-0.37***
(Booking-Expedia)/Expedia			-0.25***	-0.20***
Channel Availability	No	Yes	Yes	Yes
Observations	2307513	2307513	2307513	2307513
Pseudo R^2	0.11	0.12	0.12	0.12
Log-Likelihood	-1017972.67	-1008722.92	-1008732.13	-1008686.70

Heteroscedasticity-robust standard errors not reported.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Regressions of the ranking with rank-ordered logit. We first present the results of the rank-ordered logit estimator and start with the ranking of Booking.com (Table 5). In specification (1) one can see that hotels with more stars, higher consumer ratings, a higher number of ratings, close to public transport, breakfast included, and free cancellation are ranked higher at Booking.com. These findings are consistent with the expectation that OTAs rank hotels of higher quality better. Moreover, hotels with more rooms and chain

[58] The results for both Booking.com and Expedia are robust to only including hotels for which we observe the prices on the direct channel on Kayak and Expedia and are thus not driven by the absence of deviations.

hotels are ranked better. It is not as straightforward that – other things equal – larger hotels and chain hotels are ranked higher. It might be that these hotels have more offers available at the OTA, are thus booked more often, and therefore ranked better by the OTA. In addition, hotels with more bookings at Booking.com in the last three days have a better ranking. Finally, hotel offers with a higher price on Booking.com have a better rank. These results persist across all four specifications.

We comment now on specification (4), which is consistent with the results of specifications (2) and (3). One can see that if the hotel's direct channel price or the price at Expedia is lower than the price at Booking.com, this leads to a worse ranking position. For the non-linear, rank-ordered logit estimates, we can only interpret the size of the coefficients relative to other coefficients of the regression. We observe that the effect of lower direct channel prices is stronger than the effect of lower Expedia prices. A 10 percent deviation on the direct channel has about the same effect as a decrease in (0.37/10)/0.12=0.308 points (out of 10) in the user rating.

We find similar effects for the Expedia ranking. The main difference is that higher prices on Expedia itself lead, on average, to lower rankings at Expedia. The possibility to pay on arrival improves the ranking. Apparently, on average, independent hotels also receive a lower ranking. Interestingly, the effect of different prices on Booking.com is about two times larger than the effect of different prices on the direct channel. A 10 percent deviation in the price on the direct channel has a comparable effect to a decrease in (0.40/10)/0.41=0.098 points (out of five) in the user rating.

Table 6: Determinants of Expedia ranking – Rank-ordered logit results

	(1)	(2)	(3)	(4)
Price at Expedia/100	-0.16***	-0.18***	-0.18***	-0.17***
Sponsored Ad	0.55***	0.57***	0.57***	0.57***
Hotel Stars	0.46***	0.40***	0.40***	0.40***
User Rating at Expedia	0.34***	0.41***	0.41***	0.41***
Number of User Ratings/100	0.05***	0.04***	0.04***	0.04***
Number of Rooms/100	-0.01***	-0.00	-0.00	-0.00
Chain Hotel	-0.13***	-0.12***	-0.13***	-0.13***
Pay Later	0.02***	0.02***	0.03***	0.03***
# Bookings Last 2 Days	0.04***	0.04***	0.04***	0.04***
(Expedia-Direct)/Direct		-0.62***		-0.40***
(Expedia-Booking)/Booking			-0.84***	-0.82***
Channel Availability	No	Yes	Yes	Yes
Observations	2457164	2457164	2457164	2457164
Pseudo R^2	0.04	0.06	0.06	0.06
Log-Likelihood	-1025005.56	-996150.79	-995504.57	-995474.98

Heteroscedasticity-robust standard errors not reported.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Summary. In the rank-ordered logit regressions, lower prices on the direct channel or at the other OTA lead to worse ranking positions at both Booking.com and Expedia.

Linear regressions of the ranking with hotel fixed effects. To rule out that our findings are driven by unobserved heterogeneity between hotels, we estimate a linear model with hotel fixed effects. Consequently, all time-constant hotel characteristics have been excluded.[59] To better account for time-varying effects, we include measures for local seasonality and booking-month fixed effects.

Table 7: Determinants of Booking.com's ranking – OLS with hotel fixed effects

	(1)	(2)	(3)	(4)
Price at Booking/100	4.73***	4.84***	4.88***	4.92***
Preferred Partner	74.80***	74.81***	74.81***	74.80***
User Rating at Booking	-3.50	-3.85*	-3.89*	-3.88*
# Bookings Last 3 Days	0.01*	0.01**	0.01**	0.01**
% Hotels Sold Out	40.66***	40.58***	40.56***	40.56***
Google Trends	-0.07***	-0.07***	-0.07***	-0.07***
(Booking-Direct)/Direct		-11.94***		-9.76***
(Booking-Expedia)/Expedia			-8.74***	-7.93***
Channel Availability	No	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Observations	2307513	2307513	2307513	2307513
R^2	0.91	0.91	0.91	0.91

Heteroscedasticity-robust standard errors not reported.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7 shows the results for Booking.com. Higher prices and higher commission rates (in the form of participating in the Preferred Partner Program) lead to a better ranking. An increase in the user rating leads to a worse ranking, while the number of bookings at the OTA in the last three days continues to lead to a better ranking.[60] As for seasonality, a higher share of sold out rooms and a lower Google Trends search volume is correlated with a better ranking. Most importantly, the relationship between more aggressive pricing on competing channels and a worse ranking position is also significant when restricting attention to changes of the variables within hotels over time. This is comparable for

[59] We also exclude further offer characteristics such as "Breakfast" and "Free Cancellation" for Booking.com and also "Pay Later" for Expedia, as these variables have a very low variation within hotels.

[60] The low variation of the user rating across time can explain this inconclusive result. On Booking.com, the average rating increased by 0.1 points, while on Expedia it worsened by 0.04 points. The 10th and 90th percentile of changes on Booking.com is also -/+ 0.1 points, while on Expedia the corresponding figures are -0.2/+0.1.

the Expedia ranking (Table 8). The results between OTAs differ for the variable Google Trends and the user rating.[61]

Table 8: Determinants of Expedia's ranking – OLS with hotel fixed effects

	(1)	(2)	(3)	(4)
Price at Expedia/100	-12.12***	-11.80***	-11.63***	-11.52***
Sponsored Ad	1.71*	1.71*	1.71*	1.70*
User Rating at Expedia	3.17**	3.12**	3.18***	3.18**
# Bookings Last 2 Days	0.67***	0.67***	0.66***	0.66***
% Hotels Sold Out	140.45***	138.65***	137.87***	137.95***
Google Trends	0.17***	0.17***	0.17***	0.17***
(Expedia-Direct)/Direct		-41.89***		-33.99***
(Expedia-Booking)/Booking			-40.81***	-38.86***
Channel Availability	No	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Observations	2457164	2457164	2457164	2457164
R^2	0.92	0.92	0.92	0.92

Heteroscedasticity-robust standard errors not reported.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Summary. When eliminating time-constant unobserved heterogeneity by including hotel fixed effects, lower prices on the direct channel or at the other OTA still lead to worse ranking positions at both Booking.com and Expedia.

