Competition in the digital era : evidence from the hotel industry

Morgane Cure

NNT : 2020IPPAG013

Thèse de doctorat

INSTITUT POLYTECHNIQUE DE PARIS

ENSAE

IP PARIS

Concurrence à l'ère du numérique : exemples dans l'industrie hôtelière

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'École nationale de la statistique et de l'administration économique

École doctorale n°626 de l'Institut Polytechnique de Paris (ED IP Paris)
Spécialité de doctorat : Sciences économiques

Thèse présentée et soutenue à Palaiseau, le 9 décembre 2020, par

MORGANE CURE

Composition du Jury :

Marie-Laure Allain Directrice de recherche, CNRS et Ecole Polytechnique (CREST)	Présidente
Philippe Gagnepain Professeur des Universités, Université Paris 1 Panthéon-Sorbonne	Rapporteur
Thierry Pénard Professeur des Universités, Université Rennes 1	Rapporteur
Thibaud Vergé Professeur, ENSAE Paris (CREST)	Directeur de thèse

Remerciements

Si je devais résumer l'effet ressenti lorsqu'on se lance dans une thèse, je dirais que c'est un peu comme arriver au sommet d'un plongeoir, dont la hauteur est proportionnelle au nombre d'années d'étude déjà réalisées (environ cinq mètres), s'élancer, puis sauter, sans avoir la certitude de savoir nager. Épilogue.

Pour l'ascension des marches jusqu'au sommet, échauffement indispensable avant de prétendre sauter, je remercie le Magistère d'économie de Paris 1 puis l'ENSAE Paris, leurs enseignants-chercheurs ainsi que leurs équipes pédagogiques, de m'avoir donné la possibilité d'y suivre un enseignement de grande qualité. Ces formations ont représenté pour moi un véritable tremplin et m'ont permis d'acquérir les connaissances nécessaires pour accéder aux études doctorales.

Mes remerciements vont également au GENES, plus particulièrement à Francis Kramarz, ex-directeur du Centre de Recherche en Économie et Statistiques et à Laurent Linnemer, ex-directeur du Laboratoire d'Économie Industrielle, qui m'ont offert le financement nécessaire au début de cette thèse, un cadre et des conditions de travail inégalés, à l'instar des chercheurs qui y exercent et qui font de cet établissement un si beau centre olympique. Le laboratoire ne serait pas le même sans le soutien quotidien du personnel administratif et je remercie à ce titre, tout particulièrement, Pascale Deniau, Murielle Jules, Eliane Madelaine, Lyza Racon, Arnaud Richet, Fanda Traoré et Edith Verger pour leur appui.

Je remercie chaleureusement Marie-Laure Allain, Philippe Gagnepain et Thierry Pénard d'avoir accepté d'être les juges de ce plongeon et de mon acclimatation dans le grand bain, d'avoir lu attentivement l'ensemble de mes travaux pour me permettre d'y apporter plus de profondeur. Marie-Laure, en participant à chacun de mes comités de suivi de thèse depuis trois ans, en plus d'être une spécialiste des relations verticales qui inondent les trois chapitres de cette thèse, est le parfait témoin de mon évolution de jeune chercheuse. Philippe, qui a été mon professeur d'économie des réseaux il y a six ans au Magistère d'économie est un spécialiste des méthodes d'économétrie appliquée dont plusieurs sont exploitées dans cette thèse. Thierry, qui ne me connaissait peut-être pas avant la lecture de ces chapitres, est un fin connaisseur de l'économie du numérique. Ses recherches et publications sur la tarification, les plateformes collaboratives, leurs systèmes de réputation et le lien à tisser avec la politique de la concurrence, font de lui l'un des chercheurs français avec lequel je partage le plus d'intérêts.

Une fois préparée et arrivée au sommet, le jury est prêt, il n'y a plus à hésiter, il faut se lancer et sauter. L'océan des connaissances se dévoile bien plus grand qu'il n'y parait et l'apprentissage de la nage bien plus technique qu'anticipé. Dans les longues phases d'apnée, comment être sûre de ne pas perdre pied ? De nombreux soutiens étaient présents pour m'accompagner.

Ma reconnaissance pour Thibaud Vergé, mon directeur de thèse, est immense. Il a eu l'un des rôles les plus difficiles à jouer, devant m'apprendre à nager dans le courant de la recherche sans trop se faire éclabousser ! Merci d'avoir su jouer ce rôle avec brio. Je lui suis également très reconnaissante de m'avoir fait confiance dès la pataugeoire du Master en acceptant d'encadrer mon mémoire, ma thèse, puis en me proposant de coécrire à ses côtés. Il a toujours été de bons conseils, présent dans les moments de doute, sans pour autant prendre les décisions ou faire le travail à ma place. Ces quelques lignes ne sauraient refléter ma gratitude. J'espère que cette première médaille ne sera que le début d'une longue carrière d'entraîneur-maître-nageur pour lui.

J'ai également eu le bonheur de constater qu'une fois dans le grand bain, certains doctorants avaient déjà quelques années d'expérience qu'ils n'ont pas hésité à partager. A cet égard, je remercie tout particulièrement Julien Monardo et Hugo Molina, pour leur accueil, leur bienveillance et leurs nombreux conseils sur les fondements théoriques et la mise en pratique de l'économétrie structurelle (ou brasse coulée). Je remercie également Arthur Cazaubiel, mon co-auteur et relayeur du deux fois cinquante mètres, avec lequel je ne compte plus les longueurs à débattre de nos différentes techniques de nage. Je dois à chacun d'eux une grande partie des connaissances acquises durant cette thèse.

Ce plongeon a également bénéficié de tous les conseils bienveillants des chercheurs du CREST, experts du quatre fois quatre nages et en particulier : Marie-Laure Allain, Christophe Béllego, David Benatia, Pierre Boyer, Claire Chambolle, Philippe Choné, Romain de Nijs, Xavier d'Haultfoeuille, Roxana Fernández, Laurent Linnemer, Sébastien Mitraille, Said Souam et Michael Visser.

Je remercie mes (autres) co-auteurs, Matthias Hunold, Bjørn Olav Johansen, Reinhold Kesler, Ulrich Laitenberge et Thomas Larrieu qui ont accepté de me laisser nager à leurs côtés malgré mon peu d'expérience. Ces collaborations m'ont, entre autre, permis d'accéder à des bases de données sans lesquelles aucun des travaux de cette thèse n'auraient pu être réalisés.

Sur le papier, cette thèse, ce grand plongeon dans le monde de la recherche, avait tous les éléments pour se dérouler parfaitement. Sur la dernière ligne droite, l'arrivée de la fatidique troisième année, l'eau a commencé à se troubler. La direction des études de l'ENSAE a été une chouette bouée à laquelle m'arrimer.

Je tiens à remercier en ce sens Pierre Biscourp, directeur de l'ENSAE, Lionel Wilner puis Corentin Trévien, successivement directeurs des études, ainsi que Philippe Choné, responsable de la séquence microéconomie, d'avoir encouragé et soutenu mon recrutement sur le poste (très mal dénommé) d'assistante d'enseignement en microéconomie. Après deux années de thèse à temps complet, loin d'être une parenthèse, ce poste m'a permis de pouvoir faire aboutir ce projet. Il a été une façon de se diversifier sans pour autant s'égarer.

Merci également à Laurent Davezies et Rosalinda Solotareff, actuels directeur des études et directrice adjointe-responsable des Masters, pour la confiance qu'ils m'ont accordée dès leur arrivée, pour l'attention qu'ils ont toujours porté au bien être de l'équipe et pour l'intérêt dont ils ont témoigné pour mes travaux de recherche.

Mes remerciements vont aussi à mes collègues assistants d'enseignement, en particulier Nicolas Baradel, Jules Depersin, Christophe Gaillac, Jérémy l'Hour, Lucie Neirac, Fabien Perez, Arthur Thomas, Jérôme Trinh et Wissal Sabbagh, assistants de choc avec qui, pour l'esprit collectif qui a régné, il a toujours été agréable de travailler ces quelques années. Par mon engagement, j'espère avoir pu rendre à l'école une partie de ce qu'elle m'a apporté.

Mes journées à l'ENSAE auraient été bien moins agréables si je n'avais pas eu la chance d'y rencontrer Sarah Teriitaumihau, avec qui je partage bien plus que des cafés, une véritable amitié. Je la remercie d'être toujours partante pour mes grands défis sportifs, pour nos discussions à toutes heures de la journée et pour le soleil tahitien que sa bonne humeur diffuse au quotidien. Je lui suis également reconnaissante, ainsi qu'à Audrey Mallet, pour la relecture attentive de certaines parties de cette thèse.

Après la préparation et le plongeon, l'eau trouble puis l'oxygène des alliés, la dernière longueur, à laquelle j'ai souvent pensé durant ces quatre années, est arrivée. C'est un peu le meilleur moment, celui où le nageur essoufflé (mais pas à bout de souffle) aperçoit le mur du bassin et d'un doigt le touche pour sceller son destin.

Dans les gradins, j'aimerais remercier mes collègues doctorants en économie industrielle, comme la baleine bleue, espèce rare et désormais menacée. Il a été divertissant à mon tour de les voir plonger puis parfois accoster le bureau 4071, un peu dépeuplé. J'espère leur avoir réservé un accueil aussi chaleureux qu'à mon arrivée. Rémi Avignon, Morgane Guignard et Alexis Larousse, je reste là, à vos côtés, pour vous encourager.

Je garde un excellent souvenir des moments partagés avec l'équipe de football du ManCRESTer United, en particulier mon 11 préféré : le capitaine Lucas Girard, les défenseurs Gwen-Jiro Clochard et Fabien Perez, le goal Yannick Guyonvarch, les tantôt latéraux, tantôt attaquants, Reda Aboutajdine, Rémi Avignon, Aurélien Bigo, Thomas Delemotte, Etienne Guigue, Jérémy Hervelin et Martin Mugnier.

Je remercie également tous les autres doctorants que j'ai croisés et avec qui j'ai échangé, en particulier Antoine Bertheau, Guidogiorgio Bodrato, Léa Bou Sleiman, Pauline Carry, Badr-Eddine Cherief Abdellatif, Pierre-Edouard Collignon, Antoine Ferey, Germain Gauthier, Clémence Lenoir, Esther Mbih, Boris Muzellec, Sandra Nevoux, Elio Nimier-David, Mathis Oliva, Louis Pape, Emmanuel Paroissien, Berengère Patault, Elia Pérennès, Anasuya Raj, Lionel Riou-Durand, Emilie Sartre, Yuanzhe Tang, Clémence Tricaud, Antoine Valtat, Ao Wang et Benjamin Walter. Ces moments ont été (trop) rares, mais toujours d'une grande qualité.

Je remercie aussi l'ensemble de mes collègues et amis du 3ème étage qui ont partagé et égayé mes repas et pauses café à l'espace convivialité. Loin d'être anodins, j'ai beaucoup apprécié ces moments où j'ai pu sortir la tête de l'eau et vous découvrir, en particulier avec Elisabeth Andreoletti-Cheng, Nicolas Apoteker, Michèle Arger, Marie-Christine Baker, Luciana Barbera, Stéphanie Breuil, Emmanuel Cavadini, Emmanuelle Chesseret, Carole Delamare, Oumi Djoumoi, Sylvie Floquet, Anne Grillot et ses petits pots de rillettes, Guillaume Hû, Christine Imbert-Courtois, Marylène Jubertie, Christophe Lagarde, Audrey Mallet, Françoise Moreels, Karine Pain, Agnes Pélage, Thierry Raussin, Sri Srikandan, Patrick Tartrou et Edith Verger.

Enfin, dans les tribunes, il y a vos fans, ceux qui crient toujours plus fort que les autres, qui sont toujours là pour vous encourager, même s'ils ne connaissent rien à la natation. Je remercie pour cela mes parents, Marie & Patrick, qui m'ont soutenue sans avoir toujours cru en l'utilité de savoir plonger de si haut. Je remercie en particulier ma mère pour tout ce qu'elle a sacrifié au profit de mon éducation et pour son ouverture d'esprit. J'espère la rendre fière.

Merci également à mes soutiens de l'ombre : Sabrina, Chloé, Josette, Virginie, Sylvie et beaucoup d'autres. Leur présence et leurs textos ont toujours su maintenir mon moral à flots. Je remercie enfin Daniel, qui a été le premier à croire en mes chances d'entrer à l'ENSAE Paris puis d'être acceptée en thèse. Il est le plus grand de mes supporters. Je le remercie pour sa motivation, son inspiration, la liberté qu'il ma donnée et sa patience à mes côtés depuis ces 8 années. Cette thèse, c'est aussi un peu la sienne.

À toutes celles et ceux qui ont été les petits brassards bien gonflés m'ayant permis de ne pas couler,

MERCI.

Contents

Introduction

Contexte

Les technologies de l'information et de la communication (TIC), dont Internet, ont radicalement changé les comportements de consommation. En 2019, 63 % des utilisateurs d'Internet dans l'Union européenne ont effectué un achat en ligne[1]. C'est deux fois plus qu'il y a dix ans. Le recours aux ventes en ligne est inégalement réparti entre les industries. Ce type de vente est plus spécifiquement développé dans le commerce de détail et dans les activités d'hébergement et de restauration. Près de la moitié des entreprises spécialisées dans ces domaines ont reçu des commandes par Internet en France en 2019[2].

Parallèlement à l'accroissement des ventes en ligne, une diversité de modèles économiques s'est développée. Certains vendeurs physiques ont adapté leur entreprise au nouvel environnement numérique en intégrant leur propre site de vente en ligne à leur expérience client. Centralisant à la fois le processus productif et les organes de distribution, ces vendeurs sont dits "intégrés verticalement" (Hart et al., 1990). D'autres commerçants s'appuient sur de nouveaux acteurs numériques pour leur distribution en ligne : les plateformes d'intermédiation. Par la quantité considérable d'informations qu'elles rassemblent, ces plateformes orientent et simplifient le choix des consommateurs, tout en offrant aux vendeurs un accès centralisé à la demande. Beaucoup de ces plateformes sont multifaces, générant des externalités de réseaux indirectes entre plusieurs catégories d'utilisateurs (les "faces"). La valeur de l'offre ou du produit vendu augmente avec le nombre d'usagers de la plateforme (Rochet and Tirole, 2003; Armstrong, 2006) qui est alors valorisée pour son intermédiation.

Dans un modèle de vente au détail traditionnel, les intermédiaires achètent au prix de gros un produit ou un service à un fournisseur puis le revendent aux consommateurs au prix de détail. L'intermédiaire se rémunère pour ses services grâce à la différence entre ces deux prix. Avec l'émergence de la vente en ligne, le modèle tradition-

1 ICT Access and Usage by Households and Individuals, OECD Telecommunications and Internet Statistics (database), OECD.Stats

2 ICT Access and Use by Businesses, OECD Telecommunications and Internet Statistics (database), OECD.Stats

nel de revente a peu à peu laissé sa place au modèle d'agence (Foros et al., 2014).

Dans ce dernier, le prix final payé par le consommateur n'est plus fixé par l'intermédiaire mais par le fournisseur qui rémunère dans un second temps l'intermédiaire pour ses services. La règle de partage des revenus préempte la vente et prend généralement la forme de frais de commission, c'est-à-dire un pourcentage fixe du prix final payé. Le secteur de l'immobilier, dans lequel le vendeur fixe le prix de son bien et en reverse un pourcentage à l'agence immobilière, est particulièrement caractéristique du modèle d'agence.

Une relation économique dans laquelle le prix de détail est fixé par le fournisseur et non pas par les intermédiaires, comme dans le modèle d'agence, représente un terrain fertile pour différents cas de restrictions verticales (Mathewson and Winter, 1984). Cette situation rappelle en particulier celle des prix de revente imposés (*Resale Price Maintenance* - RPM), dans laquelle le fournisseur indique un prix fixe ou minimum en deçà duquel les intermédiaires ne doivent pas vendre aux consommateurs sous peine de non approvisionnement. En France et dans l'Union européenne[3] ces pratiques sont interdites *per se* en raison de la pression inflationniste sur les prix qu'elles sont susceptibles de générer (Rey and Vergé, 2010).

Les industries du numérique sont particulièrement concernées par certaines restrictions verticales. Avec le développement d'intermédiaires comme les plateformes et les places de marché (*marketplaces*), de nouvelles pratiques contractuelles comme les clauses de la nation la plus favorisée (*Most Favor Nation clauses* - MFN, (Johnson, 2017)) ont fait leur apparition. Ces pratiques inquiètent les autorités de concurrence car des analyses suggèrent qu'elles s'apparenteraient à des prix de revente imposés (*"RPM as its worst"*, (Fletcher and Hviid, 2014)), dans leur forme la plus sévère. Quel que soit l'acteur qui fixe le prix, les clauses de la nation la plus favorisée et les prix de revente imposés ont pour point commun l'existence d'une relation verticale de distribution. En outre, les économistes observent que les cas de prix de revente imposés les plus dommageables, c'est à dire facilitant la collusion ou l'éviction de concurrents en aval, impliquent toujours une relation horizontale entre les prix finaux payés par les consommateurs ; résultat plus généralement observé dans les cas de clauses de la nation la plus favorisée.

Cette clause tire son origine du commerce international et fait référence à la situation économique d'un pays bénéficiant des meilleures conditions commerciales accordées par son partenaire, par exemple les tarifs les plus bas, les barrières commerciales les moins nombreuses et les quotas d'importation les plus élevés. Son importance est telle qu'elle est inscrite comme article premier de l'Accord général sur les

[3] Conformément à l'article L 442-5 du Code de commerce et au règlement 330/2010 de la Commission Européenne du 20/04/2010.

tarifs douaniers et le commerce (*General Agreement on Tariffs and Trade* - GATT, 1947)[4] régissant le commerce des marchandises et signé par la majorité des pays internationalement reconnus à ce jour. Dans le cadre du numérique, cette clause permet aux plateformes d'intermédiation d'exiger d'un fournisseur qu'il ne vende pas son produit à un prix inférieur sur tout autre canal de vente, qu'il s'agisse d'une autre plateforme en ligne ou de son canal de vente directe. Autrement dit, l'intermédiaire qui impose cette clause exige une politique tarifaire au moins aussi compétitive que tout autre canal de vente utilisé par le fournisseur. Les clauses de la nation la plus favorisée sont également connues sous le nom de clauses de parité de prix (*Price Parity Clauses* - PPC), appellation résumant l'effet susceptible de se produire si la clause est exigée par l'ensemble des plateformes d'intermédiation utilisées par le fournisseur. Lorsque chacun des intermédiaires impose une clause de la nation la plus favorisée, l'entreprise amont fixe des prix de détail identiques sur tous ses canaux de vente, ce qui se traduit par une absence de différenciation tarifaire, c'est-à-dire une parité sur le prix. Cette dernière correspond à l'élément horizontal qui rendrait les clauses de la nation la plus favorisée plus dommageables que les prix de revente imposés : les prix sont à la fois plus élevés en aval mais également les mêmes sur tous les canaux de vente du fournisseur.

Les toutes premières théories du préjudice à l'encontre de telles clauses suggèrent que les effets anti-concurrentiels proviennent d'un assouplissement de la concurrence entre les intermédiaires, d'une augmentation des taux de commission et des prix finaux (Johnson, 2017; Wang and Wright, 2016) ainsi que de l'éviction de nouvelles plateformes entrantes (Boik and Corts, 2016) et de la réduction de l'investissement des plateformes déjà actives sur le marché (Wang and Wright, 2020). Les clauses de parité de prix réduisent les incitations pour les plateformes à se concurrencer entre elles, à travers les taux de commission, car elles ne peuvent pas s'attendre à ce que les fournisseurs baissent leurs prix sur des plateformes moins chères. Des commissions plus élevées se répercutent sur les prix payés par les consommateurs. En outre, même si les nouvelles plateformes entrantes facturent des commissions moins élevées, elles ne peuvent pas rivaliser sur les prix et gagner des parts de marché auprès des opérateurs historiques et renoncent par conséquent à venir concurrencer le marché. D'autres travaux suggèrent au contraire que les clauses de parités n'affectent pas les prix finaux, notamment lorsque les commissions ne sont pas linéaires (Rey and Vergé, 2016). Dans le cas où la concurrence entre fabricants est suffisamment importante, elles peuvent même mener à des prix plus faibles (Johansen and Vergé, 2017). En l'absence d'un large consensus sur leurs effets, les clauses de parité de prix ont fait l'objet de plusieurs investigations menées par les autorités de concurrence dans différents secteurs et pays.

[4] World Trade Organisation

Régulation

L'un des premiers cas de concurrence les mettant en cause dans un secteur du numérique remonte au début des années 2010 sur le marché des livres électroniques (dits "*E-books*"). À cette époque, l'entreprise américaine Amazon était le vendeur dominant le marché avec son produit *Kindle E-reader*. Amazon fonctionnait via un modèle traditionnel de revente et les principaux éditeurs, parmi lesquels Harper Collins et Hachette, protestaient contre sa politique de prix. Ils considéraient que le prix de revente des livres choisi par Amazon était trop bas et à même de concurrencer leur canal de vente physique en librairie. En lançant l'iPad, Apple avait obtenu de ces éditeurs qu'ils revoient leurs accords avec Amazon et qu'ils acceptent un modèle d'agence avec une commission de 30% sur le prix de vente. D'autres clauses imposaient à chaque éditeur de fixer, sur la plateforme d'Apple, le prix le plus bas proposé à tout intermédiaire concurrent, soit une clause de parité de prix. Apple et les éditeurs s'étaient vus reprocher, par la Commission Européenne, des pratiques concertées et s'étaient engagés, en plus de rompre les contrats qui les liaient, à ne pas appliquer de clauses de parité de prix pendant une durée de cinq ans[5]. Le sujet est réapparu dans ce secteur en juin 2015 lorsqu'Amazon a exigé des éditeurs de livres électroniques qu'ils notifient toutes conditions générales de vente plus favorables (ou alternatives) qu'ils proposent à d'autres intermédiaires et qu'ils les rendent disponibles sur sa plateforme. Suspecté d'abus de position dominante, Amazon s'est engagé en 2017 à ne plus appliquer ou introduire de clause de parité pendant cinq ans[6].

Le cas des livres numériques n'est pas isolé. Plus généralement, en octobre 2012, l'office britannique de la concurrence (*Office of Fair Trading* - OFT[7]) a ouvert une enquête à l'encontre d'Amazon qui, sur le marché de la vente de détail, exigeait qu'un commerçant ne puisse pas vendre un produit sur son propre site ou sur une autre plateforme de vente en ligne à un prix inférieur à celui proposé sur la place de marché Amazon Marketplace. L'OFT[8] et l'autorité de concurrence allemande, le Bundeskartellamt[9], ont clôturé leur enquête respective en 2013 suite à l'accord d'Amazon de supprimer la clause.

La très large audience touchée par un géant des sites marchands, tel qu'Amazon, premier site internet visité en France en 2020[10], accroît les craintes quant aux répercussions négatives des effets anti-concurrentiels que pourraient générer les clauses de parité de prix. Si la numérisation des marchés ne révolutionne pas les pratiques, elle

5 European Commission decision (2012) ; Case COMP/AT.39847-E-BOOKS
6 European Comission decision (2017) ; Case AT.40153 E-book MFNs and related matters (Amazon)
7 Le 1er Avril 2014, l'OFT et la comission de la conccurence (Competition Comission - CC) ont été abolies et leurs fonctions transférées à la CMA.
8 OFT press release (2013) - "*OFT welcomes Amazon's decision to end price parity policy*"
9 Bundeskartellamt press release (2013) - "*Amazon abandons price parity clauses for good*"
10 Fédération du e-commerce et de la vente à distance (FEVAD)

complexifie grandement leur analyse. Dans le cas des clauses de parité, alors qu'il s'agit généralement d'apprécier la concurrence entre opérateurs physiques de revente presque identiques, dans le domaine du numérique, différents canaux, tels que les plateformes d'intermédiation et le canal de vente directe en ligne du fabricant, co-existent. Afin de prendre en compte la diversité des acteurs, deux types de clauses de parité de prix ont été distingués par les autorités : les clauses étendues et les clauses restreintes. Les premières exigent du vendeur qu'il publie sur la plateforme le même prix ou un prix inférieur à celui proposé sur tout autre canal de vente en ligne. En comparaison, les clauses de parité de prix restreintes sont moins exigeantes et ne considèrent dans les canaux concurrents à la plateforme que le canal de vente directe. Un vendeur appliquant la clause de parité restreinte est contraint de publier sur la plateforme le même prix ou un prix inférieur à celui publié sur son propre site internet.

L'industrie hôtelière

L'industrie hôtelière illustre parfaitement cette séparation avec des cas impliquant à la fois des agences de voyages en ligne (*Online Travel Agencies* - OTAs), parmi lesquelles Booking.com et Expedia, mais également le canal de vente directe des hôteliers. En 2015, plusieurs agences de voyage en ligne ont imposé des clauses de parité étendues aux hôteliers, leur interdisant de vendre une chambre moins cher sur tout autre canal de vente en ligne que la plateforme. La substitution entre les différents canaux de vente et, *in fine*, la définition du marché ont été des éléments importants du débat, les autorités concluant finalement que les ventes directes des hôteliers n'appartenaient pas au même marché que les ventes effectuées par le biais des agences de voyages en ligne. Les autorités ont en effet estimé que les agences de voyage en ligne offrent un ensemble de services qui inclut la recherche et la comparaison ainsi que la possibilité de réserver en ligne, alors que les sites internet des hôtels ne proposent que la possibilité de réserver. Si les clauses étendues réduisent la concurrence entre plateformes, les clauses restreintes permettent, dans ce cadre, d'éviter une situation dans laquelle un consommateur chercherait sur une plateforme puis se rendrait sur le canal direct de l'hôtelier pour acheter moins cher (dit le "*showrooming*"). Elles sont donc considérées comme nécessaires à la soutenabilité du modèle économique des plateformes. En avril 2015, les autorités de concurrence française, italienne et suédoise ont simultanément accepté les engagements proposés par Booking.com à transformer leurs clauses de parité de prix étendues en clauses de parité restreintes[11]. En France, la loi pour la croissance, l'activité et l'égalité des chances économiques de juillet 2015[12] (dite "loi Macron") a finalement interdit toutes les formes de clauses de parité dans

11 Report on the monitoring exercise carried out in the online hotel booking sector by EU competition authorities in 2016. Lien.

12 Loi n° 2015-990 du 6 août 2015 pour la croissance, l'activité et l'égalité des chances économiques.

l'industrie hôtelière, décision également adoptée par l'Autriche en novembre 2016, l'Italie en août 2017[13] et la Belgique en juillet 2018[14]. En Allemagne, le Bundeskartellamt a directement interdit les clauses étendues et restreintes imposées par HRS.com (décembre 2013)[15] et Booking.com (décembre 2015)[16]. Si ces décisions sont courantes dans l'Union Européenne, les clauses dans le secteur de l'hôtellerie n'ont pas toujours fait l'objet de régulation comme aux États-Unis ou en Amérique du Sud.

En plus de représenter un cas typique pour les clauses de parité de prix, l'industrie hôtelière est un cas d'étude intéressant car particulièrement propice au développement d'acteurs intégralement numériques (dits *"pure players"*). En plus des plateformes de réservations d'hôtels en ligne pré-citées, cette industrie est marquée par l'émergence de sites de comparaison d'offres (dits "*meta-search platforms*"- MSP ou "*Price Comparison Website*" - PCW - lorsqu'ils ne comparent que les prix), comme Kayak.fr ou Trivago.fr. Ces plateformes collectent et rassemblent les prix proposés par les différents canaux de vente pour un seul et même produit ou service et, à la différence des agences de voyage en ligne, ne permettent pas un achat direct mais proposent une redirection vers le canal de vente sélectionné par le consommateur. Ces sites soulèvent plusieurs problématiques concurrentielles intéressantes.

Premièrement, tout comme les agences de voyage en ligne, ils peuvent faire l'objet de clauses de parités de prix. Ainsi en 2015, l'Autorité de la concurrence et des marchés britannique (*Competition and Markets Authority* - CMA) a publié le rapport final de son enquête ouverte en septembre 2012, à l'issue de laquelle furent interdits les clauses de parité de prix étendues et les comportements équivalents sur le marché de l'assurance automobile[17]. Jusqu'alors, les sites de comparaison obligeaient les assureurs à proposer des prix qui n'étaient pas plus élevés que ceux proposés sur tout autre site, y compris leur propre site internet.

D'autre part, ces plateformes posent une question plus générale quant à la composition des algorithmes qu'elles utilisent afin de classer les offres, que ce soit les hôtels entre eux ou les différents canaux de vente pour un hôtel donné. Si l'argument tarifaire était initialement le principal critère du classement, d'autres éléments entrent désormais en compte, ce qui remet en cause la pertinence et l'objectivité de ce qui est présenté aux consommateurs. En janvier 2020 la commission australienne de la concurrence et de la consommation (*Australian Competition and Consumer Commission* - ACCC) a ainsi conclu que Trivago avait induit les consommateurs en erreur en indiquant que son site les aiderait à identifier les tarifs les moins chers disponibles pour un hôtel donné alors que son algorithme de classement accordait un poids prépondérant

13 Communiqué de presse - HOTREC Hospitality Europe
14 Communiqué de presse - Union des métiers et des industries de l'hôtellerie.
15 Bundeskartellamt – Décision n° B9-66-10 du 20 décembre 2013
16 Bundeskartellamt – Décision n° B9-121-13 du 22 décembre 2015
17 CMA Decision (2015) ; Private Motor Insurance Market Investigation

à la commission (au coût par clic) payée par le canal de vente à Trivago[18].

Enfin, dans les cas très particulier de l'hôtellerie, la structure verticalement intégrée des sites de comparaison de prix et des agences de voyage pose question. En effet, la plateforme Kayak a été rachetée en 2013 par le groupe Priceline également propriétaire d'agences de voyage en ligne comme Booking.com ou Agoda.com. La même année, Expedia Group a acquis Trivago qui détient aussi plusieurs agences de voyage comme Expedia ou Hotels.com. Chacun des sites de comparaison de prix référence ainsi des agences de voyage affiliées au même groupe que le sien, des agences qui appartiennent au groupe concurrent, des agences de voyage indépendantes et parfois le site de vente directe de l'hôtelier. En plus de mettre en avant les canaux de vente qui paient les commissions les plus importantes, les sites de comparaison de prix pourraient également être tentés de favoriser les agences de voyage affiliées au même groupe, flouant les consommateurs qui pensent généralement que les canaux les plus visibles sont ceux offrant les prix les plus compétitifs.

Contributions

A travers l'exemple de l'industrie hôtelière, cette thèse étudie certains des nouveaux enjeux soulevés par la numérisation de l'économie.

Le premier chapitre porte sur l'estimation du degré de substitution entre les canaux de distribution d'une chaîne d'hôtels. Il s'agit d'un paramètre crucial dans la définition du marché retenue par les autorités de concurrence, mais également d'un indicateur permettant d'évaluer le rapport de force entre hôteliers et intermédiaires. Sur le marché de la vente en ligne des hôtels, composé des deux plus grandes agences de voyage (Booking.com et Expedia) et du canal direct de distribution de la chaîne, la question est de savoir si la vente directe constitue une alternative crédible à la vente via les agences de voyage en ligne. Autrement dit, si un hôtelier décide d'arrêter son référencement sur une plateforme, qu'adviendra t-il des consommateurs qui réservaient l'hôtel par ce canal ? Les données utilisées pour répondre à cette question contiennent les prix et volumes vendus par une chaîne d'hôtels scandinave et se différencient ainsi de la plupart des articles précédents sur le secteur (Hunold et al. (2018); Larrieu (2019); Mantovani et al. (2020)) qui reposent sur des données de prix affichés collectées sur Internet. L'atypicité de ces données permet d'avoir recourt à l'élaboration d'un modèle structurel de demande dans lequel la décision d'achat d'un consommateur est micro-fondée. Cela permet notamment de s'interroger sur le nombre de canaux parcourus lors du processus de choix et d'achat des clients. En particulier, une hypothèse simplificatrice retenue est qu'un consommateur cherche puis

[18] ACCC website *"Trivago misled consumers about hotel room rates"*

achète via le même canal. Les résultats d'estimation indiquent que lorsqu'un hôtelier décide volontairement de ne plus être référencé sur Booking.com, une première moitié des consommateurs (53%) qui réservaient l'hôtel via la plateforme, se tourne vers un hôtelier concurrent tandis que parmi la seconde moitié qui continue de venir à l'hôtel, trois fois plus de clients ont acheté via Expedia que sur le site de l'hôtel. Ces résultats suggèrent qu'en moyenne, après avoir choisi leur destination, les consommateurs sont plus fidèles à une plateforme qu'à un hôtel et que la concurrence entre hôteliers est forte. Cela confirme, en outre, que le pouvoir de négociation d'un hôtelier face à une plateforme leader de la vente en ligne reste relativement faible.

Les hypothèses adoptées et la qualité du modèle sont testées. Dans ce but, un événement particulier ayant eu lieu en Norvège en 2012 est exploité : le non renouvellement des accords entre plusieurs grandes chaînes hôtelières et Expedia suite à des litiges sur les termes de ces contrats (principalement des problèmes de clauses de parité et de commissions). Les résultats de simulation du modèle précédemment estimés sont ainsi confrontés aux données réellement observées, dans le cas où la plateforme boycottée serait Expedia. Après avoir contrôlé par l'évolution des caractéristiques de la demande au cours du temps, les prédictions sur les prix et les parts de marché sont très proches des données réellement observées. Cela permet de conclure que si les hypothèses retenues ne sont pas parfaitement représentatives du processus de choix des consommateurs, elles autorisent une approximation fidèle des comportements observés.

Le faible pouvoir de négociation laissé aux hôteliers face aux agences de voyage en ligne illustre l'emprise que des plateformes comme Booking.com ou Expedia prennent sur le marché. Si à l'origine elles représentaient un moyen pour l'hôtelier d'atteindre davantage de demande, elles sont aujourd'hui devenues un outil de recherche incontournable pour les consommateurs, à tel point que certains leur sont fidèles et limitent leur connaissance du marché aux produits ou services qui y sont référencés. Désormais, pour appartenir à l'ensemble de choix des consommateurs, l'hôtelier se doit à la fois d'être visible sur ces plateformes mais également d'y être bien positionné : les consommateurs sont en effet plus enclins à rechercher des informations détaillées (donc cliquer) sur les hôtels les mieux classés (Ursu (2018)). Tirant partie de cette course à la visibilité, Booking.com a développé le programme *Partenaire Préféré* offrant à ses adhérents un meilleur classement sur son site internet en échange d'un taux de commissions plus élevé et de l'application par l'hôtelier d'une clause de parité étendue.

Le second chapitre de cette thèse évalue l'effet de l'adoption du programme original *Partenaire Préféré* sur les volumes de ventes et les prix des canaux en ligne (Booking.com, Expedia et le site internet) des hôtels qui y ont souscrit. Afin d'identifier l'impact d'un tel programme, il est préalablement nécessaire d'identifier les principales composantes à l'origine de la formation des prix et de définir la demande qui s'adresse à un hôtel. La stratégie tarifaire de l'hôtelier est modélisée sous la forme d'une régression hédonique dans laquelle le prix est défini comme une fonction linéaire des caractéristiques du service : à la fois celles de la chambre (gamme, nombre de personnes, durée, etc.) et celles de l'hôtel (équipements, localisation, etc.). En comparaison des travaux déjà existants sur le secteur de l'hôtellerie (Thrane (2007); Law et al. (2011); Abrate and Viglia (2016)), la modélisation de ce chapitre inclut les caractéristiques spécifiques des réservations, comme le canal de distribution utilisé ou la possibilité d'annulation et de remboursement sans frais, ainsi que les taux d'occupation anticipés et réels, variables clés dans la prise de décision tarifaire d'un hôtelier. Dans une telle estimation, la difficulté réside dans la simultanéité de la détermination du prix et de la demande, l'un influençant l'autre. En réponse à ce problème, dans la régression de prix, le taux d'occupation est instrumenté par le prix des mêmes vendeurs sur un autre segment de marché tandis que dans la régression de la demande, le prix de la chambre d'un hôtelier est instrumenté par le coût salarial ainsi que par le prix d'une chambre identique dans le même pays mais dans une ville différente.

Dans un second temps, les effets du programme *Partenaire Préféré* (adopté par 6 hôtels parmi 22) sur les prix et les volumes de vente sont estimés par un modèle de différence de différences. Le recourt à cette méthode d'estimation nécessite la création d'un groupe de contrôle, comprenant des hôtels aux caractéristiques proches mais n'ayant pas adopté le programme. Le faible échantillon d'hôtels observés conduit à rejeter la méthode d'appariement par score de propension et à privilégier une sélection de 8 hôtels, parmi les 16 hôtels non traités, par la méthode des plus proches voisins. Les résultats d'estimation montrent que l'adoption du programme conduit à la hausse des prix de tous les canaux de vente. Cette hausse a plusieurs origines. Elle est d'abord liée aux commissions plus élevées payées par les hôtels du programme *Partenaires Préférés* qui se répercutent dans les prix finaux. D'autre part, elle peut provenir de l'accroissement de la demande lié au gain de visibilité : les hôteliers anticipent qu'il est possible de tarifer plus cher sans perdre de consommateurs. Enfin l'application de la clause de parité inhérente au programme incite les plateformes à fixer des commissions plus élevées (Boik and Corts (2016)) qui se répercutent de nouveau dans les prix affichés. Du côté de la demande, les résultats montrent que l'adoption du programme accroît le volume des ventes réalisées via les plateformes tandis qu'elle semble réduire ceux du site internet de l'hôtel. Le gain de visibilité sur Booking.com bénéficierait ainsi également à Expedia mais cannibaliserait une partie des ventes sur le site direct.

La dernière partie de ce second chapitre compare des réservations en tout point identique et montre que, si la clause de parité semble généralement être appliquée sur les tarifs remboursables, l'adoption du programme coïncide pour ses hôteliers adhérents avec l'application d'une clause de parité de prix étendue sur les réservations non remboursables, qui n'était jusqu'alors pas mise en oeuvre. Le gain de visibilité et la hausse des commissions générés par le programme étant des éléments communs à ces deux types d'offre, un modèle de triple différence est utilisé afin d'isoler l'effet de l'application de la clause de parité contenue dans la hausse du prix observée suite à l'adoption du programme. Les résultats suggèrent que l'application de la clause de parité n'a pas d'effet significatif sur le niveaux de prix de chaque canal en Suède et au Danemark, confirmant les premières conclusions de Mantovani et al. (2020). D'autre part, lorsqu'il sont significatifs, ce qui est le cas en Norvège, les résultats indiquent que la clause conduit bien à la convergence des prix identifiée dans la littérature théorique mais qu'elle s'effectue en particulier par une baisse des prix sur Booking.com, plateforme à l'origine du programme. Ces estimations viennent ainsi nuancer les effets anti-concurrentiels des clauses de parité de prix et la principale théorie du préjudice développée par les autorités de concurrence jusqu'à aujourd'hui.

Du point de vue du consommateur, un hôtelier engagé dans le programme *Partenaire Préféré* est signalé par un pouce jaune à la droite de son nom. Si trop peu de consommateurs remarquent puis s'interrogent sur la signification de cet icône, une proportion encore plus faible a conscience de ses conséquences, notamment sur les prix qui sont affichés. Si la numérisation des marchés permet aux consommateurs de chercher et d'acheter plus rapidement, elle ne résout pas le manque de transparence, voire crée également de l'opacité. Un autre exemple est celui des opérations de fusions et d'acquisitions de petites start-ups par de plus gros groupes, menant à une concentration croissante des marchés numérisés, dont le tourisme en ligne ne fait pas exception. Pour preuve, aux Etats-Unis, 95% des agences de voyage en ligne appartient à deux groupes et 74% des consommateurs l'ignorent[19]. Si les entreprises se sont longtemps développées horizontalement, en rachetant leurs concurrentes, des stratégies verticales ont récemment été mises en oeuvre. Les groupes sont maintenant propriétaires à la fois de plusieurs agences de voyage en ligne mais également de sites de comparaison de prix des canaux de distribution, dont ces agences font partie. Ces opérations d'acquisition posent des questions de neutralité et poussent à s'interroger sur les incitations des sites de comparaison de prix à toujours mettre en avant les offres les moins chères.

Le troisième chapitre est consacré à l'étude de l'impact de l'intégration verticale de Kayak et de plusieurs agences de voyage en ligne comme Booking.com ou Agoda.com au sein du groupe Booking Holding. Kayak propose aux consommateurs deux types

[19] American Hotel & Lodging Association

de classement : un classement vertical des hôtels pour une destination donnée et, pour chaque hôtel, un classement horizontal des différents canaux de vente proposant une offre, c'est à dire un prix. La question est de savoir si Kayak intègre dans ces deux algorithmes une préférence de classement pour les agences de voyage qui appartiennent au groupe Booking Holding, sa maison-mère. Les données utilisées pour ce chapitre ont été collectées automatiquement par un robot, formant une base de plus de 17 millions d'observations. Celle-ci contient les résultats de recherche pour plus de 2 000 requêtes lancées sur Kayak, en particulier le classement de 1 784 hôtels distincts à Paris et les prix et classements de 828 canaux de vente différents (plateformes d'intermédiation mais aussi sites internet directs des hôtels). L'intérêt de ces données est qu'elles représentent fidèlement la page internet vue par un consommateur, sans que les classements automatiquement affichés par Kayak ne soit biaisés par la redondance d'une adresse IP ou la présence de cookies.

L'analyse est menée autour de deux conjectures. La première stipule que pour un hôtel donné (classement horizontal), Kayak met en avant les canaux de vente appartenant au même groupe. La seconde ajoute que pour une recherche donnée (classement vertical), Kayak met avant les hôtels pour lesquels ses canaux de vente sont les plus compétitifs, i.e. les moins chers. Concernant le classement horizontal, des statistiques préliminaires montrent que, toutes choses égales par ailleurs (en particulier le prix), pour un hôtel donné, les offres des agences de voyage du groupe Booking Holding sont plus souvent mises en avant que celles des autres canaux de vente. Ces résultats sont confirmés par des modèles de régressions linéaires à effets fixes, sur les hôtels et les requêtes, contrôlant, en outre, l'effet intrinsèque lié à la popularité des canaux de vente et des hôtels. Du côté du classement vertical, les résultats des régressions à effets fixes montrent qu'un hôtel pour lequel le groupe Expedia est le vendeur le plus compétitif a un classement dégradé (d'en moyenne sept positions) sur la page internet, confirmant ainsi les conjectures formulées.

