



HAL
open science

Essais on e-commerce platforms

Arthur Cazaubiel

► **To cite this version:**

Arthur Cazaubiel. Essais on e-commerce platforms. Economics and Finance. Institut Polytechnique de Paris, 2020. English. NNT : 2020IPPAG005 . tel-02913868

HAL Id: tel-02913868

<https://theses.hal.science/tel-02913868>

Submitted on 10 Aug 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



INSTITUT
POLYTECHNIQUE
DE PARIS

NNT : 2020IPPAG005

Thèse de doctorat



Essais sur l'économie des plateformes

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'Ecole 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 26 juin 2020, par

ARTHUR CAZAUBIEL

Composition du Jury :

Saïd SOUAM Université Paris Nanterre	Rapporteur
Wilfried SAND-ZANTMANN Toulouse School of Economics	Rapporteur
Marie-Laure ALLAIN Ecole Polytechnique	Présidente
Philippe CHONÉ Ecole nationale de la statistique et de l'Administration économique	Directeur de thèse

Remerciements

Je souhaiterais exprimer en quelques mots toute ma reconnaissance pour les personnes qui ont accompagné au long de ce périple.

En premier lieu, mon directeur de thèse, Philippe CHONÉ, qui a su trouver la distance optimale pour me laisser apprendre de mes erreurs sans jamais me laisser douter de trouver un jour la solution. Mes travaux seraient bien différents sans sa bienveillance et son aide pour rendre plus lisible mon premier chapitre. J'ai beaucoup appris de nos conversations devant son tableau blanc, et je n'oublierai plus de vérifier la multiplicité de mes équilibres à l'avenir.

Je tiens également à remercier les membres de mon jury, Wilfried SAND-ZANTMANN et Saïd SOUAM pour leur rôle de rapporteur, et Marie-Laure ALLAIN pour ses commentaires toujours très pertinents.

Mes coauteurs, Morgane CURE, Thibaud VERGÉ et Bjørn Olav JOHANSEN pour les deux papiers écrits ensemble. Morgane, ta redoutable efficacité nous aura épargné bien des embûches, et tu as su insuffler une vraie énergie dans nos projets, merci ! Au delà de ces considérations, j'espère avoir l'opportunité de retravailler avec toi.

Une thèse est aussi une aventure, surtout quand on la cumule avec un poste d'assistant de microéconomie à l'Ensaë ou un poste à l'Insee. Ces deux premières expériences professionnelles furent particulièrement riches, et je tenais à remercier toutes les personnes avec qui j'ai pu commencer à apprendre à travailler. Merci à Lionel, Corentin et Corinne de m'avoir parfois laissé délaissé mon travail pour ce document, avec toujours beaucoup de bienveillance. La team-assistant bien sûr, Malika, Vincent, Vanda, Romain, Jérémy, Jérôme, Christophe, Morgane, Lucie. On a beaucoup râlé, mais on aimait bien ça. Les autres services de la DE, naturellement. La division Commerce également. Elodie, garde moi une place en région, j'arrive.

Je voudrais également remercier tous mes co-doctorants au CREST-LEI Alessandro, Jiekai, Julien, Hugo, Etienne, Emmanuel, Manuel, Christophe, Alexis, Morgane, Rémi, Louis et les chercheurs qui ont toujours accepté de prendre du temps pour m'écouter : Alessandro, Christophe, Michael, Romain, Laurent, Thibaud, Roxana, Xavier, Philippe. Que ceux que j'ai pu oublier m'excusent !

Je remercie également ma famille, qui a appris (sauf Quiterie !) à ne plus poser la question délicate de l'avancement de cette thèse.

A Marion, pour nos longues discussions où ton bon sens détruisait mes modèles imaginaires, et tes encouragements incessants pour en reconstruire de nouveaux - et pour conclure cette thèse. Merci pour ton soutien indéfectible. Gaspard, sans la qualité remarquable de ton sommeil et ta capacité à jouer seul, cette soutenance aurait lieu en 2022.

Introduction

Cette thèse traite de l'économie des plateformes à travers trois chapitres indépendants. Le premier développe une modélisation théorique autour des ventes exclusives, ou ventes flash. Cette pratique, très courante sur internet, consiste à proposer à des consommateurs une offre à prendre ou à laisser aujourd'hui, sans possibilité d'en bénéficier demain. Le second chapitre, écrit avec Morgane CURE, Thibaud VERGÉ et Bjørn Olva JOHANSEN évalue la substituabilité des différents canaux de distribution de chambres d'hôtels en Scandinavie, notamment entre le site internet de l'hôtel, Booking, et Expedia. Nous analysons également la décision d'une chaîne d'hôtels de boycotter un canal de vente. Enfin, le troisième chapitre, avec les mêmes coauteurs, analyse l'implémentation d'une nouvelle stratégie par Booking auprès de ses hôteliers, et tente d'en déterminer les différentes composantes. Si ces travaux diffèrent par leurs approches -théorique, empirique structurelle, et empirique non-structurelle- et leurs champs d'application -les ventes exclusives, ou la distribution numérique des chambres d'hôtels, ils traitent différents aspects relatifs à l'économie numérique.

Le poids croissant du numérique dans l'économie est largement documenté, notamment par [Insee \(2019\)](#). Entre 2000 et 2016, le volume de valeur ajoutée généré par les sociétés TCSI¹ croit de 35% en Union Européenne, et 25% en France. Simultanément, les consommateurs se tournent de plus en plus vers le numérique : 80% des ménages Français disposent d'un accès à internet en 2017, contre 58% en 2008. Cette transition de l'offre et de la demande vers les canaux numériques a modifié les interactions entre différentes parties prenantes.

Un type d'acteur particulièrement visible a émergé au cours des années 2000 : les plateformes, définis par [Tirole \(2016\)](#) :

Votre carte Visa, votre PlayStation familiale, le moteur de recherche Google Chrome [...] renvoient au modèle de marché "biface", c'est à dire un marché où un intermédiaire permet à des vendeurs et des acheteurs d'interagir.

Cette intermédiation se construit notamment sur les externalités que génèrent chaque face du marché sur les autres. Ces plateformes, et notamment leur croissance, ont été considérablement étudiées par la littérature, qui a mis en évidence des stratégies de type *divide and conquer*. Les différents chapitres de cette thèse analysent deux plateformes:

- le premier chapitre s'intéresse aux sites de déstockage en ligne, comme *Veepee*, ex *vente-privée*.

Ces sites mettent en relation des consommateurs avec des vendeurs qui cherchent à écouler les

¹Technologie, contenu et support de l'information. Ce champ défini par l'Insee correspond approximativement aux « technologies de l'information et de la communication » (TIC) et aux « contenus et supports » définis par l'OCDE.

stocks restants de leurs collections éphémères passées. La plateforme se rémunère à travers une commission sur les vendeurs, ou en achetant directement le stock à des grossistes en dés-stockage². Ces sites ont une stratégie commerciale atypique, basée notamment sur l'exclusivité (il est nécessaire de s'inscrire, gratuitement, sur le site pour pouvoir accéder aux produits), la rapidité (les ventes sont de courte durée, avec des stocks limités), la régularité (tous les jours, de nouvelles ventes sont ouvertes) et l'imprévisibilité (les produits ne sont dévoilés que le jour de l'ouverture de la vente). Le travail que je développe propose une justification à cette imprévisibilité, dans ce contexte fondamentalement dynamique.

- les autres chapitres traitent des agences de voyage en ligne (OTA *Online Travel Agency*) comme *Booking.com* ou *Expedia*. Ces agences de voyage en ligne intermédièrent les consommateurs et les hôteliers. Elles offrent aux consommateurs un algorithme de recherche d'hôtels disponibles sur leur site, des informations sur chaque hôtel, et se rémunèrent sur les réservations effectuées via des commissions de 15% à 25%. Certaines pratiques commerciales exercées par ces plateformes ont engendrés de nombreuses discussions dans la littérature économique, notamment sur le caractère pro ou anti-concurrentiel des clauses de parité de prix³. Le second chapitre de cette thèse traite de cette problématique, à travers l'analyse d'une caractéristique clé du marché : la substituabilité des consommateurs entre canaux de réservation. Le chapitre trois analyse une autre pratique commerciale, le *Booking.com Preferred Program*, qui permet aux hôteliers de bénéficier de davantage de visibilité notamment.

La numérisation de l'économie contraint tous les acteurs à s'adapter et les plateformes qui ont émergés développent de nouvelles stratégies, avec les consommateurs finaux, ou entre fournisseurs et plateformes. Le reste de cette introduction présente mes différentes contributions plus en détail.

²Voir Chapsal et al. (2014) pour plus d'éléments

³Les clauses de parité de prix consistent en une obligation contractuelle de ne pas proposer le même bien à un tarif inférieur sur un autre canal de vente.

Chapitre un : une modèle théorique de vente flash

Le chapitre premier de cette thèse explore un modèle théorique de ventes-flash. Le principal et premier acteur en France qui fonde sa stratégie d'entreprise sur les ventes flash est *Veepee*, ex *Vente-privee.com*. En 2019, et depuis de nombreuses années, ce site internet se situe dans le top 5 des sites commerçants les plus visités⁴, devant d'autres géants internationaux comme *Booking.com*, et 20 millions de comptes clients créés en France. Initialement centré sur le prêt-à-porter, *Veepee* a diversifié son offre, entre voyages, divertissements, équipement du foyer, mais propose toujours une offre très abondante en équipement de la personne (prêt-à-porter, chaussures, maroquinerie, etc.). *Veepee* fonctionne sur le principe d'une succession de *ventes exclusives*, de courte durée, avec des stocks limités. L'idée sous-jacente est que les produits actuellement disponibles ne le seront plus demain. Un concurrent de *Veepee*, *Showroomprive.com* a d'ailleurs pour slogan "Il est urgent de se faire plaisir". Ce premier chapitre propose une justification à cette stratégie commerciale.

Ce travail s'inscrit dans la littérature de non-révélation de l'information, car la stratégie de *Veepee* consiste à ne pas dévoiler aux consommateurs les produits qu'il vendra demain, alors même qu'il dispose de cette information. Cette communication de l'information des consommateurs est naturellement relié aux modèles de recherche, comme [Stahl \(1989\)](#). Dans ce contexte, l'*obfuscation* est une stratégie mise en place par les entreprises pour diminuer l'intensité concurrentielle. D'autres modèles s'appuient sur une sophistication différentielle entre consommateurs. Les articles de [Ellison and Wolitzky \(2012\)](#) montrent que les entreprises peuvent utiliser les *add-on* pour accroître leurs profits⁵. L'*obfuscation* consiste pour cette littérature à ne pas dévoiler aux consommateurs naïfs l'existence d'achats supplémentaires, ce qui permet aux entreprises de se différencier et de minimiser l'intensité concurrentielle dans ce marché.

Ce chapitre s'inscrit dans cette littérature de l'*obfuscation*, mais en présence d'un monopole, comme [Petrikaitė \(2018\)](#). Cet auteur montre que cette stratégie de révélation de l'information permet de réduire la concurrence qu'un monopole se fait à lui-même, à la [Bulow \(1982\)](#). A la manière de [Petrikaitė \(2018\)](#), je modélise un monopole vendant plusieurs produits. Toutefois, ma modélisation est fondamentalement dynamique, car elle se base sur des biens durables, avec une succession temporelle de ventes telle que les consommateurs ne peuvent pas acheter le bien de première période à la seconde.

S'il existe également des coûts à revenir sur le marché, je montre que la stratégie d'*obfuscation* peut être profitable pour un monopole. En effet, le coût à revenir demain voir les produits disponibles a un impact très différent selon la stratégie informationnelle du vendeur. Si le vendeur dévoile toute

⁴Source : Fédération de la Vente à Distance (FEVAD)

⁵L'exemple canonique des imprimantes montre que, si certains consommateurs ne prennent pas en compte le coût des cartouches d'impression, alors, deux entreprises identiques en concurrence frontale peuvent se différencier et maximiser ainsi chacune leur profit. Une entreprise choisira de réduire sa marge sur l'imprimante et de l'accroître sur ses cartouches, tandis qu'une autre fera l'inverse, de telle manière à être au total un peu moins chère que la première. Les consommateurs qui anticipent qu'il faut prendre en compte le coût des cartouches d'encre lors de l'achat d'une imprimante choisiront la plus chère à l'achat, mais la plus économique à l'utilisation totale, tandis que les consommateurs qui oublient de prendre en compte ces coûts supplémentaires se reporteront vers une imprimante moins chère, mais avec des cartouches plus coûteuses. Il est particulièrement intéressant de noter qu'aucune entreprise n'a ici intérêt à "éduquer" les consommateurs. L'existence d'une sophistication différente entre consommateurs est exploitée via la présence d'une tarification non immédiatement visible à tous. Sans cette fraction de consommateurs, les deux entreprises entreraient en concurrence frontale et ne pourrait pas se différencier.

l'information disponible, cette friction impacte négativement son profit. Mais, si le vendeur choisit d'*obfusquer*, les coûts à revenir demain ont un impact non-monotone sur le profit. Ils augmentent dans un premier temps le profit du vendeur, puis le diminuent. Il existe un ensemble de valeurs des coûts de recherche tels que l'*obfuscation* est préférée à la révélation.

Le mécanisme économique expliquant ce résultat est très simple. En l'absence d'information sur le bien futur, les consommateurs forment des attentes sur l'utilité qu'ils retireront de la seconde période. Comme ils ne disposent d'aucun élément distinctif, tous les consommateurs forment les mêmes attentes, et disposent d'une espérance d'utilité à attendre CS_2 . Cette utilité à attendre est directement impactée par les coûts à revenir: $CS_2 - r$. Dans le cas limite où ces coûts sont exactement égaux à CS_2^- , les consommateurs disposent d'une espérance d'utilité nette des coûts à revenir nulle. Les consommateurs vont donc prendre des décisions de manière séquentielle : ils choisissent d'acheter le bien de première période sans prendre en compte le futur (car celui-ci amène une utilité nulle); dans un second temps, ceux qui n'ont pas acheté le bien de première période reviennent sur le marché (car ils sont indifférents à revenir ou non) et achètent éventuellement le bien de seconde période. Cette séquentialité dans la prise de décision peut notamment être exploitée par le vendeur pour maximiser son profit.

Les consommateurs sont parfaitement rationnels mais c'est la stratégie de dévoilement de l'information couplée à une friction du marché, qui génère des comportements ressemblant à des consommateurs rationnellement limités. Je mets notamment en évidence un parallèle entre myopisme (c'est à dire l'incapacité à prendre en compte les événements du futur dans nos choix actuels) et les coûts à revenir quotidiennement sur le marché. Ce premier chapitre montre donc que, dans certains contextes, la stratégie informationnelle d'une firme peut contraindre des consommateurs parfaitement rationnels à se comporter comme s'ils ne l'étaient pas.

De manière intéressante, tous les consommateurs ne sont pas également impactés par la stratégie du vendeur. En effet, les consommateurs qui ont une préférence très grande (respectivement faible) sur le bien de première période sont peu sensibles à l'information éventuellement disponibles : ils comptent acheter (respectivement ne pas acheter) le premier bien. En revanche, l'information est très sensible pour les consommateurs avec une valuation moyenne pour le bien de première période. Cette valeur différente de l'information peut justifier l'existence de stratégies mixtes plus complexes : par exemple, le monopole pourrait proposer à ces clients d'acquérir l'information sur les biens futurs. Certaines filiales de *Veepee* opérant en Amérique du Sud, comme *Privalia*, proposent déjà ce service, et permettent à des clients *premium* d'accéder en avance aux ventes exclusives.

Chapitre deux : Substitution entre canaux numériques

La croissance des canaux numériques dans la vente de biens et services a été pendant de nombreuses années une question récurrente pour les autorités de la concurrence. Plus récemment, face à l'importance de ces canaux, se pose également la question des restrictions verticales qui existent entre plateformes et fournisseurs, notamment les clauses de parités de prix (CPP) pour les plateformes de réservations de voyage en ligne (OTA). Ces clauses empêchent l'hôtelier de proposer des prix plus faibles sur d'autres canaux. Les conséquences concurrentielles de telles clauses restent complexes à déterminer précisément car elles impliquent différentes forces économiques, ce qui a conduit de nombreuses autorités de la concurrence à adopter des positions différentes en Europe.

L'argument en faveur de telles clauses est double : d'une part, les plateformes permettent aux consommateurs de diminuer drastiquement leurs coûts de recherche. L'existence d'une clause tarifaire permet aux consommateurs d'effectuer une unique requête pour accéder à tous les prix de tous les produits. En l'absence de clause de parité de prix, cette économie des coûts de recherche n'est plus valable, car les consommateurs doivent désormais effectuer des recherches supplémentaires pour trouver éventuellement moins cher directement sur le site internet de chaque hôtel. D'autre part, les OTA investissent dans des technologies qui permettent aux consommateurs de trouver l'hôtel qui leur convient le mieux, et offrent ainsi une visibilité à des hôtels qui ne réalisent pas les mêmes investissements qu'eux (notamment dans la publicité en ligne). Sans clause de parité, les hôtels pourraient profiter de cette visibilité accrue sans rétribuer la plateforme pour ce surcroît de clientèle qu'elle génère.

Les arguments contre les clauses de parité de prix sont également nombreux. Ces clauses empêchent les hôteliers de répercuter les coûts d'acquisition différents à leurs clients : les clients qui achètent directement à l'hôtel paient le même prix que ceux qui réservent via la plateforme, alors que ces derniers coûtent plus chers à l'hôtel. En conséquence, il est contraint d'augmenter son prix pour tous les clients, y compris ceux qui n'utilisent pas la plateforme. La concurrence entre plateformes est également impactée. Les incitations pour une plateforme à diminuer ses taux de commission sont plus faibles si une partie de cette diminution est répercutée vers la plateforme concurrente.

Une notion fondamentale dans la littérature théorique est la substituabilité des consommateurs entre canaux de réservation. Lorsque la substituabilité est élevée, les consommateurs changent volontiers de canal d'achat. Ainsi, le pouvoir de marché des plateformes est faible, car si elles augmentent leurs prix, les hôteliers peuvent crédiblement boycotter une plateforme avec des pertes de revenus limitées.

Ce chapitre exploite une base de données fournies par une grande chaîne hôtelière en Norvège, pour estimer la substitution entre canaux de vente numérique (*Booking.com*, *Expedia*, et le site de la chaîne d'hôtels). Par opposition à la littérature pré-existante, la connaissance des volumes achetés nous permettent enfin de mettre en place des méthodes d'estimations structurelles. Cet travail ne peut pas prétendre estimer l'impact concurrentiel des clauses de parité de prix, car il ne prend pas en compte la négociation qui existe entre la chaîne hôtelière et les plateformes de réservation.

Nous exploitons dans un second temps les estimations de la demande pour déterminer les coûts marginaux sous-jacents à chaque canal de vente pour l'hôtel. Nous montrons que les taux moyens de commission des différentes plateformes diffèrent : ils sont de 16% en moyenne sur *Booking.com* et de

19% sur *Expedia*.

Ces coûts marginaux nous enfin de simuler un boycott d'un canal de vente. Si la chaîne boycotte une plateforme, seule la moitié de ses consommateurs qui réservaient via ce canal vont se reporter sur un autre canal de vente. Les trois-quarts d'entre eux réserveront via l'autre plateforme, et le dernier quart via le site de la chaîne d'hôtel. Si la chaîne choisit de fermer son propre site de réservation, seuls 15% de ses consommateurs se reporteront vers des plateformes.

Nous exploitons enfin le fait que la chaîne ait effectivement boycotté *Expedia* pendant 18 mois pour comparer nos prédictions avec la réalité.

Ce second chapitre a été publié dans l'*International Journal of Industrial Organization* en 2020 sous le titre *Substitution between online distribution channels: Evidence from the Oslo hotel market*.

Chapitre trois : Évaluation d'une stratégie de *Booking.com*

Ce travail s'insère également dans la littérature sur les relations verticales qui existent entre plateformes de réservation et les hôteliers. De manière générale, la vision des plateformes par les hôteliers est ambivalente. Ces plateformes permettent à des hôtels d'être exposés à une demande internationale sans réaliser des investissements en visibilité, mais ces plateformes peuvent également cannibaliser les ventes directes aux consommateurs. Les hôteliers craignent également de perdre le contact avec leurs clients car ils ne disposent d'aucune de leurs données personnelles, le taux de commission de ces plateformes est parfois jugé excessif, et les hôtels se retrouvent en situation de forte concurrence les uns contre les autres, pour être le plus visible sur la plateforme.

Cette concurrence pour la visibilité notamment mise en évidence par [Hunold et al. \(Forthcoming\)](#)⁶ peut être exploitée par la plateforme à travers des offres particulières aux hôtels, comme le *Preferred Partner Program (PPP)* analysé dans ce chapitre.

Si les termes exacts de chaque *PPP* sont négociés avec chaque hôtel et restent secrets, il semble que le *PPP* fonctionne de la manière suivante :

1. Certains hôtels sont invités à devenir membres, notamment si les notes des consommateurs sont suffisantes
2. Les hôtels qui acceptent sont soumis à des obligations supplémentaires :
 - Le taux de commission des réservations augmente de 1% à 2% (il est d'environ 15% pour les hôtels standards)
 - Les hôtels doivent respecter formellement des clauses de parité de prix.

En contrepartie, les hôtels voient leur visibilité s'accroître :

- Un pouce jaune est ajouté à côté de leur nom
- La visibilité de ces hôtels serait accrue. Booking indique que ces hôtels *PPP* ont 65% de visiteurs en plus et 35% de réservations supplémentaires.

Ce programme impacte donc de nombreuses variables qui sont à même d'affecter la stratégie des hôtels : l'accroissement du taux de commission (susceptible d'augmenter les prix) ; l'implémentation de la clause de parité de prix (effet incertain) ou encore l'accroissement de visibilité (susceptible d'augmenter la demande, et donc les prix des hôtels).

Les contributions de ce chapitre sont de plusieurs ordres. Tout d'abord une régression hédonique des prix permet de fournir le cadre d'analyse et d'estimer la stratégie de prix des hôteliers. Dans un second temps, nous analysons plus en détail le programme *PPP* pour les hôteliers. Enfin, nous montrons que le respect de la clause de parité diffère selon les hôtels et le type de réservation. Cette hétérogénéité peut être exploitée pour déterminer l'impact de la clause de parité de prix sur les prix.

La régression hédonique de prix est un outil exploité abondamment par la littérature pour évaluer l'impact de certaines caractéristiques observables sur le prix ([Court \(1939\)](#), [Monty and Skidmore](#)

⁶Les auteurs montrent l'impact significatif du ranking dans Booking.com sur le prix des autres canaux de vente (notamment le site de l'hôtel et les autres plateformes).

