



HAL
open science

Valeur ajoutée du ciblage publicitaire

Rémi Devaux

► **To cite this version:**

Rémi Devaux. Valeur ajoutée du ciblage publicitaire. Economies et finances. Université Paris sciences et lettres, 2023. Français. NNT : 2023UPSLM015 . tel-04210777

HAL Id: tel-04210777

<https://pastel.hal.science/tel-04210777>

Submitted on 19 Sep 2023

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.

THÈSE DE DOCTORAT

DE L'UNIVERSITÉ PSL

Préparée à Mines Paris-PSL

Valeur ajoutée du ciblage publicitaire
The Added Value of Targeted Advertising

Soutenue par

Rémi DEVAUX

Le 4 avril 2023

Ecole doctorale n° 543

**Sciences de la Décision, des
Organisations, de la Société
et de l'Échange (SDOSE)**

Spécialité

Economie

Composition du jury :

Marc, BOURREAU
Professeur, Télécom Paris

*Président
Rapporteur*

Grazia, CECERE
Professeur, IMT Business School

Rapporteur

Anne, PERROT
Inspection Générale des Finances

Examineur

Olivier, BOMSEL
Professeur, Mines Paris - PSL

Directeur de thèse

« Toute publicité dit le produit (c'est sa connotation), mais elle raconte autre chose (c'est sa dénotation); c'est pourquoi on ne peut que la ranger aux côtés de ces grands aliments de nutrition psychique [...], que sont pour nous la littérature, le spectacle, le cinéma, le sport, la Presse, la Mode : en touchant le produit par le langage publicitaire, les hommes lui donnent du sens et transforment ainsi son simple usage en expérience de l'esprit. »

Roland Barthes, *Le message publicitaire, rêve et poésie* (1963)

Remerciements

La science est une aventure collective. Hélas, cette thèse ne sera associée qu'à mon nom. Elle n'aurait pourtant pas vu le jour sans le soutien de ceux que je souhaite remercier ici.

Je ne serai jamais assez reconnaissant envers Oliver Bomsel de m'a ouvert les portes de la recherche en économie. Olivier m'a mis sur la piste de cette thèse avec deux textes : un article de Gary Becker et Kevin Murphy ainsi que les *Mythologie* de Roland Barthes. Ces deux références m'ont passionné, peut-être obsédé, et me poursuivront encore longtemps. J'y ai notamment appris que la publicité produisant un bien intangible : le sens. Du sens, Olivier en a donné à tous les niveaux de cette thèse. Il m'a aussi offert l'opportunité de faire connaître – publier, dirait-il – mes recherches auprès du grand public, ce grand Autre auquel les médias donnent accès. En bref, Olivier a endossé beaucoup rôles : directeur de thèse, co-auteur, *sparring partner*, et je lui en serai toujours redevable.

Cette thèse n'aurait pas non plus été possible sans l'implication d'*Ekimetrics*. A ce titre, je remercie chaleureusement Jean-Baptiste Bouzige pour sa confiance et la liberté qu'il a accordé à cette thèse. Merci aussi à Nicolas Chesneau, qui a toujours été attentif au bon déroulé de mes travaux et à mon intégration chez *Ekimetrics*.

Je remercie Pierre Fleckinger et Margaret Kyle, membres de mon comité de suivi de thèse pour leur conseil avisé sur mes travaux. Un grand merci également à Barbara pour son aide dévouée et sa bonne humeur. Elle a fait du CERNA ma seconde maison.

Sur un plan plus personnel, je dois beaucoup à mes collègues et amis doctorants grâce à qui ces trois ans fut aussi amusant que stimulant. Merci donc à Victor, Simon, Félix, Hugo, Amélie, Guillaume, Tiphaine et Nassireddine. En outre, je ne peux cacher que je suis particulièrement redevable d'Amélie et Guillaume pour leurs heures passées à relire et corriger les différentes parties de cette thèse.

Je souhaitais également remercier les brillants stagiaires avec qui j'ai eu le plaisir de travailler sur les chapitres de cette thèse : Albane, Alix et Benjamin.

Je dois beaucoup à Philippe de m'avoir appris à encaisser et rendre les coups. Le club Cenvint, ses leçons, m'ont aidé à mener à bien cette thèse. Elles m'épauleront encore longtemps.

Enfin, un grand merci à mes proches, amis et familles pour leur soutien. Une pensée particulière pour Pauline, présente quotidiennement dans les hauts et les bas de ces trois années, et encore bien d'autres.

Ce qui suit vous est dédié.

Résumé

La publicité en ligne abaisse considérablement les coûts de ciblage du consommateur. Cette thèse étudie les apports et limites de la publicité en ligne au travers de 4 études empiriques basées sur des données d’annonceurs. Le **Chapitre 1** montre que les liens sponsorisés (publicités *search*) bénéficient largement des publicités hors-ligne. Je montre qu’en augmentant son activité publicitaire hors-ligne d’1%, une marque génère jusque’à 0.95% de clics supplémentaires sur ses liens sponsorisés. Le **Chapitre 2** porte sur la substituabilité des publicités hors-ligne et numériques. À l’aide d’un modèle translog, je conclus que les publicités hors-ligne et numériques sont des substituts limités. Les formats publicitaires numériques (*display* et *search*) sont en revanche fortement substituables. Dans le **Chapitre 3**, je m’intéresse aux asymétries d’information quant au placement des publicités en ligne. Je montre que les contrats en coût-par-impression (CPM) incitent moins les régies à rendre la publicité visible que les contrat de coût-par-vue. L’achat programmatique — basé sur des intermédiaires publicitaires — expose l’annonceur à une visibilité et une audience de moins bonne qualité que l’achat en direct. En outre, l’adéquation des publicités aux contenus des sites engendre des taux de clic 69% supérieurs au ciblage des consommateurs indépendamment du contexte. Enfin, si les inventaires premium ne sont pas plus cliqués à court terme, ils semblent chasser les mauvais espaces du marché standard. Les effets de contexte sont également abordés dans le **Chapitre 4**. À partir de modèles en double-différences et d’une estimation de contrefactuels, je montre que la circulation des contenus controversés et la dégradation de l’image de Facebook durant le boycott de juillet 2020 ont altéré la valeur des publicités sur la plateforme. Les publicités Facebook ont enregistré 5000 à 10 000 clics de moins, comparées aux autres formats *display* de la marque. Leur prix a également baissé. Cette thèse conclut que la publicité en ligne est davantage un complément de la publicité traditionnelle qu’un substitue. Enfin, elle plaide pour une meilleure contextualisation de la publicité. Celle-ci apparaît aujourd’hui essentielle à l’heure où la régulation limite l’usage des données personnelles à des fins publicitaire.

Mots clés : Publicité, Econométrie Appliquée, Ciblage, Externalités

Abstract

Online advertising significantly lowers the costs of targeting individuals. This thesis studies the contributions and limitations of online advertising through 4 empirical studies based on advertiser data. **Chapter 1** shows that sponsored links (*search* ads) benefit greatly from offline ads. I show that when increasing its offline advertising activity by 1%, a brand generates up to 0.95% additional clicks on its sponsored links. **Chapter 2** focuses on the substitutability between offline and digital ads. Using a translog model, I find that offline and digital ads are limited substitutes. Digital ad formats (*display* and *search*) are highly substitutable. **Chapter 3** focuses on information asymmetries in the placement of online ads. I show that cost-per-impression (CPM) contracts do not provide incentives for advertisers to make ads visible compared to cost-per-view contracts. Programmatic buying - based on advertising intermediaries - exposes the advertiser to a lower visibility and audience quality compared to direct buying. In addition, matching ads with website content results in 69% higher click-through rates than ads that only target consumers regardless of context. Finally, while ads bought from premium inventories are not more clicked, it seems to be driving out low-quality ad spaces from the standard inventories. Context effects are also discussed in **Chapter 4**. Using differences-in-difference and counterfactual estimations, I show that the circulation of controversial content and the degradation of Facebook's credibility during the July 2020 boycott altered the value of ads on the platform. From June to July 2020, Facebook ads recorded 5,000 to 10,000 fewer clicks compared to the brand's other display campaigns. Their price also dropped. This thesis concludes that online advertising is more a complement to traditional advertising than a substitute. I also advocate for a better contextualization of advertising. This appears to be essential as regulation limits the use of personal data for advertising purposes.

Keywords : Advertising, Applied Econometrics, Targeting, Externalities

Table des matières

| | |
|--|-------------|
| Remerciements | ii |
| Résumé | iii |
| Abstract | iv |
| Table des matières | iv |
| Liste des figures | viii |
| Liste des tableaux | ix |
| Introduction de la thèse | xi |
| 1 L'économie de la publicité | xi |
| 2 Médias et publicité : une économie des récits | xii |
| 3 Publicité en ligne : organisation industrielle | xiii |
| 3.1 Internet : quoi de neuf? | xiii |
| 3.2 Splendeurs et misères de la publicité en ligne | xiv |
| 4 Contributions de la thèse | xvi |
| 1 Externalities across Advertising Markets | 1 |
| 1 Introduction | 2 |
| 2 Related Works | 4 |
| 2.1 Advertising Externalities | 4 |
| 2.2 Cross-media Effects | 4 |
| 2.3 Advertisers and Publishers' Profit | 5 |
| 3 Research Hypotheses & Data | 6 |
| 3.1 Conceptual Framework | 6 |
| 3.2 Data Sources | 6 |
| 3.3 Descriptive Evidences | 8 |
| 4 Identification Strategy | 11 |
| 5 Results and Discussion | 12 |
| 5.1 Adstock Carryover Rates | 12 |
| 5.2 Mechanism | 14 |
| 5.3 Robustness Checks | 16 |
| 6 Implications | 17 |
| 6.1 Managerial Implications for Advertisers | 17 |
| 6.2 Free-riding and Market Power | 18 |
| 6.3 Advertising Media Market | 18 |

| | | |
|----------|--|-----------|
| 7 | Conclusion | 19 |
| 8 | Appendix | 21 |
| 8.1 | Additional Descriptive Statistics | 21 |
| 8.2 | Facebook Weekly Data Conversion | 22 |
| 8.3 | Grid Search Algorithm | 22 |
| 8.4 | Instrumental Variable Approach | 23 |
| 8.5 | Clicks and Cost Elasticities in Value | 24 |
| 8.6 | Full Regression Results (Including Ad Budget) | 25 |
| 8.7 | Robustness Chekcs | 28 |
| 2 | Better Together ? Complementarity and Substitution between Offline and Online Advertising | 31 |
| 1 | Introduction | 33 |
| 2 | The Model | 36 |
| 2.1 | Theoretical Framework | 36 |
| 2.2 | Functional Form | 36 |
| 2.3 | Deriving Elasticities | 37 |
| 3 | Data and Estimation | 38 |
| 3.1 | Data Sources | 38 |
| 3.2 | Descriptive Statistics | 39 |
| 3.3 | Estimation Strategy | 39 |
| 4 | Results | 43 |
| 4.1 | Summary of the Findings | 43 |
| 4.2 | Analysis by Industry | 45 |
| 4.3 | Robustness Checks | 46 |
| 5 | Implications and Concluding Words | 48 |
| 6 | General Appendix | 50 |
| 6.1 | Shephard's Lemma | 50 |
| 6.2 | From GRPs to Impressions | 50 |
| 7 | Model robustness | 50 |
| 7.1 | First-stage Results | 50 |
| 7.2 | Correcting for Autocorrelation | 51 |
| 7.3 | Concavity Checks | 52 |
| 7.4 | Dealing with Unobserved Prices | 53 |
| 8 | Additional Results | 55 |
| 8.1 | Four-inputs Model | 55 |
| 3 | Display Advertising : How Context Matters ? | 57 |
| 1 | Introduction | 59 |
| 2 | Related Works | 61 |
| 2.1 | The Effect of (Contextually) Targeted Advertising | 61 |
| 2.2 | Ad Viewability | 62 |
| 2.3 | Incentive Misalignment in Online advertising | 63 |
| 3 | Data | 64 |
| 3.1 | Dataset 1 : Ad Pricing and Viewability | 64 |
| 3.2 | Dataset 2 : Ad Context and Effectiveness | 66 |
| 4 | Econometric Analysis | 67 |
| 4.1 | General Approach | 67 |
| 4.2 | Case Study 1 : The Effect of Pricing on Viewability | 68 |
| 4.3 | Case Study 2 : The Price and Effectiveness of Context | 69 |

| | | |
|----------|--|------------|
| 4.4 | Robustness Checks | 71 |
| 5 | Conclusion | 71 |
| 6 | Appendix | 74 |
| 6.1 | The iterative OLS (iOLS) Estimation | 74 |
| 6.2 | Robustness Checks | 75 |
| 6.2.1 | Other Dependent Variable Specifications | 75 |
| 6.2.2 | Results | 76 |
| 6.3 | Full Regression Results | 77 |
| 6.4 | Heterogeneous Effect across Devices | 79 |
| 4 | Should I Stay or Should I Go ? Social Media Advertising in Times of Boycott | 80 |
| 1 | Introduction | 82 |
| 2 | Background & Data | 85 |
| 2.1 | The Boycott : Timeline and Expected Effects | 85 |
| 2.2 | Data and Descriptive Evidence | 86 |
| 3 | Identification Strategy | 91 |
| 4 | Results | 92 |
| 4.1 | Facebook Ad Performance and Prices Dropped during the Boycott | 92 |
| 4.2 | Content Moderation Did not Fully Offset the Boycott Nuisance | 93 |
| 4.3 | The Ad Boycott's Media Coverage is Negatively Correlated to Facebook's Ad Performance and Price | 95 |
| 4.4 | Robustness Checks | 96 |
| 5 | Concluding Words | 97 |
| 6 | Appendix | 98 |
| 6.1 | Summary Statistics | 98 |
| 6.2 | Effect of Boycott on Population Exposed | 99 |
| 6.3 | Facebook Click Rates : 2019 vs 2020 | 101 |
| 6.4 | The IFect Estimator | 101 |
| 6.5 | Robustness Tests | 103 |
| 6.5.1 | An Equivalence Test for No Pre-trend | 103 |
| 6.5.2 | Placebo Checks | 104 |
| 6.5.3 | Heterogeneous Treatment Effects | 105 |
| 6.6 | Effect by Month | 106 |
| 6.7 | Effect of Press Coverage on Clicks | 107 |
| 6.8 | Full Regression Results | 107 |
| | Conclusion et perspectives | 109 |
| 1 | Changement de conjoncture | 109 |
| 2 | Résumé des chapitres et implications | 110 |
| 3 | Pistes pour les recherches à venir | 111 |
| | Bibliographie | 113 |

Liste des figures

| | | |
|------|--|-----|
| 1 | Intermédiaires dans de la publicité programmatique (DEPARTMENT OF JUSTICE, 2023) | xiv |
| 1.1 | Google Clicks Variation by Country | 9 |
| 1.2 | Google vs Facebook Clicks through Time | 10 |
| 1.3 | Google+Facebook Ads Clicks and Offline Advertising Stock | 10 |
| 1.4 | Effect of Past Google Clicks on Media Investment by Semester | 17 |
| 1.5 | Effect of Past Semester’s Google Clicks on Media Investment by Semester | 29 |
| 2.1 | Evolution of advertising expenditures by brand | 40 |
| 2.2 | Average media mix by industry | 41 |
| 2.3 | Evolution of log(price) by media. Prices averaged by month | 41 |
| 2.4 | Cross-price elasticities of demand by brand and year | 47 |
| 2.5 | Own-price elasticities by brand and year | 48 |
| 2.6 | Cross-price elasticities and their 95% confidence intervals after deleting observations where all media are not observed. Data are grouped by industry | 54 |
| 2.7 | Cross-price elasticites with four inputs (offline brodacst, offline printed, display and search media) | 56 |
| 3.1 | Intermediaries in Programmatic Advertising (DEPARTMENT OF JUSTICE, 2023) | 60 |
| 3.2 | Density Distributions | 65 |
| 3.3 | Effect of Ad Inventory Type on Click-Through Rate Depending on whether Publisher Sells Only on One Marketplace (Open or Private) or on Both | 67 |
| 3.4 | Evolution of λ_δ | 75 |
| 4.1 | Average Number of Individuals Exposed to the Brand Ads across Age Groups per Week | 87 |
| 4.2 | Average 18-54 years old Women Exposed to the Brand Ads per Day | 88 |
| 4.3 | Evolution of Clicks across Ad Networks in 2020 | 89 |
| 4.4 | Facebook vs Other Networks Ad Outcomes : Boycott vs other Periods | 90 |
| 4.5 | Effect of the boycott on Facebook and Instagram clicks (IFeCt estimates) | 94 |
| 4.6 | Click-through Rate (CTR) by Age (Panel A) and Platforms (Panel B) | 99 |
| 4.7 | Ad Exposure (Panel A) and Click-Through-Rate (Panel B) by Gender | 100 |
| 4.8 | Evolution of Facebook ads (Facebook and Insta) click rate in 2019 and 2020 | 101 |
| 4.9 | Equivalence Test Results | 103 |
| 4.10 | Placebo estimation | 104 |

Liste des tableaux

| | | |
|------|--|----|
| 1.1 | Relationship with Related Empirical Studies | 7 |
| 1.2 | Carryover parameters estimated vs Reported in the Literature | 12 |
| 1.3 | Effect of Offline, Online and Competing Adstocks on Google and Facebook Ads Outcomes | 14 |
| 1.4 | Results of the Mechanism Analysis | 15 |
| 1.5 | Correlation Matrix | 21 |
| 1.6 | Summary Statistics by Brand | 21 |
| 1.7 | Summary Statistics by Country | 22 |
| 1.8 | First Stage Results from Equation (1.1) | 24 |
| 1.9 | Effect of Offline Adstock dA on Google Clicks η and cost ξ | 24 |
| 1.10 | Effect of Offline, Online and Competing Adstocks on Google and Facebook Ads Outcomes | 25 |
| 1.11 | Results of the Mechanism Analysis | 26 |
| 1.12 | Effect of Offline, Online and Competing Adstocks on Google and Facebook Ads Outcomes | 27 |
| 1.13 | Robustness to Fixed Effects | 28 |
| 1.14 | Robustness to Other Specifications and Estimators | 28 |
| 1.15 | Results when Controlling for Ad Impressions | 30 |
| 2.1 | Estimation of the 3SLS equation system | 44 |
| 2.2 | First-stage result | 51 |
| 2.3 | Durbin-Watson tests before and after specifying standard errors as in equation (2.9) | 52 |
| 2.4 | Results of the concavity test | 53 |
| 2.5 | Prices regressions results | 54 |
| 3.1 | Viewability rate by pricing contract | 65 |
| 3.2 | Summary Statistics of both Datasets | 66 |
| 3.3 | Econometric results : effect of pricing and programmatic buying on ad viewability | 69 |
| 3.4 | Econometric results : effect on context on click rates | 71 |
| 3.5 | Robustness test testing other specifications | 76 |
| 3.6 | Results for the Popular Fix and IHS Transformations | 77 |
| 3.7 | Econometric results : effect of pricing and programmatic buying on ad viewability | 77 |
| 3.8 | Econometric results : effect on context on click rates | 78 |
| 3.9 | Results per Device Type | 79 |
| 4.1 | Timeline of The 2020 Facebook Ad Boycott | 85 |
| 4.2 | Differences-in-Differences results : Boycott Effects | 93 |

| | | |
|------|---|-----|
| 4.3 | Differences-in-Differences Results : Moderation Effects | 95 |
| 4.4 | Effect of Media Coverage on Facebook Ad Outcomes | 96 |
| 4.5 | Summary Statistics | 98 |
| 4.6 | Placebo tests. The table show the DiD results when the treatment starts at week $t - x$ instead of t | 105 |
| 4.7 | Weights associated to the DiD model (4.1) | 105 |
| 4.8 | Effect of Month on Facebook Ad Outcomes | 106 |
| 4.9 | Effect of Media Coverage on Facebook Ad Outcomes | 107 |
| 4.10 | Differences-in-Differences results with Control Variables | 108 |
| 4.11 | Differences-in-Differences results for Moderation with Control Variables | 108 |

Introduction

« *La publicité est un aspect proéminent de la vie économique* ». C'est en ces termes que BAGWELL, 2007 introduit son *analyse économique de la publicité* : 125 pages recensant les opinions des économistes sur la publicité. Celle-ci est en effet une représentation proéminente, pour ne pas dire obsédante, de l'économie. Les dépenses mondiales en publicité ont dépassé les 800 milliards de dollars en 2022 (EMARKETER, 2023). Un institut d'étude estime que l'américain moyen rencontre au moins 4,000 messages commerciaux par jour (R. MARSHALL, 2015). Si beaucoup aimeraient éviter la publicité, ils profitent toutefois des contenus et services gratuits qu'elle finance. En tant que fait social et économique majeur, la publicité est le réceptacle de réflexions critiques, idéologiques, empiriques, mais plus rarement scientifiques.

1 L'économie de la publicité

L'économie s'est très tôt attelée à définir une théorie de la publicité. On en trouve les premières traces chez A. MARSHALL, 1919 qui distingue les publicités dites *constructives* des *combatives*. Alors que les premières informent et facilitent le choix du consommateur, les secondes martèlent le nom de la marque pour dérober des parts de marché aux concurrents. Il en découle naturellement que là où la première est souhaitable, la seconde est nuisible.

Les économistes se sont longtemps focalisés sur la dimension *persuasive* de la communication commerciale : la publicité est accusée de manipuler les préférences du consommateur au profit du produit annoncé (BRAITHWAITE, 1928). La publicité engendre des effets anti-concurrentiels : réduction de l'élasticité-prix, barrières à l'entrée du marché, loyauté des consommateurs à la marque ou encore concentration des plus grandes firmes. Des études économétriques tendront à confirmer que la publicité favorise les firmes en place sur un marché (COMANOR et WILSON, 1974).

À partir des années 1960, l'économie prend la mesure de l'information et de son poids dans l'organisation des marchés (OZGA, 1960). Pour STIGLER, 1961, les consommateurs ne sont pas totalement informés et doivent engager des efforts de recherche pour connaître le prix des produits sur le marché. La publicité atténue les coûts de recherche en fournissant directement l'information au consommateur. Dans son acception informative, la publicité est donc socialement souhaitable car elle éclaire le choix du consommateur et stimule la concurrence. Cela dit, on peut objecter que beaucoup de publicités ne sont pas informatives. Pour NELSON, 1974, la publicité est toujours indirectement informative : la décision d'une firme de communiquer sur ses produits agit comme un signal auprès du consommateur. Par exemple les firmes efficaces – celles

qui réalisent le plus d'économies d'échelle – ont plus intérêt à communiquer que les autres¹.

La théorie persuasive n'explique pas pourquoi ni comment le consommateur se laisse manipuler par la publicité. Quant à la théorie informative, elle peine à justifier la raison d'être de certaines publicités. L'analyse novatrice de BECKER et MURPHY, 1993 répond à ces deux limites en considérant la publicité comme un bien en soi. Celui-ci est complémentaire au produit vendu : si le consommateur trouve la publicité utile, elle augmente la valeur du produit. À l'inverse, s'il ne l'apprécie pas, la publicité est un mal qui réduit sa demande pour le bien annoncé. Cette analyse a l'avantage de traiter les individus comme rationnels. La publicité entre alors dans la fonction d'utilité du consommateur, lequel est libre de l'apprécier au même titre que les autres produits (STIGLER et BECKER, 1977). Sachant que différents groupes de consommateurs n'apprécient pas la publicité de la même manière, celle-ci est un moyen de discrimination par les prix.

2 Médias et publicité : une économie des récits

La théorie de BECKER et MURPHY, 1993 inscrit la publicité dans l'économie des médias. D'une part, la publicité est une marchandise vendue à prix nul, voire négatif lorsqu'elle subventionne les émissions et contenus éditoriaux auxquels elle s'adosse. D'autre part, la publicité est elle-même un média. Elle en partage du moins les caractéristiques économiques. Pour les firmes, la production d'une publicité mobilise d'importants coûts fixes échoués. S'agissant des consommateurs, la publicité est un bien d'expérience : sa valeur n'est connue qu'après l'avoir vue. En conséquence, l'effet de la publicité est incertain. La publicité peut magnifier ou dégrader l'image du produit ; amuser comme agacer le consommateur. Au même titre que les produits culturels tels les livres, musiques ou films, la publicité est un *nobody knows product* (CAVES, 2000).

La difficulté à caractériser l'utilité d'une publicité provient de sa dimension narrative : la pub produit de l'information, certes, mais aussi des récits. Si l'économie a pris la mesure de l'information, elle fait peu de cas des récits. Et pour cause, les deux catégories s'opposent fondamentalement. Si l'information revêt un caractère objectif, les récits sont sujets à des valorisations hétérogènes et subjectives. Il n'existe pas de critère unanime pour les hiérarchiser : les récits sont différenciés horizontalement. SHILLER, 2017 est un des premiers économistes à consacrer une étude à l'économie des récits. Il y postule qu'un récit est l'expression d'une histoire ou d'une suite d'évènements, dont la diffusion importe à l'économiste. ELIAZ et SPIEGLER, 2020 vont plus loin en définissant les récits comme des explications causales entre des évènements. Leur rôle est de susciter des corrélations de long-terme. Par exemple, l'idée que le commerce extérieur engendrerait du chômage est un récit dont le but est de relier les variables « importations » et « destruction d'emplois ».

L'idée s'applique volontiers à la publicité. Les récits mis en scène dans la communication

1. Le raisonnement est que les firmes avec les plus fortes économies d'échelle ont le plus intérêt à augmenter leur demande par la publicité. Le consommateur gagne à acheter auprès de ces firmes car, étant donné l'efficacité de leur structure de production, elles offrent les prix les plus intéressants. L'argument est contestable. Comme BAGWELL, 2007 le fait remarquer, les marques les moins intéressantes ont intérêt à faire plus de publicité pour se différencier de concurrents mieux placés. En outre, les produits low-cost bénéficient tout autant de la publicité en raison de leur coûts marginaux faibles.

des firmes ont pour but d'associer un produit avec une idée, ou une représentation : qualité, performance, prestige, réussite sociale, amitié, beauté... Dans sa campagne « *Tout pour la daronne* », Winamax crée une relation entre son service de pari sportif, la réussite sociale et l'esprit de famille. Au plan économique, les associations de long-terme tissées par les récits publicitaires sont internalisées par la marque. L'identité d'une marque (son nom, son logo) catalyse les dépenses d'image et empêche d'autres d'en bénéficier (LANDES et POSNER, 2003)

3 Publicité en ligne : organisation industrielle

La dimension créative de la publicité (coûts échoués et incertitude sur la demande) sécrète une organisation industrielle particulière. Pour être certain qu'une campagne rapporte plus qu'elle ne coûte, les marques ont recours à des études de marché *ex-ante*, elles optimisent la création et le placement de leurs messages et en calculent l'efficacité *ex-post*. Les instituts de sondage, agences média et firmes de mesure concourent à assurer une réception optimale de la publicité.

3.1 Internet : quoi de neuf ?

La publicité en ligne s'appuie sur une innovation fondamentale : la baisse des coûts de ciblage du consommateur (GOLDFARB, 2014). Celle-ci rebat les cartes de l'industrie. À l'étude du contenu publicitaire, se substituent les débats sur le ciblage des campagnes. La cible, plus que le message, apparaît alors comme le facteur déterminant de l'efficacité d'une publicité.

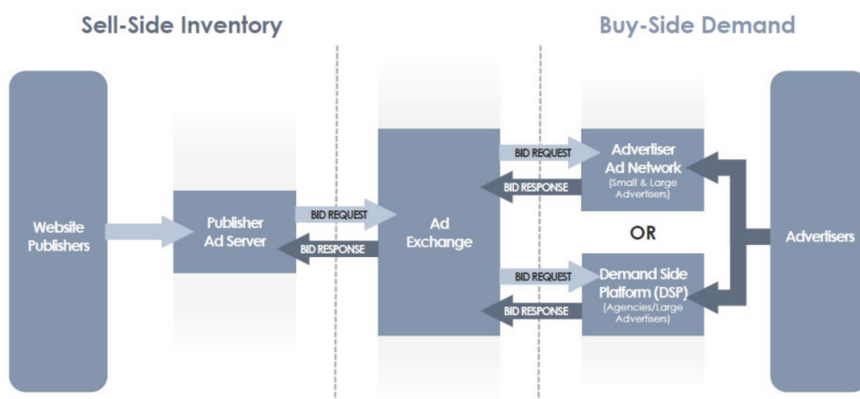
L'abaissement des coûts de ciblage bouleverse également les enjeux de mesure. Annonceurs et régies peuvent dorénavant observer la réaction du consommateur à un message : clic, requête, achat immédiat ou différé. Nourris des données d'utilisateurs, les algorithmes de ciblage sont également capables de calculer *ex-ante* la probabilité de clic ou d'achat d'un utilisateur (nous y reviendrons).

La baisse des coûts de ciblage induit deux transformations fondamentales de l'industrie publicitaire : le recours aux enchères et aux intermédiaires.

En ligne, un même espace publicitaire peut être vendu à plusieurs annonceurs selon l'audience ciblée par ceux-ci. Le ciblage peut s'opérer selon le contexte (contenu de la page web), la démographie de l'utilisateur, son comportement de navigation ou selon son historique d'achat. Du point de vue de l'organisation industrielle, l'enjeu est de parvenir à tarifier tous ces espaces à leur juste valeur, sachant que le consentement à payer des annonceurs est très hétérogène selon les cibles en jeu (LEVIN et MILGROM, 2010). Les enchères, par lesquelles les annonceurs dévoilent leur valorisation pour chaque segment ciblées, s'imposent comme le modèle dominant d'allocation de la publicité en ligne.

La multiplication des espaces rend caduque la vente en gré à gré issue de la publicité traditionnelle. Sur internet, annonceurs et éditeurs ont recours à des transactions « programmatiques » appelant une série d'intermédiaires techniques chargés de délivrer le message de l'annonceur à l'utilisateur final : *ad servers, demand/supply-side platforms, ad exchanges, ad networks, data management platforms...*

FIGURE 1 – Intermédiaires dans de la publicité programmatique (DEPARTMENT OF JUSTICE, 2023)



L’agencement de ce réseau d’intermédiaires est cartographié dans la Figure 1. Plutôt que de fournir une description exhaustive des différents acteurs, je m’attarderai sur les conséquences économiques de cette organisation industrielle.

3.2 Splendeurs et misères de la publicité en ligne

D’abord, la présence d’intermédiaires spécialisés le long de la chaîne réduit les coûts de transaction entre annonceurs et éditeurs. Certains intermédiaires permettent par exemple à un annonceur de se connecter à de nouveaux inventaires. D’autres fournissent aux éditeurs des données leur permettant de vendre des segments d’audience aux annonceurs. En somme, sans cette chaîne, il serait impossible d’organiser autant d’enchères en si peu de temps pour tous les espaces et segments demandés par les annonceurs. Toutefois, les problèmes d’incitation et d’opacité constituent le revers de la médaille de cette industrie.

Intérêts divergents En effet, les intermédiaires ne sont pas toujours incités à agir dans l’intérêt de leur délégataire. Par exemple, un intermédiaire en haut de chaîne (coté annonceur) peut refuser d’enchérir pour l’annonceur si celui-ci ne mise pas assez (BALSEIRO et al., 2021). Ce résultat surprenant provient de ce que plus un intermédiaire est en haut de la chaîne, plus il devra rémunérer les suivants.

De plus, les modèles tarifaires des intermédiaires incitent souvent ceux-ci à biaiser le ciblage de la publicité. Par exemple, Facebook cible les publicités aux utilisateurs les mieux à même d’enregistrer l’objectif défini par l’annonceur (GORDON, ZETTELMEYER et al., 2019 ; GORDON, MOAKLER et al., 2022). Ainsi, sur les bannières reciblées, FRICK et al., 2022 montrent-ils qu’un intermédiaire payé à l’achat adressera la publicité à des consommateurs qui auraient acheté le produit quoi qu’il arrive. Ce biais de sélection interroge l’effet réel de la publicité en ligne et son utilité pour les annonceurs. Elle soulève également des enjeux de mesure : en ligne, l’exposition à la publicité est endogène à son efficacité (G. JOHNSON, 2022). Dans certains cas, la sélection opérée par les plateformes mène à des discriminations sur la base de stéréotypes (CECERE, JEAN, MANANT et al., 2018 ; CECERE, JEAN, LE GUEL et al., 2020).

Manque de transparence L’opacité est le second handicap de la publicité en ligne. L’enchevêtrement d’intermédiaires rend peu transparent le placement final du message. Cela entraîne des problèmes de fraude publicitaire et de contextualisation.

Le terme d’*ad fraud* désigne l’ensemble des techniques destinées à tromper l’annonceur et gaspiller ses dépenses publicitaires. Puisque la publicité en ligne est souvent rémunérée au clic ou à la page vue, des parties ont intérêt à manipuler ces indicateurs. La fraude peut concerner la visibilité des publicités : les bannières et vidéos (publicité *display*) sont empilées les unes sur les autres ou placées en des endroits incongrus de la page web (BOUNIE et al., 2017). De même, certains sites peuvent stimuler du trafic non-humain pour gonfler les vues facturées aux annonceurs (GORDON, JERATH et al., 2021). Dans la publicité sur les moteurs de recherche (dîtes *search*) la fraude porte davantage sur des clics frauduleux engendrés par des robots (WILBUR et Y. ZHU, 2009).

Enfin, une source importante d’opacité porte sur le contexte dans lequel la publicité apparaît. Parce qu’ils ciblent des populations plutôt que des contextes, les annonceurs savent rarement sur quel site ou page web leur publicité est affichée. En conséquence, les marques risquent d’abîmer leur image en s’associant à des contenus ou des éditeurs controversés (SHEHU et al., 2021). Le contexte dans lequel la publicité se met en scène structure son récit. Les annonceurs n’hésitent plus à boycotter les plateformes en ligne lorsque celles-ci peinent à maîtriser le contexte d’affichage de leurs publicités. L’*Adpocalypse* de YouTube en 2017, le boycott #StopHateForProfit de Facebook en 2020 ou encore la fuite des annonceurs de Twitter fin 2022 en sont des exemples récents.

Concentration et antitrust Le marché de la publicité en ligne est souvent critiqué pour sa concentration. En France, le duopole Google et Facebook capte en effet 75% des dépenses publicitaires en ligne (PERROT et al., 2022). En sus d’être des plateformes de service, Google, Facebook ou encore Amazon sont des régies publicitaires intégrant tout ou partie de l’écosystème décrit en Figure 1.

Parce qu’elles intègrent verticalement la vente d’espace, les grandes plateformes peuvent s’ériger comme des goulots d’étranglement (BACACHE-BEAUVALLET et BOURREAU, 2022) pour les annonceurs souhaitant accéder aux consommateurs. Les résultats théoriques et empiriques montrent ainsi que plus une plateforme possède une audience captive, plus elle est à même d’imposer ses prix aux annonceurs (S. P. ANDERSON et al., 2018 ; GENTZKOW et al., 2022).

En outre, l’intégration des plateformes et intermédiaires publicitaires permet aux premières de vendre leur inventaire exclusivement par le biais de leurs outils publicitaire : YouTube est uniquement distribué par les outils de Google, Twitch par la régie Amazon etc. Enfin, parce qu’elles mesurent elles-mêmes l’efficacité de leurs publicités, les plateformes sont juge et partie.

Ces éléments sont au cœur des préoccupations du régulateur (en France, voir par exemple PERROT et al., 2022). D’autant que la régulation de la publicité en ligne s’inscrit plus généralement dans l’encadrement des grandes plateformes en ligne (BOURREAU et PERROT, 2020 ; O. BOMSEL et DEVAUX, 2022). Le RGPD, Digital Market Act, ou encore le récent procès de Google intenté par le Department of Justice Américain sont des exemples récents de régulation

impactant fortement la publicité en ligne.

4 Contributions de la thèse

Cette thèse rassemble quatre études empiriques sur la publicité en ligne. Chaque chapitre s'appuie sur les données d'annonceurs obtenues grâce au partenariat avec *Ekimetrics*. Les chapitres 1 et 2 de la thèse sont dédiés aux interactions entre les marchés publicitaires hors-ligne et numériques. Les chapitres 3 et 4 se concentrent sur le fonctionnement de la publicité en ligne plus précisément.

Dans le **Chapitre 1**, j'étudie l'effet des campagnes publicitaires hors-ligne et numériques sur la performance des liens sponsorisés (*search*) des firmes. J'utilise un modèle de régression linéaire avec effets fixes pour estimer l'effet du stock de publicité (hors-ligne et en ligne) d'une marque sur les clics enregistrés par ses publicités *search*. L'endogénéité entre les dépenses publicitaires et la réaction de la demande (RUTZ et WATSON, 2019) est atténuée par des variables instrumentales utilisant les dépenses publicitaires hors-ligne et numérique dans d'autres régions. Je trouve qu'en augmentant son stock de publicité hors-ligne d'1% (soit en moyenne 720€), les marques génèrent +0.43% clics sur leurs liens sponsorisés (ce qui correspond en moyenne à 134 clics). Une analyse additionnelle montre que ces clics viennent de nouvelles requêtes générées par les utilisateurs. En augmentant le volume de clics sur les liens sponsorisés, la publicité hors-ligne en augmente également le prix, étant donné que les espaces sont tarifés au coût-par-clic. Les campagnes numériques de la marque en revanche tendent à réduire les performances des publicités *search*.

Ces résultats montrent que lorsqu'une firme investit en publicité hors-ligne, elle rend sa marque plus attractive et donc plus recherchée sur les moteurs de recherche. Ce mécanisme profite in fine à Google qui, payé à la performance, tarifie davantage l'annonceur. En tant que goulot d'étranglement sur la recherche en ligne, Google bénéficie donc des efforts de promotion des marques, y compris lorsque ceux-ci sont générés hors de sa plateforme. Nous ne trouvons pas de tels effets sur les campagnes Facebook par exemple.

Le **Chapitre 2** étudie la capacité des annonceurs à substituer leurs dépenses publicitaires hors-ligne et numériques. Une partie de la littérature académique tend à démontrer que publicités hors-ligne et numériques sont substituables du point de vue des annonceurs (GOLDFARB et TUCKER, 2011e). Cependant, les arguments basés sur les élasticités prix des firmes à la publicité sont rares. Dans ce chapitre, je construis une fonction de demande des firmes en publicité, puis j'en dérive des élasticités propres et croisées de trois médias : le hors-ligne, le display (bannières et vidéos) et le *search* (liens sponsorisés). J'estime ce modèle sur les données d'achat d'espace de neuf grands annonceurs français répartis dans trois industries différentes (hôtellerie, boissons non-alcoolisées et télécommunications). Les possibilités de substitution entre médias diffèrent selon l'industrie. Globalement, publicités hors-ligne et numériques sont des substituts limités. Les annonceurs sont inélastiques au prix des publicités hors-ligne, lesquelles semblent être un actif essentiel des marques. L'élasticité au prix du *search* est également modérée en raison du pouvoir de marché de Google. J'en conclus que le marché publicitaire est très fragmenté, à la fois entre les industries et médias.

Le **Chapitre 3** s'attaque aux asymétries d'information dans le marché de la publicité display. Comme souligné précédemment, beaucoup de publicités ne sont jamais vues en raison d'un mauvais placement ou parce qu'elles sont exposées à un trafic non-humain. Certains espaces sont également adossés à des contenus non-pertinents ou dégradants pour la marque et le produit. La plupart du temps, les annonceurs ne savent pas où leur publicité apparaît et si elle est effectivement vue. Dans ce chapitre, j'utilise deux bases de données pour étudier comment les contrats, méthodes d'achat ou techniques de ciblage affectent l'efficacité des publicités display. Dans une première étude, j'utilise les données de 2300 espaces display achetés par une marque de produit paramédicaux. À partir d'une analyse de régression, je montre que la tarification des contrats au coût-par-impression (CPM) et le recours à l'achat programmatique réduisent tous les deux la visibilité des publicités. Ce résultat est intuitif : en CPM, la régie n'est pas incitée à rendre la publicité visible puisqu'elle est payée quoi qu'il arrive. Quant à l'achat programmatique, les intermédiaires privilégient souvent le prix de l'espace au détriment du placement de la publicité.

Dans une seconde étude, j'estime l'efficacité du ciblage contextuel. À partir de données de campagnes d'une marque de parfum, je trouve que l'adéquation des publicités au contenu des pages résulte dans de meilleurs taux de clic que le ciblage des consommateurs. Comme beaucoup d'annonces ne sont jamais cliquées (39%), j'utilise le nouvel estimateur de BELLÉGO et al., 2022 prenant en compte les zéros dans les régression log-linéaires.

Enfin, le Chapitre 4 étudie l'effet des contenus controversés sur l'efficacité des publicités en ligne. En juillet 2020, plus de 1,000 grandes marques ont suspendu leurs campagnes sur Facebook et Instagram du fait d'un trop grand nombre de contenus polémiques sur la plateforme. En utilisant les données d'une marque n'ayant pas boycotté les publicités Facebook, nous estimons l'effet des contenus polémiques et du boycott sur ses publicités. Les résultats d'une estimation en double différences (DiD) font état d'une chute significative des clics et du prix des publicités Facebook pendant le boycott, par rapport aux autres campagnes *display* de la marque. Afin d'estimer l'effet continu des contenus et du boycott, j'utilise l'estimateur contrefactuel de L. LIU et al., 2022 basé sur un modèle à effets fixes interactifs (BAI, 2009). Les résultats suggèrent que les clics et le prix des campagnes débute en juin (effet des contenus polémiques) puis s'accroît particulièrement en juillet, le mois où les marques ont boycotté la plateforme. Le mois de juillet correspond également au mois où le boycott a été le plus médiatisé. Nous trouvons en effet une corrélation négative entre le nombre d'articles de presse publiés sur le boycott et l'efficacité et le prix des espaces. Cette étude conclut que les contenus d'un média, mais surtout sa crédibilité, sa marque éditoriale, affectent l'effet et le prix des publicités.

Chapitre 1

Externalities across Advertising Markets

Abstract

This chapter investigates externalities generated by offline advertising campaigns on the performance of online advertisements of the same brand. Using advertising data on a panel of firms in the hotel industry, we estimate how offline, display, and competing ad campaigns impact the effectiveness of Google and Facebook ads. We find a positive effect of traditional mass-media campaigns on Google clicks. Advertising from competitors does not affect Google ads performance but it increases their prices, suggesting keyword poaching. Further analyses hint that as a gatekeeper over online search, Google free-rides on externalities generated by other advertising media. We do not find that Facebook ads benefit from such externalities.

Résumé

Ce chapitre étudie les externalités de la publicité hors ligne sur la performances des publicités numérique. À partir de données d'un panel d'annonceurs dans l'hôtellerie, nous estimons l'impact des campagnes publicitaires hors ligne, en ligne et de communication sur l'efficacité des publicités Google et Facebook des marques. Nous trouvons un effet positif des campagnes traditionnelles sur les clics des publicités Google. La publicité des concurrents n'impacte pas la performance des annonces Google, mais elle en augmente leurs prix. Cela suggère que les marques « braconnent » les mot-clés de leur concurrent sur Google. Des analyses plus approfondies indiquent qu'en tant que goulet d'étranglement sur la recherche en ligne, Google profite des externalités générées par les efforts de promotion des marques. Nous ne trouvons pas que les publicités Facebook bénéficient de telles externalités.