Des tests de robustesse complètent cette analyse. D'abord, les effets de visibilité au sein d'un même groupe sont différenciés. Il est observé qu'au sein du groupe Booking Holding, la plateforme Booking.com est particulièrement mise en avant. En effet, Agoda.com, appartenant au même groupe, est relativement moins visible à prix donnés. Cela peut s'expliquer par le business modèle d'Agoda.com qui s'apparenterait davantage à de la revente qu'à de l'agence. D'autre part, cette plateforme est très peu utilisée en France, ce qui expliquerait que le groupe préfère mettre en avant une plateforme plus populaire. Concernant le classement vertical, les hôtels semblent être moins bien classés lorsque Expedia.fr est la plateforme proposant le prix le moins cher. Cela n'est pas surprenant, cette dernière étant la principale plateforme concurrente de Booking.com. Le second test de robustesse porte sur l'affiliation d'un hôtel à une chaîne. Il est intéressant de constater que si Kayak semble défavoriser les agences de voyage qui n'appartiennent pas à son groupe sur le classement horizontal, le site propre des hôteliers est, lui, relativement protégé, à condition qu'il ne s'agisse pas d'un hôtel appartenant à chaîne. De même, si le site internet d'un hôtel affilié à une chaîne est le canal de vente le plus compétitif, alors cet hôtel est déclassé verticalement. Ces observations s'expliquent par une politique des sites de comparaison de prix visant à encourager les hôteliers indépendants à se référencer sur Kayak sans passer par l'intermédiaire des agences de voyage en ligne. Enfin, des tests de robustesse suggèrent que les biais de classement observés varient au cours du temps, notamment avec l'évolution de la régulation liée aux clauses de parité des prix sur l'intervalle temporel couvert par la base de données. En juillet 2015, Booking.com s'est engagé à passer ses clauses de parité étendues à restreintes puis en août 2015, l'application de loi Macron a aboli toutes les clauses de parité de prix dans le secteur de l'hôtellerie. Toutefois, l'existence d'évènements concomitants, tels que les attaques terroristes de 2015 et 2016 à Paris, empêche de conclure à une relation de cause à effet et pousse à considérer ces résultats avec mesure.

Une difficulté supplémentaire rencontrée dans ce chapitre réside dans l'absence d'information quant à la profitabilité des différents canaux de vente pour Kayak. S'il est admis que les commissions payées par les agences de voyage en ligne pour apparaître sur les sites de comparaison de prix ont un impact sur les classements, les valeurs des commissions, susceptibles de varier selon les hôtels mais également au cours du temps, restent à ce jour indéterminées. Les résultats observés sont donc à interpréter avec prudence, compte tenu de la présence d'éventuelles commissions, qui, si elles existaient, aggraveraient le biais estimé. Bien que les résultats soient spécifiques à l'intégration verticale au sein de Booking Holding, leur intérêt est renforcé par leur application au-delà de l'industrie hôtelière, dès lors que se pose la question du référencement; comme récemment dans le cas dans l'abus de position dominante de Google Search[20] ou encore pour les investigations à l'encontre de la place de marché d'Amazon dans le cadre de l'attribution de sa *buy box*[21].

En conclusion, certains des résultats des chapitres sont mis en regard afin d'avancer trois des principaux axes de réflexion à retenir après la lecture de cette thèse : la question de la substitution entre canaux de vente, celle des effets des clauses de parité de prix et enfin, la question de la transparence des algorithmes de classements.

20 European Comission decision (2017) ; Case AT.39740 — Google Search (Shopping)

21 European Comission (2019) *"Antitrust: Commission opens investigation into possible anti-competitive conduct of Amazon"*

Note

Les trois chapitres de cette thèse sont des articles de recherche indépendants. Cela explique pourquoi certaines informations sont redondantes et que le terme "article" est parfois utilisé en lieu et place du terme "chapitre".

Les deux premiers chapitres de cette thèse ont été co-écrits avec Arthur Cazaubiel (CREST, ENSAE Paris, Institut Polytechnique de Paris), Bjørn Olav Johansen (University of Bergen) et Thibaud Vergé (CREST, ENSAE Paris, Institut Polytechnique de Paris, University of Bergen).

Le troisième chapitre de cette thèse a été co-écrit avec Matthias Hunold (University of Siegen), Reinhold Kesler (University of Zurich), Ulrich Laitenberger (Telecom Paris, Institut Polytechnique de Paris) et Thomas Larrieu (CREST, École polytechnique, Institut Polytechnique de Paris).

Bibliography

Abrate, G. and Viglia, G. (2016), ‘Strategic and tactical price decisions in hotel revenue management’, *Tourism Management* **55**, 123–132.

Armstrong, M. (2006), ‘Competition in two-sided markets’, *The RAND Journal of Economics* **37**(3), 668–691.

Boik, A. and Corts, K. (2016), ‘The effects of platform most-favored-nation clauses on competition and entry’, *Journal of Law and Economics* **59**, 105–134.

Fletcher, A. and Hviid, M. (2014), ‘Retail price mfns: Are they rpm ’at its worst”, *ESRC Centre for Competition Policy Working Paper* pp. 14–5.

Foros, U., Kind, H. J. and Shaffer, G. (2014), ‘Turning the page on business formats for digital platforms: Does the agency model soften competition?’, *Available at SSRN 2317715* .

Hart, O., Tirole, J., Carlton, D. W. and Williamson, O. E. (1990), ‘Vertical integration and market foreclosure’, *Brookings papers on economic activity. Microeconomics* **1990**, 205–286.

Hunold, M., Laitenberger, U. and Schlütter, F. (2018), ‘Evaluation of best price clauses in hotel booking’, *International Journal of Industrial Organization* **61**, 542–571.

Johansen, B. O. and Vergé, T. (2017), ‘Platform price parity clauses with direct sales’, *University of Bergen* .

Johnson, J. P. (2017), ‘The agency model and mfn clauses’, *The Review of Economic Studies* **84**(3), 1151–1185.

Larrieu, T. (2019), ‘Pricing strategies in online market places and price parity agreements: evidence from the hotel industry’, *mimeo* .

Law, R., Ye, Q. and Zhang, Z. (2011), ‘Determinants of hotel room price: An exploration of travelers’ hierarchy of accomodation needs’, *International Journal of Contemporary Hospitality Management* **23**(7), 972–981.

Mantovani, A., Piga, C. A. and Reggiani, C. (2020), 'Online platform price parity clauses: Evidence from the eu booking. com case', *Available at SSRN 3381299* .

Mathewson, G. F. and Winter, R. A. (1984), 'An economic theory of vertical restraints', *The RAND Journal of Economics* pp. 27–38.

Rey, P. and Vergé, T. (2010), 'Resale price maintenance and interlocking relationships', *The Journal of Industrial Economics* **58**(4), 928–961.

Rey, P. and Vergé, T. (2016), 'Secret contracting in multilateral relations'.

Rochet, J.-C. and Tirole, J. (2003), 'Platform competition in two-sided markets', *Journal of the european economic association* **1**(4), 990–1029.

Thrane, C. (2007), 'Examining the determinants of room rates for hotels in capital cities: the oslo experience', *Journal of Revenue and Pricing Management* **5**(4), 315–23.

Ursu, R. M. (2018), 'The power of rankings: Quantifying the effects of rankings on online consumer search and purchase decision', *Marketing Science* **37**(4), 530–552.

Wang, C. and Wright, J. (2016), 'Platform investment and price parity clauses', *Available at SSRN 2848574* .

Wang, C. and Wright, J. (2020), 'Search platforms: Showrooming and price parity clauses', *the RAND Journal of Economics* **51**(1), 32–58.

Chapter 1

Substitution Between Online Distribution Channels: Evidence from the Oslo Hotel Market *

> *“But the challenge of market definition isn’t only about geographic markets. The changes that we’re going through – especially digitisation – are also creating new challenges for defining product markets – for working out which products consumers are willing to substitute for each other."*
>
> - Margrethe Vestager (2019), *Commissaire européenne à la concurrence*

* Published under the following reference: Cazaubiel, A., Cure, M., Johansen, B. O. & Vergé, T. (2020), ‘Substitution between online distribution channels: Evidence from the oslo hotel market’, *International Journal of Industrial Organization* **69**, 102577. Online access here.

1 Introduction

Retail e-commerce sales have been rapidly growing over the last 20-25 years. According to Statista, online sales will reach 2.8 trillion US dollars worldwide in 2018, having almost doubled in the last three years. In some markets such as music, books or travel, a large majority of sales are now made online rather than offline. Even groceries are now more commonly bought online.

The rapid growth of online retailing has led economists and competition agencies to look at the importance and impact of multi-channel distribution, and at the degree of substitution between online and offline sales.[1] Among others, Gentzkow (2007) and Pozzi (2013) analyze the cannibalization effects of online distribution on offline sales. Gentzkow (2007) shows that the introduction of a digital version of the *Washington Post* reduced sales of the print edition. Pozzi (2013) concludes that the introduction of an online shopping service by a large US grocery retailer had a limited cannibalization effect on brick-and-mortar sales while increasing total revenues. Another important question has been to identify whether online retailing has led consumers to benefit from increased competition, i.e., to focus on across-firm substitution (see for example Prince (2007), Duch-Brown et al. (2017) and Ellison and Fisher Ellison (2018)).

Substitution between online and offline distribution is also an important issue for competition authorities. In merger control, delineating product markets is essential to assess the competitive impact of mergers and this now frequently involves identifying whether online sales should be part of the same relevant market as offline sales.[2] The role of online sales and the interaction between brick-and-mortar, click-and-mortar, and pure online players has also been a major issue when revising the European rules applicable to vertical agreements.[3] Many cases involving restraints related to online sales have been evaluated by competition agencies in the last decade: restriction of online sales in selective distribution networks [e.g., Pierre Fabre (France, 2007 and CJEU, 2011)], dual pricing or resale price maintenance [e.g., BSH (Germany, 2013) and United Navigation (UK, 2015)], exclusive territories or geo-blocking [e.g., Sector inquiry into e-commerce (European Commission, 2016)].[4]

More recently, the policy debate has shifted to the impact of specific types of ver-

1 For a review of the early literature, see Lieber and Syverson (2012).

2 See for example recent cases in traditional retailing [e.g., Picwic/Toys'R'Us (France, 2019)], mobile payments [e.g., Telefonica UK/Vodafone UK/Everything Everywhere (European Commission, 2012)] or sales of books [e.g., Ahold/Flevo (European Commission, 2012)].

3 See Commission Regulation 330/2010 of 20 April 2010 on the application of Article 101(3) of the Treaty on the Functioning of the European Union to categories of vertical agreements and concerted practices, *Official Journal of the European Union*, L102, pp. 1-7, and *Guidelines on Vertical Restraints*, Commission Notice, C(2010) 2365.

4 For a detailed review of competition issues and cases in Europe, see Friederiszick and Glowicka (2016).

tical restraints in online retailing, restraints usually related to the role of third-party platforms. Recent cases have involved restrictions imposed by manufacturers on online retailers with respect to the use of third-party platforms [e.g., Coty (Germany, 2014 and CJUE, 2017) or Adidas and Asics (Germany, 2014)], and by platforms on suppliers with respect to pricing, such as price parity (or MFN) clauses [e.g., eBooks (European Commission, 2017), Amazon (UK and Germany, 2013)].

Throughout Europe, platform price parity clauses have been the subject of several investigations in the market for online booking platforms/online travel agencies (OTAs). Such price parity clauses imposed by a platform to suppliers constrain the supplier's ability to freely set prices on different distribution channels. A wide price parity clause covers all potential channels, that is, the clause prevents a supplier from selling a product at a price lower than the price charged on the platform imposing it (and this applies anywhere else including on the supplier's own website). When all platforms used by a supplier impose wide price parity clauses, the supplier has to set the same price everywhere (it may only sell at a higher price on its own website). By contrast, a narrow price clause only constrains the price set for the supplier's direct sales: the supplier can freely set prices on different platforms, but it cannot sell on its own website at a lower price than the price set on the platform imposing the constraint. Price parity clauses thus limit the supplier's ability to set low prices for direct sales. In addition, when the clauses are wide, they may also lead to uniform prices on all platforms. Competition authorities in Europe consider that wide price parity clauses reduce incentives for platforms to compete on commission rates because they cannot expect suppliers to lower prices on cheaper platforms.

In Germany, the Bundeskartellamt prohibited price parity clauses imposed by HRS (December 2013) and Booking (December 2015). In April 2015, the French, Italian and Swedish competition agencies simultaneously accepted commitments offered by Booking to remove any availability requirements from their contracts and to switch from wide to narrow price parity clauses.[5] Although it did not formally offer commitments to competition agencies, Expedia announced similar changes to its contracts throughout Europe.[6,7]

Market definition has been an important part of the debate, with agencies ultimately concluding that the hotels' direct sales do not belong to the same market as

5 See Decision of 15 April 2015 by the Swedish Competition Authority in Case 596/2013.

6 The French (2015), Austrian (2016) and Italian (2017) parliaments have since voted in favor of legislation prohibiting any form of price parity (or price control by the platforms) for hotel room bookings.

7 In October 2015, the Swiss Competition Commission prohibited the use of wide price parity clauses by Booking, Expedia and HRS but allowed them to adopt narrow price parity clauses. In September 2016, the Australian Competition and Consumer Commission accepted commitments offered by Expedia and Booking to amend the price and availability parity clauses in their contracts and to switch from wide to narrow price parity clauses.

sales made through OTAs. Authorities have indeed taken the view that OTAs offer a bundle of services that includes search and comparison as well as the possibility to book online, whereas hotels' websites only offer the opportunity to book. They also concluded that hotels view OTAs more as a complement than as a substitute to their own direct sales.

The issue of substitution between online channels also has important theoretical implications when considering the effects of price parity clauses. Boik and Corts (2016) (in a context with a monopolist supplier) and Johnson (2017) (with competing suppliers) both show that when suppliers sell through competing platforms, price parity clauses lead to higher commissions and thus higher final prices.[8] However, their results rely on the assumptions that the platform commissions either are linear tariffs (i.e., a fixed price per sale) or based on revenue-sharing. Once these assumptions are relaxed, the effects of price parity clauses may well be different. For example, Rey and Vergé (2016) show that with non-linear commissions, price parity clauses do not affect final prices, but only affect the division of profits. Johansen and Vergé (2017) consider linear commissions but assume that suppliers can also reach final consumers directly. In such a setting, price parity clauses have an ambiguous effect on commissions, final prices, and suppliers' profits. In particular, when inter-brand competition (i.e., competition between suppliers) is sufficiently fierce, price parity clauses may well lead to lower commissions and prices, while simultaneously increasing suppliers' and platforms' profits. However, their result relies on the assumption that it is a viable option for a supplier to delist from one of the platforms. This requires that, when delisting from a platform, a sufficiently large share of the lost sales are indeed recaptured through the direct channel and not exclusively through the rival platforms.[9]

In this paper, we use an exhaustive database of bookings in 13 Oslo hotels (all belonging to the same chain) to evaluate the degree of substitution between online distribution channels, including the two largest OTAs (Booking and Expedia) and the chain's own online distribution channel. We can then try to check whether selling directly constitutes a credible alternative to selling through OTAs. Contrary to recent papers that have focused on the effects of price parity clauses in this industry by using scrapped price data from metasearch engines (see, e.g., Hunold et al. (2018), Mantovani et al. (2020) and Larrieu (2019*b*)), we use a large dataset of actual bookings to

[8] Larrieu (2019*a*) allows for balanced negotiations between suppliers and platforms and obtains qualitatively similar results.

[9] See also Edelman and Wright (2015), Wang and Wright (2016) and Wang and Wright (2020) who show that, despite reducing the risk of free-riding by platforms ("showrooming"), price parity clauses usually lead to higher prices or inefficient investment. However, in their search setting, delisting from a platform is never a profitable strategy for suppliers: Because all sales are made through the most efficient platform in equilibrium, a supplier would lose all of its consumers by not listing on this platform.

estimate a nested logit demand model that allows us to evaluate substitution patterns between online distribution channels. Our results suggest that, while a substantial share of consumers seem to be loyal to the OTAs, and would switch to the other hotels (i.e., our "outside good") in case of the hotel chain's decision to delist from a platform (or after a substantial price increase by the hotel chain on the same platform), the chain's direct sales channel remains a reasonably credible alternative to the OTAs. Still, among the consumers that would continue to book a room at the same hotel (after the hotel's decision to delist from one of the OTAs), only a minority (about one in four) would book directly from the hotel rather than from the competing OTA.

We then use the demand estimates to uncover the hotels' marginal costs through a structural model of price competition with differentiated products. We thus solve the system of first-order conditions, in a Bertrand-Nash model where hotels compete in prices, each hotel setting prices for each channel it uses: we thus consider an agency model where hotels keep control of the final prices and pay commissions to OTAs that they use as service providers. We can then use these marginal cost and demand estimates to run counterfactual simulations. In particular, we simulate the effects of a common decision by the 13 hotels to stop using on of the distribution channels (e.g., delisting from Expedia's platforms). Making use of the actual chain's decision to delist from Expedia, we can compare simulated and actual effects of such an event on prices and market shares. In that sense, we try to contribute to the debate on the effectiveness of structural IO models initiated by Peters (2006), Angrist and Pischke (2010) and Nevo and Whinston (2010).[10] Comparing the simulated and observed outcomes, we observe discrepancies in terms of prices and market shares. Following Peters (2006), we thus try to identify sources for these differences and see how to improve the counterfactual simulation. Accounting for changes in the product characteristics changes the simulated outcome and provides results that are comparable to the effects of the actual delisting decision.

The rest of the paper is organized as follows. After presenting our dataset and the specific context in which the 13 hotels operated during the sample period (Section 2), we proceed to the estimation of our nested logit demand model and derive substitution patterns between online distribution channels (Section 3). We then use the estimated demand parameters and a structural pricing model to obtain per-channel marginal costs (Section 4). We then perform a counterfactual analysis and compute equilibrium prices and market shares assuming that all hotels decide to stop selling through one channel. Taking advantage of the hotels' decision to delist from Expe-

[10] For recent evidence on the accuracy of merger simulation methods, see among others, Weinberg (2011), Weinberg and Hosken (2013), Björnerstedt and Verboven (2016) and Miller and Weinberg (2017).

dia in 2013, we then compare the simulated outcome to the observed data (Section 5). Section 6 concludes.

2 Data and Context

2.1 Data

We use an exhaustive dataset of all bookings made over almost four years in 13 hotels located around Oslo (Norway).[11] These hotels all belong to one of the leading hotel chains active in Norway.

Our initial dataset includes more than 1.2 million observations (i.e., bookings). This dataset has been directly extracted from the hotel chain's information system. It includes all bookings made by consumers through all distribution channels between January 2013 and November 2016. For each booking, we observe:

- The booking date as well as the arrival and departure dates. This allows us to compute the length of stay as well as advance purchase (i.e., how many days prior to arrival the room has been booked).
- The room type (e.g., standard, superior, junior suite, ...).
- The number of guests.
- The channel through which the room was booked.
- The price paid by the consumer as well as the rate code associated with the tariff.

We use our exhaustive dataset and existing information on the number of rooms at each hotel to compute occupancy rates at any point in time. Specifically, we compute the variable $OR_{h,t,x}$, which is the occupancy rate at hotel h at date t, computed at date $t-x$ (i.e., x days in advance). As x becomes smaller and we get closer to the date t, we thus expect the occupancy rate $OR_{h,t,x}$ to increase. More formally, $OR_{h,t,x}$ is the number of bookings made at date $t-x$ and earlier, for all stays that include a night at date t, divided by the total number of rooms at the hotel. We compute these occupancy rates for all values of x between 0 and 30 (i.e., we compute the occupancy rate daily up to one month before arrival). The ratio $1-OR_{h,t,x}$ thus indicates which proportion of the rooms (for a stay at date t) were still available x days in advance.

Although we use the full dataset to compute occupancy rates, we carry out our econometric analysis on a subset of bookings that we consider to be homogeneous

[11] Our hotels are located either in the municipality of Oslo or close to the city boundaries, with the exception of two airport hotels (at Oslo-Gardermoen Airport).

Channel	2013		2014		2015		2016	
	#	%	#	%	#	%	#	%
Offline	197,617	91.0%	190,462	89.3%	192,510	84.8%	185,471	82.0%
Online	19,528	9.0%	22,923	10.7%	34,433	15.2%	40,684	18.0%
Direct Online (DON)	8,571	43.9%	11,275	49.2%	17,952	52.1%	17,910	44.0%
Booking (BOO)	10,957	56.1%	11,648	50.8%	13,419	39.0%	11,663	28.7%
Expedia (EXP)	0	0.0%	0	0.0%	3,062	8,9%	11,111	27.3%
Total	217,145	100%	213,385	100%	226,943	100%	226,155	100%

Table 1.1: Share of bookings made through the offline and online channels

enough. We restrict attention to bookings made for one or two guests, for one room only, for no more than a week and exclusively for standard or superior rooms (thus excluding business rooms or suites). In addition, we only consider bookings made at most 30 days prior to arrival. As a first step, this helps to ensure that the bookings we observe for a specific date are mostly made under the same regime (i.e., either during or after the boycott described in section 2.2). Yet, for the first 30 nights after the boycott started, and for the first 30 nights after it ended, we still observe that some of the reservations (for a specific night) were made during the boycott, while the rest of the reservations (for the same night) were made either before or after the boycott. Thus, to ensure that all bookings are made under the same regime, we want to exclude all reservations made for any of the 30 first nights after the boycott had started, and all reservations made during the boycott, that have an arrival date that falls after the hotel has started listing again on Expedia. This whole selection process eliminates about 28 % of the observations that account for about 50 % of the hotels' revenue, leaving us with 885,249 observations[12].

Finally, we are only interested in the substitution between online sales channels, and more specifically between the chain's own booking platform (which we refer to as the direct online channel) and the two largest online travel agencies, namely Booking and Expedia. As shown in Table 1.1, the three online channels account for between 9 % (in 2013) and 18 % (in 2016) of all bookings (average of 13% over the full period January 2013–November 2016). Although we use information from bookings made through other channels[13] as instruments in our demand estimation, we essentially focus on the 117,760 online bookings. Table 1.1 also shows that, among online bookings, the direct channel accounts for nearly half of the sales, Booking accounts for about 40 % on average (with a share above 50 % during the boycott period but closer to 30 % in 2016). Expedia's overall market share is just over 12 % of all online bookings but this is biased because of the long period during which our 13 hotels decided not to list

12 We provide in Appendix – see Table 1.18 – a detailed breakdown of the elimination stages specifying the share of bookings and revenues for each stage of the process.

13 Although other bookings are made through different types of booking channels such as travel agencies or B2B contracts, we refer to such bookings as made "offline."

their rooms on Expedia (see section 2.2). In 2016, Expedia's market share was closer to Booking's market share.

Channel	Booking	Direct	Expedia	Offline
Price (NOK)	1,123	1,024	1,279	1,074
Advance	9.2	9.6	8.0	7.3
Nights	1.7	1.5	1.5	1.5
Persons	1.4	1.3	1.3	1.2
Superior room	6%	10%	10%	13%
Week-end	34%	31%	31%	22%
Occupancy rate	85%	80%	86%	80%

Table 1.2: Summary statistics of booking characteristics

Table 2.5 presents some summary statistics of booking characteristics (for all online and offline bookings). Overall, it appears that online prices are lower on the hotels' own websites (about 100 NOK $\sim$ 13\$) than on Booking or Expedia (offline prices are somewhere in between). Consumers tend to book earlier online than offline (conditionally on booking less than a month prior to arrival). Given that most bookings are made relatively late (just over one week before arrival on average), it is not surprising that the occupancy rate as seen at the date of the booking (i.e., proportion of rooms already booked) is relatively high, between 80 % and 86 % on average. We also observe that online bookings include weekend nights more often than offline bookings. This should not be surprising, as our dataset includes corporate rates, and bookings made using these corporate tariffs are all part of the offline bookings. Finally, rooms are booked for one to two nights and for 1.3 persons on average.

We also collected some hotel characteristics, and this additional data includes:

- Number of rooms.
- Precise hotel location (as well as distance from city center and Oslo-Gardermoen Airport).
- Star rating as well as existence of specific amenities (bar, restaurant, fitness and/or wellness center).
- Consumer reviews have been scrapped from TripAdvisor. For our 13 hotels, these reviews have been collected daily for the whole period. Each day, we observe for each hotel the last five ratings (on a 1-to-5 scale), the current average rating and the total number of reviews to date.

Table 2.4 presents summary statistics of the hotel characteristics. Our sample includes only 3 and 4-star hotels (the majority are 3-star hotels) that are relatively large

	Mean	Median	Min	Max
Number of rooms	195	164	103	435
Star rating	3.3	3	3	4
Last TripAdvisor Rating (1-5 scale)	3.8	3.8	3.4	4.3
Bar	0.62	–	0	1
Restaurant	0.54	–	0	1
Fitness/Wellness	0.38	–	0	1
Distance to city center (km)	7.6	1.0	0.5	36.2
Distance to airport (km)	33.5	36.9	4.4	37.8

Table 1.3: Summary statistics of hotel characteristics

(about 200 rooms on average, all above 100 rooms). Hotels located in the city center tend to be smaller and centrally located, whereas the two hotels located in the vicinity of Oslo-Gardermoen Airport are the largest (both with more than 200 rooms).

2.2 Context: Delisting From Expedia

During the year 2012, several large hotel chains active in Norway decided not to renew their agreements with Expedia following disputes over the terms of these contracts (most prominently the issues of rate parity and commission fees). First Hotels was the first chain to pull out its inventory from Expedia's platforms and was soon followed by some of the other leading chains such as Nordic Choice, Rica Hotels (later acquired by Scandic), Scandic Hotels and Thon Hotels. By the end of 2013, some of these chains had signed new contracts with Expedia and had started listing again on Expedia's various platforms. Nordic Choice (the largest chain in Scandinavia with more than 160 hotels) reported that Expedia had accepted to cut its commission rate to less than 15 %, a level similar to Booking's commission rate (reported to be around 15 % on average in Europe) and to drop the price parity requirement.[14]

The chain that owns the 13 hotels in our dataset cut its ties with Expedia at the end of 2012, and its inventory stopped appearing on Expedia's platforms as of January 1, 2013. The "boycott" ended in 2015, after almost 3 years, when the hotels started listing again on Expedia's platforms in September and October 2015. Our almost four years of observations thus cover this boycott period (from January 2013 to September/October 2015) as well as a period during which the hotels were listing on Expedia's platforms (from September/October 2015 to November 2016).

For the first month of the dataset (January 2013) none of the 13 hotels are listing on Expedia. However, at the end of the boycott we observe that the different hotels start listing again on Expedia's platforms on different dates. We therefore identify for each hotel the date for which we start observing bookings made through Expedia, and then

[14] See press reports at NewsinEnglish.no and Hotel News Now.

we use this date as the end of the boycott for that hotel.[15]

As mentioned in the previous section, to ensure that all bookings are made under the same regime, we exclude from our sample all bookings with an arrival date in January 2013 (first month of our dataset), as well as all bookings with an arrival date within the first month after each hotel's decision to list again on Expedia. This helps to ensure that all the bookings we observe for a given arrival date are comparable. For example, for a given hotel, if the boycott ended on September 10, 2015, we consider two separate periods for that hotel: The boycott period, which includes all bookings with an arrival date between February 1, 2013 and September 9, 2015, and the post-boycott period, which includes all bookings between October 10, 2015 and November 30, 2016. Table 1.4 shows the date (formally week) that we identify as the end of the boycott for each of the 13 hotels.

Hotel	End of Boycott
Hotel 1	October 29, 2015
Hotel 2	September 10, 2015
Hotel 3	September 10, 2015
Hotel 4	October 15, 2015
Hotel 5	October 15, 2015
Hotel 6	October 15, 2015
Hotel 7	October 8, 2015
Hotel 8	October 15, 2015
Hotel 9	December 24, 2015
Hotel 10	October 22, 2015
Hotel 11	January 8, 2016
Hotel 12	October 15, 2015
Hotel 13	December 24, 2015

Table 1.4: Identifying the end of the boycott period

This long boycott period (33-34 months out of 47 months for which we have data) explains Expedia's low market share (about 12 % of the online bookings). Now that we have precisely identified the boycott period, we can compute markets shares (restricting attention to our three online distribution channels) for the boycott and post-boycott periods separately. Table 1.5 shows each channel's market share during the two periods. Note that we cannot infer from these numbers which distribution channels (if any) were affected by Expedia's return after the boycott, as the market shares do not tell us anything about the underlying volumes. In the next sections, we propose to carefully analyze substitution patterns between these three distribution channels.

15 Formally, we require that all least three bookings are made during the week through Expedia to consider that the hotel is listing again. We check evolution of each hotel's sales through Expedia between July and November 2015, and this methods seems to perfectly identify the boycott end.

Channel	Boycott	Post-boycott
Direct Online (DON)	47 %	42 %
Booking (BOO)	53 %	31 %
Expedia (EXP)	–	27 %

Table 1.5: Online distribution market share for each channel

3 Demand Estimation

In this section, we focus on the final period of our dataset, during which hotels all list on Expedia (as well as on Booking and on their own website). Using a nested logit demand model, we estimate demand on all three online channels during that period and evaluate substitution patterns between online channels.

3.1 Specification

We consider the following multi-level nested logit model (we discuss the outside option below):

1. Consumers first decide whether to buy through one of the OTAs (Booking and Expedia) or directly (through the chain's booking platform). There are thus two groups, the platforms ($g = P$) and the direct sales ($g = D$).

2. If a consumer buys directly, s/he decides in which of the 13 hotels to book. If s/her buys through the platforms, s/he chooses which of the two platforms to use before deciding in which of the 13 hotels to book. Within the group $g = P$, we thus consider two subnests (or distribution channels d), Booking ($d = B$) and Expedia ($d = E$).

On top of the outside option, a consumer can choose among 39 products as a product is a combination of hotel and distribution channel. There are 13 different hotels and three distribution channels (direct sales, Booking and Expedia). An alternative - and simpler - specification could have to assume that the consumer directly decides which distribution channel to use (removing one nesting level). The nesting structure that we propose has the advantage to allow consumers using one platform, say Expedia, to be more likely to switch to another OTA, here Booking, rather than booking directly in the event of delisting from Expedia (or simply a price increase on Expedia).

Consumer i's conditional indirect utility when buying product j in group g and subnest d (with the convention that $d = D$ if $g = D$) at time t (i.e., for a stay starting during week t) is thus given by:

$$u_{ijt} = \underbrace{X'_{jt}\beta - \alpha p_{jt} + \xi_{jt}}_{\equiv \delta_{jt}} + \zeta_{igt} + (1-\sigma_g)\zeta_{igdt} + (1-\sigma_g)(1-\sigma_d)\varepsilon_{ijt}, \tag{1.1}$$

where product j is the combination of a hotel and a distribution channel (i.e., subnest), i.e., $j = (h,d)$. The first part of the function, δ_{jt}, is the mean utility for product j at time t. The mean utility depends on observed characteristics that are included in the vector X_{jt}, which consists of booking characteristics (type of room, advance booking (in days), proportion of weekend travelers, occupancy rate at the time of booking, etc.), and hotel characteristics that may be time-invariant (distance from the city center or Oslo-Gardermoen airport, star rating, restaurant, bar, wellness/fitness center) or not (TripAdvisor ratings). The mean utility also depends on the price of product j at time t, p_{jt}, and on unobserved (to the econometrician but not to consumers) time-specific product characteristics, ξ_{jt}. For the outside good, we normalize this mean utility to zero, i.e., $\delta_{0t} = 0$ for all t.

The other terms consist of deviations from the mean utility and include three random terms: ζ_{igt} is an individual-specific unobserved preference shock common to all products in group g, ζ_{igdt} is an individual-specific unobserved preference shock common to all products in subnest d and ε_{ijt} is an individual/product-specific preference shock. Finally, the nesting parameters are σ_g and σ_d that should satisfy $0 < \sigma_g < \sigma_d < 1$.

If the random terms have distributions that give rise to the nested logit form, the market share system can then be inverted (see, e.g., Berry (1994)) to obtain the following equation for product j in group g and distribution channel d at time t:

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \delta_{jt} + \sigma_g \ln\left(\frac{s_{dt}}{s_{gt}}\right) + \sigma_d \ln\left(\frac{s_{jt}}{s_{dt}}\right), \tag{1.2}$$

where s_{jt} is the market share of product j at time t, s_{0t} is the overall market share of the outside good, s_{dt} is the overall market share of the products in subnest (or distribution channel) d and s_{gt} is the overall market share of the products in the nest (or group) g.

To compute the outside good's market share, we adopt the following strategy: starting with monthly data for the total number of hotel rooms booked in Oslo[16], we divide by four to obtain the total number of rooms booked on average each week for that particular month. We then multiply by the share of online bookings observed each week for our thirteen hotels, to estimate the total size of the online booking market for rooms in Oslo in that particular week. Finally, we multiply by the proportion of three

[16] We use the number of guest nights by month and county for hotels and similar establishments as published by Statistics Norway (*Statistik Sentralbyrå*): `https://www.ssb.no/en/overnatting`.

or four-star hotels in Oslo, i.e., 70%. In Appendix 1.C, we confirm that our results are robust to variations in the outside goods' market share by varying this last multiplier (share of three and four star hotels) between 50% and 90%.

3.2 Instruments

The exercise relies on our ability to consistently estimate equation (1.2). Unfortunately, prices and market shares are endogenously determined and likely to be correlated with product-specific demand shocks that are included in the error terms. Three types of instruments are commonly used to solve such endogeneity problems in demand model estimations: marginal cost shifters, characteristics of rivals' products, and prices in other markets.[17]

Cost shifters are a first common set of instruments. The idea is that costs affect the prices charged to consumers (thus marginal cost shifters and prices are correlated), and that they are uncorrelated with (unobserved) demand shocks. We have therefore collected hourly wages in Norway between 2012 and 2016, and use them as one set of instruments (weighted by the number of rooms to account for hotel size).[18]

We then follow Bresnahan (1987) and assume that demand for a given product (i.e, a hotel in a specific channel) depends not only on the product's own characteristics but also on the characteristics of competing products. However, these characteristics are not likely to be correlated with unobserved demand shocks, because hotels cannot quickly adjust their characteristics (such as star rating and amenities) in response to short-term shifts in demand. We thus use as instruments TripAdvisor ratings of competing hotels in the same market, which are characteristics that change over time. More specifically, we consider, at any point in time, the average across the twelve competing hotels of the average rating for each hotel (for all reviews), the average rating of the last five reviews, and the total number of reviews.

In addition, following Hausman (1996) and Nevo (2001), we instrument the price of a specific product with the average price of other products sold by the same seller. In our case, a seller corresponds to a specific hotel, and we thus use the average price of rooms sold offline to instrument online prices. Prices in different distribution channels are likely to be correlated because they are directly affected by common demand and cost shocks. Moreover, the exclusion condition requires that prices set offline do not affect demand in the online channels. This condition is likely to hold because offline prices essentially consist of walk-in prices, B-2-B contracted tariffs, and offers to travel agents.

17 See for example Bresnahan (1987) and Hausman (1996).

18 These are seasonally adjusted average total earnings paid per employed person per hour, including overtime pay and regularly recurring cash supplements (reported on a quarterly basis). The data has been collected from OECD statistics: https://stats.oecd.org/.

All of the above instruments vary between hotels and over time but do not vary between distribution channels. We are thus looking for an instrument that differs between platforms and can nevertheless affect channel specific market-shares, i.e., affect demand for a particular platform. We thus use Google trends indices for the different distribution channels as popularity indices. According to Statistics Norway (see Table 1.6 below), about 85% of hotel guests in Norway come from Scandinavian countries.

Country	Number	Share	Cum.
Norway	1,142,560	75%	75%
Denmark	74,858	5%	80%
Sweden	69,224	5%	85%
Others	227,535	15%	100%
Total	1,514,177	100%	

Table 1.6: Nationality of the guests for hotels in Norway (in February 2015)

We thus constructed our Google trend indices by focusing on these three countries (Norway, Sweden and Denmark). For each country, we downloaded the Google Trends weekly indices for the key words "Booking.com", "Expedia" and "XXX" between January 2013 and December 2016.[19] When the index for one platform is high, potential customers search more on that platform and this should increase each hotel's sales through that platform relative to competing online channels.

Overall, we thus use four different sets of instrumental variables which vary between hotels or online distribution channels and over time:

- Cost shifters: Hourly wage multiplied by number of rooms [quarterly].
- Characteristics of competing hotels in the same market: TripAdvisor ratings [daily].
- Supplier's prices in other markets for the same good: Offline prices [daily].
- Google trend indices (for three different countries) for keywords identifying the different platforms [weekly].

3.3 Results

In the following, we combine the four types of instruments and estimate the multi-level nested logit model given by equation (1.2). For the estimation, we restrict attention – for each hotel – to the period during which the hotel was listing rooms on Expedia's platforms. Results of these estimations are given in Table 1.7 for specifications including different types of instruments.

[19] "XXX" corresponds to the name of the hotel chain.

	(1)	(2)	(3)	(4)
α	-0.0028	-0.0017*	0.0010	0.0009*
	(0.003)	(0.001)	(0.001)	(0.000)
σ_g	1.2873	0.9952***	0.3138	0.3429***
	(0.776)	(0.188)	(0.166)	(0.098)
σ_d	1.8648	1.1073***	0.6039***	0.6260***
	(1.706)	(0.149)	(0.138)	(0.090)
Instruments:				
Google Trend	X	X	X	X
Cost		X	X	X
Competitor characteristics			X	X
Prices in other market				X
F-Stat:				
p_{jt}	189	211	129	126
$\ln(s_{jt}/s_{gt})$	<1	63	37	35
$\ln(s_{dt}/s_{gt})$	6,386	4,815	2,768	2,421
N	1,923	1,923	1,923	1,923

Heteroscedastic-consistent standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 1.7: Demand model estimation

Concentrating on the last specification (column (4) in the above Table) which includes all four types of instrumental variables, we see that our estimates satisfy the requirements for the model to be consistent. The nesting parameters are indeed positive and lower than one, and they also verify $\sigma_d > \sigma_g$ (estimates are statistically different). This suggests that consumers view platforms Expedia and Booking are closest substitutes, and direct sales are as a more distant substitute. The price parameter (α) also has the expected sign.[20]

Using our multi-level nested logit, we can compute own-price and cross-price elasticities of demand. The elasticity of the demand for product j with respect to the price of product k (where product j is part of group g and subnest d) is given by:

$$\varepsilon_{jk} = \alpha p_j s_j \left(1 - \frac{\mathbb{1}_{k=j}}{(1-\sigma_d)s_j} + \frac{\left(\sigma_d - \sigma_g\right)\mathbb{1}_{k\in d}}{(1-\sigma_d)\left(1-\sigma_g\right)s_d} + \frac{\sigma_g \mathbb{1}_{k\in g}}{\left(1-\sigma_g\right)s_g} \right) \tag{1.3}$$

Table 1.8 reports the average elasticities at the product level (i.e., the average over time and for the 13 hotels). All own-price elasticities are negative and equal to about 1.5 in absolute value. Consumers are thus quite price-sensitive and react to price changes by switching channel and/or hotel. Cross-price elasticities are small, especially across products that are not in the same nest or in the same group, suggesting that consumers tend to switch to hotels outside our sample (other brands) rather than within our sample. This also suggests that substitution between platforms and direct sales is rather limited.

20 The results of the first-stage estimations are presented in Appendix 1.B.

Channel Group Nest	ε_{jj}	ε_{jk} $g_j = g_k$ $d_j = d_k$	ε_{jk} $g_j = g_k$ $d_j \neq d_k$	ε_{jk} $g_j \neq g_k$ $d_j \neq d_k$
Booking	-1.594 (0.032)	0.144 (0.001)	0.023 (0.000)	0.002 (0.000)
Direct	-1.384 (0.029)	0.146 (0.001)	–	0.002 (0.000)
Expedia	-1.632 (0.033)	0.150 (0.001)	0.022 (0.000)	0.002 (0.000)

Table 1.8: Elasticities estimates

We then estimate elasticities of substitution between channels, that is, we compute the impact on total sales for the 13 hotels of an identical price increase for all 13 hotels in a given distribution channel. For example, we compute the relative change in sales on Booking (for all 13 hotels) when all hotels increase the price in their direct channel by 1 %. Results for these "aggregate elasticities" at the channel level are presented in Table 1.9.

Channel	Booking	Direct	Expedia
Booking	-1.315 (0.009)	0.023 (0.000)	0.240 (0.001)
Direct	0.019 (0.000)	-0.902 (0.008)	0.018 (0.000)
Expedia	0.254 (0.001)	0.023 (0.000)	-1.422 (0.009)

Table 1.9: Aggregate elasticities estimates

Own-price elasticities are (in absolute value) only slightly smaller than products' (i.e., hotels × distribution channel) own-price elasticities. They are also larger for OTAs (1.32 for Booking and 1.42 for Expedia) than for direct sales (0.90). In addition the cross-price elasticities are rather small especially between an OTA and direct sales. This was to be expected given our nesting structure, but the order of magnitude is significantly different in the two cases. For instance, when prices for our 13 hotels increase by 1 % on Booking, total sales on Booking for the 13 hotels decrease by 1.32 % and they increase by only 0.24 % on Expedia and 0.02 % on the chains' website. When prices increase on Booking, consumers almost do not switch to the direct channel and a small minority book through a rival OTA. Most consumers actually "leave the market", most likely by booking in different hotels. The situation is very different when prices increase on the chain's website (i.e., direct channel). Following a price increase of 1 %, demand on the chain's website decreases by 0.91 %, and increases only marginal through the OTAs (+0.02 % on Booking or on Expedia). Once again, this suggests that following a price increase consumers tend to switch to other hotels, i.e., price competition between hotels is rather fierce and substantially more important than between distribution channels (i.e., consumers are more loyal to a distribution channel than to a specific hotel or hotel chain).

4 Supply Estimation

We now use the results of the demand estimation together with a structural model of price competition with differentiated products to uncover the hotels' marginal costs for each distribution channel. For each hotel h and each online distribution channel d, we estimate the total marginal cost ($\gamma_{h,d}$), which includes the "production" cost but also the channel-specific distribution costs (including commissions paid to the online travel agencies Expedia and Booking). Among the different estimated demand models, we now focus on the model that includes all four sets of instruments, i.e., the model corresponding the last column in Table 1.7.