Discussion. The regression results are qualitatively similar across estimation methods (rank-ordered logit and fixed effects) and, with at least one interesting exception, also across OTAs (Booking.com and Expedia). Across all specifications, there is a positive relationship between the measures of hotel quality and the ranking position.[62] This is consistent with all three theories of how the OTA might calculate the ranking.

Being a Preferred Partner at Booking.com typically means that the hotel has a higher commission rate than other hotels. The positive coefficient indicates that hotels paying a higher commission have a better ranking. The same holds for the positive coefficient of "Sponsored Ad" in the Expedia regressions (it also means that the hotel offer is more visible). Taken together, this indicates that both Booking.com and Expedia do not use a pure "best match ranking," but also rank hotels offers with a higher profitability better.

[61] We also find that while the level of customer rating for the same hotel on Booking.com and Expedia are highly correlated, changes do not correlate across platforms (r=.03) in our sample. This might explain the different finding.

[62] There is one exception, which is likely due to a lack of variation across time (see footnote 60).

The positive coefficient for the number of hotel bookings at the OTA in the last days could be taken as an indication that the ranking is not based purely on best match value for consumers (determined in principle by quality measures and price), but also on the booking likelihood and thus the profitability of the hotel offer for the OTA. However, we are aware that our quality measures might not fully capture hotel quality, so that the number of bookings might still be informative of this.

While the coefficient of the own price is positive in the Booking.com regression, it is negative in the Expedia regression. Our theory predicts a negative coefficient of the hotel's price at the OTA in the case of a best match ranking, but – conditional on controlling for the booking likelihood – a positive coefficient in the case of standard profit maximization as a higher price increases the OTA's margin. For the same reason, also in the case of the direct conditioning of the ranking on the prices of other channels and when controlling for price differentials, a positive coefficient of the hotel's price at the OTA is conceivable. However, it might as well be negative if – in a longer term perspective – the OTA also cares about attracting consumers. Taken together, the negative own OTA price coefficient seems to suggest that Expedia accounts for the price with more emphasis on the match value, whereas Booking.com might put more emphasis on a higher margin.[63]

The coefficients of the relative differentials between the OTA price and prices at the other (major) OTA as well as the hotel website are negative and significant across all specifications. This strongly suggests that both OTAs do not employ a pure best match ranking. If one believes that we are able to fully control for the booking likelihood, then this would mean that both OTAs do not just employ standard profit maximization, but condition directly on prices at competing channels. However, although we control for the number of previous bookings, various quality measures and the OTA price, we cannot say with certainty that we fully capture the booking likelihood at the OTA. Nevertheless, our results are at least consistent with a direct conditioning of the ranking on prices at other channels.

7.2 Robustness checks

In the last subsection we showed that controlling for many other factors, OTAs condition their ranking on the hotel's pricing across channels. Following up on the results with hotel fixed effects, in this subsection we first investigate whether this conditioning is different in countries with and without price parity clauses being allowed. Second, we employ

[63] A caveat applies, as one should also keep in mind that we might have imprecise measures of the booking likelihood, which should be affected by the OTA price as well.

different specifications for the price differences across channels. Finally, we study the timing of the relation between ranking and lower prices elsewhere. The respective results are shown in Tables 14, 15, and 16 in Appendix C.

Different PPC regimes across jurisdictions. As outlined in section 3, price parity clauses in the agreements between OTAs and hotels have been prohibited in various countries. Making the ranking decision dependent on the prices of other sales channels may then act as a substitute to PPCs for OTAs to discourage hotels from setting lower prices on other channels. Our dataset covers the period from mid-July 2016 until the end of January 2017, whereas most prohibitions of price parity clauses had already taken place before or after.[64] Therefore, we can only use cross-sectional variation to study differences across regimes. For this, we distinguish between hotels in countries where PPCs were prohibited (Germany and France) and where they were allowed in the analyzed time frame (the remainder).[65] Starting from the regressions in Tables 7 and 8 with hotel fixed effects, we have built an indicator "PPC" for countries with PPCs and included interactions with the relative differences of the OTA price and the other channel's price. The results are displayed in column (1) of Table 14 for Booking.com and Table 15 for Expedia. For both platforms we find that, generally, the relative difference of the OTA price toward the direct channel and the competing OTA price has a negative impact on the ranking in all countries, irrespective of the PPC regime. For Booking.com we see no significant difference in the way in which the platform treats deviations on the direct channel in PPC countries, whereas there is a weakly significant and positive effect for deviations on Expedia, indicating that it is more lenient with deviations on its competitors in countries with PPCs. For Expedia, one can see that the interaction terms are both strongly significant, positive, and high in magnitude. This means that Expedia is much more lenient toward deviations on both the direct channel and Booking.com in PPC countries. However, a statistical test for the difference of coefficients reveals that Expedia still assigns worse ranking positions to deviating hotels in countries with PPCs, but only to a smaller degree. These findings are in line with Proposition 1 that, regardless of whether a PPC is enforced, an OTA always has an incentive to condition rankings on other sales channel prices.

[64] During our observation period, only Austria announced a ban of PPCs, starting January 1, 2017. Italy announced that it planned to ban PPCs, however, the decision only became effective after our observation period. For more information, see also section 3.

[65] To be precise, in Germany Booking.com and HRS (the German OTA incumbent) were prohibited from using any PPCs. Expedia, however, only announced that it would narrow down its PPCs to the direct channel, while the German competition authority had an open case against Expedia on that matter. We conjecture that Expedia did not enforce its narrow PPC with force anymore.

Alternative specifications of deviations. In the main regression analyses we measure deviations on competing sales channels by calculating the relative difference of the OTA price, i.e., "[(OTA Price) - (Other Price)]/(Other Price)." In the following, we discuss the robustness of this approach.

- Taking the relative difference allows us to compare hotels with different absolute average price levels and decreases the potential spurious correlation between the price and the deviation. In column (2) of Table 14 for Booking.com and Table 15 for Expedia we show that we find qualitatively similar results when including the absolute price difference between the OTA and the competing sales channel.
- Based on the prediction that conditioning rankings on the prices of other channels gives hotels incentives to set not only equal, but also lower prices at the OTA, we implicitly assumed that OTAs account for positive and negative deviations in the same way. In column (3) of Tables 14 and 15 we distinguish between upward and downward deviations. We find that an OTA takes both of these deviation types into account.
- Finally, with our specification we assumed that OTAs take all levels of deviation in the same way into account. To test this, we partitioned both positive and negative deviations into whether they were above or below 10 percent and built corresponding indicator variables included in column (4) of Tables 14 and 15. We find for – both positive and negative – deviations above 10 percent that the ranking is significantly altered, whereas for any deviations below 10 percent the significance of the relationship is often weak.[66] This finding makes sense, as small deviations also occur due to data transmission errors, e.g., slow updating, and OTAs therefore might only focus on major deviations.