1 Introduction

Online advertising now accounts for the majority of media spending : with a growth rate of 29% in 2021, it amounted to 63% of total advertising expenditure (EMARKETER, 2021). A number of economic studies have argued that online advertisements generate efficiency by dramatically decreasing the costs of targeting consumers and measuring ad effects¹. However, these arguments may fail to consider externalities between advertising media. For example, offline ad campaigns may affect online ads outcomes by providing information and narratives on the product to a large mass of consumers. This idea is consistent with the complementary view of advertising in which advertising increases the product's utility for consumers who value the ad (BECKER et MURPHY, 1993). As a complement, advertising can raise the desirability of a product or a brand, hence generating externalities on online behaviors such as product search, ad exposure or clicks.

Such externalities are important. On the one hand, online advertising slots are often priced based on effectiveness metrics such as clicks or purchases. Yet, these outcomes may be partly generated by offline media campaigns that raise awareness for a brand or its product. On the other hand, the existence of significant externalities between both media may suggest that offline and online advertising are two distincts, complementary, markets. The definition of a relevant advertising market is still in debate in the literature and underpins the action of regulators on questions such as offline media mergers (CROSS, 2022).

While cross-media effects have been widely studied in the advertising literature, we treat the question from an industrial organization perspective. Indeed, since online ads are often priced on the basis of effectiveness metrics such as clicks or conversions, any effect of offline media on online ads outcomes would entail changes in online ad platforms' revenues. Because vertically integrated ad platforms such as Google or Facebook have a strong market share over single-homing users, they may act as competitive bottlenecks for brands and thus free-ride on their promotional efforts.

In this study, we quantify the existence of externalities across advertising media in a special industrial context : the market for hotels. We leverage firm-level data from five advertisers belonging to an international hotel group to study how a brand's offline and online display campaigns impact Google and Facebook advertising outcomes. We also consider the effect of competitors' ad spending. Using a fixed-effect regression with instrumental variables, we find offline investments have a positive impact on the effectiveness of Google search ads. For example, increasing the stock of offline advertising by 1% increases clicks on Google ads by 0.67%. In addition, we find a negative effect of display ads on Google clicks, suggesting that online ads compete for users' attention. Similar results are found for Facebook ads but they remain statistically insignificant.

The presence of offline-to-online effects opens the path to a more important question : who benefits from such externalities ? Further analyses show that by increasing the volume of searches and the propensity to click, offline advertising increases the overall cost allocated to Google ads. In the long run, the increase in Google advertising performance (clicks) negatively affects

1. On the economic nature of online advertising, see for example (GOLDFARB, 2014)

the offline share of advertising budget. Although they do not decrease clicks on search ads, competitors' ad spendings rise their prices, suggesting that firms compete in auctions to buy their competitors' branded keywords (SAYEDI et al., 2014; DESAI et al., 2014; SIMONOV et al., 2018). This strategy, known as *brand poaching*, consists in a brand buying a well-known competitor's Google keyword in order to free-ride its notoriety. For example, a London-based hotel brand could buy the keyword "Airbnb London" to appear in the latter's search results. Or vice-versa.

The literature on offline-online advertising effects is abundant and our contribution is both conceptual and empirical. First, the study demonstrates the existence of offline-to-online externalities affecting not only advertising performances, but also ad prices, and the media budget share. Second, to the best of our knowledge, we are the first to document the joint effect of offline ads on both Google and Facebook – two large online advertising networks. Third, we study the simultaneous externalities generated by offline, online and competing ads by using a novel instrumental variable which exploits advertising on foreign markets.

Our results have several implications. (i) First, they suggest that online advertising's return on investments (ROI) may be biased in the presence of externalities between offline and online ads. Given the positive effect of traditional media campaigns on search advertising outcomes, the effectiveness of the latter is likely to be over-estimated. (ii) Second, as an online search monopoly, Google seems able to free-ride on such externalities. Indeed, the increase in queries and clicks generated by offline ads translates into additional revenues for Google since search ads are priced based on clicks (cost-per-click model). (iii) Third, brand poaching creates a prisoner dilemma for advertisers, resulting in increased search advertising costs. Brand poaching is not Google's privilege as other search platforms like Amazon allow advertisers to buy branded keywords. We argue that this strategy should be regulated. Points (ii) and (iii) are related to the literature on competitive bottlenecks in media markets : Google have a monopoly power over advertisers looking to reach Google's single-homing users (ARMSTRONG, 2006 ; KAISER et WRIGHT, 2006 ; PRAT et VALLETTI, 2022). By gatekeeping access to consumers, platforms such as Google or Amazon benefit from a brand's promotional efforts and even allow its competitors to free-ride it through keyword-poaching. (iv) Finally, this study could suggest that offline and online advertising are complements rather than substitutes. While offline campaigns provide information and narratives to a mass of consumers, online search ads guide consumers toward the purchase.

Although we only identify offline-to-online externalities in the hotel industry, our results are consistent with similar studies in other sectors. In particular, we believe that such externalities exist in any industry where firms heavily advertise offline to consumers searching and/or buying online (e.g. apparel, electronics, events). However, the magnitude and direction of the effects may be different across industrial contexts.

The present paper is organized as follows. In Section 2, we review the past literature on advertising externalities, cross-media effects and their impact on the advertising industry. Section 3 sums up the research question and introduces the data used. Section 4 presents some descriptive evidence followed by the econometric methodology used to identify the presence of cross-media

externalities. The results and mechanisms are discussed in detail in Section 5, while Section 6 addresses the implications for the advertising market.

2 Related Works

2.1 Advertising Externalities

Advertising is an important vector of externalities. In their model, BECKER et MURPHY, 1993 consider advertisement as a good which is complementary to the advertised product. Advertising increases or decreases the utility for the product depending on consumers' taste for the ad. Advertising increases product's attractiveness for consumers who value the ad. This shift in consumer's utility may generate in additional product sales. However, the increase in product's valuation may also result in effects that are not internalized in transactions.

For example, TV advertising campaigns may induce consumers to propagate word-of-mouth which in turn increases consumer's awareness and consideration for the product (ONISHI et MANCHANDA, 2012; FOSSEN et SCHWEIDEL, 2017). LEWIS et NGUYEN, 2015 found displaying an advertising banner on Yahoo!'s homepage increased search queries for the brand advertised and its competitors in the insurance and tablet markets. It also raised clicks on complementary services such as online distributor or review sites.

2.2 Cross-media Effects

While advertising acts as a complementary commodity to the advertised product and produces externalities, it can be distributed offline as well as online. The relevant market definition of advertising is still in debate in the literature. While both offline and online advertising ensure the common economic function of providing information and narratives about products, they employ different targeting, pricing and measurement technologies (EVANS, 2009; GOLDFARB, 2014). On the one hand, a body of early research relied on theoretical (BERGEMANN et BONATTI, 2011) and experimental (GOLDFARB et TUCKER, 2011e; GOLDFARB et TUCKER, 2011d) settings to demonstrate that offline and online ads were substitutes. On the other hand, many advertiser-level studies have tended to demonstrate the existence of positive cross-media effects in generating demand, which supports the view that these are complements.

NAIK et PETERS, 2009 provide a comprehensive review of cross-media effects. Furthermore, they show that, in the case of a car manufacturer, offline and online ads generated higher returns when they were released simultaneously. Other studies have focused on the effect of television advertising on online search outcomes. Descriptive researches based on Google queries found that television advertisements aired during highly watched live sports events generated immediate surges in searches for the brands and products advertised (ZIGMOND et STIPP, 2010; LEWIS et REILEY, 2013). Going further in this analysis, JOO, WILBUR, COWGILL et al., 2014; JOO, WILBUR et Y. ZHU, 2016 empirically demonstrated that television advertising for financial services resulted in immediate queries for the brands advertised, while decreasing generic queries. Adding sales to the equation, other studies have quantified a positive effect of television adver-

tisements on online searches, traffic and purchases, with strong heterogeneous effects depending on the advertising content (LIAUKONYTE et al., 2015; GUITART et STREMERSCHE, 2021). More aggregated analysis has found that offline media generated online sales by increasing search ad impressions (i.e. more queries) and online purchases for a high-end clothing retailer (DINNER et al., 2014).

Fewer studies have investigated the effect of television advertising on the effectiveness of social media ads. Whereas TV advertising increases the effectiveness of unpaid social media posts, it does not enhance the performance of paid ads (A. KUMAR et al., 2016; V. KUMAR et al., 2017).

All of these empirical studies suggest that online media outcomes are affected by offline advertising externalities.

2.3 Advertisers and Publishers' Profit

Cross-media externalities yield implications for both advertisers and publishers. On the advertiser-side, they impact advertising strategies. Media may produce external effects that ultimately benefit other media. This assumption has key implications for advertising media competition. Indeed, advertisers may end up allocating most of their budget to media that free-ride other media's external effects. This problem is exacerbated online, where advertisers are charged for each consumer's response to their ad (e.g. cost-per-click (CPC) or cost-per-acquisition (CPA) pricing).

A wide range of literature on attribution models considers that advertisers' media investments are a function of their previous ad effectiveness measures. For example, JORDAN et al., 2011 show that when an advertiser buys impressions from multiple publishers and does not consider externalities between ads, it ends up allocating most of its budget to publishers closer to the demand². Similarly, BERMAN, 2018 finds that when externalities exist between publishers, an advertiser's chosen attribution model constitutes a strategic choice that directly impacts both its own profit and that of the publishers. An empirical descriptive analysis has also proved that advertisers' attribution modeling has an impact on ad prices and in fine on consumer welfare (TUCKER, 2013).

However, these attribution studies only consider externalities between online ads. Moreover, they do not study the case of asymmetric pricing among ads generating externalities. For example, let us consider the case of an advertiser purchasing a television campaign priced on its expected audience and a search slot priced on a cost-per-click (CPC) basis. By generating search clicks, the television advertisement would simultaneously increase the effectiveness and the cost of the search engine advertisement and revenues. Meanwhile, by increasing advertisers' notoriety, television encourages competitors to poach the keywords used by the brands advertised (SAYEDI et al., 2014). Eventually, the advertiser could end up either losing its search paid slot or keeping it for a higher advertising cost. In both cases, the search engine benefits from the higher competition in the auction. In this paper, we consider cross-media effects in a context where the media (offline and online) pricing model is asymmetric in two dimensions. Both the

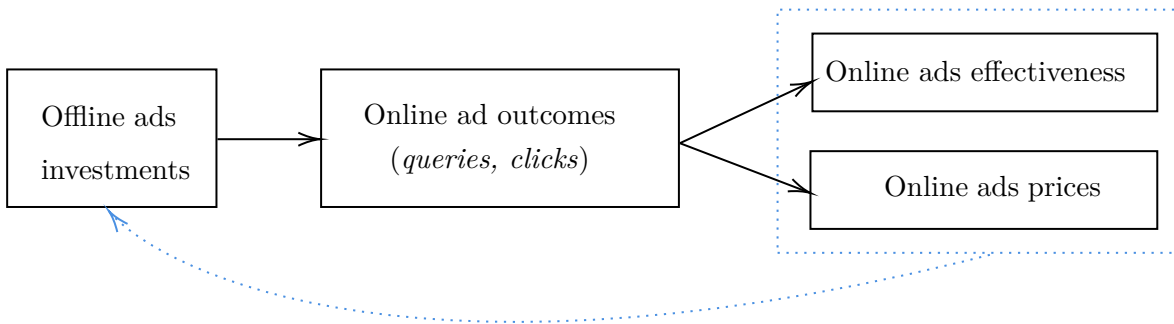
2. This phenomenon is known as last-touch attribution.

commodity sold (audience vs performance) and the allocation design (over-the-counter contract vs auctions) differ here.

3 Research Hypotheses & Data

3.1 Conceptual Framework

This paper fits in with the literature on cross-media effects by studying how advertisements on a given media impact the effectiveness of other media. These effects are considered as externalities. Indeed, referring to cross-media effects as "synergies" supposes that the benefit is shared between the entities which generate the effect. As highlighted in the literature, a brand's television advertisement is likely to affect search queries and sponsored link clicks, increasing the search engine's revenues. Similarly, when a brand's display ad fosters search queries for its competitors (LEWIS et NGUYEN, 2015), the effect is not internalized by any contract.



Rather than focusing on online sales or conversions, we analyze the key internalization mechanism of cross-media externalities, i.e. clicks. Indeed, the media synergy literature suggests that both online and offline media only generate externalities in favor of the advertisers. In doing so, it ignores the fact that while offline media are remunerated on an audience model (e.g. GRP), online media revenues directly depend on their outcomes (e.g. clicks). Hence, the effect of cross-media externalities on online advertising prices is rarely dealt with appropriately. In the next section, we present the data used to assess these effects.

3.2 Data Sources

We use advertising data from three brand pertaining to a global hotel group. Each of the three brands carries specific brand elements and their membership of the group is not signaled in their name or their logo. They differ in both price and quality : we distinguish between low-cost, mid-range⁻ and mid-range⁺ brands. The distinction between mid-range chains depends on the location and room prices. The mid-range⁺ brand offers more expensive rooms in locations closer to places of interest (e.g. downtown, airports).

The brands operate in two countries : the United Kingdom and Germany. However, the mid-range⁻ brand has a very low advertising activity in Germany³ and is thus excluded from

3. In addition to a low level of spending, the firm does not advertise offline in Germany at all. This reduced advertising activity is due to a low number of hotels supplied by the chain abroad.

TABLE 1.1 – Relationship with Related Empirical Studies

| Authors | Media of interest | | | Effect modeled | | | Features | |
|-----------------------------|-------------------|--------|---------|----------------|---------|---------|---|--|
| | Search | Social | Offline | Competitors | Adstock | Pricing | Mechanism | |
| NAIK et PETERS, 2009 | No | No | Yes | No | No | No | No | |
| DINNER et al., 2014 | Yes | No | Yes | Yes | Yes | No | No | |
| LIAUKONYTE et al., 2015 | Yes | No | TV only | No | No | No | No | |
| JOO, WILBUR et Y. ZHU, 2016 | Yes | No | TV only | No | Yes | No | Effect on TV on search, keyword choice and click | |
| A. KUMAR et al., 2016 | No | Yes | TV only | No | Yes | No | Effect on TV \times social media on consumer spending, buying behavior and profit | |
| KIREYEV et al., 2016 | Yes | No | No | No | No | Yes | Effect on display and search ads on each others impressions, clicks, and cost | |
| This study | Yes | Yes | Yes | Yes | Yes | Yes | Effect of offline ads on Google click propensity, queries and budget | |

Note : the "Pricing" column indicates whether the study investigated the impact of media synergies on the online advertising cost. The last column "Mechanism" refers to whether the paper explained the mechanism by which cross-media effects appear : e.g. do offline ads increase search volumes or the propensity to click ? Does it increase the volume of online sales or the average purchase value ?

the analysis. We end up with 5 firms : 2 brands in both countries and 1 brand in the United Kingdom only. All variables are reported on a weekly basis from January 18th 2016 to September 2nd 2019. Thus, the dataset is organized along three dimensions : brand (b) \times country (c) \times time (t) for a total of 950 observations.

The hotel industry is particularly suited to the study of offline-to-online effects. Indeed, while large hotel chains significantly advertise offline, consumers mainly search and book rooms online. Approximately two-thirds of travel industry revenues are generated online (STATISTA, 2021a). Thus, advertising campaigns simultaneously take place on several media, and offline ads are very likely to generate externalities online. These characteristics of the hotel industry make it easier to measure meaningful effects.

Media Spending The dataset provides ad expenditure in euros for several media – both offline (TV, radio, cinema, press, outdoor) and online (display and video advertisements). Google search and Facebook ad budgets are excluded from online investments. The aggregated media mix is generally balanced between offline and online media campaigns in the amount invested. However, both media exhibit different investment patterns : while offline campaigns often take place in specific periods, online campaigns are conducted throughout the year. Thus, the fact that offline investments are often null leads to a lower share of advertising budget in firms’ media-mix.

Google and Facebook Data Using data collected from Google and Facebook’s respective advertising tool, we retrieve the number of clicks and impressions recorded by consumers living in the firm’s country. We will exploit this specificity further in the identification strategy. While Google data are available for all brands in both countries for most of the period ($N=890$), Facebook data are only reported for the two German hotel brands ($N=269$).

Competitors data Competitors’ weekly spending by country and media are retrieved from a Nielsen database for each brand in each country.

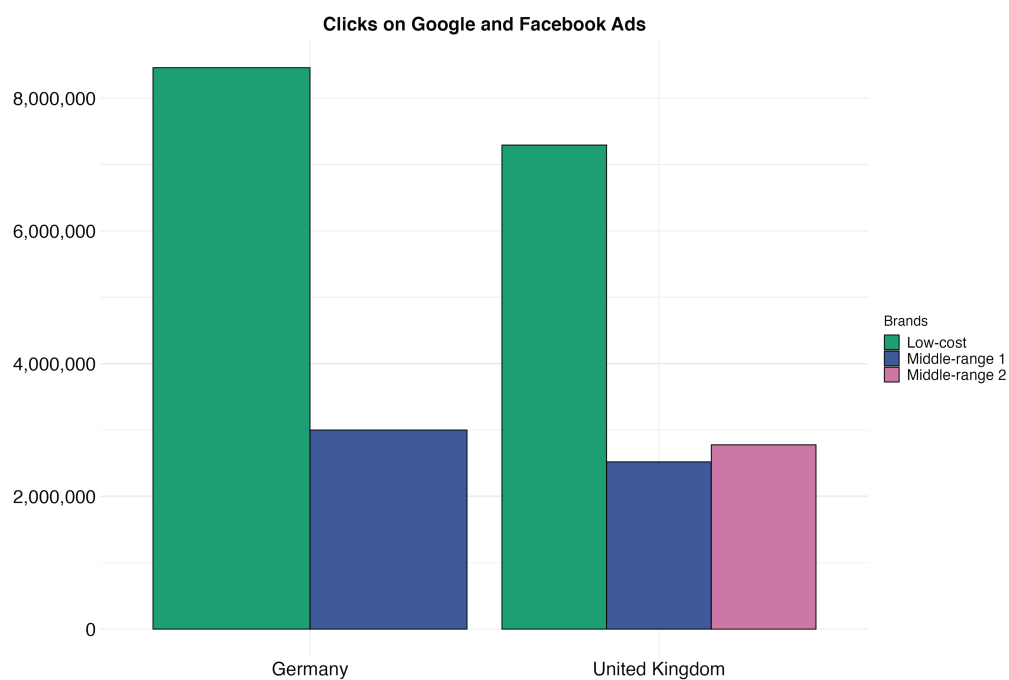
3.3 Descriptive Evidences

The main summary statistics are reported by brand in Table 1.6 and by country in Table 1.7. Demand and advertising variables exhibit strong standard deviations, suggesting significant heterogeneity among observations. As depicted in Figure 1.1, the difference in clicks across countries is moderate, however it differs strongly across brands (especially low-cost vs mid-range ones). Seasonal patterns are also important as shown in Figure 1.2.

Figure 1.3 is more directly related to our research question. It displays a positive correlation between Google+Facebook clicks and the stock of offline advertising. The latter is defined as the sum of ad spendings across time, discounted by a factor $1 - \lambda$. Figure 1.3 displays an offline adstock with a carry-over rate of $\lambda = 0.85$. We explain latter how we choose this parameter.

It is also worth noting that a number of confounding factors may affect the relationship between offline advertising and online demand behavior. Offline and online investments may

FIGURE 1.1 – Google Clicks Variation by Country



be correlated since firms are more than likely to coordinate their advertising strategies simultaneously on several media. Moreover, ad spending and consumer behavior exhibit common seasonality, making it difficult to infer a causal relationship of the former on the latter. This classic endogeneity issue has been widely identified in the advertising literature. In the next subsection, we provide our estimation strategy to identify cross-media externalities.

FIGURE 1.2 – Google vs Facebook Clicks through Time

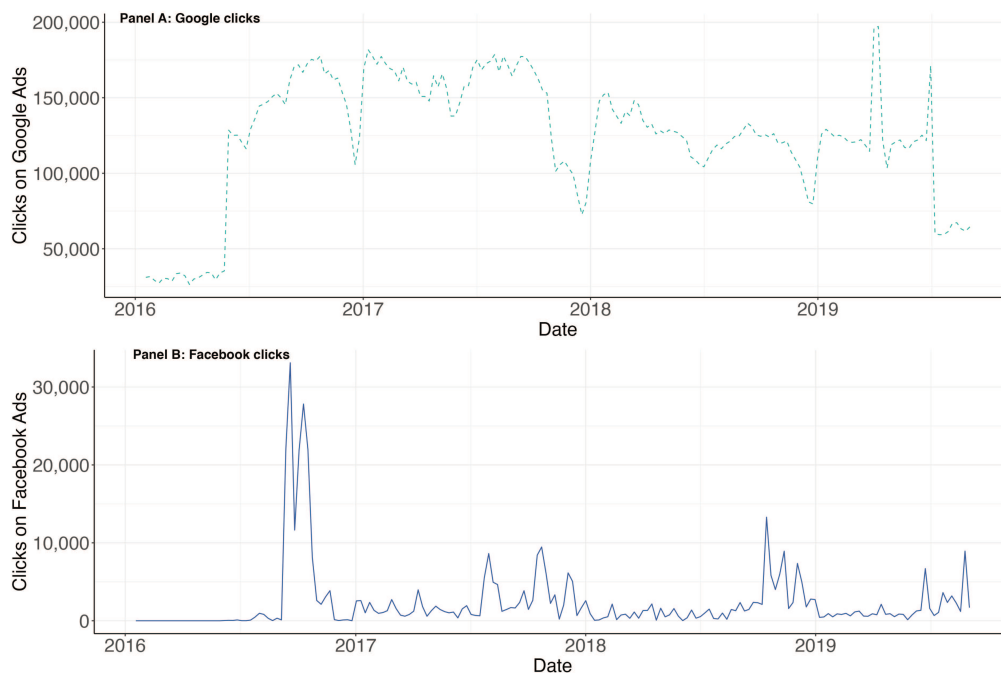
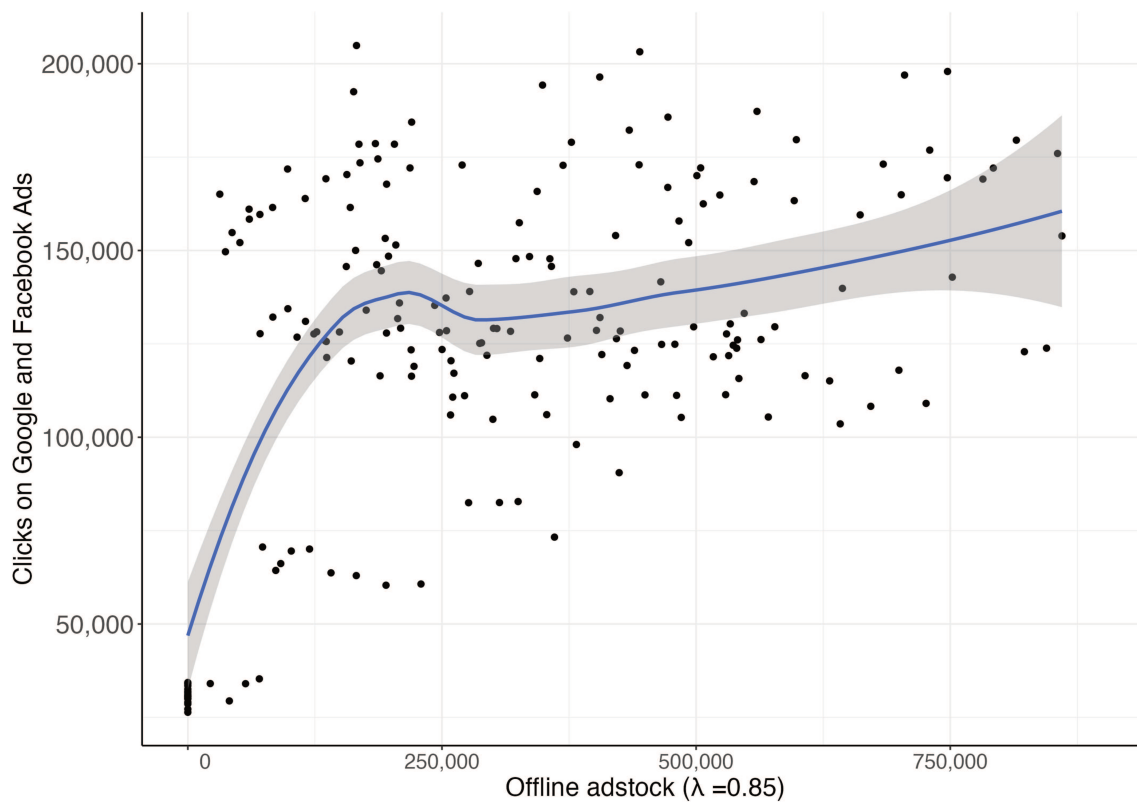


FIGURE 1.3 – Google+Facebook Ads Clicks and Offline Advertising Stock



4 Identification Strategy

We seek to estimate whether the effectiveness of Google and Facebook ads is impacted by externalities from other media campaigns, especially offline. We measure online effectiveness by the total number of clicks Y_{bct} , the brand b in country c recorded on its ads in week t . Clicks approximate consumers' utility for the ad : they are more related to sales than impressions and have been used in previous empirical researches to approximate advertising effectiveness (TUCKER, 2014; JEZIORSKI et SEGAL, 2015; SHEHU et al., 2021). In addition, clicks are one the main pricing instruments on which advertisers are charged (in CPC contracts). In order to identify the effect of the offline advertising activity on online ad effectiveness, we specify the following log-log fixed effects regression :

$$\log(Y_{bct}) = \sum_m \beta_m \log(1 + A_{mbct}) + \gamma X_{bct} + \xi_b + \mu_c + \delta_t + \varepsilon_{bct} \quad (1.1)$$

Our feature of interest, A_{mbct} , is media m 's (offline, online, competitors) advertising stock for brand b in country c on week t . As the number of clicks primarily depends on the campaign size, we control for X_{bct} , the log of amount spent by the advertiser on Google (or Facebook), in euro. The model also relies on a set of fixed effects (FEs) to account for unobserved heterogeneity. The brand and country fixed effects ξ_b , μ_c capture time-invariant characteristics (e.g. national market characteristics, brand positioning, specific advertising strategies). Seasonality is captured by month FEs while general long-term effects are accounted by implementing year FEs, both represented by the vector δ_t . Although relatively straightforward, this model raises additional estimation challenges.

Adstock Parameter First, the adstock function is to be specified. We implement a linear decay stock specification : $A_{mbt} = a_{mbct} + \lambda_m A_{mbc,t-1}$ where the media-specific carryover parameter $\lambda_m \in (0, 1)$ is to be set. We follow the literature and estimate carryover rates by conducting a grid search. We run model (1.1) with different λ_m , $m \in \{\text{offline, online}\}$ and choose the pair of rates under which the sum of squared residuals (SSR) is minimized (as in DINNER et al., 2014). A full description of the method is given in appendix 8.3. Parameters returned by the procedure are reported in Table 1.2.

Endogeneity Endogeneity between demand and advertising is a common issue in marketing models (RUTZ et WATSON, 2019). Indeed, advertising investments are neither random nor independent from clicks : an omitted variable could affect both the decision of firms to advertise and the propensity of consumers to click on hotel ads (e.g. firms anticipating their demand). To attenuate this bias, we use an instrumental variable approach.

For a given country c , we look for an instrument that affects a firm's advertising expenditure without being directly correlated to the endogeneity source or demand. It is tempting to consider using competitors' ad spending as a valid instrument. However, competitors are also very likely to advertise according to their anticipation of the demand. This is a problem since, by definition, competitors target the same demand as the brands we study here (hereafter, *fo-*

cal brands). Our idea is to instrument the spending of our focal brands by advertising from competitors in a foreign country c' . Practically, the advertising stock of competing hotel brands in the United Kingdom will be used to instrument German hotels' advertising and vice-versa. Indeed, advertising investments in the domestic and foreign markets should be correlated : firms in the European hotel industry share common cost and demand characteristics which may underlie similar advertising strategies. As they do not target the same demand, we believe that foreign competitors' spendings attenuate the endogeneity between advertising and demand in the domestic market. Especially since we only consider clicks by consumers located in the domestic market. The exploitation of marketing variables in foreign regions or close non-competing markets has been proven to provide effective instruments (NEVO, 2001 ; CHINTAGUNTA et al., 2006 ; VAN HEERDE et al., 2013). We implement the instrument using a 2SLS approach. The first-stage equation and results are detailed in Appendix 8.4.

Standard Errors Finally, we estimate the model using heteroskedasticity-robust standard errors clustered at the week-year level (52 groups). Our intuition behind this choice is twofold. First, our standard errors may be correlated across date for each year and brand due to the seasonality of an unobserved component. Second, the treatment (i.e. advertising expenditure) is allocated at the date-level : firms allocate their advertising investments for each period. Because the five firms observed in the data belong to the same hotel group, their advertising strategies may be correlated through time. Especially because they are likely to work with a common advertising agency, brands may share the same media planning strategy. Third, the continuous treatments effects β_m are likely to vary across time periods. Fourth, the number of time periods is far larger than the number of units (brands, countries). These four observations advocate for the clustering of standard errors at the date-level (ABADIE et al., 2022).

5 Results and Discussion

5.1 Adstock Carryover Rates

Parameters λ_m^* estimated from the grid search procedures are reported in Table 1.2. As we can see, these rates are globally consistent with those previously estimated in the literature⁴.

4. Although having $\lambda_{\text{online}} > \lambda_{\text{offline}}$ can seem counter-intuitive, HE et al., 2018 also found a greater advertising carry-over effect online. This can be related to brands' high online share of budget, highlighted before. This also coincides with the fact the hotel industry is a sector where sales mainly take place online.

TABLE 1.2 – Carryover parameters estimated vs Reported in the Literature

| Paper | Media studied | λ_{Offline} | λ_{Online} |
|--|-----------------------------|----------------------------|---------------------------|
| DUBÉ et al., 2005 ; SHAPIRO et al., 2021 | TV | .90 | – |
| HE et al., 2018 | TV & Online | .69 | .70 |
| DINNER et al., 2014 | Offline & Online | .89 | .84 |
| This study | Offline & Online | .85 | .90 |

Table 1.3 provides the results of model (1.1) for the coefficients of interest. The KP Wald F-Stats and first stage F-Stat both indicate a good validity of the instrument. Google and Facebook models display a good adjusted R^2 .

Our results demonstrate that the existence of statistically significant externalities generated by offline ads impact on the effectiveness of Google ads. A 1% increase in the offline adstock generates a 0.67% increase in Google ad clicks. This finding is consistent with previous studies on offline-search behaviors showing that TV ads generate a significant amount of immediate and delayed search queries (REILEY et al., 2010; LEWIS et REILEY, 2013; DINNER et al., 2014; LIAUKONYTE et al., 2015; JOO, WILBUR, COWGILL et al., 2014; JOO, WILBUR et Y. ZHU, 2016).

Surprisingly, the effect of online ads (display, video) on search clicks is negative. This echoes contradictory findings in the literature⁵. Thus, negative externalities between online ads do exist – confirming the theoretical literature previously described – but they seem heterogeneous among industries. In our case, we have established several reasons for this result. First, negative externalities between online ads may be a consequence of a competition for clicks and attention among online media. Second, display ads generally come later in the purchase decision process. A consumer targeted by a banner ad may have already shown interest in the advertised brand or product category and thus be less willing to search (since he already knows about the product).

Competitors ad investment do not seem to decrease search clicks. Finally, we do not find any significant effect of media ad expenditures on Facebook clicks. This echoes V. KUMAR et al., 2017 that found no interaction effect of television and social media ads on sales. Facebook do not act as a gatekeeper on product search, which makes it less easy for Facebook to benefit from brand building effects. Moreover, Facebook data are more limited : we have less observations and clicks exhibit a stronger standard deviations.

Our results show the existence of strong externalities from offline advertising impacting online search ads. This effect may be generated through different channels : increase in consumer’s search volume vs in propensity to click. In the next section, we provide additional models to disentangle the different effects governing our results.

5. DINNER et al., 2014 found a negative non-significant effect of a luxury clothing retailer’s display ads on the click rate for search ads. In contrast, KIREYEV et al., 2016’s study on a bank showed that display impressions often increased search effectiveness while search ads decreased display performance. The study by LEWIS et NGUYEN, 2015 also showed that display advertising could trigger consumers’ search for competing brands, explaining these ambiguous findings.

TABLE 1.3 – Effect of Offline, Online and Competing Adstocks on Google and Facebook Ads Outcomes

| | Google clicks | | | Facebook clicks | | |
|--------------------------------|----------------------|-----------------------|-------------------|---------------------|-------------------|-------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\log(A_{\text{offline}})$ | 0.219*** (0.0472) | 0.302*** (0.0474) | 0.673* (0.333) | -0.167* (0.0765) | 6.379 (28.86) | 0.460 (0.313) |
| $\log(A_{\text{online}})$ | | -0.288*** (0.0445) | -0.236 (0.141) | | -12.79 (56.91) | -1.718 (1.018) |
| $\log(A_{\text{competitors}})$ | | | -0.551 (0.373) | | | 0.888 (0.678) |
| Ad spending | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Country & Brand FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Month & Year FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 875 | 875 | 875 | 299 | 299 | 299 |
| Adjusted R-squared | 0.996 | 0.994 | 0.965 | 0.975 | -2.881 | 0.922 |
| KP LM-stats | 22.68 | 34.88 | 4.312 | 30.94 | 0.0496 | 4.115 |
| KP-Wald F-Stats | 43.41 | 16.17 | 1.357 | 58.29 | 0.0233 | 1.397 |

*p<0.05 ; **p<0.01 ; ***p<0.001

Note : all the regressions include brand, country, month-of-the-year and year fixed effects. Robust standard errors clustered by week-year. Facebook data were only available for United Kingdom firms : country FEs dropped

5.2 Mechanism

Two mechanisms may explain the positive effect of offline ads on Google clicks. On the one hand, traditional mass-media ads may increase the volume of searches, and thus clicks (extensive margin), by informing or reminding consumers about the existence of the brand. This *search lift* phenomenon has been extensively studied in the literature (see Section 1.1). On the other hand, offline ads may increase consumers' utility for the brand and thus induce more consumers to click for a constant number of impressions (intensive margin). Thus, if offline ads affect the volume of searches and/or users' propensity to click, how does this translate into advertising prices ?

We answer these questions by running model (1.1) on four new alternative Google dependent variables : (i) the brand's Google impressions, which measure the number of brand-related searches, (ii) clicks on search ads conditional on ad impressions, which measure propensity to click, and (iii) Google ads' cost-per-click, a measure of search ad prices.

Results reported in Table 1.4 confirm our two hypotheses : offline advertising increases both the brand-related query volume (column (1)) and the advertising click rate on Google (column (2)). The first effect is higher and more significant than the second. Finally column (3) shows that media campaigns, offline as well as online, do not affect the cost-per-click in a significant

manner.

TABLE 1.4 – Results of the Mechanism Analysis

| | Google Ads | | |
|--------------------------------|-------------------------|---------------------|----------------------|
| | log(Impressions) (1) | log(clicks) (2) | log(CPC) (3) |
| $\log(A_{\text{offline}})$ | 0.921* (0.458) | 0.133* (0.0676) | -0.173 (0.0966) |
| $\log(A_{\text{online}})$ | -0.344 (0.191) | -0.0644 (0.0430) | 0.0168 (0.0570) |
| $\log(A_{\text{competitors}})$ | -0.773 (0.517) | -0.0705 (0.0671) | 0.188*** (0.0513) |
| Control | Ad spending | Ad impressions | – |
| Country & Brand FEs | ✓ | ✓ | ✓ |
| Month & Year FEs | ✓ | ✓ | ✓ |
| Observations | 875 | 875 | 875 |
| Adjusted R-squared | 0.951 | 0.726 | -1.835 |
| KP LM-stats | 4.312 | 2.232 | 11.12 |
| KP-Wald F-Stats | 1.357 | 0.780 | 3.468 |

*p<0.05; **p<0.01; ***p<0.001

Note : columns (1) to (3) show the effect of brands' own and competitors ad spendings on other online ad outcomes : number of time a user is exposed to a Google ad in column (1) ; Google ads' clicks conditional on ad impressions in column (2) ; and the cost-per-click for Google ads in column (3). All regressions control for brand, country, month-of-the-year and year fixed effects. All adstocks are instrumented by competitor's ad spendings in foreign regions. Robust standard errors clustered by week-year

The positive and significant effect of competing adstock on cost-per-click impression is particularly interesting as it supports the existence of *keyword poaching* strategies. Two options are possible here. First, the focal brand may advertise on keywords that are also demanded by competitors : generic keywords such as “hotel Berlin” or competitors' brand name such as “Airbnb Berlin”. In both cases, an increase in competitors' ad spendings increase the price of the keyword because competitors bid higher to acquire it. A second option may be that the focal brand's keyword is being poached by its competitors. And as a result, the focal brand may still win its keyword (competitor ad spendings do not decrease clicks) but they a higher second-price in the ad auction.

In the long-run, the overstated effectiveness of search ads may augment the budget allocated to Google ads at the expense of other media. To test for this effect, we aggregate our data at the semester-year level and analyze the effect of past clicks on Google and the offline budget share in our data by running the following model :

$$\text{logit}(S_{cbt}) = \alpha + \delta_t + \beta \log(Y_{bc,t-2}) + \sum_t \phi_t (\delta_t \times \log(Y_{bc,t-2})) + \varepsilon_{bct}. \quad (1.2)$$

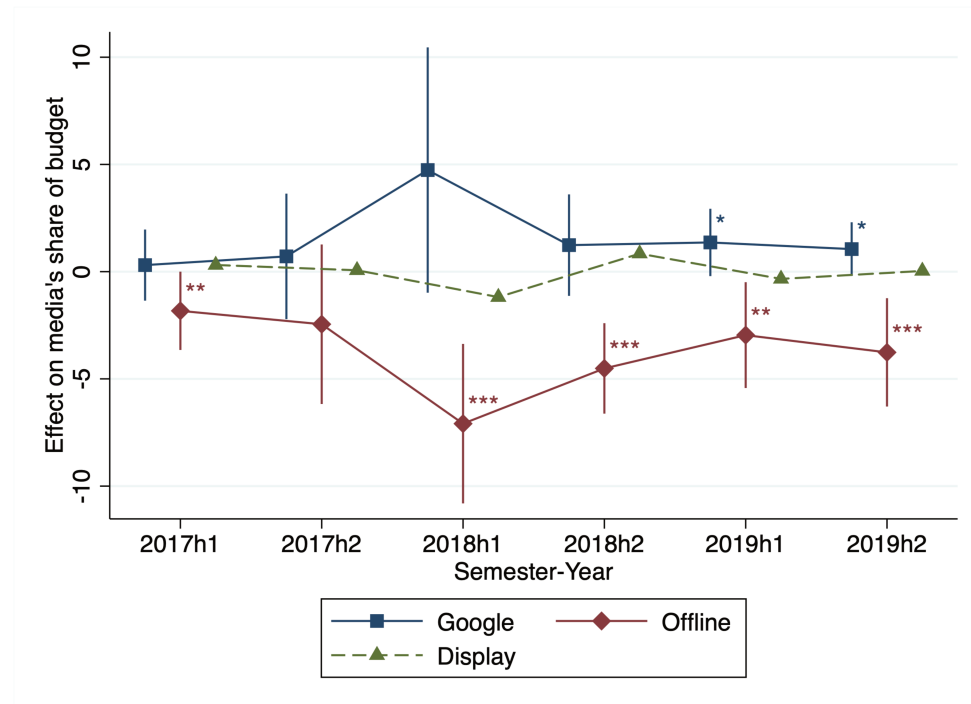
The share of budget S is estimated as a function of the clicks Y obtained two semester before⁶. Results in Figure 1.4 suggest that a high Google performance in the past year increases Google’s share of the budget at the expense of traditional media. This effect tends to get stronger with time. The effect of Google clicks on the online display budget share is not significant and does not exhibit any trend.

5.3 Robustness Checks

Several checks are performed to ensure the robustness of our findings. First, results from linear-log and OLS estimations are reported in Table 1.14 : they are consistent with the effects observed in our main estimations. Second, we run model (1.1) controlling for ad impressions in X_{bct} , rather than cost. Because the two variables are closely correlated, the two regression yield similar results. Third, an estimation using the click-through rate (CTR) as the dependent variable is performed. A log and logit transformation of the CTR provide results that are consistent with our main estimation (see Table 1.12). Finally, we also provide different checks regarding fixed effects that show the importance of accounting for unobserved time-varying characteristics (Table 1.13).

6. The reason for choosing $t - 2$ instead of $t - 1$ is that firms are likely to adjust their advertising budget on a year-on-year basis. This is especially true for offline ads which are typically bought far in advance (6 months to a year), before advertisers have observed search ads’ performance. Thus, the full effect of Google’s performance on the media budget share may be observed at least two semesters after the performance has been reported. A model run with $\log(Y_{bc,t-1})$ shows similar but less significant results (Figure 1.5), suggesting that a semester is not the appropriate time span to observe a substitution effect.

FIGURE 1.4 – Effect of Past Google Clicks on Media Investment by Semester



Note : marginal effects of $\log(Y_{bc,t-2})$ estimated from equation (1.2). 95% confidence intervals displayed. Standard errors robust to heteroskedasticity and auto-correlation used. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

6 Implications

In the previous sections, we found that offline, online and competing ad campaigns all affect search advertising outcomes. When the effect is positive, Google free-rides on these externalities : as an integrated ad platform, Google can charge more clicks. For their part, advertisers end up paying a cost to Google each time a click is initiated by an offline ad campaign. This cost may not be wasteful if the brand would not have sold a room in the absence of a Google ad. However, this hypothesis may not always hold. For example, a search for a brand or product name is not random and it is very likely that such a query results in a sale, independent of the presence of a Google Ad. Some brands may decide to under-invest offline and poach their competitors' online branded keywords. In this section, we present the implications of such findings for the advertising industry.

6.1 Managerial Implications for Advertisers

The presence of externalities between advertising media yields implications for firms' advertising strategies. Indeed, our results suggest that when an advertiser invests offline, this not only increases the effectiveness of its online campaigns but also its cost. We provide an estimation of such effects in Appendix 8.5. We find that the typical marginal increase in offline adstock ($\approx \text{€}720$) results, in average, in 257 additional clicks on Google sponsored links, which ultima-

tely cost around €68 to the advertiser. Thus, for every €10.5 invested in offline advertising, Google captures €1. As discussed earlier, offline-online effects have already been highlighted in the marketing literature, but their effects on ad pricing and media substitution have not been investigated.

Our results suggest that the computation of Google campaigns ROIs may be overestimated in the presence of externalities because search ads are addressed to consumers already exposed to firms' promotional efforts (including offline campaigns). However, firms may continue to invest in search advertisements because of the opportunity cost of letting their competitors poach their keywords, as suggested by our results.

6.2 Free-riding and Market Power

Our results suggest that Google free-rides on advertising externalities generated by other media campaigns. As the main gateway to online search, Google benefits from brands' promotional efforts. This set-up is similar to vertical relationships in which a manufacturer's advertising generates vertical externalities for the retailer of the product (MURRY, 2017).

Google acts as a monopolist retailer located between multi-homing advertisers willing to reach single-homing users. As a competitive bottleneck, Google allows brands to reach its single-homing users through advertising (ARMSTRONG, 2006; KAISER et WRIGHT, 2006). PRAT et VALLETTI, 2022 recently showed that online *attention brokers* can leverage their market power to reduce the number of advertising slots, and in turn increase ad prices. This is done at the expense of the social welfare with brands getting less ad slots and consumers potentially paying the higher ad rates in the final product price.