We consider here an agency model where hotels keep control of the final prices (including prices charged through OTAs) and pay commissions to OTAs that are simply service providers (but do not acquire rooms for resale). Given that our 13 hotels all belong to same chain, we have to decide whether to focus on a centralized (i.e., a single agent sets the prices for the 13 hotels maximizing the chains' profit) or decentralized (i.e., hotels set prices independently maximizing the hotel's individual profit) pricing model. The estimated aggregate own-price elasticity of demand (-0.902 for the direct sales channel) is consistent with decentralized pricing but not with centralized pricing. In addition, even though the chain has used a central pricing manager towards the end of our sample period, this was not the case early on and hotels may still have some freedom to adjust prices. We thus focus on decentralized pricing and treat hotels as independent, in the sense that each hotel sets prices (one for each channel) independently maximizing the hotel's individual profit.[21]

The system of first-order conditions (to solve for the Nash-Bertrand equilibrium or to obtain the profit maximizing prices of the single-agent) is then given by:

$$\mathbf{s}(\mathbf{p}) - \Theta \odot \nabla s(\mathbf{p}).(\mathbf{p} - \gamma) = 0 \quad \Longleftrightarrow \quad \gamma = \mathbf{p} + (\Theta \odot \nabla s(\mathbf{p}))^{-1}.\mathbf{s}(\mathbf{p}), \tag{1.4}$$

where $\mathbf{s}(\mathbf{p})$ represents the vector of market shares, $\mathbf{p}$ and γ are the vector of prices and marginal costs, $\nabla s(\mathbf{p})$ is the Jacobian matrix of partial derivatives of market shares, Θ is the ownership matrix[22] and the symbol $\odot$ represents the element-by-element matrix product. Given our assumption of decentralized pricing, the ownership matrix is a 39×39 block matrix, each block being a 3×3 submatrix, such that all the elements of the diagonal blocks are equal to one and all elements of the non-diagonal blocks are equal to zero.

[21] In practice, managers' salaries and/or bonuses may be directly linked to their hotel's financial performance, and they may have some freedom to adjust the prices that are recommended by a central entity.

[22] See e.g., Berry et al. (1995) or Björnerstedt and Verboven (2016).

An alternative approach is to impose additional structure on marginal costs. Rather than assuming different marginal costs for different channels, one simply assumes a common marginal cost for all distribution channels. This allows us to additionally estimate the commission rates paid by the hotels to each OTA.[23] In this setting, the system of first-order condition is then given by:

$$(1-\boldsymbol{\tau}).\mathbf{s}(\mathbf{p}) - \Theta \odot \nabla s(\mathbf{p}).((1-\boldsymbol{\tau}).\mathbf{p} - \boldsymbol{\gamma}) = 0, \tag{1.5}$$

where $\boldsymbol{\tau}$ is the vector of commission rates (such that $\tau_d = 0$ for direct sales), and where the vector of marginal costs $\boldsymbol{\gamma}$ is now such that $\gamma_{h,d} = \gamma_h$ for every channel d.

Structure		No		Yes	
Channel	Price	Marg. Cost	Margin	Marg. Cost	Commission
Direct online	1,176	719	41.4%	719	–
Booking	1,334	873	36.9%	719	16%
Expedia	1,366	905	35.9%	719	19%

Table 1.10: Average marginal cost per channel (in NOK)

Table 1.10 reports the average marginal costs (and commission rates) derived, using our estimated demand parameter, from equations (1.4) and (1.5). We first observe that higher prices coincide with higher marginal costs and lower margins, and that selling directly is the cheapest option for the hotel. Selling through the OTAs (rather than directly) adds a cost of 154 NOK for Booking and 186 NOK for Expedia on average, that is, about 12 % and 14 % of the prices charged through these two channels.

We obtain similar results when we impose structure on the marginal cost and try to recover the OTAs' commission rates: These commissions (about 16 % for Booking and 19 % for Expedia) seem in line with rates that are regularly mentioned for OTAs; around 15% for Booking (sometimes higher in large cities), and closer to 20 % for Expedia.

Because hotel pricing really is a dynamic optimization problem, due to the combination of capacity constraints (fixed number of rooms to be sold each day) and anticipated fluctuations in demand over time (seasonality, concerts, sports events, etc), one may worry that our static structural model does not allow us to estimate true marginal costs (and thus commission rates). The worry is that, when computing the marginal cost at each date, we actually capture the true marginal cost as well as the opportunity cost (or option value) of having a room booked a given day rather than closer to the arrival date.

[23] Even if the hotel faces specific distribution costs for its online sales, we cannot identify them separately from the "production cost". What we identify is thus the cost differential between selling through a given OTA (i.e., commission paid to the OTA) and selling directly.

To test the robustness of our estimation[24], we derive the marginal costs using our system of first-order conditions given by equation (1.4), but restricting attention (for each hotel) to bookings made less than 5 days before arrival (rather than less than 30 days) and for dates for which at least 10 % of the hotel's rooms are still available at the arrival date (i.e., $OR_{h,t,0} \leq 90$ %). If a hotel still has a sufficient number of rooms available this close to the arrival date, dynamic optimization should be less of an issue, and the optimization problem should be identical to a static pricing problem. Results are presented in Table 1.11.

Selection	No		Yes	
Channel	Marg. Cost	Margin	Marg. Cost	Margin
Direct online	719	41.4%	653	43.2%
Booking	873	36.9%	740	40.2%
Expedia	905	35.9%	823	37.9%

Table 1.11: Average marginal cost per channel (in NOK)

As we should have expected, once we restrict attention to late bookings for date with late availability of rooms, estimated marginal cost tend to be slightly lower but remains of the same order of magnitude, the difference varying from 66 NOK (for Direct online) to 133 NOK (for Booking). However, it confirms that revenue management plays a non-negligible role.

5 Simulated vs. Actual Effects of Delisting

In this section, we evaluate the effects of removing one distribution channel on prices charged by hotels on the active channels as well as on the different channels' market share. Given that our dataset includes an actual "boycott" of Expedia by our 13 hotels for a relatively long period of time, we take advantage of the data to compare the predicted outcome to the actual outcome and determine the reasons for the observed differences.

5.1 Counterfactual analysis: removing one distribution channel

When hotels decide not to sell their rooms through Expedia, they each choose two prices (one for direct sales and one for sales through Booking) rather than three. Formally, we now solve a reduced version of the system of equations given by (1.4) where we remove the 13 first-order conditions corresponding to prices for sales through Expedia, and replace market shares (and terms of the Jacobian matrix) relative to Expedia

[24] See also Appendix 1.D.

by 0. We thus solve a system of 26 first-order equations for 26 prices (each hotel now sets two prices) given by:

$$\tilde{\mathbf{s}}(\tilde{\mathbf{p}}) - \tilde{\Theta} \odot \nabla \tilde{s}(\tilde{\mathbf{p}}).(\tilde{\mathbf{p}} - \tilde{\gamma}) = 0, \tag{1.6}$$

where the different vectors and matrices are now limited to prices and market shares related to Booking and the direct sales. The ownership matrix is also reduced (now 26×26), blocks being 2×2 submatrices.

When performing the counterfactual simulations, we implicitly assume that commission rates (with the remaining OTAs) are unaffected. It could however be the case that when the chain delists from on OTA (e.g., Expedia), the rival OTA (e.g., Booking) is able to modify its commission. Yet, we have reason to believe this may not be a major problem in our case. Firstly, when the chain decided to boycott Expedia because of the high commission rates charged by Expedia (as well as because of price parity clauses), they did not boycott Booking, reportedly because it accepted not to enforce price parity clauses and already asked for much lower commission rates. In addition, in European countries where price parity clauses were used (and enforced by OTAs) and had to be removed in 2015 (following antitrust investigations), the monitoring exercise carried out by the European Commission (and national competition agencies) suggests that commission rates charged by OTAs was not affected.[25] It thus seems reasonable -– as a first approach — to take the commission rates as exogenous during the whole period.

Given the estimated demand parameters from our demand analysis as well as the marginal costs derived from the structural estimation, we solve the system of equations for the new equilibrium price vector $\tilde{\mathbf{p}}$ and then derive the corresponding market shares $\tilde{\mathbf{s}}(\tilde{\mathbf{p}})$ (through simulations). From these new market shares, we observe how consumers modify their demand choices when the hotels stop using one online channel. Although this differs from looking at switching following a small but significant change in price and we also include the chain's pricing reaction (i.e., change of equilibrium prices through the other channels, although this effects appears to be quite limited), we refer in what follows to diversion ratios between online distribution channels. Formally, for any channel $\hat{d} \neq d$, we define the diversion ratio from channel d to channel $\hat{d}$:

$$DR_{d \to \hat{d}} \equiv \frac{\Delta s_{\hat{d}}}{|\Delta s_d|} = \frac{\tilde{s}_{\hat{d}} - s_{\hat{d}}}{s_d}.$$

This diversion ration $DR_{d \to \hat{d}}$ thus corresponds to the fraction of sales lost by dropping distribution channel d that are recaptured through channel $\hat{d}$. These estimated diversion ratios are presented in Table 1.12.

[25] See the Report on the monitoring exercise carried out in the online hotel booking sector by EU competition authorities in 2016.

Delisting from (d)	Expedia	Direct	Booking
$D_{d \to \text{Direct}}$	15%	-	13%
$D_{d \to \text{Booking}}$	43%	11%	-
$D_{d \to \text{Expedia}}$	-	5%	34%
$D_{d \to \text{Outside option}}$	42%	84%	53%

Table 1.12: Estimated diversion ratios

Not surprisingly given the nesting structure of our demand model, results are very different when deciding to delist from an OTA than when deciding to stop selling directly. When hotels delist from an OTA, they recapture a significant share of the consumers as only about half of the consumers switch to the outside good, i.e., book in a different hotel (possibly on the platform on which our hotels have stopped selling). The other half who continue to book a room from the hotel chain switch to a large extent to the rival OTA (this is the case for about 3 out of 4 consumers who keep booking in one of the 13 hotels). When the hotels decide to stop selling directly, they only recapture a small share of the lost consumers as 84% of the consumers who used to buy directly switch to the outside good (i.e., most likely book a room in a different hotel, directly or maybe even through an OTA). Among the few consumers who remain loyal to the chain, about two-third book through Booking and one-third through Expedia.

The high diversion ratios to the outside good suggest that inter-brand competition (i.e., competition between hotels) is an important factor and that consumers tend to be more loyal to a distribution channel than to a hotel (or even a chain). This is even more true for consumers who have a preference to buy directly (rather than through an OTA). Among those who tend to favour OTAs, consumers seem to have a slight preference for Booking than for Expedia (this is consistent with a higher market share for Booking than for Expedia).

In the particular case of a delisting of Expedia (which occurred in practice between the end of 2012 and the second semester of 2015), we estimated the distribution of diversion ratios (to Booking, direct sales and the outside good) through bootstrap. Out of 10,000 iterations, we kept only observations with coherent demand estimates and diversion ratios (i.e., 3,961 observations with non negative diversion). Figure 1.1 provides details of the distributions of these diversions ratios following a decision to stop listing on Expedia's platforms. Although, some iterations generate extreme values, results appear consistent: we observe a high diversion ratio to the outside good, and among those consumers who keep booking a room in one of the 13 hotels included in our sample, a large majority does so through Booking rather than through the chain's website.

Finally, we compute the simulated impact on consumer surplus as well as on hotels

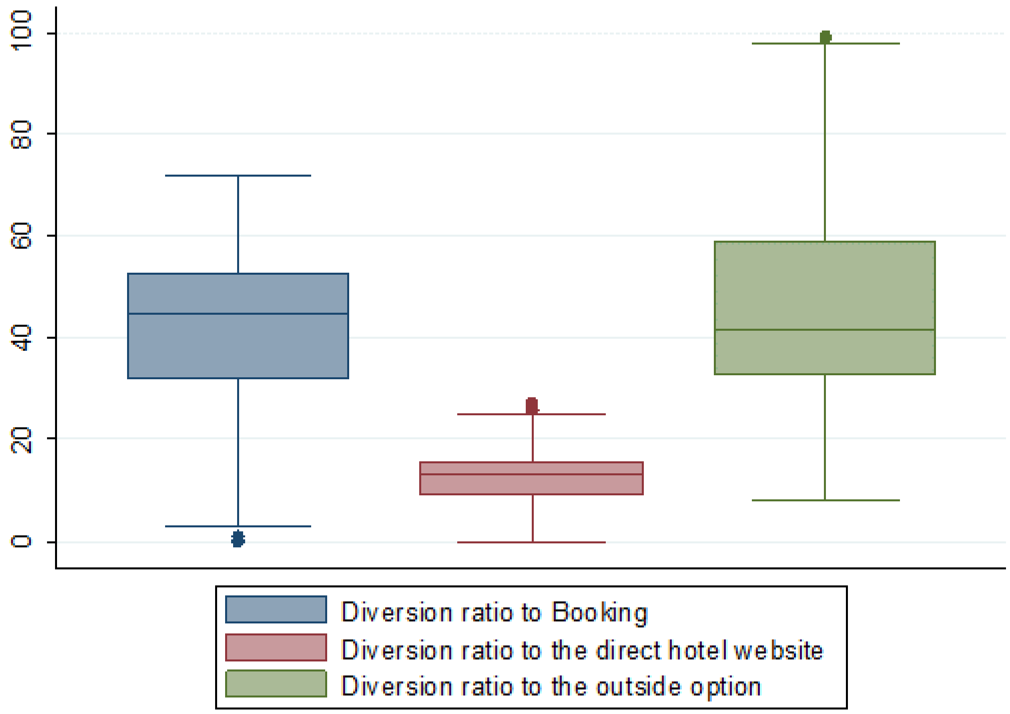


Figure 1.1: Box-plot of estimated diversion ratio (in %)

profits and Booking's revenues. These measures are obviously only partial as we focus on those consumers who – in the absence of delisting – would have booked (online) a room in one of the 13 hotels included in the sample. The impact on hotels' profit is also limited to the impact on the profit generated by online sales, and the impact on Booking's revenue is limited to the revenue generated on sales for the 13 hotels included in the sample. Results from the counterfactual simulation are presented in Table 1.13. Figures in the first column ("Observed") are the values using the estimated demand parameters and marginal costs and figures in the second column ("Delisting") are the values in our simulated counterfactual scenario where the hotels all stop listing on Expedia. Figures in the last column simply measure the relative change between the two. All values are measured in thousands NOK per week.

×1,000 NOK	Observed	Delisting	Δ
Consumers	1,117	991	-21.3%
Hotels	444	398	-10.1%
Booking	34	51	+45.3%

Table 1.13: Welfare effects of delisting from Expedia (average weekly levels)

Based on this simulated counterfactual scenario, it appears unsurprisingly that consumers and hotels are harmed by the boycott. A large share of consumers switch to other hotels ("outside good") or to a second-best distribution channel and do not benefit from better prices (new equilibrium prices are almost identical to the initial prices). Hotels pay lower commissions (i.e., faces lower marginal costs) because Expedia was the most expensive distribution channel and thus earn higher profits on sales recaptured through the two remaining channels, but the share of consumers lost to rival hotels is too large to be compensated by the higher margins. Finally, Booking's revenue increases because it captures an important share of Expedia's original sales.

5.2 Comparing predicted and actual outcomes

Because our dataset includes an actual boycott of Expedia, we can compare the predicted outcome to the actual outcome and, more importantly, try to determine the reasons for the observed differences. We first compare the predicted and actual effects of the boycott on prices charged by hotels through Booking and their own website. The average predicted and actual prices are reported in Table 1.14. Because the boycott period is relatively long (January 2013 - September/October 2015), it is possible that demand for the direct channel or one of the OTAs has evolved over time (for example because consumers got accustomed to booking hotel rooms through online platforms). To limit such effects, we propose two comparisons between the predicted outcome (based on about one year of data post-boycott) and the observed outcome: in the first case, we keep the whole boycott period ("Whole period"); while in the second case, we restrict attention to the bookings made for the last year of the boycott only ("Last year").

Channel	Observed		Counterfactual
	Whole period	Last year	
Booking	1,196 (-10.88%)	1,247 (-7.08%)	1,341 (+0.64%)
Direct	1,063 (-10.07%)	1,130 (-4.42%)	1,175 (-0.04%)

Table 1.14: Observed and predicted prices

Whereas our counterfactual simulation predicts almost no change in the prices charged by the hotels on the chain's website and a small increase in the prices charged on Booking, prices observed for these distribution channels during the actual boycott period (February 2013 - September 2015) were actually about 10 % lower than once hotels started listing again on Expedia (September/October 2015 - November 2016). The predicted prices are thus much higher than the actual prices. The difference is slightly lower once we restrict the observed boycott period to the last year, but even in this case observed prices were about 4 to 7 % lower during this year than they were after the boycott ended.

The same observation can be made for the distribution channels' market shares (conditional on buying online) that are reported in Table 1.15. The model seems to predict the outside good's market share quite well, but the split of the online sales between Booking and the direct channel is inaccurately predicted unless we restrict our attention to the last year of the boycott.

Channel	Observed		Counterfactual
	Whole period	Last year	
Booking	56%	49%	47%
Direct	44%	51%	53%
Outside Good	95%	94%	95%

Table 1.15: Comparison on market shares

The discrepancies between predicted and observed outcomes do not necessarily mean that our structural model is not well-suited to perform a sensible counterfactual analysis. It does however suggest that it cannot be used without caution to predict the outcome of a delisting decision for example. In the line of Peters (2006), we try to identify a possible explanation for these discrepancies and focus on changes in the observed "product characteristics", here characteristics of the different bookings such as type of room or advance booking for example (i.e., changes in the X's). In general, when performing counterfactual simulations, these parameters are assumed to remain constant. However, if there are good reasons to believe that characteristics may have changed, the simulation will always yield an incorrect outcome if these changes are not accounted for.

Table 1.16 reports average characteristics of bookings during the post-boycott period (September/October 2015 - November 2016), that is, during the period that we used to estimate our demand model, as well as the average booking characteristics observed during the boycott period for the whole period and for the last year only.

Control variables	Channel	Post-Boycott	Boycott	
			Whole period	Last year
Occupancy rate	Booking	75.4%	86.5%	88.3%
	Direct	72.7%	79.6%	83.7%
Days in advance	Booking	9.3	9.5	9.3
	Direct	10.1	11.8	10.8
Superior rooms	Booking	10.1%	7.8%	8.8%
	Direct	14.7%	11.8%	13.5%
Week-end	Booking	34.2%	31.1%	32.2%
	Direct	32.6%	33.7%	28.7%

Table 1.16: Average booking characteristics during and after the boycott period

It appears that booking characteristics were slightly different during the boycott period (whether we focus on the whole period or only on the last year) when compared to the post-boycott period. For example, occupancy rates (at the time of booking) were about 10 percentage points higher on average, consumers used to book fewer superior rooms and were booking less often for week-end nights.

Rather than using characteristics of the post-boycott observations to simulate the counterfactual equilibrium, we thus use the actual booking characteristics during the boycott period. Results for these simulations are reported in the second column of Table 1.17.[26]

		Counterfactual		Observed
Correction		*No*	*Yes*	(whole period)
Price	Booking	1,341	1,196	1,196
	Direct	1,175	1,063	1,063
Market share	Booking	47%	55%	56%
	Direct	53%	45%	44%
	Outside Good	95%	94%	95%

Table 1.17: Simulated and predicted outcomes

Using the product characteristics observed during the boycott (rather than the post-boycott characteristics) clearly improves the accuracy of the simulated results as the boycott-period prices and market shares are now very precisely estimated. Once we correct for changes in product characteristics, our structural model can thus be used to predict quite well the outcome of delisting.

6 Conclusion

In this paper, we use an exhaustive dataset of bookings for 13 hotels in Oslo to estimate a (structural) demand model and evaluate the degree of substitution between different online distribution channels. We conclude that, for each online distribution channel (i.e., two large OTAs as well as the chain's own website), the own-price elasticities of demand are relatively large, meaning that consumers tend to be price sensitive. In addition, cross-price elasticities are significantly lower, which suggests that a large share of consumers would rather switch between hotels (and thus to the outside good in our specification) than switch distribution channel. On average, consumers thus seem more loyal to a platform than to the hotels, and inter-brand competition seems fierce enough. However, our analysis also shows that among those consumers who are willing to switch distribution channel following a price increase on one OTAs' platform (around 50 to 60 %), a large majority would rather book through the rival

[26] We simulated two different counterfactual scenarios: one using all observations during the boycott period, the second restricting attention to the last year of the boycott period. However, because the results are almost identical - identical market shares and prices that differ only by less than 2 NOK, we only report one set of results (using data for the whole period).

OTA than directly from the hotel. It thus appears that, OTAs are closer substitutes to other OTAs than are direct sales.

Our analysis cannot directly be used to evaluate the competitive effects of price parity clauses imposed by OTAs on hotels, as we would first need to estimate a structural model allowing for bargaining between hotels and OTAs over commission rates (to evaluate the impact of price parity clauses on commissions). It suggests, however, that direct sales are a credible alternative to OTAs, because a significant share of consumers would stay loyal to the hotel if the hotel were to stop listing on one of the OTAs (such as Expedia for example). Therefore, from a theoretical point of view, one cannot simply assume that suppliers (hotels in our case) cannot directly and efficiently reach final consumers. It thus cannot be presumed that platform price parity clauses would necessarily harm consumers and/or hotels in this market.

Because our dataset covers a period that includes an actual decision to delist from Expedia's platforms, we have been able to compare the simulated and actual effects of such an event. Given the discrepancies between the simulated and observed effects on prices and market shares, one may be tempted to conclude that we either did not use the correct demand model, or, pushing it even further, that structural IO models cannot accurately be used to predict outcomes of counterfactual experiments (such as strategic decisions to stop using some distribution channels or, as more commonly used, to evaluate the competitive effects of a potential or notified merger). We have, however, been able to identify a plausible possible reason for these discrepancies, namely changes in product characteristics over time. Once we account for the changes in product characteristics, we observe that the simulated and actual outcomes (in terms of prices and market shares) are very similar. We thus believe that structural IO models can be reasonably accurately estimated and used to perform sensible counterfactual experiments. However, one needs to proceed with caution and account for all important changes that may affect the simulated outcome.

Appendices

1.A Selected Observations

Table 1.18 summarizes our data selection process. We start with 1,235,106 (online) bookings that account for about 2,91 billion NOK in revenue for the 13 hotels. Restricting attention to standard and superior rooms eliminates about 10% of the booking that account for 26% of the hotels' revenue. The share of excluded revenue is substantially higher than the share of excluded bookings as we eliminate bookings for more expensive rooms including suites. We then focus on bookings for less no more than 7 days. This eliminates only 1% of the bookings but again a larger share of the revenue (13%) as we exclude more expensive bookings (more nights). Concentrating on bookings that include only one room removes a very small number of bookings and a very small share of revenue (less one 1% in both cases). Focusing on bookings for one or two guests eliminates about 3% of the remaining bookings and about 4% of the revenue. Finally, we remove very early bookings (i.e., made more than a month in advance): this eliminates about 17% of the remaining rooms that account for about 19% of the remaining revenue (i.e., early booking have on average similar prices than late bookings).

Restriction	Revenue	Evolution	Bookings	Evolution
-	2,91	-	1,235,106	-
Standard and superior rooms	2,16	-26%	1,116,204	-10%
$\leq$ 7 days	1,88	-13%	1,104,210	-1%
One room	1,88	$\leq$-1%	1,099,862	$\leq$-1%
1 or 2 guests	1,80	-4%	1,064,158	-3%
Advance $\leq$ 30 days	1,45	-19%	885,249	-17%
All		-50%		-28%

Table 1.18: Restrictions on bookings (and impact on revenue)

1.B First-stage Estimates

Table 1.19 presents the results of the first-stage estimations. Instruments globally enter with the expected sign.

	p_{jt}	$\ln\left(\frac{s_{jt}}{s_{ht}}\right)$	$\ln\left(\frac{s_{ht}}{s_{gt}}\right)$
Instruments :			
Prices in other market	324.39***	-0.76***	-0.00
	(37.995)	(0.176)	(0.031)
Competitor last 5 grade	-2087.67**	5.87	-0.40
	(787.869)	(3.451)	(0.666)
Competitor grade	-8251.07**	18.31	0.85
	(2940.730)	(14.123)	(2.592)
Competitor comments	-3766.08***	8.99	0.29
	(1011.974)	(4.832)	(0.882)
Cost	-1261.62***	7.33***	-0.03
	(148.598)	(0.715)	(0.134)
Google Trend Norway	-77.53***	-0.01	0.30***
	(11.881)	(0.055)	(0.010)
Google Trend Sweden	31.09*	-0.04	0.02
	(13.130)	(0.064)	(0.013)
Google Trend Denmark	14.03	0.04	-0.18***
	(14.021)	(0.068)	(0.013)
N	1,923	1,923	1,923

Heteroscedastic-consistent standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 1.19: First-stage estimation results for instruments variables

For instance, instruments that were expected to be correlated with the product prices are significant and have the expected sign: the cost shifter -– computed as the number of room multiplied by the hourly wage — is negatively correlated with the price, which is consistent with the idea that large hotels that have higher fixed costs benefit from economies of scale, leading to lower prices. Similarly, prices in other markets (i.e., offline prices) are positively correlated with online prices, and better competitors' characteristics or higher ratings lead to lower online prices for the reference hotel (consistent with the idea that hotels set lower prices when they face tougher competition).

Our channel varying instruments (i.e., the Google trend indices) work more or less well. We expected a positive correlation between the index for an OTA and this OTA's relative market share among OTAs. This work well for Norway (and to a smaller extend for Sweden) but less so for Denmark.

1.C Robustness checks on the outside good

In this section, we show that our results are not extremely sensitive to the methodology adopted to compute the outside good's market share. Until now, we have decided to estimate, for each week, the number of bookings made online in 3 and 4-star hotels in Oslo. We thus compute the outside good's share based on number of booked made that month in Oslo, divided by four (to obtain weekly values) and multiplied by the share of online bookings (in our sample for that particular week) and by 0.7 (share of 3 and 4-star hotels in Oslo).

Share of 3/4-star hotels	50%	60%	70%	80%	90%
α	0.0009*	0.0009*	0.0009*	0.0009*	0.0009*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
σ_g	0.3429***	0.3431***	0.3429***	0.3425***	0.3419***
	(0.099)	(0.098)	(0.098)	(0.098)	(0.098)
σ_h	0.6252***	0.6259***	0.6260***	0.6257***	0.6252***
	(0.090)	(0.090)	(0.090)	(0.090)	(0.090)
Instruments:					
Google Trend	X	X	X	X	X
Cost shifter	X	X	X	X	X
Competitor characteristics	X	X	X	X	X
Prices in other market	X	X	X	X	X
N	1,923	1,923	1,923	1,923	1,923

Note: Heteroscedastic-consistent standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$.

Table 1.20: Demand model estimation

We now confirm that results are relatively robust by varying the last multiplier (and thus the outside good's share) between 0.5 and 0.9. Estimates of the nested-logit parameters are displayed in Table 1.20 for different values of the multiplier. We observe that estimates remain almost unchanged.

We also compute the estimated diversion ratios (following a decision by all hotels to delist from Expedia) for these different shares of 3 and 4-star hotels (or outside good's market share). Once again, diversion ratios – that we report in Table 1.21 – remain almost unaffected. Results are thus robust to (reasonable) changes in the outside good's market share.

Share of 3 and 4-star hotels	50%	60%	70%	80%	90%
$D_{\text{Expedia}\to\text{Direct}}$	15.78%	15.50%	15.32%	15.19%	15.11%
$D_{\text{Expedia}\to\text{Booking}}$	42.90%	43.01%	43.03%	43.07%	43.07%
$D_{\text{Expedia}\to\text{Outside option}}$	41.33%	41.50%	41.65%	41.74%	41.82%

Table 1.21: Estimated diversion ratios

1.D Capacity constraints

To further confirm that capacity constraints may not be such a major issue, we now compare the proportion of days without sales during and after the Expedia boycott. If capacity constraints were a crucial issue (especially for online prices) and affects pricing strategies, it should have been made worse when hotels decided to list again on Expedia (end of our sample period used for the demand estimation). In that case, we should expect the decision to list on Expedia's platform to reduce the proportion of days for which we observe bookings on different channels, especially close to the arrival date. But the data shows – see Figure 1.2 – that the decision to use Expedia's platforms did not really affect these proportions: actually, the only change occurs for direct sales in the last few days before the arrival date, but the effect goes in the unexpected direction as we observe an increase in the number of days where bookings are made after hotels decide to join Expedia's platforms.

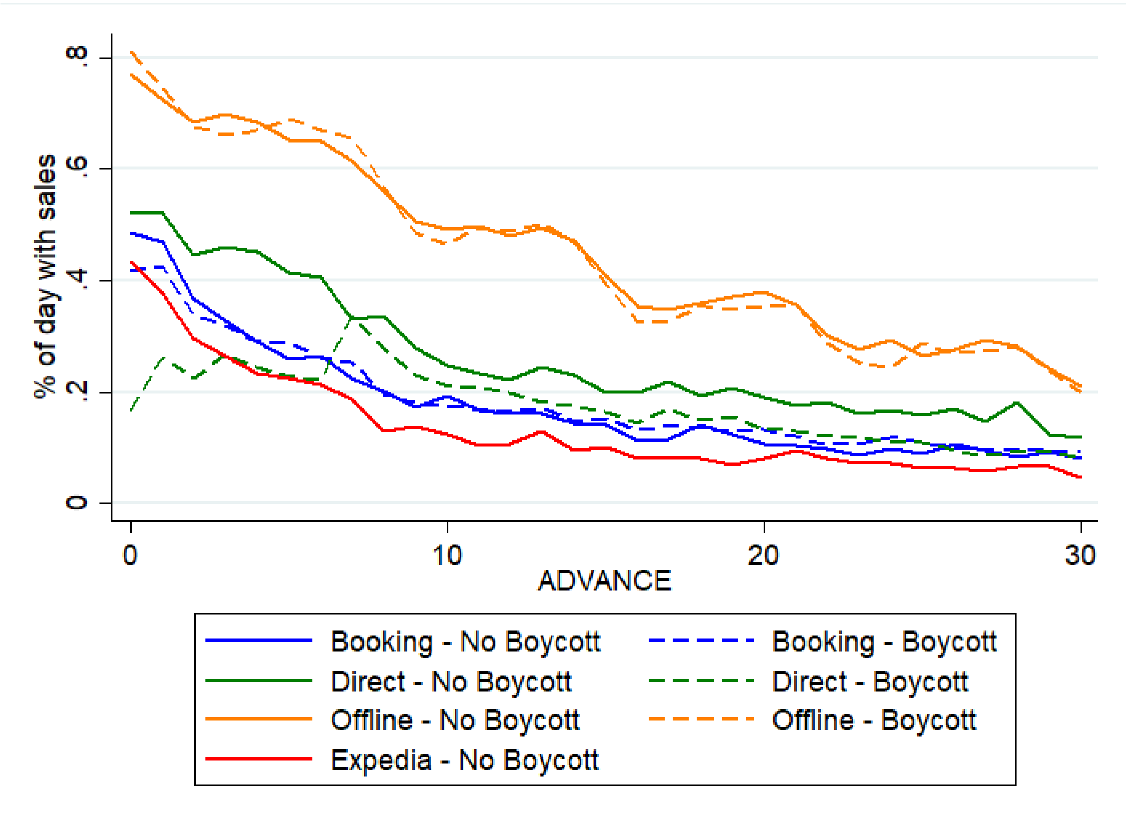


Figure 1.2: Proportion of days with sales during/after the boycott

Bibliography

Angrist, J. and Pischke, J.-S. (2010), 'The credibility revolution in empirical economics: How better research design is taking the con out of econometrics', *Journal of Economic Perspectives* **24**(2), 3–30.

Berry, S. (1994), 'Estimating discrete-choice models with product differentiation', *Rand Journal of Economics* **25**(2), 242–262.

Berry, S., Levinhson, J. and Pakes, A. (1995), 'Automobile prices in market equilibrium', *Econometrica* **63**, 841–890.

Björnerstedt, J. and Verboven, F. (2016), 'Does merger simulation work? evidence from the swedish analgesics market', *American Economic Journal: Applied Economics* **8**, 125–164.

Boik, A. and Corts, K. (2016), 'The effects of platform most-favored-nation clauses on competition and entry', *Journal of Law and Economics* **59**, 105–134.

Bresnahan, T. (1987), 'Competition and collusion in the american automobile industry: The 1955 price war', *Journal of Industrial Economics* **35**(4), 457–82.

Duch-Brown, N., Grzybowski, L., Romahn, A. and Verboven, F. (2017), 'The impact of online sales on consumers and firms. evidence from consumer electronics', *International Journal of Industrial Organization* **52**(2), 30–62.

Edelman, B. and Wright, J. (2015), 'Price coherence and excessive intermediation', *Quarterly Journal of Economics* **130**, 1283–1328.

Ellison, G. and Fisher Ellison, S. (2018), 'Match quality, search, and the internet market for used books', *NBER Working Paper 24197* .

Friederiszick, H. and Glowicka, E. (2016), 'Competition policy in modern retail markets', *Journal of Antitrust Enforcement* **4**(1), 42–83.

Gentzkow, M. (2007), 'Valuing new goods in a model with complementarity: Online newspapers', *American Economic Review* **97**(3).

Hausman, J. (1996), Valuation of new goods under perfect and imperfect competition, *in* T. Bresnahan and R. J. Gordon, eds, 'The Economics of New Goods', National Bureau of Economic Research.

Hunold, M., Kessler, R., Laitenberger, U. and Schlütter, F. (2018), 'Evaluation of best price clauses in hotel booking', *International Journal of Industrial Organization* **61**, 542–571.

Johansen, B. r. O. and Vergé, T. (2017), 'Platform price parity clauses with direct sales', *University of Bergen Working Papers in Economics 01/17* .

Johnson, J. (2017), 'The agency model and mfn clauses', *The Review of Economic Studies* **84**(3), 1151–1185.

Larrieu, T. (2019*a*), 'Most favoured nation clauses on the online booking market', *mimeo* .

Larrieu, T. (2019*b*), 'Pricing strategies in online market places and price parity agreements: evidence from the hotel industry', *mimeo* .

Lieber, E. and Syverson, C. (2012), Online vs. offline competition, *in* M. Peitz and J. Waldfogel, eds, 'Oxford Handbook of the Digital Economy', Oxford University Press.

Mantovani, A., Piga, C. A. and Reggiani, C. (2020), 'Online platform price parity clauses: Evidence from the eu booking. com case', *Available at SSRN 3381299* .

Miller, N. and Weinberg, M. (2017), 'Understanding the price effects of the millercoors joint venture', *Econometrica* **85**, 1793–1791.

Nevo, A. (2001), 'Measuring market power in the ready-to-eat cereal industry', *Econometrica* **69**(2), 307–342.

Nevo, A. and Whinston, M. D. (2010), 'Taking the dogma out of econometrics: Structural modeling and credible inference', *Journal of Economic Perspectives* **24**(2), 69–82.

Peters, C. (2006), 'Evaluating the performance of merger simulation: Evidence from the u.s. airline industry', *The Journal of Law & Economics* **49**(2), 627–649.

Pozzi, A. (2013), 'The effect of internet distribution on brick-and-mortar sales', *The Rand Journal of Economics* **44**(3), 569–583.

Prince, J. (2007), 'The beginning of online/retail competition and its origins: An application to personal computers', *International Journal of Industrial Organization* **25**(1), 139–156.

Rey, P. and Vergé, T. (2016), 'Secret contracting in multilateral relations', *Toulouse School of Economics Working Papers 16-744* .

Wang, C. and Wright, J. (2016), 'Platform investment and price parity clauses', *Available at SSRN 2848574* .

Wang, C. and Wright, J. (2020), 'Search platforms: Showrooming and price parity clauses', *the RAND Journal of Economics* **51**(1), 32–58.

Weinberg, M. (2011), 'More evidence on the performance of merger simulations', *American Economic Review* **101**, 51–55.

Weinberg, M. and Hosken, D. (2013), 'Evidence onthe accuracy of merger simulations', *Review of Economics and Statistics* **95**, 1584–1600.

Chapter 2

The Impact of Online Platforms' Preferred Partner Programs on Consumer Prices

"Hors du programme, une grande partie des hôteliers applique la parité restreinte sans y être obligée, par souci de simplicité ou de performance. Nous ne l'indiquons pas pour eux, pourquoi l'indiquerions-nous pour les établissements préférés ?"

- Vanessa Heydorff (2017), *Directrice France de Booking.com*

1 Introduction

The last decade has seen a massive increase in usage of online intermediaries by buyers and sellers. Sellers frequently try to reach consumers by selling through their own direct sales channels (i.e., own website) but also through marketplaces such as Amazon MarketPlace or eBay. Traditional brick-and-mortar retailers have also started selling competitors' products on their own marketplaces (e.g., Walmart or BestBuy in the U.S., Zalando or Fnac-Darty in Europe). For some products, it is now common for consumers to search and buy products and/or services through online platforms (e.g., books or music), price comparison websites (e.g., insurance contracts) or online travel agencies and meta-search sites (e.g., travel, accommodation).

The speedy development of online intermediaries had important consequences in terms of business models and contractual relationships, the traditional retailer being for instance replaced by a service provider using an agency model. It also raised some new consumer protection and antitrust issues.

For a long time, traditional grocery retailers have been selling products under private labels competing with branded products that they also carried, without much antitrust scrutiny. Authorities – especially in Europe – are now increasingly wary of dominant vertically integrated online platforms acting as gatekeepers for most suppliers. The European Commission is indeed expected to issue very soon a Statement of Objection charging Amazon regarding the treatment of third-party sellers on its marketplace. Some commentators even argue that it should not be allowed to be both a player (selling its own private label products) and the referee (setting the rules on the marketplace).

In June 2017, the European Commission fined Google 2.42 billion euros for abusing its dominance by giving an illegal advantage to its own comparison shopping product (Google Shopping). The investigation concluded that Google was systemically giving prominent placement to its own service, placing Google's comparison shopping results above the results of Google's generic search algorithm.

Platforms sometimes impose price parity clauses to sellers. These clauses limit the sellers' ability to freely set prices on different channels, notably preventing them from offering lower prices on their own sales channel. These clauses have led to multiple investigations by competition authorities. Almost ten years ago, Amazon decided to stop requiring sellers to offers their best prices on the marketplace after the UK's Office of Fair Trading and Germany's Bundeskartellamt initiated investigations. Although, the decision did not directly concentrate on price parity clauses, the combination of a switch from the retailer model used by Amazon (for e-books) to the agency model used by Apple (iBookstore) and price parity requirements imposed to publishers were at the heart of the Apple iBookstore decision. The UK Competition and Markets Authority also investigated the competitive effects of price parity clauses imposed by price comparison websites to car insurance providers.

Online travel agents (OTAs), such as Booking.com or Expedia, have become essential intermediaries between hotels and travellers. According to European competition agencies, OTAs now account for about 70% of online room bookings. OTAs perform a double role. They provide a unified search platform aggregating the supply of a large number of hotels, thus helping consumers to select a hotel. But they also give consumers the opportunity to book directly through the platform, thus competing with the selected hotel's own website to complete the booking.

The leading OTAs, Booking.com and Expedia (and to a lower extent HRS in Germany), have been at the heart of multiple investigations by competition agencies regarding the use of price parity clauses (PPCs). Prior to the investigations, OTAs commonly imposed wide PPCs: hotels were required to offer the best deal on the given platform and were not allowed to propose a lower price through any other channel, including its own website.

Competition agencies considered that such clauses reduced the platforms' incentives to compete over commission fees, ultimately harming consumers. They thus widely considered wide PPCs to be anti-competitive. The theory of harm put forward by competition agencies is relatively straightforward: when PPCs are generalized (i.e., used by all leading platforms), each hotel is forced to set the same price everywhere. This uniform price is thus based on the average commission / distribution cost incurred by the hotel. If a platform unilaterally increases its commission, the hotel faces a higher cost and thus uniformly increases all prices. Starting from the competitive commission (set in the equilibrium without PPCs), a platform thus has an incentive to unilaterally increase its commission as it does not affect its market share (and only marginally affects the total sales). As commissions are strategic complements, all commissions, and therefore all consumer prices, increase in equilibrium.

With a few exceptions, competition agencies however considered that free-riding by hotels was a possible issue: hotels could indeed use the OTAs as "showrooms" where consumers search for the best match, offering lower prices on their own website ultimately driving consumers away from the platforms to avoid paying the OTAs' commissions. In April 2015, the French, Italian and Swedish competition agencies thus accepted commitments offered by Booking.com to abandon wide PPCs and revert instead to narrow PPCs that only constraint the hotel as regard the price charged on its own website. Booking.com also announced that it would revert to narrow PPCs in all EEA countries and, in July 2015, Expedia announced that it would do the same.[1] But

[1] Similar decisions were adopted by competition authorities in Australia, New-Zeland or Switzerland for instance.

in some countries, competition agencies (Germany in the HRS and Booking cases) or legislators (France (2015), Austria (2016), Italy (2017, Belgium (2018)) went further and banned all PPCs imposed by OTAs.

Possibly as a response to antitrust investigations and bans on PPCs, several large OTAs have started proposing voluntary preferred partner programs. Booking.com introduced its Preferred Partner Program offered to a small number of selected hotels (see section 4.1). Hotels joining the program benefit from increased visibility as they appear ahead of the other hotels in the search results. Booking.com Preferred Partner program is one of its improved visibility programs along Genius or Visibility Booster.[2] The central features of this specific program are that hotels that voluntary join the program accept to pay slightly higher commissions to Booking and also commit to price parity (i.e., to offer their best deal through Booking.com) in exchange for improved visibility (better position ranking).[3]

Improved visibility or search ranking is of crucial importance for suppliers on platforms or marketplaces as it directly affects the supplier sales. Focusing specifically on the hotel industry, Ursu (2018) shows that the ranking of offers on OTAs affects consumers' choices: consumers are indeed more likely to look for detailed information (i.e., to click) on hotels with a better ranking. Preferred partner programs thus have a direct impact on demand for hotels and may therefore affect prices charged by hotels through different channels. Whether we expect this prominence effect to lead to higher or lower prices is however uncertain. Better ranking implies that hotels will usually observe a positive shift in demand. It is for instance claimed by Expedia that *"one-third of bookings are for hotels in the top 1 spot"* and that *"the first five of the top 30 positions in search results capture 65% of the bookings."* This may push hotels to higher prices as they can afford to propose slightly more expensive deals when they anticipate a positive shift in demand. However, the literature on search (and more specifically on ordered search) has shown that prominent seller tend to offer lower prices than less prominent sellers (see for instance Arbatskaya (2007) and Armstrong (2017)), although these prices may still be higher than with random search: this is because sellers that are listed below know that they only face consumers whose match values with the prominent sellers are low if they are still searching. They thus face less competition and can afford to offer higher prices.

In addition, because these programs often include price parity clauses, they have

[2] Expedia also offers an improved visibility program called *Accelerator* through which hotels can improve their ranking by paying a higher commission. However, as for the Booking Visibility Booster program, this is a program that is commonly used for short periods of time by hotels, when they need to fill rooms or plan to offer discounts to a specific audience.