(2003), Thrane (2007) par exemple⁷). Nous mettons en évidence l'importance de deux variables dans les prix : l'appartenance au programme *PPP* et le type de réservation (remboursable ou non remboursable). Nous montrons notamment que, une fois contrôlé par le taux d'occupation, le *lead-time*⁸ n'a que peu d'influence sur le prix.

La comparaison des hôtels *PPP* aux autres, par canal de vente et par pays permet de mettre en évidence une augmentation des prix de 5% à 13% selon le canal et le pays. Pour les volumes, les effets sont plus complexes à interpréter : les hôtels *PPP* voient leurs réservations augmenter sur *Booking.com* dans des proportions très variables selon les pays (+7% (non-significatif) en Norvège à +65% au Danemark). D'autres part, les réservations sur *Expedia* semblent également positivement influencées tandis que les réservations sur le canal direct diminuent. Ces éléments suggèrent une forme de substituabilité entre ces canaux.

Nous remarquons enfin que le *PPP* n'a pas affecté de la même manière toutes les réservations. En effet, la clause de parité de prix n'a été appliquée pour les réservations dites non-remboursables. Pour autant, nous n'avons pas eu connaissance de taux de commission différents entre ces réservations et les gains en visibilité s'appliquent à toutes les réservations uniformément. Nous exploitons donc cette différence de traitement dans une triple différence pour identifier l'impact causal de la clause de parité de prix sur les prix. Nos résultats montrent des coefficients faibles, non-significatifs, voire négatifs, ce qui est en ligne avec Mantovani et al. (2018).

Ce chapitre permet donc de mieux comprendre les enjeux entre vendeurs (hôteliers) et une plateforme (*Booking.com*), et notamment la capacité que possède la plateforme à mettre en concurrence ses vendeurs, et à extraire une partie de la rente de ceux-ci.

⁷Toutefois, la prise en compte de stratégies de tarification dynamique dans cette industrie est peu prise en compte par cette littérature.

⁸Le *lead-time* correspond à la durée entre la date de réservation et la date d'arrivée

Références

- BULOW, J. I. (1982): "Durable-goods monopolists," *The Journal of Political Economy*, 314–332.
- CHAPSAL, A. ET AL. (2014): "Three essays in competition policy," Ph.D. thesis, Universitat Pompeu Fabra.
- COURT, A. T. (1939): "Hedonic Price Indexes with Automotive Examples," *The Dynamics of Automobile Demand*.
- ELLISON, G. AND A. WOLITZKY (2012): "A search cost model of obfuscation," *The RAND Journal of Economics*, 43, 417–441.
- HUNOLD, M., R. KESLER, AND LAITENBERGER (Forthcoming): "Hotel Rankings of Online Travel Agents, Channel Pricing, and Consumer Protection." *Marketing Science*.
- INSEE (2019): *L'économie et la société à l'ère du numérique*, Insee Références.
- MANTOVANI, A., C. PIGA, AND C. REGGIANI (2018): "Much ado about nothing? Online platform price parity clauses and the EU Booking.com case," *mimeo*.
- MONTY, B. AND M. SKIDMORE (2003): "Hedonic Pricing and Willingness to Pay for Bed and Breakfast Amenities in Southeast Wisconsin," *Journal of Travel Research*, 42, 195–199.
- PETRIKAITĖ, V. (2018): "Consumer obfuscation by a multiproduct firm," *The RAND Journal of Economics*, 49, 206–223.
- STAHL, D. O. (1989): "Oligopolistic pricing with sequential consumer search," *The American Economic Review*, 700–712.
- THRANE, C. (2007): "Examining the determinants of room rates for hotels in capital cities: the Oslo experience," *Journal of Revenue and Pricing Management*, 5, 315–23.
- TIROLE, J. (2016): *Économie du bien commun*, Presses universitaires de France.

Contents

1	Strategic Obfuscation	1
1	Introduction	2
2	Model	4
3	Profit under Obfuscation and Revelation	5
3.1	Obfuscating	5
3.2	Disclosure	7
	Without price commitment	7
	With price commitment	10
4	Decision to obfuscate	10
4.1	Comparisons of the profits	10
4.2	A dynamic interpretation of obfuscation	12
5	Conclusions	13
1.A	Full resolution under obfuscation	14
1.B	Full resolution under revelation without commitment	14
1.C	Profit comparisons under revelation or under obfuscation	15
2	Substitution between Channels	17
1	Introduction	18
2	Data and Context	21
2.1	Data	21
2.2	Context: Delisting From Expedia	24
3	Demand Estimation	25
3.1	Specification	25
3.2	Instruments	27
3.3	Results	28
4	Supply Estimation	30
5	Simulated vs. Actual Effects of Delisting	32
5.1	Counterfactual analysis: removing one distribution channel	33
5.2	Comparing predicted and actual outcomes	36
6	Conclusion	38
2.A	Selected Observations	40
2.B	First-stage Estimates	40

CONTENTS

2.C	Robustness checks on the outside good	41
2.D	Capacity constraints	42
3	Preferred Partner Program	46
1	Introduction	46
2	Data	49
2.1	Presentation	49
2.2	Descriptive statistics	49
	Rate and channel identification	49
	Other variables	52
2.3	Additional data	52
2.4	Data quality in price explanation	54
3	Price analysis	55
3.1	Identification strategy	55
3.2	Instruments	56
3.3	Results	58
4	On the effect of the Preferred Partner Program	59
4.1	Volume definition	61
4.2	Identification strategy	61
4.3	Common Trend Assumption	62
4.4	Results	63
5	On the effect of Price parity application	65
5.1	General application of the price parity clause	65
5.2	Wide price parity effect	67
6	Conclusions	68
3.A	Samples comparisons	70
3.B	Preferred Establishments Program	71
3.C	Preferred Partner Program adoption	72
3.D	Model comparisons	73
3.E	Hotel selection	74
3.F	Details on the matching procedure	76
3.F.1	General matching	76
3.F.2	Type of offer	76
3.F.3	Price parity application by platform, offer and contract	78
3.G	Common trend for triple difference	79

Chapter 1

Strategic obfuscation or disclosure by a monopoly when today's substitutes are tomorrow's products

A seller decides either to disclose or to obfuscate its future products when it is costly for its consumers to return in the following period. Under revelation, these return costs are a pure loss for the seller. Under obfuscation, they can turn consumers into more myopic agents. This difference of action generates a range of return costs where obfuscation is preferred by the seller to revelation.

Keywords. Multiproduct-firm, monopoly, obfuscation, intertemporal price discrimination

JEL Classification. L12 D21 D42 D83

1 Introduction

In September 2009, The Economist shed light on the emergence of a new business model in the e-commerce, the *Online Exclusive Sales Market*, at the expense of traditional luxury retailers. This new flourishing market consists of websites (e.g. Veepee, HauteLook, Rue La La, Net-A-Porter. . .) selling high quality products, mostly in the ready-to-wear industry, to registered consumers. In barely ten years, the world leader, Veepee, almost reached 2 billions euro of sales revenues in 2015. Empirical studies of such websites are rare, but one can be found in [Helmets et al. \(2015\)](#). One specificity of this market is the very short duration of product availability (between one day and one week), with a very high turnover. Every day, new sales are opened and old ones are closed, whatever the remaining stock. This contributes, as stated by The Economist, to "*make shopping an urgent and competitive daily activity*". Even though the planning of sales is determined weeks in advance, consumers only discovers the new products at the opening date of the sales. Vanessa Friedman, in the New York Times in 2014, rephrases this idea as "*Get it cheap now before it disappears!*". Another specificity is that shopping on these websites is definitively time-consuming, as it requires consumers to log in everyday, at the opening hour of the sale, 7 or 9am. Interestingly, Veepee, the first-mover of this industry, adopted this specific business model before the entry of its competitors, suggesting that it is profitable even without competition.

This paper provides a rationale for the choice of this business model, using two key elements: consumers' information about their future preference and costs of returning to the seller. I show how a monopoly can obfuscate information to minimize the level of competition with itself in a dynamic framework. The baseline model developed here is a two-period game where one product is sold in each period. Consumers have a unit demand and have to choose between the two goods. They incur exogenous return costs. To maximize its profit the platform has two possible strategies: (i) disclosure: announce all the products in the first period, so that consumers can choose the product yielding the higher surplus or (ii) obfuscation: hide the second period's good so that consumers only have expectations of the future product. This modelization seems especially appropriate to describe situations where firms cannot fully determine return costs, but still commit on an information strategy, as in the *Online Exclusive Sale Market*.

I show that the choice of the information strategy depends on the level of the return costs. The role played by these is drastically different in each strategy. Under disclosure, they are perfectly equivalent to a cost of production for the seller, and a pure loss. Under obfuscation, they lower the waiting option value, convince consumers to disregard the future and help the seller to price discriminate. One result of this paper is to show that there exists an intermediate range of return costs such that obfuscation is preferred to revelation. An intuition for this is that obfuscation forces all consumers to share the same expectations of the future. If the return costs are close to the waiting option value, it is as if consumers behave myopically. Because of this myopia, the seller can easily price discriminate between the two periods. Under revelation, she would have faced strategic consumers and high production costs. Another result of this paper is to state that, if return costs were freely chosen by the seller, she would set strictly positive return costs, associated to an obfuscation strategy.

This paper contributes to the growing literature of firm's marketing strategies, such as rebates, behavioural based price discrimination or dynamic pricing. Many markets, such as hotels or restau-

rants, are using exploding offers, especially on the Internet. This work provides a credible rationale for obfuscation of future products, based on costs of coming back and on intertemporal price differentiation.

Optimal intertemporal pricing decisions of a monopoly has been extensively studied since [Coase \(1972\)](#)'s durable good seminal paper. A limit case of my modelization, with perfect correlation of valuations, reproduces [Coase's](#) result: the monopoly cannot gain more than a static one period monopoly. Unlike this field of literature, my point of interest is the choice of the information strategy by the monopoly (disclosure or obfuscation) in presence of frictions.

The question of communicating information to potential sellers has been extensively analyzed. Most of the theoretical literature about obfuscation has focused on competition. Some models, *à la* [Stahl \(1989\)](#) are models of incomplete search. A proportion of consumers are without search costs. This heterogeneity generates some randomization within the pricing strategy of the firms. In this context, obfuscation is interpreted as the proportion of consumers with search costs. Thus, increasing obfuscation is directly related to softening the competition, as firms compete only for the consumers without search costs. [Wilson \(2010\)](#) modifies [Stahl \(1989\)](#). [Wilson](#) considers obfuscation as a modification of the search costs. Another way of softening competition is [Ellison and Wolitzky \(2012\)](#). With the crucial assumption of consumers convex search costs, the introduction of search costs increases the marginal cost of future searches and therefore reduces the competition. [Armstrong and Zhou \(2011\)](#) develop a setting in which firms can generate search frictions to be prominent. Other models assume a differentiated sophistication between the consumers as in [Ellison \(2003\)](#) or [Gabaix et al. \(2006\)](#). This heterogeneity in degree of consumer rationality can be exploited with, for example, add-on. In all this literature, obfuscation allows firms to reduce competition. Other papers have tried to justify obfuscation for a monopoly. One approach is to use obfuscation to reduce costs. [Shin \(2005\)](#) or [Taylor \(2014\)](#) show that obfuscation helps a monopoly differentiate between high and low valued consumers. Therefore the monopoly can better allocate its selling efforts. Another approach is to use sequential search models.

Most closely related is [Petrikaitė \(2018\)](#). She models a static multi-product monopoly setting prices and search costs, and shows that obfuscation can be profitable, through the minimization of sales externalities. In contrast, my modelization is fundamentally dynamic, with one product per period, instead of multiple products and only one period. This dynamic framework allows me to draw a parallel between consumer myopia and obfuscation, and not minimization of sales externalities. Crucially, instead of search costs, I introduce return costs r as there is no possible recall.

Another widely developed literature in the optimal revelation of information is to be found in the auction literature. Following [Milgrom and Weber \(1982\)](#) seminal paper, many authors considered the level of information revealed to the potential bidders, as [Esó and Szentes \(2007\)](#) or, in a working paper [Budish and Zeithammer \(2011\)](#). Without any friction for the potential bidders (no reserve price, costs of bidding), an unpublished work by [Budish and Zeithammer](#) find the differential equations of the bidding strategies. The main result is that the seller's expected gain are higher when revealing the information. My work can be seen as an extension and adaptation of their model, where return costs soften this strong result. When the monopoly is allowed to sell information to its consumers, partial revelation is indeed the optimal strategy.

The rest of this paper is structured in the following way. Section 2 describes the basic assumptions

of the model. I solve it in Section 3, first under obfuscation and then under revelation. The results and comparisons are displayed in Section 4. Concluding remarks are provided in Section 5.

2 Model

A mass one of consumers has unit demand for one durable good existing in many distinct varieties. These varieties could represent different brands of the same product: Nike and Adidas running shoes for instance. A unique seller has two varieties $i \in \{1, 2\}$ of this good. Each consumer k has a personal instantaneous utility ϵ_{ik} , an idiosyncratic matching value between consumer k and variety i . These valuations are private information, whereas ϵ_{ik} 's distribution functions are common knowledge. Each variety is sold at a distinct period: the first variety is sold at the first period, and the second at the second period. There is no possibility for a consumer to purchase variety i in period t if $i \neq t$. The net utility of consumers buying a given product, from the first period's point of view, is the following:

$$\begin{aligned} (1 + \delta)\epsilon_{1k} - p_1 & \text{ for variety 1} \\ \delta(\epsilon_{2k} - p_2) & \text{ for variety 2} \end{aligned}$$

where $\delta \in [0, 1]$ represents the consumer's discount factor. Purchasing the first variety provides the instantaneous utility for two periods. Consumers also incur return costs r to come back in the second period to purchase the second variety. Contrary to Petrikaitė (2018) and most of the search model literature, it is impossible for consumers to buy variety 1 in period 2. This lack of recall motivates Vanessa Friedman's quote: the first variety has '*disappeared*'.

To maximise its profit, the seller chooses its prices p_1 and p_2 . It also chooses whether to disclose the identity of the next product to be sold (disclosure) or not (obfuscation). In this latter case, consumers know a second product is going to be sold, but fail to know their valuation for it: they can only have expectations about the future.

The timing of the game is the following:

$t = 0$ The monopoly chooses whether to disclose or to obfuscate.

if discloses

$t = 1$ The monopoly chooses p_1 . Consumers then discover ϵ_1 , ϵ_2 , p_1 and r . They choose whether they buy variety 1 and whether they come back.

$t = 2$ The monopoly chooses a price p_2 . Consumers find p_2 and incur r . They choose whether they buy variety 2.

if obfuscates

$t = 1$ The monopoly chooses p_1 . Consumers then find ϵ_1 , p_1 and r . They choose whether they buy variety 1 and whether they come back.

$t = 2$ The monopoly chooses a price p_2 . Consumers find p_2, ϵ_2 and incur r . They choose whether they buy variety 2.

To ensure tractability and readability, the following assumptions and notations are made. Consumers idiosyncratic matching values are independent and follow a uniform distribution: $\forall i, f_i =$

$\mathcal{U}_{[0,1]}$. Additionally, the number of varieties is very large, such that the expected valuation for an unknown second variety follows also a uniform distribution. To rule out any incentive to advance sales from the second period to the first, the seller is assumed not to discount the future.

Let's define $\tilde{p}_1 = \frac{p_1}{1+\delta}$ to have a per period price. Because of the dynamic nature of the game, consumers may form expectations of the future. Let's call p_2^e the expected second period price, when relevant.

3 Profit under Obfuscation and Revelation

This section explains the driving forces of the model and establishes a set of normative results between obfuscation and revelation. For readability, we drop all k subscripts but keep the distinction between the two distribution functions f_1 and f_2 .

3.1 Obfuscating

One possible strategy for the seller is not to reveal the second variety at the beginning of the first period. To properly solve by backward induction this model, one has first to determine the second period demand.

Consumers coming back in the second period are those expecting the surplus of waiting to be bigger than the first period surplus. Mathematically:

$$\begin{aligned} \delta \left(\int_{p_2^e}^1 (\epsilon_2 - p_2^e) f_2(\epsilon_2) d\epsilon_2 - r \right) &\geq \max((1 + \delta)(\epsilon_1 - \tilde{p}_1), 0), \\ \Leftrightarrow \begin{cases} \int_{p_2^e}^1 (\epsilon_2 - p_2^e) f_2(\epsilon_2) d\epsilon_2 \geq r \\ \tilde{p}_1 + \frac{\delta}{1+\delta} \left(\int_{p_2^e}^1 (\epsilon_2 - p_2^e) f_2(\epsilon_2) d\epsilon_2 - r \right) \geq \epsilon_1 \end{cases} \end{aligned}$$

These inequalities state that there exists a maximum value of the return costs \bar{r} compatible with the existence of a second period demand. This maximum value is such that the net expected second period consumer's surplus is null:

$$\bar{r} = \int_{p_2^e}^1 (\epsilon_2 - p_2^e) f_2(\epsilon_2) d\epsilon_2 = CS_2(p_2^e)$$

For $r > \bar{r}$, there is no second period demand. Consumers expect to have a negative surplus if they come back, and thus prefer to leave the market. On the other hand, if return costs are smaller, i.e. $r \leq \bar{r}$, some consumers choose to come back in the second period. These clients are the ones with a relatively low first period matching value, i.e.:

$$\begin{aligned} \epsilon_1 < \tilde{p}_1 + \frac{\delta}{1+\delta} \left(\int_{p_2^e}^1 (\epsilon_2 - p_2^e) f_2(\epsilon_2) d\epsilon_2 - r \right) &\Leftrightarrow \\ \epsilon_1 < \tilde{p}_1 + \frac{\delta}{1+\delta} (CS_2(p_2^e) - r) &= \tilde{\epsilon}_1. \end{aligned}$$

Consumers returning to the seller are the ones with a first period valuation lower than the per-period price \tilde{p}_1 adjusted by an expected gain of the future. This expected gain of the future is directly related to the time discount factor δ . All consumers with $\epsilon_1 < \tilde{\epsilon}_1$ come back in the second period and the potential second period demand is given by $F_1(\tilde{\epsilon}_1)$.

The second period profit writes:

$$\Pi_2 = F_1(\tilde{\epsilon}_1)p_2(1 - F_2[p_2]).$$

Thus, the optimal second period price is given by the monopoly price. Indeed, consumers choose in the first period to come back or not based on their first period valuation ϵ_1 . As there is no correlation between this valuation and the second period's valuation, the distribution of consumers coming back is equal to the distribution of consumers ex-ante.

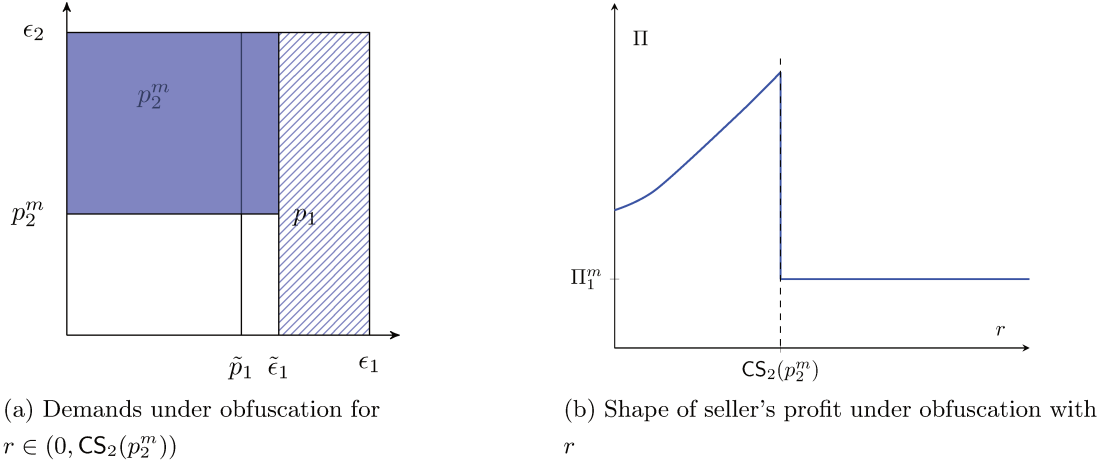
The total profit of the seller under obfuscation writes:

$$\Pi^O(p_1, r) = p_1(1 - F_1[\tilde{\epsilon}_1]) + F_1[\tilde{\epsilon}_1]\Pi_2^m.$$

Appendix 1.A determines $p_1^*(r)$ and proves that $p_1^*(r) > \Pi_2^m$. To conclude on the resolution under obfuscation, the influence of r on the profit with obfuscation is given by Lemma 1.

Lemma 1: *Under obfuscation, the shape of the seller's profit $\Pi^O(r)$ is given by Figure 1.1b. The profit is increasing with r , then abruptly decreasing at $r = CS_2(p_2^m)$, and constant afterwards.*

Figure 1.1: Demands and profit under obfuscation



Proof. There are three cases to consider. For $r > CS_2(p_2^m)$, there is no second period demand, and the total profit is simply Π_1^m .

For $r < CS_2(p_2^m)$, the intertemporal profit of the seller writes:

$$\begin{aligned} \Pi^O(p_1, r) &= p_1(1 - F_1[\tilde{\epsilon}_1]) + F_1[\tilde{\epsilon}_1]\Pi_2^m \\ &= p_1 + F_1[\tilde{\epsilon}_1](\Pi_2^m - p_1). \end{aligned}$$

We also know that the first period price is strictly higher than the second period profit Π_2^m . As visible on the Figure 1.1a, the seller offers the first period variety to all consumers, but the ones with $v_1 < \tilde{\epsilon}_1$ prefer to wait. These waiting consumers don't buy the first period variety ($-p_1$) and purchase the second variety at a price p_2^m with probability $(1 - F_2(p_2^m))$, generating a net per-consumer-profit of $(-p_1 + \Pi_2^m)$. As, in equilibrium, $p_1 > \Pi_2^m$, it is of the seller's interest to minimize the proportion of consumers waiting. Thus, an increase of r increases the profit.

There are multiple outcomes for $r = \text{CS}_2(p_2^m)$, depending on the fraction of consumers coming back. As there is no correlation between the valuations of the varieties, all consumers share the same expectations about the future $\text{CS}_2(p_2^m)$ and take the decision not to come back independently. As in the search literature, the introduction of heterogeneity between consumers, through a small correlation between the valuations or different sensibilities to the search procedure would smooth this result. \square

The non-monotonicity of the profit with respect to the return costs highlights the two effects of return costs. With very high return costs, consumers have a negative surplus of coming back, and thus, there is no second period demand anymore. The seller loses its opportunity to sell an additional variety. With lower return costs, when r increases, the option value to wait for consumers decreases. In the extreme case where $r = \text{CS}_2(p_2^m)^-$, this option value is null and consumers choose to purchase the first period variety if and only if $\epsilon_1 > \tilde{p}_1$ and come back otherwise. Differently stated, they act as if they were myopic. Their time discount δ plays no role anymore.