Timing of the relation between ranking and lower prices elsewhere. As an additional robustness check, we analyze to which extent past and future differences of the OTA price and the prices on other channels explain the ranking of a hotel on the OTA. This also addresses the question of reverse causality in the sense that hotels might react in their pricing to changes in rankings. For this we calculate past and future deviations for several observation days around the date of ranking and include them in our regressions with hotel fixed effects. We restrict the data to include only observations from days which

[66] We also tried other thresholds, such as distinguishing deviations of "zero to below 5 percent," "5 to below 10 percent" and "10 percent and more" but found that the coefficients for the low deviations were equally often not or only weakly significantly different from zero.

are less than one week away plus the number of time periods we are interested in.[67] This means that the seventh lag may not be more than 14 days before. To facilitate this computation, we only make use of search results where the booking date coincides with the travel date (price deviations also occur most frequently for these observations).

- In Table 16, we report the results for Booking.com (1–3) and Expedia rankings (4–6). Specifications (3) and (6) include the results of interest, while specifications (1) and (4) are useful for comparison to specification (4) in Tables 7 and 8. Specifications (2) and (5) are added to show changes in the determinants when the dataset is restricted to observations where the past and future deviations are observable. Additionally, we restrict our sample to hotels which have a direct channel on Kayak and which are also listed on the respective other OTA. As reflected in specification (1) of Table 16, only studying search results for which the booking and the travel date coincide makes the impact of deviating on the direct channel insignificant for the Booking.com rankings.
- For Expedia, we see that the ranking depends on the current deviations and not on future values of deviations (as would be the case if the hotel reacted to a bad ranking by deviating). Depending on the number of periods we control for, we sometimes observe that past deviations also have significant coefficients. This is the case in the current specification with seven periods (deviations on the direct channel in $t-4$ and Booking.com in $t-2$, $t-3$, $t-4$, and $t-5$ are significant). We find similar results for Booking.com, where the respective ranking also depends on the current deviation on Expedia, plus one instance of future deviations (deviation on Booking.com in $t+1$) and past pricing behavior on the direct channel (deviation on the direct channel in $t-7$).

Summary. Hotel rankings are affected by prices on other channels, both in countries with PPCs and countries without PPCs, but possibly more in the latter. We further find that i) the effect of a worse ranking is robust to different specifications of the functional form of the deviations, ii) OTAs take both downward and upward deviations on other channels into account, and iii) OTAs are mostly concerned about substantial deviations with a magnitude of around 10 percent and more. Finally, the strongest correlation between ranking and lower direct prices is contemporaneous. We do not find a systematic pattern of significant correlations between price differences today with rankings in the future.

[67] Note that due to technical reasons we did not obtain data for every day of our observation period. In order to account for outliers and bad sampling, we excluded hotels with less than 30 observations which only make up for a negligible fraction.

7.3 Characteristics of hotels that price differentiate across channels

We now investigate which hotels price differentiate, and under which circumstances they do so. We run a set of descriptive regressions of the form

$$deviation_{i,t} = \beta' X_{i,t} + \gamma' Z_{i,t} + \xi_l + \eta_t + \varepsilon_{i,t}, \tag{16}$$

where the dependent variable $deviation_{i,t}$ is an indicator variable taking the value 1 if the price of hotel i on the OTA (Booking.com or Expedia) is above the price on another channel (the hotel website or the other OTA), and else 0. ξ_l is a fixed effect for the location l where hotel i is located and η_t is a booking-month fixed effect.

To better understand the short-run decision of hotels to undercut prices, we additionally exploit the within-variation by substituting location fixed effects ξ_l with hotel fixed effects ξ_i, yielding

$$deviation_{i,t} = \beta' X_{i,t} + \gamma' Z_{i,t} + \xi_i + \eta_t + \varepsilon_{i,t}. \tag{17}$$

Tables 17 and 18 in Appendix D show the regression results for deviations on the direct channel and Expedia, respectively, toward Booking.com. Specifications (1) to (3) explain cross-sectional differences in hotel behavior while specification (4) employs hotel fixed effects and therefore allows for within-variation interpretation.

In specification (1) one observes that independent hotels are more likely to deviate both on the direct channel and Expedia. However, most of the characteristics affect the deviation decision on the two channels differently. Higher priced and small hotels at Booking.com are more likely to deviate on the direct channel. Conversely, hotels which have joined Booking.com's Preferred Partner Program, have more stars and a worse rating, deviate more likely on Expedia.

In specification (2), we add information on the percentage of rooms sold out on Booking.com in the same city and on hotel-individual demand. We observe that once rooms in a city get scarcer, hotels tend to deviate more on the direct channel but not on Expedia. The more bookings one hotel had in the last days on Booking.com, the less likely it was to deviate on both the direct channel and Expedia. In specification (3), we additionally control for the time horizon between the search and the travel date. One can see that deviations in the direct channel are more likely if searching for a hotel on the day of travel, while for Expedia we then see fewer deviations.

Finally, in specification (4) we add hotel fixed effects. As most variables are time-constant, they drop out in this specification. For the direct channel, we find that hotels deviate more if they have higher current prices, the share of booked hotels increases, and the search and travel date coincide, while they deviate less with prior bookings made on the platform. We do not find any significant effects for the enrollment in the Preferred Partner Program and for increases in the user rating; however, these variables also only have a small variation in the observation period. As for Expedia, deviations are more likely to occur with a higher share of sold out hotels and fewer bookings made on the platform. Additionally, one observes a positive and weakly significant coefficient for the Preferred Partner Program enrollment and deviations are more likely to occur when the distance to the travel date is seven days.

The findings regarding the determinants of deviations on the direct channel toward Expedia are similar, as can be seen in Table 19. The main differences are that elevated commission fees – as reflected by a sponsored advertisement – negatively influence the decision to deviate. However, this result is not confirmed in specification (4). In specifications (1) to (3) we find that the number of user ratings affects the decision to deviate negatively. The findings with respect to deviations on Booking.com toward Expedia (Table 20) are similar, such that independent hotels are more likely to undercut. Additionally, prices are lower on Booking.com compared to Expedia if the travel date is close to the booking date and for hotels that are small and have worse ratings.

Summary. We find that hotels price differentiate in a plausible way. For instance, hotels set the direct channel price lower if demand is high and the travel date is close to the booking date.