[3] HRS (a large OTA in German-speaking countries) introduced a similar program called "Top Quality Seal". Hotels that join the program must give HRS wide price parity. See the European Commission's report on the monitoring exercise in the online booking sector (para. 40).

additional effects on hotels' pricing incentives. Independently of the impact of such clauses on hotels' costs (i.e., commissions paid to platform), hotels are prevented to offering better deals for direct sales. When setting uniform prices, one should thus expect hotels to increase the prices set for direct sales and lower the prices offered for sales through more expensive platform.

Finally, prominence and price parity both affect the platform's commission rate. Because prominence usually leads to higher profits for the prominent firm than for the competitors (even if they are ex-ante symmetric), hotels are willing to pay for prominence (Armstrong and Zhou, 2011). In addition, price parity clauses may also reduce competition between platforms and lead to higher commissions (Boik and Corts, 2016). We understand that Booking.com's Preferred Partner Program (BPPP) usually led hotels to pay higher commissions, which would be consistent with those theories.[4] Because commissions are variable costs for hotels, we expect hotels to pass-through part of the additional cost to consumers in the form of higher room prices.

The overall effects of the program on room prices is thus ambiguous. Due to prominence, we may expect hotels that join the program to increase their prices relative to prices set by non-joiners. In addition, increased commissions should also push prices upwards for all types of sales (through Booking or for direct sales). However, due the non-discrimination constraint (price parity clause), prices should decrease through Booking and increase through the direct channel. Overall, the impact of the program is most likely ambiguous for prices set for sales through Booking but we could expect the program to lead to higher direct prices. There is however an uncertainty linked to the effect of prominence on prices.

In this paper, we specifically focus on the impact of Booking.com's Preferred Partner Program (BPPP) on consumers prices. Using exhaustive booking data from 22 Scandinavian hotels (belonging to the same chain), we first run a simple hedonic price regression trying to evaluate the factors that affect prices charged by hotels. Results are unsurprising and consistent with earlier similar work. This preliminary analysis also shows that prices are significantly and substantially lower in hotels that ended-up joining Booking.com's Preferred Partner Program. However, this does not imply that the program led hotels to decrease prices but could simply be due to an important selection bias: hotels that joined the program were not randomly selected but are hotels that tend to be located in more competitive local markets.

To go further, we then try to compare the price changes for hotels that joined the program to the evolution of prices for comparable hotels. Rather than relying on a full

[4] In practice, we observe that hotels that join the program pay higher commissions than the other hotels. This does not however tell us whether the commissions would have been lower – at least for hotels joining the program – in the absence of the program.

synthetic control approach (Abadie and Gardeazabal, 2003; Abadie et al., 2010), we construct a control group by selecting hotels that are have similar characteristics and face similar competition conditions than the hotels included in the treatment group. We then estimate the impact of the BPPP on prices and quantities relying on standard difference-in-differences methods. We estimate that joining the BPPP led hotels to increase prices in all online distribution channels (+4% to +7% in Denmark, +3% to +7% in Sweden, the impact being even larger in Norway). This price increase can be consistent with increased commissions (imposed by Booking.com) and price parity clauses (one component of BPPP) as well as with a positive shift in demand pushing the hotel to extract more surplus from consumers. We complement this by evaluating the impact of the BPPP on volumes sold through the different online channels. Results show that sales increased through OTAs, more so through Booking.com that through Expedia whereas online direct sales were reduced but to a much smaller extent. Overall, online sales increased making it more likely that a positive shift in demand is the most consistent explanation for the price increases.

Finally we take advantage of differences in compliance with price parity clauses to disentangle the specific effect of price parity. We indeed observe that, whereas "standard" hotels (part of our control group) have almost never complied with price parity for non-refundable offers, "preferred"hotels (our treatment group) have changed behaviour after joining the program. Both types of hotels seem to have always applied price parity for flexible tariffs. We use this difference in compliance for different types of tariffs (non-refundable vs. flexible) for different hotels (treatment vs. control) to estimate the effect of price parity on prices through a triple-difference approach. Our analysis shows that new compliance with price parity (for treated hotels and non-refundable tariffs) pushed preferred hotels to lower prices charged for non-refundable offers through Booking but only in Norway. We otherwise find no statistically significant effect of price parity on prices, result that is consistent with Mantovani et al. (2020) who find no (medium-term) effect of the legal ban on narrow price parity clauses in France and in Italy.

Related Literature.

To the best of our knowledge, there are no empirical research articles directly evaluating the impact of visibility programs on supplier's pricing incentives in different online channels. Hunold et al. (2020) have however looked at the determinants of hotels' rankings in OTAs' search results. They find that hotels appear lower on an OTA's search results when they charge lower prices on other channels, especially on their own website. For instance, a 10% price difference reduction on the hotel's own channel relative to the price charged on Booking.com has the same effect as a reduction of user rating on Booking.com of about 0.3 point (on a 0-10 scale). They also show that joining Booking.com's preferred partner program (or Expedia's Sponsored adds) has a massive positive impact on the hotel's ranking. These results are also perfectly in line with the theory literature on ranking algorithm biases by profit maximizing search engines (e.g., Hagiu and Jullien (2011)).

Still focusing on pricing issues, Lu et al. (2015) evaluate how the introduction of a new sales channel (e.g., online direct sales by a hotel) leads to a reduction in the pricing of intermediaries (i.e., OTAs) suggesting that OTAs and direct sales are substitutable.[5]

Our paper also indirectly relates to the growing literature evaluating the monetary value for consumers (in terms of increased indirect utility) of a hotel's improved ranking on OTAs. Ursu (2018) finds that consumers are more likely to click on a better ranked hotel when looking for detailed information. However, conditional on looking at the detailed information, the ranking position does not influence the consumer's booking decision. She also estimates that the position effect is worth less than $3.2. Several authors such as Ghose et al. (2012, 2014), Koulayev (2014), Chen and Yao (2016) and De los Santos and Koulayev (2017) had earlier suggested that the monetary value of improved ranking was higher, with estimates for a one-position improvement ranging from $3 to $35.

In this paper, we try to evaluate the impact of visibility boosters on consumers prices. Our analysis thus related to the theory literature on firm prominence. In a sense, a visibility booster makes a hotel joining the program more prominent (i.e., consumers are more likely to click to look for more detailed information about better ranked hotels, and therefore more likely to buy). Behavioral models such as Salant (2011) or Zhou (2011) suggest that consumers with bounded-rationality tend to favour prominent options. Originating with Armstrong et al. (2009)[6], authors have shown that firms may benefit from prominence (and are thus willing to pay for prominence) despite charging lower prices in equilibrium. However, platforms may end up charging higher commissions to extract the additional profits generated through prominence. Higher fees in turn lead to higher consumer prices, so that the overall effect on consumer surplus may be ambiguous. In addition, results may be affected when search is ordered (i.e., consumer start by looking at a specific suppliers) rather than random (see Arbatskaya (2007) or Armstrong (2017)).

In addition, the specific program that we analyze also re-introduces price parity requirements, thus generating additional effects on consumer prices (directly but also indirectly through the possible impact on commission fees). Most of the theoretical

5 See also Cazaubiel et al. (2020) who analyse substitution patterns between OTAs and direct sales by hotels.

6 See also for instance, Armstrong and Zhou (2011) and Rhodes (2011)

literature has insisted on the anti-competitive effects of such clauses. Because an increase in one platform's commission leads to uniform price increases on all platforms (due to the price parity requirements) and thus does not affect market shares – conditionally on suppliers selling through the same platforms, a platform has unilateral incentives to increase its commission. This leads to higher equilibrium commissions and therefore to higher consumer prices (see, e.g., Boik and Corts (2016) and Johnson (2017)).[7] Johansen and Vergé (2017) allow suppliers to endogenously decide on which platforms to sell. They show that, when inter-brand competition (i.e., competition between suppliers) is sufficiently fierce and de-listing is a credible option, price parity clauses may actually lead to lower commissions and thus to lower final prices.[8] Calzada et al. (2019) show that although the threat of delisting may force platforms to reduce commissions when they impose PPCs, consumers are always worse-off even when the clauses prevent show-rooming and increase hotels' participation.

A few authors have tried to empirically evaluate the effect of price parity clauses on consumer prices, essentially looking at the hotel industry.[9] However, because many factors may have influenced demand and therefore prices at the time price parity clauses have been removed by OTAs, these authors have often looked at the impact on price dispersion rather than on price levels (see for instance Hunold et al. (2018), Larrieu (2019), Ennis et al. (2020) and Mantovani et al. (2020)). Findings show that the price dispersion increased after the ban on price parity clauses, with the price charged on the hotel's website now more likely to be lower than the prices charged elsewhere. The European Commission recently published a study carried out in the context of the evaluation of current rules regarding vertical contracts that suggest that the complete ban on price parity clauses introduced by legislators in Austria, Belgium and Italy led to price reductions of the order of 3 to 4%. It also confirms that price dispersion increased and that the lowest price proposed for a given hotel may have decreased by about 10 to 12%.[10]

The rest of the paper is organized as follows. In section 2, we present the context

7 In settings where consumers sequentially search for the best match, Edelman and Wright (2015),Wang and Wright (2016) and Wang and Wright (2020) show that, despite reducing the risk of free-riding by sellers on platforms' investment ("showrooming") and thus ensuring the platforms' viability, price parity clauses usually harm consumers because of higher prices or inefficient investment.

8 In the context of secret contracting over non-linear commissions, Rey and Vergé (2016) show that price parity clauses do not affect consumer prices but only affect the division of surplus between suppliers and platforms.

9 Although not directly evaluating the effect of price parity clauses but also the switch from the wholesaler to the agency model, De los Santos and Wildenbeest (2017) show that the settlement reached by publishers and the U.S. Department of Justice regarding the retailers' ability to freely set prices for e-books led to (average) price reductions of 8% at Barnes & Noble and 18% at Amazon.

10 Support studies for the evaluation of the VBER, European Commission, 2020.

and the data that we use in this paper. We then propose a hedonic price analysis (section 3) that suggests that average prices decreased after Booking.com' Preferred Partner Program was introduced. In section 4, we take advantage of the fact that only some of the hotels in our dataset have joined the program to estimate, using a difference-in-difference method, the impact of the program on consumer prices but also on quantities. Finally, in section 5, we propose a triple-difference approach to estimate more specifically the impact of the price parity clause included as part of the visibility booster program. Section 6 concludes.

2 Data and Context

2.1 Data

We use an exhaustive dataset of all bookings made over a 29-month period in 22 hotels. These hotels all belong to one of the leading hotel chains active in Scandinavia and are located in major cities in Sweden (11 hotels), Denmark (6) and Norway (5). Although the hotels belong to the same chain, they are not located in same geographical markets (i.e., different cities) except for the two hotels located in Copenhagen. We can therefore consider each hotel as a separate entity that is responsible for its own pricing strategy. Even if the pricing strategy was set centrally, the central revenue manager would ultimately solve separate pricing problems for the different hotels.

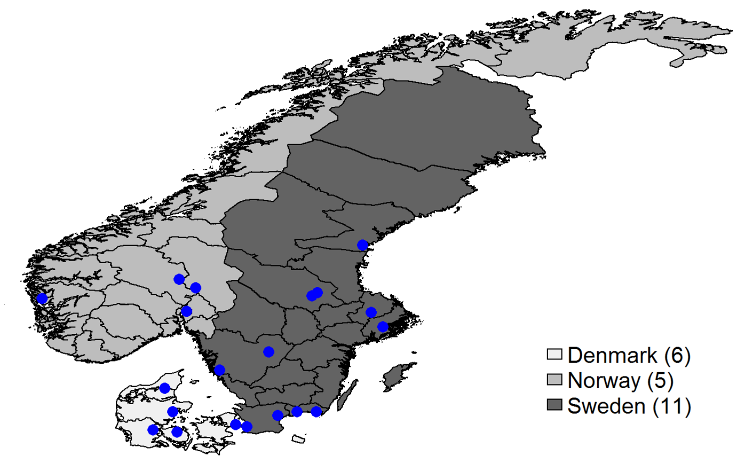


Figure 2.1: Hotels' location

The dataset has been provided by the hotel chain and data has been directly extracted from its information system. The original dataset includes 824,048 bookings made by consumers through all distribution channels between March 2014 and July 2016. For each booking, we observe:

- The booking date as well as the arrival and departure dates. This allows us to compute the length of stay as well as the lead time (i.e., how many days prior to arrival the room has been booked).
- The room type (e.g., standard, superior, junior suite, ...).
- The number of guests.
- The channel through which the room was booked.
- The price paid by the consumer as well as the rate and the market codes associated with the booking.

Although our analysis focuses on online individual bookings (i.e., bookings made through the hotel's website or through OTAs), we use our exhaustive dataset and existing information on the number of rooms at each hotel to compute occupancy rates at any point in time. More precisely, for each hotel h and each arrival date t, we compute the occupancy rate $OR_{h,t,x}$ at date $t-x$ (i.e., x days before date t) as the ratio between number of bookings made at date $t-x$ or earlier for a stay including a night at date t and the total number of rooms in the hotel. $OR_{h,t,0}$ thus indicates which proportion of hotel h's rooms are effectively occupied at date t, whereas $1-OR_{h,t,x}$ indicates which proportion of the rooms (for at stay at date t) were still available for purchase x days prior to arrival. We compute occupancy rates for all values of x between 0 and 180 (i.e., we compute the occupancy rate daily up to six months before arrival, given that all bookings are made less than 6 months in advance).

In addition, we collected hotel characteristics and consumer reviews. The additional data includes for each hotel:

- Number of rooms.
- Precise hotel location (as well as distance from city center and nearest airport).
- Star rating as well as existence of specific amenities (bar, restaurant, fitness and/or wellness center).
- Consumer reviews collected daily from TripAdvisor for the whole period. For each hotel and each day, we observe the last five ratings (on a 1-to-5 scale), the current average rating and the total number of reviews.
- The number of hotels from the same city that are listed on Booking.com.

2.2 Rate identification

Bookings can be made by different type of consumers (e.g., leisure vs. business), through different distribution channels (e.g., online vs. offline, OTAs vs. hotel website, Global distribution systems (GDS) vs. B2B contracts) and at different rates (e.g., rack rate or negotiated tariffs).

Channel	# Observations	Share of bookings
Offline	408,785	50%
B2B	189,583	46%
Individual	119,171	29%
Travel agency	100,031	24%
Online travel agency	237,712	29%
Booking	122,683	52%
Expedia	99,174	42%
Others	15,315	6%
Global distribution system	98,670	12%
Hotel website	79,421	10%
Total	824,048	100%

Table 2.1: Market share repartition by channel

Table 2.1 provides the market shares (based on numbers of bookings) by distribution channel. About half of the bookings fall into the "offline" category that includes individual bookings that can be made by consumers over the phone or directly at the hotel's front-desk (about 14% of all bookings), bookings through traditional travel agents that negotiate directly with the hotels (12%) and bookings made directly by business at negotiated rates (23%). GDS (e.g., Amadeus, Galileo, Pegasus, Sabre or Worldspan) account for about 12% of all bookings. Finally, "online sales" account for 39% of all bookings. These online bookings can be made directly through the hotel's (or hotel chain's) website (10%) or through on online travel agents (29%), with platforms linked to Priceline (most notably Booking.com) and Expedia (e.g., Expedia.com or Hotels.com) being the two prominent groups of such OTAs.

Our data includes detailed distribution channels and market categories but also precise rate codes that have been used to determine the price paid for a given booking. This information allows us to classify bookings according to the associated rate into four different groups: *public rates* for individual travellers paying an unqualified or non-corporate negotiated rate, *contracted rates* for individual travellers paying a contracted rate (e.g., a given discount rate relative to the best-available-rate or BAR), *group leisure rates* for leisure related travel for groups and *group business rates* for business travel for groups.

In Table 2.2, we present for each of the four rate categories, the share of bookings made through the different distribution channels. We observe that all bookings associated with group rates (leisure or business) are made "offline". We see however, that leisure tariffs tend to be booked through travel agents whereas business tariffs are mostly linked to B2B contracted rates. Bookings for individual rates are mostly made online through OTAs and direct channels for public rates (67% of such bookings) but almost exclusively offline (61%) or through GDS (33%) for negotiated rates.

Channels by rate category	Individual Public	Individual Negotiated	Group Business	Group Leisure
Offline	28%	61%	100%	100%
B2B	5%	34%	88%	19%
Individual	18%	12%	3%	8%
Travel agents	5%	15%	9%	73%
Online travel agencies	52%	1%	-	-
Booking	27%	0%	-	-
Expedia	22%	0%	-	-
Others	4%	1%	-	-
Global Distribution Systems	5%	33%	-	-
Hotel website	15%	5%	-	-
Total	100%	100%	100%	100%

Table 2.2: Proportion of distribution channels by rate category

In Table 2.3, we now look at the decomposition across rate categories for each distribution channel. If offline bookings are made using all four categories of rate categories – with differences as individual bookings use mostly public rates whereas B2B bookings use contracted rates (individual or group rates) and travel agents use mostly contracted rates for individual or leisure groups, online bookings and bookings through GDS use almost exclusively one rate category. Online bookings are almost exclusively related to public tariffs (this is the case for all bookings made through Priceline or Expedia platforms) whereas a large majority (78%) of the bookings through GDS use contracted rates. This important difference plays an important role in our choice of instrumental variables.

2.3 Summary statistics

Table 2.4 presents summary statistics of the hotel characteristics. Our sample includes only 3 and 4-star hotels that are relatively large (about 143 rooms on average, minimum of 65 rooms). Hotels are all centrally located (maximal distance to the city geographic center is less than 2 kilometers) but a large proportion of them are quite distant from the closest large airport[11]. Finally, hotels in our sample face very differ-

[11] We restricted attention to airports with at least 10 different destinations. The closest airports to our 22 hotels are thus: Aalborg (AAL), Billund (BLL) and Copenhagen (CPH) in Denmark, Göteborg (GSE), Malmö (MMX), and the two Stockholm airports (Arlanda (ARN) and Bromma (BMA)) in

Rate categories by channel	Individual Public	Individual Negotiated	Group Business	Group Leisure	Total
Offline	31%	35%	22%	12%	100%
Business	12%	42%	41%	5%	100%
Individual	70%	25%	3%	3%	100%
Travel	22%	34%	8%	36%	100%
Online travel agencies	99%	1%	-	-	100%
Booking	100%	0%	-	-	100%
Expedia	100%	0%	-	-	100%
Others	78%	22%	-	-	100%
Global Distribution Systems	22%	78%	-	-	100%
Hotel website	86%	13%	-	-	100%

Table 2.3: Proportion of rate categories by distribution channel

ent degrees of competition. Each hotel faces on average 35 competitors (i.e., hotels located in the same city and listed on Booking), 12 of which are in a similar star-rating category (i.e., 3 and 4-star hotels). But this varies substantially withing our sample, as one hotel only faces 6 competitors (and only one 3 or 4-star hotel) while at the other extreme one hotel may face as much as 619 competitors (including 132 of a similar star-rating category).

	Mean	Median	Min	Max
Number of rooms	143	124	65	300
Star rating	3.5	3.5	3	4
Bar	0.77	–	0	1
Restaurant	0.68	–	0	1
Fitness/Wellness	0.68	–	0	1
Distance to city center (km)	0.6	0.5	0.1	1.9
Distance to closest airport (km)	70	96	6	345
Number of competitors[1]	35	121	6	619
Number of similar competitors[2]	12	33	1	132

[1] Number of available accommodations listed on Booking.com and located in the same city (as of early 2019).
[2] Number of available 3-star and 4-star rated accommodations only.

Table 2.4: Summary statistics of hotel characteristics

Summary statistics of booking characteristics are presented by major categories of distribution channels in Table 2.5. Because our data includes hotels in countries using different currencies (Danish, Swedish or Norwegian krone), prices for hotels in Denmark and Sweden have all been converted to Norwegian krone using the daily exchange rate at the time of booking.

As expected, prices tend to be higher for online bookings as this category includes

Sweden and Bergen (BGO) and Oslo (OSL) in Norway.

mostly individual travellers paying standard rates whereas contracted and group rates are usually lower. We also observe that average prices are almost identical for OTAs and direct online sales (i.e., hotel website). Individual online bookings also tend to be made earlier (28 to 35 days in prior to arrival compared to a lead time of 17 to 23 days on average for GDS and offline bookings), for more travellers (1.6 to 1.7 compared to 1.1 to 1.4), slightly more often for standard rather than superior rooms and more often for stays starting during week-ends (about 40% of bookings compared to 5% to 26% for GDS and offline bookings).

Channel	OTA	Hotel website	GDS	Offline	All
Price (in NOK)	1,239	1,237	1,180	1,081	1,135
Lead Time (days)	35	28	17	23	26
Nights	1.7	1.6	1.7	1.8	1.7
Persons[1]	1.7	1.6	1.1	1.4	1.5
Standard room	77%	73%	80%	83%	80%
Week-end[2]	42%	39%	5%	26%	29%
Occupancy rate (booking)[3]	43%	41%	47%	48%	46%
Occupancy rate (final)[4]	77%	73%	77%	75%	76%

[1] Number of adults and children.
[2] Week-end represents the proportion of stays that start on a Friday or a Saturday.
[3] Occupancy rate (booking) is evaluated the day a booking is made (i.e., indicates the proportion of rooms that are no longer available for the requested arrival date at the time of booking).
[4] Occupancy rate (final) is evaluated at the arrival date (i.e., realized occupancy rate at date t).

Table 2.5: Summary statistics of booking characteristics

Finally, our booking data includes detailed rate codes that provide additional information on the type of tariff that has been proposed to the consumer. This is particularly true for online bookings, i.e., bookings made directly through the hotel website or through the large OTAs that are Booking and Expedia. In particular, through the rate code, we can identify whether the tariff chosen by the consumer was refundable or not. Flexible tariffs are either refundable (i.e., the consumer leaves his/her credit card details at the time of booking and/or pays directly but may cancel the trip and be fully refunded – at least if the cancellation does not occur too late) or the consumer pays directly at the hotel when checking-in (or checking-out). When booking a non-refundable offer, the consumer pays immediately and cannot be reimbursed in the event that s/he cannot travel.

Table 2.6 shows that, except for the direct sales (made through the hotel's website), we can identify whether the tariff was refundable or not for more than 97% of observations. This share is lower for direct sales but we still manage to identify the type of tariffs for more than 80% of bookings. The table also confirms that prices are usually significantly lower for non refundable bookings, and that a majority of bookings are flexible on OTAs (about 60% on Booking and 54% on Expedia) but about two-thirds of direct bookings are non-refundable. Either hotels have tried to differentiate by offering more non-refundable (but cheaper) offers on their own website rather than through the OTAs, or consumers that look for cheaper deals are also willing to book more often through the hotel directly.

Channel	Booking		Expedia		Hotel website	
Tariff	Price	Share	Price	Share	Price	Share
Flexible	1,369	58%	1,241	52%	1,413	28%
Non refundable	1,315	39%	1,059	45%	1,211	55%
Not identified	1,285	2%	1,174	3%	1,038	17%

Table 2.6: Average prices by type of tariff and channel

3 Price Analysis

So far, the hedonic price literature (dating back to Court (1939)) has been used to quantify the impact of room or hotel characteristics on the price of bed and breakfast rooms (see, e.g., Monty and Skidmore (2003)) or hotel rooms (see, e.g., Abrate and Viglia (2016) for hotels in Milan and Thrane (2007) for hotels in Oslo). Some authors have tried to adapt such methods to the evolution of the industry and the growing importance of internet bookings. For instance, Law et al. (2011) include customer reviews in the characteristics. However, we are not aware of hedonic price analyses that account for specific characteristics of the consumer bookings, such as the distribution channel used to book or the occupancy rate or other proxies for demand that may affect the pricing decision.

3.1 Hedonic price regression

In this section, we run a simple hedonic price regression but include these additional characteristics. In particular, on top of hotel characteristics (such as star rating, consumer reviews, location, ...) and room characteristics (such as single / twin - double / triple room, breakfast included or not), we include booking characteristics that may directly affected by the hotel's revenue management strategy.

More specifically, for a booking i made for a stay in hotel h at time t, we estimate the following price equation:

$$\ln P_{iht} = \alpha + (\beta.X_h + \gamma.Y_{ht}) + \delta.Z_{iht} + \varepsilon_{iht}. \tag{2.1}$$

The hotel characteristics consists of time invariant characteristics (X_h) such as star rating, distance to the city center and to the nearest international airport, amenities ratings (i.e., presence of a bar, restaurant and/or fitness/well-being centre), number of competitors (listed accommodations on Booking.com in the same city) but also of time varying characteristics (Y_{ht}). We include two characteristics that vary over time. The first is linked to customer reviews and consists of a re-scaled (on a 0 to 10 scale) Tripadvisor rating (based on current average rating at the time of booking). The second is a dummy variable identifying whether, at the time of booking, the hotel was a member the Booking.com's Preferred Partner Program.[12]

Booking characteristics include usual variables such as the type of room (standard vs. superior), the number of travellers, the booking length (number of nights), whether or not the trip starts on a week-end (arrival on Friday or Saturday). We also include dummies identifying the channel that has been used to book (Booking, Expedia or the hotel's / chain's website) as well as the type of tariff (non-refundable vs. flexible).

Finally, to account for variations in demand that may directly affect prices through the hotel's pricing strategy, we control for seasonality (including a time trend for booking date and dummies for week and month of arrival) but also for lead time and occupancy rate. The lead time is the difference between the arrival date and the booking date, i.e., it indicates how many days prior to arrival the booking has been made. Hotels often use tariff types that may be available only more than a week, two weeks or a month prior to arrival. We also account for the occupancy rate at the time of booking (i.e., proportion of rooms that were no longer available when the booking was made) but also at the time of arrival (i.e., proportion of rooms that were occupied on a given date). Hotels usually use complex pricing strategies to account for the fact they face capacity constraints but variable demand at each date. This type of revenue management strategy (common in other industries with similar characteristics such as airline or train tickets) usually involves prices that depend on the observed occupancy rate, the expected demand but also the remaining time before arrival (i.e., lead time).[13] We view the realized occupancy rate as a proxy for the expected demand for a given date, while the occupancy rate at the time of booking gives an indication of how early consumers are likely to book for a given date.

We estimate the price equation (2.1) separately for hotels located in different countries, i.e., we run three separate regressions for hotels located in Denmark, Norway and Sweden, to account for country-specific factors. In addition, we allow error terms to be correlated within a group of observations and thus compute the standard errors

[12] See Section 4.1 for a more detailed description of the program.

[13] See for instance Gallego and Van Ryzin (1994) for a theoretical analysis, as well as Escobari (2012) or Alderighi et al. (2015) for reduced-form empirical analysis and Williams (2017) or Cho et al. (2018) for structural estimations.

clustering by hotel and reservation dates. Error terms are thus assumed independent across hotels and dates of booking but not within each group.

3.2 Instruments

Many of our variables may be endogenous and simultaneously determined with prices. This could in principle be the case for some of the hotel characteristics such as the number of rooms, the star rating, the presence of specific amenities. However, these are characteristics are more difficult to modify in a short period of time. Therefore, we treat them as exogenous.

Some other variables are however more likely to be endogenous " in a short run" approach. This is in particular the case for the two occupancy rates that are included among the booking characteristics (in Z_{iht}). The occupancy rate can be seen as a demand proxy, and because prices and demand influence each other, they are determined simultaneously. To solve this endogeneity problem, we use an instrumental variable approach.

To construct our instruments, we use bookings made through GDS. As we have already discussed earlier, we expect online bookings and booking through GDS to belong to independent market segments. As shown previously in Table 2.3, bookings made through GDS fall mostly in one single rate category, namely, *Individual Negotiated Rates* (78%), whereas online bookings (made directly through the hotels' website or through OTAs), the rates are mostly *Individual Public Rates* (86% and 100% respectively). Because they belong to independent markets (i.e., unrelated demands), online prices and prices on GDS should not adjust simultaneously. However, both types of bookings contribute to the hotel's occupancy rate.

Using our complete dataset, we can compute, for each hotel and each arrival date t, the number of rooms booked through GDS at date $t-x$ (i.e., x days prior to the arrival date t) for stay including a night at date t. To make it comparable to the hotel's occupancy rate $\text{OR}_{h,t,x}$, we simply divide that number of GDS bookings by the hotel's number of room. We thus obtain a figure $\text{GDS}_{h,t,x}$ that indicates the proportion of the hotel's rooms that have been booked through GDS before date $t-x$ for a stay starting on night t. Exactly like for occupancy rates, we compute these proportion for all values of x between 0 and 180.

We then instrument the hotel's occupancy rate at the time of booking (i.e., $\text{OR}_{h,t,x}$ or OR_{Book} in the tables below) by the corresponding proportion of rooms sold through GDS before date $t-x$ (i.e., $\text{GDS}_{h,t,x}$ or GDS_{Book} in the tables below). Similarly, we instrument the realized occupancy rate (i.e., $\text{OR}_{h,t,0}$ or OR_{Final}) by the corresponding proportion of rooms sold through GDS for date t (i.e., $\text{GDS}_{h,t,0}$ or GDS_{Final}).

The coefficients associated to these excluded instruments in the first stage regres-

	Sweden		Denmark		Norway	
	OR_{Book}	OR_{Final}	OR_{Book}	OR_{Final}	OR_{Book}	OR_{Final}
GDS_{Book}	1.747***	0.071***	2.018***	0.231***	2.025***	0.025
	(0.03)	(0.02)	(0.04)	(0.03)	(0.04)	(0.04)
GDS_{Final}	-0.459***	1.010***	-0.511***	1.039***	-0.430***	1.266***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
F-Stat	2,352	3,154	1,270	1,545	1,691	1,775
APF-Stat	3,523	5,758	2,804	2,953	2,980	3,135
N	104,834	104,834	73,256	73,256	47,978	47,978

Notes: Clustered (by hotel and by reservation data) standard errors in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.7: First-stage estimation results

sion are displayed in Table 2.7. As we expected, the occupancy rate and the corresponding share of rooms sold through GDS (either at the time of Booking or realized at the arrival date) are positively and significantly correlated. Our instruments are thus correlated with potentially endogenous variables. In addition, the high values of the F-statistic figures suggest that our instruments are indeed relevant.

3.3 Results

Using the two instruments introduced in the previous section that correct for endogeneity issues related to the occupancy rates, we estimate the hedonic price equation (2.1). We present separately the estimated coefficients for the hotel characteristics (Table 2.8) and for the booking characteristics (Table 2.9). We estimate the price equation separately for each country, and present the results of the OLS and IV regressions.

Consumers seem to be sensitive to the hotel's location. If the distance to the airport has a significant impact on prices, its economic magnitude remains relatively small. However, the distance to the city-center has an economically meaningful (and significant) effect on prices. Doubling the distance from the city-center (from the mean of 0.5km to 1km) generates a price drop between 1% (in Sweden) and 3.5% (in Norway). As expected, we also note that the star-rating has a meaningful impact on prices: 4-star hotels tend be 10% (Sweden) to 26% (Denmark) more expensive than 3-star hotels. Reputation (as measure by the TripAdvisor review ratings) also matters, one extra point on the 0-to-10 scale leads to a price increase that remains smaller in Norway (+1.5%) and Denmark (+2.5%) than in Sweden (+7%). Having special amenities is also positively valued by consumers, one extra service (e.g., bar, restaurant or fitness/well-being center) adding between 8% (Denmark) and 18-19% (Sweden, Norway) to the room price.

	Ordinary Least Squares			Instrumented variables		
	Sweden	Denmark	Norway	Sweden	Denmark	Norway
Distance /airport	-0.000***	-0.002***	-0.006***	-0.000***	-0.003***	-0.007***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Distance /center	0.002	-0.040**	-0.092***	-0.016**	-0.040**	-0.070***
	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
Amenity score	0.111***	0.042***	0.166***	0.179***	0.082***	0.193***
	(0.01)	(0.02)	(0.00)	(0.01)	(0.02)	(0.00)
Stars	0.099***	0.234***		0.103***	0.257***	
	(0.01)	(0.03)		(0.01)	(0.03)	
Popularity score	0.056***	0.009*	-0.005	0.069***	0.025***	0.015**
	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
Competitors	0.001***			0.000		
	(0.00)			(0.00)		
Preferred Partner	-0.173***	-0.021	-0.478***	-0.297***	-0.089***	-0.699***
	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
N	104,834	73,256	47,978	104,834	73,256	47,978

Clustered (by hotel and by reservation date) standard errors in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.8: Estimated effects of hotel characteristics

More interestingly, hotels that join Booking.com's Preferred Partner Program tend to charge lower prices. The magnitude of the effect appears extremely large especially in Sweden (-30%) and Norway (-70%). This may seem both counter-intuitive (sign) and extreme (magnitude). Indeed the program is associated with higher commissions paid to Booking, enforcement of price parity clauses and higher visibility for the hotel. One may initially believe that all this should lead to higher prices. However, some effects might go in the "unexpected" direction: when competition between hotels is fierce, price parity may lead to lower commissions and therefore possibly lower prices (Johansen and Vergé (2017)). In addition, more visible firms may also charge lower prices as suggested by the theory on prominence (Armstrong et al. (2009)). More important, as we thoroughly discuss in the next section, hotels do not randomly join the program. Joining the program may indeed be the consequence of increased competition in the hotel's local market. In this case, we should expect, competition to push prices down and this may well explain this high negative coefficient in the hedonic price regression (i.e., selection bias). Although we control for the intensity of competition in the hotels' local markets (through the number of competitors located in the same city that are listed on Booking.com), this variable has either no significant effect on prices (Sweden) or is co-linear to some other variables (Denmark, Norway). This seems to confirm that there is a strong correlation between the intensity of competition faced by the hotel and the likelyhood to joined the program.

Focusing now on booking characteristics, we again find rather usual effects. For instance, consumers face lower rates for week-end trips (-5% [Denmark] to -20% [Swe-

	Ordinary Least Squares			Instrumented Variables		
	Sweden	Denmark	Norway	Sweden	Denmark	Norway
Weekend	-0.156***	-0.026***	-0.069***	-0.196***	-0.048***	-0.114***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Superior room	0.173***	0.193***	0.233***	0.169***	0.181***	0.210***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Persons	0.086***	0.074***	0.124***	0.093***	0.085***	0.135***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Nights	-0.005***	-0.000	0.003**	0.000	0.005***	0.008***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Booking	-0.029***	0.011***	-0.010***	-0.031***	0.006**	-0.020***
(ref: direct)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Expedia	-0.021***	-0.064***	-0.063***	-0.027***	-0.066***	-0.075***
(ref: direct)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Non refundable	-0.134***	-0.138***	-0.156***	-0.132***	-0.134***	-0.155***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
OR_{Book}	0.258***	0.247***	0.146***	0.145***	0.386***	0.108***
	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)
OR_{Final}	0.111***	0.165***	0.201***	0.585***	0.373***	0.691***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)
Lead time	0.003***	0.001***	0.001***	0.001***	0.002***	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
N	10,4834	73,256	47,978	104,834	73,256	47,978

Clustered (by hotel and by reservation date) standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.9: Estimated effects of booking characteristics

den]) and this is easily explained by the fact that our hotels tend to be located in large cities and belong to a well-known chain that caters for an important business crowd during weekdays. As expected, superior rooms are about 20% more expensive than standard rooms, and adding one traveller increases the price by 9% to 14%, or roughly the price of breakfast in this hotel category. Staying more than one night has a significant but economically meaningless effect on the price paid by consumers.

More interesting are the effects of the booking channel, type of tariff or occupancy are on the room price. Prices tend to be lower on OTAs (Booking or Expedia) than on the hotel's (or chain's) website. In Sweden, prices are roughly 3% lower on OTAs, in Norway they tend to be lower on Expedia (-7%) than on Booking (-2%). Denmark looks slightly different, as Booking is marginally more expensive than the hotel's website (+0.6%) whereas Expedia is quite cheaper (-6.6%). Non-refundable offers are on average 13% to 15% cheaper than refundable ones, which is totally unsurprising as they do not offer the same flexibility for consumers. We discuss more thoroughly the type of offers and the difference between OTA and direct prices in the next sections.

Finally, occupancy rates positively affect prices paid by consumers. Adding 10 percentage points to the occupancy rate at the time of booking (this rate is just over 40%

on average), means that consumers face prices that are 1% (Norway, Sweden) to 4% (Denmark) higher. Higher occupancy rate at any point in time means that room has filling faster and this unsurprisingly leads the hotel managers to increase room prices for the remaining rooms. In the same vein, prices tend to be higher for dates at which the hotels are busier: adding 10 percentage point to the realized occupancy rate increases the room prices by about 4% (Denmark) to 7% (Norway) at any time. Once again, this is unsurprising: hotels can charge higher prices when demand is higher with risking to see an important reduction in the number of guests. It is commonly thought that booking earlier allows consumers to obtain better deals. However, this is due to the combination of two effects: the actual effect of buying earlier (linked to the lead time) and the effect of buying before others, i.e., at a period when the occupancy rate is still low (effect of OR_{Book}). Controlling for this second effect as well, it would appear that buying earlier, as measured by the lead time, leads to higher prices although the effect is relatively small.

4 Impact of the Preferred Partner Program

As we discussed in the previous section, the hedonic price regression is not well-suited to identify the effect of Booking.com's Preferred Partner Program (BPPP) on prices. Joining the program is not an exogenous decision and, in addition, we should expect the program to have potentially different effects on prices in different online distribution channels. In this section, we look further into the effects of this program.

4.1 Booking's Preferred Partner Program

In 2015, possibly in response to bans on best price (or rate parity) clauses in many countries, Booking.com introduced a new program that was proposed to selected hotels. Hotels satisfying eligibility criteria based on customer ratings, number of nights sold through Booking.com, price attractiveness (relative to prices charged through other online channels) were invited to join BPPP.

Hotels that decided to join the program, had to accept to pay higher commission rate (about 2 to 3 percentage points above the standard rate) and to commit to conform to the *"Parity of Rates and Conditions"* requirement. This imposes that the hotel offers through Booking.com *"equal or better rates for the same accommodation, services and equipment of equal of better quality, and restrictions or conditions that are equal or more advantageous in comparison to what this hotel is offering on its own online platform, publishes online on other platforms or displays in an online advertisement (including meta search engines)."*[14] We discuss in great details compliance with the rate parity condi-

[14] Translation from the French version of the conditions of use of the Preferred Establishment Pro-

tions in the next section.

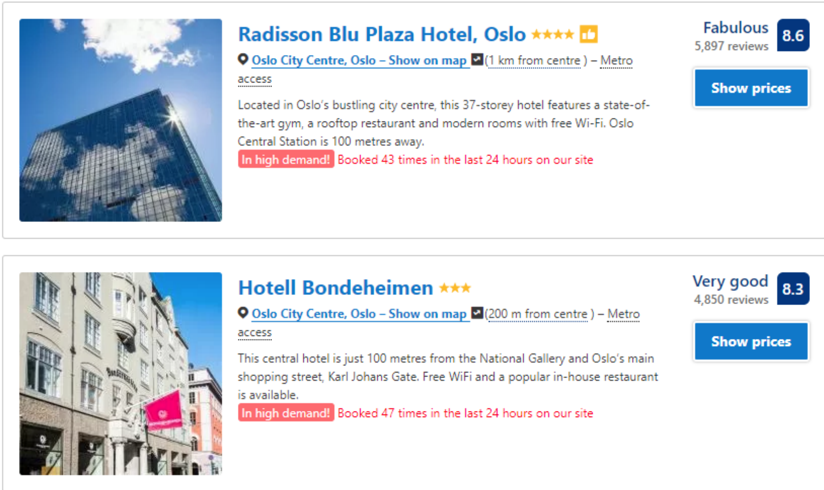


Figure 2.2: Thumb-up symbol identifying a preferred Trip provider

In return, the *"preferred Trip Providers"* benefit from increased visibility as they appear ahead of the rest of the hotels in the search results and are marked with a thumbs-up symbol as shown in Figure 2.2.[15] According to Booking.com, hotels that join the program enjoy a 65% increase in the number of visits and a 35% increase in the number of bookings through the platform.

Joining the BPPP is rather a long-term strategic decision for a hotel, in the sense that hotels cannot decide to enter or exit the program extremely often. To increase the hotel's visibility for a shorter period of time, Booking.com proposes others programs such as the *Visibility Booster* that allows more flexibility in terms of timing and targeted consumers.

One may wonder whether the program really improves visibility for hotels. Indeed, if too many hotels are allowed to join the program, hotels are in the same position as before as they continue to compete for better position with all competitors. It is difficult to know exactly how many hotels have joined the program, especially at the period covered by our data. Today, Booking.com's website mentions that *"Preferred is an exclusive Programme that gives greater visibility to our top 30% of partners."*

In January 2019, we identified for a number of Scandinavian cities, the number of hotels listed on Booking.com as well as the number of hotels that had joined the program at that time. Figures are presented for 12 large cities in which some of our hotels are located in Table 2.10. We observe that in the cities in which the hotels from our sample that joined the program (in bold), the share of hotels that have joined the

gram as available here.

15 See detail at Booking.com

City	Country	All hotels	Preferred	% Preferred
Copenhagen	Denmark	**620**	**78**	**13%**
Odense		40	12	30%
Aarhus		44	11	25%
Aalborg		34	11	32%
Oslo	Norway	**254**	**29**	**11%**
Bergen		**173**	**27**	**16%**
Lillehammer		35	2	6%
Göteborg	Sweden	**108**	**18**	**17%**
Stockholm		258	68	26%
Uppsala		36	5	14%
Skövde		11	0	0%
Falun		21	2	10%

Table 2.10: Preferred Partner Program adoption

program in 2019 remains below 17%. Therefore, in these cities, hotels that join the program have a strong chance to get a major visibility boost.

Within our dataset, 6 of the 22 hotels have joined the program during our sample period. We obtained the exact dates at which each hotel joined the program from the hotel chain. Hotels joined at different times during the first semester of 2015 but they all remained in the program until the end of our observation period. Among the 6 hotels that joined the program (Table 2.11), 2 are located in Denmark (Copenhagen), 3 in Norway (Bergen and Oslo) and 1 in Sweden (Göteborg).

Country	Preferred	Standard	Total
Denmark	2	4	6
Norway	3	2	5
Sweden	1	10	11
Total	6	16	22

Table 2.11: Program adoption

These hotels have been specifically selected by the chain because they are located in large cities where they face tough competition (i.e., high number of competing hotels listed on Booking.com) and the chain wanted to guarantee high visibility on search results.

4.2 Empirical Strategy and Control Group

Our goal is to evaluate the effects of joining the BPPP on prices charged by hotels through different distribution channels. Because the decision to join the program is not exogenous, we cannot simply run a difference-in-difference analysis using hotels that did not join the program as a control group. Our strategy consists in constructing a control group by selecting hotels that were not included in the BPPP by the chain but that are reasonably comparable to the 6 hotels that were selected to join the program.

The 22 hotels were all eligible to join the BPPP (they all meet the same quality requirements, achieve good customer ratings, generate enough sales through Booking.com, ...). Therefore except for possibly different hotel characteristics that we already account for in our pricing equation, they essentially differ through the degree of competition that they face in their local markets.