In this paper, we highlight another dimension of competitive bottlenecks in advertising : the ability of a monopoly ad platform to capture a brand's offline promotional efforts. We also show that Google's ability to auction off trademark keywords may rise ad prices. This practice has been discussed, both in courts and papers (O. P. BOMSEL, 2013; DLAPIPER, 2015). Our results empirically support theoretical models in which poaching creates a prisoner dilemma for brands, which ultimately benefits to the search engine (DESAI et al., 2014). A study on the effect of poaching on market concentration, retail prices, and media revenues would be welcome to enlighten the welfare effect of this practice.

By contrast, Facebook may not be able to free-ride on advertising externalities because the platform does not have a monopoly power over display advertisements⁷ (STATISTA, 2021b).

6.3 Advertising Media Market

Empirical studies have shown that outdoor and mail advertising restriction increases online ad spending, concluding that both markets are substitutes (GOLDFARB et TUCKER, 2011e). However, more recent papers on television and print have provided evidence that offline and online ads are likely to be complementary (CHANDRA et KAISER, 2014; HE et al., 2018). We supplement this empirical literature by showing that online advertising performance and revenues also

7. In 2019, the entire Facebook group controlled less than half of display ad spendings according to STATISTA, 2021b.

depend on the activity in the offline advertising market. In particular, we argue that offline and online advertising may be vertical and complementary markets. In particular, branded search queries critically rely on the brand equity built through promotional efforts such as offline advertising. Offline ads massively provide information and narratives about products, which then tend to initiate consumer searches and clicks online.

However, the results depicted in Figure 1.4 speak against this hypothesis : the offline advertising activity decreases with Google clicks. Based on our results, we can give three reasons for brands to advertise on Google at the expense of traditional media. First, as explained, brands may be poached and over-invest in search ads to defend their branded keywords. Second, investing in offline brand building may increase search ad costs and competitors' incentive to poach. This can encourage advertisers to avoid traditional media campaigns. Third, they can poach keywords from notable advertisers instead of advertising in mass-media. In all cases, advertisers under-invest offline.

7 Conclusion

In this study, we consider advertising as a commodity that impacts consumers' preferences for products. When a firm invests in offline advertisements, it produces a complementary good that tends to increase consumers' utility for its product. Google benefits from this incremental utility, as offline campaigns induce consumers to look for brands online and click more on sponsored links. Moreover, the search ad auction design creates incentives for competitors to free-ride well-known brand's notoriety. As search ads prices depend on queries or clicks, these cross-media externalities turn into additional revenues for the search engine.

No such effect was found to be statistically significant for Facebook ads. The reason may be that, as a monopolist search gatekeeper, Google is able to charge firms for the access to online consumers. The search engine benefits from firms' promotional efforts which attract consumers online. This effect is particularly strong in the travel industry, where most sales take place online.

This study suffers from both technical and theoretical limitations. First, Facebook data are limited in the number of observations and quality, which did not help identification. Second, using clicks to measure externalities can lead to over- or underestimating advertising effects. On the one hand, it is unclear whether clicks eventually lead to sales for advertisers (BLAKE et al., 2015). On the other hand, externalities between advertisements may produce effects beyond clicks (ZENETTI et al., 2014).

Moreover, the present paper only focuses on the measure of a certain type of advertising externalities in a given industrial context. However, the difficulty to generalize advertising effects is the curse of advertising research.

The technical and theoretical limitations of this study open the path to further modeling and policy topics. In particular, the study may be replicated on a more diverse set of periods and industries. Indeed, offline-to-online effects, poaching and media substitution are all likely to vary across industries. Conversions and online purchase data could also be used in order to observe whether cross-media effects effectively lead to sales.

Further analyses could tackle how under-investment in offline advertising induces collateral effects on the quality of copyrighted works outside traditional media such as news, movies or documentaries. Similarly, the effect of poaching firms' competition, retail prices, and media revenues needs to be studied to gauge whether poaching is detrimental.

8 Appendix

8.1 Additional Descriptive Statistics

TABLE 1.5 – Correlation Matrix

| | Google clicks | Google imp | Google | Fb clickss | Fb imp | facebook | Online | Offline | Competitors | F comp on | F comp off | Off share |
|---------------------|---------------|------------|-----------|------------|----------|----------|---------|------------|-------------|-----------|------------|-----------|
| Google clicks | 1 | | | | | | | | | | | |
| Google imp. | 0.981*** | 1 | | | | | | | | | | |
| Google cost | 0.851*** | 0.870*** | 1 | | | | | | | | | |
| Fb clicks | 0.202*** | 0.176*** | 0.176*** | 1 | | | | | | | | |
| Fb imp | 0.208*** | 0.174*** | 0.142** | 0.682*** | 1 | | | | | | | |
| Fb cost | 0.240*** | 0.209*** | 0.171** | 0.827*** | 0.879*** | 1 | | | | | | |
| Online | 0.165** | 0.148** | 0.0916 | 0.274*** | 0.300*** | 0.361*** | 1 | | | | | |
| Offline | 0.103 | 0.109* | 0.159** | 0.0922 | 0.0690 | 0.115* | 0.108* | 1 | | | | |
| Competitors adstock | -0.165** | -0.153** | -0.128* | -0.0537 | -0.0390 | -0.0322 | -0.0397 | -0.0000317 | 1 | | | |
| F comp offline | 0.00659 | -0.00950 | -0.0734 | 0.226*** | 0.106* | 0.185*** | 0.0815 | 0.0328 | 0.0227 | 1 | | |
| F comp online | -0.338*** | -0.319*** | -0.326*** | 0.0115 | -0.0189 | 0.0269 | -0.0586 | -0.0209 | 0.240*** | 0.0113 | 1 | |
| Offline share | 0.0612 | 0.0634 | 0.0905 | 0.215*** | 0.147** | 0.271*** | 0.0685 | 0.529*** | 0.0472 | 0.195*** | 0.0137 | 1 |

Note :

*p<0.1; **p<0.05; ***p<0.01

TABLE 1.6 – Summary Statistics by Brand

| Variable | Low cost | | Mid range ⁺ | | Mid range ⁻ | |
|-------------------------|--------------------|---------|------------------------|--------|------------------------|--------|
| | Mean | SD | Mean | SD | Mean | SD |
| Google clicks | 44506 | 16838 | 15133 | 7184 | 29003 | 19398 |
| Google impressions | 168512 | 63620 | 75171 | 35454 | 119248 | 68881 |
| Google cost | 9819 | 4194 | 4632 | 2767 | 7081 | 4363 |
| Facebook clicks | 1878 | 4514 | 661 | 1354 | 1236 | 3306 |
| Facebook impressions | 232301 | 548045 | 45199 | 67712 | 133553 | 390583 |
| Facebook cost | 496050 | 1152347 | 128650 | 240842 | 302145 | 830293 |
| Online | 14538 | 35737 | 4621 | 6175 | 9304 | 25414 |
| Offline | 7174 | 40240 | 4795 | 23998 | 5918 | 32663 |
| Competitors | 42813 | 53200 | 103270 | 156487 | 74721 | 123036 |
| Offline adstock | 47779 | 93191 | 31960 | 47466 | 39430 | 73054 |
| Online adstock | 144646 | 174037 | 46043 | 38943 | 92605 | 132238 |
| Competitors off adstock | 183877 | 158578 | 487238 | 557004 | 343984 | 445168 |
| Competitors on adstock | 188007 | 200624 | 279320 | 172137 | 236200 | 191393 |
| IV on adstock | 145695 | 197303 | 119233 | 105641 | 131729 | 156128 |
| Competitors off adstock | 265395 | 391618 | 1307203 | 744885 | 815238 | 797222 |
| Offline share of budget | 0 | 0 | 0 | 0 | 0 | 0 |
| Observations | 380 | | 380 | | 190 | |
| Period covered | 01/2016 to 09/2019 | | | | | |

TABLE 1.7 – Summary Statistics by Country

| Variable | DE | | United Kingdom | | Total | |
|-------------------------|--------------------|--------|----------------|--------|--------|--------|
| | Mean | SD | Mean | SD | Mean | SD |
| Google clicks | 28990 | 18888 | 22091 | 15568 | 24851 | 17298 |
| Google impressions | 118226 | 67228 | 100393 | 57952 | 107526 | 62410 |
| Google cost | 6307 | 3749 | 7032 | 4253 | 6597 | 3972 |
| Facebook clicks | 1171 | 3229 | 0 | 0 | 468 | 2120 |
| Facebook impressions | 126524 | 381309 | 0 | 0 | 50609 | 248822 |
| Facebook cost | 302145 | 830293 | 0 | 0 | 302145 | 830293 |
| Online | 8814 | 24821 | 8488 | 8430 | 8618 | 16991 |
| Offline | 5607 | 31817 | 12983 | 37771 | 10032 | 35676 |
| Competitors | 75875 | 121360 | 183322 | 258282 | 140343 | 220575 |
| Offline adstock | 37355 | 71645 | 85660 | 114586 | 66338 | 102387 |
| Online adstock | 87732 | 130357 | 80891 | 40742 | 83627 | 88277 |
| Competitors off adstock | 333805 | 435549 | 1096578 | 951106 | 791469 | 870590 |
| Competitors on adstock | 246286 | 195903 | 156394 | 276811 | 192351 | 251417 |
| IV on adstock | 133690 | 152193 | 238315 | 194191 | 196465 | 185716 |
| Competitors off adstock | 842785 | 806528 | 282608 | 367547 | 506679 | 645114 |
| Offline share of budget | .08 | .24 | .15 | .31 | .12 | .28 |
| Observations | 570 | | 380 | | 950 | |
| Period covered | 01/2016 to 09/2019 | | | | | |

8.2 Facebook Weekly Data Conversion

The heterogeneous lengths of social ad campaigns is an issue we have to deal with since all our explanatory features are reported on calendar weeks. Let Y_{bp}^F , be the clicks recorded on Facebook ad campaigns for brand b in a $p \in \mathcal{P}$ associated to a length of $d(p)$ days. We convert data from heterogeneous period length into calendar weeks as following :

$$Y_{bt}^F = \sum_{p \in \mathcal{P}} \{ [Y_{bp}^F / d(p)] \times d(p \cap t) \} \quad (1.3)$$

The terms inside the brackets correspond to the daily-average Facebook features. We then multiply it by the number of days for which period p overlaps week t , i.e. $0 \leq d(p \cap t) \leq 7$. The same calculation is used for Facebook impressions and costs.

8.3 Grid Search Algorithm

A grid search algorithm is used to estimate the carryover parameter of our adstock function. Two parameters are to be estimated : $\lambda_{\text{offline}}^*$ and $\lambda_{\text{online}}^*$. The grid search procedure consists in running the regression model with different carryover-rate and then selecting the optimal λ^* s which minimize the error of the model. Formally, we run the following simple model :

$$\log(Y_{cbt}) = \underbrace{\sum_m \beta_m \log(1 + A_{mbct}) + \gamma X_{bct} + \xi_b + \mu_c + \delta_t + \varepsilon_{bct}}_{\widehat{Y}_{cbt}(\lambda_m)}$$

In the equation above, Y_{cbt} is the sum of Google and Facebook clicks while X_{bct} contains the log of Google+Facebook impressions and a linear time trend. $A_{mbct}(\lambda_m)$ is the adstock variable of media m given a carryover rate λ_m . The algorithm then chooses the best pair of $\lambda_m^* \in (\lambda_{\text{offline}}^*, \lambda_{\text{online}}^*)$ which minimizes the Sum of Squared Residuals (SSR).

$$\lambda_m^* = \arg \min_{\lambda_m} \sum_{b,c,t} [Y_{cbt} - \widehat{Y}_{cbt}(\lambda_m)]^2$$

The algorithm searches for $\lambda_m \in (0.05, 0.95)$ with a pas of 0.05. Once the adstock rate obtained, we compute the respective adstocks $A_{mbct} = \lambda_m^* A_{mbc,t-1} + a_{mbct}$ and proceed to the instrumentation strategy.

8.4 Instrumental Variable Approach

In regression (1.1), the adstock of media m is instrumented by the adstock of competitors in the foreign market on that same media. Formally, adstocks of firms located in the United Kingdom will be instrumented by competitors in the German market and vice-versa. Similarly, the competitors' spendings in foreign market are instrumented by the focal brands' spending in the domestic market. The first-stage is as following :

$$\log(A_{bct}^m) = \phi^m \log(C_{bc't}^m) + \theta_b^m + \psi_c^m + \rho_t^m + \nu_{bct}^m$$

where $C_{b,c',t}^m$ is the adstock of b 's competitors in the foreign market c' ; θ_b^m , ψ_c^m and ρ_t^m are brand, country and time fixed effects as in the main model. The results of the first stage are given in the following table.

TABLE 1.8 – First Stage Results from Equation (1.1)

| | A_{offline} | A_{online} | $A_{\text{competitors}}$ |
|--------------------------|----------------------|----------------------|--------------------------|
| C_{offline} | 0.416*** (0.0349) | | |
| C_{online} | | 0.576*** (0.0472) | |
| $C_{\text{competitors}}$ | | | 0.425*** (0.0560) |
| Observations | 950 | 950 | 950 |
| Adjusted R-squared | 0.919 | 0.945 | 0.960 |
| F-statistic | 1317.8 | 1105.7 | 11718.3 |

*p<0.1; **p<0.05; ***p<0.01

Note : regression include brand, country, month-of-the-year and year fixed effects. Robust standard errors in parentheses

8.5 Clicks and Cost Elasticities in Value

Our results suggest that a marginal increase in offline advertising dA rises the clicks and costs of Google campaigns. To have an approximation of such effects in values, we convert our elasticities β_m in values with the following back-of-the-envelope equations :

$$\underbrace{\eta \equiv \frac{\exp(\beta_{\text{offline}} - 1)}{100} \times \overline{Y_{bct}}}_{\Delta \text{Clicks (in value)}} \quad \underbrace{\xi \equiv \eta \times \overline{\text{Cost-per-click}_{bct}}}_{\Delta \text{Cost (in value)}} \quad \underbrace{dA = \frac{\overline{A_{\text{offline},bct}}}{100}}_{\Delta \text{Offline adstock (in value)}}$$

where \bar{x} denotes the sample mean of variable x . Variables are drawn from the sample and regression coefficients. We use the coefficient of our preferred specification : column (3) of Table 1.3. Results are reported in Table 1.9. The interval for the amount of the advertising stock at which the effect takes place is large and hard to approximate.

 TABLE 1.9 – Effect of Offline Adstock dA on Google Clicks η and cost ξ

| | $\Delta \text{Clicks } (\eta)$ | $\Delta \text{Google Cost } (\xi)$ | $\Delta \text{Offline Adstock } (dA)$ |
|----------|--------------------------------|------------------------------------|---------------------------------------|
| Mean | 257 | €68 | €720 |
| Std. Dev | (166) | (84) | (1,047) |

8.6 Full Regression Results (Including Ad Budget)

TABLE 1.10 – Effect of Offline, Online and Competing Adstocks on Google and Facebook Ads Outcomes

| | Google clicks | | | Facebook clicks | | |
|--------------------------------|----------------------|-----------------------|------------------------------|----------------------|-------------------|-----------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\log(A_{\text{offline}})$ | 0.219*** (0.0472) | 0.302*** (0.0474) | 0.673* (0.333) | -0.167* (0.0765) | 6.379 (28.86) | 0.460 (0.313) |
| $\log(A_{\text{online}})$ | | -0.288*** (0.0445) | -0.236 (0.141) | | -12.79 (56.91) | -1.718 (1.018) |
| $\log(A_{\text{competitors}})$ | | | -0.551 (0.373) (0.373) | | | 0.888 (0.678) (0.678) |
| $\log(\text{Ad spending})$ | 1.044*** (0.0363) | 1.261*** (0.0232) | 1.721*** (0.394) | 0.882*** (0.0560) | 1.176 (1.679) | 1.070*** (0.198) |
| Country & Brand FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Month & Year FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 875 | 875 | 875 | 299 | 299 | 299 |
| Adjusted R-squared | 0.996 | 0.994 | 0.965 | 0.975 | -2.881 | 0.922 |
| KP LM-stats | 22.68 | 34.88 | 4.312 | 30.94 | 0.0496 | 4.115 |
| KP_Wald F-Stats | 43.41 | 16.17 | 1.357 | 58.29 | 0.0233 | 1.397 |

*p<0.05; **p<0.01; ***p<0.001

Note : all the regressions include brand, country, month-of-the-year and year fixed effects. Robust standard errors clustered by week-year. Facebook data were only available for United Kingdom firms : country FEs dropped

TABLE 1.11 – Results of the Mechanism Analysis

| | Google Ads | | |
|---------------------------------|------------------|-------------|----------|
| | log(Impressions) | log(clicks) | log(CPC) |
| | (1) | (2) | (3) |
| log(A_{offline}) | 0.921* | 0.133* | -0.173 |
| | (0.458) | (0.0676) | (0.0966) |
| log(A_{online}) | -0.344 | -0.0644 | 0.0168 |
| | (0.191) | (0.0430) | (0.0570) |
| log($A_{\text{competitors}}$) | -0.773 | -0.0705 | 0.188*** |
| | (0.517) | (0.0671) | (0.0513) |
| log(Google cost) | 2.092*** | | |
| | (0.550) | | |
| log(Google impressions) | | 1.065*** | |
| | | (0.0716) | |
| Country & Brand FEs | ✓ | ✓ | ✓ |
| Month & Year FEs | ✓ | ✓ | ✓ |
| Observations | 875 | 875 | 875 |
| Adjusted R-squared | 0.951 | 0.726 | -1.835 |
| KP LM-stats | 4.312 | 2.232 | 11.12 |
| KP-Wald F-Stats | 1.357 | 0.780 | 3.468 |

*p<0.05; **p<0.01; ***p<0.001

Note : columns (1) to (3) show the effect of brands' own and competitors ad spendings on other online ad outcomes : number of time a user is exposed to a Google ad in column (1) ; Google ads' clicks conditional on ad impressions in column (2) ; and the cost-per-click for Google ads in column (3). All regressions control for brand, country, month-of-the-year and year fixed effects. All adstocks are instrumented by competitor's ad spendings in foreign regions. Robust standard errors clustered by week-year

TABLE 1.12 – Effect of Offline, Online and Competing Adstocks on Google and Facebook Ads Outcomes

| | Google clicks | | | Facebook clicks | | |
|--------------------------------|----------------------|-----------------------|------------------------------|----------------------|-------------------|-----------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\log(A_{\text{offline}})$ | 0.219*** (0.0472) | 0.302*** (0.0474) | 0.673* (0.333) | -0.167* (0.0765) | 6.379 (28.86) | 0.460 (0.313) |
| $\log(A_{\text{online}})$ | | -0.288*** (0.0445) | -0.236 (0.141) | | -12.79 (56.91) | -1.718 (1.018) |
| $\log(A_{\text{competitors}})$ | | | -0.551 (0.373) (0.373) | | | 0.888 (0.678) (0.678) |
| $\log(\text{Ad impressions})$ | 1.044*** (0.0363) | 1.261*** (0.0232) | 1.721*** (0.394) | 0.882*** (0.0560) | 1.176 (1.679) | 1.070*** (0.198) |
| Country & Brand FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Month & Year FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 875 | 875 | 875 | 299 | 299 | 299 |
| Adjusted R-squared | 0.996 | 0.994 | 0.965 | 0.975 | -2.881 | 0.922 |
| KP LM-stats | 22.68 | 34.88 | 4.312 | 30.94 | 0.0496 | 4.115 |
| KP_Wald F-Stats | 43.41 | 16.17 | 1.357 | 58.29 | 0.0233 | 1.397 |

*p<0.05; **p<0.01; ***p<0.001

Note : all the regressions include brand, country, month-of-the-year and year fixed effects. Robust standard errors clustered by week-year. Facebook data were only available for United Kingdom firms : country FEs dropped

8.7 Robustness Chekcs

TABLE 1.13 – Robustness to Fixed Effects

| | Google clicks | | | Facebook clicks | | |
|--------------------------------|-----------------------|-----------------------|--------------------|--------------------|-------------------|--------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\log(A_{\text{offline}})$ | 0.214*** (0.0463) | 0.229*** (0.0288) | 0.673** (0.333) | -0.0708 (0.136) | 0.0626 (0.180) | 0.460 (0.313) |
| $\log(A_{\text{online}})$ | -0.167*** (0.0477) | -0.206*** (0.0249) | -0.236* (0.141) | -0.393 (0.591) | -0.692 (0.757) | -1.718* (1.018) |
| $\log(A_{\text{competitors}})$ | 0.0270 (0.0178) | 0.146*** (0.0401) | -0.551 (0.373) | 0.0717 (0.347) | 0.558 (0.550) | 0.888 (0.678) |
| Brand & Country FEs | No | Yes | Yes | No | Yes | Yes |
| Month & Year FEs | No | No | Yes | No | No | Yes |
| Observations | 875 | 875 | 875 | 299 | 299 | 299 |
| Adjusted R-squared | 0.996 | 0.996 | 0.965 | 0.962 | 0.952 | 0.922 |
| F-statistic | 69304.8 | 71742.1 | 9180.3 | 1802.6 | 1374.2 | 474.9 |
| KP LM-stats | 43.14 | 62.03 | 4.312 | 2.804 | 2.374 | 4.115 |
| KP_Wald F-Stats | 17.02 | 20.94 | 1.357 | 0.955 | 0.800 | 1.397 |

Note :

*p<0.1; **p<0.05; ***p<0.01

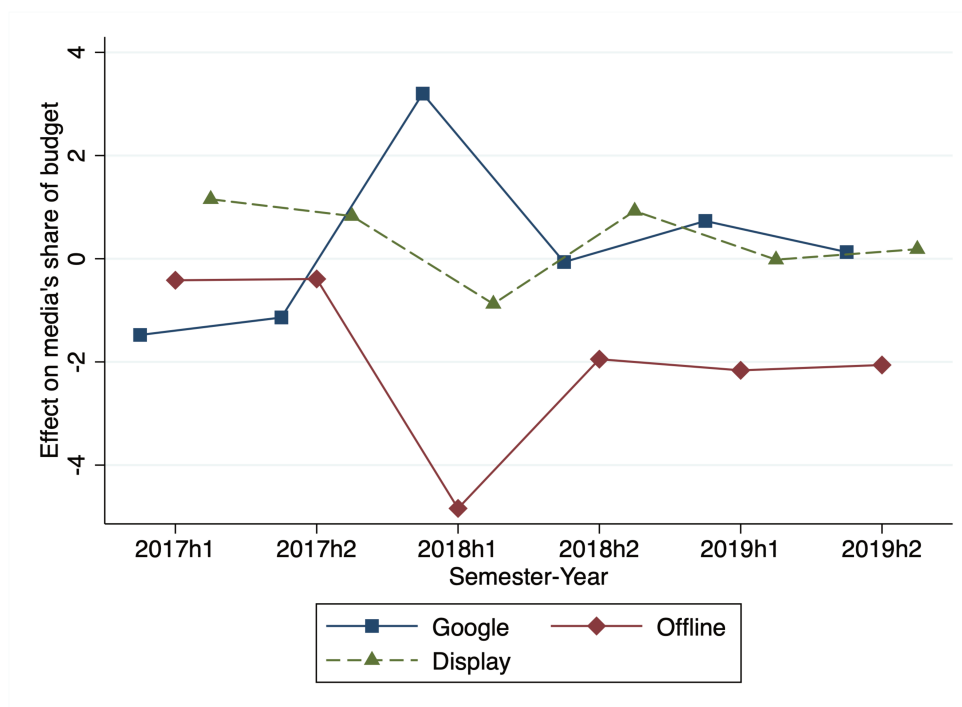
TABLE 1.14 – Robustness to Other Specifications and Estimators

| | Linear-log specification | | | |
|----------------------------|--------------------------|----------------------|-------------------|-------------------|
| | Google clicks | | Facebook clicks | |
| | IV | OLS | IV | OLS |
| | (1) | (2) | (3) | (4) |
| $\log(A_{\text{offline}})$ | 859.1*** (105.1) | 3571.4*** (294.2) | -34.95 (51.27) | -226.6 (126.6) |
| Observations | 875 | 875 | 299 | 299 |
| Adjusted R-squared | 0.953 | 0.878 | 0.727 | 0.716 |
| F-statistic | 1658.6 | 1040.9 | 20.44 | 18.30 |
| KP LM-stats | | 45.04 | | 32.99 |
| KP-Wald F-Stats | | 111.1 | | 66.37 |

Note :

*p<0.1; **p<0.05; ***p<0.01

FIGURE 1.5 – Effect of Past Semester’s Google Clicks on Media Investment by Semester



Note : marginal effects of $\log(Y_{bc,t-1})$ estimated from equation (1.2). 95% confidence intervals displayed. Standard errors robust to heteroskedasticity and auto-correlation used. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

TABLE 1.15 – Results when Controlling for Ad Impressions

| | Google Ads Outcome | | |
|--------------------------------|------------------------------|----------------------|----------------------|
| | log(Clicks) | log(CTR) | logit(CTR) |
| | (1) | (2) | |
| $\log(A_{\text{offline}})$ | 0.643*** (0.192) | 0.189** (0.0606) | 0.192** (0.0594) |
| $\log(A_{\text{online}})$ | -0.448** (0.154) | -0.0801* (0.0386) | -0.0839* (0.0385) |
| $\log(A_{\text{competitors}})$ | 0.598*** (0.0648) | -0.234** (0.0731) | -0.218** (0.0719) |
| Google imp | 0.0000119*** (0.00000331) | | |
| Observations | 875 | 875 | 875 |
| Adjusted R-squared | 0.969 | 0.879 | 0.826 |
| KP LM-stats | 18.59 | 8.959 | 8.959 |
| KP-Wald F-Stats | 7.639 | 3.421 | 3.421 |

*p<0.05; **p<0.01; ***p<0.001

Note : columns (1) to (3) show the effect of brands' own and competitors ad spendings on Google clicks (1), log of CTR (2) and logit-transformed CTR (3). All regressions control for brand, country, month-of-the-year and year fixed effects. All adstocks are instrumented by competitor's ad spendings in foreign regions. Robust standard errors clustered by week-year

Chapitre 2

Better Together ? Complementarity and Substitution between Offline and Online Advertising

Abstract

In this chapter, we investigate whether firms use offline and online ads as complements or substitutes in order to generate sales. While previous works studied substitution patterns between traditional media, studies taking Internet advertising into account are more scarce. In this paper, we estimate the demand of advertising and then assess own- and cross-price elasticities between offline, display and search ads. The model is estimated on 2016-2020 data from French advertisers in three industries. We find different media substitution patterns across these sectors but offline and online advertisements generally appear as limited substitutes. Brands are relatively inelastic to offline and search ad prices. This study concludes that (i) advertising strategies are heterogeneous across industries and (ii) that the advertising market is more fragmented than it was thought to be.

Résumé

Dans ce chapitre, nous cherchons à savoir si les entreprises utilisent les publicités hors ligne et en ligne comme des compléments ou substituts pour générer des ventes. Si les travaux antérieurs ont étudié la substituabilité des médias traditionnels, les études prenant en compte la publicité en ligne sont plus rares. Dans cet article, nous estimons la demande de publicité des firmes, puis nous dérivons les élasticités-prix propres et croisées entre publicités hors ligne, display et search. Le modèle est estimé sur les données d'annonceurs français opérant dans trois industries entre 2016 et 2020. La substitution des médias varie entre chaque secteur mais les publicités hors ligne et en ligne apparaissent généralement comme des substituts limités. Les marques sont relativement inélastiques aux prix des publicités hors-ligne et search. Cette étude conclut que (i) les stratégies publicitaires sont hétérogènes entre les industries et (ii) que le marché publicitaire est plus fragmenté qu'on ne le pensait.

1 Introduction

In September 2022, TF1 and M6, the two largest French private-owned free TV networks, officially gave up their merger plan announced nearly 5 months earlier (DAMS, 2022). After investigations, the French competition authority had rejected the view that TF1-M6 competes against online advertising sellers. Ruling that TV advertising was a relevant market on its own, the authority considered that the market power of the merged entity would be too strong, raising concerns on a potential increase in TV ad prices. Restrictive conditions had hence been established for the merger to proceed, which TF1 and M6 declined only to give up on their plan.

This decision sheds light on an old affair : the dependence of firms to advertising and the definition of advertising's relevant market. Offline and online advertisements have often been considered as substitutes by researchers and marketers (GOLDFARB et TUCKER, 2011e). However it might only be a mere speculation since recent economic studies are scarce on this topic. The key questions are thus : are advertisers elastic to ad prices? Can they substitute between offline and online advertising in order to generate sales? And thus, do traditional and Internet advertisement belong to the same relevant market?

Historically important in advertising, the complements-or-substitutes question has regained interest since the advent of Internet. Theoretical studies on optimal ad spendings typically consider advertising as a unique entity. But obviously, whether offline and online campaigns benefit from each other in generating sales yields strong implications for brands' advertising strategies. The structure of advertising demand is also important in an antitrust perspective : firms' elasticities to ad prices or relevant market for advertising are critical insights for regulators.

In the traditional media industry, most offline ad support are seen as substitutes to each other. In the case of France, the introduction of new advertising medias, such as posters, radio and TV advertising, has always decreased the incumbent medias' share of ad revenue (MARTIN, 2016), hinting substitution. However, internet advertising is a game changer because its targeting, pricing and industrial organization are very different from the traditional ad industry. Advertisers' price elasticities to online media becomes of interest as a root cause of market power and concentration on the ad-funded media market. Early during the emergence of online advertising, antitrust authorities argued that offline and online media belonged to distinct markets. In a 2008 court case, the European Commission ruled out that online and offline advertising were distinct markets (EUROPEAN COMMISSION, 2008). The year before, the Federal Trade Commission similarly stated that search advertising was a relevant market on its own (FEDERAL TRADE COMMISSION, 2007). However, these early market definitions have been challenged by researches in the last decade. These recent studies often rely on natural experiments rather than authentic measures of competition like cross-price elasticities to measure advertising substitution. Thus, in the absence of empirical proof the market definition remains open. In this study, we investigate whether search, display and traditional offline ad sellers compete against each other or not. To qualify the relevant market, we focus on demand-side substitution, which is the ad-buyer's (firm) viewpoint. We use a translog cost function, in the form first introduced by BERNDT et CHRISTENSEN, 1973 and then assess price elasticities. Data used are based on 2016-2020 ad

spendings from French advertisers operating at the national scale in three different industries : hotels, soft drinks and telecommunications. These sectors are heavy ad-spenders both offline and online.

We find that substitution patterns are heterogeneous across industries. In most cases, offline is distinct from search and substitutable to display. Display and search, the two components of digital advertising, are rather substitutes. Own-price elasticities also give interesting insights on media market power : brands are generally inelastic to offline and search ad prices. While the offline media market is national and competitive, the search ad market is dominated by Google, which raises question about the latter's monopoly power over firms. This study concludes that (i) advertising strategies and price elasticities are heterogeneous across industries and (ii) that the advertising market is more fragmented than it was thought to be in previous studies. The first point is of importance : when assessing ad-funded media mergers, regulators should maintain that offline and online ads do not belong to the same relevant market.

This paper brings several contributions to the literature and policy debate. First, we do not focus on a sole industry as previous papers, but consider three different sectors of the economy. Second, we refine the empirical evidences on offline-online substitutions by distinguishing between search and display advertisements. To the best of our knowledge, no study investigated substitution patterns between search and display ads¹. Nevertheless, if Google sponsored links were to compete with Facebook and Instagram ads, it would yield strong implications for competition in the online ad market. Finally, we show that media substitution is an industry-dependant phenomenon and that a pooled estimation which disregards industry characteristics may result in misleading price elasticities and substitution patterns.

Related Works This paper addresses different streams of the literature. First, our study relates with general models of optimal advertising allocation such as DORFMAN et STEINER, 1954 ; NERLOVE et ARROW, 1962 ; BECKER et MURPHY, 1993. These theoretical models highlight that a firm's optimal ad expenditures should depend on its demand's elasticities to price and advertisement. Yet, demand elasticity to prices and ads is very likely to vary across industries, perhaps across brands. Moreover, these models only consider a representative firm spending on a representative advertising medium. In BECKER et MURPHY, 1993's view, advertising act as a complementary good which increase or decrease the value of the advertised product according to consumer's valuation for the ad. In practice, each firm may create its own complement using different media investments (like television advertising, direct mail, online search ads). Whether firms can substitutes between the different inputs generating the advertising complement remains to be seen.

Second, our findings lie in the literature on pricing power in advertising. Models of two-sided markets stressed that media are competitive bottlenecks : when viewers single-home, the media has a monopoly power over advertisers seeking to reach the viewers (ARMSTRONG, 2006). PRAT et VALLETTI, 2022 consider the case of an online attention broker, whose market power allows to shrink ad supply in order to extract additional profits. Extending the model to multi-homing

1. DEVAUX et O. BOMSEL, 2022 nevertheless hint that display ads compete with search ads for user's attention.

users, recent studies showed that a media’s advertising price decrease with viewer’s propensity to consume other media (S. P. ANDERSON et al., 2018; GENTZKOW et al., 2022).

Third, the literature on media substitution and competition is closely related to our research. Industrial organization studies provide many empirical evidence about offline media substitution (SELDON et JUNG, 1993; Robert B EKELUND et al., 1999; Robert B. EKELUND et al., 2000; SELDON, JEWELL et al., 2000; SILK et al., 2002; FRANK, 2008). These analyses were based on variations in advertising prices and expenditures. While slightly differing in results, they unanimously concluded that partial advertising bans were ineffective given advertisers’ ability to shift ad expenditures from banned media toward other offline supports. With the emergence of online advertising, many empirical studies stressed it might as well be a substitute to traditional advertising. EVANS, 2009 discussed offline-online substitution in a early review and highlighted that substitutability would depend on campaign objectives and demand behavior. GOLDFARB et TUCKER, 2011a; GOLDFARB et TUCKER, 2011d found that bans in outdoor and direct mail advertisements increased the effectiveness and expenditures of online advertising. Additional studies argue that online advertising decreased newspapers ad prices and expenditures (SEAMANS et F. ZHU, 2014; SRIDHAR et SRIRAM, 2015). Conversely, other empirical evidences suggest that offline and online ads may be complements. From the advertiser’s standpoint, advertising on both media can create synergies and generate higher returns than the sum of their individual effectiveness (for instance, see NAIK et PETERS, 2009). CHANDRA et KAISER, 2014 showed that the Internet increased advertising prices in printed magazines which operated a companion website. It therefore seems that advertisers value more printed ads when they can use them conjointly with online ads, which suggests complementarity effects.

The current literature on media substitution has two main limitations. First, it focuses on a single specific offline media (print, mail, outdoor) or a particular industry (often food or alcohol) which isn’t indicative of the whole advertising market. For instance, in Western Europe, TV advertising represented 46% of offline ad expenditures during the 2016-2020 period, while outdoor only amounted to 12% (STATISTA, 2022a). In addition, these studies do not estimate substitution patterns by exploiting variations in media prices. Instead they only investigate reactions to advertising restrictions or the introduction of the Internet. As substitution effects are changes in quantities in reaction to a change in prices, measuring a change in quantities in reaction to an unavailability of one input is not informative. Indeed, the adjustments following the ban or the introduction of an input doesn’t yield any information on the reaction of the quantities consumed to changes in prices, hence they can not measure substitution effects. In an attempt to overcome these problems, HE et al., 2018 estimate cross-price elasticities between TV, magazines and online advertising between 2005 and 2011. They find that, while both offline media are close substitutes, they are both used as a complement to online ads. Nevertheless, the online advertising market significantly changed since 2011. Improvement in targeting technologies and regulatory changes are very likely to have affected advertising media substitutions. Yet, researches investigating advertising media substitution using recent data are too scarce. Using 2016-2019 data, DEVAUX et O. BOMSEL, 2022, find that offline advertising campaigns influences search ads’ effectiveness, hinting complementarity. By increasing user’s propensity to click on

sponsored link, traditional ads also increase the cost of search ads since the latter are paid on a cost-per-click basis. However they do not exploit price variations, which prevents them from deriving substitution accurately. This paper seeks to fill this gap.

2 The Model

2.1 Theoretical Framework

Our approach is based on the theory of cost functions. Our starting hypothesis is that production of goods and advertising involve two different and separable processes. While the quantity of goods produced by firms depends on the prices and quantity of commodities used in the production process, the quantity of advertising bought depends on media prices and sales target. Our assumption here is that the advertising behaviour is decided after, and according to, production. Thus, the goal of a firm is to minimize the cost C allocated to advertising *given the quantity q of goods they wish to sell* :

$$\begin{aligned} \arg \min_{\mathbf{X}} C &= \mathbf{p}\mathbf{X} \\ \text{s.t } \psi(\mathbf{X}) &= q. \end{aligned} \tag{2.1}$$

The total advertising cost is specified as the product of the vector of advertising message bought on different media \mathbf{X} and its unit price \mathbf{p} . The function $\psi(\mathbf{X})$ is an advertising effectiveness function relating the quantity of advertising message to sales. For the translog estimation, we assume this function is quasi-concave and twice differentiable.

2.2 Functional Form

The total cost can be expressed as a translog cost function of the $i = 1, \dots, n$ media prices, quantities sold q and a linear time trend T (BERNDT et CHRISTENSEN, 1973) :

$$\begin{aligned} \ln(C) &= \alpha + \sum_{i=1}^n \beta_i \ln(p_i) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \gamma_{ij} \ln(p_i) \ln(p_j) \\ &\quad + \theta \ln(q) + \sum_{i=1}^n \theta_i \ln(q) \ln(p_i) \\ &\quad + \delta T + \sum_{i=1}^n \delta_i T \ln(p_i) \\ &\quad + \mu + \tau + \varepsilon_c. \end{aligned} \tag{2.2}$$

All variables are at the firm-time level but we drop indexes for simplicity. We included brands and quarter-year fixed effects with μ and τ to account for unobserved firms and time-varying shocks in ad spendings respectively.

For this cost function to be consistent with the economic theory, it must meet two strong conditions. First, it must be concave in prices. This condition will be discussed further later. Second, the function must be homogeneous of degree one in prices, which requires the following

restrictions on the coefficients :

$$\sum_i \beta_i = 1, \text{ and } \sum_i \gamma_{ij} = \sum_j \gamma_{ji} = \sum_i \theta_i = \sum_i \delta_i = 0. \quad (2.3)$$

Using Shephard's lemma, one can find that the differentiation of cost equation(2.2) with respect to media i 's price is equal to the share of that media in the advertiser's budget (see appendix 6.1) :

$$\frac{\partial \ln(C)}{\partial \ln(p_i)} \equiv S_i = \beta_i + \sum_j \gamma_{ij} \ln(p_j) + \theta_i \ln(q) + \delta_i T + \varepsilon_i. \quad (2.4)$$

Equation (2.4) makes it clear that the share of a media in the advertising budget depends on all media prices (including its own price), the quantity sold $\ln(q)$, a linear trend T and a constant β_i . It also shows that the interaction coefficient γ_{ij} measures the change in i 's input share of budget following an increase in j 's price. Whether media i 's share of budget increases with quantity sold is measured by θ_i . Coefficient δ_i measures the general trend effect of time on media i 's share.

β_i can be interpreted as medium i 's equilibrium share in total budget. The equilibrium hereby refers to a situation in which all prices and the quantity sold are equal to 1. This equilibrium share can be explained by the natural dependence of the advertiser to a media. For example, large automobile advertisers are likely to have an incompressible share of TV in their advertising budget. For this share to be interpreted, we have to add a novel constraint to the model : $\beta_i \geq 0$. This constraint can be achieved by normalizing quantities sold and prices to their mean. Indeed, when the quantity, prices and trend do not vary from the mean, only β_i remains.

2.3 Deriving Elasticities

In equations (2.2) and (2.4), the cross-price coefficients γ_{ij} cannot be directly interpreted as elasticities. Indeed, they only measure a change in inputs shares $\partial S_i / \partial p_j$, while elasticity usually measure a change in quantity X in reaction to a change in price p . A fairly good measure of the ij 's price elasticity is : $\partial X_i / \partial p_j$. The literature has discussed extensively which elasticities of substitution should be used for which purpose (BLACKORBY et RUSSELL, 1989 ; STERN, 2011 ; FRONDEL, 2011)². The choice of the elasticity family is important as it comes with assumptions on substitution patterns. For example, Morishima elasticities tend to consider that all inputs are all substituable. Because we allow advertising media to be substitutes, complement or perfectly different inputs, we rely on cross-price elasticities (hereafter CPE) which measure the relative variation in the quantity of one input following a change in the price of another input. It can be computed easily using regression coefficients and input shares :

2. BLACKORBY et RUSSELL, 1989 advocate strongly in favor of Morishima elasticities against Allen elasticities. However, FRONDEL, 2011 pointed out that the Morishima elasticities tend to classify all inputs as substitutes and hence recommended to use cross-price elasticities. STERN, 2011 concluded there is not *one best* elasticity and therefore recommended that each study should use an elasticity fitted to the situation.

$$\eta_{ij} = \frac{\partial \ln(X_i)}{\partial \ln(p_j)} \equiv \frac{\gamma_{ij} + S_i S_j}{S_i}. \quad (2.5)$$

Using CPE, media i and j can be classified as substitutes if $\eta_{ij} > 0$; complements if $\eta_{ij} < 0$; and independent if η_{ij} approaches 0.

3 Data and Estimation

3.1 Data Sources

We use data between 2016 and 2020 from French advertisers operating on a national scale in three different consumer industries : hotels, soft drinks and telecommunications. These brands are big advertisers : in total, their market share in each industry is always greater than 33%. The three industries studied here are interesting for several reasons. First, they represent markets with different characteristics (e.g. selling channels, media mix, seasonality). Our data thus represent a wider diversity of media substitution than the datasets traditionally focused on food or alcohol markets. Second, the brands in these sectors massively invest in advertising, offline as well as online and hence constitute good case studies for offline-online ads substitution.

As we will see, our panel is unbalanced : all firms do not advertise on the same period. For each brand and month, our data provides : the amount spent in advertising on media i , the quantity of individual impressions³ bought and the quantity sold by the firm. To compute prices we divide expenditures by the quantity of impression and then multiply the result by 1000 to obtain the cost-per-thousand impressions (or CPM).

The dataset reports data on 9 advertising media, both offline and online. For the rest of the analysis, we will group them into three categories : offline, search and display. We have several reasons to do so. First, it reduces the number of observations where at least one input is not consumed. Indeed, firms do not invest on cinema or social media ads at each period of the panel. Grouping them with similar advertising media fixes the problem. Second, the translog cost function requires to estimate many parameters. For n media, the functional form in equation (2.2) implies the computation of at least $(n \times (n + 7))/2$ coefficients. Thus, using 9 advertising inputs requires the computation of at least 72 coefficients. With 3 inputs, the number of price coefficients to be estimated is reduced to 15. Thirdly, the aim of this study is to provide insights on : (i) the substitutability between offline and online media and (ii) the substitutability between online ads (search and display). It only makes sense to group offline media altogether. Complementarity and substitution among offline media have already been widely studied in the past (as discussed in Section 1). We also group different kinds of display format such as online video, social media ads and online banners together.

3. Impressions quantify the number of people exposed to an advertising message. Conversely, Gross Rating Points (GRPs) measure the share of the band's population of interest effectively exposed to the ad. Because TV and radio media reported GRPs instead of impressions, we had to convert GRPs into impressions in order to retrieve prices per 1000 impressions. The procedure is detailed in appendix 6.2.