8 Conclusion

We address the question of whether OTAs rank hotels worse in their default search results if these charge lower prices at other OTAs or on their own website. Our conceptual starting point is a "best match ranking" that gives a more prominent position, at the top of the list, to the hotel offers with the best match values. This matters because consumers see and book prominent offers more often than less visible offers down in the list of search results (Ursu, 2018). We then demonstrate that it can be profit-maximizing for an OTA to deviate from a best match ranking (we label these rankings *biased*), both when the OTA employs a standard profit maximization strategy and when it directly conditions the ranking on prices at competing channels.

If the OTA employs a standard profit maximization for given hotel prices, it is best off when presenting the hotel with the highest product of hotel price, commission rate, and booking likelihood at the OTA (conversion rate) most prominently. Lower prices of the hotel on competing channels reduce the booking likelihood if consumers that found the hotel at the OTA then consider booking it elsewhere. Consequently, hotels with lower prices on other channels are less visible and are thus less frequently booked at the

OTA. Anticipating this affects the pricing of hotels and can reduce price dispersion across channels.

An OTA may also employ a policy of conditioning its ranking directly on the price differentials of hotels across channels and may in particular rank hotels with a lower price elsewhere worse. A more drastic down-ranking might not even be in the short-term interest of the OTA, as described above. However, such a policy can discipline hotels and thus lead to much less aggressive prices elsewhere. If it is very effective, it can even lead to equal prices across channels like a (successfully enforced) price parity clause. This drastic down-ranking can ultimately raise the OTA's profits even more than the standard profit maximization.

Building on this theory, we have empirically investigated whether the position of a hotel in the search results of Booking.com and Expedia depends on the prices charged by the hotel on other channels. We find that for a given price at an OTA (either Booking.com or Expedia), a lower price at the other OTA or on the hotel's website leads to a worse ranking position. We obtain these results with both a rank-ordered logit model and linear regressions with hotel fixed effects. We find this relationship both in countries with and without PPC clauses. It is also robust to different functional forms of the price differences and is significant for both downward and upward deviations. It is more pronounced for substantial price differences of around 10% and above. When using specifications with lags and leads of the price differences across channels, we do not see a systematic pattern of significant correlations between price differences in the future with rankings today (which could be an indication of reverse causality). Finally, we also find that hotels price differentiate in a plausible way. For instance, hotels set the direct channel price lower if demand is high and the travel date is close to the booking date.

In summary, our results suggest that OTAs make the ranking of their recommended search results dependent on factors that are relevant for the OTA to maximize its profit, but arguably not to maximize the match value of customers. In particular, we obtain the relationship between lower prices on competing channels and a worse ranking when controlling for other factors, such as the number of bookings of the hotel at the OTA in the previous days, which should already be highly correlated with the conversion rate at the OTA. This raises the question of whether OTAs only optimize their recommended ranking with respect to their conversion rate in the short term, or whether they directly condition their ranking on the hotels' prices on competing channels. If one believes that we are able to fully control for the booking likelihood, then this would suggest that both Booking.com and Expedia do not just employ standard profit maximization but condition directly on prices at competing channels. However, although we control for the number of previous bookings, various quality measures, and the OTA price, we cannot say with certainty that we fully capture the booking likelihood at the OTA. Nevertheless, our results are consistent with a direct conditioning of the ranking on prices at other channels. Interestingly, for Expedia we find a larger effect of lower prices on competing channels on the ranking than for Booking.com. This could indicate that Booking.com is more careful because it committed to competition authorities in Europe that "Booking.com's ranking algorithm will not take into account directly whether an accommodation refuses to enter into or does not comply with Price Parity [...]."[68]

Our study complements a hotel survey conducted by various European competition authorities. In their report of April 2017, the authorities state that "for the 21% of respondents that did price differentiate between OTAs, the most frequent reason given was to increase the hotel's visibility on a particular OTA (for example, its display ranking)."[69] According to the report, "many hotels mentioned measures taken by OTAs to 'penalize' unwanted behavior by hotels – such as price and/or availability differentiation – without relying on parity clauses." These measures would include "'downgrading' the hotel (lowering its ranking) in the OTA's search results."[70]

Our analysis cannot provide a definite conclusion on what kind of OTA ranking of hotels is socially optimal. Yet, we do see two potential risks of an OTA's ranking optimization being based on measures such as the prices of competing offers, which by themselves are unrelated to the intrinsic consumer value of the offers presented by the OTA.

One potential risk is that such a ranking optimization can have effects comparable to those of price parity clauses. PPCs have been forbidden as practices that might lead to excessively high commission rates by competition authorities and legislators in various European countries, whereas they are legal in the US (and other parts of the world).[71]

Another potential risk is that frequent deviations of an OTA from a best match ranking can lead to a low search quality for consumers. However, even if one establishes such a relationship with certainty (and we do not claim to do so), the policy implications are not clear. On the one hand, regulating in which order an ordinary retailer should present the available products seems far-fetched. On the other hand, how large online information providers present their information is arguably of relevance. In this context, it is interesting to note the different decisions of competition authorities in the EU and the US with respect to the search bias allegations against Google. Their conclusions range from the European's high fine of €2.42 billion for Google having allegedly biased its

68 See section 4 of the Booking.com commitments to the Swedish competition authority of April 2015 (last accessed August 31, 2018). Booking.com made comparable commitments in France and Italy.

69 See paragraph 10 of the Report on the Monitoring Exercise carried out in the Online Hotel Booking Sector by EU Competition Authorities in 2016 (last accessed August 31, 2018).

70 See paragraph 50 of the report, see footnote 69.

71 See section 3 for details.

organic search results with the objective of placing competitive offers in an unfavorable way[72] to the FTC's statement that changes in Google's search algorithm have a legitimate business justification could plausibly be viewed as an improvement in the overall quality of Google's search results.[73]

In terms of consumer protection, it seems desirable that consumers are made aware of how such rankings, which are called "Recommended" or "Our top picks," are being computed. For instance, consider that consumers know that "recommended" offers do not (necessarily) appear high in the list because they offer the best value for money, but that the offers might possibly be ranked high because they maximize the expected profit of the OTA. Consumers with this awareness might pay less attention to such a ranking and make better choices. In response to this, it might become desirable for OTAs to publicly commit to rank offers in their "recommended" list only with respect to consumers' value for money. This could lead to both a higher search quality and might eliminate the potentially anti-competitive effects of disciplining hotels in their pricing through the ranking of search results.

72 See the press release by the European Commission, *Antitrust: Commission fines Google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service* (last accessed August 31, 2018).

73 See Statement of the Federal Trade Commission Regarding Google's Search Practices In the Matter of Google Inc. FTC File Number 111-0163 January 3, 2013 (last accessed August 31, 2018).

References

Beggs, Steven, Scott Cardell, and Jerry Hausman, "Assessing the Potential Demand for Electric Cars," *Journal of Econometrics*, 1981, *17* (1), 1–19.