Ideally we would have liked to perform propensity score matching (see, e.g., Rosenbaum and Rubin (1983)), estimating first the probability to adopt the program (i.e., being included in the treatment group) given hotel characteristics. We would have then associated to each treated hotel, hotels that could be considered closest matches that would then have been included in the control group. However, standard probit and logit models can only be properly estimated using a sufficient number of observations. Given the small number of hotels in our sample (22 hotels in total) and our strategy to conduct the analysis country by country, we cannot use this matching strategy. For instance, we only have 5 observations (i.e., hotels) in Norway and 6 in Denmark.

We rely instead on a nearest neighbour approach, method that can be more easily applied to small samples. For each treated hotel, we compute a distance to each non-treated hotel, restricting attention to hotels located in the same country.[16] We base our analysis on a large set of hotel characteristics already used in our hedonic price analysis. We keep the following variables to identify closest hotels within the same country: the number of competitors, distance to the airport and to the city center, number of rooms, star rating, number of amenities (e.g., bar, restaurant, fitness, etc.), city size.

To compute distance, we start by centering and reducing our variables (i.e., for each variable, we substract the mean and divide by the standard-deviation), thus obtaining reduced variables that are now comparable. For each pair of hotels, the compute the Euclidian distance (i.e., each variable is given the same weight) between hotels using the reduced variables.

Table 2.12 displays the estimated distances for each possible couple of hotels depending on the set of matching variables that we use. We first use only the number of competitors (first column) and observe that extremely helpful as we cannot really identify a best "control" for each hotel. We then successively add more controls (i.e., matching variables) such as star-rating, number of rooms, distance to the city center, distance to the closest airport, number of amenities and finally city size. We present the results for some of these combinations in columns 2 to 5. The grey cells present

[16] An alternative strategy could have been to include all non-treated hotels in the control group, controlling for all possible hotel characteristics in the econometric estimation. If we can guarantee that we control for all relevant variables, this should provide consistent results (see, e.g., Kim and Singal (1993)). Using this alternative method, we found results to be qualitatively similar.

the hotel that seem to be closest to our treated hotel. In the last column, we identify the hotels that we include in our control group.

			Matching Variables					
Country	Treated	Not treated	Comp. Dist.	# rooms Stars Comp. Dist.	Airport Center # rooms Stars Comp. Dist.	Amen. Airport Center # rooms Stars Comp. Dist.	City Size Amen. Airport Center # rooms Stars Comp. Dist.	Control
Sweden	SEGOTHOG	SESKOBIL	0.285	1.316	10.456	10.456	15.129	X
		SEKARCAR	0.274	12.238	15.231	15.231	20.043	
		SEKRICHR	0.262	17.643	20.145	21.165	25.884	
		SEBORBRA	0.257	16.887	21.715	22.735	27.099	
		SESUNSTR	0.246	3.032	16.072	16.072	20.580	
		SEFALGRA	0.230	3.748	8.513	8.513	13.026	X
		SEKARSTA	0.194	15.044	21.198	21.198	25.829	
		SEUPPLIN	0.157	10.184	10.256	11.276	15.337	X
		SEMALJOR	0.079	15.158	15.160	16.180	19.691	
		SESTOREI	0.682	7.889	9.654	9.654	11.376	X
Norway	NOOSLMIL	NOHAMVIC	1.850	7.247	7.868	6.440	17.318	
		NOLILBRE	1.454	1.625	5.840	4.487	9.933	X
	NOOSLGRI	NOHAMVIC	1.850	4.041	5.420	17.047	6.711	
		NOLILBRE	1.454	1.640	4.487	9.920	4.500	X
	NOBERMAR	NOHAMVIC	0.836	1.547	3.900	4.920	5.148	
		NOLILBRE	0.578	1.716	4.946	4.946	4.972	X
Denmark	DKCOPKOF	DKKOLHOT	10.954	14.915	16.494	25.673	23.339	
		DKAALAAL	10.414	14.859	14.998	19.078	15.357	X
		DKODEGRA	10.202	10.418	12.520	21.699	13.915	X
		DKAARATL	10.062	10.081	17.712	21.792	15.800	X
	DKCOPH27	DKKOLHOT	10.954	16.142	15.508	16.528	32.484	
		DKAALAAL	10.414	14.806	15.292	15.292	19.143	X
		DKODEGRA	10.202	11.377	11.705	12.724	22.889	X
		DKAARATL	10.062	12.906	15.799	15.799	21.792	X

Table 2.12: Selecting our control groups (one for each country)

Our sample includes 11 hotels in Sweden. Only one of them has joined the program, we thus have ten potential candidates for the control group. To select hotels to be included in the control group, we use the following “elbow rule”: we rank the computed distance between each candidate and the treated hotel, and look for a significant drop between two consecutive distances. Using this simple rule, we identify four hotels to be included in the control group (i.e., we “cut” where the distance suddenly increases from 15.3 to 19.7). We also observe that using different sets of matching variables to compute distances does not affect the composition of the control group.

For Norway, we have three treated hotels and only two candidates for the control group. We observe that one of the candidates is close to all three treated hotels (NOLILBRE). The second candidate is also relative close but the distance increases substantially for one of the treated hotels. We thus decided to exclude it from the control group, group that thus includes only one hotel.

In Denmark, the sample includes two treated hotels (both located in Copenhagen) and four candidates for the control group. When we use all the matching variables, the distances are very similar for three of the candidates (between 19.1 and 22.9) but substantially larger for the last one (32.5). We thus opted to keep only three hotels in the control group. However, when we do not account for city size, distance are comparable for all for hotels. We thus could have opted to include all four in the control group. We present the results with this alternative control group in Appendix 2.A, showing that results are qualitatively similar.

Based on this analysis, we thus constructed control groups that include four hotels in Sweden (for one treated hotel), one in Norway (for three treated hotels) and three in Denmark (for two treated hotels).

In order to test the robustness of our analysis and the validity of our control group, we check hotel characteristics and degree of local competition are the main factors influencing the decision to include a given hotel in the BPPP. In our analysis, we control for as many hotel characteristics as possible (including information on the degree of market competition). In addition, we based the construction of our control group on these characteristics (used to compute the "distance" between hotels).

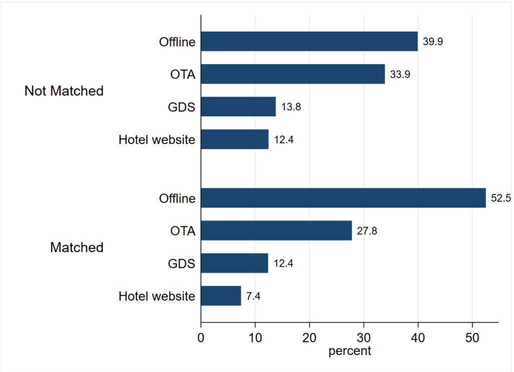
We thus essentially need to check that characteristics that influence prices are all observable and accounted for in our pricing equation. To do this, we identify identical bookings, i.e., bookings with identical characteristics, and compare the prices paid by the different consumers. We consider that our observables include all characteristics that influence prices.

Our dataset includes a total of 824,048 bookings that can be divided into 501,639 unique reservations (i.e., bookings that cannot be matched) and 322,409 bookings that can be matched to another one.[17] Matched bookings correspond to 94,452 different types, 63% of them being observed twice (i.e., we match two identical bookings), 16% being observed three times, 22% being observed four times or more.

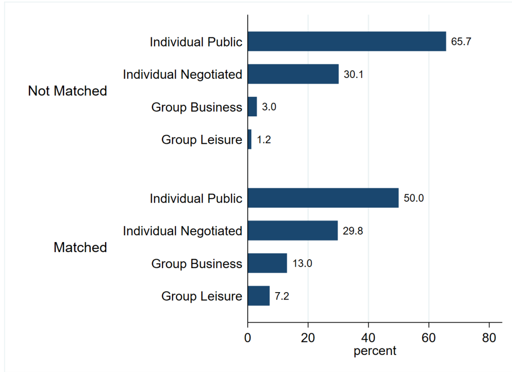
In Figure 2.3 and Table 2.13, we show that this matching procedure does not generate any selection bias and that matched and unmatched bookings are similar, i.e., the distributions of characteristics remain comparable.

We then restrict attention to matched bookings and compare prices for each pair of matched bookings. We observe that matched booking prices are identical in about 98% of cases (see Table 2.14). We also observe that when prices differ, the price difference is rather large. One possible explanation for such large price differences – in very rare cases – may be related to add-ons paid by consumers on their final bill but that do not reflect a real room price difference. Overall, this analysis confirms that bookings with identical characteristics also have the same price and that we do not miss any important variable that may affect room prices.

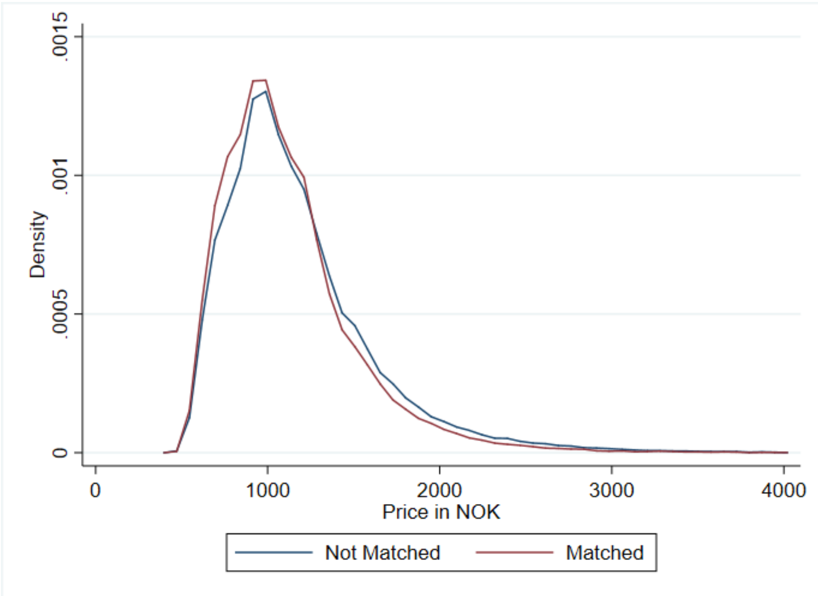
[17] This analysis has been done by looking at all bookings in our initial dataset, not by restricting attention to online bookings only.



(a) Channel



(b) Rate



(c) Price

Figure 2.3: Samples comparisons

Mean (Std. Error)	Not matched	Matched
Lead Time [days]	25 (38)	27 (38)
Nights [days]	1.8 (1.8)	1.7 (1.6)
Persons	1.6 (0.7)	1.4 (0.6)
Standard room	75%	85%
Week-end	31%	27%
Refundable offer	21%	17%
Occupancy rate [booking]	45% (29%)	45% (29%)
Occupancy rate [final]	73% (29%)	77% (28%)

Table 2.13: Sample comparisons

# occurrences	# types	Proportion	% identical prices	Price difference
2	58,654	62%	97.2%	18.0 %
3	14,911	16%	97.4%	21.8%
4+	20,887	22%	98.0%	35.6%
All	94,452	100%	98.1%	34.3%

Table 2.14: Pairwise price comparison

4.3 Effect on prices

Having now selected our control groups, we estimate - for each distribution channel and each country - the following difference-in-difference equation for log-prices:[18]

$$\ln P_{iht} = \alpha + \beta.X_h + \gamma.Y_{ht} + \delta.Z_{iht} + \zeta \mathbb{1}_{Treated} + \eta \mathbb{1}_{After} + \theta \mathbb{1}_{After \times Treated} + \varepsilon_{iht} \tag{2.2}$$

For each country, we run separate regressions for the three distribution channels. However, prices set for sales through different sales channels by a given hotel are not independent as they are derived from the same maximization. OTAs indeed operate under the agency model meaning that hotels keep control of the prices they charge for their rooms even if the rooms are sold through an intermediary such as Booking.com or Expedia. OTAs do not control prices and generate revenues through the commission fee charged to the hotel for each room booking it generates. One may also wonder whether prices are set at the hotel level, or centrally at the chain level. However, as hotels in our sample are located in different cities (with the exception of our two hotels located in Copenhagen), that is in different local markets. Even if prices are set centrally so as to maximize the chain's profit, the maximization program essentially boils down to separate maximization programs for the different hotels, the central manager maximizing each hotel's profit separately.

It is therefore highly likely that error terms can be correlated across channels. To account for this correlation, we rely on the SURE procedure introduced by Zellner (1962) in the standard OLS case to estimate *"Seemingly Unrelated Regression Equations"*. However, because we also account for endogeneity of the occupancy rate variables, we first use instrumental variables and estimate SURE with two-stage least squares rather than OLS. Doing so requires a three-step procedure. We first regress the endogenous variables (OR_{Book} and OR_{Final}) on all exogenous variables and we compute the endogenous variables' predicted values. We then estimate the system of equations by least squares, replacing the endogenous right-hand-side variables with their predicted values. Finally, we compute the estimated coefficients of the variance-covariance matrix for the

18 The observables are the same as in the hedonic price regression (i.e., hotel and booking characteristics) except the "preferred" dummy that is now excluded from Y_{ht} (as it is now the "Treated" dummy). We additionally include as a control variable the city size, introduced in the control group construction.

residuals from the second estimation step, and re-estimate the pricing equations using the SURE procedure.

We therefore perform a three-stage least-squares estimation (3SLS) that corrects for any correlation across error terms for the per-channel pricing equations (Zellner and Theil, 1962). Estimating coefficient through 3SLS (combining IV and SURE) rather than through IV only, we obtain more efficient and consistent parameter estimates. Table 2.15 presents the results of the 3SLS (i.e., IV + SURE) estimation procedure. The results for the OLS, SURE and IV estimations are presented in Appendix 2.B.

Country	Sweden (N=2,714)			Denmark (N= 3,022)			Norway (N=1,944)		
Channel	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
Estimation procedure: **3SLS (IV + SURE)**									
After	-0.025**	-0.006	0.013	-0.095***	-0.140***	-0.116***	-0.015	-0.182***	-0.124***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.04)	(0.04)	(0.04)
Treated	5.051***	4.774***	4.892***	0.225***	0.146***	0.167***	0.239***	0.231***	0.200***
	(0.30)	(0.32)	(0.36)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)
After × Treated	0.066***	0.028**	0.040**	0.049***	0.072***	0.036***	0.021	0.124***	0.105***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.03)	(0.04)	(0.03)
F-Stat									
OR_{Book}		78.95			88.82			45.08	
OR_{Final}		98.99			45.25			14.79	

Standard errors (clustered by hotel and date) in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.15: Difference-in-Difference estimates (prices)

Results show that joining Booking.com's Preferred Partner Program unambiguously led hotels to increase prices on all channels relative to prices charged by non-joiners (i.e., hotels within the control group). This thus confirms that the negative (sometimes large) effect estimated using our hedonic price regression was essentially due to selection bias. Hotels that join the program are located in highly competitive markets and this explains why prices are on average lower for this selection of hotels. However, once we compare price for these hotels with hotels located in similar markets, this selection bias disappear and we can then better evaluate the true effect of joining the visibility booster program. We focus first on Denmark ans Sweden and look at our variable of interest, i.e., the interaction term "*After* × *Treated*". We observe similar results with hotels that joined the BPPP increased prices relatively uniformly in all online distribution channels: prices increased by about 4% on Expedia, 5% to 7% on Booking and 3% to 7% for direct sales.

The generalized price increase may be due to a combination of factors. Firstly, the commission paid to Booking increased due to the BPPP as hotels are willing to pay for prominence (better ranking). In addition, commissions paid to OTAs may have increased as a side effect of the introduction of price parity clauses (reduced competition between OTAs on commissions). Under price parity, higher commissions imposed by Booking.com affect all prices and not simply the price set on that platform.

However, joining the program improves the hotels' visibility. If demand increases unconditionally simply because hotels are now better ranked, hotels may be able to increase prices (and profits) without losing too many customers. This would then partially explain the generalized price increase.

Finally, the introduction of price parity may have forced hotels to increase prices for direct sales and to decrease prices set on more expensive platforms as a standard effect of removing price-discrimination between channels. This could thus explain why, at least in Denmark, the price increase has been larger for direct sales than for sales through OTAs.

Overall, it appears that prominence is unlikely to have led to price decreases thus contradicting some of the theoretical results (e.g., Armstrong et al. (2009) or Armstrong (2017)) and rather suggests that either the increased cost or the shift in demand are the cause of the generalized price increase.

Results are slightly less convincing for Norway, also they again that prices have increased for direct sales and for sales through Expedia, and even more than in Sweden and Denmark as prices increased by as much as 10% to 12%. The difference could however partially related to our difficulties to constitute a good control group (only one hotel as a control).

4.4 Effect on quantities

Depending on the rationale behind the price increase, the hotel can expect a different impact on its sales. Prices increases due to increased costs and the reintroduction of price parity clauses should led to lower sales on all channels. But it would then be compensated, at least on Booking.com, as the hotel also benefit from increased visibility according to the platform. Overall effect might then be ambiguous on Booking.com but we would expect lower sales for the other two channels. If the hotel simply took advantage of increased demand (due to increased visibility on Booking.com) to increase prices on Booking.com, we should also expect sales to have increased through that channel, possibly cannibalizing sales through other online channels. To understand better the effects of the BPPP, we estimate in this section, the effect of the BPPP on online sales, adopting the same difference-in-difference approach as before. This time, we want the estimate - for each distribution channel (and each country) - the following equation:

$$Q_{ht} = \alpha + \beta.\bar{X}_h + \gamma.\bar{Y}_{ht} + \delta.\bar{Z}_{ht} + \zeta \mathbb{1}_{Treated} + \eta \mathbb{1}_{After} + \theta \mathbb{1}_{After \times Treated} + \varepsilon_{ht} \tag{2.3}$$

where Q_{ht} is the proportion of available rooms booked through the relevant channel for an arrival at date t in hotel h.[19] To make "quantities" comparable across different hotels, we divide the total number of bookings made through that channel by the hotel's total capacity. We control for time-invariant hotel characteristics $\bar{X}_h$ that are the same as before except that we now exclude the hotel's capacity, as well as for time-varying hotel characteristics $\bar{Y}_{ht}$ which now essentially consist of TripAdvisor ratings: because bookings with an arrival date equal to t have been made at different times, we compute the average rating over all bookings made for date t. In the similar vein, the vector $\bar{Z}_{ht}$ now includes the average characteristics of all bookings made for date t (same booking characteristics than in Z_{iht}) but also includes the average price paid by consumers having made a reservation for date t through the relevant channel.[20]

As we regress quantities on average prices, we again face an endogeneity problem. We thus need to instrument the average price in each equation (one per channel) and use two different instruments. The first instrument is common to all three channels but country-specific and is simply the hourly wage in each country.[21] The idea is simply that cost-shifters may be good instruments as they affect hotel prices but uncorrelated with unobserved demand shocks. The second instrument is then hotel and channel-specific, but also different for different types of tariffs. To instrument the average price charged by hotel h for a booking at date t through channel c, we use the average price charged for bookings at date t through the distribution channel by hotels located in the same country but in different cities than hotel h. We also restrict attention to prices charged for the same type of tariffs (refundable vs. flexible). Here the intuition is again that hotels in the same country may share common cost shocks that affect their pricing strategies but because they are located in different cities, they do not compete directly for the same set of consumers.

As for prices, it is likely that the error terms of our three quantity equations (one for each channel) are correlated and we therefore estimate the equations through 3SLS (combining IV and SURE). The results are presented in Table 2.16.

Looking again at Denmark and Sweden first, we observe that joining the BPPP has a positive impact on quantities sold through OTAs, with a larger impact on Booking.com butt a negative impact on direct sales. Sales through Booking.com increased by 3 and 7 percentage points in terms of occupancy rate in Sweden and Denmark respectively,

[19] In Appendix 2.C, we discuss an alternative way to define quantities that provides similar estimation results.

[20] As we estimate one equation per channel, we have omitted the subscript c everywhere in equation (2.3).

[21] These are seasonally adjusted average total earnings paid per employed person per hour, including overtime pay and regularly recurring cash supplements (reported on a quarterly basis). The data has been collected from OECD statistics.

Country	Sweden (N=2,837)			Denmark (N= 3,334)			Norway (N=2,246)		
Channel	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
	Estimation procedure: **3SLS (IV + SURE)**								
After	-0.035***	-0.005**	-0.013***	0.022***	0.002	-0.001	-0.056***	-0.018***	-0.071***
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Treated	-0.669***	-0.168***	-0.421***	0.000	0.040***	0.097***	0.026**	0.060***	0.126***
	(0.14)	(0.06)	(0.11)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
After × Treated	0.031***	-0.006**	0.009*	0.073***	-0.010***	0.019***	-0.002	0.003	0.030***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)
F-Stat Price	181.31	194.82	124.44	82.04	88.25	57.51	49.79	50.73	64.35

Standard errors (clustered by hotel and date) in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.16: Difference-in-Difference estimates (quantities)

but only by 1 and 2 points of occupancy for sales through Expedia. Direct sales decreased in both countries but the effect remained no larger than 1 percentage point. This confirms that the visibility program indeed increases sales through Booking.com but, although these additional sales partly come at the expense of other online channels, online sales increased overall. Despite the uniform price increase for online sales, hotels that have joined the program have indeed managed to increase online sales overall.

Results for Norway are once again less convincing, as only the impact on sales through Expedia is statistically significant (similar in magnitude to what is observed in other countries).

5 Impact of Price Parity

As discussed in the previous section, one essential element of the BPPP was to restore price parity clauses. Such clauses had been imposed by OTAs until 2015 but were then partially or totally banned by competition agencies or national legislations in many countries, especially in Europe. In 2015, Booking.com offered to the French, Italian and Swedish competition authorities to switch from wide (linking all prices) to narrow price parity clauses (linking only the OTA's price to the direct online price). They also unilaterally modified their contracts in the European Economic Area (including our three countries) and switched from wide to narrow clauses, unless even narrow clauses were banned (like in France [Loi Macron, summer 2015], Germany [Bundeskartellamt's decisions against HRS and Booking.com], Austria [by law after 2016], Italy [by law after 2017] or Belgium [by law as of 2018]). The introduction of the BPPP was potentially one way to bypass the ban and reintroduce – at least for some selected hotels – wide price parity clauses.

Competition agencies have claimed that price parity clauses eliminated competition over commission fees between platforms, thus increasing the hotels' distribution costs and therefore harmed consumers through higher room prices. However, there is very little empirical evidence on the effects of price clauses. In this section, we try evaluate the effects of price parity clauses, disentangling the effect of improved visibility (through the BPPP) and of the price parity clause itself. To do this, we take advantage of variations in compliance with price parity as we discuss below.

5.1 Compliance with price parity clauses

In this section, we evaluate compliance with price parity separately for preferred and for standard hotels (i.e., depending on whether the hotels joined or not the BPPP) and separately for refundable and non-refundable tariffs.

To evaluate compliance, we start by comparing prices charged by a hotel through two different distribution channels (Booking, Expedia or Direct [hotel's website]) for identical bookings. We thus start by identifying identical bookings, i.e., pairs of bookings that have identical characteristics except for the distribution channel: these are thus bookings for identical types of rooms in the same hotel, booked for the same night on the same day, for the same number of travellers and same number of nights. In addition, bookings are considered identical only if the type of offer - i.e., non-refundable offer or flexible tariff - is the same for both bookings. For two identical bookings made through two different platforms, we compare the prices paid by consumers. To check compliance with a price parity imposed by Booking.com (as part of the program), we can impose either a strict rule or a softer rule. The strict rule imposes that the direct price has to be at least equal to price charged through Booking.com. A softer rule allows the hotel to offer a limited discount on its own direct channel (e.g., ± 10%).

Figure 2.4 present the probability that two identical bookings, the reservation made through Booking.com was not more expensive than the one made directly through the hotel's website. We distinguish between refundable (yellow) and non-refundable (blue) offers, and between standard (full line) and preferred hotels (dashed line). The vertical red line indicates the introduction of the BPPP. We observe that all hotels have always applied price parity rather strictly (for more than 80% of all bookings) for refundable offers but hotel's behavior has changed over time regarding flexible tariffs: whereas standard hotels have never strictly complied with price parity for such tariffs (at most for 20% of bookings), preferred hotels started to apply price parity once they joined the BPPP.

To ensure that this change in compliance with price parity is not driven by our very strict definition of price parity, we also drew in Figure 2.5 the same graph for a much softer compliance condition, allowing the direct price to be at most 10% lower than the price charged through Booking.com. We observe that the two figures are almost identical. This means that when two identical bookings correspond to different prices,

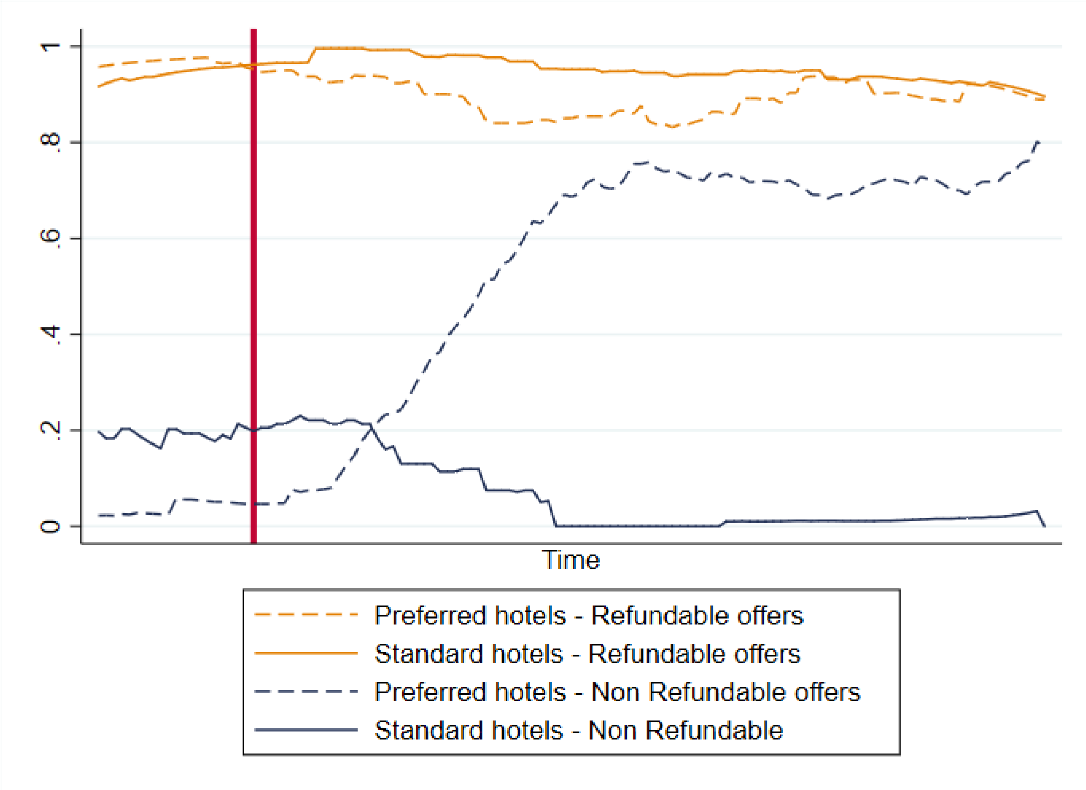


Figure 2.4: Price parity compliance – Booking.com vs. Direct sales

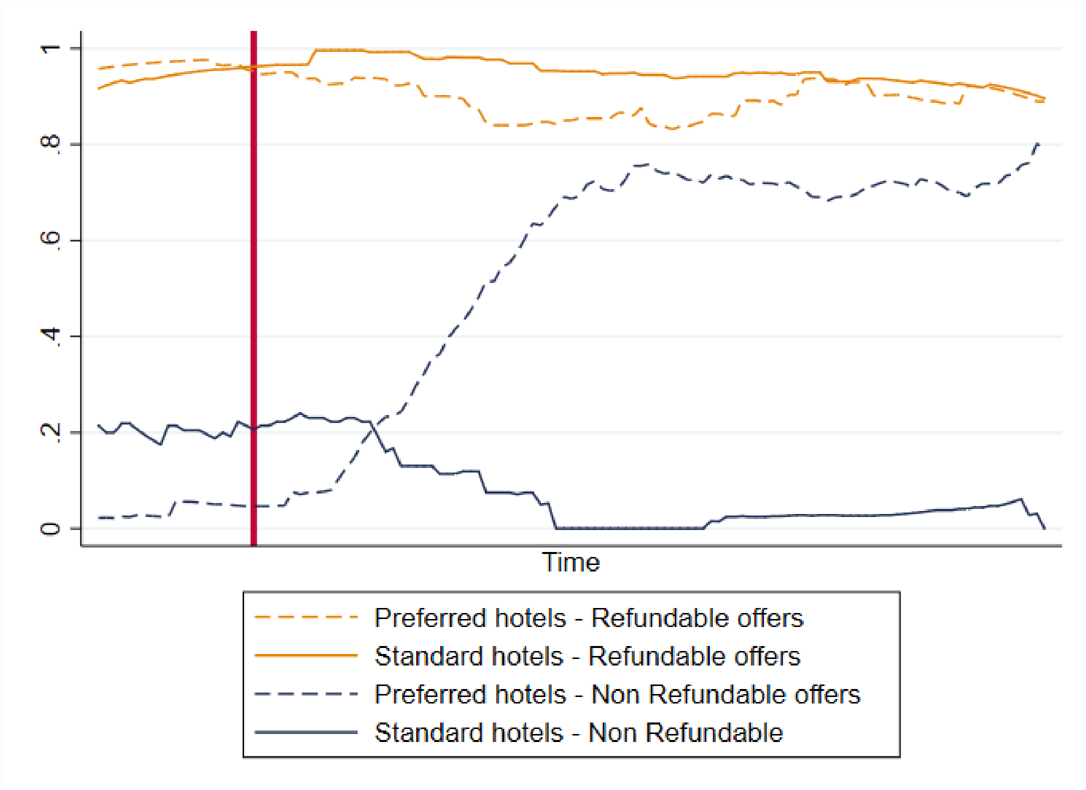


Figure 2.5: Price parity compliance (±10%) – Booking.com vs. Direct sales

the difference is usually quite important (exceeding 10%).

5.2 Empirical strategy

As we have just seen, it appears that before the BPPP was introduced, hotels were complying with (or simply self-applying) price parity across online distribution channels for refundable tariffs but not for non-refundable tariffs. Hotels that did not join the BPPP did not change behaviour regarding price parity once the program was introduced (if anything they started offering discount on almost non-refundable direct bookings), whereas hotels that joined the program started to comply with price parity for non-refundable tariffs as well.

We use this distinction between refundable and non-refundable tariffs for the two types of hotels to estimate the effect of the price parity clause on prices. This is feasible because of the way the Booking's Preferred Partner Program works and search results are presented. We understand from our discussion with the chain that provided the data, that hotels pay the same commission for refundable and non-refundable offers. Joining the BPPP thus similarly affects both types of tariffs. Moreover, the search algorithm ranks hotels not tariffs. Preferred hotels are thus presented higher on the search results list and consumers have to click on the hotel to see a list of offers (different types of rooms, with or without breakfast, refundable or flexible tariff).

The improved visibility and increased commission associated with the BPPP thus affect all types of offers at preferred hotels (relative to standard hotels), whereas the price parity element directly affects the non-refundable offers for preferred hotels only.

We thus estimate a triple difference price equation that look similar to our double difference equation (2.2) but now includes interactions between the *After* and/or *Treated* dummies with the dummy for the type of offer, i.e., $\mathbb{1}_{NR}$ that takes value one if and only the tariff is non-refundable (remember that this dummy variable was already included in Z_{iht}). As for our difference-in-difference analysis, we run separate regressions for the three distribution channels and for the three countries. The modified pricing equation for observation i for a booking at hotel h for a stay at time t (thus omitted per-channel and per-country subscripts for all variables and coefficients) is therefore:

$$\begin{aligned} \ln P_{iht} =& \alpha + \beta.X_h + \gamma Y_{ht} + \delta.Z_{iht} + \zeta \mathbb{1}_{Treated} + \eta \mathbb{1}_{After} + \theta \mathbb{1}_{After} \mathbb{1}_{Treated} \\ & + \kappa \mathbb{1}_{Treated} \mathbb{1}_{NR} + \lambda \mathbb{1}_{After} \mathbb{1}_{NR} + \mu \mathbb{1}_{After} \mathbb{1}_{Treated} \mathbb{1}_{NR} + \varepsilon_{iht} \end{aligned} \tag{2.4}$$

and our main parameter of interest is now the coefficient μ measuring the impact of the BPPP on prices of non-refundable tariffs for hotels that joined the program. As before we need to take care of the potential endogeneity issues with our occupancy rate variables (included in Z_{iht}) and do this using the same instruments as in the earlier estimations of price equations (2.1) and (2.2). We again estimate three pricing equations (on for each distribution channel) for which prices are derived from a common profit maximization program. As for our difference-in-difference approach we rely on the SURE estimators. Combining the two approaches, IV and SURE, we estimate equation (5.2) through three-stage least squares (3SLS).

5.3 Results

Table 2.17 present the results of the different estimation methods for the triple difference approach. Results are not very conclusive as we essentially get non statistically significant coefficients.

The only significant coefficient, suggest that prices have substantially decreased for our treated hotels in Norway for sales through Booking. This would suggest that the de facto ban on price discrimination had the expected effect (prices decrease where they were initially more expensive). Results suggest that flexible tariff offers decreased by 18% in that channel. The other signs are also as expected (positive for the direct sales and negative for Expedia) although the coefficient are not significant. Compliance has no statistically significant effect on prices for flexible offers sold though Expedia or the hotels' direct sales channels. In the case of Sweden and Denmark, we find no statistically significant effect on online prices. This last result is in line with Mantovani et al. (2020) who find no (medium-term) effect of the legal ban on narrow price parity clauses in France and in Italy.

Country	Sweden (N=2,714)			Denmark (N= 3,022)			Norway (N=1,944)		
Channel	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
OLS	0.092*	0.066	-0.040	-0.125***	-0.055	-0.021	-0.180**	0.106	-0.134
	(0.05)	(0.06)	(0.05)	(0.03)	(0.04)	(0.04)	(0.09)	(0.09)	(0.08)
SURE	0.039	0.062	-0.015	-0.062**	-0.031	-0.019	-0.162**	0.100	-0.098
	(0.05)	(0.06)	(0.05)	(0.03)	(0.04)	(0.03)	(0.08)	(0.08)	(0.07)
IV	0.076	0.064	-0.029	-0.117***	-0.037	-0.039	-0.217**	0.113	-0.132
	(0.05)	(0.06)	(0.06)	(0.03)	(0.04)	(0.04)	(0.09)	(0.08)	(0.08)
3SLS	0.026	0.060	-0.003	-0.050	-0.022	-0.026	-0.181**	0.098	-0.095
	(0.05)	(0.06)	(0.05)	(0.03)	(0.04)	(0.03)	(0.08)	(0.08)	(0.07)
F-Stat									
OR_{Book}		70.12			79.32			40.15	
OR_{Final}		88.51			40.31			13.39	

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.17: Triple Difference results on prices

Our analysis thus suggests that the positive impact of the BPPP on prices is not due to the price parity constraint imposed by Booking.com but rather to the other components of the program (higher commission and increased visibility). It also suggests that when deciding to comply with price parity (for flexible tariffs), the hotels chose to restore parity by lowering the price on Booking.com rather than by increasing the direct price.

6 Conclusion

In this paper, we focused on the impact of Booking.com's Preferred Partner Program (BPPP) on consumers prices using exhaustive booking data from hotels belonging to the same Scandinavian chain. Using a difference-in-differences approach, we showed that joining the program led "preferred" hotels to increase prices through all channels. This first important result also confirms that a simple price regression is not sufficient to evaluate the effect of this program as hotels do not randomly join the program (in our case, the chain picked hotels that were located in very competitive markets). In addition, we showed that quantities sold through OTAs increased when joining the program, including through Expedia even if to a lower extent, but direct sales decrease. However, the small decrease in volumes sold directly is more than compensated by the increased sales through OTAs. These results suggest that joining the program generates a positive shift in demand for the hotel, not only on Booking.com but potentially also on competing platforms, supporting the idea that "showrooming" between OTAs may be at play here.

Finally, taking advantage of differences in compliance with price parity clauses, we evaluated the specific effects of price parity clauses (one important component of Booking.com's program) and showed that price parity conditions led to price reductions being offered by the hotels for non-refundable offers through Booking.com but had no significant effect on prices of such offers through other online channels. This result thus contradicts the theory literature on the anticompetitive effects of price parity clauses and the main theory of harm developed by competition agencies in the various cases related to price parity clauses imposed by the leading OTAs (Booking.com, Expedia or to a more limited extent HRS).

Appendices

2.A Alternative Control group for Denmark

2.A.1 Effect on prices

Country	Sweden (N=2,714)			Denmark (N= 2,405)			Norway (N=1,944)		
Channel	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
Estimation procedure: **3SLS (IV + SURE)**									
After	-0.025**	-0.006	0.013	-0.161***	-0.188***	-0.199***	-0.015	-0.182***	-0.124***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)
Treated	5.051***	4.774***	4.892***	-0.863***	-0.717***	-1.188***	0.239***	0.231***	0.200***
	(0.30)	(0.32)	(0.36)	(0.19)	(0.23)	(0.24)	(0.04)	(0.04)	(0.04)
After × Treated	0.066***	0.028**	0.040**	0.128***	0.134***	0.123***	0.021	0.124***	0.105***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)	(0.04)	(0.03)
F-Stat									
OR_{Book}		78.95			80.14			45.08	
OR_{Final}		98.99			41.69			14.79	

Standard errors (clustered by hotel and date) in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.18: Difference-in-Difference estimates (prices)

2.A.2 Effect on quantities

Country	Sweden (N=2,837)			Denmark (N= 2,621)			Norway (N=2,246)		
Channel	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
Estimation procedure: **3SLS (IV + SURE)**									
After	-0.035***	-0.005**	-0.013***	0.012	-0.006	0.006	-0.056***	-0.018***	-0.071***
	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Treated	-0.669***	-0.168***	-0.421***	-0.184*	-0.123**	0.167**	0.026**	0.060***	0.126***
	(0.14)	(0.06)	(0.11)	(0.10)	(0.05)	(0.08)	(0.01)	(0.00)	(0.01)
After × Treated	0.031***	-0.006**	0.009*	0.088***	0.001	0.014**	-0.002	0.003	0.030***
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
F-Stat Price	181.31	194.82	124.44	72.49	73.01	49.49	49.79	50.73	64.35

Standard errors (clustered by hotel and date) in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.19: Difference-in-Difference estimates (quantities)

Country	Sweden (N=2,682)			Denmark (N= 2,234)			Norway (N=1,763)		
Channel	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
Estimation procedure: **3SLS (IV + SURE)**									
After	-0.006**	-0.002	-0.003	0.005	-0.001	0.008**	-0.027***	-0.010**	-0.043***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)
Treated	-0.037	-0.055	-0.223***	-0.173***	-0.089**	0.029	0.019**	0.044***	0.099***
	(0.10)	(0.04)	(0.07)	(0.06)	(0.04)	(0.06)	(0.01)	(0.00)	(0.01)
After × Treated	0.006	-0.004**	0.005*	0.053***	0.002	0.015***	0.018**	0.006	0.029***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)
F-Stat Price	152.02	148.68	94.30	50.60	43.65	35.73	27.86	29.94	43.77

Standard errors (clustered by hotel and date) in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.20: Difference-in-Difference estimates (alternative quantities)

2.A.3 Triple difference - Effect of price parity

Country	Sweden (N=2,714)			Denmark (N= 2,405)			Norway (N=1,944)		
Channel	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
OLS	0.092**	0.066	-0.040	-0.129***	-0.059	-0.045	-0.180**	0.106	-0.134
	(0.05)	(0.06)	(0.05)	(0.04)	(0.05)	(0.04)	(0.09)	(0.09)	(0.08)
SURE	0.039	0.062	-0.015	-0.068**	-0.044	-0.038	-0.162**	0.100	-0.098
	(0.05)	(0.06)	(0.05)	(0.03)	(0.05)	(0.04)	(0.08)	(0.08)	(0.07)
IV	0.076	0.064	-0.029	-0.124***	-0.044	-0.065	-0.217**	0.113	-0.132
	(0.05)	(0.06)	(0.06)	(0.04)	(0.05)	(0.04)	(0.09)	(0.09)	(0.08)
3SLS	0.026	0.060	-0.003	-0.062**	-0.038	-0.049	-0.181**	0.098	-0.095
	(0.05)	(0.06)	(0.05)	(0.03)	(0.05)	(0.04)	(0.08)	(0.08)	(0.07)
F-Stat									
OR_{Book}		70.12			71.64			40.15	
OR_{Final}		88.51			37.25			13.39	

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.21: Triple Difference results on prices

2.B Model comparisons

Country	Sweden (N=2,714)			Denmark (N= 3,022)			Norway (N=1,944)		
Channel	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
Estimation procedure: **OLS**									
After	0.024**	0.036***	0.051***	-0.060***	-0.108***	-0.065***	-0.011	-0.162***	-0.128***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.04)	(0.03)
Treated	-0.726***	-0.814***	-0.436***	0.271***	0.191***	0.236***	1.544***	1.537***	1.576***
	(0.06)	(0.06)	(0.06)	(0.01)	(0.01)	(0.01)	(0.14)	(0.16)	(0.14)
After × Treated	-0.001	-0.014	0.008	0.033***	0.061***	0.019*	0.019	0.100***	0.100***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.03)	(0.04)	(0.03)
Estimation procedure: **SURE**									
After	0.021*	0.036***	0.051***	-0.059***	-0.105***	-0.064***	-0.016	-0.166***	-0.132***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.04)	(0.03)
Treated	-0.754***	-0.802***	-0.441***	0.268***	0.190***	0.233***	1.566***	1.547***	1.591***
	(0.06)	(0.06)	(0.06)	(0.01)	(0.01)	(0.01)	(0.13)	(0.16)	(0.13)
After × Treated	0.010	-0.017	0.006	0.035***	0.061***	0.019*	0.023	0.102***	0.101***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.03)	(0.04)	(0.03)
Estimation procedure: **IV**									
After	0.019	0.033**	0.047***	-0.097***	-0.148***	-0.126***	-0.014	-0.164***	-0.133***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.04)	(0.03)
Treated	-0.763***	-0.953***	-0.308***	0.226***	0.139***	0.158***	1.520***	1.510***	1.585***
	(0.07)	(0.07)	(0.07)	(0.02)	(0.02)	(0.02)	(0.14)	(0.16)	(0.14)
After × Treated	-0.001	-0.013	0.004	0.048***	0.073***	0.039***	0.019	0.100***	0.102***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.03)	(0.04)	(0.03)
Estimation procedure: **3SLS (IV + SURE)**									
After	-0.025**	-0.006	0.013	-0.095***	-0.140***	-0.116***	-0.015	-0.182***	-0.124***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.04)	(0.04)	(0.04)
Treated	5.051***	4.774***	4.892***	0.225***	0.146***	0.167***	0.239***	0.231***	0.200***
	(0.30)	(0.32)	(0.36)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)	(0.04)
After × Treated	0.066***	0.028**	0.040**	0.049***	0.072***	0.036***	0.021	0.124***	0.105***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.03)	(0.04)	(0.03)
F-Stat									
OR_{Book}		78.95			88.82			45.08	
OR_{Final}		98.99			45.25			14.79	

Standard errors (clustered by hotel and date) in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.22: Difference-in-Difference Estimations (Prices)

2.C Effect on quantities - Alternative specification

We previously defined the quantity Q_{ht} as the number of bookings made through the relevant channel for an arrival at date t in hotel h (divided by the hotel's total capacity). Rather than focusing on booking made for arrival at a given date t, an alternative approach would be to focus on all bookings made on the same day τ (i.e., all observations with identical values for $\tau \equiv t - x$). In this alternative specification, some of our control variables are now common to all bookings. This is for instance the case for the TripAdvisor reviews or prices that depend on the booking date not on the arrival date.