3.2 Descriptive Statistics

Our data span from 2016 to 2021. Because the Covid crisis may have deeply impacted brand's selling channels, consumer behaviors and advertising media strategies, we discarded observations after March 2020. A first striking evidence is the heterogeneity of brand's ad spending across periods (Figure 2.1). In the soft drinks industry ad investments peak and then decrease, while in telecommunications brands seem to maintain ad effort throughout the period.

Looking at this first evidence, an important question is thus : to what extent are firms in those different industries comparable? As mentioned, the three brands more or less differ in seasonality. Going more into details, Figure 2.2 highlights a strong heterogeneity in the media mix of each sector. In particular, we can see brands in the hotel industry rely a lot on search ads. This is because hotels are very dependant of online sales : two thirds of the revenues in the travel market are generated online (STATISTA, 2021a). Hotel rooms are mainly sold on platforms such as Google Hotels or Booking and search ads increase the rank in platforms' listing. Conversely, soft drinks rely a lot on mass-media advertisement and search ad expenditures are marginal. This is natural given that in 2021, 97.7% of soft drinks were sold offline (STATISTA, 2022c). The high offline share is also explained by the strong brand equity of firms on this market. Finally, telecommunication brands are located halfway, with a strong offline component in the media mix and around one third of ad budget spent online, equally distributed between search and display. This reflects that electronic devices and mobile subscription can be purchased offline as well as online.

Equally interesting is the evolution of prices by media given in Figure 2.3. Since 2016, the price of offline impressions followed a downward trend, consistent with previous researches (GENTZKOW, 2014). Ad prices have risen online, both on display and search on the period. However, search ads experienced a stronger increase. A reason may be found in ad network's market power : as a monopoly, Google captured most of the Internet advertising growth, while the display markets is more competitive and advertiser's valuation for online ads have been distributed across different media and ad platforms (e.g. Facebook, Youtube, Amazon, Criteo, Teads, small ad sellers etc).

3.3 Estimation Strategy

We estimate both the total cost equation (2.2) and shares equations (2.4) (offline, display, search) under constraints (2.3). Because shares sum to one, the errors ε_i are jointly determined. To avoid this problem, we arbitrarily exclude a share of the system (here the search share)⁴. The coefficients from the search equation can be recovered from the total cost equation given the constraints. This estimation procedure has been widely used in the literature (SELDON et JUNG, 1993; SELDON, JEWELL et al., 2000; SILK et al., 2002; FRANK, 2008; HE et al., 2018). However, many issues remain to be solved at this point.

4. No matter what share of the equation is dropped, the estimated parameters do not change (GREENE, 2003)

FIGURE 2.1 – Evolution of advertising expenditures by brand

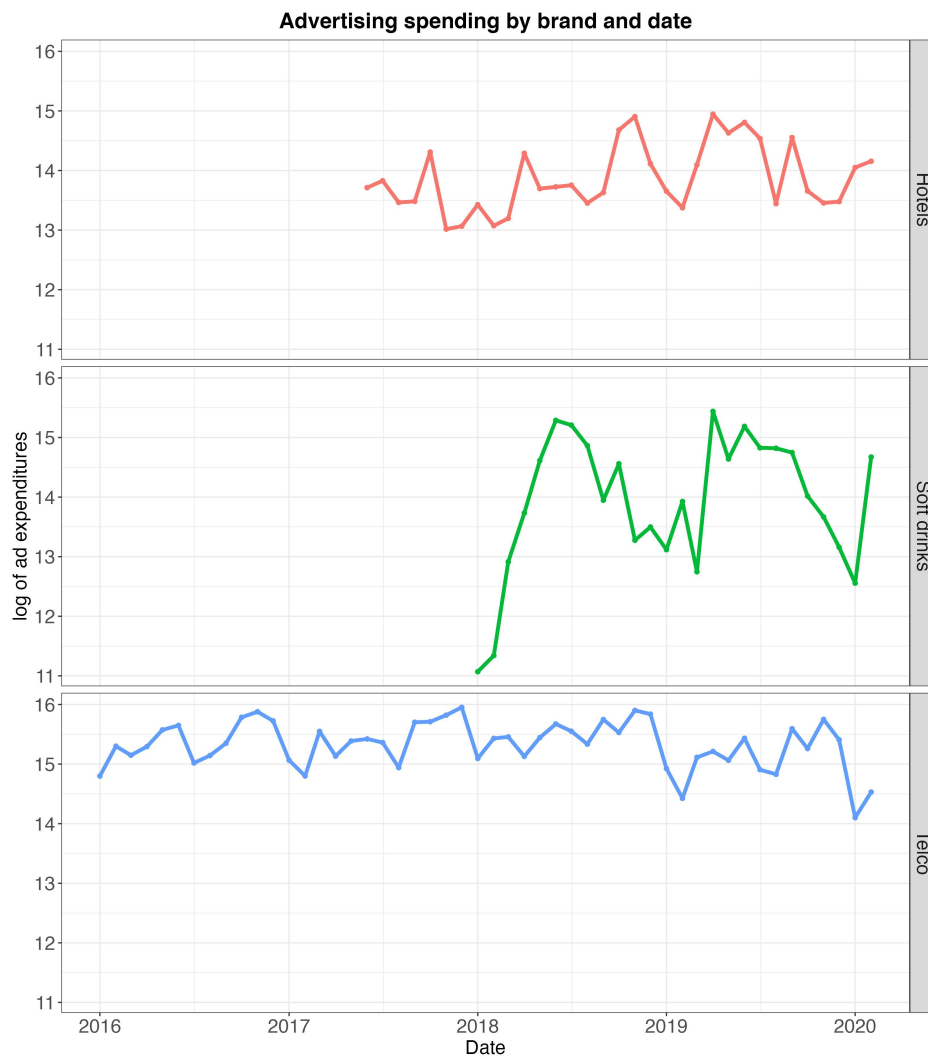


FIGURE 2.2 – Average media mix by industry

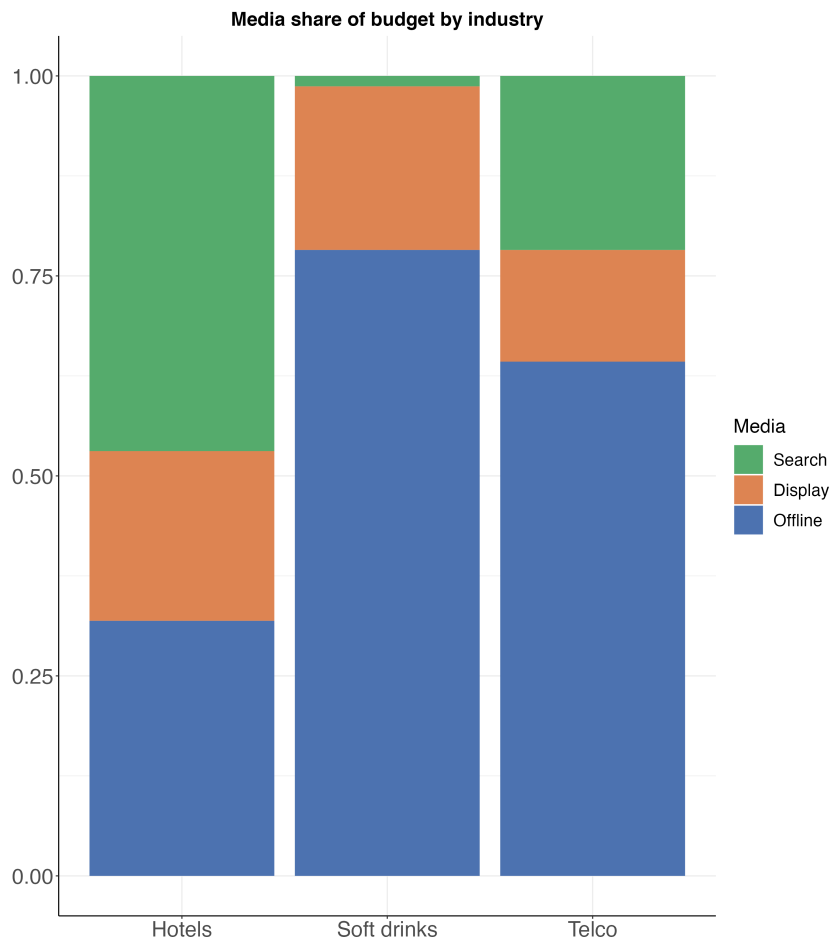
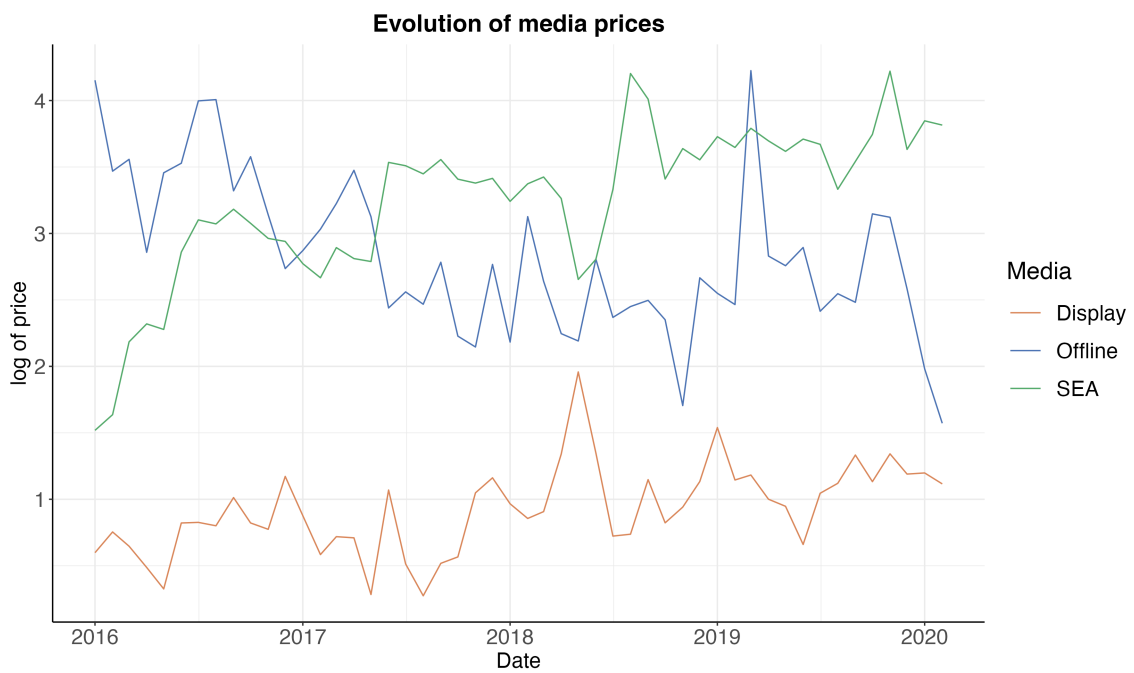


FIGURE 2.3 – Evolution of log(price) by media. Prices averaged by month



Endogeneity A weakness of the translog model (equations (2.2) and (2.4)) is the endogeneity between the total advertising spending C and quantity sold q . For the translog model to be theoretically valid, we assume that firms advertise accordingly to the quantity they want to sold (observed via q). However, the inverse is also true : sales depend on current and past advertising expenditures. This endogeneity problem has been widely documented in the marketing literature (RUTZ et WATSON, 2019). To address this issue, we use instrumental variables (IV). In equations (2.2) and (2.4), the sales q_{kt} for brand k in month t are instrumented by the log of consumer price index (CPI) in the brand's industry (travels, food and home equipment). Similar instruments have been used in previous papers (SELDON, JEWELL et al., 2000; FRANK, 2008). We assume that the CPI explains the natural demand for goods, independently from advertising. However, the CPI in a given industry may be correlated with consumer's price elasticity, which directly influence firms' propensity to spend in advertising (as illustrated by the well-known theorem from DORFMAN et STEINER, 1954). In such a case, the exclusion restriction assumption would be violated. We thus also employ the global Producer Price Index (PPI) as an instrument for sales. The intuition is that the PPI impacts a firms' supply independently from demand-side characteristics such as price and advertising elasticities.

In both cases, the first-stage F-Stat suggests that CPI and PPI are relevant instruments (see appendix 7.1). Because unobservables shocks may impact both the total ad cost and the share of each media, we use a system by three-stage least square (3SLS) estimator to account for both endogeneity and correlation between errors ε s across equations (HENNINGSEN et HAMANN, 2007).

Auto-correlation A common drawback of the translog estimation is that it does not account for errors correlation through time. However, cost and share equations at month t may be correlated with errors in the previous periods. This is likely to occur since advertising effects are inter-temporal and thus continue to affect sales in the following period. Auto-correlation biases the consistency of the results by violating the independence of errors. Durbin-Watson tests performed on cost and share equations revealed strong and positive auto-correlation of order 1. We correct this bias by implementing past errors in the cost and share equations (BERNDT et SAVIN, 1975). The procedure is detailed in appendix 7.2. Durbin-Watson tests made after the correction suggest auto-correlation has been mitigated.

Concavity A strong assumption of translog models is that the cost function (2.2) is concave in prices. Although this hypothesis is necessary for the isoquants to be convex, in practice it is rarely met. Two papers apart (SELDON et JUNG, 1993; SELDON, JEWELL et al., 2000), the concavity of the cost function has not been investigated in an advertising context. We employ the approach of DIEWERT et WALES, 1987 to check local and global concavity of our cost function (see appendix 7.3). The model satisfies concavity conditions at the reference point and 34% to 68% of the point are concave. However missing shares and prices harms the concavity of the model.

Missing prices We only observe ad prices if the brand bought ads from media i during a given period. Else, prices are missing. Treating missing prices as null is a problem : on the one hand $\ln(0)$ is not defined, on the other hand it would be the basic economic principle that if the price of a good is zero, then it should be highly demanded. We address this problem by estimating counterfactual prices as a function of firms' characteristics, month, year and competitors' ad spending. The method is detailed in appendix 7.4. As an alternative check, we also estimate the model at the industry-level excluding missing prices. Finally, we run the translog model after aggregating all industries together. All methods result in similar elasticities (see Table 2.6).

Elasticities of substitution's standard errors As we can see in equation (2.5), cross-price elasticities depend on media shares' of budget. Because the latter vary at each data point, CPE change with firms and time. GREENE, 2003 supports that it makes sense to compute elasticities at the sample mean of the variables. However, doing so we potentially discard heterogeneous variations inside the sample. We make the choice to compute the elasticities using shares at the brand-year level to take into account variation across years and brand inside the sample. Observations for which at least one $S_i = 0$ are dropped. A last challenge consists in estimating confidence intervals for our elasticities. Because the CPE is a non-linear transformation of regression coefficients γ_{ij} , deriving its variances is a non-trivial affair. We rely on R. G. ANDERSON et THURSBY, 1986's method and use the mean of observed shares to compute standard deviations and confidence intervals⁵.

4 Results

Media substitution patterns are very likely to be an industry-dependant phenomenon, as indicated by previous descriptive evidences. Even though fixed effects capture the heterogeneity in ad expenditures, the way advertisers substitute media is very likely to rely on industrial characteristics such as selling channels, regulatory constraints, degree of concentration or consumer's price elasticities. The model is thus estimated for each industry separately with brand and quarter-year fixed effects. Regression results are reported in Table 2.1. Cross and own-price elasticities are plotted in Figure 2.4 and Figure 2.5 respectively.

4.1 Summary of the Findings

As expected, cross-price elasticities are heterogeneous across sectors. Search and offline do not appear to be substitutes : η_{ij} are small and non-significant, indicating independent markets.

In most cases, offline and display are substitutes. A notable exception is the telecommunication sector where both media are classified as complements. This case is of interest. The particularity of brands in this sector is that they do not invest on print media (press, magazines, outdoor ads). Instead, they only advertise on audiovisual channels (mostly TV but also radio and cinema). Thus, the offline-display substitution pattern observed for other firms may be driven mostly by printed media which more naturally substitute with online banners. For

5. The formula used are the ones reported on the page 653 of the article (section V).

TABLE 2.1 – Estimation of the 3SLS equation system

| Variable | Coef. | Hotels | Soft drinks | Telco |
|------------------------------|---------------|-----------------|----------------|-----------------|
| Intercept | α | 11.20 (0.61)*** | 4.77 (2.31)* | 12.73 (0.25)*** |
| Offline price | β_o | 0.21 (0.08)** | 0.46 (0.08)*** | 0.29 (0.06)*** |
| Display price | β_d | 0.10 (0.06) | 0.41 (0.07)*** | 0.01 (0.06) |
| SEA price | β_s | 0.69 (0.10)*** | 0.12 (0.08) | 0.70 (0.06)*** |
| Sales (normalized) | θ | 0.13 (0.81) | 0.62 (0.54) | 0.67 (0.13)*** |
| Trend | δ | -0.03 (0.10) | -0.35 (0.17)* | -0.03 (0.04) |
| Offline price ² | γ_{oo} | 0.06 (0.05) | 0.02 (0.03) | 0.37 (0.05)*** |
| Display price ² | γ_{dd} | -0.01 (0.03) | -0.08 (0.03)* | -0.05 (0.05) |
| Search price ² | γ_{ss} | 0.10 (0.05) | 0.00 (0.02) | -0.15 (0.05)** |
| Offline \times Search | γ_{os} | 0.02 (0.03) | 0.03 (0.02) | -0.23 (0.04)*** |
| Offline \times Display | γ_{od} | -0.08 (0.04)* | -0.05 (0.02)** | -0.14 (0.04)*** |
| Display \times SEA | γ_{ds} | -0.01 (0.03) | 0.05 (0.02)* | 0.28 (0.05)*** |
| Sales \times Offline price | θ_o | 0.11 (0.05)* | 0.01 (0.02) | 0.18 (0.03)*** |
| Sales \times Display price | θ_d | -0.04 (0.03) | 0.05 (0.02)** | -0.13 (0.02)*** |
| Sales \times Search price | θ_s | -0.08 (0.06) | -0.06 (0.02)** | -0.05 (0.02)* |
| Trend \times Offline price | δ_o | 0.01 (0.00)* | 0.00 (0.00) | 0.00 (0.00)* |
| Trend \times Display price | δ_d | 0.01 (0.00)*** | -0.01 (0.00) | -0.00 (0.00) |
| Trend \times Search price | δ_s | -0.01 (0.00)*** | 0.00 (0.00) | -0.00 (0.00)* |
| Brand FEs | | ✓ | ✓ | ✓ |
| Quarter-Year FEs | | ✓ | ✓ | ✓ |
| McElory R ² | | 0.541 | 0.278 | 0.763 |
| Observations | | 192 | 345 | 294 |

*** $p < 0.01$; ** $p < 0.05$ * $p < 0.1$

o =offline, d =display, s =search.

Note : Standard errors reported in parentheses after the coefficient. Fixed effects and lagged error terms are omitted from the table

example, brands may easily substitute between advertising in Vogue printed and buying banners on Vogue.com as the price ratio between both media changes. Hotels and soda which rely a lot more on printed ads exhibits strong positive display-offline elasticities which supports this idea⁶.

Finally, display and search appear to be substitutes in all sectors. The elasticities range from small (hotels) to fairly high (telecommunication). The size of confidence intervals in the soft drink industry can be explained by the low share of display and search, which generate more variability in the computation of elasticities' standard errors.

Own-price elasticities are reported next in Figure 2.5 : they are all negative, which is consistent with the standard microeconomic theory. What is striking is the relatively low elasticity to offline ad prices. For notorious advertisers, traditional ads seem to be an essential input in maintaining brand equity. The elasticity to search ad prices are also low (except in the telecommunication industry). Google's monopoly power on the search ad market is likely to make advertisers inelastic to search ad prices. Moreover brands are motivated by the opportunity cost of letting a competitor buy their brand keyword (i.e. *brand poaching*, see SAYEDI et al., 2014; DEVAUX et O. BOMSEL, 2022; BHATTACHARYA et al., 2022. Finally, advertisers are very elastic to display ad prices. This is understandable since display campaigns have a dual role : produce information and narratives on the brand through *branding* campaigns, or foster consumer's conversion via *performance* ads. While the first role can be also ensured by search ads, offline campaigns assume the branding role. Thus, our intuition is that as the price of online display increases, brands are able to substitutes away on search or offline ads, depending on their objective (branding vs performance) and budget.

4.2 Analysis by Industry

The previous discussion of the results ignores the industrial characteristics of the sector considered. To explain the contradictory results displayed above, the authors documented on each brand's advertising strategy through internal documents or interview with each brand's marketing experts. Here are the case-by-case analysis for the resulting elasticities.

Hotels As explained earlier in section 3.2, the hotel industry relies a lot on online platforms and search ads to generate sales (see also Figure 2.2). However, and surprisingly, search does not seem to substitute to offline. Our explanation is the following : for consumers to search and book hotels, brands need to generate information and quality signals on the rooms, which can be done through offline campaigns. Thus, the low, non-significant offline-search elasticities denote independence, perhaps, complementarity between both media. The high dependence of hotels to offline and search is reflected in their respective low elasticity to offline and search elasticities.

For display, ad investments are allocated between branding and performance campaigns. As explained earlier, while the first type of campaigns is rather complementary to search, the second one is a substitute. Those two contradictory effects of display campaigns may explain

6. Moreover, a four-inputs model separating audiovisual from printed media has been fitted in appendix 8.1. The results show printed ads are more elastic to display prices than audiovisual advertisements, supporting our hypothesis

why offline-display elasticities are significant in some periods but not in others.

Finally, display and SEA appears to be weak substitutes. Globally, our results echo DEVAUX et O. BOMSEL, 2022 which hinted a weak search-display substitution in the European hotel market.

Telecommunications Brands in the telecommunication industry exhibit a rather unambiguous offline-display complementarity. As we explained, this particularity can be imputed to the high audiovisual share in the industry's media mix. During discussion with marketers in charge of the media plan, it appeared TV ads were mainly used for branding purposes while display ads (banners and social media) were used for performance purposes. Since the effectiveness of performance ads may be enhanced by branding efforts and vice-versa, the two kind of ads should be complements. This supports the offline-display complementarity identified by the model.

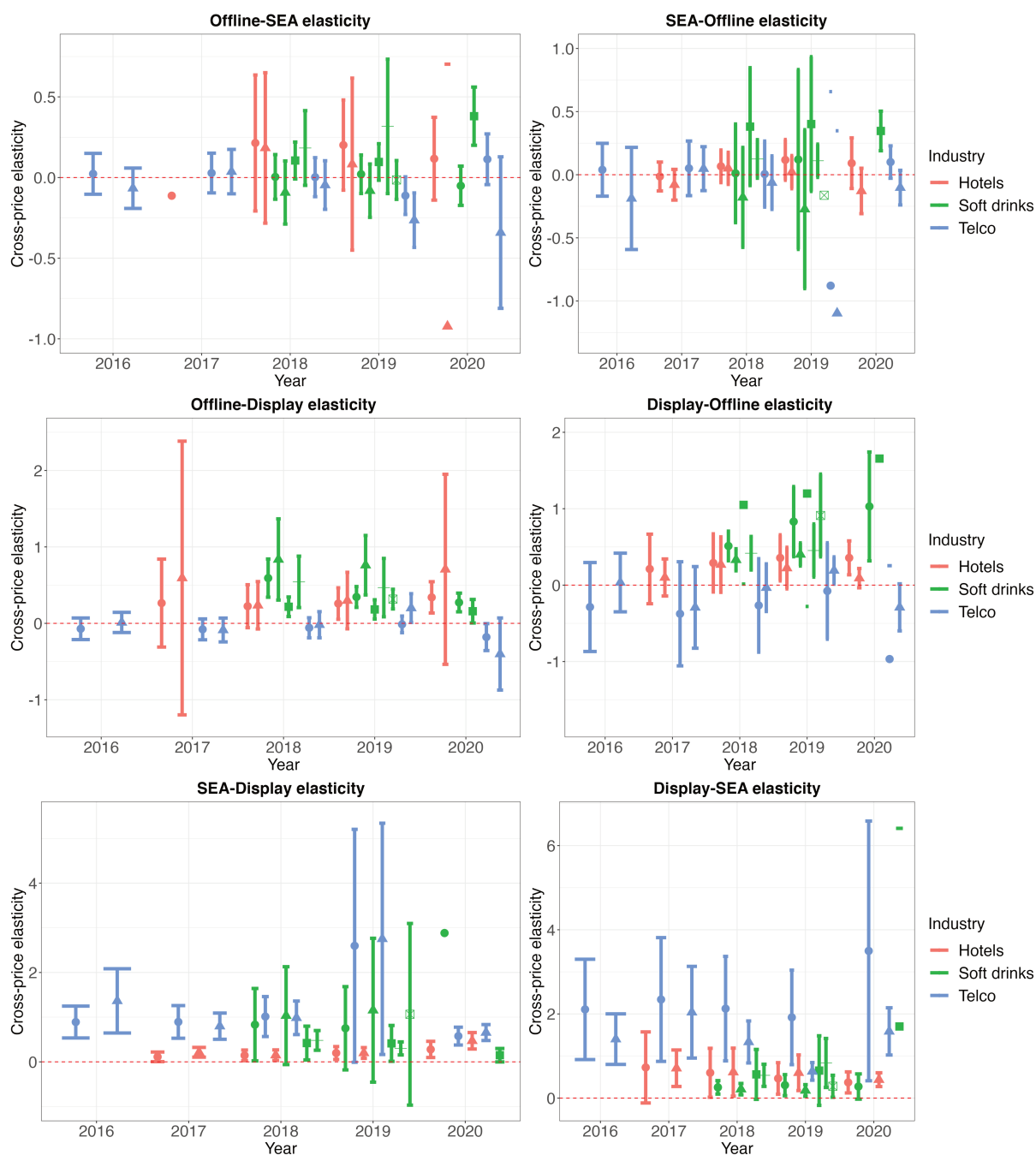
Soft drink As mentioned earlier, in the soft drink industry, most sales happen offline. Thus, the role of online performance campaign is limited in this sector. While offline and display can generate awareness and narratives, point-of-sale advertisements can help to convert in-shop customers. Naturally, offline and search appears to be independent in this context. Offline and display are classified as substitutes since both media are used in branding rather than performance purposes. Finally, search and display appear as rather independent advertising media. It is however hard to tell whether SEA and banners are complement or substitutes given their respective low-share of advertising budget. However, most elasticities are positive and significant under a 10% confidence interval.

4.3 Robustness Checks

Some robustness checks are provided to ensure the quality of the results. First and foremost, we challenged the reserve prices estimation by running models where missing prices are discarded. In all cases, the elasticities which result from the model yield similar results summarized in Figure 2.4.

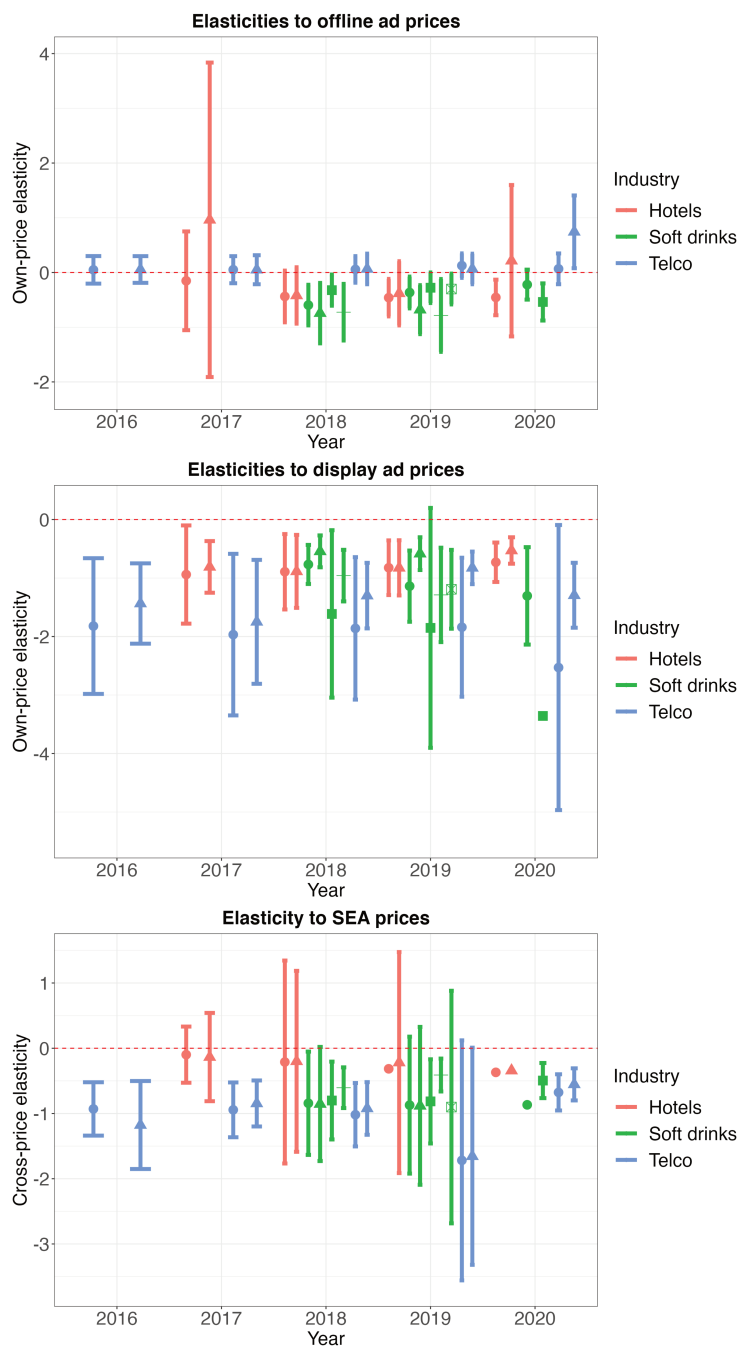
We also challenge the split operated between offline, display and search media. We run a model where offline media is split into broadcast (audiovisual) and printed (newspapers, magazines and outdoor display) media. Because all brands do not advertise on both broadcast and printed media at the same time, we had to aggregate all firms together in order to deal with the missing price issues discussed earlier. Results are displayed in Figure 2.7. It shows that (i) printed and broadcast media are significant substitutes, (ii) display is a closer substitutes to printed than broadcast ads and (iii) neither broadcast nor printed media were substitutes to search ads.

FIGURE 2.4 – Cross-price elasticities of demand by brand and year



Note : Offline-display elasticities (top left), Offline-Search elasticities (top right) and Display-Search elasticities (bottom left) are reported with their 95% confidence intervals. Each color represents an industry and each shape is a brand. The confidence interval of the point is not represented when it is too large

FIGURE 2.5 – Own-price elasticities by brand and year



Note : Elasticities to offline (top) ad prices, display ad prices (middle) and search prices (bottom) are reported with their 95% confidence intervals. Each color represents an industry and each shape is a brand. The confidence interval of the point is not represented when it is too large

5 Implications and Concluding Words

Our study sheds light on advertising media substitution patterns in the digital area. The results presented above depict a fragmented advertising market. Despite the heterogeneity among

substitution phenomena, we conclude that (i) search and offline advertising are separate markets, (ii) display and (printed) offline ads are substitutes and (iii) display and search ads can be used as substitutes. These results are of importance for at least two reasons. First, they temper some previous studies which argued that offline and online ads were substitutes without relying on cross-price effects. Second, they provide useful insight on advertising media concentration by showing that offline and online ad markets are rather separate. To the best of our knowledge, we are the first to investigate substitutions between search and display ads. The fact that they seem to be substitutes has strong implications for competition and concentration among digital ad-based platforms.

However our study has limitations. First, our sample only includes big advertisers which constitutes a bias. It is common knowledge that online advertising is mainly used by small and medium enterprises. Second, display ads are not homogeneous and can serve different purposes. Especially, display advertisements can be used to induce immediate sales (for instance retargeting campaigns) or build brand image (via YouTube or Instagram ads). While our study hints display and search ads are substitutes, this may only be the case for the first kind of online display campaigns.

Globally, the translog cost function framework offers a flexible and robust way to estimate cross-price effects in advertising. Future research would gain from being replicated on more diverse and aggregated data such as the one reported by Kantar or Nielsen. Advertising substitution research is still a promising topic given its importance in media concentration.

6 General Appendix

6.1 Shephard's Lemma

Taking the derivation of equation (2.2) w.r.t to p_i yields :

$$\frac{\partial \ln(C)}{\partial \ln(p_i)} = \frac{\partial C}{\partial p_i} \frac{p_i}{C} \quad (2.6)$$

Derived from microeconomic theory, Shephard's lemma states that the derivative of the cost w.r.t to the price of an input gives the optimal quantity of that input to use. In other word, deriving the total advertising cost w.r.t to media i 's price results in X_i , the optimal quantity of advertising message to produce.

$$\frac{\partial C}{\partial p_i} = X_i \quad (2.7)$$

Substituting (2.7) into (2.6) yields the share S_i of input i in the firm's budget :

$$\frac{\partial \ln(C)}{\partial \ln(p_i)} = X_i \frac{p_i}{C} \equiv S_i = \beta_i + \sum_j \gamma_{ij} \ln(p_j) + \theta_i \ln(q) + \delta_i T. \quad (2.8)$$

6.2 From GRPs to Impressions

Advertising prices are obtained by dividing impressions with expenditures. However, TV and Radio data does not provide impressions but Gross Rating Points (GRPs). GRPs measure the number of impressions delivered to the targeted population out of 100 individuals. A GRP of 20 means that out of 100 individuals exposed to an ad, 20 impressions were made to the targeted population (e.g. 15-29 years old people). In order to convert GRP into impressions, we employ the following formula :

$$M_{it} \approx \frac{\text{GRP}_{it}}{100} \times P_i \times R_i,$$

where P_i is the total French population targeted by the firm watching media i (e.g. the 24-49 years old) and R_i the target's share of audience on the media. Both variables are retrieved from the INSEE and CNC respectively. An advertiser having a GRP of 20 on a population of 20M which is watching TV at 30% will get : $(20/100) \times 20M \times 30\% = 1.2M$ impressions.

7 Model robustness

7.1 First-stage Results

We instrument the normalized quantity of sales of brand k in period t by the following equation

$$\ln(q_{kt}) = \pi_0 + \pi_z \ln(Z_{kt}) + \pi_x \mathbf{X}_{kt} + \mu_k + \tau_t + \epsilon_{kt}$$

where Z_{kt} is the log of consumer or producer price index in the industry of brand k in month t ; X is the vector of other covariates present in the second stage; μ and τ are brand and quarter-year fixed effects. As we can see in Table 2.2, both price indexes can be considered as a valid instrument for most industries given its level of significance and the overall first-stage F-Stat (STOCK et YOGO, 2005).

We employ the PPI to instrument the sales q in the hotel and telecommunication industries. Because PPI is not a relevant instrument for the soft drinks sector, we employ the CPI as an instrument solely in this industry.

TABLE 2.2 – First-stage result

| | Effect on $\ln(q)$ | | | | | |
|-------------------|----------------------|-----------------------|---------------------|----------------------|-------------------|----------------------|
| | Consumer Price Index | | | Producer Price Index | | |
| | Hotels | Soft drinks | Telco | Hotels | Soft drinks | Telco |
| $\ln(\text{CPI})$ | -0.039 (0.038) | -21.983*** (5.974) | -0.958** (0.464) | | | |
| $\ln(\text{PPI})$ | | | | -2.380*** (0.593) | -0.289 (4.551) | -1.149*** (0.135) |
| Observations | 64 | 115 | 98 | 64 | 115 | 98 |
| Adjusted R^2 | 0.942 | 0.761 | 0.970 | 0.952 | 0.769 | 0.972 |
| F Statistic | 35.355 | 14.109 | 87.230 | 43.115 | 14.682 | 96.021 |

Note :

*p<0.1; **p<0.05; ***p<0.01

Note : the table display regression results with heteroskedasticity-robust standard errors. It controls for the covariates of model (2.2) (q excluded). Brand and quarter fixed effects are included.

7.2 Correcting for Autocorrelation

Auto-correlation may appear in equation (2.2) and (2.4) because of the effect of omitted variables from the past period. This is likely to be the case in advertising. In particular, past advertising may affect current sales and investment (adstock effect). Following the literature, we correct for auto-correlation by implementing the lagged error terms ε s in both (2.2) and (2.4) (SELDON et JUNG, 1993; SELDON, JEWELL et al., 2000; SILK et al., 2002; FRANK, 2008; HE et al., 2018). In other words, errors from equation (2.2) will be specified as a function of past error terms in the total cost and two share equations :

$$\varepsilon_{c,t} = \rho_{c,c}\varepsilon_{c,t-1} + \rho_{c,o}\varepsilon_{o,t-1} + \rho_{c,d}\varepsilon_{d,t-1} + \xi_{c,t} \quad (2.9)$$

Where the ρ s are first-order auto-correlation coefficients and ξ_t is an independent error term. Similarly, the error terms from the share equation i will include cost equation's past errors as well as lagged errors from all the share equations (2.4) (including its own) will be specified as a function of past error terms :

$$\varepsilon_{i,t} = \rho_{i,c}\varepsilon_{c,t-1} + \rho_{i,j}\varepsilon_{j,t-1} + \rho_{i,i}\varepsilon_{i,t-1} + \varepsilon_{j,t-1} + \xi_{i,t}. \quad (2.10)$$

Durbin-Watson tests made before the treatment indicate the presence of a positive auto-correlation in both share equations at the 5% thresholds. Tests made after the correction exhibit high p-value, indicating auto-correlation has been mitigated.

TABLE 2.3 – Durbin-Watson tests before and after specifying standard errors as in equation (2.9)

| Equation | Hotels | | Soft drinks | | Telco | |
|--------------------|----------|-------|-------------|-------|----------|-------|
| | Before | After | Before | After | Before | After |
| Cost equation stat | 2.616 | 2.398 | 1.921*** | 2.008 | 2.431 | 2.454 |
| p-value | 0.753 | 0.409 | 0.035 | 0.078 | 0.653 | 0.675 |
| Offline share stat | 1.595*** | 1.899 | 1.410*** | 2.231 | 1.632*** | 1.942 |
| p-value | 0.017 | 0.177 | 0.0002 | 0.803 | 0.014 | 0.243 |
| Display share stat | 1.754 | 1.954 | 1.398*** | 2.206 | 1.517*** | 2.049 |
| p-value | 0.073 | 0.241 | 0.0001 | 0.763 | 0.003 | 0.375 |

Note : p-values are presented below the statistic : ***p<0.05

7.3 Concavity Checks

For the cost function to be concave with respect to prices, the Hessian matrix of the total cost function $\nabla^2 C$ has to be negative semi-definite. The $(i, j)^{\text{th}}$ element of the matrix is :

$$H_{ij} = \gamma_{ij} + S_i S_j - \mathbb{1}\{i = j\} S_i,$$

where S_i is the share of input i . This matrix is different at each point, given that S_i and S_j vary in the data. The matrix will be negative semi-definite only if all its eigenvalue are inferior or equal to 0. Using the estimated parameters γ_{ij} we can compute H_{ij} at each point of the data and then test for concavity. First, we check concavity at the reference point by simply replacing the S_i and S_j by β_i and β_j . Then, for each observations of the data, we check (S_i then becomes S_{ikt}) whether all the eigenvalues are inferior or equal to zero. Then, we investigate concavity on the subset of the data where the S_i are all positive.

As we can see, the reference point satisfies the concavity constraint. However it is not the case of all the data points. The share of observations satisfying the concavity condition greatly increases when we deleted observations where at least one input share is null. Thus $S_i = 0$ harms the concavity of the model. We explain how we deal with missing prices and share in the next section of the appendix.

TABLE 2.4 – Results of the concavity test

| Sample | % of concavity |
|-----------------|----------------|
| Reference point | 1/1 = 100% |
| Full sample | 103/304 = 34% |
| All $S_i > 0$ | 103/164 = 63% |

The percentage indicates the share of observations for which $\max(\nabla^2 C) < 0$ in the data. Eigenvalues were rounded to 0 if they were below 0.01.

7.4 Dealing with Unobserved Prices

The more straightforward remedy to missing prices is to delete zero-price observations. However, the drawbacks of this method however are (i) a risk of selection bias and (ii) a decrease in the internal validity of our results. We employ several alternative methods to deal with zero prices.

Counterfactual prices First, we use information on observed prices to estimate counterfactual prices. This is done by specifying the following regression for each media :

$$\widehat{\ln(p_{ikt})} = \alpha_i + \beta_i \ln(A_{kt}) + \mu_{ik} + m_{it} + y_{it} + \nu_{ikt},$$

where A_{kt} is the sum of competitors' expenditures. The hypothesis behind this model is that competitors affect firm's k prices by two channels. First, if brands are price taker, the ad investment of competitors are correlated with media prices. This is likely to be the case offline. Second, there are situations where competitors' spending may impact a firm's ad prices. This is typically the case in online ad auctions whenever firms target the same contexts. For each media i , we estimate the model on the subset of observations where prices are non-null. We then predict prices on the subset of remaining observations.

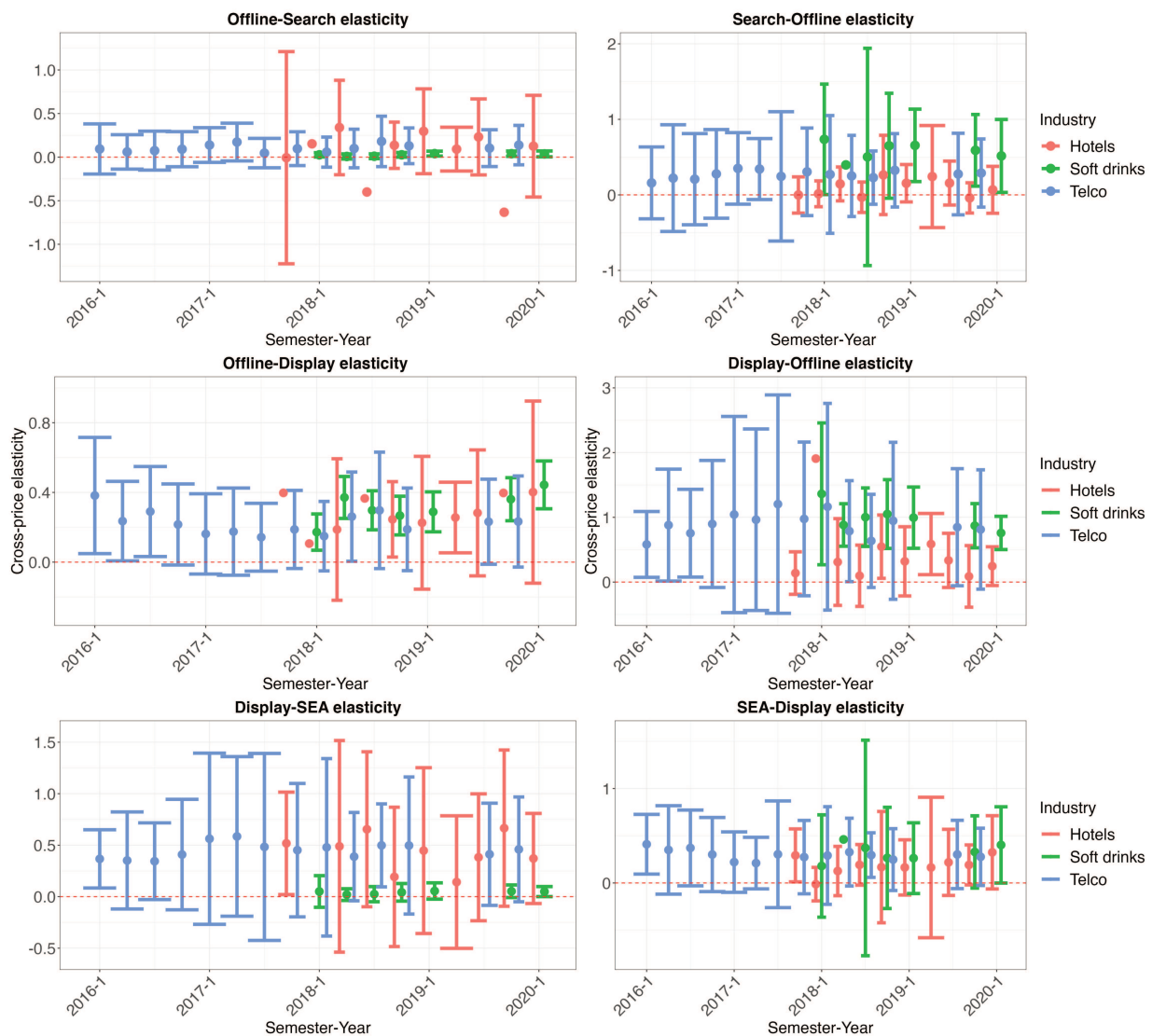
Deleting zero-input observations As a first robustness check, we also run the equation system after deleting all observations for which the quantity of advertising bought is null. However, doing so we introduce a selection bias by only keeping weeks for which brands advertise on all media. To counter this issue, we aggregate all brands together inside each industry⁷. The underlying assumption here is that the brands inside an industry substitutes media in a similar fashion. This is plausible given brand elasticities are similar inside each industry (see Figure 2.4). Corresponding elasticities are summarized by semester-year in Figure 2.6. Although confidence intervals are large (due to the lesser number of observations), they are consistent with our main estimates.