Blake, Thomas, Chris Nosko, and Steven Tadelis, "Returns to Consumer Search," in "Proceedings of the 2016 ACM Conference on Economics and Computation - EC '16" ACM Press 2016.

Boik, Andre and Kenneth S. Corts, "The Effects of Platform Most-Favored-Nation Clauses on Competition and Entry," *The Journal of Law and Economics*, 2016, *59* (1), 105–134.

Chen, Yuxin and Song Yao, "Sequential Search With Refinement: Model and Application With Click-Stream Data," *Management Science*, 2016, *63* (12), 4345–4365.

Corniere, Alexandre De and Greg Taylor, "Integration and Search Engine Bias," *The RAND Journal of Economics*, 2014, *45* (3), 576–597.

Edelman, Benjamin G. and Julian Wright, "Price Coherence and Excessive Intermediation," *Quarterly Journal of Economics*, 2015, *130*, 1283–1328.

Ghose, Anindya, Panagiotis G. Ipeirotis, and Beibei Li, "Designing Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowdsourced Content," *Marketing Science*, 2012, *31* (3), 493–520.

_ , _ , and _ , "Examining the Impact of Ranking on Consumer Behavior and Search Engine Revenue," *Management Science*, 2014, *60* (7), 1632–1654.

Hagiu, Andrei and Bruno Jullien, "Why Do Intermediaries Divert Search?," *The RAND Journal of Economics*, 2011, *42* (2), 337–362.

_ and _ , "Search Diversion and Platform Competition," *International Journal of Industrial Organization*, 2014, *33*, 48–60.

Hair, Joseph F., William C. Black, Barry J. Babin, and Rolph E. Anderson, *Multivariate Data Analysis: A Global Perspective*, Upper Saddle River, NJ: Pearson, 2010.

Hannak, Aniko, Gary Soeller, David Lazer, Alan Mislove, and Christo Wilson, "Measuring Price Discrimination and Steering on E-commerce Web Sites," in "Proceedings of the 2014 conference on Internet Measurement Conference" ACM 2014, pp. 305–318.

Hunold, Matthias, "Best Price Clauses: What Policy as Regards Online Platforms?," *Journal of European Competition Law & Practice*, 2017, *8* (2), 119–125.

— **and Johannes Muthers**, "Resale Price Maintenance and Manufacturer Competition for Retail Services," *The RAND Journal of Economics*, 2017, *48* (1), 3–23.

— **, Reinhold Kesler, Ulrich Laitenberger, and Frank Schlütter**, "Evaluation of Best Price Clauses in Online Hotel Bookings," *International Journal of Industrial Organization (forthcoming)*, 2018.

Inderst, Roman and Marco Ottaviani, "Competition Through Commissions and Kickbacks," *The American Economic Review*, 2012, *102* (2), 780–809.

Johansen, Björn O. and Thibaud Vergé, "Platform Price Parity Clauses with Direct Sales," *Working Paper*, 2017.

Johnson, Justin P., "The Agency Model and MFN Clauses," *The Review of Economic Studies*, 2017, *84* (3), 1151–1185.

Koulayev, Sergei, "Search for Differentiated Products: Identification and Estimation," *The RAND Journal of Economics*, 2014, *45* (3), 553–575.

los Santos, Babur De and Sergei Koulayev, "Optimizing Click-Through in Online Rankings with Endogenous Search Refinement," *Marketing Science*, 2017, *36* (4), 542–564.

Lu, Qiang, Yupin Yang, and Ulku Yuksel, "The Impact of a New Online Channel: An Empirical Study," *Annals of Tourism Research*, 2015, *54*, 136 – 155.

Mantovani, Andrea, Claudio A. Piga, and Carlo Reggiani, "The Dynamics of Online Hotel Prices and the EU Booking.com Case," *NET Institute Working Paper No. 17-04*, 2017.

Marden, John I., *Analyzing and Modeling Rank Data, Volume 64 of Monographs on Statistics and Applied Probability*, Chapman & Hall, London, 1995.

Punj, Girish N. and Richard Staelin, "The Choice Process for Graduate Business Cchools," *Journal of Marketing Research*, 1978, *15* (4), 588–598.

Raskovich, Alexander, "Retail Buyer Power Through Steering," *Economics Letters*, 2007, *96* (2), 221–225.

Ronayne, David and Greg Taylor, "Competing Sales Channels," *Oxford Economics Discussion Paper Number 843*, 2018.

Shen, Bo and Julian Wright, “Why (Don't) Firms Free Ride on an Intermediary's Advice?,” *International Journal of Industrial Organization (forthcoming)*, 2018.

Teh, Tat-How and Julian Wright, “Steering by Information Intermediaries,” *Working Paper*, 2018.

Ursu, Raluca M., “The Power of Rankings: Quantifying the Effect of Rankings on Online Consumer Search and Purchase Decisions,” *Marketing Science*, 2018, *37* (4), 530 – 552.

Wals, Francisca and Maarten P. Schinkel, “Platform Monopolization by Narrow-PPC-BPG Combination: Booking et al.,” *International Journal of Industrial Organization (forthcoming)*, 2018.

Wang, Chengsi and Julian Wright, “Search Platforms: Showrooming and Price Parity Clauses,” *Working Paper*, 2017.

A Appendix: Proofs

Proof of Lemma 4. The expected channel revenue $p_i^{OTA} q_i^{OTA}$ is maximized by a price that solves

$$q_i^{OTA} + p_i^{OTA} \left(\frac{\partial q_i^{OTA}}{\partial p_i^{OTA}} + \frac{\partial q_i^{OTA}}{\partial p_i^{DIR}} \right) = 0. \tag{18}$$

The partial derivatives of π_i^C are given by

$$\frac{\partial \pi_i^C}{\partial p_i^{OTA}} = (1 - c_i)\, q_i^{OTA} + \Big((1 - c_i)\, p_i^{OTA} - b \Big) \frac{\partial q_i^{OTA}}{\partial p_i^{OTA}} + \Big(p_i^{DIR} - b \Big) \frac{\partial q_i^{DIR}}{\partial p_i^{OTA}} = 0, \tag{19}$$

$$\frac{\partial \pi_i^C}{\partial p_i^{DIR}} = q_i^{DIR} + \Big(p_i^{DIR} - b \Big) \frac{\partial q_i^{DIR}}{\partial p_i^{DIR}} + \Big((1 - c_i)\, p_i^{OTA} - b \Big) \frac{\partial q_i^{OTA}}{\partial p_i^{DIR}} = 0. \tag{20}$$