This time, we estimate - for each distribution channel (and each country) - the following equation:

$$\widetilde{Q}_{h\tau} = \alpha + \beta . X_h + \gamma Y_{h\tau} + \delta . \widetilde{Z}_{h\tau} + \zeta \mathbb{1}_{Treated} + \eta \mathbb{1}_{After} + \theta \mathbb{1}_{After \times Treated} + \varepsilon_{h\tau} \qquad (2.5)$$

where $\widetilde{Q}_{h\tau}$ is the number of bookings made in hotel h through the relevant channel at date τ. To make "quantities" comparable across different hotels, we divide the total number of bookings by the hotel's total capacity. We control for time-invariant hotel characteristics X_h and TripAdvisor rating ($Y_{h\tau}$) at date τ, as well as for the average characteristics of all bookings ($\widetilde{Z}_{h\tau}$) made at date τ. We use similar instruments as for the quantities based on arrival date t. The first instrument is now the country-specific wage at the reservation date τ. The second instrument is simply the average price for all bookings made at date τ through the same channel by hotels located in the same country but in different cities than hotel h. The results are presented in Table 2.23.

Country	Sweden (N=2,682)			Denmark (N= 2,841)			Norway (N=1,763)		
Channel	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
	Estimation procedure: **3SLS (IV + SURE)**								
After	-0.006**	-0.002	-0.003	0.011***	0.005**	0.004	-0.027***	-0.010**	-0.043***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)
Treated	-0.037	-0.055	-0.223***	0.059***	0.016***	0.070***	0.019**	0.044***	0.099***
	(0.10)	(0.04)	(0.07)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
After × Treated	0.006	-0.004**	0.005*	0.042***	-0.004*	0.021***	0.018**	0.006	0.029***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)
F-Stat Price	152.02	148.68	94.30	63.55	43.47	56.39	45.61	29.94	43.77

Standard errors (clustered by hotel and date) in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.23: Difference-in-Difference estimates (alternative quantities)

Looking first at Denmark, we again observe that joining the BPPP has a positive impact on quantities sold through OTAs, with a larger impact on Booking.com (+4 percentage points of occupancy rate) than on Expedia (+2 percentage points of occupancy rate). The effect on direct sales is still negative and statistically significant but it is very limited (less than one point of occupancy rate). Results are this very similar to the initial specification based on quantities measured at arrival date.

Results for Sweden are relatively similar with the difference that the effects on sales through OTAs are now much smaller and not always statistically significant.

Finally, for Norway, we find a statistically significant impact on sales through OTAs but slightly smaller for Booking (+2 percentage points) than for Expedia (+3 points). Effect on direct sales is now marginally positive but is not statistically significant.

Bibliography

Abadie, A., Diamond, A. and Hainmueller, J. (2010), 'Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program.', *Journal of the American Statistical Association* **105**(490).

Abadie, A. and Gardeazabal, J. (2003), 'The economic costs of conflict: A case study of the basque country.', *The American Economic Review* **93**(1).

Abrate, G. and Viglia, G. (2016), 'Strategic and tactical price decisions in hotel revenue management', *Tourism Management* **55**, 123–132.

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

Arbatskaya, M. (2007), 'Ordered search', *Rand Journal of Economics* **38**, 119–126.

Armstrong, M. (2017), 'Ordered consumer search', *Journal of the European Economic Association* **15**, 989–1024.

Armstrong, M., Vickers, J. and Zhou, J. (2009), 'Prominence and consumer search', *Rand Journal of Economics* **40**(2), 209–233.

Armstrong, M. and Zhou, J. (2011), 'Paying for prominence', *Economic Journal* **121**(556), F368–F395.

Boik, A. and Corts, K. (2016), 'The effects of platform most-favored-nation clauses on competition and entry', *Journal of Law and Economics* **59**, 105–134.

Calzada, J., Manna, E. and Mantovani, A. (2019), 'Platform price parity clauses and segmentation', *UB Economics Working Papers E19/387* .

Cazaubiel, A., Cure, M., Johansen, B. O. and Vergé, T. (2020), 'Substitution between online distribution channels: Evidence from the oslo hotel market', *International Journal of Industrial Organization* **69**, 102577.

Chen, Y. and Yao, S. (2016), 'Sequential search with refinement: Model and application with click-stream data', *Management Science* **63**(12), 4345–4365.

Cho, S., Lee, G., Rust, J. and Yu, M. (2018), 'Optimal dynamic hotel pricing.', *mimeo* .

Court, A. T. (1939), 'Hedonic price indexes with automotive examples', *The Dynamics of Automobile Demand* .

De los Santos, B. and Koulayev, S. (2017), 'Optimizing click-through online rankings with endogenous search refinement', *Marketing Science* **36**(4), 542–564.

De los Santos, B. and Wildenbeest, M. (2017), 'E-book pricing and vertical restraints', *Quantitative Marketing and Economics* **15**(2), 85–122.

Edelman, B. and Wright, J. (2015), 'Price coherence and excessive intermediation', *Quarterly Journal of Economics* **130**, 1283–1328.

Ennis, S., Ivaldi, M. and Lagos, V. (2020), 'Price parity clauses for hotel room booking: Empirical evidence from regulatory change', *TSE Working Paper 20-1106* .

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

Gallego, G. and Van Ryzin, G. (1994), 'Optimal dynamic pricing of inventories with stochastic demand over finite horizons.', *Management Science* **40**, 999–1020.

Ghose, A., Ipeirotis, P. G. and Li, B. (2012), 'Designing ranking systems for hotels on travel search engines by mining user-generated and crowsourced content', *Marketing Science* **31**(3), 493–520.

Ghose, A., Ipeirotis, P. G. and Li, B. (2014), 'Examining the impact of ranking on consumer behavior and search engine revenue', *Management Science* **60**(7), 1632–1654.

Hagiu, A. and Jullien, B. (2011), 'Why do intermediairies divert search?', *Rand Journal of Economics* **42**(2), 337–362.

Hunold, M., Kesler, R. and Laitenberger, U. (2020), 'Rankings of online travel agents, channel pricing, and consumer protection', *Marketing Science* **39**, 92–116.

Hunold, M., Laitenberger, U. and Schlütter, F. (2018), 'Evaluation of best price clauses in hotel booking', *International Journal of Industrial Organization* **61**, 542–571.

Johansen, B. O. and Vergé, T. (2017), 'Platform investment and price parity clauses', *Working Papers in Economics 01/17, University of Bergen* .

Johnson, J. (2017), 'The agency model and mfn clauses', *The Review of Economic Studies* **84**(3), 1151–1185.

Kim, E. H. and Singal, V. (1993), 'Mergers and market power: Evidence from the airline industry', *The American Economic Review* pp. 549–569.

Koulayev, S. (2014), 'Search for differentiated products: Identification and estimation', *Rand Journal of Economics* **45**(3), 553–575.

Larrieu, T. (2019), 'Pricing strategies in online market places and price parity agreements: evidence from the hotel indusrtry', *mimeo* .

Law, R., Ye, Q. and Zhang, Z. (2011), 'Determinants of hotel room price: An exploration of travelers' hierarchy of accomodation needs', *International Journal of Contemporary Hospitality Management* **23**(7), 972–981.

Lu, Q. S., Yang, Y. and Yuksel, U. (2015), 'The impact of a new online channel: An empirical study', *Annals of Tourism Research* **54**, 136–155.

Mantovani, A., Piga, C. A. and Reggiani, C. (2020), 'Online platform price parity clauses: Evidence from the eu booking. com case', *Available at SSRN 3381299* .

Monty, B. and Skidmore, M. (2003), 'Hedonic pricing and willingness to pay for bed and breakfast amenities in southeast wisconsin', *Journal of Travel Research* **42**, 195–199.

Rey, P. and Vergé, T. (2016), 'Secret contracting in multilateral relations', *TSE Working Paper 16-744* .

Rhodes, A. (2011), 'Can prominence matter even in an almost frictionless market?', *Economics Journal* **121**(556), F297–F308.

Rosenbaum, P. R. and Rubin, D. B. (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika* **70**(1), 41–55.

Salant, Y. (2011), 'Procedural analysis of choice rules with applications to bounded rationality', *American Economic Review* **101**(2), 724–748.

Thrane, C. (2007), 'Examining the determinants of room rates for hotels in capital cities: the oslo experience', *Journal of Revenue and Pricing Management* **5**(4), 315–23.

Ursu, R. M. (2018), 'The power of rankings: Quantifying the effects of rankings on online consumer search and purchase decision', *Marketing Science* **37**(4), 530–552.

Wang, C. and Wright, J. (2016), 'Platform investment and price parity clauses', *Available at SSRN 2848574* .

Wang, C. and Wright, J. (2020), 'Search platforms: Showrooming and price parity clauses', *the RAND Journal of Economics* **51**(1), 32–58.

Williams, K. R. (2017), 'Dynamic airline pricing and seat availability.', *Cowles Foundation Discussion Paper 2103R* .

Zellner, A. (1962), 'An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias.', *Journal of the American Statistical Association* **57**, 348–68.

Zellner, A. and Theil, H. (1962), 'Three stage least squares: Simultaneous estimate of simultaneous equations.', *Econometrica* **29**, 54–78.

Zhou, J. (2011), 'Reference dependence and market competition', *Journal of Economics & Management Strategy* **20**(4), 1073–1097.

Chapter 3

Vertical Integration of Platforms and Product Prominence

"The trouble is, it's not easy to know exactly how those algorithms work. How they've decided what to show us, and what to hide. And yet the decisions they make affect us all."

- Margrethe Vestager (2017), *Commissaire européenne à la concurrence*

1 Introduction

Online platforms are an essential part of e-commerce. Their emergence has promised substantial advantages for sellers and, in particular, consumers, especially in terms of transparency on offers and prices, low search and distribution costs, and better matches between supply and demand. Nowadays, platforms use complex algorithms for presenting product information to consumers. The way they function is often opaque and raises questions for both market participants and policy makers. In particular, the choice and transparency of ranking criteria have become the subject of intense policy debates.

Online hotel booking is particularly interesting as it mixes complex ranking algorithms of online travel agencies (OTAs) and meta-search platforms (MSPs) together with market concentration and vertical integration; Booking Holding and Expedia Group have both popular OTA websites (such as Booking.com and Expedia) as well as MSPs (Kayak and Trivago).

The relationship between hotels and OTAs has come under scrutiny with the different national policies in Europe regarding the price parity clauses (PPCs)[1] and academic research. Much less visible in the debate is another important link in the chain of this industry: meta-search platforms such as Kayak, which gather a large part of the offers from different hotel booking websites and thus enable a price comparison both across hotels and across sales channels for a given hotel.

Against the expectation that a key promise of price comparison websites is to show consumers the best offers on the market, Booking Holding's acquisition of Kayak in 2013 raised neutrality concerns about the ranking algorithm of sales channels and products on the price comparison website. Some observers argued that Kayak may have the incentive to promote Booking Holding OTAs rather than the cheapest ones. Instead, the CEO of Booking Holding claimed: "We won't bias Kayak search results."[2]

In January 2020, the Australian Competition and Consumer Commission concluded that the meta-search platform Trivago had misled consumers by indicating that its website helped them to identify the cheapest rates available for a given hotel while its ranking algorithm gave preponderant weight to the cost per click paid by the sales channel.[3]

In this article, we study the impact of the vertical integration between OTAs and MSPs on the ranking algorithm by MSPs of hotels and sales channels. For this, we use data collected between October 2014 until September 2017 from the meta-search platform, Kayak, for hotels in Paris. The data comprises information from about 1,800

[1] See, for instance, the report on the "Monitoring exercise carried out in the online hotel booking sector by EU competition authorities in 2016."

[2] See the report on the statement.

[3] See the press release by the ACCC.

Parisian hotels on room availabilities and prices on up to 22 different sales channels and different time horizons adding up to 17 million observations. We distinguish the horizontal ranking of sales channels for a given hotel and the vertical ranking of hotels for given reservation and arrival dates (see Figure 3.1).

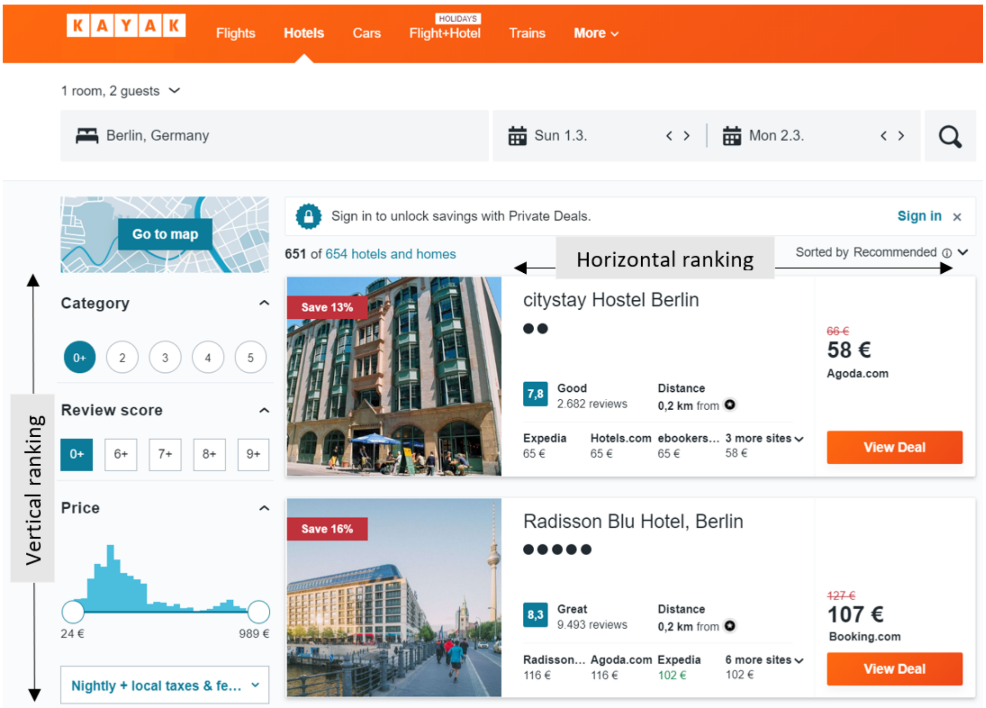


Figure 3.1: Kayak rankings

Our analyses of the horizontal ranking indicate that OTAs of Booking Holding are more often position leaders (i.e., highlighted sales channels on hotel offers) than price leaders (i.e., among the cheapest sales channels). Using linear fixed effects regressions on the hotel- and request-level that also account for prices and popularity measures, we additionally show that OTAs of the Booking Holding have a higher probability than any other OTA to be among the visible providers and to be the highlighted sales channel. For the vertical ranking, our results suggest that hotels are ranked worse in the Kayak search results when the Expedia Group is the cheapest sales channel. We provide various robustness checks. First, we distinguish sales channels within groups and show that the two major OTAs, Booking.com and Expedia, drive the main results. Second, we use changes in the regulation of PPCs to estimate its potential impact on the Kayak ranking algorithm. Finally, we distinguish hotels affiliated to chains from independent ones and show that for the horizontal ranking, the direct channel of independent hotels is put more prominent compared to Booking Holding, while for the vertical ranking independent hotels present on Kayak with a direct website are favored as well.

The rest of the paper is structured as follows: Section 2 covers the related literature and Section 3 provides background information on hotel meta-search platforms. In Section 4, the empirical strategy is developed and Section 5 introduces the data, descriptive statistics, and preliminary evidence. Section 6 is devoted to the econometric results and Section 7 complements these with robustness checks. Section 8 provides a conclusion.

2 Related Literature

One relevant strand of theoretical literature studies the decisions of intermediaries to bias product presentations by making certain products more prominent than others (Raskovich, 2007; Inderst and Ottaviani, 2012; Hagiu and Jullien, 2011, 2014; De Corniere and Taylor, 2014; De Cornière and Taylor, 2019; Hunold and Muthers, 2017; Shen and Wright, 2019). 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. De Corniere and Taylor (2014) show that integrated search engines distort search results, but the overall welfare effect is unclear. For instance, the integrated search engine can have a strong incentive to generate demand. Similarly, De Cornière and Taylor (2019) study biased recommendations of intermediaries and show that if the payoff functions of sellers and consumers are conflicting, bias can harm consumers. Hagiu et al. (2020) analyze an intermediary's dual role of being both a reseller and marketplace. They contrast a ban of this practice to other policies, which restrict the imitation of products by third parties or steering towards the intermediary's product. Using the example of a streaming platform, Bourreau and Gaudin (2018) and Drugov and Jeon (2019) study incentives to bias recommendations to consumers towards vertically integrated content.

In our paper, the acquisitions by the Booking Holdings and Expedia Group of Kayak and Trivago, respectively, are viewed as a vertical integration of two platforms. Such acquisitions can have a negative impact on downstream competitors in the form of foreclosure or sabotage. The literature on sabotage has explored the incentives of vertically integrated suppliers to sabotage downstream activities of rivals, depending on the degree of downstream competition. There are two types of sabotages: demand-reducing and cost-increasing sabotages (Mandy and Sappington (2007)). In case of a dominant MSP like Kayak operating upstream[4], the theory on sabotage suggests that its vertical integration with a downstream sales channels (Booking.com) could nega-

[4] See Rey and Vergé (2016) for this type of vertical relationship in which upstream sales channels sell services to downstream providers setting consumer final prices.

tively affect Expedia.fr' sales either by lowering its volume of intermediated sales for hotels or by increasing its cost per click to be referenced on the MSP. The relative profitability of these two alternatives depends on the type of competition between sales channels. Furthermore, in an extreme case of sabotage, MSPs biasing search results may lead to competing sales channels being less often promoted than integrated OTAs on MSPs. This can also be interpreted as "partial" foreclosure, leading to negative effects on downstream competition and consumers' welfare (Choi and Yi (2000); Rey and Tirole (2007)).

Related empirical literature highlights the importance of rankings of intermediaries in the context of online hotel booking. Chen and Yao (2016); De los Santos and Koulayev (2017); Koulayev (2014); Ghose et al. (2012, 2014) study how rankings affect consumer choices 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). Ursu (2018) exploits a random variation in the ranking of hotels at the OTA Expedia and finds lower, yet still significant, position effects. She finds that consumers click more often on an offer that is ranked better to obtain detailed information on it. However, conditional on seeing the detailed information, the ranking position does not influence the booking behavior of consumers.

Firms have incentives to engage in search discrimination and steering by showing different offers or rankings to customers depending on some observable characteristics, such as location or income level. Mikians et al. (2012, 2013); Hannak et al. (2014) collected data on various e-commerce websites and provide empirical evidence on search discrimination. In addition, in the case of the hotel industry, the Wall Street Journal[5] reported in 2012 that the travel agency Orbitz showed more expensive hotel offers to Mac users than to PC users. Furthermore, a CMA inquiry carried out in 2018[6] shows that hotel rankings on Expedia were different for direct access compared to consumers redirected from cashback website. Search discrimination makes difficult the analysis of algorithm biases which would require total absence of consumer customization to be identified.

Our work is also related to the recent theoretical literature on the competitive effects of price parity clauses of intermediaries, such as OTAs (Edelman and Wright, 2015; Boik and Corts, 2016; Johnson, 2017; Wang and Wright, 2020; Johansen and Vergé, 2017; Ronayne et al., 2018; Wals and Schinkel, 2018; Mantovani et al., 2018; Hunold et al., 2020). Hunold et al. (2020) demonstrate that OTAs may condition the rankings of hotels in their search results on prices these hotels set elsewhere and by this achieve the same effects as a price parity clause. Their empirical evidence is consistent with OTAs conditioning their rankings on prices on other channels.

5 *"On Orbitz, Mac Users Steered to Pricier Hotels"* (2012), online access here.

6 See the CMA report on pricing algorithms here.

This paper relates more broadly to, for example, Zhang et al. (Forthcoming), who show that a recommender system that maximizes firm's profits lowers consumer surplus and welfare using a large-scale field experiment in video-on-demand. More generally, our work also contributes to empirical studies on algorithmic bias. For example, Lambrecht and Tucker (2019) run career ads on Facebook intended to be gender neutral, which are, however, delivered to men more often. They provide suggestive evidence that women are more expensive to show ads to and the algorithm thus optimizing costs.

3 Industry Background

Online hotel booking and meta-search. Hotels can be booked through numerous distribution channels, both offline and online. In the last decade, the latter distinctively gained in popularity (Cazaubiel et al., 2020). Most website-based hotel bookings take place on OTAs according to the the European hotel association HOTREC.[7] OTAs pool offers of different hotels and display them through a ranking in response to a user's search request, typically composed of the destination, period, and amount of people travelling. Similar to OTAs, hotel MSPs gather offers of different hotels. In addition, various online sales channels are displayed for each hotel, which usually comprise several OTAs and the hotel's own direct channel (see Figure 3.2). Thus, they provide a comparison service on a more aggregate (meta-)level without actually selling hotel rooms or posting prices themselves (Hunold et al., 2018).

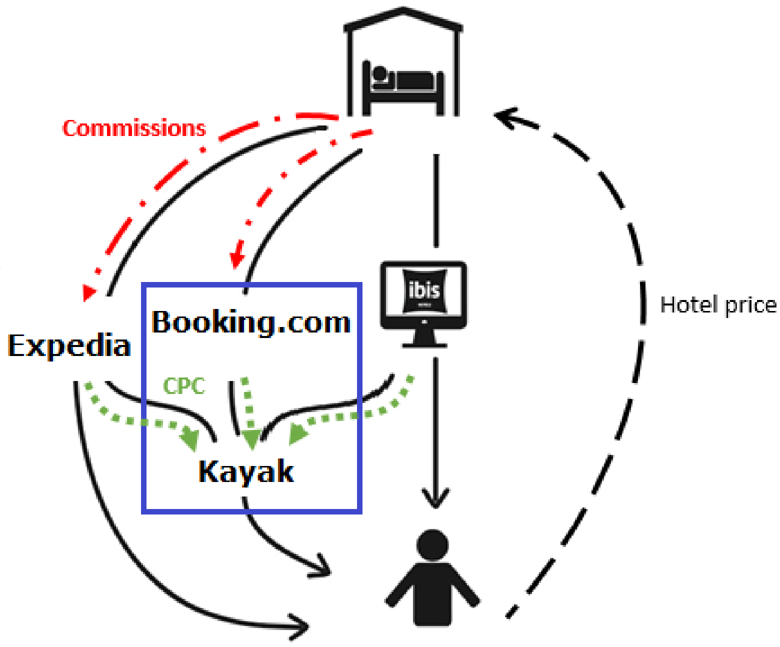
Revenue generation in meta-search. Revenues of hotel MSPs are generated by sending referrals to (actual) sales channels and advertising placements on the website.[8] The revenues are realized either once a user clicks on a referral and an advertisement or upon completion of the travel. Advertisers, be it OTAs or hotels, typically make bids for these placements (see Figure 3.2 in green). These bids can be made dependent on various characteristics, such as the user's location, device, and dates.[9] According to the sector inquiry (SI) of the German competition authority (Bundeskartellamt) on comparison websites, the 14 surveyed hotel meta-searchers report that the vast majority of their revenue comes from OTAs and the most frequent remuneration modes are cost per order and cost per click (CPC), the latter being the most important ranging from a fraction of a cent to several euros per click.[10] However, industry reports suggest a

[7] See the European Hotel Distribution Study 2018.

[8] See for Kayak, as an example, the annual report of the Booking Holdings.

[9] This bidding overview for Hotel ads on Google provides more information and we expect other hotel meta-search websites to work similarly.

[10] See Bundeskartellamt's sector inquiry.



CPC: Cost per click

Figure 3.2: Money flows in the vertical hotel industry

recent trend of OTAs pulling back from MSPs, while hotels turn to them.[11]

Hotel meta-search players and sizes. Almost every hotel meta-search visit by a user leads to a redirect to a sales channel.[12] Thus, a presumably large share of bookings originates at hotel meta-search websites, resulting in revenues for hotel meta-searchers of roughly 200 million euros in Germany compared to more than 800 million euros for OTAs (in 2017). This importance of hotel meta-search websites is also demonstrated in 316 million visits from November 2016 to October 2017 compared to 1.2 billion for OTAs in Germany. This resonates well with global website rankings, where multiple hotel meta-search websites are among the Top 50 in the travel category and are shown to be important referrers of the major online travel agents.[13] Hotel meta-search is considerably concentrated with a few large players. For Germany, in terms of visits and revenues, Trivago has a share of more than 50 percent, followed by Google, TripAdvisor, and Kayak.[14]

Display of offers on meta-search platforms. MSPs typically use two rankings to organize their websites. First, regarding the vertical ranking, hotel offers appear in a

11 See an article on skift.com.
12 This fact and the following numbers are taken from Bundeskartellamt's sector inquiry.
13 See the Alexa website ranking.
14 See Bundeskartellamt's sector inquiry.

certain order from top to bottom of the website (as do OTAs). Second, different from OTAs, MSPs display different sales channels for each hotel. We refer to the order of them as horizontal ranking. For Kayak (and similarly others), this horizontal ranking can be further divided into a prominent sales channel, three further visible offers, as well as the remaining offers that are somewhat hidden and denoted by "x more sites" (see Figure 3.1).

According to the Bundeskartellamt's SI, Kayak states that it ranks hotels worse if their average earning potential is low. The SI also reveals that meta-search websites make the horizontal ranking of sales channels for one hotel dependent on bids, especially when prices of the respective hotel are the same. Furthermore, the cheapest sales channel is not necessarily shown most prominently, but rather the expected revenue is accounted for.[15] In contrast, the Kayak website states that it voluntarily "*highlight[s] one main provider based on criteria such as customer popularity or ratings*".

The hotel meta-searchers in the inquiry further claim that in only 80 percent of the cases one of the cheapest sales channels is the most prominent. This has implications as a significant share of users is reportedly clicking on the prominent spot even though cheaper options exist (thereby steering customers).

As a result, there may be tension between a revenue-focused and customer-oriented presentation of hotel and sales channels prices. In this respect, it should be noted that following the Kayak takeover, the CEO of the parent company of Booking.com argued that Kayak "will not bias" search results.[16]

Regarding the display of offers, Trivago has been found breaching the Australian Consumer Law for misleading customers, as rankings were made dependent on the highest cost per click fees paid and did not present the cheapest rates for consumers prominently.[17] More broadly, our research complements past and ongoing cases about self-preferencing (e.g., Google Shopping, Amazon Buy Box).

Vertical integration of online hotel booking and meta-search. The leading OTAs Booking.com and Expedia each acquired a major hotel meta-search platform (Kayak and Trivago, respectively). Both acquisitions took place in 2013 with values of 1.8 billion and 632 million US dollars. They led to a vertical integration of OTAs and MSPs (see Figure 3.2 in blue for the Booking Holding case) in addition to an already present interdependence among OTAs due to other (horizontal) acquisitions by the two major OTAs (also described in Section 5.2).[18]

The Bundeskartellamt's SI expresses concerns regarding vertical integration of OTAs and MSPs as this could result in self-preferencing on the hotel meta-search website

15 See p. 94 in Bundeskartellamt's sector inquiry.
16 See the report on the statement.
17 See the press release by the ACCC.
18 See the press releases by Booking.com and Expedia.

with respect to its own OTAs and thereby steer users.[19] However, the German competition authority concludes that the survey among OTAs and MSPs did not reveal any indications for this bias.[20]

In our empirical analysis, we will investigate possible self-preferencing in the (actual) search results of the MSP Kayak.

4 Empirical strategy

4.1 Hypotheses and model

Booking Holding's MSP Kayak derives direct revenues from two main sources: advertising and referrals (see Section 3). Revenues from referrals are collected by a CPC scheme in which hotels pay anytime a consumer clicks on offers by the sales channels in the search results. In addition, the Booking Holding derives revenues, whenever a hotel is booked over its OTAs, such as Booking.com. We want to study whether Kayak takes this vertical benefit into account when it presents offers through its search results.

Our empirical approach is two-fold. First, we investigate how Kayak decides for each hotel offer which sales channels to make prominent (horizontal ranking). Second, we investigate whether Kayak takes the pricing and other factors related to the sales channels into account when deciding about which hotels to list first (vertical ranking). Let us describe both approaches in more detail.

Horizontal ranking. The sales channels visible at Kayak differ between hotels. From the perspective of a typical consumer, the price and the sales channels' popularity (and quality) should be the main determinants for this ranking. However, it is known that providers can also become more prominent, if they pay more for the clicks they get (see Section 3). We are interested in testing the following hypothesis.

Conjecture 1. *Other things equal, the sales channels affiliated to Booking Holding have a higher probability to be visible and are more likely to be a position leader.*

For this purpose, we estimate the following linear probability model for an offer of sales channel s in hotel h for a request r:

$$Y_{hrs} = X_{hrs}\beta + \alpha_h + \gamma_r + \varepsilon_{hrs}. \tag{3.1}$$

[19] See p. 38 in Bundeskartellamt's sector inquiry.
[20] See p. 50 in Bundeskartellamt's sector inquiry.

Y_{hrs} is an indicator variable, taking the value one if the sales channel s for hotel h in request r is a position leader and zero otherwise; X_{hrs} are explanatory variables, α_h denotes the hotel fixed effect and γ_r the request fixed effect. X_{hrs} includes the log-price of the sales channel, the number of price leader(s) for the offer, the group affiliation of the sales channel, as well as sales channel and hotel popularity that varies across time. The hotel fixed effects take time-invariant hotel characteristics into account, such as the number of stars, amenities, chain affiliation, or location. The group affiliation of the sales channel takes Booking Holding as the reference category. Therefore, if conjecture 1 is satisfied, we should observe a negative and significant coefficient associated to other groups, suggesting that sales channels not affiliated to Booking Holding have a lower probability to be prominent.

Vertical ranking. For a given search request, Kayak provides a list of the available hotel offers. From the consumer's perspective, the rank of a hotel in that list should be better if the hotel's gross match value for the average consumer is higher. This value should increase in the number of stars, the user rating, free breakfast, and so on. Given the gross value, the sales channels' prices should negatively affect the ranking as, other things equal, a higher price should mean a lower net match value for the average consumer. If Kayak is maximizing its short-term revenues from cost per click fees, it should incorporate the likelihood of a click (which might depend negatively on prices and positively on quality) as well as the cost per click fees.

We want to test the following hypothesis.

Conjecture 2. *Other things equal, hotels which have higher prices on Booking Holding channels than on other sales channels are more likely to be ranked worse in the Kayak search results.*

We proceed similar to Hunold et al. (2020) and estimate the following linear model for a hotel h in a request r:

$$R_{hr} = Z_{hr}\kappa + \alpha_h + \gamma_r + \epsilon_{hr} \tag{3.2}$$

R_{hr} is the ranking position of a hotel h in the Kayak search results. Z_{hr} includes the minimum log-prices of the sales channels available for this hotel, the group affiliation of the price leader, as well as the average total popularity available for this hotel and hotel popularity that both vary across time. Compared to the horizontal ranking, we additionally control for sales channel availability for the hotel including group availability dummies and the number of sales channels. Hotel (α_h) and requests (γ_r) fixed effects are the same as in the horizontal ranking analysis. The group affiliation takes the form of several dummies, one for each group, taking the value one if the group is one of the price leaders. Therefore, if conjecture 2 is satisfied, we should observe a positive and significant coefficient associated to the dummy variable of the Expedia Group and/or other OTAs. This would suggest that if a group other than Booking Holding is among the price leaders, then the hotel is ranked worse in the Kayak search results. Equation 3.2 implicitly states that the hotel ranking depends on price leadership irrespective of which channels are visible. However, one may argue that competing price leaders threaten the OTAs of the Booking Holding, only if they appear visible on the offers, as consumers generally do not click on the sub-menu to see the remaining offers. We assume the horizontal and vertical ranking to be jointly determined. Therefore, we do not include visibility measures in the vertical ranking equation to avoid endogeneity issues.[21]

4.2 Identification

Unobserved demand shocks. A concern could be that a demand shock may impact both the price of sales channels and their ranking on the MSP. For instance, an increase in demand could lead to a stock out of cheaper hotels, such that only hotels with a higher price remain in the list of search results and subsequently get a better ranking. We deal with this concern by adding request fixed effects γ_r which capture the effects linked to the combination of the booking and arrival date (and by this also the booking horizon).

Unobserved heterogeneity in hotel popularity. If the MSP expects higher revenues from a hotel, it has an incentive to rank the hotel better – other things equal. A high commission income can be either due to a higher CPC paid by the hotel (see later) or a higher likelihood of a hotel being clicked (measured by the click-through-rate, CTR). For instance, if hotels with a lower CTR typically have lower prices on other sales channels affiliated with the MSP, we could get a spurious negative correlation between a good ranking position and the price markup relative to other channels. We deal with this unobserved heterogeneity across hotels by removing time-constant unobserved heterogeneity between hotels through the inclusion of hotel fixed effects. Current deviations in hotel popularity are captured by controlling for short-term consumer ratings from TripAdvisor for this hotel.

Unobserved heterogeneity in sales channel popularity. Similar to the argument of unobserved hotel popularity, the MSP should take into account that consumers not

[21] We could instead jointly estimate it through a system of equations 3.1 and 3.2 with a seemingly unrelated regression equation model (Zellner, 1962).

only value a low price, but might also have a preference for specific sales channels. The popularity of sales channels could follow seasonal patterns or unobserved trends initiated by marketing activities of the channels. If sales channels affiliated with the MSP are increasingly popular and their popularity is then associated with higher prices, a good ranking position of the MSP could then be spuriously correlated with the common ownership of the sales channels. To mitigate this concern, we use data about the current relative search volume of each sales channel on Google in France as a measure for the different popularity and associated CTR.

Unobserved channel and hotel-specific CPC. In its decision about the horizontal and vertical ranking, the MSP also takes CPC payments by the sales channels into account. These might vary between hotels as well as across time and might affect both the pricing across channels and the ranking position of the hotel at the meta-search website. In particular, as a higher CPC implies a higher distribution cost for the hotel when selling through the direct channel, the hotel might also increase the direct channel price. Better visibility of the direct channel would then be driven by the higher CPC and bias our "direct channel" coefficient upwards in equation 3.1. Similarly, it could be that OTAs negotiate different CPCs for specific types of hotels. To deal with this potential problem, we control for changes in the price of the hotel as they should account for changes in the distribution costs possibly reflected by changes in the CPC. Furthermore, we assume that CPCs for independent hotels, at least, are relatively constant over time at Kayak, as independent hotels have to make use of an intermediary to list their rooms on the MSP. As the CPC conditions for chain hotels could be more flexible, we run the analyses separately for independent and chain hotels in subsection 7.3. This also provides insight into potential different CPCs paid by OTAs depending on the hotel type.

5 Data

In this section, we present the data set and its main characteristics along with a classification of sales channels and a conceptualization for the display of prices and offers.

5.1 Data collection

For our analysis we rely mainly on data on hotel and channel rankings on Kayak, and prices that hotels post on different channels. As control variables, we need data on the characteristics of hotels and channels, that can explain their attractiveness for consumers as well as data on their determinants of the profitability of an hotel offer for Kayak.

The data was collected as in Larrieu (2019) from October 2014 until September 2017 on Kayak.com for 863 hotels in Paris between April 10, 2014 to July 1, 2017. [22] In particular, 2,375 search requests were made for 410 distinct reservation dates for one night for two people with different time horizons (mainly 4, 14, 30 and 180 days before arrival).[23] We then amended the data set by two additional sources. First, we collected TripAdvisor information on hotels' characteristics and reviews over time. Second, we retrieved time series data from Google Trends for our observation period to approximate the sales channel popularity (see Appendix 3.A for a description of how this data was collected).

5.2 Descriptive statistics

In the following, we describe the hotels in the dataset as well as characteristics of the dataset used for the analyses.

Variable	# Obs.	Min	p50	Mean	Max	SD
Stars	1,784	0	3	3.2	5	0.9
Chain	1,784	0	0	0.3	1	-
# Rooms	1,784	1	45	76	1,093	98
# Reviews	1,716	1	142	173	4,659	225
Score over 5	1,716	2.6	4.3	4.3	4.7	0.2

Table 3.1: Hotel characteristics

Hotels in the data set. Our data set contains 1,784 distinct hotels which are in the greater Paris area. In Table 3.1 we provide some average characteristics. Hotels in our data set have on average 76 rooms and 3 stars. Overall, 30% of the hotels are affiliated to a chain, the most prominent ones being Ibis Hotels, Best Western, and Mercure (Appendix 3.B - Table 3.16). As revealed in Table 3.17 (Appendix 3.B), hotels from one to three stars have on average 62 rooms, while four- and five-star hotels have on average 104 rooms.[24] For the time-varying hotel characteristics, we report the mean of the number of reviews and the average consumer rating on Tripadvisor at the moment of the reservation as a measure for hotel popularity. At the reservation date, hotels in our data set received, on average, 173 reviews with an average rating of 4.3 of 5 stars.

[22] Search results were collected every day from 6am to 8am using a web scraping program from a Windows desktop. IP addresses were randomized in each iteration using a list of French IPs located in the region of Paris. For each iteration, the cache of our browser was cleared from all cookies and historical searches to appear as a new user without any personalization that may affect the Kayak ranking algorithm.

[23] The data set is not balanced since not all existing reservation dates were queried with all possible time horizons. However, 80% of reservation dates were queried with at least five distinct time horizons. Other time horizons were collected in order to account for intertemporal price discrimination following revenue management and are kept in the analysis.

[24] This is consistent with the INSEE French statistics on the hotel industry, see INSEE website.

Ranked hotel and channel offers. In total, our data set contains more than 17 million observations. The average price for one night is 183 €[25] and negatively correlated with the time horizon.[26] For each request, we observe between 27 and more than 1,222 hotels with an average of 933 hotels, which are linked to between 1 and 22 different online sales channels. On average, for each hotel, 10 sales channels display an offer, while there are over 828 distinct providers.

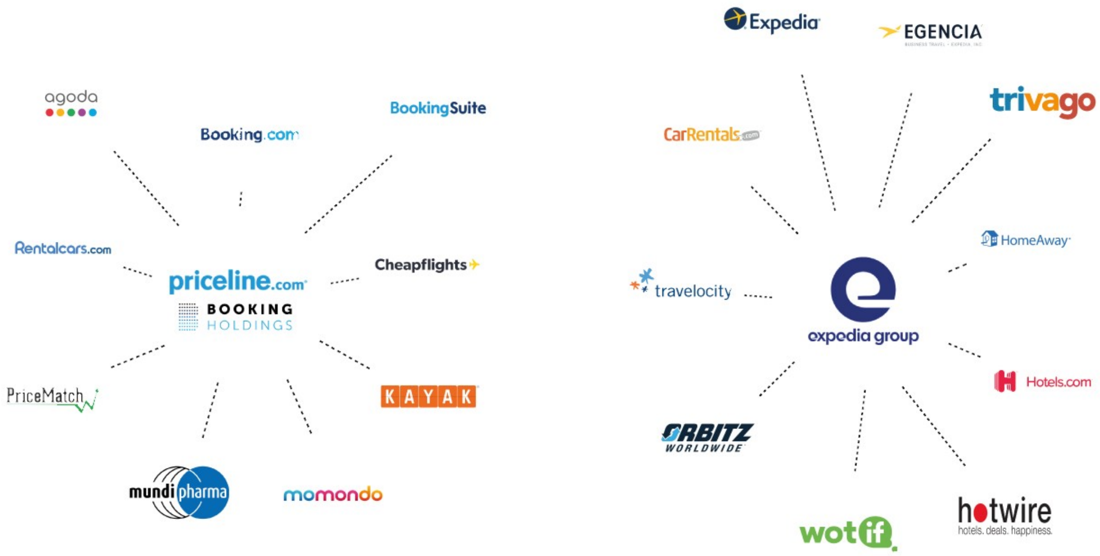


Figure 3.3: Booking Holding and Expedia Group

Sales channels in the data set and ownership. We observe 828 different sales channels. We distinguish between hotels direct channels, and online travel agencies. For the OTAs, most are linked to two different groups (see Figure 3.3). The Expedia Group owns different OTAs (Expedia.com, Classic Vacations, Hotels.com, Hotwire.com, Venere.com, and Egencia), is affiliated to some travel companies (voyage-sncf.com, Abritel HomeAway, Orbitz (including ebookers, HotelClub and CheapTickets), and Travelocity), and has Liberty Media as its parent company, which is the main shareholder of TripAdvisor. On the other hand, Booking Holdings Inc. owns and operates several travel meta-search platforms, OTAs and other travel websites, including Booking.com, Priceline.com, Agoda.com, Kayak.com, Cheapflights, Rentalcars.com, Momondo, and OpenTable. The remaining sales channels are either competing OTAs such as HRS.com and smaller ones (Presitiga, Melia, Hotelopia.com., HotelsClick, Amoma.com (bankrupt

[25] We observe in the dataset some extreme prices of up to 965,832€. To remove outliers, we restrict to prices lower than 10,000€ which is large enough for the price of an hotel room for one night in Paris even in a Palace category (and thus drop 28 observations).