7. However, we still had to drop one firm in the telecommunication sector because of too many missing prices in search advertising that would have created a selection bias

TABLE 2.5 – Prices regressions results

| | $\ln(p_{\text{off}})$ | $\ln(p_{\text{display}})$ | $\ln(p_{\text{search}})$ |
|-------------------------|-----------------------|---------------------------|--------------------------|
| Competitors' spendings | -0.117*** (0.024) | 0.011 (0.016) | -0.027*** (0.005) |
| Brand FEs | ✓ | ✓ | ✓ |
| Quarter-Year FEs | ✓ | ✓ | ✓ |
| Observations | 260 | 316 | 344 |
| Adjusted R ² | 0.215 | 0.353 | 0.341 |

FIGURE 2.6 – Cross-price elasticities and their 95% confidence intervals after deleting observations where all media are not observed. Data are grouped by industry

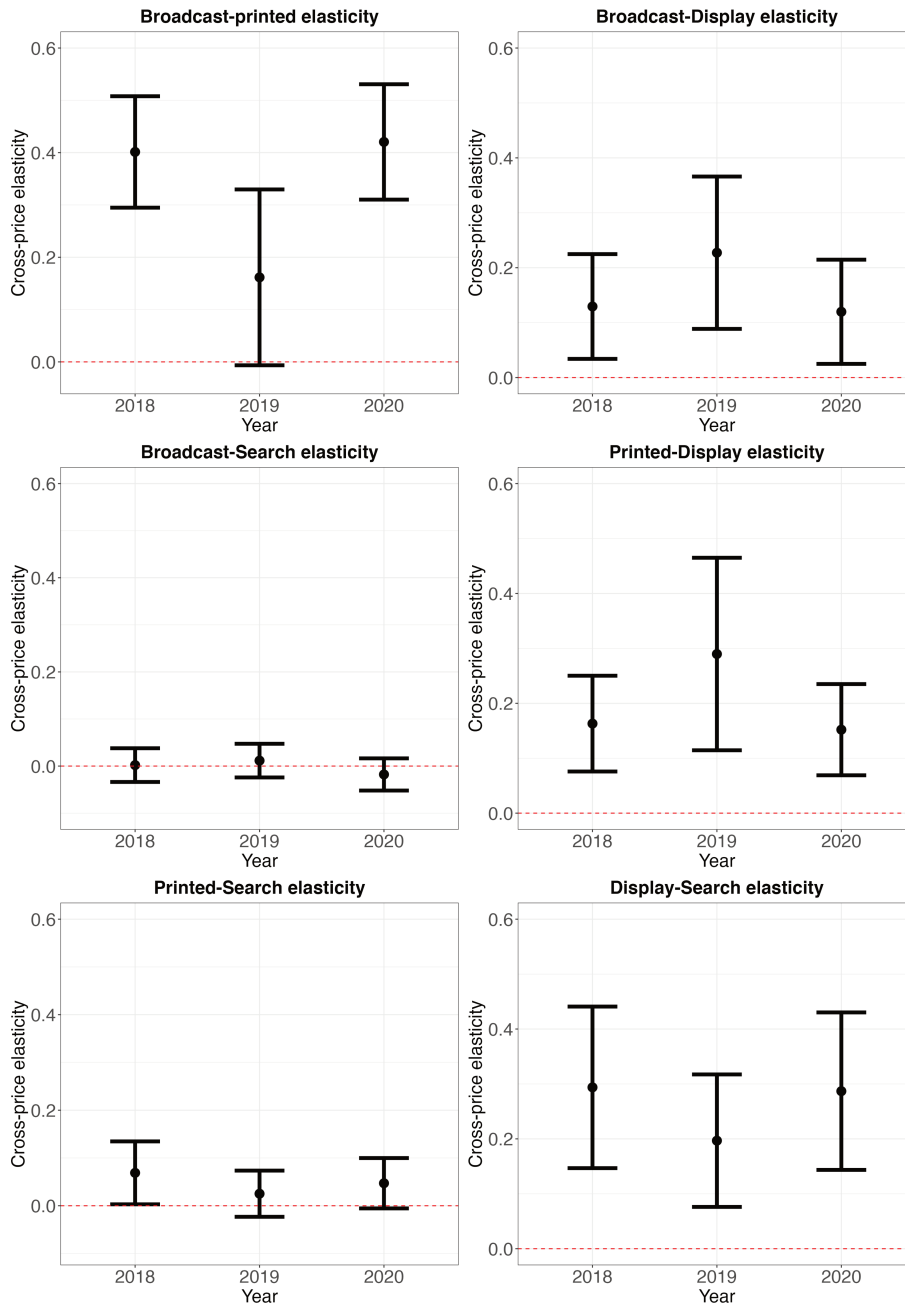


8 Additional Results

8.1 Four-inputs Model

This section presents results of a four-inputs model. The previous offline media variable is split into audiovisual, or broadcast, (TV, radio and cinema) and printed media (outdoor and press advertisement). We run the system of equation with these four inputs and apply the different correction discussed above (endogeneity, autocorrelation, missing prices). However, because all industries do not often advertise on printed and broadcast media, we end up with many zero (especially for printed media). Thus, we estimate the translog model on the whole dataset, after aggregating all brands and industries together. The results and elasticities are presented below in Figure 2.7. The model fitted well even though additional features have been added.

FIGURE 2.7 – Cross-price elasticities with four inputs (offline broadcast, offline printed, display and search media)



95% confidence intervals are reported. Results from a translog cost estimation where all brands and industries are aggregated to avoid no consumption of an input

Chapitre 3

Display Advertising : How Context Matters ?

Abstract

Online advertising relies on a complex network of intermediaries between advertisers and publishers. Often, advertisers do not know where their ads appear which can lead to placement issues. An ad may never be seen by the user, or displayed in irrelevant, perhaps degrading contexts. Cost-per-view contracts, contextual targeting and premium ad inventories are the best way to avoid placement issues. Investigating how these different mechanisms affect ad outcomes, I leverage two databases, each containing millions of ad impressions bought in 2019 and 2021. Using a first dataset on ad viewability, I find that the use of cost-per-impression (CPM) contracts reduces viewability rates by 13%, compared to view-based contracts. Programmatic advertising also exposes advertisers to less viewed ad inventories. Using a second dataset on advertising targeting, I find that matching ads with website contents yields a click rate 69% higher than individual targeting based on users' personal data. Finally, while premium ad inventories could crowd-out quality ad spaces from standard markets, we find they exert a positive effect on standard inventories.

Résumé

La publicité en ligne repose sur un complexe réseau d'intermédiaires entre annonceurs et éditeurs. Les annonceurs ne savent pas toujours où leurs publicités apparaissent, ce qui dégrade le placement des publicités. Par exemple, certaines publicités ne sont jamais vues. D'autres sont affichées dans des contextes non pertinents, voire dégradants. Les contrats de coût-par-vue effective, le ciblage contextuel et les inventaires premium sont le meilleur moyen d'éviter ces problèmes de placement. Pour étudier comment ces différents mécanismes affectent les retombées publicitaires, j'exploite deux bases de données contenant des millions d'impressions publicitaires entre 2019 et 2021. À partir d'un premier jeu de données, je constate que l'utilisation de contrats de coût-par-impression (CPM) réduit la visibilité des publicités de 13%, par rapport aux contrats basés sur les vues effectives. La publicité programmatique expose également les annonceurs à des inventaires publicitaires moins consultés. À l'aide d'un deuxième ensemble de données, je constate que l'adéquation des publicités au contenu des sites Web augmente le taux de clics des campagnes (+69% par rapport au ciblage des utilisateurs). Enfin, alors que les inventaires premium pourraient évincer les espaces de qualité du marché standard, ils exercent en réalité un effet positif sur ces derniers.

1 Introduction

Display ads consist of banners and video posted along editorial contents such as online articles, YouTube videos or social media news feeds. With search engine ads, online display is one of the two main form of online advertising. In 2021, display ad spendings amounted to a total of 220 billions of dollars worldwide (ZENITH, 2021). Display advertising consists of matching a multitude of advertisers' message with a multitude of websites supplying ad impressions (hereafter *publishers*). The pricing and allocation of ads is ensured by an auction mechanism involving a network of ad intermediaries such as demand/supply-side platforms, ad exchanges, ad networks or data-management platforms (see Figure 3.1). This industrial organization features a strong drawback : opacity. For example, brands rarely know where their ads appear on the Internet. In many cases, information such as publisher's name, page location or the content of the page are all unknown to the advertiser (PERROT et al., 2022). Display advertising is thus a fertile ground for information asymmetries. Advertising is expected to increase consumer's valuation for goods. However, the absence of contextual information may harm advertising effects by at least two manners.

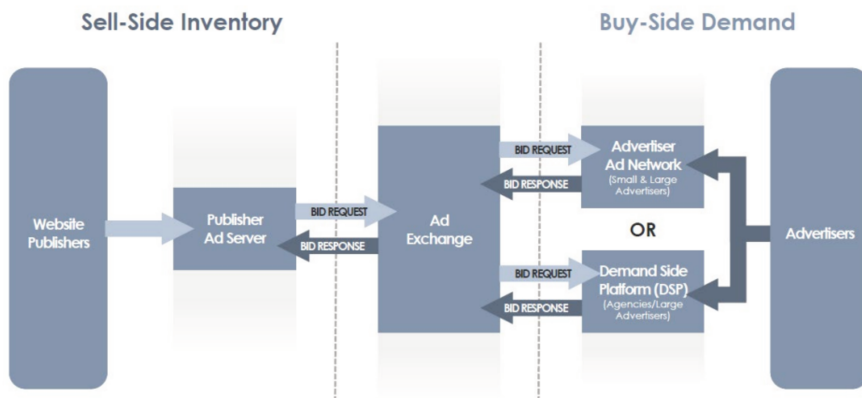
First, ads may not be seen by users. This can be the results of either *ad stacking* (i.e. stacking ads in front of each others), bad placement showing ads in part of the page that are not seen by users or more simply imputable to the non-human audience of a website¹. Second, ads can appear in bad context. Ads associated to controversial contents or low quality websites are likely to be less effective. Advertisers may even risk to hurt their brand image when appearing in bad context : something marketers often refer as *brand safety*. In recent years, notorious brands suspended their ad campaigns on major platforms such as YouTube (in 2017), Facebook (in 2020), or Twitter (in 2022) as they found that their ads appeared along controversial contents. For a long time, lab experiments have shown that media context affects advertising outcomes such as ad recall, attitudes or purchase intents (STIPP, 2018). Yet, the emergence of online advertising fostered the targeting of consumers rather than contexts. However, doing so, advertisers obliterate the advertising environment, which can externally affects their ads.

To limit the information asymmetry regarding ad placement, advertisers have three types of solutions. First, they can engage in pricing models based on viewability (CPVV : cost-per-viewable videos) or performance (CPC : cost-per-clicks). Doing so, ad platforms and publishers have incentive to engage in placement effort in order to make the ad effective.

Second, advertisers can target their advertisement based on the content of the page. By advertising in contexts relevant to its product, an advertiser can match its ads with consumers potentially interested in the product. A famous example of contextually irrelevant advertisement involves a US Airline's ad running on a YouTube video Called "US Airways Crash in Front of My House". Another benefit of contextual targeting is that it avoids the ad to be distributed on non-credible or controversial websites. Initiatives like the Facebook Ad Boycott or Sleeping Giants show that brands do not want to be exposed to controversial contexts. In order to advertise in quality contexts, brands can choose to engage in premium advertising deals with publishers via

1. As we will see, non-human ad impressions can be imputed to bots visiting the publisher's website

FIGURE 3.1 – Intermediaries in Programmatic Advertising (DEPARTMENT OF JUSTICE, 2023)



private marketplaces (PMPs). In programmatic advertising, demand- and supply-side platforms trade ad impressions on Ad Exchanges (Figure 3.1). The latter may be public – meaning that any advertising can bid on the impression – or private. PMP are private Ad Exchanges on which publishers reserve high quality ad spaces to a group of selected advertisers. Once the advertiser is invited by the publisher, it can bid on premium ad auctions against the other invited advertisers. PMPs are becoming more and more popular : in 2021, ad expenditures on PMPs amounted to 13.85 billion dollars and 53% of ads traded automatically in the US according to past forecasts (EMARKETER, 2020). However, no empirical study tackled the effectiveness of PMPs nor its impact on ads sold on the standard market.

To sum up, contexts can be classified in terms of (i) relevance regarding advertiser’s product (horizontal differentiation) and (ii) quality and credibility (vertical differentiation).

I investigate how these different pricing contract, targeting and inventories affect ad viewability and clicks in two case studies, each involving a different database containing millions of ad impressions in 2019 and 2021. The first dataset contains information on the viewability of ads bought in 2021 by a French brand selling paramedical skincare products on the Internet. The brand was buying video ads on cost-per-1,000 impression (CPM), cost-per-viewable videos (CPVV). I find that cost-per-impression (CPM) contracts deter platform’s incentives to make ads viewable : ads bought on a CPM basis are 15% less viewed than the ones bought on cost-per-view contracts. Similarly, when advertisers buy ads from intermediaries which bid on its behalf (a practice known as *programmatic* advertising) ads are 39% less viewable than when the ad is bought from direct deals. Programmatic advertising also exposes advertisers to non-human ad traffic.

The second dataset contains data on the targeting, clicks and cost of ads bought by a renowned perfume brand. Data span from June to December 2019. During that period, the brand either targeted its ads contextually or individually (i.e. based on user’s personal data). The brand also bought ads from PMPs as well as traditional open exchanges. I find that contextually targeted increased both clicks (+69%) and propensity of user to engage with the ad (+18%) compared to user-level targeted ads. Instead of crowding-out quality ads from the standard market, PMPs have positive spillovers on non-premium ads’ click rates. My results show that

publishers better price-discriminate advertisers with PMPs : brands are more willing to pay for standard ads when PMP is available. These empirical results are consistent with the recent theoretical model of W. J. CHOI et SAYEDI, 2022.

More specifically, the study brings several contributions to the existing literature. First, I empirically shed light on incentive misalignment between advertisers and ad-platforms/publishers regarding ad viewability. Second, I quantify both the effectiveness and cost of contextual targeting based on actual campaigns data. Third, to the best of my knowledge, this paper is the first to empirically document the market effect of PMPs.

Section 2 exposes the literature. Data are introduced in Section 3. The econometric approach and results are described in Section 4. Finally, the implications of this study are exposed in Section 5.

2 Related Works

2.1 The Effect of (Contextually) Targeted Advertising

Advertising is often considered as an informative or persuasive tool which shifts consumer’s utility in favor of the advertised good. The view of BECKER et MURPHY, 1993 models advertising as a complementary good of the advertised product, which is more relevant in the online context as it emphasizes the importance of targeting. In their theory, advertising is heterogeneously valued by consumers, and firms should target ads to marginal consumers in order to increase the equilibrium price. The development of internet advertising generated many theoretical studies on the effect of targeting on competition among media and firms (BERGEMANN et BONATTI, 2011; J. P. JOHNSON, 2013). Later, the literature started to draw a distinction between ads targeted contextually or at the user-level.

The first type of targeting which consists in matching ads with the media environment – e.g. topics or keywords mentioned in a webpage – appeared early in the internet (see for example SHAMDASANI et al., 2001) and was already commonly used on the offline world². Contextual targeting takes advantage of users self-selecting through online contents. Indeed, preferences for the websites visited are often correlated with preferences for advertised products (ZHANG et KATONA, 2012). For example, visiting fashion websites or contents indicate strong preferences for apparels. GOLDFARB et TUCKER, 2011b empirically showed that matching ads with relevant website categories increased purchase intent. In addition to this self-selection effect, the media environment may produce externalities on advertising reception. Such contextual effects can be positive when, for instance, ads appear around high-quality brands (DESAI et al., 2014) or trustworthy websites (AGUIRRE et al., 2015). But they can also be negative when ads are shown next to polemical contents (MADIO et QUINN, 2021) or low-quality websites (SHEHU et al., 2021). To sum up, contextual targeting offers two benefits : (i) targeting the consumers with the highest value for the advertisement and (ii) increasing ad’s perceived value by displaying

2. As EVANS, 2009 exemplified : “Consider a business that sells saltwater fishing rods to people who enjoy fly fishing. The traditional approach to matching this buyer and seller involved the creation of a magazine, such as *Fly Fisherman*, with content that attracts the relevant people.”

advertiser’s message in a favorable context. However, a drawback of contextual targeting is that it is unable to discriminate between precise profiles of consumers.

This is why advertisers mainly use leverages consumer’s personal data to address an ad to a given segment of preferences³. We will refer to this practice as *individual* targeting, which we oppose to *contextual* targeting. The goal of individual targeting is to reach the types of consumers that are supposed to be more interested by the product, given their demographics or other personal information. For instance, behavioral targeting uses consumer’s online behaviors to optimally address ads (H. CHOI et al., 2020). Retargeting, which is based on user’s past purchase intent, has been demonstrated to have limited incremental effect on consumers’ willingness to purchase products (LAMBRECHT et TUCKER, 2013; FRICK et al., 2022). While individual targeting may be more effective in matching consumer’s preferences with brand’s product, it obliterates the effect of context in consumer’s valuation of an advertisement. As described previously, relevant contexts match consumers with ad while quality context benefits to brands.

2.2 Ad Viewability

Depending on the format, between 70% and 80% of online display ads are actually viewed by the users, according to INTEGRAL AD SCIENCE, 2022. If this figure is more optimistic than the viewability rate in the early 2010s, it still implies that at least one fifth of ads are never seen by users⁴. Viewability standards are edited by the Media Rating Council (MRC). For video ads – which are of interest here – an ad impression is defined as *viewed* whenever the user saw (i) at least 50% of the ad pixels *and* (ii) two continuous seconds of the video (MEDIA RATING COUNCIL, 2014). Whether ads are visible or not is a hot topic for marketers. Yet, only a few researches tackled the determinant and consequences of ad viewability.

In their early paper, BOUNIE et al., 2017 showed that adopting ad viewability measurement technology increases advertiser’s revenues and ad prices, which also benefited to publishers. However, more visible ads decrease consumer’s utility to browse the publisher’s website. In the presence of ad-blocker, publishers thus have incentive to maintain viewability rates below a certain threshold to deter the use of ad-blockers.

UHL et al., 2020 draw the line between viewability and effectiveness, showing that for the ad to be recognized, it has to be either highly visible on the screen (in share of pixels) or shown for a certain amount of time (in seconds) to users. Thus, if 100% of visibility is not a prerequisite to achieve ad effectiveness, their study confirms that the MRC guidelines are a necessary conditions for the ad to have effects.

As pointed out by these two studies, CPM contracts do not give ad sellers any incentive to enhance ad viewability since they are paid whether the ad is shown or not. The use of cost-per-viewable video impression (CPVV) or cost-per-complete view (CPCV) should theoretically mitigate the issue.

Studying information asymmetries and ad viewability, BALOCCO et LI, 2019 showed that publishers tended to allocate the most viewable ad spaces to private contracts where they could

3. As we’ll see later, both technologies are not fully antithetical and can be used in combination.

4. As we’ll see later, the data presents an average viewability rate of 71.6%.

extract more of advertiser’s surplus. The remaining less visible impressions were sold through real-time bidding advertising auctions. The study concludes that ad viewability is a way for publisher to leverage asymmetric information on ad quality and advertising context.

Another risk associated with online ad allocation is invalid traffic. The term refers to techniques leading to inflate an ad audience (MEDIA RATING COUNCIL, 2020). It can be unintentional : for instance, impressions seen by crawlers, bots or hacked devices. Invalid traffic can also be part of ad fraud techniques : generating fake traffic make advertisers pay for impressions that are useless to them. By delegating ad placement to programmatic networks, advertisers expose themselves to invalid traffic.

2.3 Incentive Misalignment in Online advertising

Display ad serving relies on a complex chain of intermediaries. As a results, firms rarely know where their ads exactly appeared and who saw it. More importantly, the majority of ads are sold through real-time automated auctions performed by an ad platform which bids on behalf of an advertiser. The latter practice is known as programmatic advertising. While the overall organization of the online display industry reduces transaction costs in the ad buying process and allows a finer targeting of users, it also comes with several informational issues.

First and foremost, online display platforms may allocate ads in placement that are not relevant for brands, especially in programmatic buying. SHEHU et al., 2021 showed that programmatic advertising could decrease ad effectiveness, especially for premium brands’, by showing ads in low-quality websites (like flash game websites or content aggregators).

Second, there may be misaligned incentives between advertisers and ad platforms depending on the type of contracts used. Theoretical models investigated how cost-per-impression, cost-per-clicks or cost-per-action contracts induced different incentives for advertisers and ad platforms (ASDEMIR et al., 2012 ; HU et al., 2016). FRICK et al., 2022 recently demonstrated that CPA contracts induced ad platforms to target consumers with a high baseline probability to buy the product, independent of the ad. The same problem applies to ad viewability (BOUNIE et al., 2017). Because many ads on the web are not viewable (or not enough), cost-per-viewable impressions (CPVV) contacts have been implemented in place of CPM. However, no study investigates the effect of such contracts on advertising viewability or effectiveness.

Third, publishers may be involved in different selling channels. For example, they may sell ad impressions through real-time bidding auctions and private contracts. In such situation, the publisher allocate best ads on a channel at the expense of the other, reducing ad value on a part of the market (LEVIN et MILGROM, 2010 ; SAYEDI, 2018 ; BALOCCO et LI, 2019). Recently, a new selling channel has emerged : private marketplaces (hereafter PMPs). While any firm can bid on an *open* exchange, a PMP reserves premium impressions to a list of advertisers invited by a publisher. Practically, a PMP allows advertisers to avoid ad fraud, access privileged sections of a site (e.g. homepage, top banners) and use creative format better integrated into the page (native advertising). Once invited, advertisers then bid for these premium placements. The resulting cost is often higher for advertisers as publishers can specify a floor CPM in PMPs’ auctions. In a recent theoretical model, W. J. CHOI et SAYEDI, 2022 showed that private

exchanges implement an information asymmetry between invited and not-invited advertisers since the latter potentially end up with lower quality (perhaps fraudulent) impressions. A risk is that publishers selling both on a PMP and an open market allocate all their lemons to the open exchange. However, their model also finds that PMP creates a competitive pressure that could force open exchanges to enhance the quality of their ads. Even though they concentrate 53% of programmatic ad budgets, to the best of my knowledge, no empirical research explicitly tackled their effect on advertising effectiveness and prices.

3 Data

I investigate the effect of pricing and context on advertising viewability, cost and effectiveness through two case studies involving two different databases. Summary statistics for both datasets are reported in Table 3.2. I describe each dataset separately in the following sections.

3.1 Dataset 1 : Ad Pricing and Viewability

The first dataset studies online video ads from a French brand selling paramedical goods (mainly adhesive tape and cold sore solutions). This kind of products has the attribute to be sold and advertised online : low price, packaged goods, easy for consumers to assess product's characteristics. The brand studied here is distributed on Amazon and online retailers' website. It is advertised between July 2020 and March 2021 through video ads diffused on various websites (such as YouTube, Amazon or health-related publishers). The ads consist of two formats : *pre-roll* (played before video contents) and *in-read* (placed inside the content of a webpage).

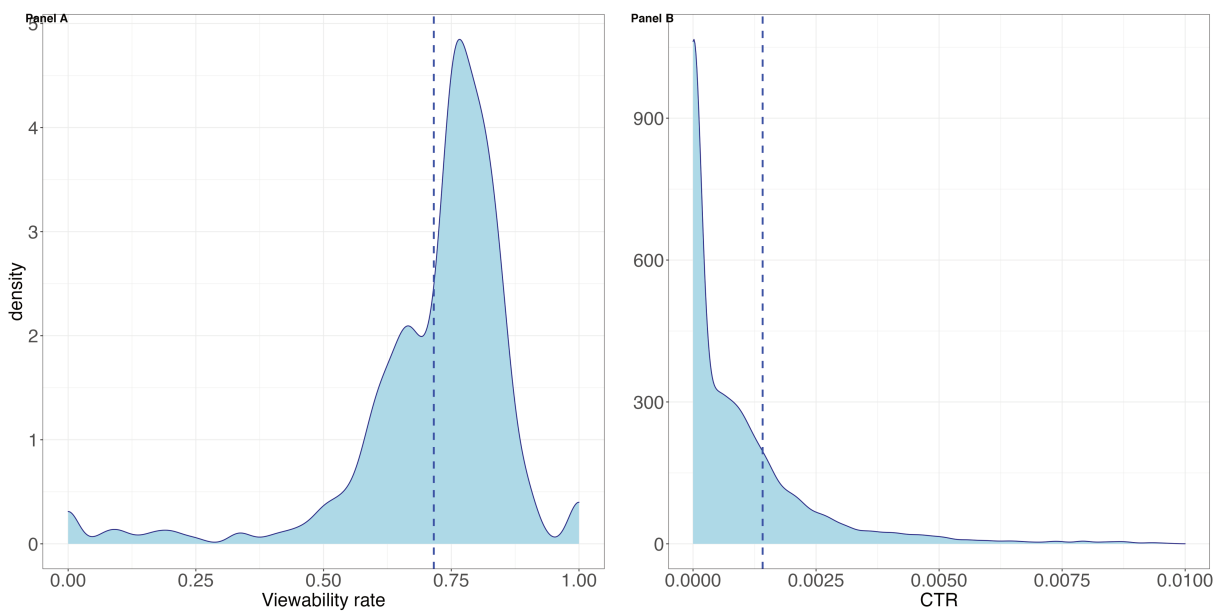
The dataset is an extract retrieved from a third-party measurement firm assessing whether the ads were eventually seen by users. For each week and ads the data contains the number of viewable impressions according to MRC's guidelines mentioned before (MEDIA RATING COUNCIL, 2014). In the data, the average view rate is of 71.6% (Figure 3.2). This is a bit lower than the 80% claimed by industry experts for video ads (INTEGRAL AD SCIENCE, 2022). As we can see on Table 3.1, the overall view rate of ads bought on CPM contracts is 15% to almost 20% lower compared to ads bought on CPVV and CPCV. This is intuitive since in CPM contracts, ad platforms do not have incentive to make ads viewable. As shown on the second plot, programmatic ads bought on CPM are never measured, although the brand did buy programmatic ads on a CPM basis.

Using ads IDs, I am able to identify some ad characteristics. The first feature of interest is the pricing. Video ads can be priced through CPM, CPVV or cost-per-complete view of the video (CPCV). As discussed earlier, these contracts give different incentives to foster ad viewability. In our data, 10% of video ads are bought on CPM, 29% on CPCV and 61% on CPVV. Second, the data allows us to identify whether the ad has been bought via programmatic advertising or not. As we can see in Table 3.2, 40% of ads are bought on programmatic.

TABLE 3.1 – Viewability rate by pricing contract

| | CPM | CPCV | CPVV |
|-----------|--------|--------|--------|
| View rate | 0.56 | 0.70 | 0.75 |
| | (0.21) | (0.13) | (0.16) |

FIGURE 3.2 – Density Distributions



Panel A exhibits the distribution of ads' viewability rate retrieved from dataset 1. Panel B shows click-through rates (CTR) distribution from Dataset 2.

TABLE 3.2 – Summary Statistics of both Datasets

| Dataset 1 : Ad Viewability | | | Dataset 2 : Ad Targeting | | |
|----------------------------|--------|--------|--------------------------|--------|--------|
| Variable | Mean | SD | Variable | Mean | SD |
| Impressions | 11,830 | 25,472 | Impressions | 12,805 | 36,298 |
| Viewable Rate | 0.72 | 0.17 | Clicks | 15.23 | 73.34 |
| Valid Traffic Rate | 0.97 | 0.14 | Contextual | 0.15 | 0.36 |
| CPM | 0.10 | 0.30 | Open | 0.52 | 0.50 |
| Programmatic | 0.40 | 0.49 | Both | 0.02 | 0.15 |

3.2 Dataset 2 : Ad Context and Effectiveness

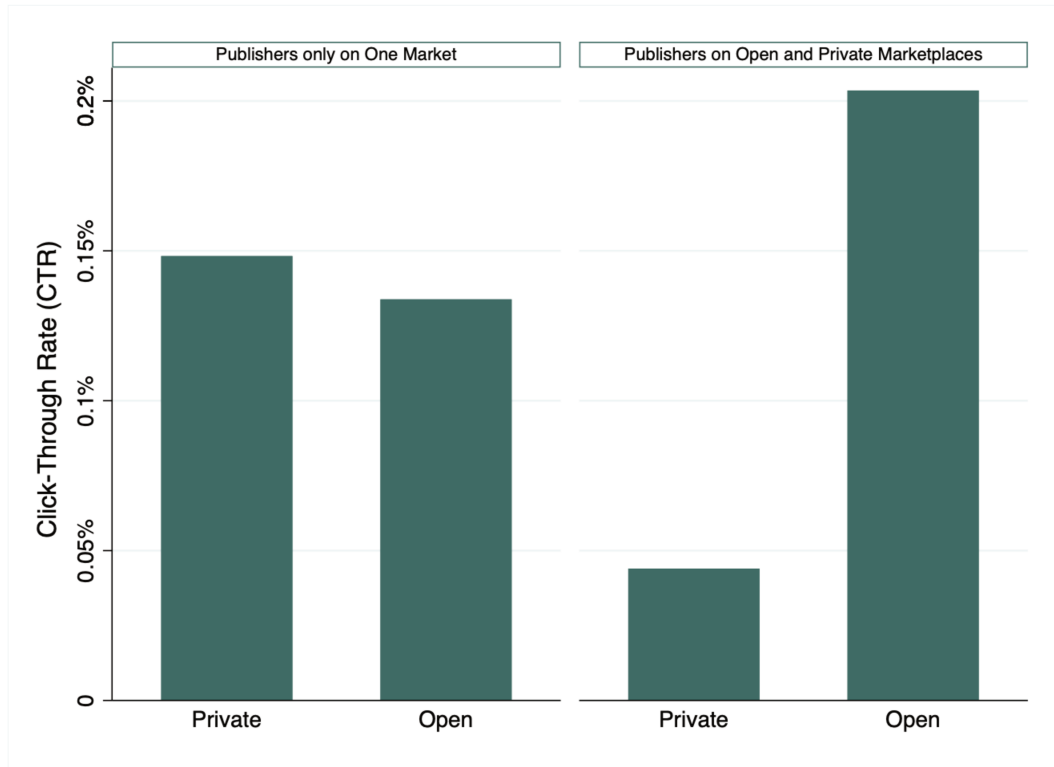
The second dataset is retrieved from two national UK advertisers commercializing perfumes under high-end fashion brands. The data consist of extracts from online advertising banners bought between June and December of 2019.

The database provides details such as advertising targeting, whether the ad is bought from an open or private exchange and the publisher (website) or group of publishers from which the ad is bought. I also have data on the device on which the ad is shown and the ad format. 35% of ads end up on desktop computers, 34% on smartphones and 31% on tablets. Several targeting technologies are identified in the data : contextual, retargeting, socio-demographic and behavioral. In the data, 15% of ads are contextually matched with the page, while the rest is targeted at the user-level.

More than half of ads (52%) are bought from open exchanges. By looking at which publishers sold ads to the brand, and on which market (open or private), we are able to identify that 2% of the ads have been bought from a publisher selling both on an open and private exchange. As shown on Figure 3.3, PMP outperforms open exchanges when comparing publishers only selling on the standard or private market. However, when publishers sell both ads on an open and private ad exchange (thus price discriminating advertisers), ads from the standard market seem surprisingly more clicked.

Along clicks, the database provides traditional statistics such as the impressions and cost associated to each advertisement. Using clicks to assess an ad effectiveness is a straightforward strategy used in previous studies (TUCKER, 2014; JEZIORSKI et SEGAL, 2015; SHEHU et al., 2021). A problem is that, as shown in Figure 3.2, many ads are never clicked. We will come back on this specificity later in the identification strategy. More generally, ads are associated to a very low click-through rate (0.14% in average).

FIGURE 3.3 – Effect of Ad Inventory Type on Click-Through Rate Depending on whether Publisher Sells Only on One Marketplace (Open or Private) or on Both



4 Econometric Analysis

4.1 General Approach

While the two datasets differ in period or industry covered, they follow a similar structure. To investigate how ad targeting, buying and pricing affect ad effectiveness as measured by views and clicks, I employ the following regression model, specified as :

$$\ln(y_{it}) = \beta X_i + \gamma Z_{it} + \theta_{c(i)t} + \nu_{it}. \quad (3.1)$$

Dependent Variable In the equation above, y_{it} is the outcome of ad i bought in week t . These outcomes can be the ratio of views, valid impressions and clicks over impressions depending on the model. Because these explained features are comprised between 0 and 1, the ideal specification of the dependent variable y would be a log or logit transformation.

Features of Interest We are interested in the relationship between ad outcomes y_{it} and the vector X_{it} which denote the ad-specific features of interest such as ad pricing targeting or whether the ad has been bought from an open exchange or a PMP. I will provide further detail on the variables in the next sections.

Other Covariates One can note that those feature of interest are time invariant : the pricing or targeting of a given ad is the same across time. To identify the causal effects of X_i , I thus control for other ad-specific characteristics such as the ad platform, ad format used or the device on which the ad appears (e.g. desktop, mobile, tablet). These controls are included in the vector Z_{it} .

Fixed Effects Because many unobservables may affect ad outcomes, I add campaign \times week \times year fixed effects θ_{ct} . Many marketing decisions (like the ad copy used or the objectives of the ad) are taken at the campaign level, thus θ_{ct} neutralizes the effect of a specific campaign across time, as well as seasonality. Because the treatments (pricing, targeting, inventory) are assigned at the ad-level, errors ν_{it} are clustered by ads (ABADIE et al., 2022).

4.2 Case Study 1 : The Effect of Pricing on Viewability

I use the regression model (3.1) to identify how pricing contracts and programmatic advertising both affect ad viewability. I regress the log of view rate on two dummies of interest X_i indicating (i) the pricing used (CPM vs CPVV/CPCV) and (ii) whether the ad has been bought by programmatic advertising. The model controls for the ad platform involved in the transaction, the format of the video (pre-roll vs in-read) as well as campaign-time fixed effects θ_{ct} . I regress the the log of valid traffic rate on the same variables. Both models are estimated by OLS.

Results are reported in Table 3.3. As expected, I find that CPM contracts are associated with a viewability rate 18%⁵ lower than other contracts (CPCV, CPVV). This is intuitive since in CPM, advertisers are charged each time their ads are loaded on the web-page, independently of their viewability. This is close to what is observed in the summary statistics (Table 3.1). Programmatic buying also seems to decrease ad viewability. It would be interesting to see the interaction effect of Programmatic ads and CPM contracts on ad viewability. However, as explained previously, programmatic ads bought by CPM were not measured by the third-party viewability vendor.

Regarding traffic rate, the effect of CPM disappears (it is not significant under a 5% confidence interval). This is expected : impression and view-based contracts do not affect the non-human traffic received by the ad. In both cases, the ad platform and publishers are remunerated, as opposed to cost-per-clicks or cost-per-action contracts which require human actions. Column (4) shows that programmatic advertising however reduces the share of human traffic by approximately 87%. Ads allocated through real-time automated auctions are more likely to be placed on sites which receive fraudulent impressions. This could be a consequence of the lack of transparency generated by programmatic advertising.

5. Marginal effects are computed by $\exp(\beta) - 1$ with β the coefficient of the variable of interest

TABLE 3.3 – Econometric results : effect of pricing and programmatic buying on ad viewability

| | log Ad Viewability Rate | | log Valid Traffic Rate | |
|-------------------------|-------------------------|-----------------------|------------------------|----------------------|
| | OLS | OLS | OLS | OLS |
| | (1) | (2) | (3) | (4) |
| CPM | -0.123** (0.0391) | -0.143*** (0.0386) | -0.263 (0.136) | -0.269 (0.136) |
| Programmatic | | -0.509*** (0.0401) | | -1.355*** (0.137) |
| Ad format | | ✓ | | ✓ |
| Ad network | ✓ | ✓ | ✓ | ✓ |
| Campaign-week FEs | ✓ | ✓ | ✓ | ✓ |
| Observations | 2363 | 2363 | 2363 | 2363 |
| Adjusted R ² | 0.261 | 0.289 | 0.137 | 0.137 |

Note : *p<0.05 ; **p<0.01 ; ***p<0.001

Note : Robust standard errors clustered at the ad-level displayed in parentheses.

4.3 Case Study 2 : The Price and Effectiveness of Context

I use the second dataset to answer two questions. I seek to investigate how (i) the type of targeting and (ii) the ad inventory affect ads' short-term effectiveness (as measured by clicks) and per-impression costs.

Ideally, I would run model (3.1) by specifying y as the log of the click rate (defined as the ratio of clicks over ad impressions). However, 39% of ads in the data are never clicked. Dropping such observations would generate a selection bias by excluding the least effective ads. Many transformations have been used in the literature (log of variable plus one, inverse-hyperbolic sine transformation...), but most of them induce strong bias (see BELLÉGO et al., 2022 for a discussion). Instead, I run three models. First, an OLS regression on the click rate observed on the subset of ads clicked is performed. Then, I complement this first regression with a second probit estimation on the probability to observe at least one click on the ad. As a third solution, I use BELLÉGO et al., 2022's novel iterated OLS (iOLS) estimation in order to deal with the presence of 0 in the log-transformed variable. The method consists in adding a positive observation-specific variation to y such as :

$$y_{it}^* = y_{it} + \Delta_{it}$$

The quantity Δ_{it} is a function of the model features and an hyper-parameter δ : $\Delta_{it} = \delta \exp(\beta X_i + \gamma Z_{it} + \theta_{c(i)t})$. The method and the selection of the hyper-parameter are discussed in appendix 6.1.

The main covariates X_i here are two dummies indicating : (i) whether the ad is contextually targeted or not and (ii) whether the ad has been bought through an open exchange. I also run

a model where I interact the open exchange dummy with an indicator *Both* taking the value 1 if the ad is bought from a publisher operating both on open and private exchanges. The latter variable captures the effect of buying an from the open market when the publisher also sells on the private market. The vector of controls Z_{it} includes the device and site on which the ad is displayed as well as the ad format. Campaign-week-year fixed effects θ_{ct} are implemented.

Columns (1) and (3) of Table 3.4 show that contextual placement yields significant positive returns in all models. Contextual targeting increases click-through rate by 29% (column (1)), the propensity of user to click at least once on the ad by 9% (column (2)) and the number of clicks by 69% (column (3)). However, I find no evidence of contextual targeting being more expensive in terms of CPM. I find a reasons for that : contextual targeting is generally less demanded by advertisers (in the data, only 15% of the ads were contextually targeted). The competition in ad auctions is lower for this kind of placement, which may lead to lower second-price bids paid by the winning advertiser.

As expected, ads bought from open exchanges are significantly cheaper ((4)). However, I do not find that open ads are significantly less effective in terms of clicks. Column (3) shows that the presence of PMPs largely increases click rate for ads on the public market. It also allows the publisher to get a higher CPM from its open ads (column (4)). Thus, not only the PMP does not crowd-out quality spaces (as it positively impacts ads bought from the open market) but it also better price-discriminates advertisers. Although the $\text{Open} \times \text{Both}$ effect on clicks is only significant under the 10% confidence interval, I performed several robustness checks which led to similar conclusions. This result is in line with W. J. CHOI et SAYEDI, 2022's model stating that, in the presence of PMPs, open exchanges had incentives to make effort in order to increase the quality of their ads. The positive effect of $\text{Open} \times \text{Both}$ on CPM suggests that publishers try to capture a part of this effort.

In this case, positive cross-externalities between ads may explain this surprising finding. Indeed, when a publisher operates a private market, it attracts quality advertisements on its site. These quality ads may increase the value of the advertising context and thus positively affect other ads. For example, a perfumer is likely to benefit from being advertised next to a Rolex banner ad. Because perfume and luxury watches do not compete, Rolex creates a positive spillover on perfume brands by generating a prestigious association between both brands.

TABLE 3.4 – Econometric results : effect on context on click rates

| | Ad Effectiveness | | | Ad Price |
|-------------------------|----------------------|---------------------------------|------------------------------------|----------------------|
| | logit(CTR) | $\mathbb{P}(\text{Clicks} = 1)$ | $\ln(\text{Clicks} + \Delta_{it})$ | $\log(\text{CPM})$ |
| | OLS | Probit | iOLS | OLS |
| | (1) | (2) | (3) | (4) |
| Contextual | 0.252* (0.101) | 0.532*** (0.175) | 0.527* (0.247) | -0.0576 (0.109) |
| Open | 0.953*** (0.164) | 0.539*** (0.134) | 0.489 (0.377) | -1.712*** (0.191) |
| Both | -2.315*** (0.201) | 0.663 (0.538) | -0.726** (0.237) | -2.377** (0.869) |
| Open \times Both | -0.469 (0.344) | 0.605 (0.480) | 1.421** (0.511) | 2.304** (0.801) |
| Device & Ad format | ✓ | ✓ | ✓ | ✓ |
| Site FEs | ✓ | ✓ | ✓ | ✓ |
| Campaign-Week FEs | ✓ | ✓ | ✓ | ✓ |
| δ | | | 1 | |
| Observations | 2755 | 4337 | 4498 | 4498 |
| Adjusted R ² | 0.428 | | | 0.757 |
| ϵ | | | 0.000000640 | |

Note :

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Note : Robust standard errors clustered at the campaign-site level displayed in parentheses. Column (1) displays effects on ads click-through rate (CTR) on the subset of clicked ads. Column (2) displays effects on the propensity to click on ads, at least once. Column (3) shows results from the iOLS estimation where δ is the value of the optimal hyper-parameter and ϵ represents the absolute difference in coefficients between the last two iteration of the iOLS model. Finally, column (4) shows effects on ads' cost-per-1,000 impressions (CPM).