As prices and thus quantities must be equal with price parity, equation (8) simplifies to

$$q_i^k + p_i \left(\frac{\partial q_i^k}{\partial p_i^k} + \frac{\partial q_i^k}{\partial p_i^{-k}} \right) - \frac{2b}{(2 - c_i)} \left(\frac{\partial q_i^k}{\partial p_i^k} + \frac{\partial q_i^k}{\partial p_i^{-k}} \right) = 0,$$

which implies a larger optimal price than (18). □

Proof of Proposition 2. In equilibrium it must hold that $\frac{\partial \pi_i}{\partial p_i^{OTA}} + \frac{\partial \pi_i}{\partial p_i^{DIR}} = 0$, which is equivalent to

$$\left(\frac{\partial \pi_i^C}{\partial p_i^{OTA}} + \frac{\partial \pi_i^C}{\partial p_i^{DIR}}\right) + \frac{\alpha}{\alpha\phi_i(r^*) + (1-\alpha)}\phi_i'(r^*)\left(\frac{\partial r^*}{\partial p_i^{OTA}} + \frac{\partial r^*}{\partial p_i^{DIR}}\right)\pi_i^C = 0. \tag{21}$$

At prices of $p_L^k = p_R^k = p_i^*(\alpha = 0)$, the left-hand side of (21) is negative because the first term is zero, but the second is negative as $\phi_i' \sum_k \frac{\partial r^*}{\partial p_i^k} < 0$. The latter follows from Lemma 4 as $p_i^*(\alpha = 0)$ is above the level that maximizes the OTA revenue of that hotel. □

Proof of Lemma 5. For $p_i^{\text{DIR}} = p_i^{\text{OTA}}$ the difference $\frac{\partial \pi_i^C}{\partial p_i^{OTA}} - \frac{\partial \pi_i^C}{\partial p_i^{\text{DIR}}}$ reduces to

$$-c_i \left[\underbrace{q_i^{OTA} + p_i^{OTA}\left(\frac{\partial q_i^{\text{OTA}}}{\partial p_i^{\text{OTA}}} + \frac{\partial q_i^{\text{OTA}}}{\partial p_i^{\text{DIR}}}\right)}_{<0 \text{ at optimal uniform price}} - p_i^{\text{OTA}} 2\frac{\partial q_i^{\text{OTA}}}{\partial p_i^{\text{DIR}}}\right].$$

At the level that results in uniform prices, the above expression must be positive. The underlined part is negative as the uniform price is higher than the price that maximizes the expected OTA revenue per interested consumer (see Lemma 4). The remainder in the brackets is also negative as $\frac{\partial q_i^{OTA}}{\partial p_i^{DIR}} > 0$, which implies that the expression is positive. This means that the hotel has an incentive to increase the OTA price relative to the direct price. □

B Appendix: Travel destinations in the dataset

Tables 9 until 13 show the travel destinations covered in our dataset which is as in Hunold et al. (2018). We include the 25 biggest German cities (Table 10), a control sample of 20 pairs of German and non-German cities along the German border (Table 11), and other countries for which we chose a composition of the 15 biggest cities and 15 most popular travel destinations with the objective to gather representative data across touristic and urban destinations for these countries.

Table 9: Countries covered in dataset

Country	Cities covered
Germany	25 biggest cities
Various	20 pairs of cities near German border
Italy	15 biggest cities and 15 tourist destinations
Sweden	15 biggest cities and 14 tourist destinations
Canada	15 biggest cities and 15 tourist destinations
France	15 biggest cities and 15 tourist destinations
Austria	15 biggest cities and 15 tourist destinations

Table 10: Germany – Top 25 cities

Germany Top 25 cities				
Berlin	Stuttgart	Leipzig	Bochum	Karlsruhe
Hamburg	Dusseldorf	Dresden	Wuppertal	Mannheim
Munich	Dortmund	Hanover	Bielefeld	Augsburg
Cologne	Essen	Nuremberg	Bonn	Wiesbaden
Frankfurt am Main	Bremen	Duisburg	Munster	Gelsenkirchen

Table 11: Twin cities along German border

Pair	German City	Non-German neighbor	Neighbor country
1	Flensburg	Kolding	Denmark
2	Puttgarden/Fehmarn	Rodby	Denmark
3	Wilhelmshaven	Groningen	The Netherlands
4	Borkum	Schiermonnikoog	The Netherlands
5	Rheine	Enschede	The Netherlands
6	Aachen	Maastricht	The Netherlands
7	Heringsdorf	Wolin	Poland
8	Greifswald	Stettin	Poland
9	Cottbus	Zielona-Gora	Poland
10	Trier	Rosport	Luxembourg
11	Monschau	Eupen	Belgium
12	Pruem	St. Vith	Belgium
13	Saarbrucken	Metz	France
14	Karlsruhe	Strasbourg	France
15	Freiburg	Basel	Switzerland
16	Konstanz	St. Gallen	Switzerland
17	Oberstdorf	Bad Ischl	Austria
18	Garmisch-Partenkirchen	Innsbruck	Austria
19	Nuremberg	Pilsen	Czech Republic
20	Dresden	Prague	Czech Republic

Table 12: Cities covered in dataset

Italy	Canada	France	Sweden	Austria
		Biggest Cities		
Rome	Toronto	Paris	Stockholm	Vienna
Milan	Montreal	Marseille	Göteborg	Graz
Naples	Vancouver	Lyon	Malmö	Linz
Turin	Calgary	Toulouse	Uppsala	Salzburg
Palermo	Edmonton	Nice	Västeras	Innsbruck
Genoa	Ottawa	Nantes	Örebro	Klagenfurt
Bologna	Québec	Strasbourg	Linköping	Villach
Florence	Winnipeg	Montpellier	Helsingborg	Wels
Bari	Hamilton	Bordeaux	Jönköping	St. Pölten
Catania	Kitchener	Lille	Norrköping	Dornbirn
Venice	London	Rennes	Lund	Wiener Neustadt
Verona	Victoria	Reims	Umea	Steyr
Messina	Saint Catharines	Le Havre	Gävle	Feldkirch
Padua	Halifax	Saint-Étienne	Boras	Bregenz
Trieste	Oshawa	Toulon	Eskilstuna	Leonding
		Tourist Destinations		
Lecce	Regina	Grenoble	Växjö	Zell am See
Viareggio	St. John's	Cannes	Lulea	Kitzbühel
Matera	Fredericton	Chambéry	Falun	Bad Hofgastein
Sanremo	Charlotte Town	Annecy	Varberg	Hermagor
Mantova	Whitehorse	Aix-les-Bains	Visby	Schladming
Vasto	Yellowknife	Menton	Ystad	Mittelberg
Merano	Niagara On The Lake	Albertville	Kiruna	Neustift
Caltagirone	Whistler	Bayeux	Strömstad	Bad Gastein
Montecatini Terme	Banff	Argelès-sur-Mer	Ronneby	Velden am Wörther See
Narni	Jasper	Chamonix	Jokkmokk	Finkenstein am Faaker See
Abano Terme	Tofino	Évian-les-Bains	Grebbestad	Kirchberg in Tirol
Ischia	Dawson City	Cavalaire-sur-Mer	Marstrand	St. Kanzian
Monte Argentario	Churchill	Saint-Gervais-les-Bains	Jukkasjärvi	Mayrhofen
San Felice Circeo	Bay of Fundy	Gruissan	Stöllet	Seefeld in Tirol
Santa Margherita Ligure	Thousand Islands National Park	Sainte-Marine		Sölden

Selection of travel destinations

For Italy, Sweden, Canada, France, and Austria we selected the travel destinations in two steps. First, we looked up the 15 biggest cities in terms of population on Wikipedia, respectively. Additionally, for each country, we collected information about popular tourist destinations from travel guides and official tourism websites. We then ordered all these destinations by population and again took the 15 biggest locations. For Italy, France, Sweden, and Canada the websites were all accessed in January and February 2016. The Austrian cities were selected in April 2016 after the Austrian competition authority announced proceedings against narrow PPCs later in 2016.