[26] The average price is strictly decreasing as the arrival date approaches, from 189€ at 6 months before the arrival to 182€, 180€ and 178€ respectively for one month, 14, and 4 days before the arrival date.

since 2019), Weekendesk, Lastminute group etc.) or linked to French national companies in the travel sector (Tablet as part of the Michelin guide, Splendia owned by the online platform Voyage-Privé.com). In addition, there is a large national player, AccorHotels.com, hosting the majority of the big brands (Ibis, Mercure, Novotel, etc.) in France. It offers hotels to appear on its own platform in exchange for a commission. Therefore, AccorHotels.com has a strategy of both offering its own brand hotels but also independent ones on its platform. We do not consider some particular offers (8%) for which Kayak is mentioned as a sales channel because we do not observe the identity of the sales channel really mediating the transaction. We finally classify sales channels depending on their group affiliation (Booking Holding and Expedia Group) distinguishing independent sales channels between online travel agencies (Other OTAs than groups) and the hotel direct website (Table 3.2).

5.3 Descriptive evidence

In this subsection, we focus on the horizontal ranking of sales channels, defining the price and position leaders in order to compare their respective occurrence by group and sales channels. Before this, we describe the concentration of offers on sales channels.

Concentration of offers on sales channels. The market is concentrated around seven large OTAs covering 60% of the offers, while 800 small sales channels only account for 5% of price offers. The large providers are well-known OTAs such as Booking.com, Expedia, Hotels.com together with national players (voyages-sncf) and smaller platforms in the travel market (Agoda, Venere). The sales channels attributed to the 3% are mainly composed of hotel websites, which are generally unique by hotel.

Group	# obs	% in obs	Sales channel	# obs	% in obs
Expedia Group	6,747,875	40%	Hotels.com	1,713,173	25%
			Expedia.fr	1,397,825	21%
			Venere.com	1,315,041	19%
			Voyages-sncf.com	995,594	15%
			Ebookers.com	708,853	11%
			Others	617,389	9%
Booking Holding	3,387,182	20%	Booking.com	1,899,278	56%
			Agoda.com	1,487,904	44%
Other OTAs	4,856,541	29%	Amoma.com	737,503	15%
			Hotelopia.com	691,228	14%
			Logitravel.fr	561,550	12%
			HotelTravel.com	479,809	10%
			Rumbo.fr	448,225	9%
			Hrs.com	316,822	7%
			Others	1,621,404	33%
Direct website	593,419	3%	...		
Kayak	1,417,157	8%	...		
Total	17,002,174	100%			

Table 3.2: Sales channels' availability and classification

Behind the apparent diversity of sales channels in the market, the two big groups account for 60% of price offers referenced by Kayak. Other OTAs play a non-negligible role since they account for 29% of the offers translating overall to the same weight as Booking Holding alone. We also note that the direct channel of the hotel is rarely available.

Position leader. For each request, Kayak as a meta-search website lists all available offers of sales channels and displays the associated price. For a given hotel, these offers are then displayed on the Kayak website by a horizontal ranking. One sales channel is highlighted and three others are visible but less prominent. Remaining sales channels are hidden in a sub-menu on which consumers have to click if they want to see additional offers (see Figure 3.4 in yellow). We define the highlighted sales channel as the position leader. It is also called the sales channel in the *buy box*, especially in the retail industry.

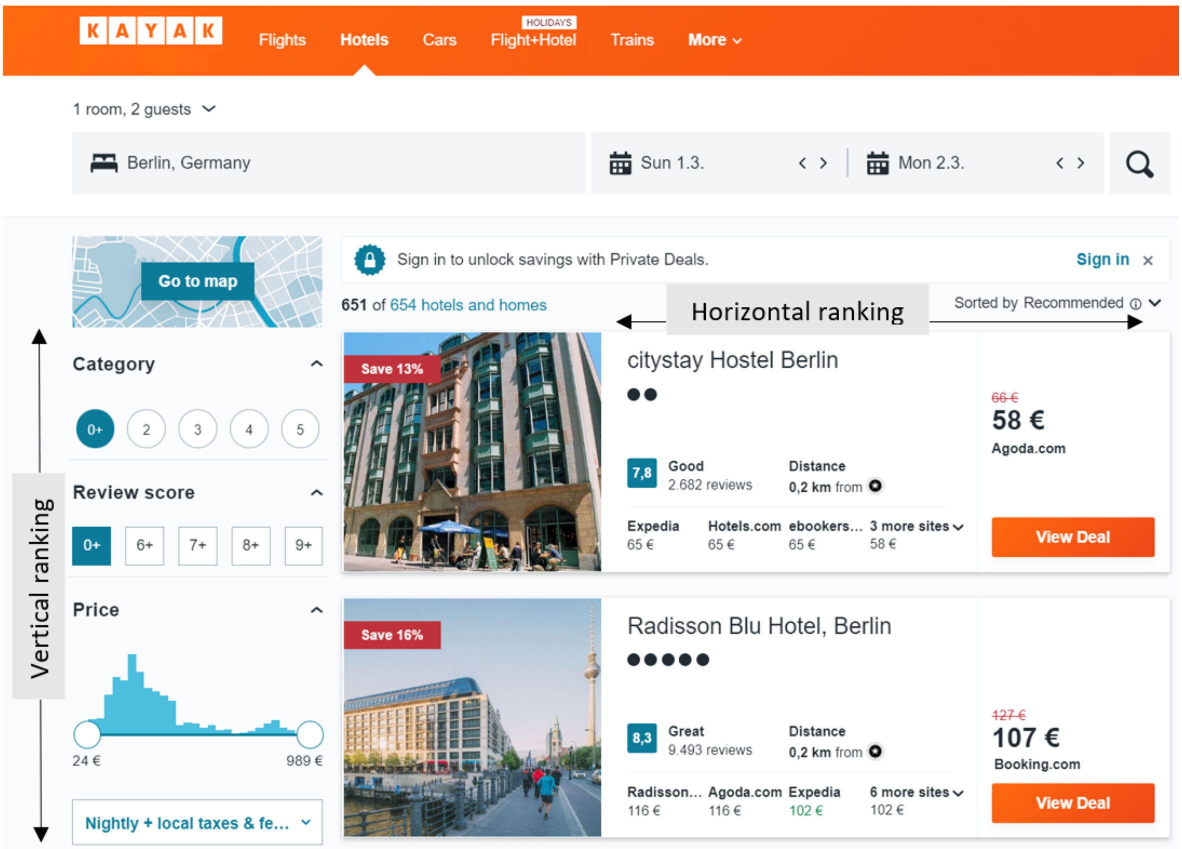


Figure 3.4: Horizontal ranking and visibility

This horizontal ranking is the result of an Kayak internal algorithm. In particular, the highlighted sales channel is not necessarily the one offering the cheapest price to book the room.

Price leader(s)

For a given hotel, we define the price leader(s) as the sales channel(s) offering the lowest price (see Figure 3.4 in green). In 62% of the hotel offers, the price leader is unique, meaning that it has a strictly lower price than others. In the remaining cases (38%) several sales channels (up to 15) offer the same cheapest price. Different price leaders can be affiliated to the same group. At this level, there is a unique group price leader in 72% of the cases (Table 3.3).

# Sales channel(s)	Freq.	Percent
1	1,303,460	62%
2	166,252	8%
3	182,218	9%
≥ 4	441,475	21%
Total	2,093,405	100%

# Group(s)	Freq.	Percent
1	1,505,549	72%
2	403,855	19%
3	151,146	7%
4	32,855	2%
Total	2,093,405	100%

Table 3.3: Number of price leader(s) at the sales channel and group levels

Even if the hotel's direct website is not often among the sales channels, it is actually one of the cheapest providers in 53% of cases, which is much more than the main online travel agencies Booking.com (37%) and Hotels.com (37%). This remains also at the group-level (Table 3.4) and the number of price leaders plays an important role for the probability to be the cheapest sales channel. Restricting to cases in which there is a unique sales channel price leader, the direct website of hotels is most often (54%) the cheapest channel compared to any other OTA. In comparison, when there are two sales channels being price leaders, Booking Holding has a much higher probability (44%) to be among them compared to other groups.

# Sales channel(s) price leader(s)	All	1	2	3	≥ 4
Direct website	53%	54%	18%	38%	68%
Booking Holding	37%	10%	44%	61%	87%
Expedia Group	32%	7%	13%	67%	89%
Other OTAs	23%	14%	16%	18%	26%

Table 3.4: Price leadership given availability by number of price leader(s)

Starting from these observations on prices, as Kayak is a price comparison website searching for the best deal, we expect price leading sales channels to be more visible than others when there are several and to be the prominent sales channel in case of a unique price leadership. In order to investigate this, we compare the price leader occurrences previously computed to the probability to be among the visible sales channels or the position leader.

Price vs Position leader

In comparison to the price leader, there is always a unique position leader. As it is the case for the price level, the position leadership only makes sense when controlling for the availability of the sales channel. Overall, in 96% of the hotel offers the position leader is one of the price leaders and this proportion increases with the number of price leaders. When the position leader is not the unique price leader, it is in favor of one of the two biggest groups in 81% of cases, with Booking Holding comprising 43%. We additionally show (in Table 3.5) that the OTAs of Booking Holding are more often (12%) the position leader than the price leader – the difference being six times greater for the Booking Holding compared to the Expedia Group. In contrast, compared to the Booking Holding, the direct website of the hotel appears less often in the first position whereas it is five times more often cheaper.

Group	Position leader	Unique Price leader	Difference
Booking Holding	22%	10%	12%
Expedia Group	9%	7%	2%
Other OTAs	11%	14%	-3%
Direct website	38%	54%	-16%

Table 3.5: Price vs Position leadership given availability

Position leader	Percent
Booking Holding	38%
Expedia Group	26%
Other OTAs	25%
Direct website	10%
Total	100%

Price leader & Not position leader	Percent
Booking Holding	13%
Expedia Group	19%
Other OTAs	17%
Direct website	4%

Table 3.6: Position leader by group

As the entire group does not necessarily reflect the case of each single provider, in Appendix 3.C - Table 3.18, we compare the share of price and position leaders at the sales channel-level. Results are amplified with Booking.com being five times more the position leader than the price leader. For other OTAs, results are less conclusive. Hotels.com and Expedia.fr are more often the position leader than the price leader, but still to a lesser degree than Booking.com while other OTAs are always less often the position leader than the price leader.

Overall, this suggests that the two biggest players are generally more often the position leaders than the price leaders, which is especially the case for Booking.com, Expedia.fr, and Hotels.com. There may be two explanations for such exposure. First, the Kayak website states that it voluntarily "*highlight[s] one main provider based on criteria such as customer popularity or ratings*". Thus, an explanation of the discrepancy could be that these particular platforms are quite popular in France. As Kayak values the popularity in the horizontal ranking algorithm, it makes them visible more often compared to a ranking based on the cheapest price. Table 3.15 in Appendix 3.A shows that Booking.com and Expedia.fr are indeed relatively more popular than others. However, this does not hold for Hotels.com. A second explanation is that sales channels (OTAs as well as hotels) may pay more to be put more prominent.

6 Estimation results

6.1 Horizontal ranking

In Table 3.7, we report the estimation results for the model described in equation 3.1. The dependent variable in columns (1) to (3) is the probability to be visible (among the four first sales channels), whereas in columns (4) to (6) it is the probability of being a *position leader* (i.e., to be in the *buy box*). In each case, we compare three models, sequentially adding popularity measures for hotels and channels.

	Linear Probability Model					
	Visible (# ≤ 4)			Position Leader (# = 1)		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(price)	-0.183***	-0.185***	-0.185***	-0.021***	-0.021***	-0.021***
	(-29.32)	(-29.03)	(-29.01)	(-12.96)	(-12.86)	(-12.45)
Price leadership						
(ref: Not price leader)						
Among price leaders	0.234***	0.234***	0.234***	0.203***	0.202***	0.201***
	(67.14)	(66.81)	(66.72)	(90.50)	(90.13)	(89.61)
Unique price leader	0.475***	0.474***	0.475***	0.749***	0.750***	0.752***
	(69.73)	(69.49)	(69.31)	(137.15)	(136.43)	(135.94)
Group						
(ref: Booking Holding)						
Direct website	-0.011	-0.010	-0.002	-0.007	-0.007	0.037***
	(-1.37)	(-1.28)	(-0.28)	(-1.32)	(-1.21)	(7.07)
Expedia Group	-0.111***	-0.112***	-0.106***	-0.106***	-0.105***	-0.071***
	(-22.09)	(-22.04)	(-19.80)	(-27.87)	(-27.36)	(-23.63)
Other OTAs	-0.096***	-0.096***	-0.087***	-0.069***	-0.068***	-0.024***
	(-15.94)	(-15.72)	(-13.53)	(-23.68)	(-23.27)	(-10.32)
Constant	1.370***	1.336***	1.326***	0.179***	0.203***	0.152***
	(44.90)	(28.37)	(28.14)	(21.32)	(9.06)	(6.78)
Request FE	yes	yes	yes	yes	yes	yes
Hotel FE	yes	yes	yes	yes	yes	yes
Hotel Popularity	no	yes	yes	no	yes	yes
Channel Popularity	no	no	yes	no	no	yes
N	15,585,011	15,089,104	15,089,104	15,585,011	15,089,104	15,089,104

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7: Visibility and Position Leadership at the group-level

As expected, a higher price reduces the probability for a sales channel to be visible on the hotel offer. Given a hotel, a sales channel has a higher probability to be visible if it offers the cheapest price together with other sales channels (multiple price leaders) and this effect is twice as large if the sales channel is the only one to offer the cheapest price. A key variable of interest, the group affiliation, takes Booking Holding as a reference category. When the sales channel belongs to any other OTA than one related to the Booking Holding, it has a lower probability to be visible. This negative effect is of the same magnitude when the sales channel is an independent OTA or one belonging to the Expedia Group. The effect is not significant if the sales channel is a hotel's direct website. Other things equal, OTAs that do not belong to Kayak's related companies, have a lower probability to be visible on the hotel offers on Kayak.

Results on visibility and position leadership are very similar. Interestingly, the price coefficient is ten times smaller when the position leadership is a dependent variable, while the coefficient associated to the unique price leadership is more important. This suggests that the price level influences the visibility of sales channels. However, in order to be the first ranked sales channel, it matters to be the cheapest one. Looking at the group coefficients, the results show that, other things equal, OTAs other than ones from the Booking Holding have a lower probability to be in the *buy box*. We additionally show that in contrast to the visibility results, once we control for channel and hotel popularity, the direct website is more likely to be position leader than OTAs from Booking Holding, suggesting that Kayak may make a distinction between the type of sales channels.

This difference of the results between other OTAs and the direct website can be explained by a policy applied by meta-search websites in favor of direct channels. For equal prices the priority can be given to the direct channel in order to encourage small independent hotels to offer through Kayak directly without relying on online travel agencies. It is true that if we restrict to cases in which the direct website is the price leader together with Expedia Group or other OTAs (except the ones belonging to the Booking Holding), the direct channel is the position leader in 53% of the cases compared to 32% and 13% of the cases for the Expedia Group or other OTAs.

Finding 1 (horizontal ranking): Other things equal, sales channels belonging to the Expedia Group or independent OTAs have a lower probability to be visible or the position leader than sales channels belonging to the Booking Holding.

6.2 Vertical ranking

We report estimation results in Table 3.8 for the model described in equation 3.2. The dependent variable in columns (a) to (e) is the ranking position of a hotel in the search results. A higher number corresponds to a worse ranking position in the Kayak search results. In columns (a) to (c), we sequentially add control variables for popularity of hotels and channels. Compared to the previous subsection, we additionally control for channel availability using group availability dummies and the number of sales channels. In columns (d) and (e), we restrict the sample to hotels having only one group being the price leader among the sales channels or multiple ones.

	Ordinary Least Square				
	Rank in search result				
# Group Price leaders(s)		All		Unique	Multiple
	(a)	(b)	(c)	(d)	(e)
min ln(price)	-1.527	-3.808	-3.849	-5.685	-9.533
	(-0.32)	(-0.78)	(-0.79)	(-1.15)	(-1.22)
# Sales Channels Price leader(s)	0.716	0.649	0.431	-0.190	3.216**
	(0.95)	(0.85)	(0.56)	(-0.14)	(2.80)
# Sales Channels	-2.231**	-2.400**	-2.979***	-2.466**	-5.899***
	(-2.87)	(-3.09)	(-3.60)	(-3.16)	(-4.66)
Group Availability dummies					
Booking Holding	-2.670	-2.551	-8.428	-4.455	-17.20
	(-0.70)	(-0.66)	(-1.92)	(-1.05)	(-1.83)
Expedia Group	4.297	4.212	2.207	2.412	13.37
	(1.05)	(1.02)	(0.52)	(0.60)	(1.28)
Direct website	5.564	6.795	7.051	-5.571	33.60***
	(1.10)	(1.33)	(1.38)	(-1.12)	(4.11)
Other OTAs	-10.30***	-10.26***	-9.640**	-10.47**	-5.031
	(-3.50)	(-3.44)	(-3.20)	(-3.23)	(-1.18)
Group Price leader dummies					
Booking Holding	-1.559	-0.646	-0.449	(Ref.)	-12.31*
	(-0.63)	(-0.26)	(-0.18)	(Ref.)	(-2.14)
Expedia Group	6.440*	7.299**	7.980**	7.713**	0.353
	(2.36)	(2.63)	(2.88)	(2.76)	(0.05)
Direct website	1.212	0.802	1.180	-11.05*	30.77***
	(0.32)	(0.21)	(0.31)	(-2.38)	(4.37)
Other OTAs	-1.063	-0.533	0.0568	-1.779	-6.031
	(-0.43)	(-0.21)	(0.02)	(-0.63)	(-1.54)
		(-0.36)	(-0.34)	(-0.78)	(0.17)
Constant	548.6***	583.1***	582.8***	620.6***	585.6***
	(21.96)	(8.40)	(8.41)	(8.67)	(6.34)
Request FE	yes	yes	yes	yes	yes
Hotel FE	yes	yes	yes	yes	yes
Hotel Popularity	no	yes	yes	yes	yes
Channel Popularity	no	no	yes	yes	yes
N	2,093,340	2,022,992	2,022,992	1,457,823	565,104

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.8: Hotel ranking at the group-level

Controlling for the number of sales channels, we show that the higher the amount, the better the ranking position of the hotel. Hotels are thus ranked better if Kayak is able to collect and provide many offers. The results also show that hotels are ranked worse for which the direct website is available and/or the price leader, but only when there are multiple price leaders. Accordingly, the evidence suggests that hotels are ranked worse for which the direct website of the hotel is a credible alternative to OTAs. At the same time, hotels having Booking Holding among the multiple price leaders are on average ranked better. Looking at the other OTAs, the results suggest that hotels are ranked worse for which the Expedia Group is a price leader, which is only present in the unique price leader specification when distinguishing in columns (d) and (e). The results also show that hotels seem to be ranked better when other OTAs are available. However, as the category "Other OTAs" is not a group by itself but gathers many independent OTAs, this result may be driven by some specific OTAs, where the very rare presence may coincide with a good position of the hotel.

Finding 2 (vertical ranking): Other things equal, hotels where an OTA of the Expedia Group has the lowest price have a worse ranking position in the Kayak search results.

7 Robustness checks

7.1 Sales channel-level analysis

In the previous section, the analyses were made at the group-level, which could disguise different signs for a given group. For instance, the Expedia Group contains highly popular platforms in France such as Voyages-sncf or Expedia, but also others that are not as popular such as Venere.com. Similarly, Booking.com and Agoda.com are not of the same importance. Despite controlling for channel popularity, estimated effects may be different as popular platforms may drive the overall effect at the group-level. For this reason, in this subsection we look at the estimation results at the sales channel-level.

In Table 3.9, we give the estimation results regarding the horizontal ranking from model 3.1 at the sales channel-level. The reference category for the sales channel affiliation is Booking.com.

The estimation on a more granular level provides more details on how the inner-group effect is composed. At the sales channel-level, one can see that the effect is driven by Booking.com being more visible, while Agoda.com as the other sales channel of Booking Holding has a lower probability to be visible with a similar magnitude than other OTAs. Therefore, it seems that Kayak puts Booking.com particularly more prominent and not necessarily other sales channels of the group. Results on visibility and position leadership are very similar.

		Linear Probability Model					
		Visible ($\# \leq 4$)			Position Leader ($\# = 1$)		
ln(price)		-0.166***	-0.168***	-0.168***	-0.008***	-0.008***	-0.008***
		(-27.83)	(-27.55)	(-27.64)	(-8.73)	(-8.95)	(-9.01)
Price leadership (ref: Not price leader)							
Among price leaders		0.275***	0.274***	0.274***	0.217***	0.217***	0.217***
		(62.16)	(61.78)	(61.51)	(93.93)	(93.18)	(93.05)
Unique price leader		0.451***	0.451***	0.445***	0.894***	0.894***	0.893***
		(69.42)	(69.09)	(67.16)	(225.72)	(225.43)	(221.62)
Sales Channel (ref: Booking.com)							
	Direct	-0.064***	-0.065***	-0.382***	-0.111***	-0.111***	-0.194***
		(-7.21)	(-7.17)	(-19.80)	(-18.97)	(-18.59)	(-15.64)
Booking Holding	Agoda.com	-0.142***	-0.145***	-0.473***	-0.142***	-0.142***	-0.228***
		(-25.93)	(-26.42)	(-25.15)	(-24.96)	(-24.76)	(-18.28)
Expedia Group	Expedia	-0.200***	-0.202***	-0.485***	-0.135***	-0.135***	-0.209***
		(-25.63)	(-25.66)	(-30.18)	(-22.55)	(-22.28)	(-18.68)
	Hotels.com	-0.036***	-0.038***	-0.363***	-0.107***	-0.106***	-0.191***
		(-6.17)	(-6.53)	(-19.75)	(-17.55)	(-17.23)	(-15.13)
	Venere.com	-0.289***	-0.291***	-0.617***	-0.188***	-0.186***	-0.272***
		(-40.74)	(-40.69)	(-31.86)	(-32.45)	(-31.91)	(-21.30)
	Voyages-sncf	-0.264***	-0.265***	-0.368***	-0.159***	-0.158***	-0.185***
		(-30.39)	(-30.49)	(-38.81)	(-28.33)	(-28.02)	(-28.58)
	(...)						
Other OTAs	Amoma.com	-0.082***	-0.083***	-0.409***	-0.123***	-0.123***	-0.209***
		(-9.19)	(-9.15)	(-20.15)	(-24.59)	(-24.34)	(-16.81)
	Hotelopia.com	-0.262***	-0.264***	-0.591***	-0.135***	-0.135***	-0.221***
		(-33.47)	(-33.69)	(-30.87)	(-27.40)	(-27.23)	(-17.72)
	Logitravel.fr	-0.045***	-0.046***	-0.371***	-0.107***	-0.107***	-0.192***
		(-4.20)	(-4.24)	(-17.39)	(-21.85)	(-21.49)	(-15.77)
	(...)						
Constant		1.344***	1.284***	1.617***	0.169***	0.146***	0.233***
		(45.89)	(26.67)	(32.91)	(29.32)	(11.46)	(13.58)
Request FE		yes	yes	yes	yes	yes	yes
Hotel FE		yes	yes	yes	yes	yes	yes
Hotel Popularity		no	yes	yes	no	yes	yes
Channel Popularity		no	no	yes	no	no	yes
N		15,585,011	15,089,104	15,089,104	15,585,011	15,089,104	15,089,104

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.9: Visibility and Position Leadership at the sales channel-level

Finding 3 (horizontal ranking): Other things equal, Booking.com has a higher probability than any other sales channel to be visible or the position leader.

For the vertical ranking of hotels (Appendix 3.D - Table 3.19), the analysis at the sales channel-level suggests that results obtained in the group analysis are particularly driven by some sales channels. First, the Expedia group effect stems from the OTAs Expedia and Venere.com. Second, the results reveal that the effect on the hotel ranking when independent OTAs are available is mainly driven by some particular platforms, for instance, Hotelopia.com. However, this is not a general result for all independent OTAs. In addition, the more detailed analysis suggests that hotels are ranked better if Voyages-sncf – one of the most popular OTAs in France – is among the available sales channels, irrespective of its price leadership. This suggests that Kayak takes into account the popularity of platforms.

Finding 4 (vertical ranking): Other things equal, hotels where Expedia has the

lowest price have a worse ranking position in the Kayak search results.

7.2 Different PPC regimes

For the past decade, numerous policy actions in the digital sphere have revolved around the hotel industry. One topic of interest have been price parity clauses, which are contractual agreements imposed by platforms to hotels to not offer lower prices on all other online channels (wide PPC) or the direct channel (narrow PPC). Booking.com committed to converting wide PPCs into narrow PPCs in July 2015 for European countries.[27] In the particular case of France, all types of PPCs have been banned by the Loi Macron[28].

In this subsection, we investigate whether the decision of banning PPCs in France had an impact on the effects estimated in the previous section. The theoretical literature (Boik and Corts, 2016) on the topic suggests higher commissions and final prices together with a reduction in the price dispersion for the different sales channels when a price parity clause is imposed. Empirically, Larrieu (2019) shows that the end of PPCs decreased average prices about 3-4% and increased price differentiation across OTAs by 1-2%. Therefore, with the end of the PPC, one may expect a higher price differentiation, especially with respect to channels competing with Booking.com that are subsequently more often the price leader than before. When PPCs were prohibited, OTAs adopted new possibilities to gain better visibility[29] in exchange for voluntary compliance to price parity. In this context, the vertical integration of OTAs and meta-search websites can be seen as an advantage. When contractual agreements no longer allow the platform to establish itself as the cheapest channel, the related meta-search website may adjust its ranking algorithm in order to highlight the affiliated channels or hide those that would be more competitive. We thus hypothesize Kayak to exploit the ranking more intensively without price parity, which can be either because hotels do not comply with it or because PPCs are banned by the law.

We consider two important dates in the process of abolishing price parity clauses in France: Booking.com committed to switching from wide to narrow price parity clauses on the 1^{st} of July 2015 and all price parity clauses were banned on the 6^{th} of August 2015 by the Loi Macron. Our period of observation ranges from the 13^{th} of October 2014 to the 18^{th} of September 2017 and therefore covers three distinct periods (Table 3.10).

In order to test this assumption for the horizontal ranking, we estimate equation 3.1 with additional interactions between the respective periods and group affiliations.

27 See the report on the "Monitoring exercise carried out in the online hotel booking sector by EU competition authorities in 2016."

28 Loi nº 2015-990 du 6 août 2015 pour la croissance, l'activité et l'égalité des chances économiques.

29 See, for instance, program details at Booking.com.

Period	Date	PPCs	# request	# hotel × request
A	10-2014 to 07-2015	All allowed	1,341	1,435,099
B	07-2015 to 08-2015	Wide banned Narrow allowed	178	200,005
C	08-2015 to 09-2017	All banned	762	458,301

Table 3.10: Periods of Different PPC Regimes

Table 3.11 displays the results.

	Linear Probability Model			
	Visible ($\# \leq 4$)		Position Leader ($\# = 1$)	
Group				
(ref: Booking Holding)				
Direct website	-0.002	0.045***	0.037***	0.0452***
	(-0.28)	(4.43)	(7.07)	(7.04)
Expedia Group	-0.106***	-0.113***	-0.071***	-0.075***
	(-19.80)	(-19.92)	(-23.63)	(-24.33)
Other OTAs	-0.087***	-0.065***	-0.024***	-0.026***
	(-13.53)	(-9.67)	(-10.32)	(-10.76)
Period B × Group				
Direct website		-0.137***		0.002
		(-8.16)		(0.17)
Expedia Group		0.109***		-0.033***
		(11.69)		(-5.32)
Other OTAs		0.071***		0.023***
		(4.29)		(3.94)
Period C × Group				
Direct website		-0.194***		-0.035***
		(-11.99)		(-3.07)
Expedia Group		0.007		0.028***
		(0.42)		(3.98)
Other OTAs		-0.184***		0.008
		(-12.91)		(1.13)
Request FE	yes	yes	yes	yes
Hotel FE	yes	yes	yes	yes
Hotel Popularity	yes	yes	yes	yes
Channel Popularity	yes	yes	yes	yes
N	15,089,104	15,089,104	15,089,104	15,089,104

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.11: Visibility and Position leadership with different PPC regimes

Regarding the direct website, results are quite similar independent of whether we take the visibility or the position leadership as a dependent variable. The evidence suggests that when PPCs were allowed (Period A), the direct channel had a higher probability than OTAs of Booking Holding to be highlighted. In contrast, when all price parity clauses were banned (Period C), the effect is the reverse. Period B seems to be a transition period with no significant effect on the position leadership while the direct channel already had a lower probability to be visible, which is strengthened in Period C. Interestingly, for the Expedia Group, even in the presence of price parity clauses, OTAs belonging to this group had a lower probability to be highlighted. With the removal of the wide PPCs, this relationship became considerably weaker.

Finding 5 (horizontal ranking): After the removal of PPCs, other things equal, direct websites of hotels have a lower probability of being visible or the position leader compared to sales channels belonging to Booking Holding.

The results of a similar analysis on the vertical ranking (Appendix 3.D - Tables 3.20 and 3.21) show that the negative effect associated to the Expedia Group being the price leader is only present in Period A. We additionally show that in the case of multiple price leaders, the negative effect linked to the direct website availability is only observed in Period A.

Finding 6 (vertical ranking): Other things equal, the removal of PPCs reduces the gap in ranking positions in the Kayak search results between hotels for which Booking Holding is one of the cheapest sales channels and hotels for which Expedia Group is the unique price leader.

Having these results, it is important to keep in mind that the periods were subject to various other events (e.g., terrorist attacks in 2015 and 2016) that could have generated effects beyond the request fixed effects. Furthermore, after the commitments by OTAs to European authorities to switch from wide to narrow price parity clauses, OTAs also started to offer new types of contracts to hotels (like the Preferred Partner Program), which imply a voluntary compliance to price parity. Therefore, even if clauses are banned by law, hotels might have been voluntarily agreeing to price parity.

Moreover, as OTAs can pay higher CPCs depending on hotels and requests, one might expect them to vary across time and thus be affected by price parity regulations. For instance, the ban of PPCs may have coincided with Booking Holding (and Expedia Group) increasing CPCs on MSPs resulting in a better horizontal (and vertical) ranking position on Kayak.

7.3 Chain affiliation

As shown in Table 3.1, 30% of hotels in the dataset are affiliated to chains. For meta-search websites and OTAs, we expect the treatment to differ when hotels are affiliated to chains. Indeed, chains may have better bargaining power and generally benefit from an increased visibility thanks to chain-level websites. For this reason, we estimate the same models including a chain parameter and compare results.

In the case of the horizontal ranking (Table 3.12), the results are quite robust with the exception of the effect related to the direct website of hotels. Interestingly, other things equal, this channel has a lower probability to be visible than one of Booking Holdings, which is mainly due to hotels affiliated to a chain. This is consistent with the policy applied by meta-search websites in favor of direct channels for small independent hotels.

	Linear Probability Model	
	Visible (# ≤ 4)	Position Leader (# = 1)
ln(price)	-0.185***	-0.020***
	(-28.87)	(-12.21)
Price leadership		
(ref: Not price leader)		
Among price leaders	0.233***	0.201***
	(65.72)	(89.26)
Unique price leader	0.474***	0.751***
	(69.39)	(135.44)
Group		
(ref: Booking Holding)		
Direct website	0.040***	0.070***
	(3.70)	(9.97)
Expedia Group	-0.091***	-0.063***
	(-16.65)	(-18.79)
Other OTAs	-0.072***	-0.017***
	(-10.57)	(-6.30)
Chain × Group		
Direct website	-0.107***	-0.077***
	(-7.10)	(-6.40)
Expedia Group	-0.063***	-0.035***
	(-7.30)	(-3.84)
Other OTAs	-0.062***	-0.030***
	(-5.69)	(-4.37)
Constant	1.324***	0.151***
	(28.03)	(6.75)
Request FE	yes	yes
Hotel FE	yes	yes
Hotel Popularity	yes	yes
Channel Popularity	yes	yes
N	15,089,104	15,089,104

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.12: Visibility and Position Leadership by chain affiliation

	Linear Probability Model					
	Visible (# ≤ 4)			Position Leader (# = 1)		
	All	Chain	Independent	All	Chain	Independent
ln(price)	-0.1851***	-0.1534***	-0.2010***	-0.0205***	-0.0090***	-0.0254***
	(-29.01)	(-13.24)	(-27.49)	(-12.45)	(-3.09)	(-13.34)
Price leadership (ref: Not price leader)						
Among price leaders	0.2337***	0.2081***	0.2447***	0.2011***	0.1736***	0.2134***
	(66.72)	(43.43)	(59.30)	(89.61)	(53.74)	(88.70)
Unique price leader	0.4747***	0.5097***	0.4641***	0.7519***	0.7990***	0.7389***
	(69.31)	(58.86)	(65.98)	(135.94)	(96.60)	(122.33)
Group (ref: Booking Holding)						
Direct website	-0.0023	-0.0622***	0.0381***	0.0369***	0.0020	0.0673***
	(-0.28)	(-5.28)	(3.54)	(7.07)	(0.26)	(9.29)
Expedia Group	-0.1056***	-0.1410***	-0.0946***	-0.0713***	-0.0798***	-0.0687***
	(-19.80)	(-15.87)	(-17.08)	(-23.63)	(-12.53)	(-20.99)
Other OTAs	-0.0873***	-0.1234***	-0.0761***	-0.0238***	-0.0296***	-0.0224***
	(-13.53)	(-10.87)	(-10.88)	(-10.32)	(-6.92)	(-8.50)
Constant	1.3260***	1.1628***	1.3998***	0.1517***	0.1110*	0.1708***
	(28.14)	(11.17)	(26.27)	(6.78)	(1.74)	(7.20)
N	15,089,104	3,996,922	11,092,174	15,089,104	3,996,922	11,092,174

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.13: Visibility and Position Leadership by chain affiliation

Finding 7 (horizontal ranking): Other things equal, the direct website of hotels

affiliated to a chain (resp. independent hotel) has a lower (resp. higher) probability to be visible or the position leader than OTAs belonging to Booking Holding.

A similar analysis is made for the vertical ranking (Appendix 3.D - Tables 3.23). Overall, when the direct channel is among the cheapest channels, independent hotels are ranked better, while chain hotels are ranked worse. Similarly, it seems consistent with the policy applied by meta-search websites in favor of direct channels of small independent hotels.

Finding 8 (vertical ranking): Other things equal, hotels affiliated to a chain (resp. independent hotels) for which the direct channel is among the cheapest sales channels have a worse (resp. better) ranking position in the Kayak search results.

8 Conclusion

We study the impact of the vertical integration between online travel agents (OTAs) and meta-search platforms (MSP) on the ranking algorithm used for the positioning of hotels and their sales channels on the MSP. We distinguish between the horizontal ranking of sales channels for a given hotel and the vertical ranking of hotels for a search request.

Our analyses of the horizontal ranking indicate that sales channels of OTAs by Booking Holding are more often position leaders (i.e. highlighted sales channel on hotel offers) than price leaders (i.e. among the cheapest sales channels). Using linear hotel and request fixed effects regressions that also account for prices and popularity measures, we additionally show that OTAs of the Booking Holding have a higher probability than any other OTA to be among the visible providers and to be the highlighted sales channel. For the vertical ranking, our results suggest that hotels are ranked worse in the Kayak search results when Expedia Group is the cheapest sales channel. We provide various robustness checks. First, we distinguish sales channels within groups and show that the two major OTAs Booking.com and Expedia drive the main results. Second, we use changes in the regulation of PPCs to estimate its potential impact on the Kayak ranking algorithm. Finally, we distinguish hotels affiliated to chains from independent ones and show that for the horizontal ranking, the direct channel of independent hotels is put more prominent compared to Booking Holding, while for the vertical ranking independent hotels present on Kayak with a direct website are favored too.

Overall, our results suggest that the ranking decision of an MSP is also affected by concerns beyond hotel and sales channel popularity. While this finding is not surprising for a profit-maximizing firm, it raises the question of whether the MSP optimizes its ranking only with respect to its own revenues (as it claims), or whether it takes the joint revenues of the integrated firm into account.

If the cost per click revenues for all sales channels and hotels were equal and constant across time, one could interpret our results such that – even controlling for differences in hotel and sales channel popularity – the MSP favors its affiliated sales channels in the ranking decisions. Based on this assumption, the observation that non-affiliated sales channels are becoming less visible (and hotels with lower prices on competing sales channels) is further evidence that the MSP uses its ranking to favor its own subsidiaries.

In practice, cost per click revenues might vary substantially across channels and time. This is also reflected in our findings that hotels which are affiliated to a chain (and therefore less likely to pay a high CPC) are, on average, ranked differently than independent hotels.

However, even if our results are solely driven by higher payments of the sales channels affiliated to the MSP, one has to bear in mind that these are payments within an organization with the same ownership. As integrated companies should have the means to compensate these payments, our results could still be consistent with the MSP favoring its own subsidiary.

In addition, results of the paper focus on Booking Holding's case. A scope for improvement would be to enrich the analysis by studying other independent MSPs or MSP affiliated to other groups. In particular, as Expedia Group, the second biggest player in the industry is also vertically integrated with several OTAs (like Expedia, Hotels.com, etc.) and the MSP Trivago, a similar analysis on this group would provide insight on whether our results are Booking-Holding-specific or more generally apply to the entire industry.

Our analysis cannot provide a definite conclusion on what sort of ranking of hotels and channels is socially optimal. However, we would like to point out two risks.

First, we find that the ranking optimization of an MSP may lead to a worse positioning of hotels with lower prices on competing sales channels, which in turn may have similar effects as a price parity clause (see also Hunold et al. (2020) in this regard). PPCs have been prohibited in many countries as they can lead to high commission rates and final prices. Second, deviations from a ranking that produces the highest match values for consumers may not only reduce consumer surplus but also allocational efficiency. However, it is unclear in this regard whether self-preferencing achieves worse results than the ranking optimization if the MSP and the OTA were not integrated. A separate MSP could have incentives to bias the search result more towards less popular sales channels. On the contrary, an MSP that is integrated with a popular OTA might improve consumers' search quality as it internalizes the profits of its subsidiary.

Empirically distinguishing between these outcomes is unfortunately beyond the scope of the present article and therefore left for future research.

Appendices

Distinct Time Horizon	Freq.	Percent	Cum.
1	2	1%	1%
2	18	4%	5%
3	51	12%	17%
4	11	3%	20%
5	62	15%	35%
6	103	25%	60%
7	122	30%	90%
8	11	3%	93%
9	30	7%	100%
Total	410	100%	100%

Table 3.14: Number of distinct time horizons by reservation date

3.A Google Trends Data

In the Kayak data, we identify 22 OTAs. For each of them, we download the associated relative search volume on Google (Google Trends) for the "search term" in France per month between 2014 and 2018.

On Google Trends, it is possible to look at searches of these keywords related to the "search term" in general, the respective website, or any other category deemed relevant by Google. As we do not observe the category "website" for all online travel agents, we use the more general query "search term" and adapt the request when needed (for instance, "Tablet hotels" instead of Tablet).

Besides these 22 OTAs, we distinguish the hotel's direct channel between large hotel chains and websites of independent hotels. We collect data from Google Trends for the 9 biggest hotel chains and normalize to zero for websites of small independent hotel. For each reservation date, we compute a popularity index (up to 100) by sales channel defined as the current Google Trends value divided by the maximum Google Trends value among the available sales channels for the request. The average popularity is higher for online travel agents than for hotel chains (Table 3.15), which is mostly driven by some very popular websites like Booking.com and voyages-sncf.fr,

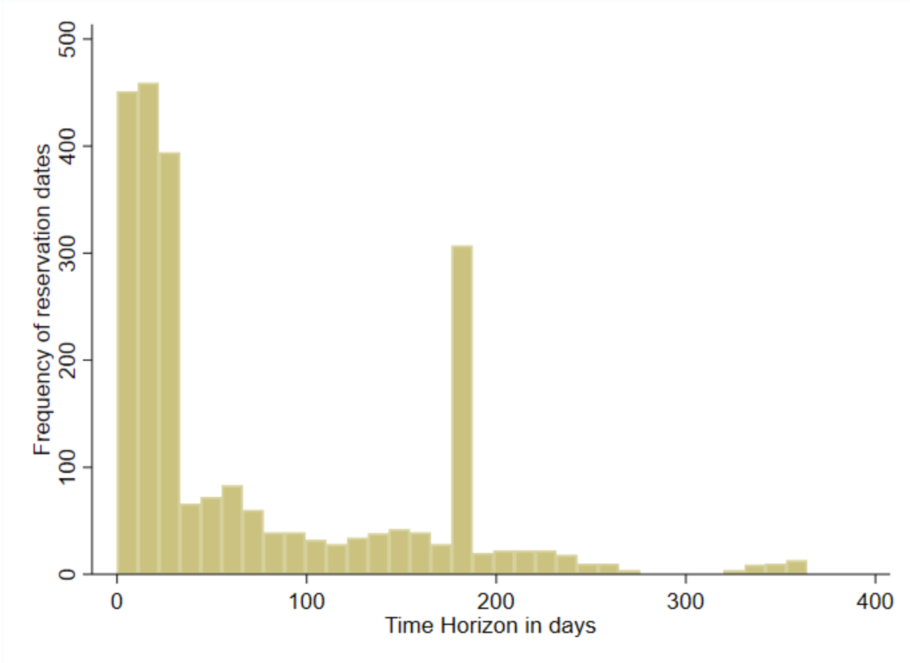


Figure 3.5: Frequency of reservation dates by time horizon

the French national railroad ticket booking website.