Obfuscation forces consumers to form expectations about the future. When the return costs are high, the net surplus of coming back is null, and perfectly rational consumers act *as if* they were myopic, even though they are not. This interpretation of obfuscation as a tool to generate some myopia in the consumers behavior differs from Petrikaitė (2018)'s point-of-view.

3.2 Disclosure

Another possible strategy of the seller is to reveal the second period variety before the first period purchase. To rule out any commitment issue, this section analyses both situations, with or without commitment power in the second period price.

Without price commitment

Without price commitment, consumers form expectations about the second price, called p_2^e . They then choose whether to come back or not comparing today's surplus and tomorrow's expected surplus. A consumer is willing to come back if and only if:

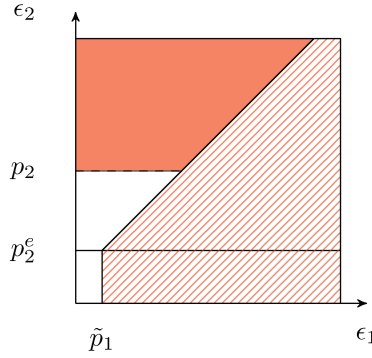
$$\delta(\epsilon_2 - p_2^e - r) \geq \max(0, (1 + \delta)\epsilon_1 - p_1)$$

Let's first analyze the specific case of $r > 0$. Consumers willing to return at the second period are the ones such that: $\epsilon_2 \geq r + p_2^e$. Say differently, the valuation for the second period variety ϵ_2 is strictly above p_2^e for all consumers returning in the second period. In this context, the seller's incentives are to set its second period price to $p_2 > p_2^e$. Without price commitment, there is no price p_2 simultaneously satisfying the consistency of consumer's expectations, the maximization of the second period's profit and the existence of a second period demand. With strictly positive return

costs and without commitment power, the monopoly cannot credibly sell in the second period. Thus, the profit without price commitment and revelation, with strictly positive return costs is given by the static monopoly price Π_1^m . This lack of commitment and its consequences have been extensively studied in the durable good literature, starting with [Coase \(1972\)](#).

Let's assume in the rest of this resolution without commitment that $r = 0$. The Figure 1.2 plots the first and second period demands on the pairs of valuations. These demands depend on three parameters: p_1, p_2 and p_2^e . The oblique line $(1 + \delta)\epsilon_1 - p_1 = \delta(\epsilon_2 - p_2^e)$ represents the indifference curve between buying now or waiting.

Figure 1.2: Demands in the revelation case without commitment if $p_2 > p_2^e$



The second period demand function writes:

$$D_2(p_2, p_2^e, p_1) = \int_{\max(p_2, p_2^e)}^1 F_1 \left[\tilde{p}_1 + \frac{\delta}{1 + \delta} (\epsilon_2 - p_2^e) \right] f_2(\epsilon_2) d\epsilon_2 + \alpha F_1(\tilde{p}_1) (F_2[p_2^e] - F_2[p_2]) \mathbf{1}_{[p_2 < p_2^e]},$$

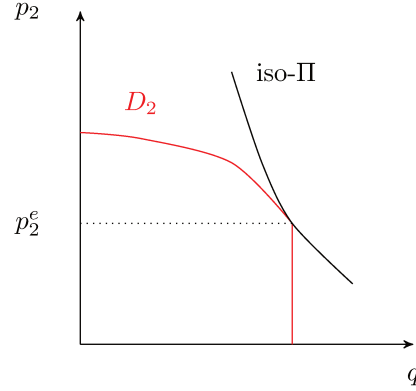
where α is the fraction of consumers choosing to return in the second period, even though their second period's valuation is below the expected price. As there is no cost of returning, α may take any value in $[0, 1]$. As soon as $\alpha < 1$, this second period demands has a kink in p_2 , for $p_2 = p_2^e$, preventing us from using standard first order conditions. Instead, as visible in Figure 1.3, we have an inequality condition between the slope of the iso-profit curve and the slope of the inverse demand function.

To be an equilibrium price p_2 must satisfy:

$$\underbrace{\frac{p_2}{D_2(p_2, p_2^e, p_1)} \Big|_{p_2=p_2^e+}}_{\text{Slope of the iso-profit curve}} \geq \underbrace{\frac{1}{-\frac{\partial D_2(p_2, p_2^e, p_1)}{\partial p_2}} \Big|_{p_2=p_2^e+}}_{\text{Slope of the inverse demand function}}.$$

Section 1.B in the appendix proves that this inequation admits a range of equilibrium prices $[\hat{p}_2(p_1), 1]$, where $\hat{p}_2(p_1)$ is such that the inequation is binding. Moreover, it also shows that the second period profit is decreasing with p_2 on $[\hat{p}_2(p_1), 1]$. In the second period, there exists therefore multiple equilibrium prices, yielding different second period profits, from 0 (with $p_2 = 1$) to a maximum and strictly positive second period profit (for $p_2 = \hat{p}_2(p_1)$).

Figure 1.3: Second Period Demand and Profit Maximisation



The seller's total profit is:

$$\begin{aligned} \Pi^{\mathcal{R}}(p_1) = & p_1 \int_{\tilde{p}_1}^1 F_2 \left(p_2(p_1) + \frac{1+\delta}{\delta} (\epsilon_1 - \tilde{p}_1) \right) f_1(\epsilon_1) d\epsilon_1 \\ & + p_2(p_1) \int_{p_2(p_1)}^1 F_1 \left(\tilde{p}_1 + \frac{\delta}{1+\delta} (\epsilon_2 - p_2(p_1)) \right) f_2(\epsilon_2) d\epsilon_2 \quad (1.1) \end{aligned}$$

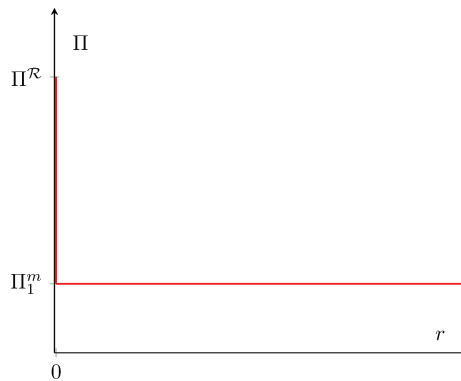
The seller chooses the optimal first period price maximizing $\Pi^{\mathcal{R}}(p_1, p_2(p_1))$.

To conclude on the resolution without price commitment, the profit of the seller is given by Lemma

2.

Lemma 2 : *Under revelation, without price commitment, the shape of the seller is given by Figure 1.4.*

Figure 1.4: Shape of seller's profit under revelation without commitment, with r



Proof. The profit is given by the following formula:

$$\Pi^{\mathcal{R}}(r) = \begin{cases} [\Pi^{\mathcal{R}}, \Pi_1^m] & \text{if } r = 0 \\ \Pi_1^m & \text{if } r > 0 \end{cases} \quad (1.2)$$

For $r = 0$, any second period price $p_2 \in [\hat{p}_2(p_1), 1]$ is an equilibrium price, and thus, the profit of the seller may take multiple values. In the worst case scenario, $p_2 = 1$, the total profit is simply equal to the static monopoly price, yielding Π_1^m . In the best scenario, $p_2 = \hat{p}_2$, the seller gains more, $\Pi^{\mathcal{R}} > \Pi_1^m$. \square

This apparent discontinuity in r is well known in the search literature and could be ruled out, by introducing some heterogeneity on the consumer side. For instance, a sensibility to the search-costs-parameter could generate some frictions solving this issue. Nevertheless, without commitment, search costs have a very strong and negative impact on the total profit.

With price commitment

With price commitment, strictly positive return costs and a second period demand can coexist. Demands are now functions of p_1 and $p_2 + r$. The profit of the seller is:

$$\begin{aligned} \Pi^{\mathcal{R}}(p_1, p_2, r) = & p_1 \int_{\tilde{p}_1}^1 F_2 \left(p_2 + r + \frac{1 + \delta}{\delta} (\epsilon_1 - \tilde{p}_1) \right) f_1(\epsilon_1) d\epsilon_1 \\ & + (p_2 + r - r) \int_{p_2 + r}^1 F_1 \left(\tilde{p}_1 + \frac{\delta}{1 + \delta} (\epsilon_2 - p_2 - r) \right) f_2(\epsilon_2) d\epsilon_2 \quad (1.3) \end{aligned}$$

The impact of r on the profit is given by Lemma 3.

Lemma 3: *The shape of the profit function with price commitment is represented in Figure 1.5b.*

Proof. Demands for the first and for the second variety are functions of p_1 and $p_2 + r$ only. Thus, if r decreases, the seller can increase p_2 by the same amount, have the same demands, but higher profits. As $r \rightarrow 1$, no consumer is willing to buy the second period product, and therefore, the total profit of the seller is Π_1^m . \square

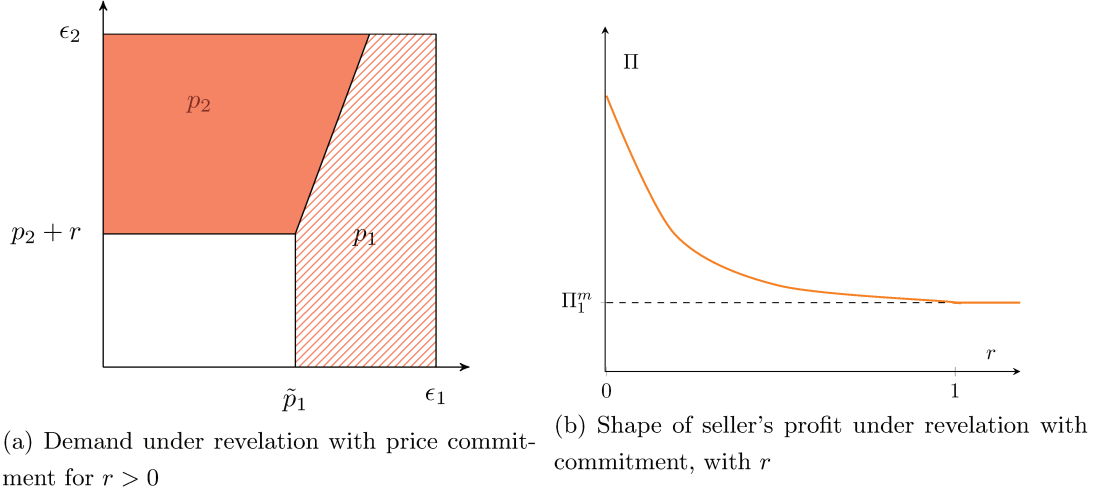
The profit under revelation is strictly decreasing with the return costs. These costs play actually the same role as production costs r for the second variety, where the final price is $p_2 + r$. Under revelation, consumers take into account the total price paid for the second variety, including return costs, whereas only p_2 is collected by the seller. The presence of return costs r , with price commitment is therefore identical to some linear production costs r for the second period product.

4 Decision to obfuscate

4.1 Comparisons of the profits

Based on the previous Lemmas, the following Propositions can be stated.

Figure 1.5: Profit under revelation



Proposition 1: *With strictly positive return costs, if the seller can't commit on prices, it will choose to obfuscate.*

Proof. Without price commitment and under revelation, the profits of the seller are equal to the static monopoly profit (Lemma 2). Under obfuscation, if $r < CS_2(p_2^m)$, the seller earns as much as under revelation by fixing $p_1 = 1$. If $r \geq CS_2(p_2^m)$, the profit is also equal to the static monopoly profit. \square

This proposition states that obfuscation can be used as a way to increase profit when the seller is unable to credibly commit on a second period price. Obfuscation is a soft commitment device.

More interestingly, even with commitment power, obfuscation is a credible strategy used by sellers. **Proposition 2 :** *If the seller can commit on prices, the seller chooses to obfuscate for an non-empty range of return costs.*

Proof. Figure 1.6 provides a graphical comparison of the profits under obfuscation or revelation with price commitment. To prove that there exists a non-empty range of return costs such that the seller prefers to obfuscate, it is sufficient to prove that:

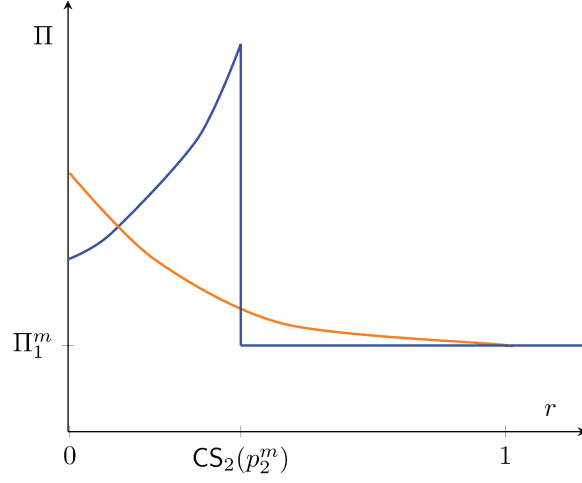
$$\Pi^O|_{r=CS_2(p_2^m)^-} > \Pi^R|_{r=0}.$$

Appendix 1.C proves this last point by brute force. As profit functions are continuous on $[0, CS_2(p^m))$, a direct application of the intermediate value theorem ends the proof. \square

Lastly, if return costs were freely chosen at no costs by the seller, Proposition 3 defines the preferred regime.

Proposition 3: *If return costs were endogenously chosen, the seller would choose to obfuscate and have $r = CS_2(p^m)$.*

Figure 1.6: Profits under Obfuscation (red) and under Revelation (blue) with price commitment



Proof. Direct consequence of the following inequation:

$$\Pi^{\mathcal{O}}|_{r=CS_2(p_2^m)^-} > \Pi^{\mathcal{R}}|_{r=0}.$$

□

These three propositions provide a rationale for the choice of obfuscation as a business model. Proposition 1 suggests that firms that cannot credibly commit to a second period price could still use obfuscation as a commitment device and thus earns higher profits. Proposition 2 states that, even with price commitment, obfuscation can be the best strategy if the return costs are in the proper range. Proposition 3 implies that firms would choose to have high return costs and obfuscation if they could freely endogenize r , as in [Petrikaitė \(2018\)](#).

4.2 A dynamic interpretation of obfuscation

The dynamic framework in this model allows me to reinterpret the previous propositions to spotlight obfuscation-generated-myopism. To do so, let's represent the demands of the seller in the two extreme cases in [Figure 1.7](#).

The seller's profit under revelation with commitment power, stated in [1.3](#), can be rewritten in the following way:

$$\Pi^{\mathcal{R}} = p_1(1 - F_1(\tilde{p}_1)) + p_2 F_1(\tilde{p}_1)(1 - F_2(p_2 + r)) + (p_2 - p_1) \mathbb{P}[(v_1, v_2) \in \mathcal{A}] \quad (1.4)$$

where

$$\mathcal{A} = \{(\epsilon_1, \epsilon_2) \mid \tilde{p}_1 \leq \epsilon_1 \leq 1; p_2 + r \leq \epsilon_2 \leq 1; \epsilon_2 \geq p_2 + r + \frac{1 + \delta}{\delta}(\epsilon_1 - \tilde{p}_1)\}.$$

[Equation 1.4](#) highlights three components in the seller's profit, when revealing. The two first terms are similar to the profit realized by the monopoly under obfuscation. The last term reflects the arbitrage possibilities consumers gain because of the revelation. Consumers in \mathcal{A} would gain a strictly positive utility if they purchased in the first period, but they prefer to wait and have the second variety.

Figure 1.7: Demands comparisons

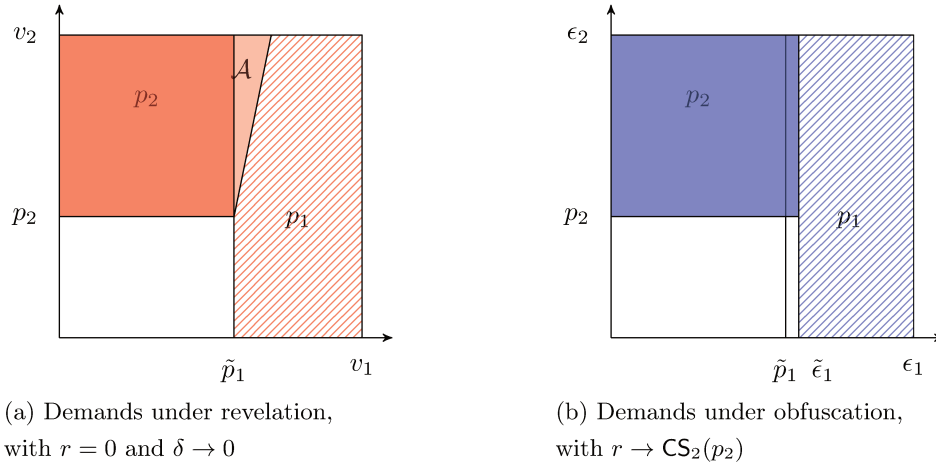


Figure 1.7 represents the demands under revelation or obfuscation, keeping p_2 and \tilde{p}_1 constant. Under revelation, the fraction \mathcal{A} of consumers choosing not to purchase the first period variety, even though they would gain a strictly positive net surplus to do so, is clearly decreasing. In the extreme case of myopic consumers, this fraction drops to zero. With perfectly myopic consumers, coming back in the second period provides no information at all about the valuation of the consumers. In this case, the shapes of the demands functions are identical to the demands under obfuscation. The similarity is perfect for $\delta = 0$ between obfuscation with $r = CS_2(p_2^m)$ and revelation with $r = 0$.

5 Conclusions

This article studies how information disclosure or obfuscation could be used by firms to increase their profit, in a dynamic framework.

I find that obfuscation is a credible business model in some situations. First, for firms lacking the ability to commit on prices, obfuscation can be used as a soft commitment instrument. Second, there might exist an intermediate range of return costs such that obfuscation is preferred to revelation. Third, as in [Petrikaitė \(2018\)](#), I find that firms would rather obfuscate than reveal if they could freely endogenize the value of the return costs r . Interestingly, these situations might give the false impression that consumers behave myopically, when they are in fact in situations where their net surplus of waiting is null.

Future research might want to look at more complex strategies, to rationalize partial disclosure strategies between consumers through *Premium services*. Such behaviors are observed in the *Exclusive sale industry*. Privalia, a platform operating in Brazil, Italy, Spain and Mexico, and recently acquired by Veepee, sells *Premium service*. Consumers with *Premium service* can access in advance to some sales, and have therefore access to the information in advance. Allowing the monopoly to sell the information it creates could soften my results.

1.A Full resolution under obfuscation

Under obfuscation, the profit writes:

$$\begin{aligned}\Pi^O &= p_1 + F_1(\tilde{\epsilon}_1)(\Pi_2^m - p_1) \\ &= p_1 + \left(\frac{p_1}{1+\delta} + \frac{\delta}{1+\delta} (\text{CS}_2(p_2^m) - r) \right) (\Pi_2^m - p_1)\end{aligned}$$

The first order condition yields:

$$p_1^* = \frac{1 + \delta + \Pi_2^m + \delta(\text{CS}_2(p_2^m) - r)}{2} = \frac{1}{2} + \frac{\Pi_2^m + \delta(\text{CS}_2(p_2^m) - r)}{2}$$

Obviously, $\tilde{p}_1^* < 1$, and the minimum value of \tilde{p}_1^* is obtained for $\delta = 0$. This minimum value, $(1 + \Pi_2^m)/2$ is higher than Π_2^m , as the second period monopoly profit cannot be higher than 1.

1.B Full resolution under revelation without commitment

This section proves that, without commitment, there exists a range of equilibrium prices satisfying the first order inequality condition.

Let's recall that the second period demand writes, for $p_2 > p_2^e$:

$$\begin{aligned}D_2(p_2, p_2^e, p_1) &= \int_{p_2}^1 F_1 \left[\tilde{p}_1 + \frac{\delta}{1+\delta} (\epsilon_2 - p_2^e) \right] f_2(\epsilon_2) d\epsilon_2 \\ &= (1 - p_2) \left(\tilde{p}_1 + \frac{\delta}{2(1+\delta)} (1 + p_2 - 2p_2^e) \right)\end{aligned}$$

Thus

$$\begin{aligned}\frac{p_2}{D_2(p_2, p_2^e, p_1)|_{p_2=p_2^e+}} &\geq \frac{1}{-\frac{\partial D_2(p_2, p_2^e, p_1)}{\partial p_2}|_{p_2=p_2^e+}} \Leftrightarrow \\ \frac{p_2}{(1 - p_2) \left(\tilde{p}_1 + \frac{\delta}{2(1+\delta)} (1 - p_2) \right)} &\geq \frac{1}{\tilde{p}_1} \Leftrightarrow \\ p_2 &\geq 1/2\end{aligned}$$

Therefore, all prices above the monopoly price $1/2$ are plausible expectations by the consumers.

Rewriting the profit under Revelation, for $r = 0$ and a fixed value p_2 gives the following first order condition:

$$2p_1 = (1 + \delta) + p_2(1 - p_2) - \frac{\delta}{2}(1 - p_2)^2, \text{ with } p_2 \in [1/2, 1]$$

Therefore:

$$p_1^*(p_2) = \frac{(1 + \delta)}{2} + \frac{p_2(1 - p_2)}{(2 + \delta)}(1 - p_2) - \frac{\delta}{4}(1 - p_2)^2$$

Thus, all profits generated by the $(p_1^*(p_2), p_2)$, $p_2 \in [1/2, 1]$ are possible profits.

1.C Profit comparisons under revelation or under obfuscation

This section proves that:

$$\Pi^{\mathcal{O}}|_{r=\text{CS}_2(p_2^n)} > \Pi^{\mathcal{R}}|_{r=0}.$$

by brute force.

The profit maximization under revelation with commitment, and $r = 0$ gives two first order conditions:

$$p_1 = \frac{(1 + \delta)}{2} + \frac{p_2(1 - p_2)}{(2 + \delta)}(1 - p_2) - \frac{\delta}{4}(1 - p_2)^2$$

$$\delta(p_2 - p_1)(1 - p_2) = \delta(1 - p_2)^2 + p_1(1 - 2p_2)$$

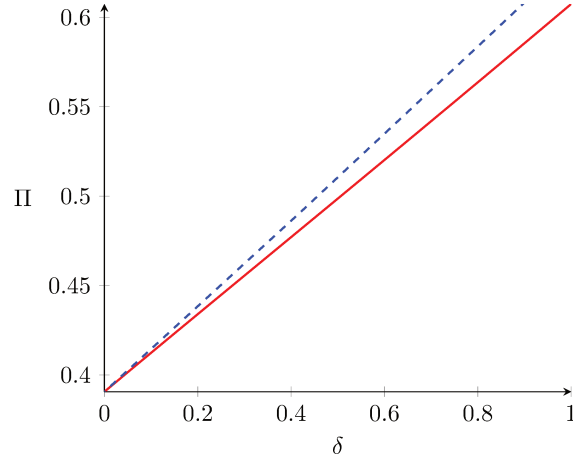
There are no tractable closed form solution in this simple case, yet numerical solutions are possible to compute. On the other hand, the solutions for the profit maximization under obfuscation with $r = \text{CS}_2(p_2^n)$ are easy to determine:

$$p_1^* = \frac{1}{8}(5 + 4\delta)$$

$$\Pi^* = \frac{(\frac{5}{8} + \frac{\delta}{2})^2}{(1 + \delta)}$$

Figure 1.8 provides an additional element: when consumer's time discount factor is null, both strate-

Figure 1.8: Profit comparison, under revelation (red-solid) or obfuscation (blue-dashed)



gies, obfuscation or revelation, are equally profitable.