The sources of the travel destinations can be found in the following table:

Table 13: Sources for travel destination selection

Country	Type	Source
Italy	Listing of health resorts	wikipedia.de
	10 most popular beaches	telegraph.co.uk
	Beyond Rome and Florence: 12 alternative Italian destinations	cnn.com
Sweden	Top 10 places in Sweden	neverstoptraveling.com
	Top 10 green Attractions	visitsweden.com
Canada	Travelers Choice	tripadvisor.com
	Tourist attractions	planetware.com
	Places to Go	de-keepexploring.canada.travel
France	The top 10 beach holidays	telegraph.co.uk
	Travelers Choice Destinations	tripadvisor.com
	16 top-rated tourist attractions in the French Alps	planetware.com
Austria	Most popular winter destinations	austriatourism.at
	Most popular summer destinations	austriatourism.at

C Appendix: Robustness checks

Table 14: Determinants of Booking.com's ranking – Additional Specifications

	(1)	(2)	(3)	(4)
Price at Booking/100	4.92***	4.95***	4.92***	4.84***
Preferred Partner	74.80***	74.82***	74.81***	74.81***
User Rating at Booking	-3.88*	-3.86*	-3.87*	-3.90*
# Bookings Last 3 Days	0.01**	0.01**	0.01**	0.01**
% Hotels Sold Out	40.56***	40.53***	40.56***	40.57***
Google Trends	-0.07***	-0.07***	-0.07***	-0.07***
(Booking-Direct)/Direct	-10.26***			
(Booking-Expedia)/Expedia	-10.76***			
PPC x (Booking-Direct)/Direct	1.02			
PPC x (Booking-Expedia)/Expedia	4.28**			
Booking-Direct		-0.05***		
Booking-Expedia		-0.04***		
(Booking-Direct)/Direct > 0			-10.51***	
(Booking-Expedia)/Expedia > 0			-9.14***	
(Booking-Direct)/Direct < 0			-7.48**	
(Booking-Expedia)/Expedia < 0			-5.77***	
I[0 > (Booking-Direct)/Direct >= -0.1]				0.69*
I[(Booking-Direct)/Direct < -0.1]				2.33***
I[0 < (Booking-Direct)/Direct <= 0.1]				0.53
I[(Booking-Direct)/Direct > 0.1]				-1.48**
I[0 > (Booking-Expedia)/Expedia >= -0.1]				-0.23
I[(Booking-Expedia)/Expedia < -0.1]				2.01***
I[0 < (Booking-Expedia)/Expedia <= 0.1]				0.06
I[(Booking-Expedia)/Expedia > 0.1]				-0.63*
Channel Availability	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Observations	2307513	2307513	2307513	2307513
R^2	0.91	0.91	0.91	0.91

Heteroscedasticity-robust standard errors not reported.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Determinants of Expedia's ranking – Additional Specifications

	(1)	(2)	(3)	(4)
Price at Expedia/100	-11.50***	-11.65***	-11.52***	-11.63***
Sponsored Ad	1.66*	1.73*	1.71*	1.71*
User Rating at Expedia	3.02**	3.19***	3.19***	3.12**
# Bookings Last 2 Days	0.66***	0.66***	0.66***	0.65***
% Hotels Sold Out	137.91***	138.10***	137.98***	138.08***
Google Trends	0.17***	0.17***	0.17***	0.17***
(Expedia-Direct)/Direct	-58.05***			
(Expedia-Booking)/Booking	-66.89***			
PPC x (Expedia-Direct)/Direct	50.16***			
PPC x (Expedia-Booking)/Booking	44.05***			
Expedia-Direct		-0.12***		
Expedia-Booking		-0.16***		
(Expedia-Direct)/Direct > 0			-38.20***	
(Expedia-Booking)/Booking > 0			-39.28***	
(Expedia-Direct)/Direct < 0			-24.36***	
(Expedia-Booking)/Booking < 0			-38.27***	
I[0 > (Expedia-Direct)/Direct >= -0.1]				0.11
I[(Expedia-Direct)/Direct < -0.1]				6.92***
I[0 < (Expedia-Direct)/Direct <= 0.1]				-1.59*
I[(Expedia-Direct)/Direct > 0.1]				-10.12***
I[0 > (Expedia-Booking)/Booking >= -0.1]				2.32***
I[(Expedia-Booking)/Booking < -0.1]				8.36***
I[0 < (Expedia-Booking)/Booking <= 0.1]				-2.14***
I[(Expedia-Booking)/Booking > 0.1]				-10.15***
Channel Availability	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Observations	2457164	2457164	2457164	2457164
R^2	0.92	0.92	0.92	0.92

Heteroscedasticity-robust standard errors not reported.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 16: Time structure of the reaction to deviations