OTAs	Index	Hotel chain	Index
Booking.com	99	Ibis	29
voyage-sncf.fr	75	Novotel	16
TripAdvisor	66	Best Western	11
Expedia.fr	15	Mercure	9
Others	≤ 5	Kyriad	9
		Others	≤ 6
All	22	All	9

Table 3.15: Sales channels' popularity index

3.B Hotel characteristics

Affiliation	Freq.	Percent	Chain name	Percent
No chain	1,206	68%	-	-
Chain	578	32%	Ibis	18%
			Best Western	11%
			Mercure	10%
			Campanile	6%
			Novotel	6%
			Adagio	5%
			Kyriad	4%
			Première Classe	4%
			Others	37%
Total	1,784	100%	Total in chain	100%

Table 3.16: Hotel Chain Affiliation

Stars	Freq.	Percent	# Rooms
Not classed	11	1%	22
1	49	3%	63
2	249	14%	62
3	876	49%	62
4	514	29%	103
5	85	5%	105
Total	1,784	100%	76

Table 3.17: Average number of rooms by category of stars

3.C Provider vs Price leaders by sales channel

Sales channel	Position leader	Unique Price leader	Difference
Hotels.com	11%	1%	10%
Expedia.fr	8%	1%	7%
Venere.om	2%	1%	1%
voyages-sncf	3%	1%	3%
Ebookers.com	17%	25%	-8%
HotelClub	20%	25%	-5%
TripAdvisor	2%	17%	-15%
CheapTickets	2%	2%	0%
Booking.com	30%	6%	24%
Agoda.com	12%	15%	-3%
Amoma.com	25%	34%	-9%
Hotelopia.com	2%	3%	0%
Logitravel.fr	18%	24%	-6%
HotelTravel.com	9%	12%	-3%
Rumbo.fr	8%	11%	-3%
Hrs.com	6%	6%	0%
Direct	38%	54%	-16%

Table 3.18: Price vs Position leader by sales channel

3.D Robustness checks for hotel ranking

		Ordinary Least Square Rank in search results		
#Price leader(s)		All	Unique	Multiple
min ln(price)		-4.067	-6.019	-9.178
		(-0.85)	(-1.16)	(-1.49)
Channel Availability dummies				
	Direct	4.224	-7.186	21.26**
		(0.84)	(-1.45)	(3.20)
Booking Holding	Booking.com	-30.04***	-19.62*	-18.11
		(-3.40)	(-2.17)	(-1.33)
	Agoda.com	3.733	6.220	3.384
		(1.17)	(1.94)	(0.76)
Expedia Group	Expedia	3.540	4.489	-6.720
		(1.14)	(1.54)	(-1.22)
	Hotels.com	0.581	3.839	-3.123
		(0.20)	(1.36)	(-0.54)
	Venere.com	3.518	-0.158	21.93***
		(1.17)	(-0.05)	(3.54)
	voyages-sncf	-24.44***	-14.27*	-21.39*
		(-3.77)	(-2.18)	(-2.08)
	(...)			
Other OTAs	Amoma.com	-1.396	-3.037	2.418
		(-0.48)	(-1.00)	(0.62)
	Hotelopia.com	14.61***	14.16***	12.26**
		(5.09)	(4.72)	(3.08)
	Logitravel.fr	3.635	1.185	8.586*
		(1.50)	(0.51)	(2.27)
	(...)			
Channel Price leader dummies				
	Direct	1.952	-9.420	35.93***
		(0.58)	(-1.92)	(6.50)
Booking Holding	Booking.com	-0.322	(Ref.)	-3.046
		(-0.17)	(Ref.)	(-1.15)
	Agoda.com	-2.632	0.921	-7.908*
		(-1.13)	(0.22)	(-2.53)
Expedia Group	Expedia	-0.917	26.38***	9.899*
		(-0.33)	(3.90)	(2.43)
	Hotels.com	1.998	-1.794	1.070
		(0.73)	(-0.35)	(0.26)
	Venere.com	9.632**	-27.68**	-4.159
		(3.28)	(-3.01)	(-0.95)
	Voyages-sncf	-5.084	4.277	-0.599
		(-1.72)	(0.54)	(-0.15)
	(...)			
Other OTAs	(...)			
Constant		572.0***	599.4***	628.2***
		(8.44)	(8.39)	(7.50)
N		2,022,992	1,264,515	758,405

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.19: Hotel ranking at the sales channel level

	Ordinary Least Square Rank in search results					
# Price leader(s)	All		Unique		Multiple	
Group Availability dummies						
Booking Holding	-8.428*	-1.535	-4.455	1.701	-17.20*	-16.46
	(-1.92)	(-0.31)	(-1.05)	(0.35)	(-1.83)	(-1.48)
Expedia Group	2.207	3.111	2.412	4.838	13.37	-2.850
	(0.52)	(0.60)	(0.60)	(1.01)	(1.28)	(-0.23)
Direct website	7.051	6.595	-5.571	-6.840	33.60***	39.59***
	(1.38)	(1.10)	(-1.12)	(-1.17)	(4.11)	(3.88)
Other OTAs	-9.640***	0.0214	-10.47***	1.252	-5.031	-2.160
	(-3.20)	(0.01)	(-3.23)	(0.32)	(-1.18)	(-0.41)
Period B × Group Availability dummies						
Booking Holding		-12.49		-10.48		-38.75
		(-1.40)		(-1.18)		(-0.91)
Expedia Group		-7.891		-13.26*		54.83**
		(-1.02)		(-1.77)		(2.40)
Direct website		-14.28		-10.99		-28.50
		(-1.10)		(-0.84)		(-1.39)
Other OTAs		-40.72***		-39.78***		-40.60***
		(-4.42)		(-4.11)		(-3.64)
Period C × Group Availability dummies						
Booking Holding		-16.41**		-15.53**		11.30
		(-2.37)		(-2.23)		(0.63)
Expedia Group		3.955		1.638		61.44***
		(0.61)		(0.25)		(3.31)
Direct website		7.540		12.94		-17.55
		(0.88)		(1.54)		(-1.32)
Other OTAs		-11.26*		-18.12***		8.508
		(-1.89)		(-2.90)		(0.98)

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.20: Hotel ranking with different PPC regimes (Availability coefficients)

	Ordinary Least Square Rank in search results					
# Price leader(s)	All		Unique		Multiple	
Group Price leader dummies						
Booking Holding	-0.449	-2.045	(Ref.)	(Ref.)	-12.31**	-12.55**
	(-0.18)	(-0.71)	(Ref.)	(Ref.)	(-2.14)	(-2.01)
Expedia Group	7.980***	10.02***	7.713***	10.61***	0.353	-0.383
	(2.88)	(3.24)	(2.76)	(3.19)	(0.05)	(-0.04)
Direct website	1.180	1.497	-11.05**	-8.343	30.77***	22.55**
	(0.31)	(0.34)	(-2.38)	(-1.57)	(4.37)	(2.55)
Other OTAs	0.0568	1.919	-1.779	0.917	-6.031	-1.737
	(0.02)	(0.68)	(-0.63)	(0.28)	(-1.54)	(-0.38)
Period B × Group Price leader dummies						
Booking Holding		-5.923		(Ref.)		28.94
		(-1.10)		(Ref.)		(0.94)
Expedia Group		-11.22**		-2.448		0.788
		(-2.24)		(-0.40)		(0.05)
Direct website		-14.13		-0.540		-13.80
		(-1.22)		(-0.04)		(-0.73)
Other OTAs		-6.814		-3.797		11.42
		(-1.00)		(-0.51)		(0.96)
Period C × Group Price leader dummies						
Booking Holding		7.492*		(Ref.)		4.305
		(1.81)		(Ref.)		(0.43)
Expedia Group		-3.400		-10.00		-3.211
		(-0.80)		(-1.63)		(-0.25)
Direct website		2.035		-10.58		29.95**
		(0.29)		(-1.11)		(2.25)
Other OTAs		-6.162		-7.057		-22.19***
		(-1.32)		(-1.13)		(-2.82)
N	2,022,992	2,022,992	1,457,823	1,457,823	565,104	565,104

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.21: Hotel ranking with different PPC regimes (Price leadership coefficients)

	Ordinary Least Square		
	Rank in search results		
# Price leaders	All	Multiple	Unique
min ln(price)	-5.345	-11.124	-5.743
	(-1.10)	(-1.42)	(-1.17)
# Sales Channels	-3.042***	-6.258***	-2.472***
	(-3.71)	(-5.04)	(-3.17)
# Sales Channel(s) Price Leader(s)	-0.036	2.372**	-0.098
	(-0.05)	(2.07)	(-0.07)
Group Availability dummies			
Booking Holding	-3.630	3.343	(Ref.)
	(-0.76)	(0.29)	(Ref.)
Expedia Group	9.953**	20.782*	7.659*
	(2.06)	(1.75)	(1.70)
Direct website	-22.135***	-20.114**	-21.228***
	(-3.79)	(-2.10)	(-3.53)
Other OTAs	-13.764***	-12.888**	-11.475***
	(-3.84)	(-2.49)	(-3.05)
Chain × Group Availability			
Booking Holding	-5.690	-33.776*	(Ref.)
	(-0.80)	(-1.81)	(Ref.)
Expedia Group	-23.519***	-14.651	-19.650***
	(-3.48)	(-0.83)	(-3.41)
Direct website	60.133***	71.867***	41.004***
	(7.08)	(5.31)	(4.78)
Other OTAs	18.845***	25.779***	7.808
	(3.12)	(3.11)	(1.18)
Group Price leader dummies			
Booking Holding	-2.867	-22.686***	(Ref.)
	(-1.07)	(-3.82)	(Ref.)
Expedia Group	6.634**	-1.522	8.408***
	(2.34)	(-0.22)	(2.70)
Direct website	-11.475**	4.981	-13.028**
	(-2.25)	(0.55)	(-2.19)
Other OTAs	-4.888*	-13.804***	-2.019
	(-1.82)	(-3.09)	(-0.63)
Chain × Group Price leader			
Booking Holding	4.792	19.110*	(Ref.)
	(1.21)	(1.81)	(Ref.)
Expedia Group	2.104	-6.178	-3.174
	(0.55)	(-0.47)	(-0.53)
Direct website	26.841***	25.295**	12.386
	(4.01)	(2.03)	(1.35)
Other OTAs	11.316***	20.615***	2.677
	(2.74)	(2.71)	(0.42)
Constant	587.430***	595.570***	613.674***
	(8.43)	(6.42)	(8.58)
N	2,022,992	565,104	1,457,823

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.22: Hotel ranking by chain affiliation

	Ordinary Least Square					
	Rank in search results					
# Price leaders	All			Unique		
	All	Chain	Independent	All	Chain	Independent
min ln(price)	-3.849	40.902***	-11.818**	-5.685	30.552***	-9.907*
	(-0.79)	(4.29)	(-2.31)	(-1.15)	(3.06)	(-1.85)
# Sales Channel(s) Price Leader(s)	0.431	0.887	1.328	-0.190	-2.451	1.020
	(0.56)	(0.67)	(1.53)	(-0.14)	(-0.77)	(0.69)
# Sales Channels	-2.979***	-6.038***	-1.973**	-2.466***	-5.083***	-1.676*
	(-3.60)	(-6.41)	(-2.10)	(-3.16)	(-5.37)	(-1.89)
Group Availability dummies						
Booking Holding	-8.428*	-3.690	-9.061*	-4.455	1.177	-6.793
	(-1.92)	(-0.49)	(-1.86)	(-1.05)	(0.15)	(-1.45)
Expedia Group	2.207	16.057***	-0.906	2.412	16.933***	-1.869
	(0.52)	(2.68)	(-0.18)	(0.60)	(2.65)	(-0.40)
Direct website	7.051	35.568***	-25.557***	-5.571	20.280***	-23.471***
	(1.38)	(5.22)	(-4.25)	(-1.12)	(2.90)	(-3.81)
Other OTAs	-9.640***	-1.896	-12.848***	-10.469***	-6.341	-11.277***
	(-3.20)	(-0.45)	(-3.69)	(-3.23)	(-1.18)	(-3.08)
Group Price leader dummies						
Booking Holding	-0.449	-9.104**	-2.553	(Ref.)	(Ref.)	(Ref.)
	(-0.18)	(-2.12)	(-0.90)	(Ref.)	(Ref.)	(Ref.)
Expedia Group	7.980***	5.587	3.588	7.713***	14.275***	5.353*
	(2.88)	(1.12)	(1.18)	(2.76)	(2.74)	(1.69)
Direct website	1.180	-1.545	-7.843	-11.049**	-5.209	-10.061
	(0.31)	(-0.32)	(-1.48)	(-2.38)	(-0.82)	(-1.63)
Other OTAs	0.057	1.541	-4.895*	-1.779	7.374	-2.438
	(0.02)	(0.36)	(-1.77)	(-0.63)	(1.46)	(-0.75)
Constant	582.752***	112.911	690.028***	620.620***	175.455	698.148***
	(8.41)	(0.74)	(9.21)	(8.67)	(0.95)	(9.20)
N	2,022,992	500,009	1,522,975	1,457,823	324,607	1,133,208

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.23: Hotel ranking by chain affiliation (All vs Unique price leader)

	Ordinary Least Square					
	Rank in search results					
# Price leaders	All			Multiple		
	All	Chain	Independent	All	Chain	Independent
min ln(price)	-3.849	40.902***	-11.818**	-9.533	46.569***	-21.464***
	(-0.79)	(4.29)	(-2.31)	(-1.22)	(3.38)	(-2.76)
# Sales Channel(s) Price Leader(s)	0.431	0.887	1.328	3.216***	3.315*	2.368*
	(0.56)	(0.67)	(1.53)	(2.80)	(1.94)	(1.71)
# Sales Channels	-2.979***	-6.038***	-1.973**	-5.899***	-7.746***	-4.873***
	(-3.60)	(-6.41)	(-2.10)	(-4.66)	(-4.98)	(-3.28)
Group Availability dummies						
Booking Holding	-8.428*	-3.690	-9.061*	-17.196*	-15.051	-13.514
	(-1.92)	(-0.49)	(-1.86)	(-1.83)	(-1.14)	(-1.12)
Expedia Group	2.207	16.057***	-0.906	13.369	25.633*	4.039
	(0.52)	(2.68)	(-0.18)	(1.28)	(1.75)	(0.30)
Direct website	7.051	35.568***	-25.557***	33.595***	47.915***	-30.024***
	(1.38)	(5.22)	(-4.25)	(4.11)	(4.36)	(-3.09)
Other OTAs	-9.640***	-1.896	-12.848***	-5.031	-6.910	-9.815*
	(-3.20)	(-0.45)	(-3.69)	(-1.18)	(-1.30)	(-1.91)
Group Price leader dummies						
Booking Holding	-0.449	-9.104**	-2.553	-12.313**	-14.887*	-16.127***
	(-0.18)	(-2.12)	(-0.90)	(-2.14)	(-1.81)	(-2.62)
Expedia Group	7.980***	5.587	3.588	0.353	-5.726	0.367
	(2.88)	(1.12)	(1.18)	(0.05)	(-0.52)	(0.05)
Direct website	1.180	-1.545	-7.843	30.765***	7.748	13.121
	(0.31)	(-0.32)	(-1.48)	(4.37)	(0.91)	(1.39)
Other OTAs	0.057	1.541	-4.895*	-6.031	-1.647	-12.206***
	(0.02)	(0.36)	(-1.77)	(-1.54)	(-0.25)	(-2.82)
Constant	582.752***	112.911	690.028***	585.619***	132.868	723.206***
	(8.41)	(0.74)	(9.21)	(6.34)	(0.91)	(6.69)
N	2,022,992	500,009	1,522,975	565,104	175,366	389,729

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.24: Hotel ranking by chain affiliation (All vs Multiple price leaders)

3.E Chain affiliation and the ban of the PPC

	Linear Probability Model			
	Visible (# ≤ 4)		Position Leader (# = 1)	
ln(price)	-0.193***	-0.0289***	-0.0178***	-0.00164
	(-29.19)	(-5.36)	(-11.12)	(-1.43)
Among price leaders	0.240***	0.297***	0.194***	0.188***
	(65.24)	(66.44)	(85.45)	(94.16)
Unique price leaders	0.511***	0.615***	0.747***	0.735***
	(77.99)	(118.15)	(129.48)	(116.22)
Group (ref: Booking Holding)				
Direct website	0.0912***	0.0510***	0.0905***	0.0985***
	(7.71)	(4.66)	(11.21)	(11.81)
Expedia Group	-0.0893***	-0.131***	-0.0661***	-0.0697***
	(-14.63)	(-21.31)	(-19.30)	(-19.62)
Other OTAs	-0.0420***	-0.105***	-0.0180***	-0.0169***
	(-5.80)	(-13.73)	(-6.75)	(-6.46)
Chain (ref: Independent)	-	0.0313**	-	0.0341***
	-	(3.17)	-	(4.83)
Period C (ref: Period A)	-	0.240***	-	-0.0207***
	-	(19.52)	-	(-4.12)
Chain × Period C	-0.0975***	-0.116***	-0.0423***	-0.0403***
	(-6.20)	(-6.77)	(-5.15)	(-4.97)
Group × Chain				
Direct website	-0.152***	-0.150***	-0.106***	-0.112***
	(-7.84)	(-7.99)	(-7.33)	(-7.71)
Expedia Group	-0.105***	-0.105***	-0.0424***	-0.0432***
	(-8.98)	(-8.86)	(-4.33)	(-4.42)
Other OTAs	-0.0935***	-0.0731***	-0.0398***	-0.0429***
	(-6.55)	(-5.13)	(-5.05)	(-5.48)
Group × Period C				
Direct website	-0.283***	-0.307***	-0.100***	-0.0976***
	(-13.54)	(-14.52)	(-6.40)	(-6.39)
Expedia Group	-0.0377*	-0.0377*	0.0151*	0.0101
	(-2.44)	(-2.40)	(2.09)	(1.41)
Other OTAs	-0.220***	-0.267***	-0.00904	-0.0109
	(-15.25)	(-18.26)	(-1.29)	(-1.62)
Group × Chain × Period C				
Direct website	0.223***	0.259***	0.145***	0.140***
	(7.25)	(8.25)	(6.72)	(6.62)
Expedia Group	0.174***	0.171***	0.0560***	0.0602***
	(8.69)	(7.93)	(5.34)	(5.69)
Other OTAs	0.139***	0.152***	0.0668***	0.0636***
	(6.98)	(7.05)	(6.68)	(6.46)
Constant	1.339***	0.501***	0.143***	0.0143*
	(27.26)	(20.72)	(6.22)	(2.19)
Fixed effects	Yes	No	Yes	No
N	14,335,370	14,335,376	14,335,370	14,335,376

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.25: Visibility and Position leadership with different PPC regimes and chain affiliation

	Ordinary Least Square Rank in search results					
Chain (ref: independent)	-	-125.7***	-	-136.3***	-	-181.7***
	-	(-5.49)	-	(-5.11)	-	(-4.70)
Period C (ref: Period A)	-	-68.43***	-	-118.0***	-	13.85
	-	(-3.97)	-	(-6.22)	(0.00)	(0.32)
Price leader dummies						
Booking Holding	-6.019*	1.791	(Ref.)	(Ref.)	-21.35***	-45.99***
	(-2.00)	(0.28)	(Ref.)	(Ref.)	(-3.30)	(-4.21)
Expedia Group	7.226*	26.81***	12.04**	26.10**	-7.250	-25.65
	(2.33)	(3.81)	(3.28)	(2.80)	(-0.93)	(-1.75)
Direct website	-13.84*	-60.39***	-13.39*	-73.19***	2.998	4.547
	(-2.38)	(-4.42)	(-1.99)	(-4.68)	(0.32)	(0.19)
Other OTAs	-6.499*	-24.24***	-1.566	-29.42***	-15.34**	-34.18***
	(-2.17)	(-3.91)	(-0.43)	(-3.50)	(-3.06)	(-3.38)
Chain × Period C	13.48	51.42*	17.66	62.28*	-26.52	18.27
	(0.85)	(2.17)	(1.00)	(2.37)	(-0.56)	(0.32)
Group × Chain						
Booking Holding	4.631	-38.27***	(Ref.)	(Ref.)	16.65	3.863
	(0.97)	(-4.17)	(Ref.)	(Ref.)	(1.41)	(0.27)
Direct website	27.70***	94.35***	17.28	128.8***	13.69	12.83
	(3.52)	(5.99)	(1.61)	(6.12)	(0.92)	(0.49)
Expedia Group	-0.926	-15.01	-10.54	22.99	4.088	-0.748
	(-0.20)	(-1.63)	(-1.45)	(1.43)	(0.25)	(-0.04)
Other OTAs	12.10*	18.74*	-0.281	39.25**	36.05***	59.03***
	(2.58)	(2.17)	(-0.04)	(2.81)	(4.08)	(4.42)
Group × Period C						
Booking Holding	9.291	-43.26***	(Ref.)	(Ref.)	4.404	-14.88
	(1.90)	(-5.75)	(Ref.)	(Ref.)	(0.37)	(-0.82)
Expedia Group	-5.403	-22.95**	-13.59	21.55	21.80	9.067
	(-1.08)	(-3.18)	(-1.92)	(1.92)	(1.51)	(0.50)
Direct website	8.074	-2.809	0.625	47.28**	17.40	-24.92
	(0.76)	(-0.20)	(0.05)	(2.82)	(0.95)	(-0.75)
Other OTAs	-1.440	7.082	-7.859	55.39***	0.789	-6.471
	(-0.26)	(0.88)	(-1.08)	(5.00)	(0.08)	(-0.47)
Group × Chain × Period C						
Booking Holding	-2.755	36.57***	(Ref.)	(Ref.)	-2.427	28.76
	(-0.32)	(3.36)	(Ref.)	(Ref.)	(-0.12)	(1.28)
Expedia Group	4.722	23.12*	13.01	-3.589	-68.25**	-23.44
	(0.54)	(2.07)	(0.96)	(-0.20)	(-2.67)	(-0.89)
Direct website	-6.961	-7.194	-17.77	-27.82	23.70	14.94
	(-0.53)	(-0.43)	(-0.98)	(-1.25)	(1.00)	(0.41)
Other OTAs	-6.181	2.885	4.152	-13.94	-36.06*	-25.11
	(-0.69)	(0.26)	(0.33)	(-0.86)	(-2.35)	(-1.38)
Fixed Effects	Yes	No	Yes	No	Yes	No

Table 3.26: Hotel ranking with different PPC regimes and chain availability (Price Leadership coefficients)

	Ordinary Least Square Rank in search results					
min ln(price)	-4.848	-66.12***	-5.641	-66.44***	-8.494	-66.62***
	(-1.01)	(-9.53)	(-1.13)	(-8.65)	(-1.15)	(-8.41)
# Sales Channels Price leader(s)	0.302	2.560	0.0633	-0.608	2.553*	10.32***
	(0.39)	(1.70)	(0.05)	(-0.19)	(2.24)	(5.68)
Group Availability dummies						
Booking Holding	-0.529	6.065	0.677	0.616	-3.453	39.04*
	(-0.10)	(0.57)	(0.13)	(0.06)	(-0.25)	(2.18)
Expedia Group	15.78**	34.97**	13.93*	34.62**	21.22	64.69**
	(2.76)	(2.99)	(2.54)	(2.88)	(1.40)	(2.66)
Direct website	-27.36***	-82.60***	-24.78***	-79.09***	-30.91**	-114.0***
	(-3.97)	(-6.67)	(-3.55)	(-6.35)	(-2.89)	(-5.65)
Other OTAs	-5.604	-27.15**	-2.312	-21.94*	-5.953	-33.85**
	(-1.36)	(-3.22)	(-0.55)	(-2.31)	(-0.99)	(-2.97)
Group × Period C						
Booking Holding	-15.64	-61.21***	-14.25	-62.68***	16.47	-83.01**
	(-1.88)	(-5.03)	(-1.73)	(-4.97)	(0.63)	(-2.58)
Expedia Group	2.157	36.80**	0.808	38.21**	17.25	-73.96**
	(0.28)	(2.69)	(0.11)	(2.73)	(0.71)	(-2.68)
Direct website	10.98	-24.82*	8.612	-28.02*	12.13	0.988
	(1.03)	(-2.10)	(0.80)	(-2.35)	(0.77)	(0.04)
Other OTAs	-23.08**	-111.1***	-23.33**	-102.0***	-24.05*	-131.8***
	(-3.23)	(-10.43)	(-3.18)	(-8.91)	(-2.21)	(-8.67)
Group × Chain						
Booking Holding	8.873	-10.06	12.87	-8.772	-22.61	-36.02
	(1.06)	(-0.55)	(1.51)	(-0.45)	(-1.01)	(-1.38)
Expedia Group	-32.41***	-14.29	-27.38**	-10.22	-40.42	-28.61
	(-3.85)	(-0.81)	(-3.24)	(-0.52)	(-1.90)	(-1.02)
Direct website	62.85***	47.24**	41.40***	31.39	85.42***	117.9***
	(6.00)	(2.76)	(4.01)	(1.78)	(4.99)	(4.70)
Other OTAs	10.34	49.70***	5.673	38.30*	-2.331	44.18**
	(1.48)	(3.46)	(0.78)	(2.20)	(-0.22)	(2.70)
Group × Chain × Period C						
Booking Holding	-19.70	0.926	-20.61	0.515	-7.930	9.195
	(-1.46)	(0.05)	(-1.50)	(0.02)	(-0.21)	(0.21)
Expedia Group	19.73	-30.75	17.25	-33.51	98.27**	67.33
	(1.60)	(-1.60)	(1.36)	(-1.65)	(2.87)	(1.82)
Direct website	-32.13*	-33.22*	-19.26	-24.71	-53.20*	-78.31**
	(-2.39)	(-2.13)	(-1.42)	(-1.57)	(-2.33)	(-2.59)
Other OTAs	35.99**	27.57	18.39	33.04	79.61***	56.35**
	(3.26)	(1.73)	(1.64)	(1.77)	(4.84)	(2.74)
Constant	579.2***	996.3***	601.0***	1011.1***	598.3***	1023.0***
	(8.53)	(28.64)	(8.43)	(25.78)	(6.83)	(20.50)
N	1,829,420	1,829,479	1,327,699	1,327,758	501,656	501,721

Table 3.27: Hotel ranking with different PPC regimes and chain availability (Availability coefficients)

Bibliography

Boik, A. and Corts, K. (2016), 'The effects of platform most-favored-nation clauses on competition and entry', *Journal of Law and Economics* **59**, 105–134.

Bourreau, M. and Gaudin, G. (2018), 'Streaming platform and strategic recommendation bias', *CESifo Working Paper 7390* .

Cazaubiel, A., Cure, M., Johansen, B. O. and Vergé, T. (2020), 'Substitution between online distribution channels: Evidence from the oslo hotel market', *International Journal of Industrial Organization* **69**, 102577.

Chen, Y. and Yao, S. (2016), 'Sequential search with refinement: Model and application with click-stream data', *Management Science* **63**(12), 4345–4365.

Choi, J. P. and Yi, S.-S. (2000), 'Vertical foreclosure with the choice of input specifications', *Rand journal of economics* pp. 717–743.

De Corniere, A. and Taylor, G. (2014), 'Integration and search engine bias', *The RAND Journal of Economics* **45**(3), 576–597.

De Cornière, A. and Taylor, G. (2019), 'A model of biased intermediation', *The RAND Journal of Economics* **50**(4), 854–882.

De los Santos, B. and Koulayev, S. (2017), 'Optimizing click-through in online rankings with endogenous search refinement', *Marketing Science* **36**(4), 542–564.

Drugov, M. and Jeon, D.-S. (2019), 'Vertical integration and algorithm bias', *mimeo* .

Edelman, B. G. and Wright, J. (2015), 'Price coherence and excessive intermediation', *Quarterly Journal of Economics* **130**, 1283–1328.

Ghose, A., Ipeirotis, P. G. and Li, B. (2012), 'Designing ranking systems for hotels on travel search engines by mining user-generated and crowdsourced content', *Marketing Science* **31**(3), 493–520.

Ghose, A., Ipeirotis, P. G. and Li, B. (2014), 'Examining the impact of ranking on consumer behavior and search engine revenue', *Management Science* **60**(7), 1632–1654.

Hagiu, A. and Jullien, B. (2011), 'Why do intermediaries divert search?', *The RAND Journal of Economics* **42**(2), 337–362.

Hagiu, A. and Jullien, B. (2014), 'Search diversion and platform competition', *International Journal of Industrial Organization* **33**, 48–60.

Hagiu, A., Teh, T.-H. and Wright, J. (2020), 'Should amazon be allowed to sell on its own marketplace?', *mimeo* .

Hannak, A., Soeller, G., Lazer, D., Mislove, A. and Wilson, C. (2014), 'Measuring price discrimination and steering on e-commerce web sites', pp. 305–318.

Hunold, M., Kesler, R. and Laitenberger, U. (2020), 'Rankings of online travel agents, channel pricing, and consumer protection', *Marketing Science* **39**(1), 92–116.

Hunold, M., Kesler, R., Laitenberger, U. and Schlütter, F. (2018), 'Evaluation of best price clauses in online hotel bookings', *International Journal of Industrial Organization* **61**, 542 – 571.
URL: *http://www.sciencedirect.com/science/article/pii/S016771871830033X*

Hunold, M. and Muthers, J. (2017), 'Resale price maintenance and manufacturer competition for retail services', *The RAND Journal of Economics* **48**(1), 3–23.

Inderst, R. and Ottaviani, M. (2012), 'Competition through commissions and kickbacks', *The American Economic Review* **102**(2), 780–809.

Johansen, B. O. and Vergé, T. (2017), 'Platform investment and price parity clauses', *mimeo* .

Johnson, J. (2017), 'The agency model and mfn clauses', *The Review of Economic Studies* **84**(3), 1151–1185.

Koulayev, S. (2014), 'Search for differentiated products: Identification and estimation', *The RAND Journal of Economics* **45**(3), 553–575.

Lambrecht, A. and Tucker, C. (2019), 'Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads', *Management Science* **65**(7), 2966–2981.

Larrieu, T. (2019), 'Pricing strategies in online market places and price parity agreements: evidence from the hotel industry', *mimeo* .

Mandy, D. M. and Sappington, D. E. (2007), 'Incentives for sabotage in vertically related industries', *Journal of regulatory economics* **31**(3), 235–260.

Mantovani, A., Piga, C. and Reggiani, C. (2018), 'Much ado about nothing? online platform price parity clauses and the eu booking.com case', *mimeo* .

Mikians, J., Gyarmati, L., Erramilli, V. and Laoutaris, N. (2012), 'Detecting price and search discrimination on the internet', pp. 79–84.

Mikians, J., Gyarmati, L., Erramilli, V. and Laoutaris, N. (2013), Crowd-assisted search for price discrimination in e-commerce: First results, *in* 'Proceedings of the ninth ACM conference on Emerging networking experiments and technologies', pp. 1–6.

Raskovich, A. (2007), 'Retail buyer power through steering', *Economics Letters* **96**(2), 221–225.

Rey, P. and Tirole, J. (2007), 'A primer on foreclosure', *Handbook of industrial organization* **3**, 2145–2220.

Rey, P. and Vergé, T. (2016), 'Secret contracting in multilateral relations'.

Ronayne, D., Taylor, G. et al. (2018), 'Competing sales channels', *University of Oxford, Department of Economics* **843**.

Shen, B. and Wright, J. (2019), 'Why (don't) firms free ride on an intermediary's advice?', *International Journal of Industrial Organization* **64**, 27–54.

Ursu, R. M. (2018), 'The power of rankings: Quantifying the effects of rankings on online consumer search and purchase decision', *Marketing Science* **37**(4), 530–552.

Wals, F. and Schinkel, M. P. (2018), 'Platform monopolization by narrow-ppc-bpg combination: Booking et al.', *International Journal of Industrial Organization* **61**, 572–589.

Wang, C. and Wright, J. (2020), 'Search platforms: Showrooming and price parity clauses', *the RAND Journal of Economics* **51**(1), 32–58.

Zellner, A. (1962), 'An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias.', *Journal of the American Statistical Association* **57**, 348–68.

Zhang, X., Ferreira, P., Belo, R. and Godinho de Matos, M. (Forthcoming), 'Welfare properties of profit maximizing recommender systems: Theory and results from a randomized experiment', *Management Information Systems Quarterly* .

Conlusion

Les contributions de cette thèse ont été présentées en introduction. Dans cette section finale, certains des résultats des différents chapitres sont mis en regard afin d'avancer trois des principaux axes de réflexion à retenir après la lecture de cette thèse : la question de la substitution entre canaux de vente, celle des effets des clauses de parité de prix et enfin, la question de la transparence des algorithmes de classements.

Substitution : quelles méthodes, quels résultats ?

La question de la substitution entre canaux de vente n'est pas nouvelle. Elle a jusqu'alors souvent opposé, dans le cadre du développement du numérique, canal de vente physique et canal de vente en ligne d'un même fournisseur. Cette thèse se concentre sur la vente en ligne et s'interroge sur la substitution entre le canal de vente direct (i.e le site internet) et les intermédiaires en ligne d'un hôtelier. Les résultats du premier chapitre suggèrent que lorsque l'hôtelier cesse d'être référencé sur Booking.com, les consommateurs se reportent à 34% vers Expedia et à 13% vers le site internet de l'hôtel, tandis que 53% achètent auprès d'un hôtel concurrent, tous types de canaux confondus. Dans le cadre du second chapitre, le modèle de différence de différences montre que l'adoption du programme accroît le volume des ventes réalisées via Booking.com et Expedia tandis qu'il réduit celui du site internet de l'hôtel.

Les résultats de ces deux chapitres peuvent, à première vue, sembler contradictoires. Le premier indique que tous les canaux sont substituables et que cette relation est plus forte entre plateformes tandis que le second indique au contraire que la substitution entre une plateforme et la vente directe domine puisque les ventes des plateformes semblent complémentaires. Il est d'abord à noter que les deux chapitres n'appréhendent pas les mêmes effets. Si le premier cherche explicitement à établir les relations de substitution entre trois canaux de vente, le second étudie l'effet de l'adoption d'un programme dont les effets ne se limitent pas à une redistribution des volumes vendus entre les différents canaux de distribution. Par exemple, le gain de visibilité permis par le programme peut attirer de nouveaux consommateurs qui n'étaient pas clients auparavant. D'autre part, les méthodes d'estimation choisies dans ces deux chapitres n'apportent pas les mêmes informations en termes de définition de marché et d'origine des consommateurs. Dans le premier chapitre, la taille du marché est définie, ce qui permet d'identifier qu'un consommateur passe de Booking.com à Expedia pour acheter la même chambre d'hôtel. En revanche, l'estimation en forme réduite du second chapitre étudie les volumes dans leur globalité. L'identité des nouveaux consommateurs achetant via Expedia est inconnue : ils peuvent aussi bien être des fidèles de l'hôtel qui achetaient auparavant sur le site direct, auquel cas la substitution entre canaux est avérée, que de nouveaux consommateurs qui séjournaient avant dans un hôtel concurrent, auquel cas il s'agit d'une substitution entre hôtels et non entre canaux de distribution. Les résultats les plus robustes concernant les relations de substitution entre canaux de vente en ligne sont donc ceux du premier chapitre.

Les clauses de parité de prix sont-elles nocives ?

Cette thèse débute sur le constat qu'il n'existe pas de consensus dans la littérature théorique sur la nocivité des clauses de parité de prix et que les premiers travaux empiriques reposent bien souvent sur des analyses des prix affichés qui ne traduisent que des effets potentiels. Ces chapitres ne prétendent ni répondre à cette question ni remettre en cause les décisions prises par les autorités de concurrence mais apportent des éléments supplémentaires aux débats.

Le premier chapitre, au travers de la substitution entre canaux, évalue le rapport de force entre les hôteliers et certains de leurs intermédiaires, les agences de voyage en ligne. Ses résultats concluent sur l'existence d'une forte concurrence entre hôteliers, certes, mais également sur l'importance des plateformes dans le processus de recherche et d'achat des consommateurs. Leur prépondérance est telle qu'un hôtel qui n'apparaît pas sur une plateforme prend le risque de ne pas appartenir à l'ensemble de choix du consommateur. Face à la popularité croissante des agences de voyage en ligne comme Booking.com ou Expedia, pour un hôtelier : se référencer, c'est exister ! La question n'est donc plus vraiment de savoir si l'hôtelier doit apparaître sur une plateforme mais plutôt de connaître le prix à payer pour que cela soit rentable de le faire. Le second chapitre, avec l'exemple du programme *Partenaire Préféré*, démontre que les engagements de Booking.com et les interdictions légales en France concernant des clauses de parité n'empêchent pas leur application volontaire de la part des hôteliers. La course à la visibilité générée par la pression concurrentielle du marché se substitue à l'obligation contractuelle imposée par la plateforme. Si les résultats indiquent que les prix et volumes vendus par les hôteliers sur les plateformes augmentent, laissant ainsi penser que l'adoption du programme est au bénéfice de l'hôtelier, le consommateur, lui, paie l'augmentation du prix final. L'un des résultats importants de ce chapitre suggère que l'ennemi serait peut être mal identifié. En effet, l'augmentation des prix liée à l'adoption du programme ne provient pas de l'auto-application de la clause de parité en elle-même mais plutôt des autres éléments inhérents au programme, comme son coût mais également sa popularité auprès des consommateurs. D'autre part les résultats du troisième chapitre indiquent qu'un hôtel proposant un prix moins cher sur un canal de vente que sur un autre, peut, par le biais de l'algorithme, voir son classement être détérioré sur le site de comparaison de prix. Ainsi la distorsion imposée par ce type de plateforme reproduirait finalement une partie des effets anti-concurrentiels identifiés dans la littérature sur les clauses de parité de prix. L'utilisation d'algorithmes de classement opaques ainsi que la mise en place de programmes pour la visibilité seraient-ils finalement équivalents, voire pire, que ce que la régulation cherche à éviter ?

Vers davantage de transparence des marchés

Beaucoup de problèmes pourraient être évités avec de meilleures transparence et connaissance du marché. Aucun consommateur n'a entendu parler des clauses de parité de prix avant 2015, ne connaît la signification du pouce jaune présent à côté du nom d'un hôtel sur Booking.com, ni n'est au courant que Booking.com et Kayak appartiennent au même groupe. Leurs comportements auraient-ils changé en le sachant ? Quelle que soit la réponse, une certitude aurait pour le moins été acquise quant à la conscience d'un choix qui ne soit pas floué par la magie du numérique.

Si les informations se diffusent lentement, de récents évènements vont dans la direction d'une meilleure transparence. A l'automne 2020, Kayak a changé la présentation de son site. La plateforme indique maintenant en haut de sa page, en caractères lisibles pour tous, la mention suivante : *"Nous combinons popularité, prix, qualité et rémunération reçue pour classer les résultats. Nous effectuons des recherches non exhaustives auprès de plus de 65 fournisseurs, avec lesquels nous avons un accord commercial"*. Dans le même ordre d'idée, depuis quelques mois maintenant, un internaute qui passe sa souris sur l'icône du pouce jaune d'un hôtelier ayant souscrit au programme peut lire : *"Cet établissement fait partie de notre Programme Partenaire Préféré. [...] Il est possible que cet établissement participe au programme en payant une commission un peu plus élevée à Booking.com"*. Cela ne peut être qu'encourageant pour l'avenir. Il est également nécessaire que les consommateurs aient les clés pour comprendre ces messages.

Enfin, il est à noter que les instances de régulation ont bien conscience des ces débats. En décembre 2020, au moment où cette thèse est soutenue, la Commission Européenne s'apprête à présenter deux textes de loi importants visant à apporter davantage de transparence et de sécurité (*Digital Service Act*) ainsi qu'à instaurer une législation *per se* et ex-ante s'appliquant aux géants des marchés du numérique (*Digital Market Act*).

L'accroissement de la transparence est une recherche collective, également poursuivie par la Recherche dont le but est d'éclairer la société. Cela peut passer par la découverte d'un vaccin, qui sauverait l'humanité de l'épidémie de COVID-19, mais également, dans de moindres mesures, par l'apport d'informations aidant aux prises de décision. J'espère, à travers cette thèse, apporter ma pierre à l'édifice et contribuer à l'éveil d'une réflexion chez celles et ceux qui la liront.

List of Tables

List of Figures

ECOLE
DOCTORALE

Titre : Concurrence à l'ère du numérique : exemples dans l'industrie hôtelière

Mots clés : Relations verticales, distribution multicanal, substitution en ligne, clauses de parité de prix, visibilité, algorithmes de classement

Résumé : La numérisation croissante de l'économie bouleverse les canaux de distribution des vendeurs et favorise l'émergence de nouveaux acteurs : les plateformes d'intermédiation. Avec elles, le modèle traditionnel de revente laisse place à un modèle d'agence et crée un terrain fertile à différents cas de restrictions verticales. La numérisation grandissante des marchés pousse les autorités de concurrence à questionner et adapter leur analyse économique des pratiques. Cette thèse se concentre sur l'industrie hôtelière qui fait l'objet de plusieurs cas d'espèces. Les pratiques contractuelles telles que les clauses de parité de prix imposées par les agences de voyages en ligne aux hôteliers ont fait l'objet de nombreuses investigations, principalement en Europe. Le premier chapitre de cette thèse développe un modèle d'estimation structurelle de la demande permettant d'évaluer le degré de substitution entre les canaux de distribution en en ligne d'une chaîne d'hôtels, élément crucial retenu dans la définition des marchés. A l'issue des différents cas de concurrence, les clauses de parité de prix ont partiellement ou totalement été interdites dans plusieurs pays. En réponse, les plateformes d'intermédiation ont développé de nouveaux programmes offrant aux hôteliers une visibilité accrue en l'échange d'une application volontaire de la clause de parité de prix. Le second chapitre de cette thèse étudie l'effet de l'adoption de ce type de programme sur les prix fixés par les hôteliers en différenciant l'effet lié à l'accroissement de la demande, permis par les gains de visibilité, de ceux liés à l'auto-application de la clause et à la hausse des commissions inhérentes au programme. Cette thèse porte également sur le lien entre les agences de voyage en ligne et un autre type de plateforme sur ce marché, les sites de comparaison de prix. Ces derniers promettent aux consommateurs l'affichage des offres les plus compétitives du marché mais les critères utilisés dans les algorithmes de classement font désormais débat. D'autre part, l'intégration verticale des certaines de ces plateformes à de plus grands groupes, en possédant déjà plusieurs, interroge leur impartialité. Le troisième chapitre de cette thèse étudie l'impact de l'intégration de Kayak et plusieurs agences de voyage en ligne (comme Booking.com) au sein du groupe Booking Holding sur les classements des hôtels et des canaux de vente affichés sur le site de comparaison de prix.

Title : Competition in the digital era: evidence from the hotel industry

Keywords : Vertical relationships, multi-channel distribution, online substitution, price parity clauses, visibility, ranking algorithms

Abstract : The growing digitalization of the economy has been disrupting the sellers distribution channels and has been favoring the emergence of new players: intermediation platforms. Meanwhile the traditional resale model gives way to an agency model and creates fertile ground for different cases of vertical restraints. The increasing digitalization of markets therefore pushes competition authorities to question and adapt their economic analysis of practices. This thesis focuses on the hotel industry which has been the subject of several specific cases, especially in Europe. Contractual practices such as price parity clauses imposed by online travel agencies to hotels have been the subject of numerous investigations. The first chapter of this thesis develops a model of structural demand estimation, allowing to assess the degree of substitution between the online distribution channels of a hotel chain, a crucial element in the market definition. Following the various competition cases, price parity clauses were partially or completely prohibited in several countries. In response, the platforms have developed new programs offering hotels an increased visibility in exchange of the voluntary compliance of price parity clauses. The second chapter of this thesis studies the effect of the adoption of this program on the prices set by the hotels separating the effects linked to the demand increase, thanks to visibility gains, from those linked to the clause compliance and fee increase linked to the program. This thesis also deals with the link between online travel agencies and another type of platforms: price comparison websites. The latter promise consumers the display of the most competitive offers on the market but the criteria used in the ranking algorithms are now debated. Moreover, their vertical integration into larger groups, which also have online travel agencies, raises questions about their impartiality. The third chapter studies the impact of the integration of Kayak and several online travel agencies (such as Booking.com) within the Booking Holding group on the ranking of hotels and sales channels displayed on the price comparison website.

Institut Polytechnique de Paris
91120 Palaiseau, France