Bibliography

- ARMSTRONG, M. AND J. ZHOU (2011): "Paying for prominence," *The Economic Journal*, 121, F368–F395.
- BUDISH, E. AND R. ZEITHAMMER (2011): "An efficiency ranking of markets aggregated from single-object auctions," .
- COASE, R. H. (1972): "Durability and monopoly," *JL & Econ.*, 15, 143.
- ELLISON, G. (2003): "A model of add-on pricing," Tech. rep., National Bureau of Economic Research.
- ELLISON, G. AND A. WOLITZKY (2012): "A search cost model of obfuscation," *The RAND Journal of Economics*, 43, 417–441.
- ESŐ, P. AND B. SZENTES (2007): "Optimal information disclosure in auctions and the handicap auction," *The Review of Economic Studies*, 74, 705–731.
- GABAIX, X., D. LAIBSON, ET AL. (2006): "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *The Quarterly Journal of Economics*, 121, 505–540.
- HELMERS, C., P. KRISHNAN, AND M. PATNAM (2015): "Attention and saliency in online markets," .
- MILGROM, P. AND R. J. WEBER (1982): "The value of information in a sealed-bid auction," *Journal of Mathematical Economics*, 10, 105–114.
- PETRIKAITĖ, V. (2018): "Consumer obfuscation by a multiproduct firm," *The RAND Journal of Economics*, 49, 206–223.
- SHIN, J. (2005): "The role of selling costs in signaling price image," *Journal of Marketing Research*, 42, 302–312.
- STAHL, D. O. (1989): "Oligopolistic pricing with sequential consumer search," *The American Economic Review*, 700–712.
- TAYLOR, G. (2014): "Browsing, Salesmanship, and Obfuscation," .
- WILSON, C. M. (2010): "Ordered search and equilibrium obfuscation," *International Journal of Industrial Organization*, 28, 496–506.

Chapter 2

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

Using an exhaustive database of bookings in one large chain of hotels active in Oslo (2013-2016), we estimate a nested-logit demand model that allows us to evaluate substitution patterns between online distribution channels. Making use of the chain's decision to delist from Expedia's platform, we compare the simulated and actual effects of such an event on prices and market shares and identify ways to improve on simulated counterfactual outcomes.

JEL Classification : D22, D43, L11, L81

Keywords: Multi-channel distribution, Pricing, Structural demand estimation, Online substitution.

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 vertical 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

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

²See for example recent cases in traditional retailing [e.g., [Picnic/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 \(European Commission, 2012\)](#)].

³See 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.

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

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 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

⁵See [Decision of 15 April 2015 by the Swedish Competition Authority in Case 596/2013](#).

⁶The 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.

⁷In 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.

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

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. \(2017\)](#) and [Larrieu \(2019b\)](#)), we use a large dataset of actual bookings to 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 one 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.

⁹See also [Edelman and Wright \(2015\)](#), [Wang and Wright \(2016\)](#) and [Wang and Wright \(2017\)](#) 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.

¹⁰For 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 \(2017\)](#).

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 Expedia 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 enough. We restrict attention

¹¹Our 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).

Table 2.1: Share of bookings made through the offline and online channels

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%
<i>Direct Online (DON)</i>	8,571	43.9%	11,275	49.2%	17,952	52.1%	17,910	44.0%
<i>Booking (BOO)</i>	10,957	56.1%	11,648	50.8%	13,419	39.0%	11,663	28.7%
<i>Expedia (EXP)</i>	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%

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 observations¹².

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 2.1, 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 channels¹³ as instruments in our demand estimation, we essentially focus on the 117,760 online bookings. Table 2.1 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 their rooms on Expedia (see section 2.2). In 2016, Expedia's market share was closer to Booking's market share.

Table 3.5 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 ~ 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

¹²We provide in Appendix – see Table 2.18 – a detailed breakdown of the elimination stages specifying the share of bookings and revenues for each stage of the process.

¹³Although 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.”

Table 2.2: Summary statistics of booking characteristics

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%

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 2.3: Summary statistics of hotel characteristics

	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 3.6 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 (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 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 2.4 shows the date (formally week) that we identify as the end of the boycott for each of the 13 hotels.

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 2.5 shows each channel's market share during the two periods. Note that we cannot infer from these numbers

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

¹⁵Formally, 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.

Table 2.4: Identifying the end of the boycott period

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

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.

Table 2.5: Online distribution market share for each channel

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

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:

$$u_{ijt} = \underbrace{X'_{jt}\beta - \alpha p_{jt} + \xi_{jt}}_{\equiv \delta_{jt}} + \zeta_{igt} + (1 - \sigma_g)\zeta_{igdt} + (1 - \sigma_g)(1 - \sigma_d)\varepsilon_{ijt}, \quad (2.1)$$

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 :

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \delta_{jt} + \sigma_g \ln\left(\frac{s_{dt}}{s_{gt}}\right) + \sigma_d \ln\left(\frac{s_{jt}}{s_{dt}}\right), \quad (2.2)$$

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 Oslo¹⁶, 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 or four-star hotels in Oslo, i.e., 70%. In Appendix 2.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 (2.2). 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.

¹⁶We 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>.

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

¹⁸These 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 2.6 below), about 85% of hotel guests in Norway come from Scandinavian countries.

Table 2.6: Nationality of the guests for hotels in Norway (in February 2015)

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%	

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 (2.2). 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 2.7 for specifications including different types of instruments.

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

¹⁹“XXX” corresponds to the name of the hotel chain.

Table 2.7: Demand model estimation

	(1)	(2)	(3)	(4)
α	-0.0028 (0.003)	-0.0017* (0.001)	0.0010 (0.001)	0.0009* (0.000)
σ_g	1.2873 (0.776)	0.9952*** (0.188)	0.3138 (0.166)	0.3429*** (0.098)
σ_d	1.8648 (1.706)	1.1073*** (0.149)	0.6039*** (0.138)	0.6260*** (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; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

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:

$$\varepsilon_{jk} = \alpha p_j s_j \left(1 - \frac{\mathbb{1}_{k=j}}{(1 - \sigma_d) s_j} + \frac{(\sigma_d - \sigma_g) \mathbb{1}_{k \in d}}{(1 - \sigma_d)(1 - \sigma_g) s_d} + \frac{\sigma_g \mathbb{1}_{k \in g}}{(1 - \sigma_g) s_g} \right) \quad (2.3)$$

Table 2.8 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.

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 2.9.

Own-price elasticities are (in absolute value) only slightly smaller than products’ (i.e., hotels \times distribution channel) own-price elasticities. They are also larger for OTAs (1.32 for Booking and

²⁰The results of the first-stage estimations are presented in Appendix 2.B.

Table 2.8: Elasticities estimates

Channel	ε_{jj}	ε_{jk}	ε_{jk}	ε_{jk}
Group		$g_j = g_k$	$g_j = g_k$	$g_j \neq g_k$
Nest		$d_j = d_k$	$d_j \neq d_k$	$d_j \neq d_k$
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 2.9: Aggregate elasticities estimates

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)

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 2.7.

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:

$$\mathbf{s}(\mathbf{p}) - \Theta \odot \nabla s(\mathbf{p}) \cdot (\mathbf{p} - \boldsymbol{\gamma}) = 0 \iff \boldsymbol{\gamma} = \mathbf{p} + (\Theta \odot \nabla s(\mathbf{p}))^{-1} \cdot \mathbf{s}(\mathbf{p}), \quad (2.4)$$

where $\mathbf{s}(\mathbf{p})$ represents the vector of market shares, \mathbf{p} and $\boldsymbol{\gamma}$ 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 matrix²² 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.

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:

$$(1 - \boldsymbol{\tau}) \cdot \mathbf{s}(\mathbf{p}) - \Theta \odot \nabla s(\mathbf{p}) \cdot ((1 - \boldsymbol{\tau}) \cdot \mathbf{p} - \boldsymbol{\gamma}) = 0, \quad (2.5)$$

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 .

Table 2.10: Average marginal cost per channel (in NOK)

Structure Channel	Price	No		Yes	
		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 2.10 reports the average marginal costs (and commission rates) derived, using our estimated demand parameter, from equations (2.4) and (2.5). We first observe that higher prices coincide with

²¹In 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.

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

²³Even 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.

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.

To test the robustness of our estimation²⁴, we derive the marginal costs using our system of first-order conditions given by equation (2.4), 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 2.11.

Table 2.11: Average marginal cost per channel (in NOK)

Selection Channel	No		Yes	
	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%

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

²⁴See also Appendix 2.D.

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

$$\tilde{s}(\tilde{\mathbf{p}}) - \tilde{\Theta} \odot \nabla \tilde{s}(\tilde{\mathbf{p}}) \cdot (\tilde{\mathbf{p}} - \tilde{\gamma}) = 0, \quad (2.6)$$

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 an 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 $\tilde{s}(\tilde{\mathbf{p}})$ (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} :

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

This diversion ration $DR_{d \rightarrow \hat{d}}$ 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 2.12.

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

Table 2.12: Estimated diversion ratios

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

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 2.1 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 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 2.13. 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

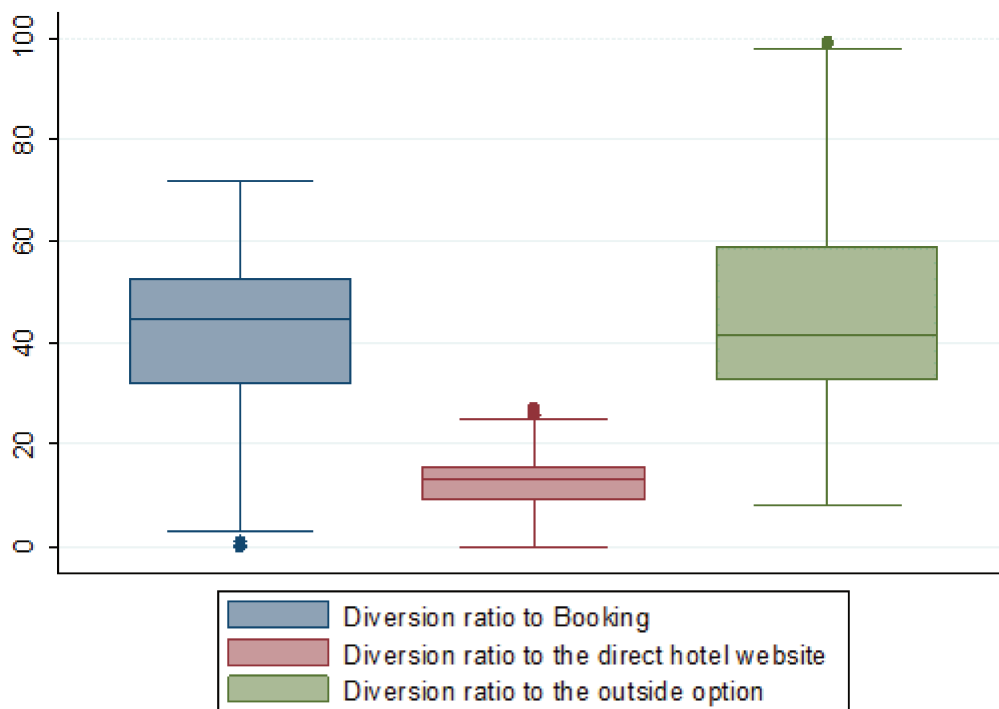


Figure 2.1: Box-plot of estimated diversion ratio (in %)

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.

Table 2.13: Welfare effects of delisting from Expedia (average weekly levels)

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

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 2.14. 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”).

Table 2.14: Observed and predicted prices

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%)

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 2.15. 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.

Table 2.15: Comparison on market shares

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

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 sug-

gest 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 2.16 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.

Table 2.16: Average booking characteristics during and after the boycott period

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%

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 2.17.²⁶

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.

²⁶We 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).

Table 2.17: Simulated and predicted outcomes

<i>Correction</i>		Counterfactual		Observed (whole period)
		<i>No</i>	<i>Yes</i>	
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%

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 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.

2.A Selected Observations

Table 2.18 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).

Table 2.18: Restrictions on bookings (and impact on revenue)

Restriction	Revenue	Evolution	Bookings	Evolution
-	2,91	-	1,235,106	-
Standard and superior rooms	2,16	-26%	1,116,204	-10%
≤ 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 ≤ 30 days	1,45	-19%	885,249	-17%
All		-50%		-28%

2.B First-stage Estimates

Table 2.19 presents the results of the first-stage estimations. Instruments globally enter with the expected sign.

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.

Table 2.19: First-stage estimation results for instruments variables

	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*** (37.995)	-0.76*** (0.176)	-0.00 (0.031)
Competitor last 5 grade	-2087.67** (787.869)	5.87 (3.451)	-0.40 (0.666)
Competitor grade	-8251.07** (2940.730)	18.31 (14.123)	0.85 (2.592)
Competitor comments	-3766.08*** (1011.974)	8.99 (4.832)	0.29 (0.882)
Cost	-1261.62*** (148.598)	7.33*** (0.715)	-0.03 (0.134)
Google Trend Norway	-77.53*** (11.881)	-0.01 (0.055)	0.30*** (0.010)
Google Trend Sweden	31.09* (13.130)	-0.04 (0.064)	0.02 (0.013)
Google Trend Denmark	14.03 (14.021)	0.04 (0.068)	-0.18*** (0.013)
N	1,923	1,923	1,923

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

2.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).

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 2.20 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 2.21 – remain almost unaffected. Results are thus robust to (reasonable) changes in the outside good's market share.

Table 2.20: Demand model estimation

Share of 3/4-star hotels	50%	60%	70%	80%	90%
α	0.0009* (0.00)	0.0009* (0.00)	0.0009* (0.00)	0.0009* (0.00)	0.0009* (0.00)
σ_g	0.3429*** (0.099)	0.3431*** (0.098)	0.3429*** (0.098)	0.3425*** (0.098)	0.3419*** (0.098)
σ_h	0.6252*** (0.090)	0.6259*** (0.090)	0.6260*** (0.090)	0.6257*** (0.090)	0.6252*** (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 2.21: Estimated diversion ratios

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

2.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 2.2 – 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.

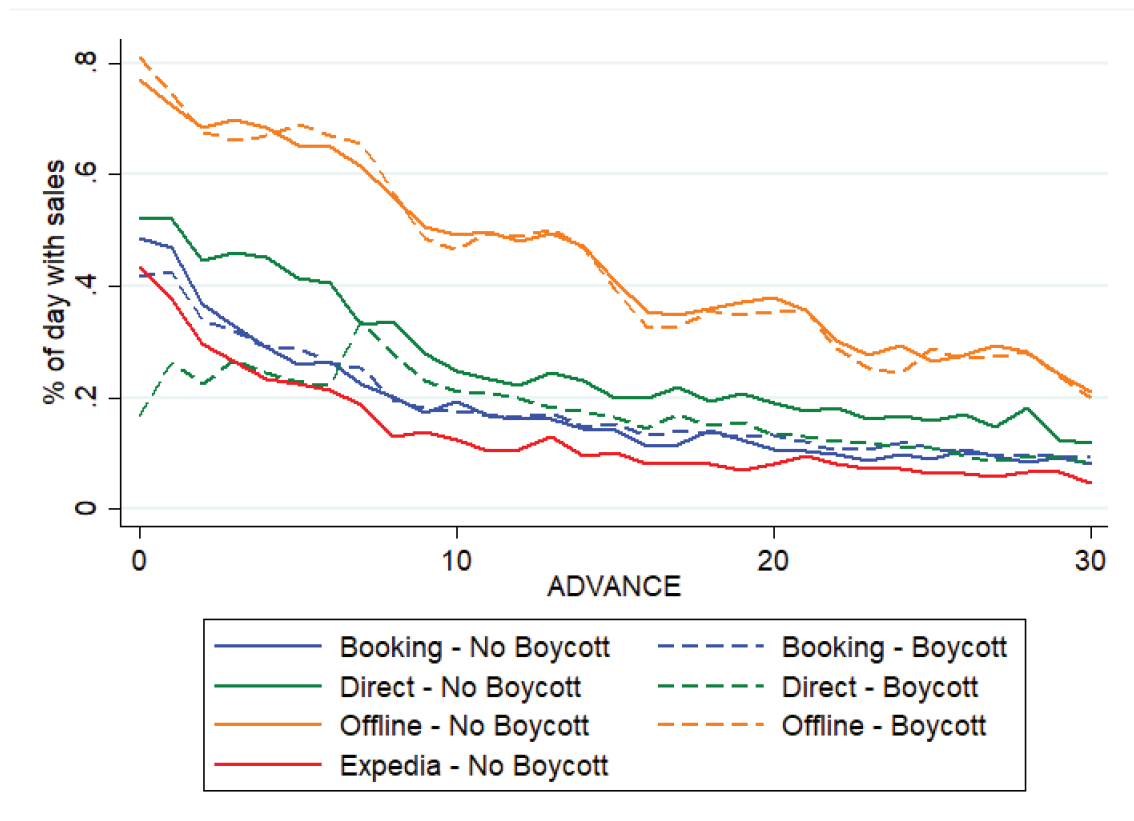


Figure 2.2: Proportion of days with sales during/after the boycott

Bibliography

- ANGRIST, J. AND J.-S. PISCHKE (2010): "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics," *Journal of Economic Perspectives*, 24, 3–30.
- BERRY, S. (1994): "Estimating Discrete-Choice Models with Product Differentiation," *Rand Journal of Economics*, 25, 242–262.
- BERRY, S., J. LEVINHSON, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–890.
- BJÖRNERSTEDT, J. AND F. VERBOVEN (2016): "Does Merger Simulation Work? Evidence from the Swedish Analgesics Market," *American Economic Journal: Applied Economics*, 8, 125–164.
- BOIK, A. AND K. CORTS (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, 457–82.
- DUCH-BROWN, N., L. GRZYBOWSKI, A. ROMAHN, AND F. VERBOVEN (2017): "The impact of online sales on consumers and firms. Evidence from consumer electronics," *International Journal of Industrial Organization*, 52, 30–62.
- EDELMAN, B. AND J. WRIGHT (2015): "Price Coherence and Excessive Intermediation," *Quarterly Journal of Economics*, 130, 1283–1328.
- ELLISON, G. AND S. FISHER ELLISON (2018): "Match Quality, Search, and the Internet Market for Used Books," *NBER Working Paper 24197*.
- FRIEDERISZICK, H. AND E. GLOWICKA (2016): "Competition policy in modern retail markets," *Journal of Antitrust Enforcement*, 4, 42–83.
- GENTZKOW, M. (2007): "Valuing New Goods in a Model with Complementarity: Online Newspapers," *American Economic Review*, 97.
- HAUSMAN, J. (1996): "Valuation of New Goods Under Perfect and Imperfect Competition," in *The Economics of New Goods*, ed. by T. Bresnahan and R. J. Gordon, National Bureau of Economic Research.

BIBLIOGRAPHY

- HUNOLD, M., R. KESSLER, U. LAITENBERGER, AND F. SCHLÜTTER (2018): "Evaluation of Best Price Clauses in Hotel Booking," *International Journal of Industrial Organization*, 61, 542–571.
- JOHANSEN, B. R. O. AND T. VERGÉ (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, 1151–1185.
- LARRIEU, T. (2019a): "Most Favoured Nation Clauses on the Online Booking Market," *mimeo*.
- (2019b): "Pricing strategies in online market places and Price Parity Agreements: evidence from the hotel industry," *mimeo*.
- LIEBER, E. AND C. SYVERSON (2012): "Online vs. Offline Competition," in *Oxford Handbook of the Digital Economy*, ed. by M. Peitz and J. Waldfogel, Oxford University Press.
- MANTOVANI, A., C. PIGA, AND C. REGGIANI (2017): "The Dynamics of Online Hotel Prices and the EU Booking.Com Case," *NET Institute Working Paper No. 17-04*.
- MILLER, N. AND M. WEINBERG (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, 307–342.
- NEVO, A. AND M. D. WHINSTON (2010): "Taking the Dogma out of Econometrics: Structural Modeling and Credible Inference," *Journal of Economic Perspectives*, 24, 69–82.
- PETERS, C. (2006): "Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry," *The Journal of Law & Economics*, 49, 627–649.
- POZZI, A. (2013): "The Effect of Internet Distribution on Brick-and-Mortar Sales," *The Rand Journal of Economics*, 44, 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, 139–156.
- REY, P. AND T. VERGÉ (2016): "Secret Contracting in Multilateral Relations," *Toulouse School of Economics Working Papers 16-744*.
- WANG, C. AND J. WRIGHT (2016): "Platform Investment and Price Parity clauses," *mimeo*.
- (2017): "Search Platforms: Showrooming and Price parity Clauses," *forthcoming, RAND*.
- WEINBERG, M. (2011): "More Evidence on the Performance of Merger Simulations," *American Economic Review*, 101, 51–55.
- WEINBERG, M. AND D. HOSKEN (2013): "Evidence on the Accuracy of Merger Simulations," *Review of Economics and Statistics*, 95, 1584–1600.

Chapter 3

Visibility, Commission fees & Parity: Evaluation of the Booking.com Preferred Partner Program

Using the reservations of 22 hotels located in three Scandinavian countries, we exploit a particular event that happened during the period covered by the databases : 6 hotels adopted the Preferred Partner Program offered by Booking while others did not. With this program, hotels basically pays a higher commission to the platform in exchange of a better visibility on the website. We show that the adoption of the program increases all online prices and increases the volume sold on platform but generates cannibalisation toward the direct channel. Using matching analysis, we show that the adoption of the program also coincides with the start in the wide price parity application on some offers. With a triple difference, we estimate the impact of the wide price parity application and show that it is associated to lower prices on Booking, questioning the general anti-competitive effect addressed to such clauses. **JEL Classification** : D22, D40, L11, L42, L81 **Keywords**: hotels, online

distribution, visibility, substitution, parity

1 Introduction

In 2016, the accommodation industry generated in Europe 88 billions euro of value added in EU¹. Yet, this sector, as many others, has known in the past decade drastic changes, notably through the emergence of intermediation platforms, whose main leaders are Priceline (Booking.com, Kayak, Agoda, ...) and Expedia (Expedia.com, Hotels.com, Hotwire.com,...).