	Booking.com			Expedia		
	(1)	(2)	(3)	(4)	(5)	(6)
Price at OTA/100	6.69***	4.74***	4.73***	-1.65*	-6.81***	-6.80***
PPP/Sponsored Ad^	72.94***	61.46***	61.41***	2.20	2.11	1.84
User Rating at OTA	-0.89	-11.08	-10.98	-5.43	0.20	-0.61
# Past Bookings°	0.08***	0.04***	0.04***	0.50***	0.14**	0.13**
\% Hotels Sold Out	19.03***	19.33***	19.32***	134.49***	136.82***	136.45***
Google Trends	0.14***	0.03	0.03	0.31***	0.14*	0.14*
(OTA-Direct)/Direct	-10.87***	-1.99	-0.91	-35.86***	-32.34***	-31.34***
(OTA-(Other OTA))/(Other OTA)	-4.17	-7.01**	-8.96***	-38.38***	-23.59***	-18.16**
Lag $t-1$ - (Dev. on Direct)			0.21			0.44
Lag $t-2$ - (Dev. on Direct)			-0.20			-0.30
Lag $t-3$ - (Dev. on Direct)			2.19			3.76
Lag $t-4$ - (Dev. on Direct)			-2.33			-9.60**
Lag $t-5$ - (Dev. on Direct)			-0.60			4.29
Lag $t-6$ - (Dev. on Direct)			-0.28			0.28
Lag $t-7$ - (Dev. on Direct)			-4.01*			-2.07
Lead $t+1$ - (Dev. on Direct)			-2.34			-3.53
Lead $t+2$ - (Dev. on Direct)			-1.83			-6.86
Lead $t+3$ - (Dev. on Direct)			-0.62			-0.67
Lead $t+4$ - (Dev. on Direct)			0.55			1.16
Lead $t+5$ - (Dev. on Direct)			-2.23			1.51
Lead $t+6$ - (Dev. on Direct)			0.19			10.34
Lead $t+7$ - (Dev. on Direct)			-2.70			0.83
Lag $t-1$ - (Dev. on other OTA)			1.80			-8.70
Lag $t-2$ - (Dev. on other OTA)			3.22			-12.13*
Lag $t-3$ - (Dev. on other OTA)			-1.10			-17.88**
Lag $t-4$ - (Dev. on other OTA)			-2.22			-13.07***
Lag $t-5$ - (Dev. on other OTA)			-1.41			-13.13**
Lag $t-6$ - (Dev. on other OTA)			0.30			-8.83
Lag $t-7$ - (Dev. on other OTA)			3.88			-9.50
Lead $t+1$ - (Dev. on other OTA)			4.62**			0.25
Lead $t+2$ - (Dev. on other OTA)			3.94			4.51
Lead $t+3$ - (Dev. on other OTA)			2.24			9.08
Lead $t+4$ - (Dev. on other OTA)			-1.48			-9.99
Lead $t+5$ - (Dev. on other OTA)			1.50			6.45
Lead $t+6$ - (Dev. on other OTA)			-1.35			1.33
Lead $t+7$ - (Dev. on other OTA)			1.00			-0.96
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	146883	38697	38697	151397	44045	44045
R^2	0.91	0.95	0.95	0.84	0.87	0.87

Heteroscedasticity-robust standard errors not reported.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: ^ For Booking.com, this is an indicator of whether the hotel is participating in the Preferred Partner Program, while for Expedia it indicates that the hotel had a sponsored advertisement. ° Past Bookings are as reported from the respective OTA; for Booking.com, this is in the last three days, while for Expedia it is for the last two days.

D Appendix: Price differentiation across channels

Table 17: Characteristics of hotels with lower prices on their websites than at Booking.com

	(1)	(2)	(3)	(4)
Average Price/100	2.77***	2.57***	2.57***	5.67***
Preferred Partner	-2.07	-1.90	-1.89	-0.87
Hotel Stars	-1.63	-1.60	-1.60	
User Rating at Booking	0.44	0.45	0.45	-3.57
Number of User Ratings/100	-0.08	-0.03	-0.03	
Number of Rooms/100	-2.06**	-1.96**	-1.98**	
Chain Hotel	-16.83***	-16.79***	-16.79***	
Google Trends	0.04*	0.05*	0.05**	-0.00
# Bookings Last 3 Days		-0.04***	-0.04***	-0.01**
% Hotels Sold Out		5.46***	4.27***	2.33***
7 days before			-2.32***	-2.21***
14 days before			-2.06**	-2.00***
Month FE	Yes	Yes	Yes	Yes
FE	City	City	City	Hotel
Observations	375758	375758	375758	375758
R^2	0.11	0.11	0.11	0.46

Heteroscedasticity-robust standard errors not reported.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 18: Characteristics of hotels with lower prices at Expedia than at Booking.com

	(1)	(2)	(3)	(4)
Average Price/100	0.31	0.29	0.28	0.05
Preferred Partner	1.48***	1.53***	1.53***	1.00*
Hotel Stars	1.21***	1.21***	1.21***	
User Rating at Booking	-1.16***	-1.18***	-1.17***	-1.10
Number of User Ratings/100	-0.04**	-0.02	-0.02	
Number of Rooms/100	-0.38	-0.33	-0.32	
Chain Hotel	-3.49***	-3.48***	-3.48***	
Google Trends	-0.04***	-0.04***	-0.04***	-0.01
# Bookings Last 3 Days		-0.02***	-0.02***	-0.01***
% Hotels Sold Out		0.41	0.65	0.82***
7 days before			0.81***	0.60***
14 days before			0.04	-0.11
Month FE	Yes	Yes	Yes	Yes
FE	City	City	City	Hotel
Observations	1622767	1622767	1622767	1622767
R^2	0.07	0.07	0.07	0.34

Heteroscedasticity-robust standard errors not reported.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 19: Characteristics of hotels with lower prices on their websites than at Expedia

	(1)	(2)	(3)	(4)
Average Price/100	3.69***	3.36***	3.40***	5.17***
Sponsored Ad	-3.68*	-3.62*	-3.56*	-0.44
Hotel Stars	-2.28	-2.17	-2.20	
User Rating at Expedia	0.00	0.16	0.13	0.06
Number of User Ratings/100	-0.18**	-0.16*	-0.16*	
Number of Rooms/100	-1.46**	-1.41**	-1.43**	
Chain Hotel	-17.41***	-17.56***	-17.56***	
Google Trends	0.02	0.01	0.02	-0.01
# Bookings Last 2 Days		-0.09	-0.08	-0.00
% Hotels Sold Out		14.52***	10.10***	8.89***
7 days before			-3.33***	-3.26***
14 days before			-2.59**	-2.46***
Month FE	Yes	Yes	Yes	Yes
FE	City	City	City	Hotel
Observations	444470	444470	444470	444470
R^2	0.09	0.09	0.09	0.45

Heteroscedasticity-robust standard errors not reported.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 20: Characteristics of hotels with lower prices at Booking.com than at Expedia

	(1)	(2)	(3)	(4)
Average Price/100	-0.24	-0.17	-0.14	-0.37**
Sponsored Ad	0.27	0.28	0.30	0.13
Hotel Stars	-0.28	-0.30	-0.32	
User Rating at Expedia	-2.96***	-2.99***	-3.01***	-4.56**
Number of User Ratings/100	-0.02	-0.02	-0.02	
Number of Rooms/100	-1.48***	-1.46***	-1.47***	
Chain Hotel	-1.83***	-1.84***	-1.84***	
Google Trends	0.06	0.06	0.06	0.05***
# Bookings Last 2 Days		-0.02	-0.02	0.03
% Hotels Sold Out		-3.58	-5.43	-3.51***
7 days before			-1.18***	-1.21***
14 days before			-1.54***	-1.66***
Month FE	Yes	Yes	Yes	Yes
FE	City	City	City	Hotel
Observations	1554516	1554516	1554516	1554516
R^2	0.10	0.10	0.10	0.34

Heteroscedasticity-robust standard errors not reported.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$