4.4 Robustness Checks

I provide robustness estimations regarding these results. I estimate model (3.1) using different transformation such as the inverse-hyperbolic sine (IHS) and the popular $\log(1 + y_{it})$ transformations. Results are displayed in Table 3.5 and 3.6 : they are all consistent with my main estimations. I also show that consumer's reaction to ads are heterogeneous across device in Table 3.9 : contextual targeting seems particularly effective on mobile devices (smartphones and tablets).

5 Conclusion

Industrial Implications In the first study, I find that CPM contracts and programmatic advertising both reduce ad viewability compared to view-based contracts and more direct ad

buying. My results advocate for the use of viewability-based contracts for advertisers, especially in programmatic advertising. The second set of results draw new implications for the targeting of advertising. While contextual matching does not rely on consumer's personal data, it yields higher returns than profiling. This finding echoes previous empirical studies which demonstrated that ads based solely on personal data were less effective. For example, GOLDFARB et TUCKER, 2011b showed contextual display ads increased purchase intent as long as it was not obtrusive. Conversely, LAMBRECHT et TUCKER, 2013 and FRICK et al., 2022 found retargeted ads proved to be ineffective except for consumers already interested in the product. Moreover, user-level targeting exposes advertisers to strict privacy regulations (GOLDFARB et TUCKER, 2011c). Restrictions regarding the use of personal data in advertising have intensified in the past few years and are likely to increase. For example, Apple and Google have recently limited the collection of user's personal data (BLOOMBERG, 2021). Meanwhile, a recent study from ADA et al., 2022 empirically documented that disclosing information about the context in which the ad appears increased advertiser's willingness-to-pay of ad impressions.

Yet, contextual targeting still represent a small share of online ad spendings. In GOLDFARB et TUCKER, 2011b's study, only 10% of the 2.8M of ads they analyzed were contextually matched with the website category. In our data, only 15% of ads are targeted accordingly to the webpage's content. However, by increasing the use of contextual advertising, advertisers and publishers may increase their own profit as well as consumer welfare, resulting in a positive-sum game. This idea is being discussed by practitioners especially since, as shown in this study, the cost-per-impression of contextual targeting is generally less expensive (SCHIFF, 2019). However, the technology underlying contextual advertising is evolving and recent researches stressed that ad products sold to advertisers as "contextual" might also rely on demographic or behavioral data (BLEIER, 2021).

Privacy has implications on market outcomes such as advertising effectiveness or publishers' revenues, with heterogeneous effects regarding firms sizes (BLEIER et al., 2020). Profiling supports a large part of advertising-based models on the internet : as BLEIER, 2021 stresses, contextual advertising may benefit more to notorious and niche site publishers, more able to monetize their context to advertisers. As a result, contextual ads may encourage publishers to produce niche and marketable contents, which more naturally attracts contextual advertising. These potential effects of contextual advertising have implications for consumer's welfare.

Concluding Words This study contributes to assess inefficiencies in the complex industry of display advertising (H. CHOI et al., 2020 ; GORDON, JERATH et al., 2021). While the choices of pricing contracts, targeting technologies and ad inventories matter for advertisers, it also yields consequences for consumers and publishers. However, my study has at least two limits.

First, ad views and clicks are not profit. If viewability is required for the ad to be effective, it is not a sufficient condition. And if clicks can lead to sales, it is not always the case. Metrics like conversion rates, online sales or purchase intent could be more appropriate. However, such individual data are rarely available. Moreover, they also come with inference challenges : display campaigns typically have small effects on sales and thus measuring their effects requires a stag-

gering amount of observations (LEWIS et RAO, 2015). The causal relationship between online ad exposure and individual sales is subject to many biases and even controlled ad experiments often fail to yield consistent estimates (G. JOHNSON, 2022).

An equally importantly limitation is that this empirical study is limited to two particular industries : paramedical goods and perfumes. This difficulty is inherent to advertising where ad effects are intrinsically related to the characteristics of the product advertised. A similar study on other industries could have yield different results. At best, this empirical study can illustrate the existence of certain market failures. But it cannot prove such inefficiencies systematically exist.

6 Appendix

6.1 The iterative OLS (iOLS) Estimation

Let us rewrite the regression equation (3.1) in the following matrix form : $y_{it} = \beta' \mathbf{X}_{it}$, where \mathbf{X}_{it} contains all the model feature and β' is the vector of associated coefficients. BELLÉGO et al., 2022's iOLS method consists in adding an individual-specific positive value to y_{it} such that :

$$y_{it}^* = \log(y_{it} + \Delta_{it}), \text{ where } \Delta_{it} = \delta \exp(\beta' \mathbf{X}_{it}).$$

In the equation above, the individual correction $\Delta_a > 0$ is estimated as a log-linear function of the predictors \mathbf{X}_{it} times an hyper-parameter $\delta > 0$. The regression parameters β' appear both on the right-hand and the left-hand sides of the model. However, they cannot be estimated simultaneously. To estimate an appropriate Δ_{it} , the iOLS algorithm follows an iterative process that can be summarized as :

Algorithm 1. iOLS $_{\delta}$ for a given $\delta \in (0, +\infty)$

1. Estimate the biased naive regression $\log(1 + y_{it}) = \beta^0 \mathbf{X}_{it} + \nu_{it}^0$.
2. Use the resulting regression coefficients β^0 to compute $\Delta_{it}^0 = \delta \exp(\beta^0 \mathbf{X}_{it})$.
3. Run a new regression $\log(y_{it} + \Delta_{it}^0) = \beta^1 \mathbf{X}_{it} + \nu_{it}^1$.
4. Compute a new $\Delta_{it}^1 = \delta \exp(\beta^1 \mathbf{X}_{it})$.
5. Repeat step 3 and 4 with $\Delta_{it}^2, \dots, \Delta_{it}^n$ until convergence.

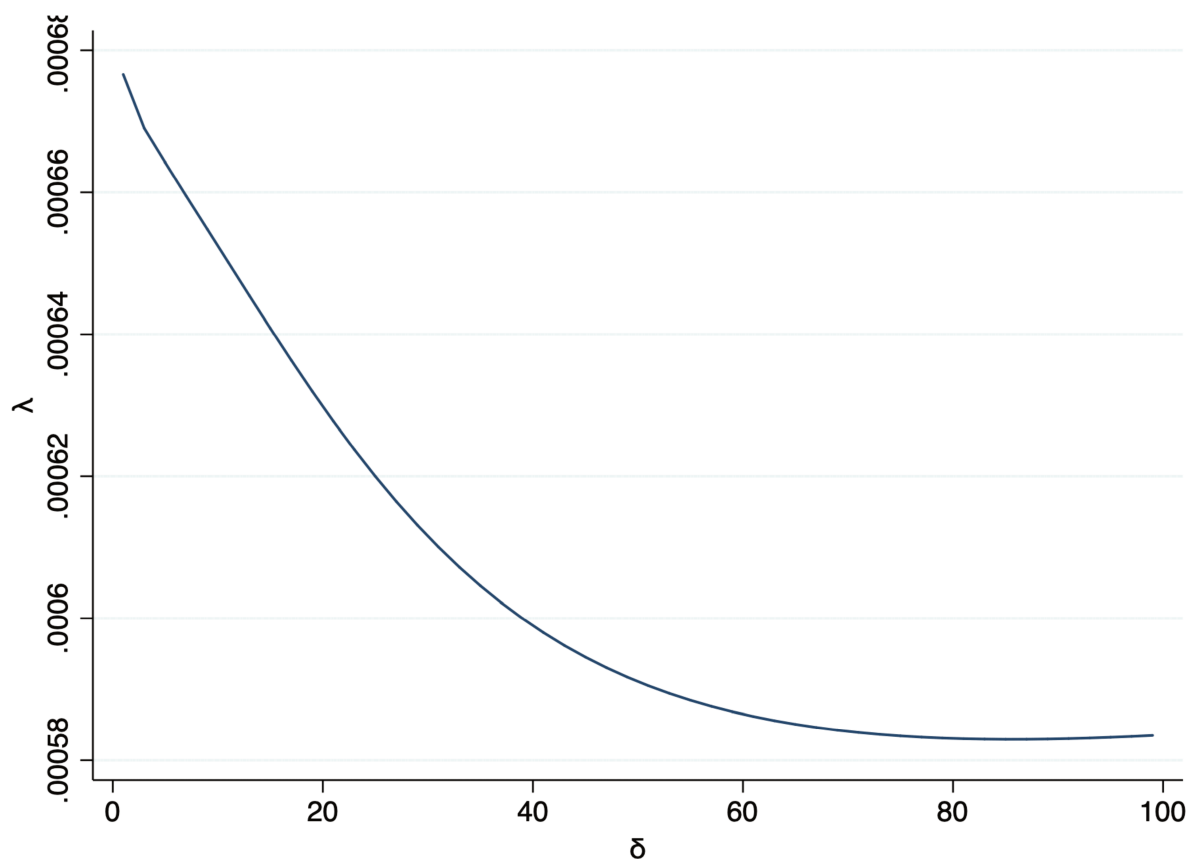
The algorithm stops when the convergence criterion $\epsilon = |\beta^t - \beta^{t-1}|$ passes below a certain threshold before or if the algorithm reach a maximum T iterations. The threshold is set to 1×10^{-8} a the maximum iteration is set to 100.

Obviously, choosing the right δ is important as it affects the set of coefficients β . I use BELLÉGO et al., 2022 data-driven method to find the best δ . The authors suggest to make a choice based on the model's λ_{δ} computed as follow :

Algorithm 2. Choosing optimal δ by computing λ_{δ}

1. Estimate $\mathbb{P}(y_{it} > 0 | \mathbf{X}_{it})$ using a logit or probit model.
2. Compute an iOLS $_{\delta}$ estimate for a given δ and recover the residuals $\hat{\nu}_{it}$
3. Compute $W_{it} = \log \{ y_{it} + \delta \exp(\beta' \mathbf{X}_{it}) - \beta' \mathbf{X}_{it} - \log(\delta) \} \frac{1}{\mathbb{P}(y_{it} > 0 | \mathbf{X}_{it})}$
4. Regress : $\log(\delta + \hat{\nu}_{it}) - \log(\delta) = \lambda W_{it} + \xi_{it}$ to obtain the estimates $\hat{\lambda}_{\delta}$
5. Repeat step 2 to 4 for each δ considered

In practice, the closer λ_{δ} is to 1, the better δ is. To find the δ , a grid search is conducted for $\delta \in (1, 100)$ with a pas of 1. For each value of δ , Algorithm 2 is processed and a λ is returned. The procedure returns an optimal $\delta = 1$ for the model based on dataset 2 (perfume). As we can see in Figure 3.4), λ is decreasing in δ . In one of their application, BELLÉGO et al., 2022 also found that λ could be monotonic with δ .

FIGURE 3.4 – Evolution of λ_δ 

6.2 Robustness Checks

6.2.1 Other Dependent Variable Specifications

In this section, I test different specification of y_{it} which also account for the presence of zeros.

The “popular fix” A popular method to deal with zeros in regressions consists in adding a positive constant to y such that : $y_{it}^* = \log(1 + y_{it})$.

Inverse hyperbolic sine (IHS) transformation I also run the model on another transformation of clicks. The IHS transformation allows to add an individual-specific term to y_{it} . It takes the following form :

$$\text{ihS}(y_{it}) = \log \left(y_{it} + \sqrt{y_{it}^2 + 1} \right)$$

This transformation is less flexible than iOLS and suffers from similar bias than the popular $\log(1 + y)$ (BELLÉGO et al., 2022). Moreover, the results are difficult to interpret. I still use it as another robustness check.

6.2.2 Results

Results from the PF and IHS transformation are reported in Table 3.5 for the models on ad viewability. All alternative specification confirms my previous results : CPM-based pricing and programmatic buying both decrease ad viewability.

TABLE 3.5 – Robustness test testing other specifications

| | log(Views+1) (1) | ihs(View Rate) (2) | ihs(Valid traffic Rate) (3) |
|-------------------------|-------------------------------|-----------------------|--------------------------------|
| CPM | -0.395*** (0.0943) | -0.192*** (0.0501) | -0.112 (0.0612) |
| Programmatic | -1.427*** (0.101) | -0.692*** (0.0521) | -0.661*** (0.0616) |
| Ad impressions | ✓ | | |
| Ad platforms | ✓ | ✓ | ✓ |
| Campaign-Week FEs | ✓ | ✓ | ✓ |
| Observations | 2363 | 2363 | 2363 |
| Adjusted R ² | 0.988 | 0.295 | 0.167 |
| <i>Note :</i> | *p<0.1 ; **p<0.05 ; ***p<0.01 | | |

Table 3.6 reports robustness checks models performed on the second dataset. The first two columns confirm my previous finding : contextual targeting increases ad effectiveness. The effect disappear in the other specification. As opposed to view, click rates are generally very low (below 1%). Thus, as opposed to logit, popular fix and IHS transformations generate small absolute value. The marginal effects measured are hence small and often below the coefficient's standard errors (column (3) and (4)). As pointed earlier in the literature, online ad effects are sometime so small that studies require a high volume of observations in order to get significant estimates. iOLS and logit transformation allows us to introduce more variability in the data and thus derive statistically significant estimates. When considering absolute value of clicks, coefficient are positive, but still too noisy given the high number of zeros in the dependant variable. Specifications (3) to (6) however confirm my surprising finding on PMPs : standard open ads are more efficient when the publisher also own a private market.

TABLE 3.6 – Results for the Popular Fix and IHS Transformations

| | log(Clicks) (1) | logit(CTR) (2) | log(1+CTR) (3) | ihs(CTR) (4) | log(1+Clicks) (5) | ihs(Clicks) (6) |
|-------------------------|----------------------|----------------------|---------------------------|---------------------------|----------------------|----------------------|
| Contextual | 0.175** (0.0792) | 0.254** (0.102) | -0.000323 (0.000585) | -0.000403 (0.000634) | 0.0368 (0.0898) | 0.0809 (0.102) |
| Open | 0.677*** (0.158) | 0.956*** (0.163) | -0.00179 (0.00128) | -0.00197 (0.00139) | 0.110 (0.0998) | 0.111 (0.115) |
| Both | -4.896*** (0.279) | -5.629*** (0.252) | -0.00237*** (0.000347) | -0.00239*** (0.000366) | -2.420*** (0.216) | -2.818*** (0.259) |
| Open × Both | -0.131 (0.328) | -0.467 (0.344) | 0.00368** (0.00154) | 0.00392** (0.00166) | 0.474** (0.219) | 0.589** (0.256) |
| Impressions | ✓ | | | | ✓ | ✓ |
| Ad Device and Format | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Site | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Campaign-Week FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 2718 | 2718 | 4439 | 4439 | 4439 | 4439 |
| Adjusted R ² | 0.710 | 0.428 | 0.0186 | 0.0158 | 0.660 | 0.668 |

Note : *p<0.1 ; **p<0.05 ; ***p<0.01

Robust standard errors clustered at the campaign-site level displayed in parentheses.

6.3 Full Regression Results

TABLE 3.7 – Econometric results : effect of pricing and programmatic buying on ad viewability

| | log Ad Viewability Rate | | log Valid Traffic Rate | |
|-------------------------|-------------------------|------------------------|------------------------|----------------------|
| | OLS (1) | OLS (2) | OLS (3) | OLS (4) |
| CPM | -0.123** (0.0391) | -0.143*** (0.0386) | -0.263 (0.136) | -0.269 (0.136) |
| Programmatic | | -0.509*** (0.0401) | | -1.355*** (0.137) |
| In Read (vs pre-roll) | | -0.0399*** (0.0105) | | -0.0123 (0.0129) |
| Ad network | ✓ | ✓ | ✓ | ✓ |
| Campaign-week FEs | ✓ | ✓ | ✓ | ✓ |
| Observations | 2363 | 2363 | 2363 | 2363 |
| Adjusted R ² | 0.261 | 0.289 | 0.137 | 0.137 |

Note : *p<0.05 ; **p<0.01 ; ***p<0.001

Note : Robust standard errors clustered at the ad-level displayed in parentheses.

TABLE 3.8 – Econometric results : effect on context on click rates

| | Ad Effectiveness | | | Ad Price |
|-------------------------|----------------------|---------------------------------|------------------------------------|----------------------|
| | logit(CTR) | $\mathbb{P}(\text{Clicks} = 1)$ | $\ln(\text{Clicks} + \Delta_{it})$ | $\log(\text{CPM})$ |
| | OLS | Probit | iOLS | OLS |
| | (1) | (2) | (3) | (4) |
| Contextual | 0.252* (0.101) | 0.532*** (0.175) | 0.527* (0.247) | -0.0576 (0.109) |
| Open | 0.953*** (0.164) | 0.539*** (0.134) | 0.489 (0.377) | -1.712*** (0.191) |
| Both | -2.315*** (0.201) | 0.663 (0.538) | -0.726** (0.237) | -2.377** (0.869) |
| Open \times Both | -0.469 (0.344) | 0.605 (0.480) | 1.421** (0.511) | 2.304** (0.801) |
| Desktop | -1.383* (0.626) | -8.173*** (0.438) | -4.904*** (1.166) | 0.377* (0.152) |
| Smartphone | -0.879 (0.601) | -7.749*** (0.399) | -4.326*** (1.139) | 0.510*** (0.121) |
| Tablets | -0.421 (0.605) | -7.528*** (0.378) | -3.609** (1.034) | 0.526*** (0.120) |
| Standard Format | -0.325 (0.366) | -0.0397 (0.368) | -0.419 (0.515) | -0.332 (0.400) |
| Impressions | | 0.715*** (0.0523) | 1.462*** (0.142) | |
| Site FEs | ✓ | ✓ | ✓ | ✓ |
| Campaign-Week FEs | ✓ | ✓ | ✓ | ✓ |
| δ | | | 1 | |
| Observations | 2755 | 4337 | 4498 | 4498 |
| Adjusted R ² | 0.428 | | | 0.757 |

Note : *p<0.05 ; **p<0.01 ; ***p<0.001

Note : Robust standard errors clustered at the campaign-site level displayed in parentheses. Column (1) displays effects on ads click-through rate (CTR) on the subset of clicked ads. Column (2) displays effects on the propensity to click on ads, at least once. Column (3) shows results from the iOLS estimation where δ is the value of the optimal hyper-parameter and ϵ represents the absolute difference in coefficients between the last two iteration of the iOLS model. Finally, column (4) shows effects on ads' cost-per-1,000 impressions (CPM).

6.4 Heterogeneous Effect across Devices

TABLE 3.9 – Results per Device Type

| | Desktop | | | Mobile (Smartphones & Tablets) | | |
|-------------------------|----------------------|--------------------|----------------------|--------------------------------|--------------------|----------------------|
| | logit(CTR) (1) | log(Clicks) (2) | log(CPM) (3) | logit(CTR) (4) | log(Clicks) (5) | log(CPM) (6) |
| Contextual | -0.0377 (0.280) | 0.604 (0.501) | -0.211 (0.161) | 0.414* (0.204) | 1.070** (0.328) | 0.00377 (0.0930) |
| Open | 0.307 (0.469) | 0.0897 (1.169) | -1.895*** (0.292) | 1.460*** (0.417) | 1.811* (0.679) | -1.632*** (0.158) |
| Both | -1.862*** (0.354) | -0.912 (0.540) | -0.785 (0.562) | -2.594*** (0.267) | -0.550 (0.393) | -2.370* (0.921) |
| Open × Both | 1.177 (0.677) | 2.001 (1.307) | 2.610** (0.884) | -1.505* (0.643) | -0.306 (1.003) | 2.133** (0.793) |
| Ad Device and Format | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Site | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Campaign-Week FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Impressions | | ✓ | | | ✓ | |
| δ | | 1 | | | 1 | |
| Observations | 950 | 1568 | 1571 | 1805 | 2923 | 2927 |
| Adjusted R ² | 0.329 | | 0.754 | 0.459 | | 0.770 |

Note :

*p<0.05; **p<0.01; ***p<0.001

Results for Desktop Devices are Displayed in Columns (1) to (3). Smartphones and Tablets Results are reported in Columns (4) to (6).

Chapitre 4

Should I Stay or Should I Go ? Social Media Advertising in Times of Boycott

Abstract

In July 2020, more than 1,000 well-known consumer brands pulled their advertisements from Facebook and Instagram as the platforms hosted an increasing number of controversial content. Such contents, along with low media credibility, are known for altering ad effects. Using data from a US brand that did not boycott Facebook Ads back then, we investigate whether Facebook and Instagram ad outcomes effectively decreased during the boycott period. Leveraging recent developments in counterfactual estimates, we find that advertising clicks and prices dropped on Facebook and Instagram during the boycott. In particular, clicks decreased in July, when the boycott received the most media coverage. Using data scraped from Google News, we find a negative correlation between the boycott's press coverage and Facebook/Instagram ads clicks and prices. The degradation of Facebook's credibility is thus likely to have affected the value of ads on its platforms. Brands hence have reasons to leave ad platforms identified as controversial.

Résumé

En juillet 2020, plus de 1000 marques notoires ont suspendu leurs campagnes sur Facebook et Instagram. La faute au nombre croissant de contenus controversés hébergés par ces deux plateformes. Les contenus polémiques et la faible crédibilité des médias sont connus pour altérer l'efficacité des publicités. À partir des données d'une marque américaine n'ayant pas boycotté les publicités Facebook en juillet 2020, nous cherchons à savoir si la valeur des publicités Facebook et Instagram ont effectivement diminué pendant le boycott. Nous constatons que les clics et le prix des publicités Facebook/Instagram ont diminué entre juin et juillet 2020. En particulier, les clics chutent en juillet, lorsque le boycott reçoit la plus grande couverture médiatique. En utilisant des données extraites de Google News, nous trouvons une corrélation négative entre la couverture médiatique du boycott et les clics et prix des publicités sur Facebook/Instagram. La dégradation de la crédibilité de Facebook est donc susceptible d'avoir affecté la valeur des annonces sur les plateformes du groupe. Les marques ont donc des raisons de quitter les plateformes publicitaires identifiées comme controversées.

1 Introduction

On June 17th, a coalition of civil rights associations named #StopHateForProfit (#SHFP) urged advertisers to suspend their campaigns on Facebook and Instagram. In their words : “*advertisements are running alongside divisive, hateful and conspiratorial content – not something that most companies want*” (GREENBLATT, 2020). Heated debates around George Floyd’s death and the upcoming US presidential elections created a deleterious context for brands on Facebook. Donald Trump and his supporters engaged in provocative posts regarding sensitive subjects like police abuse¹, which was seen by many users as hate speech. As a result, ads risked to be associated with the polemical content displayed in Facebook and Instagram users’ news feed. One thousand well-known consumer brands joined #SHFP (among which The North Face, Ben & Jerry’s, Verizon, Unilever or Coca-Cola) and suspended their advertising on Facebook and Instagram for at least the month of July 2020. On September 17th, exactly two months after #SHFP’s open letter, Facebook announced a major policy update concerning moderation of groups’ content (ALISON, 2020).

This echoes 2017’s Adpocalypse, which saw major brands pull off their video ads from YouTube because they were often displayed on extreme videos (terrorism, antisemitic content, conspiracy theories). Marketers often refer to negative association between their brand and sensitive content as *brand safety* issues.

In the two sided market of media, content producers (hereafter publishers) attract readership that may pay for their content or not. Publishers are often ad-funded : brands place advertising alongside their editorial content. Ads and media content thus form a context which is subject to cross-externalities. On the one hand, advertising can affect the perceived quality of the editorial content produced by the media². On the other hand, content produced by editors externally affect advertisements. We refer to this last phenomenon as a *context externality*. Many research already studied the effect of media content and credibility on adjacent advertising outcomes (e.g, see STIPP, 2018 for a review). The presence of context externalities is prevalent in online advertising. Indeed, with real-time bidding – the automation of media buying through real-time auctions – brands rarely know on which site their ad appears. And when they do, like when they advertise on a social media platform, they do not know along which content their ads will be displayed to the users. Display and social media advertising are hence subject to such brand safety issues since firms do not know *ex-ante* the context in which their ads appear.

The 2020 Facebook Ad Boycott offers an interesting natural experiment to study context externalities in online advertising. At least three reasons may underlie a brand’s decision to boycott Facebook and Instagram Ads. First, polemical content may impose a negative contextual externality on the brand’s advertisements. Second, the controversy associated with Facebook, as a brand and a media, may decrease the general credibility of ads displayed on the platform, even in non polemical contexts. Third, the participation in the boycott may be part of a strategic

1. Trump infamously declared in a Facebook post that “*When the looting starts, the shooting starts*” in the context of protests following the death of George Floyd.

2. For instance, the development of the advertising in the French press has long been limited as editors felt that ads (especially suspicious ones) would hurt the credibility of their articles (MARTIN, 2016)

decision : as the boycott gains importance, remaining brands may be seen as non-virtuous by consumers sensitive to hate speech issues. This last observation is another negative externality for brands that do not overtly join #SHFP. Both controversial contents and low media credibility are known for altering ad effects.

Using data from a non-boycotting US brand in the skincare industry, we investigate whether ad performance and prices effectively decreased during the boycott period. The company maintained its campaigns on Facebook and even published non-paid posts on its page on several occasions during the boycott. Taking advantage of this natural experiment setting, we use a differences-in-differences analysis to see how Facebook campaign's outcome evolved compared to similar campaigns on other display networks between June and August 2020. This time interval however covers three different events : (i) the increasing circulation of controversial contents on Facebook's platforms in June, (ii) the advertiser boycott and its media coverage in July and (iii) the end of the boycott in August and Facebook's engagement in local content moderation. Thus, we extend our analysis with an interactive fixed-effects counterfactual approach (L. LIU et al., 2022). The latter provides a treatment effect for each week of the boycott and is robust to many shortcomings of differences-in-differences estimates.

We find the number of clicks generated by ads significantly fell over the June-August 2020 period, and so did the cost-per-1,000 impressions (CPM). Both variables followed the same trend : they decreased moderately during the first weeks of June and fell even more in July. For clicks, the effect can be explained by the negative contextual externalities which associate brands with controversial content and divert attention from ads. The effect on CPM may be imputed to the decrease in brands' demand for Facebook ads which mechanically decreased prices in ad auctions. Boycott effect estimates become non-significant at the beginning of August. However, we don't know if this return to normality is attributable to Facebook's new moderation policy or to the end of the boycott.

Related Works This paper stems from several streams of the economics and marketing literature.

There is a wide literature in economics and marketing on boycotts from a consumer perspective. For instance, BONDI et al., 2022 analyze the incentives of CEOs to communicate on political issues. LIAUKONYTĖ et al., 2022 study the recent case of Goya, a food brand which praised Donald Trump in 2020, generating a boycott movement on social media. However, the literature on boycotts initiated by advertisers is lacking, with the exception of POITRAS et SUTTER, 2009. Yet, ad boycotts are common in advertising as mentioned above. The rationale for such boycotts to happen is that media content externally affects advertising perception.

This last observation is the subject of a second stream of literature to which this study refers. Online content and website credibility have been found to impact advertising effectiveness in a significant manner. These context effects are highlighted from a theoretical perspective (see for instance DESAI et al., 2014) and in empirical works (STIPP, 2018). For example, SHEHU et al., 2021 show that ads displayed on low-quality websites were less effective, in particular for premium and luxury brands. AGUIRRE et al., 2015 find that website credibility lowered

the aversion of users to targeted advertising. In their meta-analysis, LULL et BUSHMAN, 2015 conclude that, in offline media, ads never benefit from being advertised in violent and sexual programs, except when there is a congruence with the product advertised. Thus, brands are careful regarding the context attached to their ads.

Our work also relates to a growing literature on content moderation. Theoretical models show that ad-funded platforms have incentives to engage in moderation efforts whenever controversial content represent a nuisance for advertisers (MADIO et QUINN, 2021 ; Y. LIU et al., 2022). The level of content moderation is the result a trade-off between moderation costs, the effect on ad revenues and the impact on user's activity on the platform (JIMÉNEZ DURÁN, 2022). Recent empirical studies mainly tackle the effect of moderation policies on user behaviors (MÜLLER et SCHWARZ, 2022 ; JIMÉNEZ DURÁN et al., 2022). However, while these studies highlight the importance of advertiser' reaction to controversial content, empirical works on the matter are scarce.

Thus, the 2020 Facebook Ad Boycott appears as an interesting experiment to investigate the effect of unsafe content on advertising outcomes. Empirical studies of the #SHFP movement mainly adopt a market value perspective. They provide mixed evidence that joining #SHFP increased firm's stock value (AFEGO et ALAGIDEDDE, 2021 ; VILLAGRA, 2021). Thus, more empirical results are needed on this question. In particular, the rationality of brands boycotting Facebook may be further investigated. As discussed earlier, advertisers have reason to avoid Facebook while the platform was under the spotlight of controversies.

Alternatively, the boycott period may be associated with similar, perhaps better, advertising returns. Indeed, polemical content foster user's involvement on the platform, which can increase the attention dedicated to ads. Moreover, the more brands exit the platforms, the more competing firms can attract their demand. In this case, boycotting Facebook represents an additional opportunity cost for the brand.

We test which hypothesis holds true using data retrieved from a brand who did not follow the Facebook Ad Boycott.

TABLE 4.1 – Timeline of The 2020 Facebook Ad Boycott

| | |
|------------------------------|---|
| June 17 th | #SHFP official announcement. |
| June 26 th | Unilever and Coca-Cola join the boycott. |
| Aug. 6 th | Facebook removes pro-Trump and QAnon troll farms. |
| Aug. 11 th | Facebook announces a policy against antisemitic and conspiracy content. |
| Sept. 17 th | Facebook announces big update on group moderation. |

2 Background & Data

2.1 The Boycott : Timeline and Expected Effects

As explained in introduction, the boycott officially started during the middle of June. On June 17th, #SHFP called brands to stop their ad spending on Facebook during the month of July. #SHFP organizers had a meeting with Facebook executives on July 7th but they later complained about Facebook not being willing to implement the changes they advocated for (SHFP, 2020). Mid-August, however, the platform started to remove farm trolls, blackface posts and some conspiracy groups (COLLINS et COLLIER, 2020; GHAFARY, 2020; ORTUTAY, 2020). However these moderating interventions can be qualified as *local*, as they are not part of a policy update by Facebook and thus appear more arbitrary than systematic. Finally, on September 17th, exactly two months after the #SHFP’s open letter, Facebook announced a major policy update concerning group regulation (ALISON, 2020). The summary of the events is reported in Timeline Table 4.1.

Among the thousand brands that participated in the boycott, most of them are consumer brands, media, public interest associations and institutions³. The boycott was temporary, as four out of five brands declared they intended to reactivate their Facebook campaigns the next month (August 2020) (NYT, 2020a). The tension on social media platforms however started earlier, with political reactions to George Floyd’s death at the end of May. Thus, it is difficult to assess precisely when the unsafe content started to appear on the platform and impact advertisers’ campaigns. However, because the infamous post of Donald Trump dates back to May, 29th, it seems safe to assume that the effect of controversial contents began on June 1st 2020. This

3. The comprehensive list of businesses that participated in the boycott is available here : <https://www.stophateforprofit.org/participating-businesses>. The brand that we study in the empirical analysis did not appear in the list.

hypothesis relies on two expected effects.

On the one hand, the circulation of unsafe content on Facebook platforms, which preceded the boycott, may harm ad effectiveness. The tension created by controversial and extreme political narratives may have diverted user’s attention from ads. Even worse, the brands could have appeared as the financial backers of this content⁴. This hypothetical effect should have started affecting Facebook ads in the first week of June.

On the other hand, ads can suffer from the boycott itself. With the boycott’s media coverage and its extension to other brands, Facebook appears as a controversial platform. All in all, the degradation of Facebook’s brand is likely to affect the effectiveness of ads on the platform. This effect is expected to start at the beginning of July, after many brands announced they joined #SHFP.

Regarding the effect of the moderation policies, the timing is not more evident. Facebook started to ban targeted conspiracy groups like QAnon or Boogaloo from mid-August (COLLINS et COLLIER, 2020 ; GHAFARY, 2020 ; ORTUTAY, 2020). Meanwhile, advertisers who suspended their ads in July reactivated their campaigns in August. Indeed, an agency declared that, among its clients who joined the boycott, four out of five planned to go back on Facebook and Instagram in August (NYT, 2020a). The only official date we found is Facebook’s major update on groups announced on September 17th 2020 (ALISON, 2020).

To sum up, we consider the three following event and their expected effects :

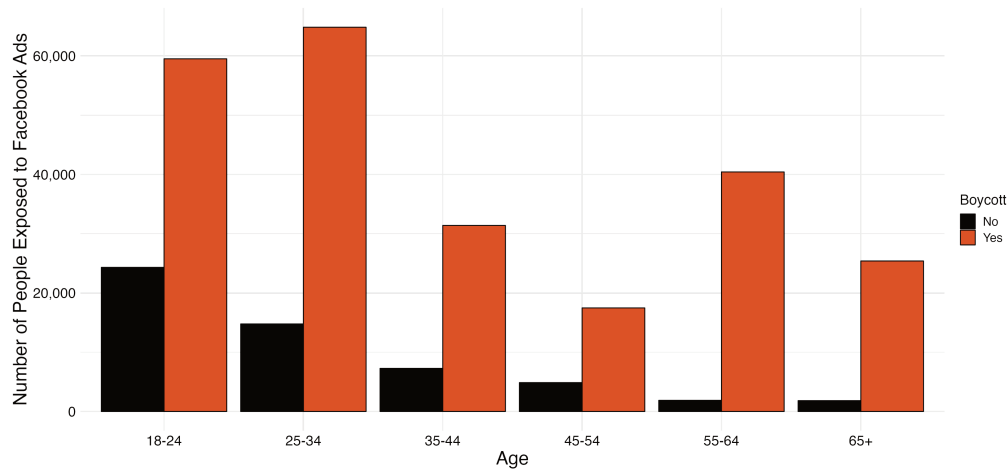
- **Negative context externality from contents** begins on June 1st. We expect the effectiveness of ads decrease because users give less attention to them or because they associate brands with controversial contents. The price of Facebook and Instagram ads is likely to remain constant since advertisers maintain their campaign.
- **Negative context externality from boycott** starts on July 1st. We expect ad effects to decrease even more as a result of controversial contents and Facebook’s deteriorated credibility. The price of ads is very likely to fall : ads will be less demanded by advertiser as a result of the boycott.
- **Positive effect from the moderation policy** begins on September 1st. We expect ad effectiveness to increase as controversial contents are regulated. The price of ads is also likely to increase : boycotting brands will be back on the platform and willing to buy Facebook Ads after the moderation policy.

2.2 Data and Descriptive Evidence

The author used data from a partner firm operating in the skincare industry. The brand mainly commercializes skin care products like scar gel, cold sore and blisters medications, or cream for stretch marks. The brand advertised on Facebook, Instagram and other display ad

4. An historian of advertising in France noticed that the political commitment of the magazine *Vu* in the early 1930s kept advertisers away : “*Vu did not succeed in its advertising breakthrough. It was the fault of the crisis, from 1931, but also the magazine’s political commitments, most advertisers hesitated, in this period of acute political and ideological conflicts, to appear as the financial backers of a politically committed publication.*” [Translated from French] MARTIN, 2016.

FIGURE 4.1 – Average Number of Individuals Exposed to the Brand Ads across Age Groups per Week



networks (like the Google Display Network or the former Verizon Media) from September 2019 to December 2020. For each media and week-year, the data reports the number of clicks, impressions, and cost associated to the ad.

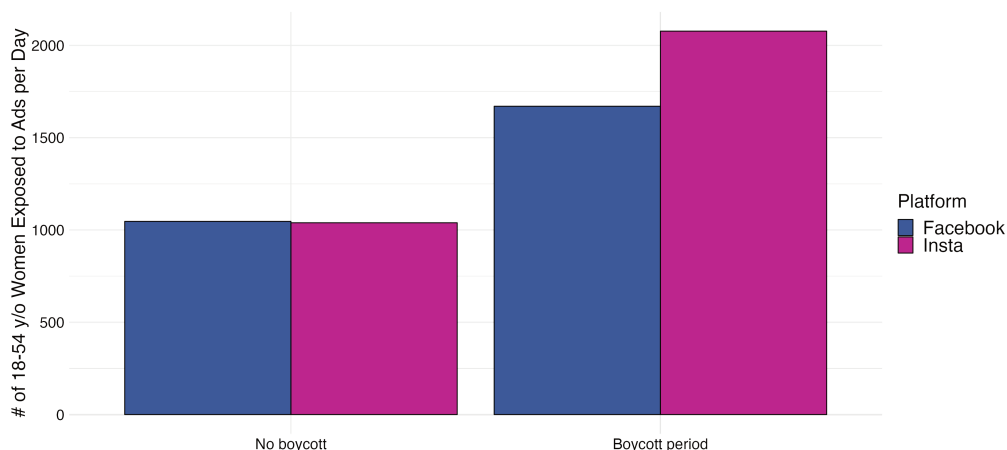
Ads are targeted to a wide audience : mainly 18 to 54 years old people with interests in skin care, beauty products, parenthood, surgery, dermatology or a sport lifestyle. The brand also targets Facebook users who liked or interacted with its page. The campaigns consist of video and static ads displayed on Facebook and Instagram. All campaigns are bought through ad auctions. Out of the USD 500,000 spent on Facebook ads, 42,8% had the objective to trigger user’s action (mainly clicks or other engagement with the ad).

A legitimate question is whether the advertiser boycott was followed by a user boycott. On an aggregated level, there is no evidence that Facebook’s number of monthly active users decreased in North America between June and September 2020 (GLOBALDATA, 2022 ; STATISTA, 2022b)⁵. However, because the focal brand targets a sub-category of Facebook users, some heterogeneity analyses regarding the audience segment are needed. As we can see in Figure 4.1, the average number of users exposed to Facebook Ads increased in all age categories during the boycott. However their distribution changed : older people represent a higher share of the audience during the boycott. The same is true for gender : both the number of men and women exploded and men became more exposed to ads during the boycott whereas, in other periods, women are more exposed (Figure 4.7). Focusing on 18-54 years-old women, the core audience of the brand, Figure 4.2 shows that they were more exposed during the boycott on both Facebook and Instagram. In Appendix, we show descriptive evidence that every audience exposed to the ads were less responsive (as measured by the click-through rate).

The advertiser did not join the boycott as it maintained its campaigns on Facebook and Instagram. It even posted 19 non-paid posts on its Facebook page between July 17th and September 17th. The author contacted the brands’ marketing experts and its advertising agency. None of the executives contacted were able to say whether not joining the boycott was a deliberate stra-

5. The lowest estimate we find reports a decrease of 0.39% in July 2020.

FIGURE 4.2 – Average 18-54 years old Women Exposed to the Brand Ads per Day



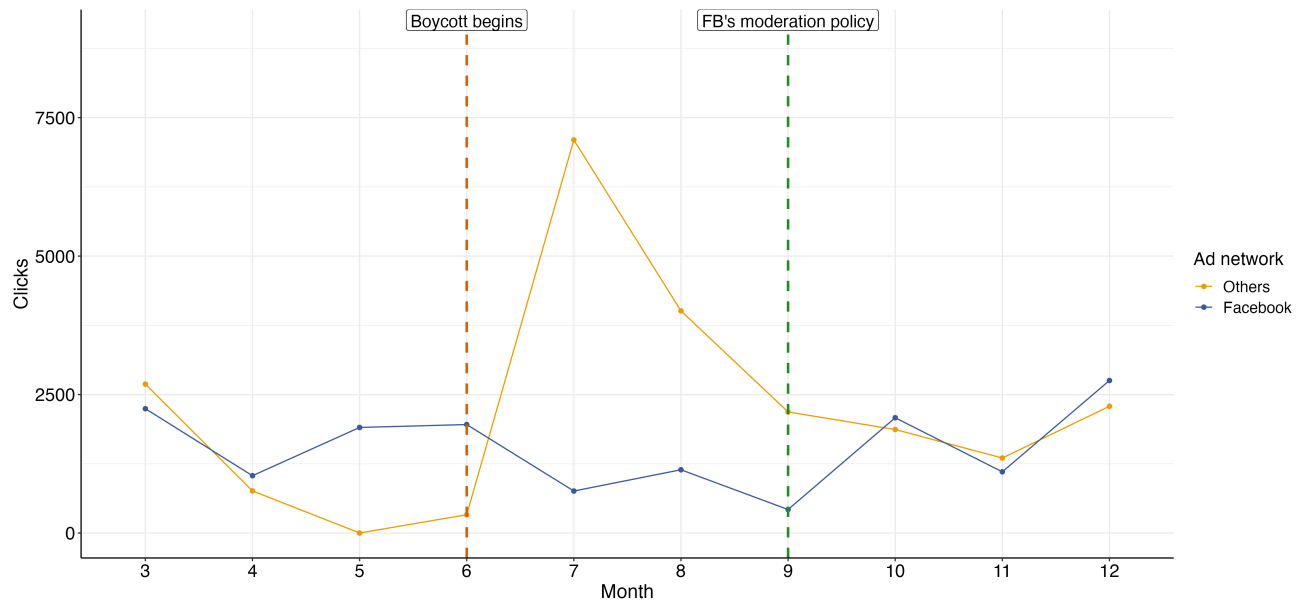
tegy from the advertiser. Some even seem not to be aware that there was a boycott at that time. Nevertheless, we can imagine some factors influencing brand’s decision to stay on Facebook at that time. First and foremost, the boycott occurred following the stay at home policies, which may have limited the quantity of products sold by the brand. With social activities coming back after May, the brand may have choose to maintain its campaign in order to compensate the Covid effect on sales. A second reason for not joining the boycott is that the brand’s perceived risk was low. As a manufacturer of skincare, blisters and scare treatments, the firm faces a natural demand on a market unaffected by political shocks. Moreover, and as opposed to brands like Ben & Jerry’s, the focal firm is not known for adopting a strong political slant.

The panel is unbalanced. Because the brand does not advertise on all platforms at all periods, we observe zero ad outcomes for some media and period of the panel. For many control ad networks (Cluep, Google, Verizon and Youtube), the pre-boycott information is limited. However, when the boycott starts, most control networks are observed. Looking at the effectiveness of ads on these different networks, Figure 4.3 shows that Facebook ads experienced a decrease in clicks while ads on other networks are characterized by an abnormal click increase which disappears after September, when Facebook starts implementing a new moderation policy.

However, one could argue that the variation in clicks is driven by a proportional variation in ad impressions. For example, the brand could have decreased investments on Facebook during the period, leading to less users seeing the ad and thus less clicks. However, as shown on Figure 4.4 (panel A) we still find a strong decrease in Facebook ad click rates during the boycott, and an increase of other ad network performance. This difference is even more striking when looking at advertising prices per-1,000 impressions. The prices fell by more than a half for Facebook during the boycott while remaining unchanged for other ad networks (Figure 4.4, panel B). This supports the idea that boycotting advertisers did not transfer their ad spending on other display ad networks – or at least, they did not target the same networks and audience than the focal brand because otherwise we would witness a rise in ad spots auction prices.