¹[https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:T2a-Sectoral_analysis_of_key_indicators,_Accommodation_and_food_service_activities_\(NACE_Section_I\),_EU-28,_2016.png](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:T2a-Sectoral_analysis_of_key_indicators,_Accommodation_and_food_service_activities_(NACE_Section_I),_EU-28,_2016.png)

The accommodation industry has ambivalent relationships with these websites. On the one hand, they provide new consumers, and help lower the search costs incurred by the online reservation process. On the other hand, this new channel might cannibalise their own channels (online and offline); they lose the knowledge of their own consumers; the level of commission paid to the platform is significant; and the hotels become increasingly dependent of their ranking in the website search algorithm. This competition for exposure on the search engine is partially exploited by the platforms, through specific contracts, aimed at increasing the hotel's visibility.

In 2015, the leading online travel agency Booking.com introduced such a program, entitled the Preferred Partner Program. This Preferred Partner Program affects the hotels in multiple ways: they are more visible on the platform, but pay higher commission fees, and have the legal obligation to implement a wide parity clause. Using a dataset of reservations over more than two years in twenty-two hotels, we evaluate extensively the impact of this program on the prices and reservations on different online channels. We also provide an identification strategy to evaluate the relative magnitude of this wide parity clause on the price changes.

We show that the adoption of the program increases all online prices set by the hotel on Booking, Expedia and its own website. This price increase on platforms can be explained by a better visibility allowed by the program on Booking and Expedia since we observe a higher number of reservations on the platform. Simultaneously, the program on Booking partially cannibalize the direct channel of the hotel for which the number of reservations decreases. For the hotel website the price increase is then mainly due to the commission rise inferred by the program. Regarding the price parity, we first show that its analysis cannot be done without taking into account the type of contract between the platform and the hotel (here Preferred Partner or not) and the type of offer selected by the consumer (refundable or not). We show that wide price parity application is associated to either no effect or lower prices on Booking, questioning the general anti-competitive effect addressed to such clauses.

To our knowledge, there is no similar evaluation of the Preferred Partner Program in the previous literature. Most closely related is [Hunold et al. \(Forthcoming\)](#), which focuses on the direct influence of the ranking on Booking and its influence on the price set on other channels (other online travel agencies or own website). Yet, the structure of our analysis differ in multiple ways. First, we work with actual reservations and not scrapped posted prices. This allows us to analyse the impact of the preferred program on prices and volumes. Second, the Preferred Partner Program is a direct choice from the hotel, when the ranking of a hotel on Booking.com is partially a black box, even to the hotels. Lastly, we decompose the impact of the Preferred Partner Program on prices, showing that multiple effects are to take into account, including the application of a wide parity clause. From the consumer's perspective, [Ursu \(2018\)](#) and [Ghose et al. \(2014\)](#) analyse the impact of ranking on search and booking behaviour.

A second wide strand of empirical and theoretical literature related to our work is the on price parity clauses. On the theoretical side, there is no clear consensus. The very first papers on the subject [Boik and Corts \(2016\)](#) and [Johnson \(2017\)](#) show that price parity clauses lead to higher commissions and thus higher final prices. However, [Johansen and Vergé \(2016\)](#) and [Johansen and Vergé \(2017\)](#) relax some previous assumptions and prove that price parity clauses have an ambiguous effect on commissions, final prices, and suppliers' profits. In [Edelman and Wright \(2015\)](#), [Wang and Wright \(2016\)](#) and [Wang and Wright \(2017\)](#) platforms provide both search and intermediation

services. They find that parity clauses usually lead to higher prices or inefficient investment. On the empirical side, [Hunold et al. \(2018\)](#) was the very first paper to provide results about the effect price parity on price dispersion. Authors show that in countries without price parity (mainly France and Germany) the direct channel price is more often below the Booking.com price compared to country with wide price parity (for instance Canada). [Larrieu \(2019\)](#) then obtained some results on the absolute effect of prices, confirming the very first theory of harms. Indeed the end of price parity agreements causes a decrease of about 3.1% to 4.5% in the average level of prices posted by hotels in Paris on the meta-search Kayak. However a more recent paper suggests that the end of the price parity in European countries had no effect on prices. [Mantovani et al. \(2018\)](#) focus on decisions taken in France, Italy, and Spain and estimate the impact of the price parity clauses on the prices posted on Booking. Their analysis suggest that the investigations and decisions did not significantly reduce prices paid by consumers. Authors infer that it may be both linked to the demand increase (because of terrorist attacks in Tunisia, Turkey and Egypt) but also to Booking's innovative strategies in the adoption of structural changes that improved the quality of its services in 2016 (Booking Suite system, the Rate Manager, etc.). This paper somewhat confirms this result since it shows that indeed Booking's program offers display incentives in terms of pricing for hotels and may blur the real effect of EU decision. We also go one step further showing that if removing price parity has no positive effect, imposing it may generate lower consumer prices. We contribute to this strand of the literature in different ways. First, we show that the application of price parity depends grandly on the type of offer (refundable or not) and the hotels. Most of the previous literature neglects these elements, leading to potentially biased estimations. We also control for the endogeneity of prices, taking into account the observed occupancy rate, something impossible with scrapped-data.

In this paper we have a focus on the hotel industry generally subjected to revenue management methods techniques ([Gallego and Van Ryzin \(1994\)](#) , [Talluri and Van Ryzin \(2005\)](#)). A literature based on dynamic adjustment and demand uncertainty is currently developing in the case of the airline industry ([Williams \(2017\)](#)) and more recently for the hotel industry ([Cho et al. \(2018\)](#)). We rely on a more reduced form analysis, taking into account the fact that an hotel manager takes prices decisions that are correlated among channels. To this end we adapt our hedonic ([Court \(1939\)](#), [Lancaster \(1966\)](#), [Rosen \(1974\)](#)) price regression to this concern using a seemingly unrelated regression equation model introduced by [Zellner \(1962\)](#). As we face some issues of endogeneity we couple this method to the instrument variable one which gives rise to a three stage least square estimation ([Zellner and Theil \(1962\)](#)).

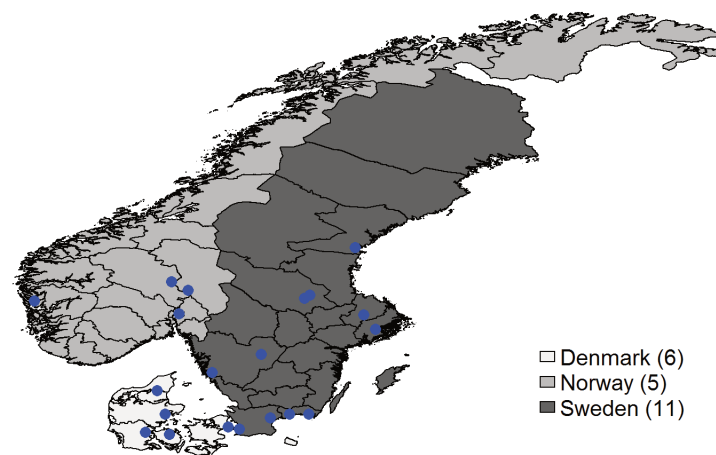
In the rest of the paper, we present the dataset and its main information, a distinctive but similar dataset used in [Cazaubiel et al. \(2020\)](#). Then, we explain the price paid by a consumer as a function of hotel characteristics and consumer choices using a hedonic model. In the third section, we estimate the effect on prices and reservation volumes of the adoption of the Preferred Partner Program by some hotels. In the fourth section we show that the adoption of the program coincides with the application of the wide price parity and we estimate its effect.

2 Data

2.1 Presentation

We use an exhaustive dataset of all bookings made over more than two years and a half in 22 hotels: 6 are located in Denmark, 5 in Norway and 11 in Sweden. These hotels all belong to one of the leading hotel chains active in Scandinavia.

Figure 3.1: Hotels' geographic localisation in Scandinavia



Our initial dataset includes around 824,048 observations (i.e., bookings). This dataset has been provided by the hotel chain and has been directly extracted from its information system. It includes all bookings made by consumers through all distribution channels between 27th February 2014 to 1st August 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 and the market codes associated with the booking.

2.2 Descriptive statistics

Rate and channel identification

Reservations are made through different channels and origins. From these two variables it is possible to precisely identify where does each consumer comes from for each reservation. Table 3.1 gives

market share distribution by channel. In 29% of cases consumers booked through an Online Travel Agency (OTA) which in 94% of cases is Booking or Expedia. In 10% of the reservations, consumers directly used the online hotel website. Therefore, overall around 39% of reservations are made online. Around 12% of bookings have been made through a Global Distribution System (GDS) which is a large computer network and a reservation tool used by travel agents that passes inventory and rates for airline seats, hotel rooms and car rentals, making the process of booking easier. The main GDS are Amadeus, Galileo, Sabre, Pegasus and Worldspan.

Table 3.1: Market share repartition by channel

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%

We observe that 50% of the consumers booked their room offline. For half of them it is through a special contract with their firm (B2B contracts), for instance a business trip. In that case, the firm directly books the room for his employee. A similar process occurs for offline travel agencies (24%) that are brick and mortar organisms that sell all-in-one packages to travellers. In that case the travel agency directly books offline the room for the client. These cases are rare, GDS remaining the main common tools for offline and online travel agencies to book their rooms. Finally, some consumers use more traditional ways for their reservations (29%). Table 3.2 shows that among them a majority (55%) sends emails or uses the telephone (37%).

Table 3.2: Offline devices used to book a room

Offline tools	# Observations	Share of bookings
Email	225,761	55%
Telephone	152,867	37%
Front Desk	29,424	7%
Fax	669	<1%
Letter	64	<1%
Total	408,785	100%

We also have an idea of the market code used by the hotel to price each reservation. Market

codes take four values classifying reservations into four groups: *public rates* for individual travellers purchasing unqualified or non-corporate negotiated rates, *contracted rates* for individual travellers purchasing contracted negotiated rates, *group leisure rates* for grouped travellers during leisure related travel and *group business rates* for grouped travellers during business related travel.

We decompose the previous channels on these market codes in Table 3.3. A first important feature is that the totality of group reservations are made offline. It seems that for these hotels, there is no possibility to book for a group reservation online or through GDS. As expected, 88% of group business reservations are made via B2B contracts offline and 73% of group leisure rates are made via travel agents contracts offline. Individual negotiated rates mainly come from offline (mainly B2B contracts) and GDS whereas individual public rates mainly come from OTAs and offline (mainly individuals).

Table 3.3: Proportion of channels by market codes

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

We then look at the proportion of these market codes on the channels in Table 3.4. Online bookings share a common feature: the majority (99% and 86%) of bookings are individual public market codes. This is clearly in contrast with GDS on which 78% are individual negotiated rates. In the case of offline reservations, proportions of market codes differ depending on the type of contracts. Individual offline reservations are mostly made of public rates (70%) and fewer negotiated rates (25%). In comparison, B2B contracts are mostly made of negotiated rates, either individual (42%) or for groups (41%). Finally offline travel agent contracts are made of a mix of market codes but fewer business rates for groups (8%).

These two tables give important conclusion regarding the market segmentation. We typically observe an online booking market made of OTA and hotel website channels associated to individual public market code pricing. In opposition, the offline channel is made of a mix of market codes while the GDS channel is dominated by negotiated tariff for non-group travellers. This suggest that in the pricing policy hosted by the chain, the online market on the one hand and the GDS market on the other hand respond to different origins and do not necessarily adjust simultaneously. For instance, the individual public rate is generally based on the Best Available Rate (BAR) that varies everyday while GDS rates are associated to an allotted number of rooms with special predetermined fares that

Table 3.4: Proportion of market codes by channels

Market codes by channels	Individual Public	Individual Negotiated	Group B2B	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 system	22%	78%	-	-	100%
Hotel website	86%	13%	-	-	100%

do not directly adjust to the public available one. In that sense the prices of these two sub-market can be correlated but not perfectly.

Other variables

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 the 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.

Descriptive statistics show that 70% of bookings have been made less than 25 days before the arrival and 100% have been made 6 months before the arrival. For this reason, we compute these 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). $OR_{h,t,0}$ 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.

Statistics on the consumers' choices and occupancy rate at the time of the reservation are summarised by channel in Table 3.5.

As expected, prices of reservations made online are bit higher compared to the other reservation channels since the tariff structure is not the same : negotiated individual tariffs (expected to be lower compared to public ones) are in only 6% associated to online channels while group negotiated tariffs are never associated to them. Note that the average price on OTA and on the hotel website are very close, around 135\$. We also observe that consumers tend to book earlier when they use an OTA and more often for a refundable contract than on the hotel website.

2.3 Additional data

We also collected hotel characteristics and this additional data include:

Table 3.5: Summary statistics of booking characteristics

Channel	OTA	Hotel website	GDS	Offline	All
Price (in NOK ¹)	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 ²	1.7	1.6	1.1	1.4	1.5
Standard room	77%	73%	80%	83%	80%
Week-end ³	42%	39%	5%	26%	29%
Occupancy rate	43%	41%	47%	48%	46%

¹ Price of bookings for hotels in Sweden and Denmark have been converted using the daily exchange rate

² Adults and children

³ Week-end represents the proportion of stays that start on Friday or Saturday

- 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 have been scrapped from TripAdvisor. For our 22 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 up to the date.

Table 3.6 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, all above 60 rooms). Hotels are located very close to the city center and pretty far from the closest important airport.²

Table 3.6: Summary statistics of hotel characteristics

	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

²We define as important, an airport with at least 10 different destinations, which means for Scandinavia a list of 19 airports. In our case, a list of 9 airports : Aalborg (AAL), Bergen (BGO), Billund (BLL), Copenhagen (CPH), Göteborg (GSE), Malmö (MMX), Oslo (OSL), Stocholm (ARN and BMA) is sufficient to cover all the cities of our hotels.

2.4 Data quality in price explanation

In the following of the paper we will focus on the price determinants. However at this stage it is not sure that the database contains all the needed information to explain a price difference between two reservations. In this subsection we test this hypothesis following a simple procedure : we identify identical reservations on all the characteristics of the database (our variables) and then compare their prices.

Our sample of 824,048 distinct bookings can be decomposed as 501,639 unique reservations, i.e. booking types that are so atypical in the mix of variables that they have been made only once by a consumer, and 94,452 distinct booking types that we recover several times, representing 322,409 different consumers. Among these 94,452 distinct booking types, 63% are observed twice, 16% are observed three times, 7% are observed four times, 4% are observed five times while 11% identify booking types observed six times or more (up to 94 times).

In Appendix 3.A we show that this matching operation do not generate any bias of selection i.e. looking at *matched* (booking types observed several times) and *not matched* (single booking) samples separately, the distribution along the main variables do not change in important magnitude. We can therefore analyse the matched observations without hiding a bias that could have been created through the selection process.

In a second step, we restrict to reservations identical according to all available variables, except for the price. We observe that in 98% of cases prices are also exactly equal. If they are not, we compute a ratio defined as the mean absolute difference between prices divided by the average price. The ratio is in average equal to 34%. These observations are pretty stable depending on the number of occurrences for the same booking type (Table 3.7).

Table 3.7: Pairwise price comparison

# Occurences	# Booking types	Proportion	% Equality	Ratio if not equal
2	58,654	63%	97.2%	18.3 %
3	14,911	16%	97.4%	20.8%
4	6,712	7%	98.0%	26.3%
5	3,555	4%	98.0%	24.5%
6+	10,620	11%	98.1%	38.0%
All	94,452	100%	98.1%	34.3%

In a less restrictive way, i.e. not looking at the share of strict equality but at a price difference lower than 200 NOK (approximately 20\$) results remain almost identical with a "same" price in 98% of cases. This indicates that prices differences, if they exist, tend to be relatively high. One explanation could be linked to some add-ons or special requests that exist in the hotel but are not identified in the database with a specific code (for instance a bottle of Champagne, or massages if there is a well-being center). The price is then not included in the room price but finally paid by the consumer.

Overall this analysis suggests that in the database, reservations with the same characteristics generally have the same price - except in 2% of cases. Hence, we do not miss any important

information that can determine prices and then should be taken into account in a price determinants analysis.

3 Price analysis

So far, the hedonic pricing literature (Court (1939)) has already been used to quantify the weights of different tangible characteristics in the prices of bed and breakfast rooms (Monty and Skidmore (2003)) but also in the hotel industry (Abrate and Viglia (2016) in Milan and Thrane (2007) in Oslo). However, very few authors tried to adapt such methods to the evolution of the industry and the growing use of revenue management techniques. If Law et al. (2011) add to the models the impact of customer reviews in New York, none of them take into account variables such as the channel used by the consumers (platform or not) and/or the role of the occupancy rate in the pricing decision.

In this section, following this literature, we run a simple hedonic price regression augmented by non tangible reservation characteristics.

3.1 Identification strategy

For each reservation i made for a resort located in country c at the insert date t , we estimate:

$$\ln \text{Price}_{ict} = \alpha_{0c} + \beta_{0c} X_{ict} + \beta_{1c} \mathbb{1}_{\text{Pref.}} + \gamma_c \mathbb{1}_{\text{NR}} + \varepsilon_{ict} \quad (3.1)$$

With X_{ict} taking into account :

- Seasonality : reservation date time trend, month of arrival dummies and day of reservation among the week dummies;
- Resort characteristics : number of stars, distance to the city center and the nearest international airport, score (between 0 and 3) taking into account the existence of amenities (restaurant, bar and/or well being-center), and a popularity score of Tripadvisor defined as a re-scaled product of the number of comments and average grades over ten.
- Reservation characteristics : arrival during the week or the week-end, type of room, number of persons in the room, number of nights and number of days between the reservation and the arrival and the channel used to make the reservation (lead time);
- Occupancy rate : we believe that hotel keepers take two variables into consideration in order to adjust their prices. They first observe the occupancy rate at the moment they set the price (occupancy rate at the insert date for a given arrival date) but also form belief on the realised value of the occupancy rate the final day, i.e. at the arrival day. These two values are not necessarily equal : at the moment a consumer is booking the occupancy rate can be low but the manager can expect the hotel to be finally full. In order to take into account this variety of indicators, we include in our explaining variables the occupancy rate at the reservation date and the realised occupancy rate at the arrival date as a proxy of the manager anticipations. We therefore infer that managers form rational expectations.

We add two important variables contained in our database that have never been taken into account in any paper about the hotel industry : the type of offer chosen by the consumer and the type of contract between the hotel and the platform.

During its reservation, the consumer has to make different choices : this may include the choice of an included breakfast or not, beds configuration (twin or join), the payment process (direct or delayed physically to the hotel) and whether the price of the room can be reimbursed or not. We call this last option, the type of offer associated to the contract. In our data we clearly identify two types of offer (refundable or not³) if the reservation is made online i.e. through OTAs and the hotel website. We take into account this information in with γ_c being the coefficient associated to making a reservation for a non refundable offer.

Furthermore, if all the hotels listed on Booking seem to be pretty similar for consumers point of view they differ in the contract they signed with the platform. Some of them may decide to take part in a special type of contract called "Preferred Partners Program" meaning at this stage a higher commission fee against a better visibility. We will come back to the precise definition of the program in Section 4). We take into account this information with β_1 being the coefficient associated to making a reservation for a member of the Preferred Partner Program.

In the estimation, we specify the standard errors allowing for intra-group correlation by hotel and reservation dates. We therefore infer that the observations are independent across hotels and day of reservation but not necessarily within these two characteristics. This affects the standard errors and variance–co-variance matrix of the estimators but not the estimated coefficients.

3.2 Instruments

The two occupancy rates variables contained in X_{ict} may be endogenous. In fact the occupancy rate is a proxy for the demand of the hotel and as prices and demand both influence each other, they are then determined simultaneously. In order to take into account this endogeneity, we use as instruments the corresponding occupancy rate for GDS. As shown previously in Table 3.4, reservations made through GDS are mainly made of Individual Negotiated Rates (78%) while on the hotel website and OTAs, the rates are, in very high respective proportions (86% and 99%), Individual Public Rates. The interpretation suggested before is that GDS and OTAs sales belongs to independent markets in the sense that their fares do not simultaneously adjust. However, consumers booking through GDS use the same room as others and then contribute to the global occupancy rate of the hotel.

However this exogeneity condition only works for online bookings (OTAs or hotel website) in comparison to GDS. In fact, offline reservations have a wide variety of rates. Especially, 31% of the offline reservations are made of individual negotiated reservations and then the instrument is not valid. For this reason, in the following we restrict the analysis to online bookings.

We then instrument the hotel global occupancy rate at the moment of the reservation for a given arrival date (OR_{TOT}) and the realised value of the occupancy rate the final arrival day (OR_{FINAL}) by their respective counterparts on GDS only $OR_{TOT_{GDS}}$ and $OR_{FINAL_{GDS}}$. The coefficients associated

³Two types of offers are frequent in the accommodation industry : refundable or not-refundable. In the case of refundable offers, consumers may book the room and decide to cancel it with full reimbursement (up to 1 or a few days before the admission date). We observe that 52% of bookings made through OTAs are for refundable offers while they represent only 28% through the hotel website.

to these excluded instruments in the first stage regression for the two types of occupancy rates are displayed in Table 3.8.

Table 3.8: First stages estimations

	Sweden		Denmark		Norway	
	OR_{TOT}	OR_{FINAL}	OR_{TOT}	OR_{FINAL}	OR_{TOT}	OR_{FINAL}
$OR_{TOT_{GDS}}$	2.02*** (0.037)	0.025 (0.036)	1.72*** (0.029)	0.034 (0.021)	2.02*** (0.040)	0.23*** (0.031)
$OR_{FINAL_{GDS}}$	-0.43*** (0.027)	1.27*** (0.027)	-0.44*** (0.020)	1.04*** (0.017)	-0.51*** (0.022)	1.04*** (0.023)
Lead time	-0.0071*** (0.00013)	0.0017*** (0.00012)	-0.0078*** (0.00017)	0.0013*** (0.00010)	-0.0081*** (0.000095)	0.0013*** (0.000098)
F-Stat	2,392	3,047	1,270	1,545	1,691	1,775
APF-Stat	3,554	5,495	2,804	2,953	2,980	3,135
N	47,978	47,978	104,834	104,834	73,256	73,256

Clustered standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We first notice that each occupancy rate is positively and significantly correlated with its value on the sub-segment GDS. This is totally consistent since the GDS market is a share of the whole market. In the same idea, the occupancy rate of GDS at the insert date day is positively correlated with the final global occupancy rate. This is not surprising since the occupancy rate of GDS at the insert date is a sub part of the global occupancy rate which is increasing as the arrival comes and then equal to the final occupancy rate the day of the arrival. However note that the coefficient is smaller.

It is very interesting to look at the second cross effect which is negative i.e a low occupancy rate the date of the reservation is associated to a high occupancy rate the day of the arrival on the GDS segment.

An interpretation can be given using the coefficients associated to the difference between the day of the arrival and the reservation : the lead time.

On the one hand, the final occupancy rate is positively correlated with the average number of days between the reservation and the arrival i.e. the more people book in advance, the fuller the hotel the day of the arrival. This is consistent with the fact that consumers anticipates that some days are fuller than others and then book more in advance.

On the other hand, in average when people book earlier (important lead time) the occupancy rate at the date of reservations is lower. This is consistent with the fact that when a reservation is made far in advance, then nobody has already booked and the hotel at that time is empty.

Therefore when the day of the arrival there is very few rooms sold through GDS (the final occupancy rate on GDS is low) it is because the average occupancy rate at the time of the reservation was high : people anticipated that it was a high demand day and booked more in advance. This reduced the number of rooms allocated to the GDS segment. In the other way, when the occupancy rate at the time of the reservation is in average low, it is because few people booked and then a larger amount of rooms is sold through GDS which increases its occupancy rate the day of the arrival.

3.3 Results

Using the instruments introduced in the previous section in order to correct for the endogeneity of the occupancy rate variables, we estimate the hedonic price equation (3.1). The estimated coefficients for our main variables of interest for the reservation characteristics are displayed in Table 3.9 and the estimated coefficients for our main variables of interest for the hotel characteristics are displayed in Table 3.10.

Table 3.9: Reservation characteristics

	Ordinary Least Squares			Instrumented Variables		
	Sweden	Denmark	Norway	Sweden	Denmark	Norway
Weekend trip	-0.16*** (0.0041)	-0.026*** (0.0026)	-0.069*** (0.0037)	-0.20*** (0.0044)	-0.048*** (0.0031)	-0.11*** (0.0058)
Superior room	0.18*** (0.0024)	0.19*** (0.0028)	0.23*** (0.0036)	0.17*** (0.0025)	0.18*** (0.0029)	0.21*** (0.0041)
Persons	0.087*** (0.0020)	0.074*** (0.0019)	0.12*** (0.0019)	0.093*** (0.0021)	0.085*** (0.0021)	0.14*** (0.0022)
Nights	-0.0040*** (0.0011)	-0.00020 (0.00094)	0.0034** (0.0015)	0.00022 (0.0013)	0.0047*** (0.0010)	0.0084*** (0.0015)
Channel - reference : hotel website						
Booking	-0.027*** (0.0023)	0.011*** (0.0027)	-0.010*** (0.0034)	-0.031*** (0.0025)	0.0058** (0.0028)	-0.020*** (0.0037)
Expedia	-0.020*** (0.0028)	-0.064*** (0.0029)	-0.063*** (0.0039)	-0.027*** (0.0029)	-0.066*** (0.0030)	-0.075*** (0.0042)
Non-refundable offer	-0.13*** (0.0028)	-0.14*** (0.0023)	-0.16*** (0.0034)	-0.13*** (0.0031)	-0.13*** (0.0024)	-0.15*** (0.0035)
OR _{TOT}	0.26*** (0.022)	0.25*** (0.0096)	0.15*** (0.013)	0.14*** (0.019)	0.39*** (0.020)	0.11*** (0.031)
OR _{FINAL}	0.13*** (0.014)	0.16*** (0.0078)	0.20*** (0.010)	0.59*** (0.018)	0.37*** (0.019)	0.69*** (0.027)
Lead time	0.0030***	0.00076***	0.0011***	0.0012***	0.0020***	0.000014
<i>N</i>	104834	73256	47978	104834	73256	47978

Clustered Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We show that people arriving on Friday or Saturday (week end trip) pay between 5 and 20% less in average. This is pretty consistent with the fact that the hotel chain we are looking at is used in a non negligible proportion for B2B stays that generally occur during the week. The hotel is less full during the week end and offers lower prices. Consumers who chose a superior room pay a higher price (between 17% and 21% more) and adding one person (children or adult) to the reservation increases the price by 8.5% to 14% which is equivalent to the price of a breakfast in an hotel of this category. The coefficient associated to the number of nights is significant but very low suggesting that the duration of the stay has a negligible impact on the price paid per night. Looking at the channel of reservation, booking through an OTA is in average cheaper in the three countries for Expedia while for Booking it is cheaper in Norway and Sweden. In the case of Denmark it seems to be more expensive to book on Booking but note that the coefficient is very small (around 0.5%). Unsurprisingly, non refundable offers are in average 13% to 15% cheaper than refundable ones since they do not offer as much flexibility for consumers who face uncertainty on their trip. A last interesting point is the relation between the price and the occupancy rates. The price paid by consumers seems to be more correlated to the expected final occupancy rate than to the occupancy rate the day of the reservation

which gives an idea on the relative performance of the two indicators for the manager. This is less true for Denmark where the two indicators seem to have an identical weight. Finally, once we control for the effect of the occupancy rate, buying in advance has a negligible effect on the price paid by a consumer.

Table 3.10: Hotel characteristics

	Ordinary Least Squares			Instrumented variables		
	Sweden	Denmark	Norway	Sweden	Denmark	Norway
Distance /airport	-0.00045*** (0.000034)	-0.0023*** (0.00045)	-0.0059*** (0.00020)	-0.00039*** (0.000038)	-0.0029*** (0.00046)	-0.0068*** (0.00021)
Distance /center	-0.018*** (0.0047)	-0.040* (0.019)	-0.092*** (0.0097)	-0.016** (0.0052)	-0.040* (0.020)	-0.070*** (0.0099)
Amenity score	0.15*** (0.0089)	0.042** (0.015)	0.17*** (0.0033)	0.18*** (0.0096)	0.082*** (0.016)	0.19*** (0.0038)
Stars	0.10*** (0.0074)	0.23*** (0.028)		0.10*** (0.0078)	0.26*** (0.028)	
Popularity score	0.086*** (0.0019)	0.0087 (0.0051)	-0.0052 (0.0078)	0.069*** (0.0023)	0.025*** (0.0053)	0.015* (0.0076)
Preferred Partner	-0.27*** (0.0056)	-0.021 (0.024)	-0.48*** (0.018)	-0.30*** (0.0057)	-0.089*** (0.024)	-0.70*** (0.021)
<i>N</i>	104834	73256	47978	104834	73256	47978

Clustered standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In terms of hotel characteristics, we observe that consumers are sensitive to the location, especially to the distance to the city-center compared to the distance to the airport : the farther from the city center the lower the price. Having special amenities is positively valued by consumers, from 8% to 19% for one additional amenity among (bar, restaurant and well being center). As expected having one more star (4 compared to 3) is correlated to a higher price. Regarding the reputation on TripAdvisor, we observe that a higher grades together with higher number of comments, giving more credibility, helps the hotel to set higher prices.

Interestingly, hotels which are involved in the Preferred Partner Program seem to have in average lower prices, between 9% less in Denmark to 70% less in Norway. This result is not really intuitive and we are pretty agnostic on the value expected. Indeed, on the one hand preferred hotels pay a higher commissions and are then expected to pass-through this additional cost into higher prices. On the other hand the program gives more visibility to the hotel which could be anticipated and converted in to higher or lower prices. Furthermore, this huge variability by country has to be nuanced. In fact, this program is adopted by some hotels - not all and not in the same proportion by country- during the period.

For these different reasons, we go through more details on the effects on this program on prices and volumes in the next section.

4 On the effect of the Preferred Partner Program

Since 2015 and the majority of ban on price parity in Europe, Booking.com reacted offering different types of contracts to hotels. Hotels can either be registered as standard hotels or be listed as preferred partners. This program works in two steps: hotels are first invited to be member and then voluntarily

decide to enter it or not.

In order to be invited to the Preferred Establishment Program, the hotel must satisfy eligibility criteria such as an average score of customer reviews equal to or greater than 7 out of 10 and a systematic and full payment of the commission within the deadlines.

Hotels that have accepted the invitation must pay a special commission percentage for all bookings made from the date of their registration in the Preferred Establishments Program. According to our information in Europe for instance, while a standard hotel pays a commission rate around 15%, an hotel part of Preferred Establishments Program has to comply with a 2 or 3% higher commission rate depending on the country.

Note that at least in France, there is a second requirement since hotels have to respect the *Parity of Rates and Conditions* with Booking.com, meaning “equal or greater rates advantageous for the same accommodation, services and equipment of equal or 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).⁴”. This is not written-proof in Scandinavian countries but we will show in Section 5 that we have good reasons to think that similar requirements apply.

In return according to Booking.com⁵, Preferred Partners, reported with a yellow thumbs up (Figure 3.4 in Appendix 3.B) on the webpage, benefit from increased visibility in the search results. The platform evaluates that preferred partners enjoy an average of 65% more visitors on their page and get 35% more bookings.

Note that for hotels being a member of this program is a long term choice in the sense hotels do not enter and exit the program everyday. For this purpose, other programs are available such like the *Visibility Booster* that allows more flexibility with a temporary boost on selected targets (consumers or time).

According to the hotel chain, among the 22 hotels we are studying, 6 hotels entered in the Preferred Establishments Program at different points in time going from 16th February 2015 to 2nd June 2015. They are located in the three countries : 2 in Denmark (over 6 hotels), 3 in Norway (over 5 hotels) and 1 in Sweden over 11. Each of these 6 hotels enters the program at an observed date in the data (Table 3.17 in Appendix 3.C) and stays into it until the end of the period. These hotels have been specifically targeted by the chain because they were located in a highly competitive location. In fact, preferred hotels are all located in cities for which the number of accommodations referenced on Booking is among the highest.

According to Booking’s website the adoption of the preferred program, thanks to a gain of visibility increases the number of bookings by 35%. On the other hand becoming a preferred hotel implies new constraints on prices because of the commission and the potential price parity compliance. In the following, after defining the volume sold by an hotel, we estimate the impact of the Preferred Partner Program adoption on hotels prices and sales.

⁴Conditions of use of the Preferred Establishment Program, available here : <https://www.artiref.com/wp-content/uploads/2015/10/hotel-preferred-booking.pdf>

⁵Booking.com

4.1 Volume definition

We define the demand by channel as the sum of reservations at a given insert date booked through a given channel. For instance, three different consumers who booked the same day for different arrival dates, will count for a demand of three this day. A different approach could be to define the demand as the sum of occupied rooms at a given arrival date. However this would prevent to precisely identify the switch in the program adoption. Indeed there are reservations starting before and ending after the program adoption, but some have been made before the adoption and start after.

If in equation (3.3) the general idea was to regress the price on the occupancy rates, here we want to regress the volume on the price and then still face an endogeneity problem. In order to solve it, we use two instruments :

- the monthly average wage by country;
- the average price by day for reservations made for the other hotels in the same country but not the same city, on the same channel and for identical types of offer. We consider here 22 hotels of the database.

These two instruments are assumed to reflect the same idea that hotels in a same country share similar functioning costs that are correlated to prices. However we restrict to hotels of a same country but not located in the same city then they do not compete for the same consumers, which makes the exogenous condition to hold.

The variable contained in X_{jct} are therefore exactly the same for the price and volume analysis except in the instrumental variables.

4.2 Identification strategy

We use the adoption of the preferred program by some hotels as a treatment effect. As explained by the hotel chain, this treatment is not independent and prevent us to use a simple difference in difference analysis between preferred and standard hotels as if being a preferred hotel was totally exogenous. The hotel that makes part of the program have been both selected by Booking and the hotel chain for particular reasons.

Our solution is to restrict our control group to some specific non preferred hotels. Under the conditional independence assumption, i.e. that apart from the observable there are no other characteristics that influence both potential prices (respectively volumes) and the choice of treatment (being preferred or not), the estimation is not biased.

First, as all our hotels belong to the same chain, we expect similar behaviour in terms of delay of payment to the platform and then in terms of invitation to become member of the preferred program or not. Second, in our data, we observe that the treatment effect is correlated with the number of competitors on Booking (Table 3.17) : the treatment group is composed by preferred hotels with more than 108 competitors on Booking (i.e. hotels 3, 4, 7, 10, 11 and 14). In order to restrict our control group to similar hotels (i.e. with a high number of competitors on Booking), we select non preferred hotels with more than 35 competitors on Booking (i.e. hotels 2, 6, 9, 18, 20 and 22). We take the threshold of 35 competitors in order to keep the three countries in the analysis. We therefore

voluntarily select 6 hotels among the 16 available that are not members of the Preferred Partner Program and then drop 62,675 observations among the 226,068 available.

4.3 Common Trend Assumption

From the previous selection, we are finally left with three treated hotels and one control hotel in Norway, one treated hotel and three control hotels in Sweden and two control hotels in Denmark.

We formally want to test the equality of the prices (respectively volumes) slopes depending on time t of the two groups before the treatment in the three countries c . As we expect the effect of the adoption of program to differ between channels, we use the thinnest time aggregating price observations by insert day t , country c and channel j . The null hypothesis tested is the equality of the coefficients in the two subgroups : treated hotels vs non-treated hotels. In practice, we run the following regression on the data before the treatment for each product :

$$\ln Y_{jct} = \delta_{0jc} + \delta_{1jc}X_{jct} + \delta_{2jc}\mathbb{1}_{\text{Pref.}} + \delta_{3jc}t + \delta_{4jc}\mathbb{1}_{\text{Pref.}}t + \epsilon_{jct} \quad (3.2)$$

With $Y = \{\text{Price; Volume}\}$.

We then test the hypothesis for Y that $H_0 : \delta_{4cj} = 0$.

Compared to the simple price regression with endogeneity computed in the previous section, here our objective is to distinguish the effects by country and channel, running independent regressions. However, the prices set on different channels by a given hotel are not totally independent since they come from the same maximisation program. We therefore take into account this hypothesis, assuming that the error terms of the three channel equation of a same country can be correlated. This boils down to estimate seemingly unrelated regressions equations (SURE) introduced by Zellner (1962) in the case of ordinary least squares.

From that point, as explained in the previous section, we take into account endogeneity linked to occupancy rate variables and then estimate SURE with a two-stage least squares and not an ordinary least square. Doing so requires a three-step procedure. First, we regress each endogenous variable on all exogenous variables in the equations and we calculate predicted values for the endogenous variables. Then, we estimate the structural system of equations by least squares, replacing the endogenous right-hand-side variables with their predicted values from the first step. Finally, we compute the estimated variances and co-variances of the residuals from the second step, and re-estimate the structural equations using the SURE method. We therefore perform a three-stage least squares (Zellner and Theil (1962)) that corrects for any cross-equation error correlations.

From two-stage least squares to three-stage least squares, we obtained consistent estimates of the parameters that are more efficient but in general the value of estimated coefficients does not vary in huge proportions. However, an important difference is that with three stage least squares we can perform a joint test of cross-equation parameters. This is particularly useful in our case since our specification takes the form of difference in difference with related equations by country. Instead of testing independent common trend assumption by couple of country and channel, we therefore need to run equation (3.2) by channel and country and then to test by country c :

$$H_0 : \delta_{4c\text{Booking}} = \delta_{4c\text{Hotel website}} = \delta_{4c\text{Expedia}} = 0$$

Adding some additional restrictions on the selected hotels⁶ (see Appendix 3.C - Table 3.17), we do not reject the null hypothesis at the 1% level for each couple of country and channel for our both variables of interest : prices and volumes (Table 3.11).

Table 3.11: H_0 p -value test by country and channel

Y Country	Prices		Volumes	
	Chi-2 Stat	p -value	Chi-2 Stat	p -value
Norway	2.099	0,5522	5.078	0,1662
Sweden	3.146	0,3696	5.449	0,1417
Denmark	2.229	0,5264	2.191	0,5338

4.4 Results

The common trend assumption satisfied, we estimate a difference in difference model by reservations made on channel j at insert date t for a resort located in country c using the form of Equation (3.1) :

$$\ln Y_{jct} = \alpha_{0jc} + \beta_{0jc}X_{jct} + \beta_{1jc}\mathbb{1}_{\text{Pref.}} + \gamma_{jc}\mathbb{1}_{\text{NR}} + \beta_{2jc}\mathbb{1}_{\text{After}} + \beta_{3jc}\mathbb{1}_{\text{After}\times\text{Pref.}} + \varepsilon_{jct} \quad (3.3)$$

With $Y = \{\text{Price; Volume}\}$.

OLS, SURE, 2SLS and 3SLS β 's coefficients are displayed in Appendix Section 3.D, notably Table 3.18 for prices and Table 3.19 for volumes. We only display the 3SLS results below in Table 3.12 for prices and Table 3.13 for volumes.

Table 3.12: Difference in Difference results on prices

Country c Channel j	Sweden ($N=1,206$)			Denmark ($N= 2,462$)			Norway ($N=1,397$)		
	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
Estimation procedure: 3SLS									
β_1	-0.333*** (0.0183)	-0.389*** (0.0177)	-0.289*** (0.0195)	0.0397 (0.0635)	-0.110 (0.0753)	-0.0798 (0.0773)	1.509*** (0.142)	1.517*** (0.168)	1.565*** (0.144)
β_2	-0.0517*** (0.0161)	-0.0420*** (0.0161)	0.0171 (0.0177)	-0.107*** (0.0141)	-0.150*** (0.0167)	-0.122*** (0.0171)	-0.0390 (0.0363)	-0.189*** (0.0433)	-0.136*** (0.0369)
β_3	0.108*** (0.0139)	0.0517*** (0.0121)	0.0787*** (0.0134)	0.0617*** (0.0123)	0.0769*** (0.0144)	0.0478*** (0.0148)	0.0440 (0.0328)	0.126*** (0.0385)	0.107*** (0.0333)
F -Stat									
OR _{TOT}	53.27			75.93			39.05		
OR _{FIN}	22.05			37.87			21.69		

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Results show that the adoption of the Preferred Partner Program by hotels on Booking has an unambiguous positive effect on prices between 5 and 13% depending on the country and the channel. These price increases may have different origins.

⁶We show in Appendix - section 3.E that this selection does not incur huge variation in estimated coefficients. However it is crucial in the common trend validity.

First, complying with the Preferred Partner Program also implies being charged a higher level of commission compared to standard hotels. Higher commissions on Booking may be reflected into higher prices on this platform but also on other channels if the so called parity applies.

Second, the price variation may be driven by the gain of visibility. As claimed by Booking, a better visibility increases the demand and then hotels could respond increasing their price on the platform in order to generate more revenue. Note that a better visibility on Booking may also affect other online sales channels : consumers searching and learning the existence of this hotel on Booking may also want to search it and convert on other channels. This effect can also be negative, more visibility and more bookings on one platform may cannibalise other channels if they are substitutes. We try to estimate this effect in the following.

Table 3.13 gives the difference results of the 3SLS method on volumes.

Table 3.13: Difference in Difference results on volumes

Country c Channel j	Sweden ($N=1,610$)			Denmark ($N= 2,294$)			Norway ($N=1,763$)		
	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
Estimation procedure: 3SLS									
β_1	0.911*** (0.167)	0.681*** (0.167)	1.632*** (0.174)	-0.378 (0.371)	-0.757** (0.362)	1.170*** (0.389)	-0.576 (0.754)	-1.863** (0.833)	-0.628 (0.886)
β_2	-0.329*** (0.0915)	0.0175 (0.0875)	0.187* (0.0967)	0.130* (0.0715)	0.0209 (0.0734)	0.0526 (0.0719)	-0.106 (0.152)	0.194 (0.150)	-0.135 (0.166)
β_3	0.180* (0.103)	-0.193* (0.101)	0.0119 (0.109)	0.659*** (0.0757)	-0.0372 (0.0754)	0.521*** (0.0755)	0.0697 (0.137)	-0.217* (0.127)	0.169 (0.148)
F -Stat									
Price _{Booking}	48.64			47.91			27.86		
Price _{Direct}	46.10			43.47			29.94		
Price _{Expedia}	47.27			35.67			43.77		

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A first interesting result is that all the coefficients associated to OTAs are positive while the β_2 associated to the direct channel are negative whatever the country. This suggests that a better visibility on Booking benefits to the platform who offer program but also to the competing platform in a lower magnitude. In that sense there is a diffusion effect through platforms visibility. On the other hand, this is not pure gain since it decreases the volume of sales on the hotel website.

Results significance are not totally the same depending on the country.

In Norway, the negative effect of the adoption is highlighted : there is no gain in visibility for OTAs while sales decrease by 22% on the hotel website, suggesting that another channel or hotel may benefit from the program. This feature can exist since we have no information about the other hotel adoption on Booking. Particularly if all the similar hotels to our chain also complied with the program at the same moment, the visibility effect cancels out and the hotel may lose the benefit of the program.

In Sweden, we still observe a loss in the hotel website sales but the substitution effect is more highlighted. Indeed the adoption of the program seems to significantly affect Booking and the hotel website in the same magnitude, about 18%. In order to know if the program really benefits to the hotel in that case, we should compare relative price and market share before the adoption.

Finally in Denmark, the situation is clearly enhanced by the adoption of the program since it significantly increases the volume sold on both OTA but leaves unchanged the sales on the hotel

website. However note that the coefficients are pretty high, about +50% on Expedia and +66% on Booking. There is no clear explanation for that but one reason could be that in Denmark few hotels adopted the program on Booking which strongly benefited to the selected ones.

Overall, the visibility effect could explain the price increase on Booking, Expedia but not really the one on the hotel website. Indeed, the substitution effect would better suggest lower prices on that channel. The main reason that could explain the prices increase on the hotel website seems therefore to be the commission level that increases for the members of the Preferred Partner Program.

Still, there is a last potential origin of these prices movement which is linked to the price parity potential application at that time. Indeed, if after the adoption of the program hotels applied a narrow price parity, then the prices of different channels were linked which could explain the price increase on the hotel website without a higher number of sales or visibility by the price increase on a platform. In section 5 we show that indeed the adoption of the Preferred Partner Program coincides with the imposition of a price parity on some type of offers.

5 On the effect of Price parity application

European competition authorities have tackled the issue of price parity for several years. Since Booking.com commitments in July 2015 to switch from wide to narrow price parity in all the European Economic Area (EEA), other decisions and laws have been adopted at the national level. In France, the Macron's Law totally banned all types of price parity (including narrow ones) and a similar law rose in Belgium in November 2017 while in Germany a competition case against Booking.com and HRS in December 2015 had the same conclusion. This is still a burning issue in Sweden: if the last competition case against Booking.com banned all type of clauses in July 2018, the decision has been cancelled in early 2019.

From this recent history, the important conclusion for the period covered in the dataset, is that for our three countries of interest the wide price parity was banned and the narrow price parity allowed. Hotels could set a price lower on their website compared to OTAs but one of these last could require the hotel to fix on it the best available price among OTAs.