This trend is not specific to the month of the year. In Appendix (Figure 4.8), we plot the evolution of click rates in 2019 and 2020 for Facebook Ads. In 2019, the CTR increased in July

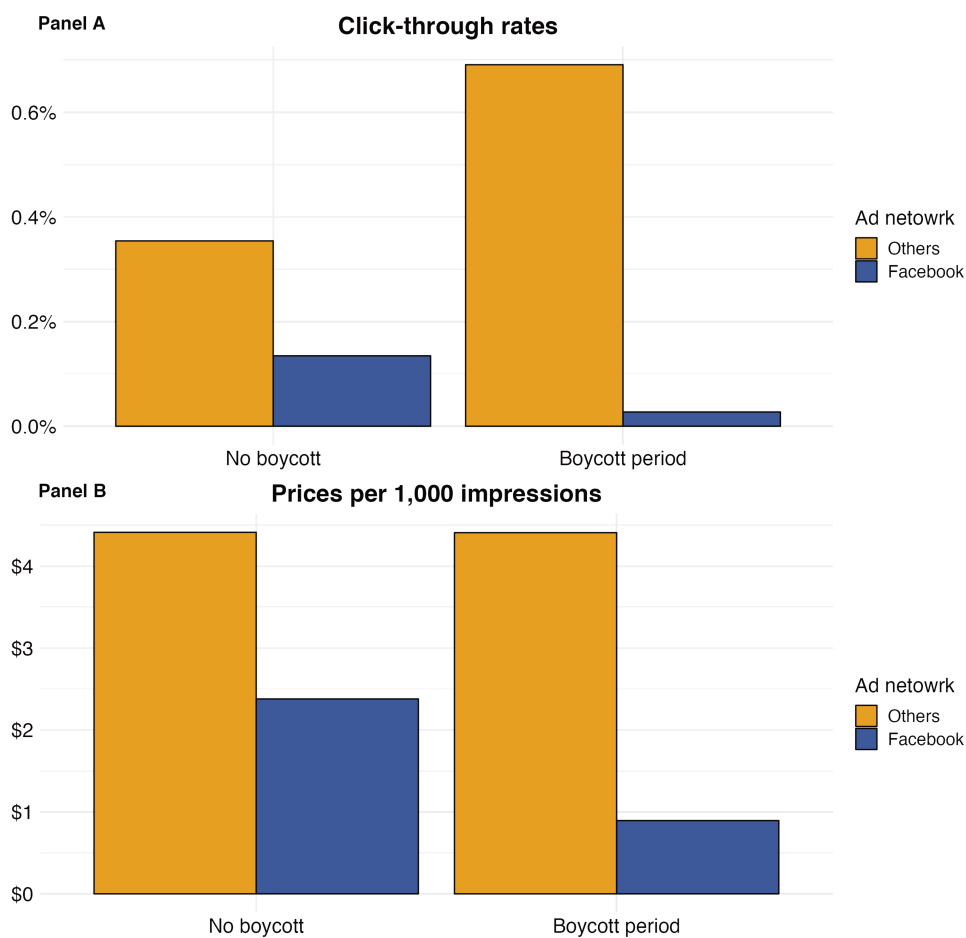
FIGURE 4.3 – Evolution of Clicks across Ad Networks in 2020



Note : to compare both groups (Facebook vs other ad networks), we present the average number of clicks recorded within each group. The blue line presents the average number of clicks recorded on Facebook and Instagram ads. The yellow line plots the average number of clicks recorded on the 4 other control ad platforms. In absolute value, the average clicks of the "others" group was 71,300 while the Facebook one averaged at 11,799 clicks

for both Facebook and Instagram, contrasting with the downward trend of July 2020.

FIGURE 4.4 – Facebook vs Other Networks Ad Outcomes : Boycott vs other Periods



Note : Panel A presents the average click rates of Facebook vs other ads during the boycott vs during other periods. Panel B presents the average cost-per-1,000 impressions for the same group and periods

3 Identification Strategy

Differences-in-Differences We take advantage of the natural experiment setting of the Facebook Ad Boycott by specifying the following Differences-in-Differences (DiD) regression :

$$Y_{it} = \beta_1 \text{Boycott}_t + \beta_2 \text{Facebook}_i + \delta(\text{Boycott}_t \times \text{Facebook}_i) + \gamma X_{it} + \alpha_i + \xi_t + \varepsilon_{it}. \quad (4.1)$$

In the equation above, Y_{it} is the number of clicks recorded on media i in week t . Boycott_t equals one between June and August 2020 included. If the boycott only started in July, its root cause (the proliferation of controversial contents) took place in June, and there has been no change in Facebook’s content policy before mid-August. Thus, this specification allows us to account for the cause and effect of the Facebook Ad Boycott. The variable Facebook_i equals 1 for $i = \{\text{Facebook}, \text{Instagram}\}$. X_{it} controls for the number of people exposed to the ad and α_i , ξ_t are media and time fixed effects respectively. The error term ε_{it} is clustered at the media level and robust to heteroskedasticity and auto-correlation. We are interested in the coefficient δ which measures the effect of advertising on Facebook during the June-August 2020 period.

Interactive Fixed-Effects Counterfactuals The advantage of specification (4.1) is its simplicity. However, it presents strong limitations. First, the treatment effect δ is constant over time. However, we do not know whether the effect of the boycott is caused by the spread of unsafe content that started in June, or to the media coverage of the boycott in July. Similarly, the timing and expected effect of the moderation policy are both unknown. Did Facebook’s local changes in August played a role in mitigating the content and boycott effects? And what about the September policy update? Dynamic treatment effect for each post-treatment periods would allow to disentangle the different effects taking place before, after and during the boycott.

Moreover, while fixed effects α_i and ξ_t account for constant changes in media and time, they do not handle unobserved time-varying confounders. Yet, unobserved media-specific components such as campaign objectives or ad design may vary over time. In addition to our DiD model, we also employ a counterfactual estimation method based on an interactive-fixed effects model (BAI, 2009; GOBILLON et MAGNAC, 2016; XU, 2017; L. LIU et al., 2022). The method, which we refer to as IFECT, uses data from the untreated observations to calibrate a counterfactual of the following form :

$$Y_{it}^N = \beta X_{it} + \alpha_i + \xi_t + \sum_{r=1}^{r^*} \lambda_{ri} f_{rt} + \varepsilon_{it}, \quad (4.2)$$

Here, Y_{it}^N corresponds to the counterfactual number of clicks Facebook and Instagram ads would have received in the absence of the boycott⁶. The factors λ_i and f_t correspond to unobserved unit and time effects respectively. The interaction of both terms allows to take into account unobserved time-varying confounders across units. The model uses $r^* \in (0, 5)$ optimal factors, with r^* determined by a cross-validation loop (XU, 2017). The estimation of the counterfactual is

6. We use the notation Y_{it}^N in reference to ABADIE, 2021’s notation in synthetic controls.

detailed in Appendix 6.4. Once Y_{it}^N is determined, the treatment effects can be computed simply by subtracting the number of clicks recorded during the boycott Y_{it} to the counterfactual :

$$\delta_{it} = Y_{it} - Y_{it}^N.$$

This approach allows us to (i) compute a treatment effect for each period and (ii) create a unique counterfactual based on the information of the untreated observations. As such, it resembles the Synthetic Control Method – hereafter SCM (see ABADIE, 2021 for a synthesis and discussion). However, IFect is more flexible than SCM as it holds for unbalanced panels and multiple treated units, while SCM does not ⁷. Moreover, while SCM relies on pre-treatment information, IFect leverages a wider set of untreated data (control groups in all periods and treated ones before treatment happens).

4 Results

4.1 Facebook Ad Performance and Prices Dropped during the Boycott

Results from the DiD model (4.1) are reported in Table 4.2. As we can see, in all specifications, the boycott decreased Facebook ad outcomes. Columns (1) to (2) display the effect of the boycott on clicks with different controls. Controlling for ad audience (column (1)) obliterates the effect of boycott on ad prices. Because the boycott may have reduced Facebook and Instagram ad prices, the drop in ad effectiveness may be compensated by a proportional drop in ad prices. Thus, brands get fewer ad clicks, but they have them at a cheaper price. Column (2) shows that when controlling for ad cost, we still find a negative effect of the boycott. In other words, for a same amount of money, the brand gets a lower volume of ad clicks on Facebook platforms during the boycott. Column (3) shows the effect on click rates while column (4) confirms that Facebook and Instagram ad prices decreased during the boycott.

Results from the IFect counterfactual model plotted in Figure 4.5 are consistent with the DiD estimates. As expected, the effect of the boycott is not constant over time. The decrease in clicks start in the first weeks of June, felling to a significant $-5,000$ in the third week. This negative effect is likely to be generated by the presence of controversial content on the platform. The decrease intensifies in July, felling to $-10,000$ clicks. This second effect may be more closely related to the boycott’s media coverage in July, which has damaged the credibility of Facebook and Instagram as a media.

An alternative explanation for this negative effect lies in the behavior of Facebook’s ad targeting algorithm. As shown in descriptive statistics (Figures 4.1, 4.2 and 4.6), the algorithm targeted the brand’s ads to unusual groups such as men or +55 years old person. An hypothesis is that in the absence of big advertisers on the platform, many impressions were to be allocated and the Facebook Ad algorithm started to target groups outside the brand’s preferred population.

On average, Facebook and Instagram campaigns would have recorded 5,000 to 10,000 additional clicks per week in the absence of the boycott. As we can see, the effect of the boycott

7. In fact, IFect can be considered as a generalization of SCM, see XU, 2017.

TABLE 4.2 – Differences-in-Differences results : Boycott Effects

| | Clicks (1) | Clicks (2) | logit(CTR) (3) | CPM (4) |
|---------------------------|----------------------|----------------------|--------------------|-------------------|
| Facebook \times Boycott | -6,839*** (1,600) | -4,513*** (1,562) | -1.24*** (0.31) | -1.77** (0.84) |
| Media FEs | ✓ | ✓ | ✓ | ✓ |
| Time FEs | ✓ | ✓ | ✓ | ✓ |
| Impressions | ✓ | | | |
| Cost | | ✓ | | |
| Observations | 236 | 236 | 236 | 236 |
| Adjusted R ² | 0.314 | 0.355 | 0.082 | 0.074 |

*p<0.1; **p<0.05; ***p<0.01

Note : Auto-correlation robust standard errors clustered at the media-level. Columns (1) and (2) estimate the model with different controls. Column (3) shows the effect on the logit-transformed click-through rate (CTR). Column (4) displays the boycott effect on the cost-per-1000 impressions (CPM)

is not significant anymore after August. Is it due to the local moderation policy of Facebook taking place in August? Or are the clicks increasing again as the boycott officially ends after July?

4.2 Content Moderation Did not Fully Offset the Boycott Nuisance

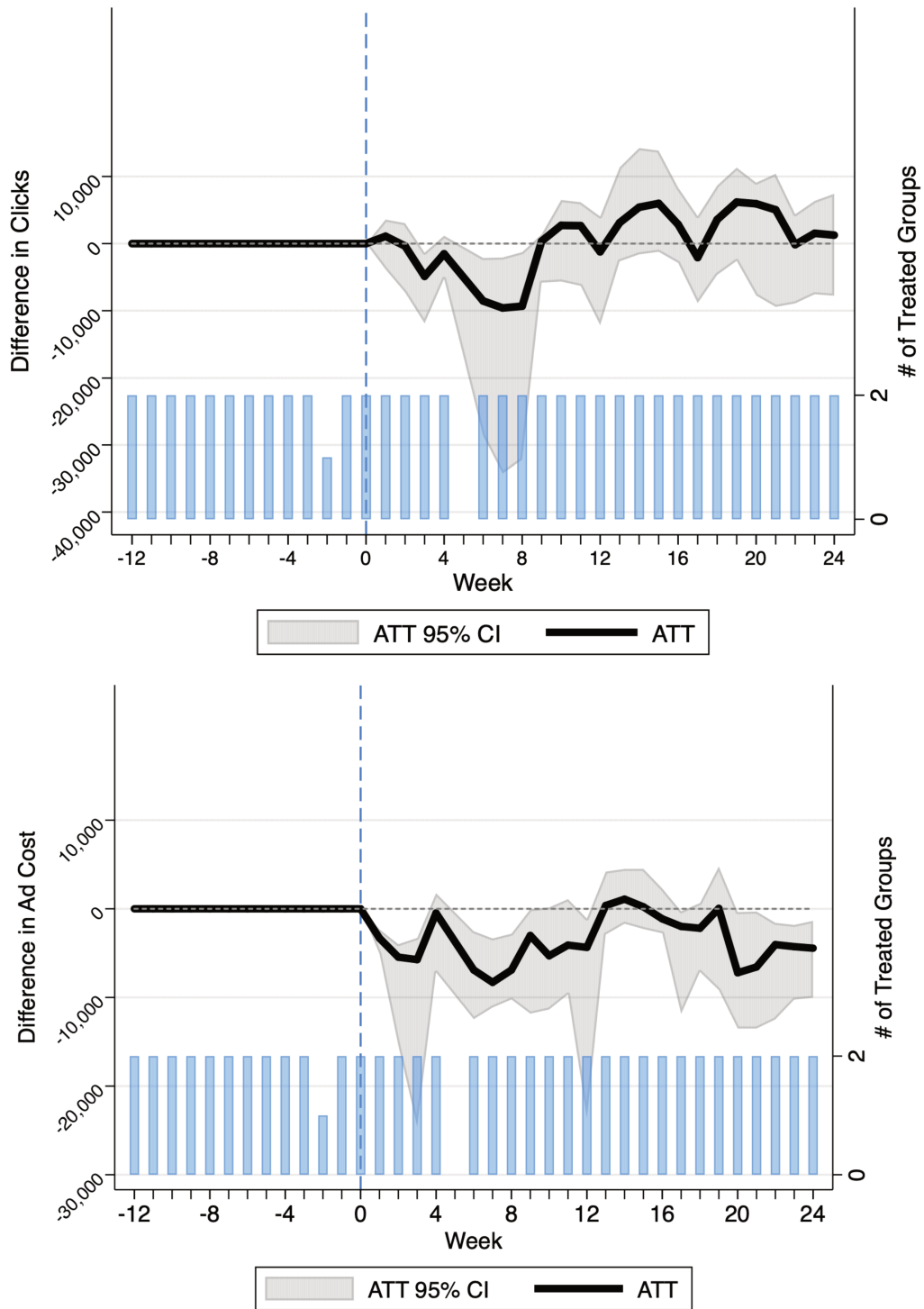
We first investigate the effect of moderation by running the DiD equation (4.1) with a different specification. Instead of the Boycott variable, we specify a Moderation dummy equal to one after September 2020 and zero otherwise. We also run a model which accounts both for the effects of boycott and moderation. We also leverage the dynamic treatment effects computed by the IFect model to investigate the long-term effect of the boycott.

The results of the DiD are reported in Table 4.3 and present mixed evidence about the effect of content moderation. On the one hand, moderation seems to exert a positive effect on clicks, regardless of the control used (columns (1) and (2)). However, the effect on the click-rate is not significant as displayed in column (3). Worse, column (4) shows that ad CPM continues to fall under the moderation policy. Finally, when taking into account both effects, we find that the content and boycott effect was stronger in magnitude than the moderation one (column (5)).

All in all, our DiD results suggest that the new content moderation policy did not fully offset the nuisance generated by the circulation of unsafe content and by the boycott itself.

Looking at the IFect results, Figure 4.5 shows that the negative effect of the boycott is becoming compensated after the 9th week. This corresponds to the moment Facebook banned conspiracy groups and content from its news feed. This is also the moment many brands stopped boycotting Facebook. The Facebook group’s update happens around the 14th week after the boycott but we do not see any clear significant effect on clicks on that period.

FIGURE 4.5 – Effect of the boycott on Facebook and Instagram clicks (IFEct estimates)



Note : Difference in clicks (up) and cost (down) conditional on ad impressions. Two-way additive and interactive fixed effects included. The treatment starts on June 1st (week 0). The left axis reports the ATT while the right axis reports the number of treated units (two in our case : Facebook and Instagram). The optimal number of factors chosen by the cross-validation procedure is $r^* = 3$.

TABLE 4.3 – Differences-in-Differences Results : Moderation Effects

| | Clicks (1) | Clicks (2) | logit(CTR) (3) | CPM (4) | Clicks (5) |
|------------------------------|---------------------|---------------------|-------------------|---------------------|---------------------|
| Facebook \times Moderation | 4,674*** (1,514) | 5,372*** (1,234) | -0.173 (0.301) | -1.94*** (0.691) | 2,789** (1,284) |
| Facebook \times Boycott | | | | | -4,698** (2,021) |
| Media FEs | ✓ | ✓ | ✓ | ✓ | ✓ |
| Time FEs | ✓ | ✓ | ✓ | ✓ | ✓ |
| Impressions | ✓ | | | | ✓ |
| Cost | | ✓ | | | |
| Observations | 236 | 236 | 236 | 236 | 236 |
| Adjusted R ² | 0.322 | 0.233 | 0.411 | 0.285 | 0.004 |

*p<0.1; **p<0.05; ***p<0.01

Note : Auto-correlation robust standard errors clustered at the media-level. Columns (1) and (2) estimate the model with different controls. Column (3) shows the effect on the logit-transformed click-through rate (CTR). Column (4) displays the boycott effect on the cost-per-1000 impressions (CPM). In column (5), the effect of the boycott and the moderation are jointly reported.

4.3 The Ad Boycott's Media Coverage is Negatively Correlated to Facebook's Ad Performance and Price

Our intuition is that the boycott's treatment effect was stronger on periods where the boycott coverage was high. To test this assumption, we scrapped the 100 first Google News pages related to the keywords "Facebook" "Ad" "Boycott" "Pull" and their publication date. 4 out of these 100 articles were irrelevant and deleted. From the 96 others, we create a variable equal to the number of press articles dealing with the Facebook Ad Boycott in week t . We then run a regression to see how these articles were correlated to Facebook Ads performance. We also used Google Trends for the "Facebook Ad Boycott" expression as another media coverage variable.

Results displayed in Table 4.4 confirm our intuition : clicks and prices decreased on Facebook and Insta ads as the media coverage increased. As an additional check, we run three DiD regressions replacing the Boycott dummy by a June, July and August dummy. Results shown in Table 4.8 are consistent with all our estimates : clicks decrease significantly in July with no clear effect in June and August.

TABLE 4.4 – Effect of Media Coverage on Facebook Ad Outcomes

| | # of Press Articles | | | | Google Trends | | |
|--------------------------|-------------------------|----------------------|-------------------------|----------------------|---------------------|---------------------|---------------------|
| | Clicks (1) | logit(CTR) (2) | Clicks (3) | CPM (4) | Clicks (5) | Clicks (6) | CPM (7) |
| # of articles | 234.066*** (84.366) | 0.005*** (0.002) | 191.945** (89.503) | -0.021*** (0.007) | | | |
| # of articles × Facebook | -288.208*** (88.873) | -0.063*** (0.019) | -261.480** (103.178) | -0.110** (0.049) | | | |
| GTrend | | | | | 26.436 (24.620) | 45.885 (48.386) | -0.008** (0.004) |
| GTrend × Facebook | | | | | -39.431 (26.222) | -68.291 (54.889) | -0.013 (0.012) |
| Media FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Week FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Impression | ✓ | | ✓ | | ✓ | ✓ | |
| Boycott sample only | | | ✓ | | | ✓ | |
| Observations | 236 | 236 | 56 | 236 | 236 | 56 | 236 |
| Adjusted R ² | 0.468 | 0.011 | 0.314 | 0.004 | 0.322 | 0.186 | -0.020 |

Note :

*p<0.1; **p<0.05; ***p<0.01

Note : the Table reports results from 7 regression models. Column (1) and (2) reports the effect of the number of articles published on the Facebook Ad Boycott on clicks and click rate respectively. Column (3) displays the effect on clicks over the boycott period only. Column (4) shows the effect on CPM. Column (5) to (7) shows similar results but using Google Trends for "Facebook Ad Boycott" as a measure of media coverage

4.4 Robustness Checks

Heterogeneous Treatment Effects Although our IFect estimation handles heterogeneous treatment effects, we check the presence of negative weights associated to our different treatment effects using CHAISEMARTIN et D’HAULTFŒUILLE, 2020’s procedure. Table 4.7 in appendix shows that only 3 negative weights have been diagnosed and that their aggregate absolute value is very small. Thus, we are not concerned with heterogeneous effects so far.

Testing for Pre-trends A drawback of the IFect model is that the estimated treatment effects could be misleading in presence of a pre-treatment trend. For example, if the number of clicks was decreasing steadily before the boycott, the boycott estimates could capture a part of this past trend. Although the absence of pre-trend is rather clear in Figure 4.5, we conduct an equivalence test for no trend before the boycott. Results show the model exhibits no trend prior to the boycott (section 6.5).

Placebo Test A second limitation of the IFect method is that the timing of the treatment could be mis-specified. In such a case, the estimates do not reflect the real treatment effect. This limitation is especially relevant in our case since, as mentioned before, we do not know precisely when the boycott exactly started to affect ads. To ensure the timing of our IFect model is

well identified, we perform a placebo test 4 weeks before the beginning of the boycott⁸. Results displayed in Appendix 6.5 show that our estimates are robust to placebo tests.

5 Concluding Words

The 2020 Facebook Ad Boycott effectively decreased the performance and prices of Facebook and Instagram ads. We find that the effect of the controversial content and boycott is U-shaped : clicks and costs decreased in June, the fall intensifies in July only to disappear in August. This trend may hide many effects : externalities from controversial content, negative image associated to Facebook, departures of well-known brands...

The results show that the controversial content indeed affected ad effectiveness and prices during the first weeks of June. However, the major decrease appeared during the boycott period (July 2020). Our additional analysis strongly suggests that the degradation of Facebook's brand image and credibility played a major role in decreasing Facebook Ads' value. This is consistent with previous studies on advertising context (especially the one of SHEHU et al., 2021)

Our study is however limited to a particular case : a brand operating in the skincare industry. Throughout this case study, we tried to show that the context matters for consumers engaging with the ad, and for advertiser's valuating ad slots. Nevertheless, the same study may have resulted in different findings in other industries or context. That may pose generalization problems. Moreover, our effectiveness variable is limited to ad clicks. Yet, clicks do not always lead to sales. Metrics like conversion rates, sales or purchase intent could be more appropriate, even though such metrics come with additional estimation challenges (LEWIS et RAO, 2015 ; G. JOHNSON, 2022).

8. We perform similar placebo test for the traditional DiD equation (4.1) by specifying a placebo boycott, starting 4 weeks earlier than the real boycott. Coefficients become small and non-significant when we declare placebo boycotts.

6 Appendix

6.1 Summary Statistics

TABLE 4.5 – Summary Statistics

| Variable | Obs | Facebook | | Obs | Non-Facebook | |
|-------------|-----|------------|------------|-----|--------------|------------|
| | | Mean | SD | | Mean | SD |
| Cost | 119 | 10255.23 | 6643.56 | 117 | 3616.01 | 2883.76 |
| Clicks | 119 | 9265.50 | 10971.56 | 117 | 2082.97 | 2555.99 |
| Impressions | 119 | 2324231.88 | 2115394.70 | 117 | 1615524.79 | 1409600.85 |
| CPM | 119 | 5.45 | 2.93 | 117 | 3.54 | 3.92 |
| CTR | 119 | 0.0045 | 0.0046 | 117 | 0.0025 | 0.0039 |

6.2 Effect of Boycott on Population Exposed

FIGURE 4.6 – Click-through Rate (CTR) by Age (Panel A) and Platforms (Panel B)

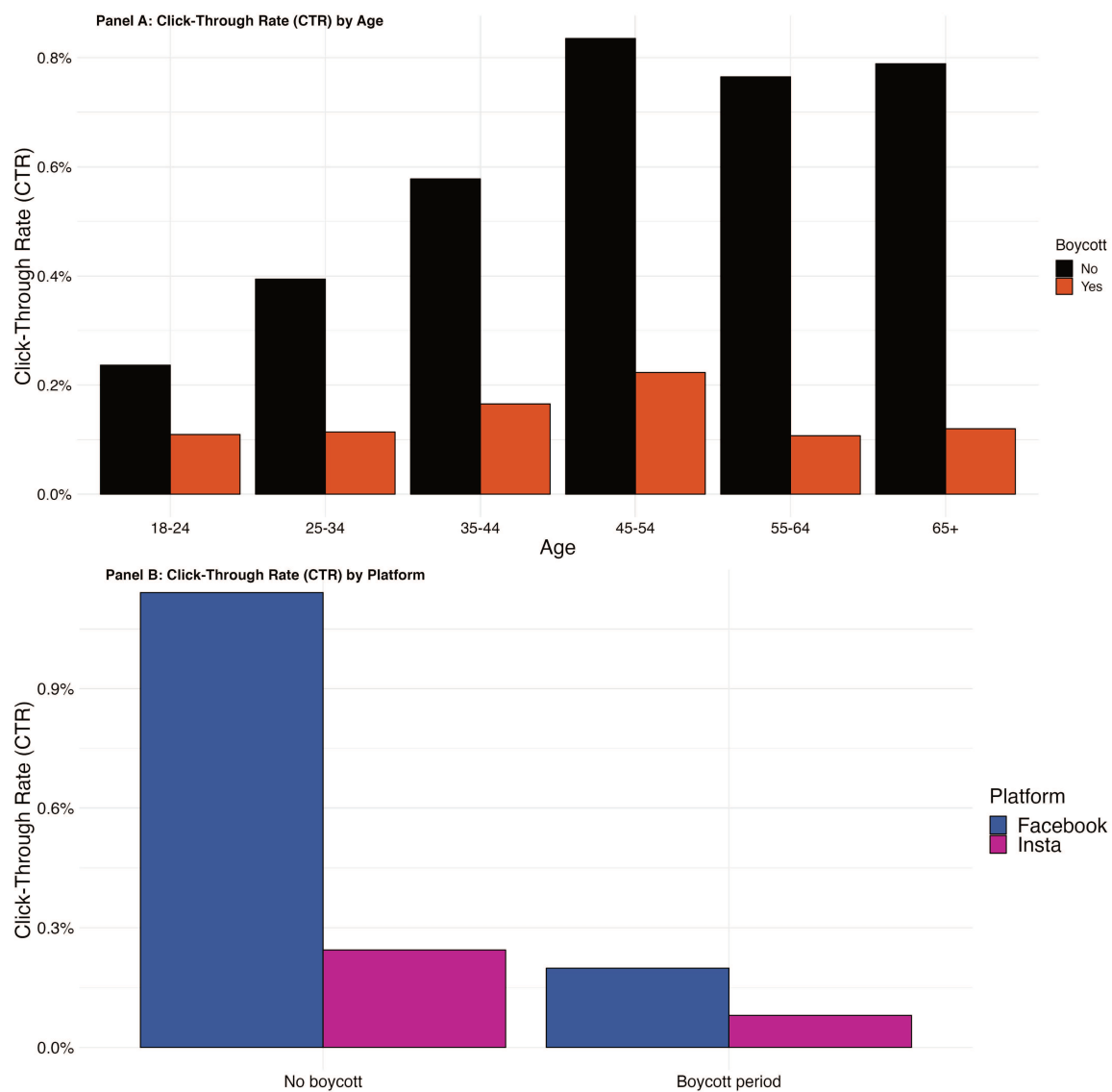
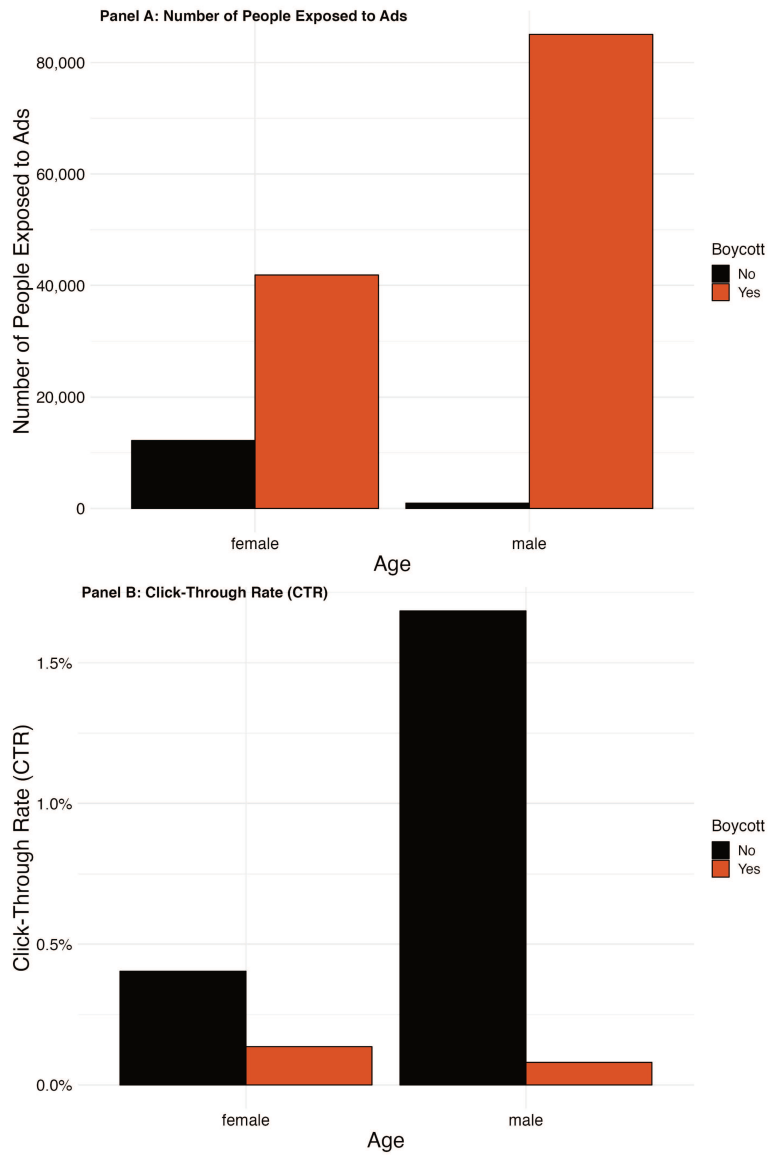
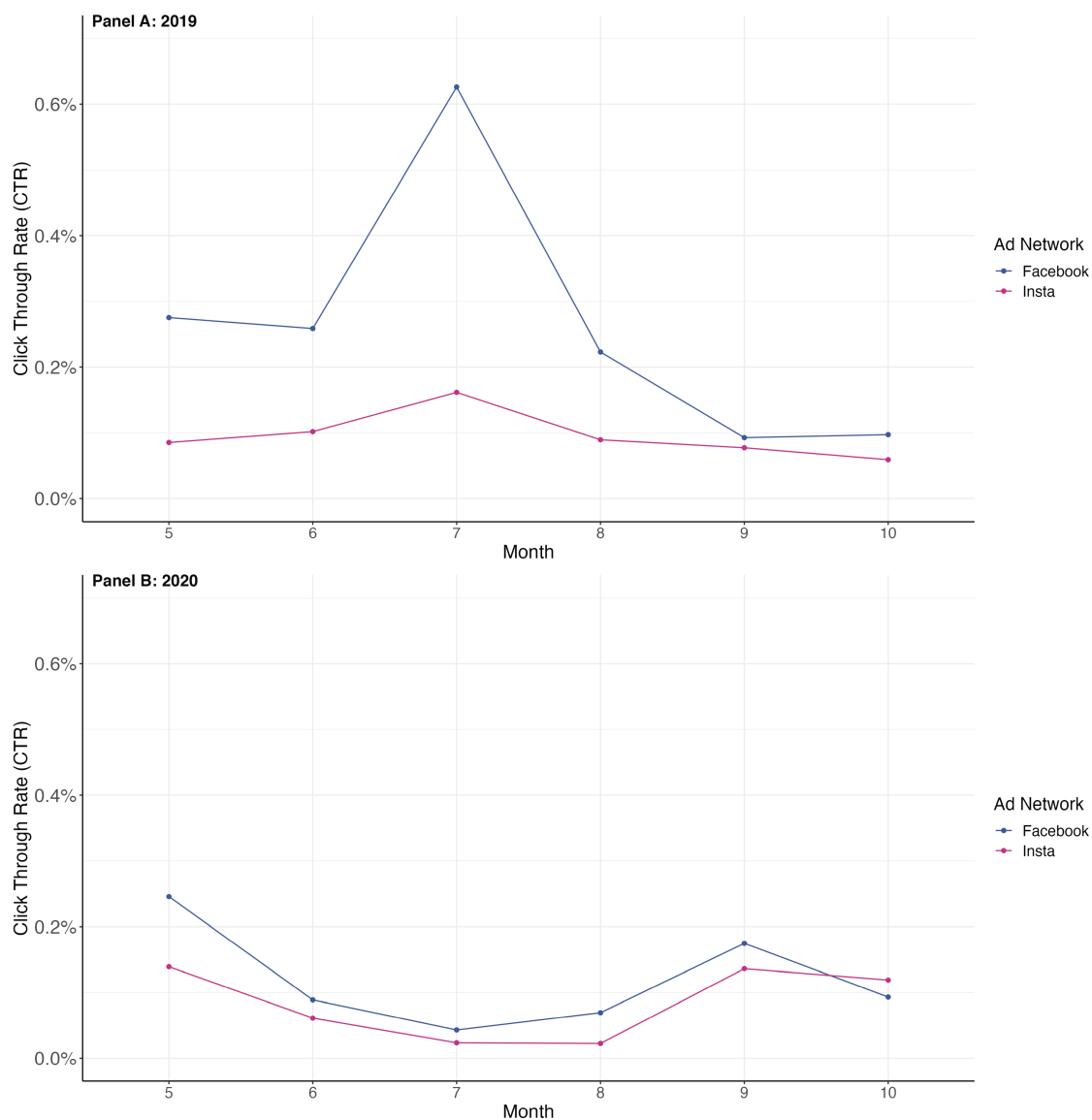


FIGURE 4.7 – Ad Exposure (Panel A) and Click-Through-Rate (Panel B) by Gender



6.3 Facebook Click Rates : 2019 vs 2020

FIGURE 4.8 – Evolution of Facebook ads (Facebook and Insta) click rate in 2019 and 2020



6.4 The IFect Estimator

Functional form Like the synthetic control method (SCM), the IFect estimator consists in computing a counterfactual outcome Y_{it}^N . In our case, Y_{it}^N is the number of clicks Facebook ads would have received in the absence of the boycott. The counterfactual takes the form of the following interactive fixed-effects model (BAI, 2009; GOBILLON et MAGNAC, 2016; XU, 2017; L. LIU et al., 2022) :

$$Y_{it}^N = \beta X_{it} + \alpha_i + \xi_t + \sum_{r=1}^{r^*} \lambda_{ri} f_{rt} + \varepsilon_{it}, \quad (4.3)$$

where X_{it} is a set of covariates, α_i , ξ_t , are unit and time unobserved effects. A set of r^*

unit-specific unobserved factors λ_{ri} are interacted with time-specific unobservables f_{rt} . This interaction allows to capture time-varying unobserved effects. The integer $r \in (0, 5)$ is the number of unit and time effects by group. The spirit of the method is that it leverages untreated outcome (control networks and pre-boycott Facebook data) to predict the counterfactual Y_{it}^N . We summarize the IFECT estimation procedure of L. LIU et al., 2022's Online Appendix in the following paragraphs :

Algorithm 1. IFECT Estimator

1. Estimate an interactive fixed-effect (IFE) model on the untreated data ($D_{it} = 0$) to obtain the estimated coefficients $\{\hat{\beta}, \hat{\alpha}_i, \hat{\xi}_t, \hat{\lambda}_i, \hat{f}_t\}$. The fitted value of this regression are denoted by \hat{Y}_{it}^0
2. Update the coefficient $\hat{\beta}$ by regressing \hat{Y}_{it}^0 from Step 1 on the control variables X_{it} . Obtain coefficients $\hat{\beta}^*$
3. Using coefficients from Steps 1 and 2, compute W_{it} , the prediction of Y of both treated and untreated units conditional on additive and interactive fixed effects :

$$W_{it} = \begin{cases} Y_{it} - X_{it}\hat{\beta}^* & \text{if } D_{it} = 0 \\ \hat{\alpha}_i + \hat{\xi}_t + \hat{\lambda}_i\hat{f}_t & \text{if } D_{it} = 1 \end{cases}$$

4. Regress W_{it} with an IFE model on all observations to obtain the updated factors $\hat{\lambda}_i^*, \hat{f}_t^*$.
5. Update fixed effects for all observations : $\hat{\alpha}_i^* = \bar{W}_i - \bar{W}$ and $\hat{\xi}_t^* = \bar{W}_t - \bar{W}$ where \bar{W}_i, \bar{W}_t and \bar{W} are unit-, time- and total-averages of W_{it} .
6. Compute the final counterfactual for $D_{it} = 1$:

$$Y_{it}^N = \hat{\beta}^* X_{it} + \hat{\alpha}_i^* + \hat{\xi}_t^* + \hat{\lambda}_i^* \hat{f}_t^*.$$

7. Compute the treatment effect for each treated units $i \in \mathcal{T}$ in each post-treatment period $t > T_0$:

$$\delta_{it} = Y_{it} - Y_{it}^N \quad \forall i \in \mathcal{T}, t > T_0$$

Cross-Validation Procedure We described the IFECT estimator with one interactive fixed effects. As shown in (4.3) however, IFECT can manage up to 5 interactive fixed effects : $\lambda_{1,i}f_{1,t}, \dots, \lambda_{5,i}f_{5,t}$. The selection of the optimal number of factors r^* is not obvious for researchers. Thus, the IFECT estimator employs the cross-validation procedure described in XU, 2017, p. 63-64.

For each $r \in (0, 5)$ the procedure holds back data from the treated units to compute the counterfactual in the pre-treatment period $Y_{i,t \leq T_0}$. The results are then used to predict the outcome on the hold-back observations. A prediction error is computed and the procedure returns the r which minimizes the overall prediction error.

6.5 Robustness Tests

6.5.1 An Equivalence Test for No Pre-trend

We test for the presence of a pre-trend prior to the boycott which could confound the treatment effect identified by the model. Typically, if the negative effect of controversial content started to affect Facebook ads before June, the boycott effects identified in models (4.1) and (4.2) would be biased. We use the equivalence test for no pre-trend developed by L. LIU et al., 2022 whose null hypothesis is :

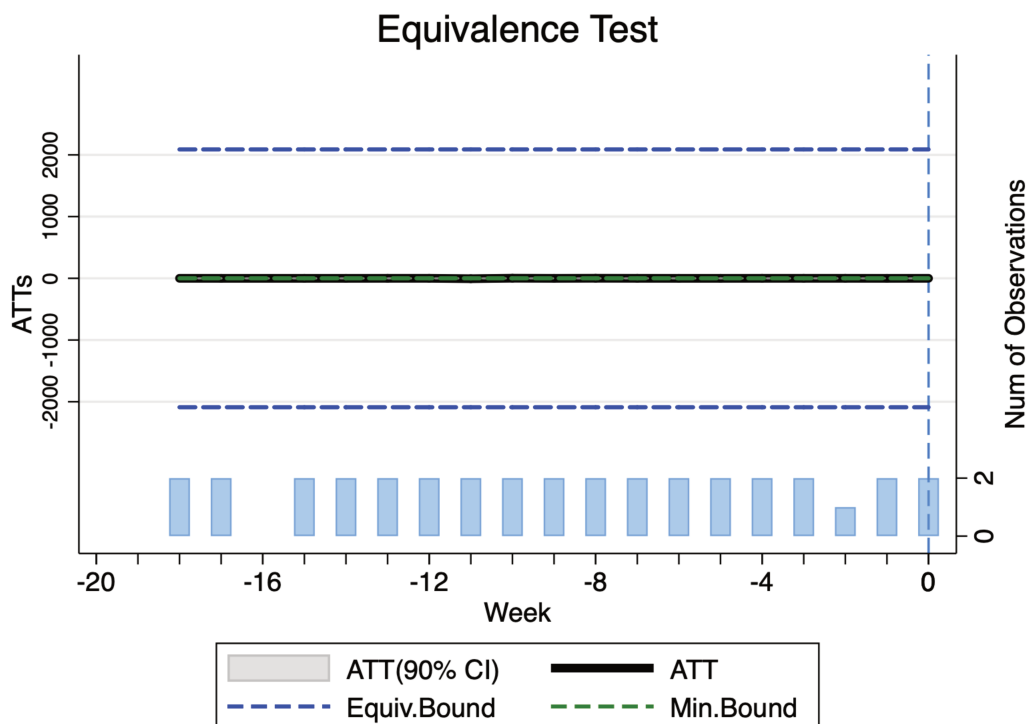
$$ATT_{t-s} \notin (-\theta, \theta),$$

where $(-\theta, \theta)$ is the 90% confidence interval testing for pre-trend. In other words, the null hypothesis states that the ATT in pre-treatment period $t - s$ lies out of the range $(-\theta, \theta)$ and thus exhibits a pre-trend. The bound of the confidence intervals are set such that $\theta = \sigma_{\varepsilon}$ which is the standard deviation of the untreated outcome's residual. The later is computed as the residual of a two-way fixed effect regression on the untreated units of the panel :

$$Y_{it} = \hat{\beta}X_{it} + \hat{\alpha}_i + \hat{\xi}_t + \hat{\varepsilon}_{it}.$$

The results of the test are presented in Figure 4.9. It shows that the treatment effect is very small in pre-treatment periods. Thus, it always lies inside the bound and we can easily reject the null hypothesis for the existence of a pre-trend.

FIGURE 4.9 – Equivalence Test Results



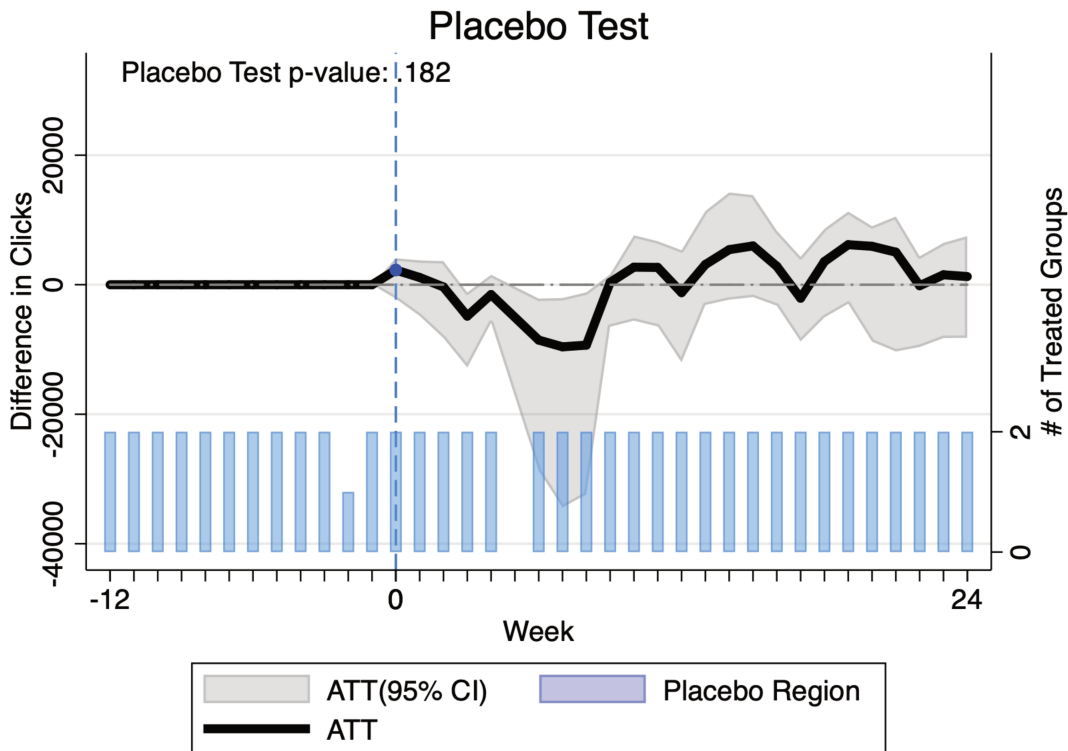
6.5.2 Placebo Checks

Similarly, we can use the equivalence test presented before to perform a placebo check. We run the model assuming the treatment occurs $s = 4$ periods before. As in the pre-trend test, the following null-hypothesis is investigated for each s :

$$ATT_{t-s} \notin (-\theta, \theta).$$

Again, if the ATT falls out of the interval, it means the placebo effect estimated is more significant than it should be. Thus, the model does not pass the test. The results from the test are reported on Figure 4.10. It appears clearly from the p-value that the placebo estimates are not statistically distinguishable from 0 ($p = 0.182$) and that they go in a direction which is the opposite of our treatment effects. Thus, our estimates pass the placebo test.