We identify two key determinants to the application of price parity: the type of offer selected by the consumer (refundable or not) and the type of hotel (Preferred Partner or not). Appendix 3.F details the procedures leading to these results.

5.1 General application of the price parity clause

In the data we observe both the standard and the Preferred Partner hotels before and after the adoption of the program. Figure 3.2 shows the narrow price parity application between the hotel website price and the Booking price i.e. the share of cases for which the price of Booking was lower or equal to the price of the hotel website for similar booking types.

We show that the narrow price parity seems to be applied by hotels in the database, especially for refundable offers. Indeed we show on the graph that the price on Booking is the leader in more than 80% of the case over all the period for these offers and whatever the contract of the hotel. On the opposite, on non-refundable offers price parity behaviours evolve across time and vary depending

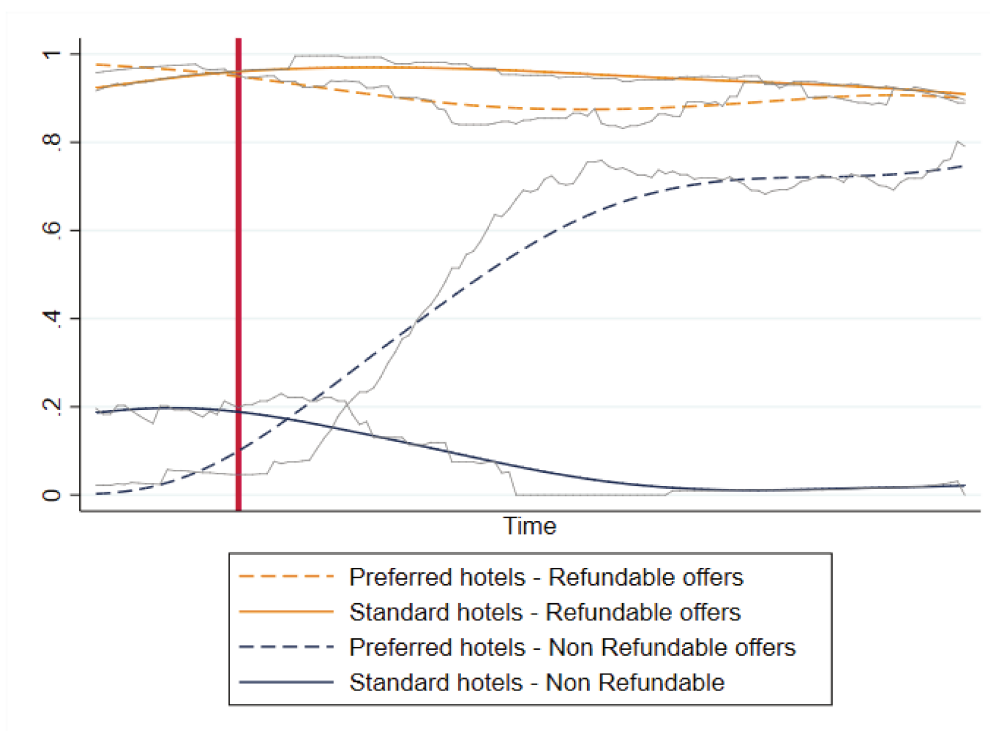


Figure 3.2: Price parity compliance on Booking by offers and types of hotels

on the adhesion to the Preferred Program. Once the Preferred Establishments Program is adopted by some hotels, these last start to self-impose a narrow price parity while the others stop to comply with it since no contractual clause require it.

At this point the question is to know whether a narrow or wide price parity is imposed on non refundable offers by preferred hotels. Indeed, Figure 3.2 shows a price parity respect between the hotel website and Booking on these offers but says nothing on the respect on Expedia for these same hotels. An identical graph for the price parity application on Expedia (Appendix 3.F.3 - Figure 3.5) shows that the hotels that adopted the Preferred Partner Program have an identical behaviour on both platforms. We therefore observe a switch from nothing to a wide price parity.

Booking website construction ranks first the hotel and then the offer within the hotel. Therefore, the visibility effect of the Preferred Partner Program should affect all types of offer, refundable or not. The hotel chain did not give us information about a differentiation in commission. Therefore we expect Preferred Partners to pay a higher commission whatever the type of offer. In comparison, as shown in Figure 3.2 the switch from nothing to wide price parity does not affect similarly the two types of contract.

This particular setting gives incentives to estimate the isolated effect of the price parity application. Indeed estimating the effect of becoming a Preferred Partner on non-refundable prices in comparison to refundable prices gives the Preferred Partner effect getting rid of the visibility and the commission effect that are common to the two types of offers.

5.2 Wide price parity effect

We run a triple difference model by reservations made on a channel j for a resort located in country c , still using the form of equation (3.1) and adding cross terms in equation (3.3) :

$$\begin{aligned} \ln \text{Price}_{jct} = & \alpha_{0jc} + \beta_{0jc}X_{jct} + \beta_{1jc}\mathbb{1}_{\text{Pref.}} + \gamma_{1jc}\text{NR} \\ & + \beta_{2jc}\mathbb{1}_{\text{After}} + \beta_{3jc}\mathbb{1}_{\text{After}\times\text{Pref.}} \\ & + \gamma_{2jc}\mathbb{1}_{\text{After}} \times \text{NR} + \gamma_{3jc}\mathbb{1}_{\text{Pref.}}\text{NR} \\ & + \tau_{jc}\mathbb{1}_{\text{After}\times\text{Pref.}}\text{NR} + \varepsilon_{jct} \end{aligned} \tag{3.4}$$

In this case, we still have to check the common trend assumption but in less restrictive way than with the difference in difference case. Indeed, in this last case we compare preferred and non preferred hotels prices for dates before treatment while in the triple difference case, we look at the trend of gaps between treated and control hotels prices, for each type of offer, and check whether they are parallel.

For that we first have to compute by type of offer, channel and country the price gap between preferred and non preferred hotels when both prices are available at the same insert date. We then compute by channel and country the difference of this price gap between refundable and non refundable offers when these two values are available. We finally plot each series by country and channel before the treatment and look at the trend (Appendix 3.G, Figure 3.6). Note that for some cases, especially in Norway, very few observations (≈ 23) are available compared to the initial sample. In the case of Sweden and Norway we have respectively 234 and 185 observations to appreciate the common trend.

In the previous section we used regressions with all the control variables described before and a test to state on the common trend assumption validity. Here the use of gap difference prevent to define the associated control variables. Graphically, the common trend assumptions are uncleaned : series are very close and there is no big change by offers but still the trends are not exactly the same. We therefore test a slightly modified version of equation (3.2) without control variable and with the gap of prices between contracts (preferred or not) by type of offers instead of the difference of prices between preferred and standard hotels as the explained variable. Using the same restriction as in the volume analysis (Appendix 3.C - Table 3.17, 7th column), we do not reject the null hypothesis at the 1% level for each couple of country and channel (Table 3.14).

Table 3.14: H_0 p -value test by country and channel - gap of prices

Country	Chi-2 Stat	p -value
Norway	6.499	0.090
Sweden	5.351	0.148
Denmark	2.781	0.427

The common trend assumption satisfied, we estimate equation 3.4 by country. We compare the estimation results for OLS, SURE, 2SLS and 3SLS τ 's coefficients are displayed in Table 3.15.

Table 3.15: Triple Difference results on prices

Country c	Sweden ($N=1,617$)			Denmark ($N= 2,462$)			Norway ($N=1,944$)		
Channel j	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
OLS	-0.0328 (0.0662)	0.0601 (0.0674)	-0.00352 (0.0621)	-0.173*** (0.0371)	-0.0394 (0.0453)	-0.0367 (0.0445)	-0.153** (0.0779)	0.118 (0.0760)	-0.0644 (0.0690)
SURE	-0.0461 (0.0584)	0.0804 (0.0618)	-0.0319 (0.0557)	-0.104*** (0.0350)	-0.0144 (0.0429)	-0.0213 (0.0415)	-0.151** (0.0755)	0.108 (0.0737)	-0.0620 (0.0666)
2SLS	-0.0782 (0.0728)	0.0702 (0.0699)	-0.00335 (0.0702)	-0.177*** (0.0380)	-0.0271 (0.0462)	-0.0452 (0.0460)	-0.167** (0.0786)	0.119 (0.0761)	-0.0579 (0.0695)
3SLS	-0.0589 (0.0604)	0.0844 (0.0619)	-0.0411 (0.0592)	-0.0975*** (0.0354)	-0.00615 (0.0432)	-0.0211 (0.0422)	-0.161** (0.0760)	0.105 (0.0738)	-0.0570 (0.0670)
<i>F</i> -Stat									
OR _{TOT}	32.70			68.55			49.29		
OR _{FIN}	18.58			33.85			23.61		

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In the case of Sweden we show no significant effect on online prices. This result is very consistent with [Mantovani et al. \(2018\)](#) that also find no effect of the EU decision in France, Spain and Italy. Recall that in 2015 the French Competition Authority, the Italian Competition Authority and the Swedish Competition Authority coordinated their investigations and adopted parallel decisions in end April. Therefore in Sweden, the ban of the wide price parity or its reintroduction through Booking tools has no effect on prices.

In Norway and Denmark, switching from no agreement to a wide price parity significantly decreases the price on Booking, by 16% in Norway and 10% in Denmark. However recall that in the case of Norway the common trend assumption was a bit less respected. This result comes against the main theoretical and empirical results so far and the antitrust decision which blamed the clause for increasing final prices and being anti-competitive. However it can easily be interpreted : going from no price parity to wide price parity, means that the OTA's prices and the hotel price convergence which can be done by a price decrease on the OTA and/or a price increase on the hotel website. The first leverage has been privileged.

6 Conclusions

This exploitation of a unique dataset from a major hotel chain located mainly in Scandadinavia provides important new insights on the relationship between hotels and online travel agencies, such as Booking.com or Expedia.

First, we underline the importance of control variables in the observation of the application of price parity clauses. The previous literature has often exploited scrapped data, with limited access to these controls. Second, we exploit the introduction of a new contract between hotels and Booking.com : the Preferred Partner Program. We analyze the impact in term of volumes and prices on differents channels through a reduced-form analysis with differences-in-differences. This program promesses seems to be satisfied : the number of reservation through Booking increases, while the number of reservations through the other channels are midly affected, positively or negatively. Prices seems to increase also on the Booking channel. Yet, these results are highly volatile across countries. It suggests that other variables, such as the proportion of competitors in the same Preferred Partner Program,

might have an impact. Lastly, we try to exploit a difference in price parity application between hotels and between offers (non-refundable reservation vs refundable). This allows us to differentiate the potential leverages behind the increase of prices to annihilate the visibility and commission fee effect and focus on the price-parity effect. Unfortunately, the statistical power tends to be low and most effects are not-significant. This could be an additional element suggesting that the price-parity effect, if anti-competitive, is small.

3.A Samples comparisons

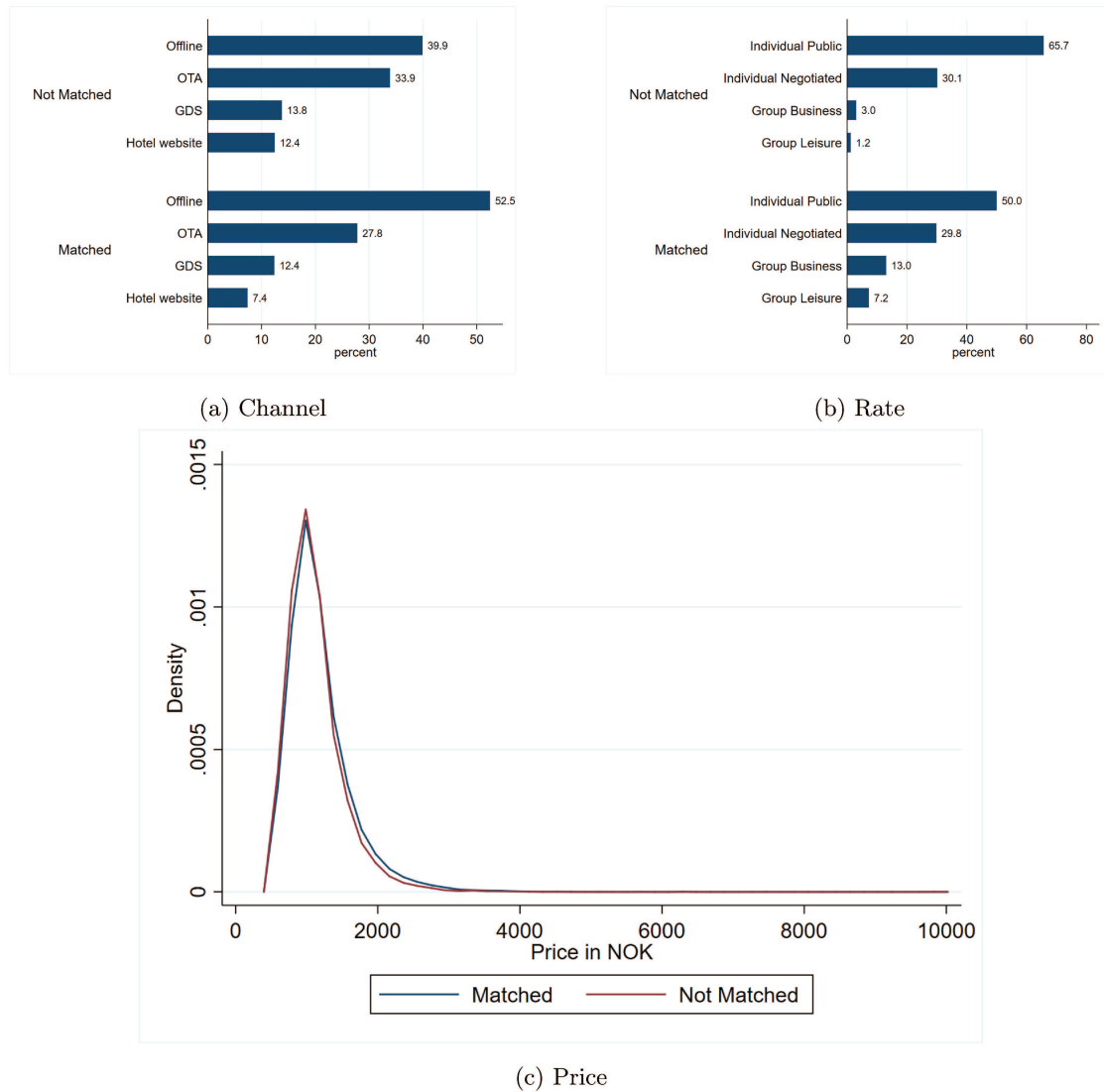


Figure 3.3: Samples comparisons

Table 3.16: Samples comparisons

	Not Matched	Matched
Lead Time (days)	25	27
Nights	1.8	1.7
Persons	1.6	1.4
Standard room	75%	85%
Week-end	31%	27%
Refundable offer	21%	17%
Occupancy rate	45%	45%

3.B Preferred Establishments Program

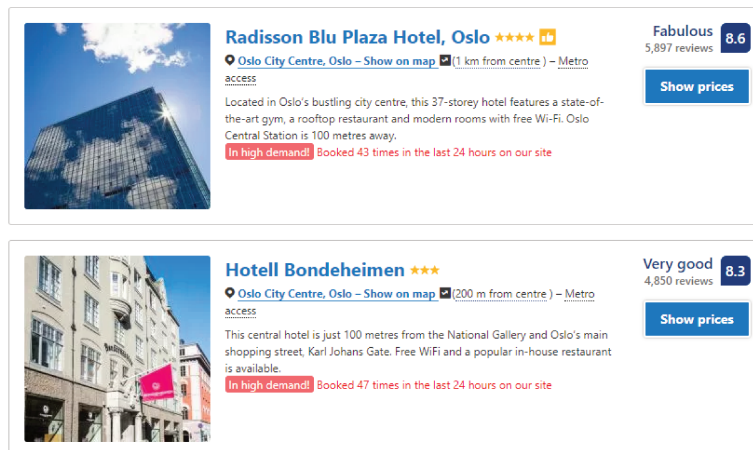


Figure 3.4: Thumb up signalling a preferred hotel

3.C Preferred Partner Program adoption

Table 3.17: Preferred hotels chain choice

	Booking Competitors ¹	Country	Type of hotel	Starting Preferred	Diff in Diff Price	Diff in Diff Volume	Triple Diff Price
Hotel 8	7/4	Norway	Standard	-	Excluded	Excluded	Excluded
Hotel 19	11/6	Sweden	Standard	-	Excluded	Excluded	Excluded
Hotel 15	13/2	Sweden	Standard	-	Excluded	Excluded	Excluded
Hotel 17	15/6	Sweden	Standard	-	Excluded	Excluded	Excluded
Hotel 12	16/4	Sweden	Standard	-	Excluded	Excluded	Excluded
Hotel 21	18/8	Sweden	Standard	-	Excluded	Excluded	Excluded
Hotel 5	19/6	Denmark	Standard	-	Excluded	Excluded	Excluded
Hotel 13	21/3	Sweden	Standard	-	Excluded	Excluded	Excluded
Hotel 16	28/9	Sweden	Standard	-	Excluded	Excluded	Excluded
Hotel 1	34/12	Denmark	Standard	-	Excluded	Excluded	Excluded
Hotel 9	35/6	Norway	Standard	-	Control	Control	Control
Hotel 22	36/17	Sweden	Standard	-	Excluded	Control	Control
Hotel 6	40/13	Denmark	Standard	-	Control	Control	Control
Hotel 2	44/17	Denmark	Standard	-	Control	Control	Control
Hotel 18	57/33	Sweden	Standard	-	Excluded	Control	Control
Hotel 14	108/51	Sweden	Preferred	17/02/2015	Treated	Treated	Treated
Hotel 7	173/37	Norway	Preferred	13/03/2015	Treated	Treated	Treated
Hotel 10	254/64	Norway	Preferred	02/06/2015	Excluded	Treated	Treated
Hotel 11	254/64	Norway	Preferred	23/03/2015	Treated	Treated	Treated
Hotel 20	258/133	Sweden	Standard	-	Control	Excluded	Excluded
Hotel 3	620/65	Denmark	Preferred	16/02/2015	Treated	Treated	Treated
Hotel 4	620/65	Denmark	Preferred	16/02/2015	Treated	Treated	Treated

¹ All accommodations listed on Booking.com in the same city/Only 3 or 4 stars hotels listed on Booking.com in the same city.

3.D Model comparisons

Table 3.18: Difference in Difference results on prices

Country c	Sweden ($N=1,206$)			Denmark ($N= 2,462$)			Norway ($N=1,397$)		
Channel j	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
Estimation procedure: OLS									
β_1	-0.355*** (0.0177)	-0.404*** (0.0174)	-0.321*** (0.0185)	0.0629 (0.0625)	-0.0754 (0.0742)	-0.0286 (0.0750)	1.513*** (0.138)	1.516*** (0.166)	1.536*** (0.142)
β_2	-0.0394** (0.0156)	-0.0307* (0.0159)	0.0320* (0.0172)	-0.0758*** (0.0122)	-0.122*** (0.0145)	-0.0741*** (0.0144)	-0.0242 (0.0359)	-0.179*** (0.0433)	-0.128*** (0.0370)
β_3	0.111*** (0.0140)	0.0576*** (0.0120)	0.0844*** (0.0130)	0.0470*** (0.0119)	0.0678*** (0.0141)	0.0324** (0.0142)	0.0342 (0.0322)	0.121*** (0.0384)	0.103*** (0.0333)
Estimation procedure: SURE									
β_1	-0.364*** (0.0173)	-0.409*** (0.0171)	-0.328*** (0.0182)	0.0482 (0.0620)	-0.0945 (0.0736)	-0.0499 (0.0741)	1.532*** (0.136)	1.524*** (0.164)	1.547*** (0.140)
β_2	-0.0425*** (0.0153)	-0.0350** (0.0157)	0.0280* (0.0169)	-0.0759*** (0.0121)	-0.119*** (0.0143)	-0.0766*** (0.0142)	-0.0286 (0.0354)	-0.180*** (0.0428)	-0.129*** (0.0365)
β_3	0.114*** (0.0135)	0.0562*** (0.0119)	0.0851*** (0.0128)	0.0495*** (0.0118)	0.0681*** (0.0140)	0.0344** (0.0140)	0.0366 (0.0318)	0.121*** (0.0379)	0.103*** (0.0329)
Estimation procedure: 2SLS									
β_1	-0.321*** (0.0189)	-0.385*** (0.0179)	-0.279*** (0.0199)	0.0557 (0.0642)	-0.0854 (0.0759)	-0.0564 (0.0783)	1.497*** (0.144)	1.513*** (0.171)	1.565*** (0.146)
β_2	-0.0496*** (0.0164)	-0.0361** (0.0163)	0.0216 (0.0180)	-0.107*** (0.0143)	-0.155*** (0.0169)	-0.125*** (0.0174)	-0.0349 (0.0368)	-0.190*** (0.0439)	-0.136*** (0.0374)
β_3	0.106*** (0.0145)	0.0513*** (0.0122)	0.0779*** (0.0137)	0.0588*** (0.0124)	0.0762*** (0.0146)	0.0480*** (0.0150)	0.0431 (0.0332)	0.129*** (0.0390)	0.110*** (0.0338)
Estimation procedure: 3SLS									
β_1	-0.333*** (0.0183)	-0.389*** (0.0177)	-0.289*** (0.0195)	0.0397 (0.0635)	-0.110 (0.0753)	-0.0798 (0.0773)	1.509*** (0.142)	1.517*** (0.168)	1.565*** (0.144)
β_2	-0.0517*** (0.0161)	-0.0420*** (0.0161)	0.0171 (0.0177)	-0.107*** (0.0141)	-0.150*** (0.0167)	-0.122*** (0.0171)	-0.0390 (0.0363)	-0.189*** (0.0433)	-0.136*** (0.0369)
β_3	0.108*** (0.0139)	0.0517*** (0.0121)	0.0787*** (0.0134)	0.0617*** (0.0123)	0.0769*** (0.0144)	0.0478*** (0.0148)	0.0440 (0.0328)	0.126*** (0.0385)	0.107*** (0.0333)
F -Stat									
OR _{TOT}	53.27			75.93			39.05		
OR _{FIN}	22.05			37.87			21.69		

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.19: Difference in Difference results on volumes

Country c	Sweden ($N=1,610$)			Denmark ($N= 2,294$)			Norway ($N=1,763$)		
Channel j	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
Estimation procedure: OLS									
β_1^d	0.785*** (0.152)	0.520*** (0.141)	1.493*** (0.163)	-0.454 (0.367)	-0.750** (0.364)	1.208*** (0.387)	-1.045* (0.615)	-1.757*** (0.558)	-1.854*** (0.647)
β_2^d	-0.317*** (0.0897)	0.0681 (0.0830)	0.193** (0.0966)	0.219*** (0.0675)	0.0418 (0.0678)	0.104 (0.0704)	-0.101 (0.154)	0.186 (0.140)	-0.0643 (0.161)
β_3^d	0.147 (0.0990)	-0.265*** (0.0910)	-0.0314 (0.107)	0.567*** (0.0709)	-0.0526 (0.0702)	0.477*** (0.0736)	0.0707 (0.138)	-0.215* (0.124)	0.116 (0.144)
Estimation procedure: SURE									
β_1^d	0.766*** (0.150)	0.517*** (0.140)	1.444*** (0.161)	-0.499 (0.363)	-0.764** (0.361)	1.147*** (0.383)	-0.980 (0.609)	-1.722*** (0.553)	-1.936*** (0.639)
β_2^d	-0.285*** (0.0883)	0.0693 (0.0819)	0.213** (0.0952)	0.205*** (0.0668)	0.0332 (0.0673)	0.0879 (0.0697)	-0.110 (0.152)	0.181 (0.138)	-0.0611 (0.159)
β_3^d	0.116 (0.0976)	-0.275*** (0.0899)	-0.0573 (0.106)	0.571*** (0.0702)	-0.0518 (0.0697)	0.481*** (0.0730)	0.0732 (0.136)	-0.210* (0.123)	0.113 (0.142)
Estimation procedure: 2SLS									
β_1^d	0.936*** (0.169)	0.714*** (0.169)	1.677*** (0.177)	-0.359 (0.374)	-0.745** (0.365)	1.245*** (0.393)	-0.767 (0.764)	-1.946** (0.844)	-0.636 (0.901)
β_2^d	-0.363*** (0.0930)	0.00544 (0.0886)	0.170* (0.0980)	0.159** (0.0723)	0.0359 (0.0741)	0.0680 (0.0727)	-0.0974 (0.154)	0.204 (0.152)	-0.133 (0.168)
β_3^d	0.214** (0.105)	-0.168 (0.102)	0.0335 (0.111)	0.635*** (0.0765)	-0.0469 (0.0760)	0.517*** (0.0762)	0.0675 (0.138)	-0.224* (0.128)	0.168 (0.150)
Estimation procedure: 3SLS									
β_1^d	0.911*** (0.167)	0.681*** (0.167)	1.632*** (0.174)	-0.378 (0.371)	-0.757** (0.362)	1.170*** (0.389)	-0.576 (0.754)	-1.863** (0.833)	-0.628 (0.886)
β_2^d	-0.329*** (0.0915)	0.0175 (0.0875)	0.187* (0.0967)	0.130* (0.0715)	0.0209 (0.0734)	0.0526 (0.0719)	-0.106 (0.152)	0.194 (0.150)	-0.135 (0.166)
β_3^d	0.180* (0.103)	-0.193* (0.101)	0.0119 (0.109)	0.659*** (0.0757)	-0.0372 (0.0754)	0.521*** (0.0755)	0.0697 (0.137)	-0.217* (0.127)	0.169 (0.148)
F -Stat									
Price _{Booking}		48.64			47.91			27.86	
Price _{Direct}		46.10			43.47			29.94	
Price _{Expedia}		47.27			35.67			43.77	

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.E Hotel selection

From Table 3.C, it is clearly observed that a trade-off has to be made in the different analysis (difference in differences on prices and volumes, triple differences). Indeed on the one hand, we would ideally keep all the hotels in the three analysis. However with all them the common trend maybe is not satisfied. In addition, the hotels selected to satisfy the common trend are not necessarily the same. Hence, there is a trade off between satisfying the common trend assumption and having the same subset of hotels in the three analysis.

There is no question in the case of Denmark since all the hotels are kept in the three analysis. The question especially arise in the case of Sweden since in the control group there is only one hotel (Hotel 20) for the difference in difference on price (restriction 1) while the two other non-treated hotels (Hotel 18 and 22) are used in the control group for the difference in difference on volume and the triple difference (restriction 2).

We show in Table 3.20 that indeed taking the same restrictions as in the control group for the difference in difference on volume and the triple difference, the common trend assumption for the

difference in difference in the price is not verified anymore in the case of Sweden and less in Norway. However Table 3.21 shows that it does not affect in huge proportion the magnitude of the coefficient.

Table 3.20: H_0 p -value test by country and channel - price - Comparison

Country	Restriction 1 p -value	Restriction 2 p -value
Norway	0,5522	0.0179
Sweden	0,3696	6.563e-28
Denmark	0,5264	0.5264

Table 3.21: Coefficient comparison on prices

Estimation procedure: 3SLS Restriction 1									
Country c	Sweden ($N=1,562$)			Denmark ($N= 2,462$)			Norway ($N=1,397$)		
Channel j	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
β_1	-0.333*** (0.0183)	-0.389*** (0.0177)	-0.289*** (0.0195)	0.0397 (0.0635)	-0.110 (0.0753)	-0.0798 (0.0773)	1.509*** (0.142)	1.517*** (0.168)	1.565*** (0.144)
β_2	-0.0517*** (0.0161)	-0.0420*** (0.0161)	0.0171 (0.0177)	-0.107*** (0.0141)	-0.150*** (0.0167)	-0.122*** (0.0171)	-0.0390 (0.0363)	-0.189*** (0.0433)	-0.136*** (0.0369)
β_3	0.108*** (0.0139)	0.0517*** (0.0121)	0.0787*** (0.0134)	0.0617*** (0.0123)	0.0769*** (0.0144)	0.0478*** (0.0148)	0.0440 (0.0328)	0.126*** (0.0385)	0.107*** (0.0333)
Estimation procedure: 3SLS - Restriction 2									
Country c	Sweden ($N=1,617$)			Denmark ($N= 2,462$)			Norway ($N=1,944$)		
Channel j	Booking	Direct	Expedia	Booking	Direct	Expedia	Booking	Direct	Expedia
β_1	0.297*** (0.0449)	0.424*** (0.0436)	0.271*** (0.0493)	0.0397 (0.0635)	-0.110 (0.0753)	-0.0798 (0.0773)	1.542*** (0.137)	1.521*** (0.158)	1.594*** (0.137)
β_2	-0.0849*** (0.0242)	-0.130*** (0.0235)	-0.00587 (0.0266)	-0.107*** (0.0141)	-0.150*** (0.0167)	-0.122*** (0.0171)	-0.0200 (0.0344)	-0.167*** (0.0398)	-0.137*** (0.0344)
β_3	0.151*** (0.0255)	0.221*** (0.0249)	0.121*** (0.0280)	0.0617*** (0.0123)	0.0769*** (0.0144)	0.0478*** (0.0148)	0.0226 (0.0311)	0.101*** (0.0357)	0.103*** (0.0310)

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.F Details on the matching procedure

3.F.1 General matching

As prices have multiple determinants, to properly estimate the price parity application, one should compare bookings that are identical in all aspects, except for the reservation channel. In our dataset, some bookings are identical in terms of dates (reservation and arrival date), number of adults, number of children, number of rooms, duration, type of room, type of offer (refundable or not) and resort but have been made through different channels of reservations: Booking, Expedia or the hotel website. When multiple identical reservations exists on a given channel, and can be compared to other bookings on an other channel, we consider all pairs of comparisons. As we want to look at strict (non)-equalities of prices between channels, averaging would add noise to our estimations.

We then compare the distribution of prices and in particular identify which channel is the price leader for each of the booking types. Table 3.22 show the results for the 34,345 matches of three prices.

Table 3.22: Price leader

All offers	Freq.	Percent	Cum.
All equal	21,164	61.62	61.62
Hotel website	3,707	10.79	72.42
Expedia/Hotel website	2,667	7.77	80.18
Booking/Hotel website	2,483	7.23	87.41
Expedia	1,814	5.28	92.69
Booking	1,672	4.87	97.56
Booking/Expedia	838	2.44	100.00
Total	34,345	100.00	

For bookings with exactly same characteristics, we observe that in 62% of cases the hotel sets the same price on the three channel, suggesting an idea of parity between channels. Nevertheless, if the prices are not equal, the hotel website is generally cheaper : overall it is the price leader (equality or not) in 87% of cases. In that sense, the narrow nor the wide price parities apply.

Only looking at the two platforms, we observe a symmetry. If their prices are not equal, they are the only price leader on the online market in 5% of total cases and a joint price leader with the hotel website in 7% of total cases.

3.F.2 Type of offer

Interestingly this result is nuanced depending on the type of offer selected by the consumer : refundable or not and the type of contract between the hotel and the platform : preferred or not. Coming back to our previous analysis of price triplet, one price by channel, we observe that only 4% (1,428 observations) of matches are for non refundable offers while all the other matched are for refundable offers. This comes from the fact, there is few booking types that are exactly the same on the three channels for non refundable offers. For this reason we adapt the analyse only looking at pair prices (the hotel website price and the price on a particular OTA or prices on Booking and Expedia). In the

case of the OTAs prices comparison (Table 3.23), this gives five times more observations (5,156) and then helps to have better conclusions.

Table 3.23: OTAs price leader by offer

Non-Refundable	Freq.	Percent	Cum.	Refundable	Freq.	Percent	Cum.
Booking	2,887	55.99	55.99	Both equal	9,399	83.44	83.44
Expedia	1,784	34.60	90.59	Expedia	1,059	9.40	92.84
Both Equal	485	9.41	100.00	Booking	807	7.16	100.00
Total	5,156	100.00		Total	11,265	100.00	

We clearly observe that a narrow price parity seems to be applied on refundable offers while it is not the case on the non refundable offers. Table 3.24 and 3.25 respectively show the price comparison between an OTA price and the hotel website.

Table 3.24: Booking vs hotel website price leader by offer

Non-refundable	Freq.	Percent	Cum.	Refundable	Percent	Cum.
Hotel website	1,113	65.74	65.74	Both equal	2,089	87.15
Booking	538	31.78	97.52	Hotel website	196	8.18
Both equal	42	2.48	100.00	Booking	112	4.67
Total	1,693	100.00		Total	2,397	100.00

Table 3.25: Expedia vs hotel website price leader by offer

Non-Refundable	Freq.	Percent	Cum.	Refundable	Freq.	Percent	Cum.
Hotel website	535	47.56	47.56	Both equal	1,289	79.18	79.18
Expedia	532	47.29	94.84	Hotel website	193	11.86	91.03
Both equal	58	5.16	100.00	Expedia	146	8.97	100.00
Total	1,125	100.00		Total	1,628	100.00	

Again, using a couple analysis rather than a triplet one helps us in these cases to have similar proportion of type of offers. Interestingly we show that on Booking for non refundable offers, more than a narrow price parity, a wide price parity seems to be applied. In fact, in 92% of cases the price on the website is higher or equal to the price on Booking. This is also the case on Expedia but in lower proportions (88%). In comparison, on non refundable contracts there is a unique price leader in respectively 98% and 95% of cases for Booking and Expedia and its not necessarily the platform.

Overall this suggests that the hotels' strategies differs between types of offers that coexist on a same web page : a narrow price parity seems to be applied on refundable contracts and even the wide price parity is banned, hotels seems to have incentives to self-impose it. In comparison, on non-refundable contracts, prices does not to seem to be constrained. As said before, our hotels sells proportionally more refundable contracts on OTA than on their website. The consumers interest on this type of offer on OTAs may explain why hotels applies more restrictive clauses for them.

To conclude this section, these results are not surprising. First because narrow price parity was allowed at that time in these countries so it is logical to see it appearing in the data for at least so offers. Second, the observation of a wide price parity application has a justification: during the period

some hotels voluntarily committed to apply it subscribing to the Preferred Establishment Program offered by Booking. For this reason it is important to take into account the type of contract between the platform and the hotel in the analysis.

3.F.3 Price parity application by platform, offer and contract

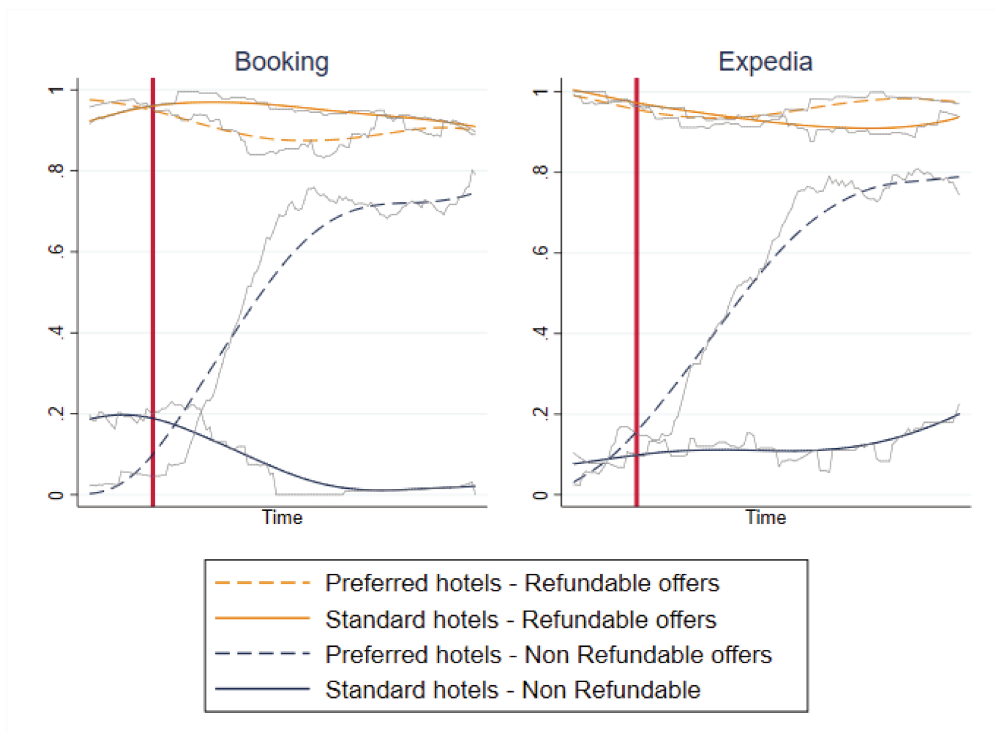


Figure 3.5: Price parity compliance by platform, offers and types of hotels

3.G Common trend for triple difference

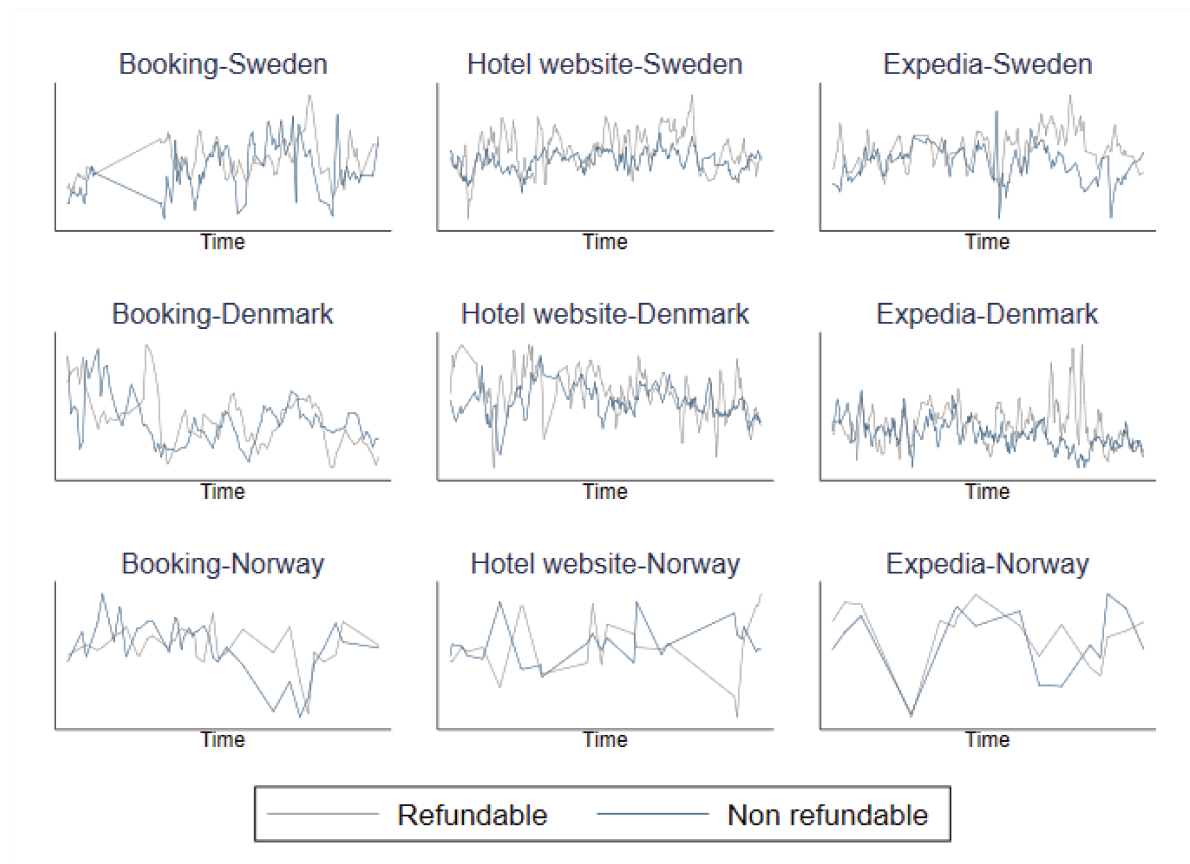


Figure 3.6: Gap of prices between type of contracts by offer, country and channel before the treatment

Bibliography

- ABRATE, G. AND G. VIGLIA (2016): "Strategic and tactical price decisions in hotel revenue management," *Tourism Management*, 55, 123–132.
- BOIK, A. AND K. CORTS (2016): "The Effects of Platform Most-Favored-Nation Clauses on Competition and Entry," *Journal of Law and Economics*, 59, 105–134.
- CAZAUBIEL, A., M. CURE, B. O. JOHANSEN, AND T. VERGÉ (2020): "Substitution between on-line distribution channels: Evidence from the Oslo hotel market," *International Journal of Industrial Organization*, 69, 102577.
- CHO, S., G. LEE, J. RUST, AND M. YU (2018): "Optimal Dynamic Hotel Pricing." .
- COURT, A. T. (1939): "Hedonic Price Indexes with Automotive Examples," *The Dynamics of Automobile Demand*.
- EDELMAN, B. AND J. WRIGHT (2015): "Price Coherence and Excessive Intermediation," *Quarterly Journal of Economics*, 130, 1283–1328.
- GALLEGO, G. AND G. VAN RYZIN (1994): "Optimal dynamic pricing of inventories with stochastic demand over finite horizons." *Management Science*, 40, 999–1020.
- GHOSE, A., P. G. IPEIROTIS, AND B. LI (2014): "Examining the Impact of Ranking on Consumer Behavior and Search Engine Revenue," *Management Science*, 60, 1632–1654.
- HUNOLD, M., R. KESLER, AND LAITENBERGER (Forthcoming): "Hotel Rankings of Online Travel Agents, Channel Pricing, and Consumer Protection." *Marketing Science*.
- HUNOLD, M., U. LAITENBERGER, AND F. SCHLÜTTER (2018): "Evaluation of best price clauses in hotel booking," *International Journal of Industrial Organization*, 61, 542–571.
- JOHANSEN, B. O. AND T. VERGÉ (2016): "Secret Contractung in Multilateral Relations," *mimeo*.
- (2017): "Platform Investment and Price Parity clauses," *mimeo*.
- JOHNSON, J. (2017): "The Agency Model and MFN Clauses," *The Review of Economic Studies*, 84, 1151–1185.
- LANCASTER, K. (1966): "A New Approach to Consumer Theory," *The Journal of Political Economy*, 74, 132–157.

BIBLIOGRAPHY

- LARRIEU, T. (2019): "Pricing strategies in online market places and Price Parity Agreements: evidence from the hotel industry," *mimeo*.
- LAW, R., Q. YE, AND Z. ZHANG (2011): "Determinants of hotel room price: An exploration of travelers' hierarchy of accommodation needs," *International Journal of Contemporary Hospitality Management*, 23, 972–981.
- MANTOVANI, A., C. PIGA, AND C. REGGIANI (2018): "Much ado about nothing? Online platform price parity clauses and the EU Booking.com case," *mimeo*.
- MONTY, B. AND M. SKIDMORE (2003): "Hedonic Pricing and Willingness to Pay for Bed and Breakfast Amenities in Southeast Wisconsin," *Journal of Travel Research*, 42, 195–199.
- ROSEN, S. (1974): "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *The Journal of Political Economy*, 82, 34–55.
- TALLURI, K. T. AND G. VAN RYZIN (2005): "The Theory and Practice of Revenue Management," *Springer, New York*.
- THRANE, C. (2007): "Examining the determinants of room rates for hotels in capital cities: the Oslo experience," *Journal of Revenue and Pricing Management*, 5, 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, 530–552.
- WANG, C. AND J. WRIGHT (2016): "Platform Investment and Price Parity clauses," *mimeo*.
- (2017): "Search Platforms: Showrooming and Price parity Clauses," *mimeo*.
- WILLIAMS, K. R. (2017): "Dynamic airline pricing and seat availability." *mimeo*.
- 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 H. THEIL (1962): "Three stage least squares: Simultaneous estimate of simultaneous equations." *Econometrica*, 29, 54–78.

Titre : Essais sur l'économie des plateformes

Mots clés : plateformes, stratégies commerciales, information révélée, substitution de la demande entre canaux numériques, agences de voyages en ligne

Résumé : Cette thèse traite de l'économie des plateformes à travers trois chapitres indépendants. Le premier développe une modélisation théorique autour des ventes exclusives, ou ventes flash. Cette pratique, très courante sur internet, consiste à proposer à des consommateurs une offre à prendre ou à laisser aujourd'hui, sans possibilité d'en bénéficier demain. Le second chapitre évalue la substituabilité

des différents canaux de distribution de chambres d'hôtels en Scandinavie, notamment entre le site internet de l'hôtel, Booking, et Expedia. Nous analysons également la décision d'une chaîne d'hôtels de boycotter un canal de vente. Enfin, le troisième chapitre analyse l'implémentation d'une nouvelle stratégie par Booking auprès de ses hôteliers.

Title : Essays on e-commerce platforms

Keywords : multi-sided platforms, marketing strategies, information revelation, online demand substitution, online travel agencies

Abstract : This dissertation deals with the economics of platforms through three independent chapters. The first one develops a theoretical model around exclusive sales, or flash sales. This practice, very common on the Internet, consists in offering consumers an offer to take today, with no opportunity to benefit tomorrow. The second chapter evaluates the substitutability

of hotel room distribution channels in Scandinavia, particularly between the hotel's website, Booking, and Expedia. We also analyze the decision of a hotel chain to boycott a sales channel. Finally, the third chapter analyses the implementation of a new strategy by Booking with its suppliers.