FIGURE 4.10 – Placebo estimation



The treatment has been voluntarily set to begin the 25th of May, one week before the date used in the main estimation (June 1st)

We also provide similar robustness checks for the regression model (4.1) by estimating fake boycott effects starting 1,2 and 3 months before the actual boycott specified in the baseline model. As we can see, results are getting smaller and non-significant.

TABLE 4.6 – Placebo tests. The table show the DiD results when the treatment starts at week $t - x$ instead of t

| | Clicks (1) | Clicks (2) | Clicks (3) | CPM (4) | CPM (5) | CPM (6) |
|-------------------------|--------------------------|---------------------|----------------------|---------------|---------------|--------------|
| DiD May-July | -9,656.4*** (3,337.1) | | | -0.9 (0.6) | | |
| DiD April-June | | -374.3 (1,143.4) | | | -0.9 (0.6) | |
| DiD March-May | | | 1,004.6 (2,083.8) | | | 0.7 (0.7) |
| Media FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Time FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Impressions | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 236 | 236 | 236 | 236 | 236 | 236 |
| Adjusted R ² | 0.368 | 0.262 | 0.269 | 0.009 | -0.013 | -0.021 |

Note :

*p<0.1; **p<0.05; ***p<0.01

6.5.3 Heterogeneous Treatment Effects

An issue with the DiD framework is that it may produce biased estimates in the presence of heterogeneous effects. CHAISEMARTIN et D’HAULTFÈUILLE, 2020 showed that the DiD estimate (δ in equation (4.1)) corresponds to a weighted average of multiple treatment effects. This is an issue when the weights are of different signs. In the presence of many negative weights, a DiD model may estimate a negative (positive) effect even though all treatment effects are positive (negative). CHAISEMARTIN et D’HAULTFÈUILLE, 2020 propose a method to estimate the weights. Results reported below show that only one negative weight has been reported and that its value is very small. Thus, we are confident our results are robust to heterogeneous treatment effects.

TABLE 4.7 – Weights associated to the DiD model (4.1)

| | Boycott (1) | Moderation (2) |
|--------------------------------|----------------|-------------------|
| Proportion of negative weights | 3/24 | 3/58 |
| Sum of negative weights | -0.000324 | -0.006385 |

6.6 Effect by Month

TABLE 4.8 – Effect of Month on Facebook Ad Outcomes

| | Clicks (1) | Clicks (2) | Clicks (3) |
|-------------------------|---------------------------|------------------------------|--------------------------|
| Facebook × June | -2,420.461 (1,669.815) | | |
| Facebook × July | | -9,718.567*** (3,502.109) | |
| Facebook × August | | | 1,052.680 (1,407.805) |
| Media FEs | ✓ | ✓ | ✓ |
| Week FEs | ✓ | ✓ | ✓ |
| Cost | ✓ | ✓ | ✓ |
| Observations | 236 | 236 | 236 |
| Adjusted R ² | 0.320 | 0.422 | 0.324 |

*p<0.1; **p<0.05; ***p<0.01

Note : the Table reports results from 3 regression models. Each model interacts a month with the Facebook dummy. Each model controls for ad budget and two-way fixed effects. Robust standard errors clustered by media

6.7 Effect of Press Coverage on Clicks

TABLE 4.9 – Effect of Media Coverage on Facebook Ad Outcomes

| | # of Press Articles | | | | Google Trends | | |
|--------------------------|-----------------------------|----------------------|------------------------|----------------------|----------------------|--------------------|---------------------|
| | Clicks (1) | logit(CTR) (2) | Clicks (3) | CPM (4) | Clicks (5) | Clicks (6) | CPM (7) |
| # of Articles | 238.496** (95.399) | 0.005*** (0.002) | 146.380 (105.928) | -0.021*** (0.007) | | | |
| # of Articles × Facebook | -422.054*** (118.250) | -0.063*** (0.019) | -205.407* (119.748) | -0.110** (0.049) | | | |
| GTrend | | | | | 8.896 (19.980) | -8.288 (42.851) | -0.008** (0.004) |
| GTrend × Facebook | | | | | -46.324* (25.716) | -8.071 (47.146) | -0.013 (0.012) |
| Media FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Week FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Impression | ✓ | | ✓ | | ✓ | ✓ | |
| Boycott sample only | | | ✓ | | | ✓ | |
| Observations | 236 | 236 | 56 | 236 | 236 | 56 | 236 |
| Adjusted R ² | 0.428 | 0.011 | 0.080 | 0.004 | 0.268 | -0.006 | -0.020 |
| <i>Note :</i> | *p<0.1; **p<0.05; ***p<0.01 | | | | | | |

Note : the Table reports results from 7 regression models. Column (1) and (2) reports the effect of the number of articles published on the Facebook Ad Boycott on clicks and click rate respectively. Column (3) displays the effect on clicks over the boycott period only. Column (4) shows the effect on CPM. Column (5) to (7) shows similar results but using Google Trends for "Facebook Ad Boycott" as a measure of media coverage

6.8 Full Regression Results

TABLE 4.10 – Differences-in-Differences results with Control Variables

| | Clicks (1) | Clicks (2) | logit(CTR) (3) | CPM (4) |
|-------------------------|----------------------|----------------------|--------------------|--------------------|
| Facebook × Boycott | −6,839*** (1,600) | −4,513*** (1,562) | −1.24*** (0.31) | −1.77** (0.84) |
| Boycott | 3,554*** (1,120) | 3,831*** (1,080) | −0.125 (0.217) | −1.069* (0.577) |
| Impressions | 0.002*** (0.0002) | | | |
| Cost | | 0.755*** (0.072) | | |
| Media FEs | ✓ | ✓ | ✓ | ✓ |
| Time FEs | ✓ | ✓ | ✓ | ✓ |
| Observations | 236 | 236 | 143 | 236 |
| Adjusted R ² | 0.317 | 0.355 | 0.298 | 0.116 |

*p<0.1; **p<0.05; ***p<0.01

Note : Auto-correlation robust standard errors clustered at the media-level. Columns (1) and (2) estimate the model with different controls. Column (3) shows the effect on the logit-transformed click-through rate (CTR). Column (4) displays the boycott effect on the cost-per-1000 impressions (CPM).

TABLE 4.11 – Differences-in-Differences results for Moderation with Control Variables

| | Clicks (1) | Clicks (2) | logit(CTR) (3) | CPM (4) | Clicks (5) |
|-------------------------|----------------------|---------------------|-------------------|---------------------|----------------------|
| Facebook × Moderation | 4,674*** (1,514) | 5,372*** (1,234) | −0.173 (0.301) | −1.94*** (0.691) | 2,789** (1,284) |
| Moderation | −4,673*** (1,045) | −5,597*** (941) | 0.097 (0.208) | 1.455*** (0.542) | −3,859*** (1,260) |
| Facebook × Boycott | | | | | −4,698** (2,021) |
| Boycott | | | | | 1,191 (1,340) |
| Impressions | 0.002*** (0.0002) | | | | 0.002*** (0.0002) |
| Cost | | 0.723*** (0.069) | | | |
| Media FEs | ✓ | ✓ | ✓ | ✓ | ✓ |
| Time FEs | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 236 | 236 | 236 | 236 | 236 |
| Adjusted R ² | 0.322 | 0.233 | 0.411 | 0.285 | 0.004 |

*p<0.1; **p<0.05; ***p<0.01

Note : Auto-correlation robust standard errors clustered at the media-level. Columns (1) and (2) estimate the model with different controls. Column (3) shows the effect on the logit-transformed click-through rate (CTR). Column (4) displays the boycott effect on the cost-per-1000 impressions (CPM). In column (5), the effect of the boycott and the moderation are jointly reported.

Conclusion et perspectives

1 Changement de conjoncture

Cette thèse débute en février 2020. La publicité en ligne représentait alors 51% des dépenses médias réalisées par les firmes, dépassant tout juste les montants alloués aux éditeurs traditionnels (BHATTACHARJEE, 2020). Le trio Google-Facebook-Amazon se taillait la part du lion en s’arrogeant 70% des revenus de la publicité numérique en 2020 (RICHAUD, 2020)). Bientôt, la pandémie dope le recours aux services en ligne, et avec eux les formats publicitaire sur lesquels ils reposent. Certes, Facebook essuie un boycott qui a érodé sa marque auprès des annonceurs. Mais celui-ci n’altère que marginalement les revenus de la plateforme, dont la manne porte sur les petites et moyennes entreprises (NYT, 2020b). La domination des “Big Tech” (Google, Facebook, Amazon) et de leurs régies publicitaires se fait alors insolente, inébranlable.

Au moment d’achever cette thèse en février 2023, le « *ralentissement de la pub en ligne* » est dans toutes les bouches. YouTube perd des revenus publicitaires pour la première fois de son histoire. Les bénéfices de Méta fondent de 41% en 2022. La faute à l’inflation qui force les marques à couper dans leurs budgets publicitaires. Mais ce ralentissement est surtout la conséquence des politiques de régulation des données personnelles. Intégré à iOS 14.5 dès avril 2021, L’App Transparency Tracking (ATT) permet aux utilisateurs d’iPhone, iPad et Apple TV de refuser la publicité ciblée. L’application vigoureuse du Règlement Général sur la Protection des Données (RGDP) en ce début d’année ébranle Méta dans la récolte des données personnelles. Les annonceurs peinant à cibler et mesurer les effets de leurs publicités, la valeur des espaces s’effrite. Et avec eux la rentabilité des services qui en dépendent.

Enfin, alors que l’antitrust s’est longtemps montré passif vis-à-vis des de la concentration des régies publicitaire en ligne⁹, le vent semble tourner. Le récent procès du *Department of Justice* contre l’activité publicitaire de Google met en lumière les différents problèmes de transparence, concentration et abus de position dominante dans le domaine de la publicité programmatique (DEPARTMENT OF JUSTICE, 2023). La menace du démantèlement pèse sur la méga-régie de Google, intégrant verticalement les services aux éditeurs (Google Ad Manager), aux annonceurs (Google Campaign Manager) et les places de marché (Google Ad Exchange).

Les différents chapitres de la thèse s’inscrivent dans cette nouvelle conjoncture et en éclaire

9. Cette quasi-complaisance s’illustre notamment par l’autorisation du rachat de DoubleClick par Google ou WhatsApp et Instagram par Facebook (O. BOMSEL et DEVAUX, 2022).

certain aspects.

2 Résumé des chapitres et implications

Chaque chapitre de la thèse a pour toile de fond l'idée que la publicité est un complément du produit vendu (BECKER et MURPHY, 1993). Dit autrement, la publicité est l'image du produit. Cette image est associée à la marque du produit, qui capte les effets des investissements publicitaires. Compris en ce sens, les effets de la publicité peuvent bénéficier à des tiers (les distributeurs du produit par exemple), donnant lieu, en l'absence de mécanismes d'internalisation, à des externalités. En tant qu'image du produit, la publicité risque également s'abîmer en s'affichant dans des contextes peu favorables. Ces différents éléments sont explicités dans les différents chapitres de la thèse.

Le **Chapitre 1** démontre que la publicité *search* profite des investissements en image, promotion et signalement des marques. Une firme augmentant ses dépenses publicitaires hors-ligne d'1% génère jusqu'à 0.95% de clics supplémentaires sur ses liens sponsorisés. En outre, l'activité des concurrents est aussi captée par le moteur de recherche puisque la concurrence en *search* augmente le prix des mots-clés disputés. Le **Chapitre 1** montre qu'une marque plus médiatisée engendre plus de requêtes et de clics, sur lesquels le moteur de recherche se rémunère. Google étant en position dominante sur la recherche, il capture la majeure partie de ces externalités.

Le **Chapitre 2** montre que publicités hors-ligne et numériques sont peu substituables du point de vue des annonceurs. À l'aide d'un modèle translog, le **Chapitre 2** montre que les élasticités-prix croisées des publicités traditionnelles et *search* étaient souvent proches de zéro, voire négatives. Le *display* et l'hors-ligne semblent substituables, mais essentiellement sur l'affichage et la presse. Quant au *search* et au *display*, les deux formats affichent des élasticités-prix croisées relativement élevées. En outre, les marques sont peu élastiques au prix des publicités hors-ligne et *search*. Si la publicité hors-ligne est, en France, un marché national et compétitif, le *search* est quant à lui dominé par Google. La dépendance des secteurs intensifs en référencement (comme l'hôtellerie) au *search* pose question quant au pouvoir de marché des régies en ligne.

La contextualisation de la publicité en ligne est étudiée dans le **Chapitre 3**. Les contrats (coût-par-impression versus coût-par-vue) ainsi que les modes d'achats (programmatique versus direct) impactent significativement la visibilité des publicités. En outre, le ciblage contextuel des annonces semble générer de meilleures performances à court-terme que le ciblage des utilisateurs. Ces résultats sont particulièrement d'intérêt dans la conjoncture actuelle. La personnalisation des annonces au profil de chaque utilisateur semble perdre en pertinence avec la récente régulation des données personnelles. L'adéquation de la publicité au contenu éditorial du support – modèle prédominant dans les médias traditionnels – apparaît donc comme un jeu à somme positive entre les annonceurs, éditeurs et utilisateurs (dont la vie privée n'est pas menacée). Cette conclusion s'impose également à la lecture du **Chapitre 4**. Celui-ci illustre les écueils d'un mauvais contexte. Le **Chapitre 4** enjoint les annonceurs à se concentrer autant sur la contextualisation des publicités que sur le ciblage de l'audience.

3 Pistes pour les recherches à venir

Les travaux inclus dans cette thèse présentent plusieurs limites ouvrant chacune des pistes pour les recherches à venir.

Par-delà les clics La première limite, commune à plusieurs chapitres de la thèse, est l'utilisation du clic comme mesure d'efficacité de la publicité en ligne. Certes, les clics peuvent conduire à des achats, et ils constituent la monnaie de la publicité en ligne¹⁰. Cependant l'objectif d'une campagne publicitaire est de générer du profit. Si les chiffres de ventes agrégés étaient accessibles pour les différentes firmes étudiées, il est difficile de les rapprocher des investissements publicitaires en ligne. En effet, les campagnes en ligne ont des effets marginaux réputés si faibles que ceux-ci sont difficiles à calculer (LEWIS et RAO, 2015 ; G. JOHNSON, 2022). D'autant plus que quantité d'éléments confondants viennent impacter les ventes, et de manière souvent plus forte que la publicité : e.g. saisonnalité, concurrence, mouvements de prix. Les clics présentent donc l'avantage d'être une mesure exogène au chercheur, directement fournis par les régies publicitaires. Il est ensuite possible d'observer les variations des clics au gré des variables d'intérêts (effets cross-média, ciblage et tarification des espaces etc...). Cependant l'usage d'autres mesures obtenues à l'échelle des campagnes comme les intentions d'achat ou conversions auraient été plus approprié. Hélas, ces données n'étaient pas disponibles dans le cadre de cette thèse. Le recours aux expérimentations à grande échelle est également à développer s'agissant des effets identifiés dans les **Chapitres 1** (exposition à des campagnes tierces), **3** (contextualisation des espaces) et **4** (effets des contenus haineux). Bien implémentées, ces expériences permettent de s'affranchir des problèmes d'endogénéité et d'identifier des effets causaux sur des populations à grande échelle.

Exploiter les réglementations Hormis le boycott de Facebook en juillet 2020, cette thèse exploite peu de chocs associés à la publicité en ligne. Les différentes réglementations liées à la régulation des données personnelles offrent pourtant un certain nombre de chocs utilisables par les chercheurs. L'ATT d'Apple est par exemple un cadre pertinent pour évaluer l'effet du "déciblage" de la publicité mobile. Et ainsi éclairer le régulateur quant à l'effet de telles politiques sur les annonceurs, éditeurs et consommateurs. Si une analyse a déjà étudié l'effet de l'ATT sur les éditeurs d'applications (KESLER, 2022), les conséquences pour les annonceurs et utilisateurs demandent encore à être découvertes. La pandémie de Covid-19 a également généré des chocs significatifs : e.g. arrêt de la publicité sur certains médias, recrudescence de l'exposition aux campagnes en ligne. Ceux-ci n'ont pourtant pas été étudiés dans le cadre de cette thèse.

Validité externe Enfin, la dernière limite de cette thèse concerne la généralisation des effets mis au jour dans les différents chapitres. L'amplitude des externalités entre médias mesurées dans le **Chapitre 1** est-elle spécifique aux hôtels, très dépendant du *search* ? Les effets positifs de la contextualisation identifiés dans le **Chapitre 3** sont-ils liés au caractère haut-de-gamme

10. Que ce soit dans les contrats de coût-par-clic ou dans l'optimisation du ciblage au click-through rate, le clic revêt encore d'une importance particulière dans le marché de la publicité en ligne.

du produit étudié? Le **Chapitre 2** montre à quel point les stratégies publicitaires des firmes varient par industrie. Ainsi, les effets mis au jour dans cette thèse gagnent à être étudiés sur d'autres secteurs, dans d'autres contextes. La création du complément publicitaire, son adéquation au produit et sa réception par le consommateur est un processus propre à chaque firme, les résultats empiriques sont difficilement généralisables. C'est toute l'originalité et la malédiction de l'économie de la publicité. Et celles de cette thèse.

Bibliographie

- ABADIE, Alberto (2021). « Using Synthetic Controls : Feasibility, Data Requirements, and Methodological Aspects ». *Journal of Economic Literature* 59.2, p. 391-425 (cf. p. 91, 92).
- ABADIE, Alberto, Susan ATHEY, Guido W IMBENS et Jeffrey M WOOLDRIDGE (2022). « When should you adjust standard errors for clustering? » *The Quarterly Journal of Economics* 138.1, p. 1-35 (cf. p. 12, 68).
- ADA, Sila, Nadia ABOU NABOUT et Elea McDonnell FEIT (2022). « Context Information Can Increase Revenue in Online Display Advertising Auctions : Evidence from a Policy Change. » *Journal of Marketing Research (JMR)* (cf. p. 72).
- AFEGO, Pyemo N et Imhotep P ALAGIDEDE (2021). « What does corporate social advocacy signal? Evidence from boycott participation decisions ». *Journal of Capital Markets Studies* (cf. p. 84).
- AGUIRRE, Elizabeth, Dominik MAHR, Dhruv GREWAL, Ko DE RUYTER et Martin WETZELS (2015). « Unraveling the personalization paradox : The effect of information collection and trust-building strategies on online advertisement effectiveness ». *Journal of retailing* 91.1, p. 34-49 (cf. p. 61, 83).
- ALISON, Tom (2020). *Our latest steps to keep Facebook groups safe* (cf. p. 82, 85, 86).
- ANDERSON, Richard G et Jerry G THURSBY (1986). « Confidence intervals for elasticity estimators in translog models ». *The Review of Economics and Statistics*, p. 647-656 (cf. p. 43).
- ANDERSON, Simon P, Øystein FOROS et Hans Jarle KIND (2018). « Competition for advertisers and for viewers in media markets ». *The Economic Journal* 128.608, p. 34-54 (cf. p. xv, 35).
- ARMSTRONG, Mark (2006). « Competition in two-sided markets ». *The RAND journal of economics* 37.3, p. 668-691 (cf. p. 3, 18, 34).
- ASDEMIR, Kursad, Nanda KUMAR et Varghese S JACOB (2012). « Pricing models for online advertising : CPM vs. CPC ». *Information Systems Research* 23.3-part-1, p. 804-822 (cf. p. 63).
- BACACHE-BEAUVALLET, Maya et Marc BOURREAU (2022). *Économie des plateformes*. La Découverte (cf. p. xv).
- BAGWELL, Kyle (2007). « The economic analysis of advertising ». *Handbook of industrial organization* 3, p. 1701-1844 (cf. p. xi, xii).
- BAI, Jushan (2009). « Panel data models with interactive fixed effects ». *Econometrica* 77.4, p. 1229-1279 (cf. p. xvii, 91, 101).
- BALOCCO, Francesco et Ting LI (2019). « LemonAds : Impression Quality in Programmatic Advertising ». *Fortieth International Conference on Information Systems* (cf. p. 62, 63).
- BALSEIRO, Santiago R, Ozan CANDOGAN et Huseyin GURKAN (2021). « Multistage intermediation in display advertising ». *Manufacturing & Service Operations Management* 23.3, p. 714-730 (cf. p. xiv).
- BECKER, Gary S et Kevin M MURPHY (1993). « A simple theory of advertising as a good or bad ». *The Quarterly Journal of Economics* 108.4, p. 941-964 (cf. p. xii, 2, 4, 34, 61, 110).
- BELLÉGO, Christophe, David BENATIA et Louis PAPE (2022). « Dealing with Logs and Zeros in Regression Models ». *arXiv :2203.11820 [econ, stat]*. Working Paper (cf. p. xvii, 69, 74, 75).
- BERGEMANN, Dirk et Alessandro BONATTI (2011). « Targeting in advertising markets : implications for offline versus online media ». *The RAND Journal of Economics* 42.3, p. 417-443 (cf. p. 4, 61).
- BERMAN, Ron (2018). « Beyond the Last Touch : Attribution in Online Advertising ». *Marketing Science* 37.5, p. 771-792 (cf. p. 5).
- BERNDT, Ernst R et Laurits R CHRISTENSEN (1973). « The translog function and the substitution of equipment, structures, and labor in US manufacturing 1929-68 ». *Journal of econometrics* 1.1, p. 81-113 (cf. p. 33, 36).

- BERNDT, Ernst R et N Eugene SAVIN (1975). « Estimation and hypothesis testing in singular equation systems with autoregressive disturbances ». *Econometrica : Journal of the Econometric Society*, p. 937-957 (cf. p. 42).
- BHATTACHARJEE, Monojoy (2020). « “Robust growth to resume in 2021” : Global digital ad spending to bounce back, reports eMarketer ». *whatsnewinpublishing.com* (cf. p. 109).
- BHATTACHARYA, Siddharth, Jing GONG et Sunil WATTAL (2022). « Competitive Poaching in Search Advertising : Two Randomized Field Experiments ». *Information Systems Research* 33.2, p. 599-619 (cf. p. 45).
- BLACKORBY, Charles et R. Robert RUSSELL (1989). « Will the Real Elasticity of Substitution Please Stand Up? (A Comparison of the Allen/Uzawa and Morishima Elasticities) ». *The American Economic Review* 79.4, p. 882-888 (cf. p. 37).
- BLAKE, Thomas, Chris NOSKO et Steven TADELIS (2015). « Consumer Heterogeneity and Paid Search Effectiveness : A Large-Scale Field Experiment : Paid Search Effectiveness ». *Econometrica* 83.1, p. 155-174 (cf. p. 19).
- BLEIER, Alexander (2021). « On the Viability of Contextual Advertising as a Privacy-Preserving Alternative to Behavioral Advertising on the Web ». *Available at SSRN 3980001* (cf. p. 72).
- BLEIER, Alexander, Avi GOLDFARB et Catherine TUCKER (2020). « Consumer privacy and the future of data-based innovation and marketing ». *International Journal of Research in Marketing* 37.3, p. 466-480 (cf. p. 72).
- BLOOMBERG (2021). « Apple and Google Are Killing the (Ad) Cookie. Here’s Why ». *Bloomberg* (cf. p. 72).
- BOMSEL, Olivier et Rémi DEVAUX (2022). *Le Nouveau Western : qui peut réfréner les géants du web ?* Cherche Midi (cf. p. xv, 109).
- BOMSEL, Olivier P. (2013). « Copyright and brands in the digital age : internalizing the externalities of meaning ». *Contemporary Economic Policy* 31.1, p. 126-134 (cf. p. 18).
- BONDI, Tommaso, Vanessa BURBANO et Fabrizio DELL’ACQUA (2022). « When to Talk Politics in Business : Theory and Experimental Evidence of Stakeholder Responses to CEO Political Activism ». *Working Paper* (cf. p. 83).
- BOUNIE, David, Morrisson VALÉRIE et Martin QUINN (2017). « Do you see what I see? Ad viewability and the economics of online advertising ». *Ad Viewability and the Economics of Online Advertising (March 1, 2017)* (cf. p. xv, 62, 63).
- BOURREAU, Marc et Anne PERROT (2020). « Plateformes numériques : réguler avant qu’il ne soit trop tard ». *Notes du conseil danalyse economique* 60.6, p. 1-12 (cf. p. xv).
- BRAITHWAITE, Dorothea (1928). « The economic effects of advertisement ». *The Economic Journal* 38.149, p. 16-37 (cf. p. xi).
- CAVES, Richard E (2000). *Creative industries : Contracts between art and commerce*. 20. Harvard University Press (cf. p. xii).
- CECERE, Grazia, Clara JEAN, Fabrice LE GUEL et Matthieu MANANT (2020). « STEM and teens : An algorithmic bias on a social media ». *Available at SSRN 3176168* (cf. p. xiv).
- CECERE, Grazia, Clara JEAN, Matthieu MANANT et Catherine TUCKER (2018). « Computer algorithms prefer headless women ». *2018 MIT CODE : Conference on Digital Experimentation* (cf. p. xiv).
- CHAISEMARTIN, Clément de et Xavier D’HAULTFÈUILLE (2020). « Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects ». *American Economic Review* 110.9, p. 2964-2996 (cf. p. 96, 105).
- CHANDRA, Ambarish et Ulrich KAISER (2014). « Targeted advertising in magazine markets and the advent of the internet ». *Management Science* 60.7, p. 1829-1843 (cf. p. 18, 35).
- CHINTAGUNTA, Pradeep K, Vrinda KADIYALI et Naufel J VILCASSIM (2006). « Endogeneity and simultaneity in competitive pricing and advertising : A logit demand analysis ». *The Journal of Business* 79.6, p. 2761-2787 (cf. p. 12).
- CHOI, Hana, Carl F MELA, Santiago R BALSEIRO et Adam LEARY (2020). « Online Display Advertising Markets : A Literature Review and Future Directions ». *Information Systems Research*, p. 21 (cf. p. 62, 72).
- CHOI, W Jason et Amin SAYEDI (2022). « Open and private exchanges in display advertising ». *Marketing Science (forthcoming)* (cf. p. 61, 63, 70).
- COLLINS, Ben et Kevin COLLIER (2020). *Facebook removes Troll Farm posing as African-American support for Donald Trump* (cf. p. 85, 86).
- COMANOR, William S et Thomas A WILSON (1974). *Advertising and market power*. 144. Harvard University Press (cf. p. xi).

- CROSS, Tim (2022). « Friends and Opponents Lay Out Arguments for TF1/M6 Merger ». *Video Week* (cf. p. 2).
- DAMS, Tim (2022). « France’s TF1 and M6 call off merger plans ». *Screen Daily* (cf. p. 33).
- DEPARTMENT OF JUSTICE (2023). *Justice Department Sues Google for Monopolizing Digital Advertising Technologies* (cf. p. viii, xiv, 60, 109).
- DESAI, Preyas S, Woochoel SHIN et Richard STAELIN (2014). « The Company That You Keep : When to Buy a Competitor’s Keyword ». *Marketing Science* 33.4, p. 24 (cf. p. 3, 18, 61, 83).
- DEVAUX, Rémi et Olivier BOMSEL (2022). « Externalities across advertising markets ». *Journal of Media Economics* 34.3, p. 152-175 (cf. p. 34, 35, 45, 46).
- DI EWERT, W. E. et T. J. WALES (1987). « Flexible Functional Forms and Global Curvature Conditions ». *Econometrica* 55.1, p. 43-68 (cf. p. 42).
- DINNER, Isaac M, Harald J HEERDE VAN et Scott A NESLIN (2014). « Driving online and offline sales : The cross-channel effects of traditional, online display, and paid search advertising ». *Journal of marketing research* 51.5, p. 527-545 (cf. p. 5, 7, 11-13).
- DLAPIPER (2015). *"Say it with flowers"... But beware of keywords!* (Cf. p. 18).
- DORFMAN, Robert et Peter O STEINER (1954). « Optimal advertising and optimal quality ». *The American Economic Review* 44.5, p. 826-836 (cf. p. 34, 42).
- DUBÉ, Jean-Pierre, Günter J. HITSCH et Puneet MANCHANDA (2005). « An Empirical Model of Advertising Dynamics ». *Quantitative Marketing and Economics* 3.2, p. 107-144 (cf. p. 12).
- EKELUND, Robert B, George S FORD et John D JACKSON (1999). « Is radio advertising a distinct local market ? An empirical analysis ». *Review of Industrial Organization* 14.3, p. 239-256 (cf. p. 35).
- (2000). « Are Local TV Markets Separate Markets ? » *International Journal of the Economics of Business* 7.1, p. 79-97 (cf. p. 35).
- ELIAZ, Kfir et Ran SPIEGLER (2020). « A model of competing narratives ». *American Economic Review* 110.12, p. 3786-3816 (cf. p. xii).
- E-MARKETER (2020). *Private Marketplace Ad Spending to Surpass Open Exchange in 2020* (cf. p. 60).
- (2021). *Worldwide Digital Ad Spending Year-End Update*. Rapp. tech. eMarketer (cf. p. 2).
- (2023). *Worldwide digital ad spend will top \$600 billion this year*. Rapp. tech. eMarketer (cf. p. xi).
- EUROPEAN COMMISSION (2008). *Case No COMP/M.4731 – Google/ DoubleClick* (cf. p. 33).
- EVANS, David S (2009). « The Online Advertising Industry : Economics, Evolution, and Privacy ». *Journal of Economic Perspectives* 23.3, p. 37-60 (cf. p. 4, 35, 61).
- FEDERAL TRADE COMMISSION (2007). *Statement of the Federal Trade Commission Concerning Google/DoubleClick* (cf. p. 33).
- FOSSEN, Beth L et David A SCHWEIDEL (2017). « Television advertising and online word-of-mouth : An empirical investigation of social TV activity ». *Marketing Science* 36.1, p. 105-123 (cf. p. 4).
- FRANK, Mark W. (2008). « Media substitution in advertising : A spirited case study ». *International Journal of Industrial Organization* 26.1, p. 308-326 (cf. p. 35, 39, 42, 51).
- FRICK, Thomas W, Rodrigo BELO et Rahul TELANG (2022). « Incentive Misalignments in Programmatic Advertising : Evidence from a Randomized Field Experiment ». *Management Science* (cf. p. xiv, 62, 63, 72).
- FRONDEL, Manuel (2011). « Modelling energy and non-energy substitution : A brief survey of elasticities ». *Energy Policy* 39.8, p. 4601-4604 (cf. p. 37).
- GENTZKOW, Matthew (2014). « Trading dollars for dollars : The price of attention online and offline ». *American Economic Review* 104.5, p. 481-88 (cf. p. 39).
- GENTZKOW, Matthew, Jesse M SHAPIRO, Frank YANG et Ali YURUKOGLU (2022). *Pricing power in advertising markets : Theory and evidence*. Rapp. tech. National Bureau of Economic Research (cf. p. xv, 35).
- GHAFFARY, Shirin (2020). *Facebook bans blackface and certain anti-Semitic conspiracy theories* (cf. p. 85, 86).
- GLOBALDATA (2022). *Number of Monthly Active Facebook Users : Global (Q4 2009 – Q1 2022)* (cf. p. 87).
- GOBILLON, Laurent et Thierry MAGNAC (2016). « Regional policy evaluation : Interactive fixed effects and synthetic controls ». *Review of Economics and Statistics* 98.3, p. 535-551 (cf. p. 91, 101).
- GOLDFARB, Avi (2014). « What is Different About Online Advertising ? » *Review of Industrial Organization* 44.2, p. 115-129 (cf. p. xiii, 2, 4).

- GOLDFARB, Avi et Catherine TUCKER (2011a). « Advertising bans and the substitutability of online and offline advertising ». *Journal of Marketing Research* 48.2, p. 207-227 (cf. p. 35).
- (2011b). « Online display advertising : Targeting and obtrusiveness ». *Marketing Science* 30.3, p. 389-404 (cf. p. 61, 72).
- (2011c). « Privacy regulation and online advertising ». *Management science* 57.1, p. 57-71 (cf. p. 72).
- (2011d). « Search engine advertising : Channel substitution when pricing ads to context ». *Management Science* 57.3, p. 458-470 (cf. p. 4, 35).
- (2011e). « Substitution between offline and online advertising markets ». *Journal of Competition Law and Economics* 7.1, p. 37-44 (cf. p. xvi, 4, 18, 33).
- GORDON, Brett R, Kinshuk JERATH, Zsolt KATONA, Sridhar NARAYANAN, Jiwoong SHIN et Kenneth C. WILBUR (2021). « Inefficiencies in digital advertising markets ». *Journal of Marketing* 85.1, p. 7-25 (cf. p. xv, 72).
- GORDON, Brett R, Robert MOAKLER et Florian ZETTELMEYER (2022). « Close enough ? a large-scale exploration of non-experimental approaches to advertising measurement ». *Marketing Science* (cf. p. xiv).
- GORDON, Brett R, Florian ZETTELMEYER, Neha BHARGAVA et Dan CHAPSKY (2019). « A comparison of approaches to advertising measurement : Evidence from big field experiments at Facebook ». *Marketing Science* 38.2, p. 193-225 (cf. p. xiv).
- GREENBLATT, Jonathan (2020). *An Open Letter to the Companies that Advertise on Facebook* (cf. p. 82).
- GREENE, William H (2003). *Econometric analysis*. Prentice Hall (cf. p. 39, 43).
- GUITART, Ivan A et Stefan STREMERSCHE (2021). « The impact of informational and emotional television ad content on online search and sales ». *Journal of Marketing Research* 58.2, p. 299-320 (cf. p. 5).
- HE, Xi, Rigoberto LOPEZ et Yizao LIU (2018). « Are Online and Offline Advertising Substitutes or Complements ? Evidence from U.S. Food Industries ». *Journal of Agricultural & Food Industrial Organization* 15.2 (cf. p. 12, 18, 35, 39, 51).
- HENNINGSEN, Arne et Jeff D. HAMANN (2007). « **systemfit** : A Package for Estimating Systems of Simultaneous Equations in R ». *Journal of Statistical Software* 23.4 (cf. p. 42).
- HU, Yu, Jiwoong SHIN et Zhulei TANG (2016). « Incentive problems in performance-based online advertising pricing : Cost per click vs. cost per action ». *Management Science* 62.7, p. 2022-2038 (cf. p. 63).
- INTEGRAL AD SCIENCE (2022). *IAS Media Quality Report Finds Ad Buyers Must Rethink Viewability Optimization in Changing Ad Market* (cf. p. 62, 64).
- JEZIORSKI, Przemyslaw et Ilya SEGAL (2015). « What makes them click : Empirical analysis of consumer demand for search advertising ». *American Economic Journal : Microeconomics* 7.3, p. 24-53 (cf. p. 11, 66).
- JIMÉNEZ DURÁN, Rafael (2022). « The economics of content moderation : Theory and experimental evidence from hate speech on Twitter ». *Available at SSRN* (cf. p. 84).
- JIMÉNEZ DURÁN, Rafael, Karsten MÜLLER et Carlo SCHWARZ (2022). « The Effect of Content Moderation on Online and Offline Hate : Evidence from Germany's NetzDG ». *Available at SSRN 4230296* (cf. p. 84).
- JOHNSON, Garrett (2022). « Inferno : A guide to field experiments in online display advertising ». *Available at SSRN 3581396* (cf. p. xiv, 73, 97, 111).
- JOHNSON, Justin P (2013). « Targeted advertising and advertising avoidance ». *The RAND Journal of Economics* 44.1, p. 128-144 (cf. p. 61).
- JOO, Mingyu, Kenneth C. WILBUR, Bo COWGILL et Yi ZHU (2014). « Television Advertising and Online Search ». *Management Science* 60.1, p. 56-73 (cf. p. 4, 13).
- JOO, Mingyu, Kenneth C. WILBUR et Yi ZHU (2016). « Effects of TV advertising on keyword search ». *International Journal of Research in Marketing* 33.3, p. 508-523 (cf. p. 4, 7, 13).
- JORDAN, Patrick, Mohammad MAHDIAN, Sergei VASSILVITSKII et Erik VEE (2011). « The multiple attribution problem in pay-per-conversion advertising ». *International symposium on algorithmic game theory*. Springer, p. 31-43 (cf. p. 5).
- KAISER, Ulrich et Julian WRIGHT (2006). « Price structure in two-sided markets : Evidence from the magazine industry ». *International Journal of Industrial Organization* 24.1, p. 1-28 (cf. p. 3, 18).
- KESLER, Reinhold (2022). « The Impact of Apple's App Tracking Transparency on App Monetization ». *Available at SSRN 4090786* (cf. p. 111).

- KIREYEV, Pavel, Koen PAUWELS et Sunil GUPTA (2016). « Do display ads influence search? Attribution and dynamics in online advertising ». *International Journal of Research in Marketing* 33.3, p. 475-490 (cf. p. 7, 13).
- KUMAR, Ashish, Ram BEZAWADA, Rishika RISHIKA, Ramkumar JANAKIRAMAN et PK KANNAN (2016). « From social to sale : The effects of firm-generated content in social media on customer behavior ». *Journal of marketing* 80.1, p. 7-25 (cf. p. 5, 7).
- KUMAR, V, JeeWon Brianna CHOI et Mallik GREENE (2017). « Synergistic effects of social media and traditional marketing on brand sales : Capturing the time-varying effects ». *Journal of the Academy of marketing Science* 45.2, p. 268-288 (cf. p. 5, 13).
- LAMBRECHT, Anja et Catherine TUCKER (2013). « When does retargeting work? Information specificity in online advertising ». *Journal of Marketing research* 50.5, p. 561-576 (cf. p. 62, 72).
- LANDES, William M et Richard A POSNER (2003). *The economic structure of intellectual property law*. Harvard university press (cf. p. xiii).
- LEVIN, Jonathan et Paul MILGROM (2010). « Online advertising : Heterogeneity and conflation in market design ». *American Economic Review* 100.2, p. 603-607 (cf. p. xiii, 63).
- LEWIS, Randall A et Dan NGUYEN (2015). « Display advertising's competitive spillovers to consumer search ». *Quantitative Marketing and Economics* 13.2, p. 93-115 (cf. p. 4, 6, 13).
- LEWIS, Randall A et Justin M RAO (2015). « The unfavorable economics of measuring the returns to advertising ». *The Quarterly Journal of Economics* 130.4, p. 1941-1973 (cf. p. 73, 97, 111).
- LEWIS, Randall A et David H REILEY (2013). « Down-to-the-Minute Effects of Super Bowl Advertising on Online Search Behavior ». *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*. EC '13. Philadelphia, Pennsylvania, USA : Association for Computing Machinery, p. 639-656 (cf. p. 4, 13).
- LIAUKONYTE, Jura, Thales TEIXEIRA et Kenneth C. WILBUR (2015). « Television advertising and online shopping ». *Marketing Science* 34.3, p. 311-330 (cf. p. 5, 7, 13).
- LIAUKONYTĖ, Jūra, Anna TUCHMAN et Xinrong ZHU (2022). « Frontiers : spilling the beans on political consumerism : do social media boycotts and buycotts translate to real sales impact? » *Marketing Science* (cf. p. 83).
- LIU, Licheng, Ye WANG et Yiqing XU (2022). « A Practical Guide to Counterfactual Estimators for Causal Inference with Time-Series Cross-Sectional Data ». *American Journal of Political Science* 0.0, p. 1-17 (cf. p. xvii, 83, 91, 101-103).
- LIU, Yi, Pinar YILDIRIM et Z John ZHANG (2022). « Implications of revenue models and technology for content moderation strategies ». *Marketing Science* (cf. p. 84).
- LULL, Robert B et Brad J BUSHMAN (2015). « Do sex and violence sell? A meta-analytic review of the effects of sexual and violent media and ad content on memory, attitudes, and buying intentions. » *Psychological bulletin* 141.5, p. 1022 (cf. p. 84).
- MADIO, Leonardo et Martin QUINN (2021). « Content moderation and advertising in social media platforms ». *Working Paper*, p. 52 (cf. p. 61, 84).
- MARSHALL, Alfred (1919). *Industry and trade*. Vani Prakashan (cf. p. xi).
- MARSHALL, Ron (2015). *How Many Ads Do You See in One Day?* (Cf. p. xi).
- MARTIN, Marc (2016). *Histoire de la publicité en France*. Presses universitaires de Paris Nanterre (cf. p. 33, 82, 86).
- MEDIA RATING COUNCIL (2014). *MRC Viewable Ad Impression Measurement Guidelines* (cf. p. 62, 64).
- (2020). *Invalid Traffic Detection and Filtration Standards Addendum* (cf. p. 63).
- MÜLLER, Karsten et Carlo SCHWARZ (2022). « The Effects of Online Content Moderation : Evidence from President Trump's Account Deletion ». *Available at SSRN 4296306* (cf. p. 84).
- MURRY, Charles (2017). « Advertising in vertical relationships : An equilibrium model of the automobile industry ». *Available at SSRN 2549247* (cf. p. 18).
- NAIK, Prasad A. et Kay PETERS (2009). « A Hierarchical Marketing Communications Model of Online and Offline Media Synergies ». *Journal of Interactive Marketing* 23.4, p. 288-299 (cf. p. 4, 7, 35).
- NELSON, Phillip (1974). « Advertising as information ». *Journal of political economy* 82.4, p. 729-754 (cf. p. xi).
- NERLOVE, Marc et Kenneth J ARROW (1962). « Optimal advertising policy under dynamic conditions ». *Economica*, p. 129-142 (cf. p. 34).

- NEVO, Aviv (2001). « Measuring market power in the ready-to-eat cereal industry ». *Econometrica* 69.2, p. 307-342 (cf. p. 12).
- NYT (2020a). *More Than 1,000 Companies Boycotted Facebook. Did It Work?* (Cf. p. 85, 86).
- (2020b). « More Than 1,000 Companies Boycotted Facebook. Did It Work? » *New York Times* (cf. p. 109).
- ONISHI, Hiroshi et Puneet MANCHANDA (2012). « Marketing activity, blogging and sales ». *International Journal of Research in Marketing* 29.3, p. 221-234 (cf. p. 4).
- ORTUTAY, Barbara (2020). « Facebook bans some, but not all, QAnon groups and accounts ». *ABC News* (cf. p. 85, 86).
- OZGA, S Andrew (1960). « Imperfect markets through lack of knowledge ». *The quarterly journal of Economics* 74.1, p. 29-52 (cf. p. xi).
- PERROT, Anne, Mathias EMMERICH et Quentin JAGOREL (2022). *Publicité en ligne : pour un marché à armes égales*. Rapp. tech. Inspection Générale des Finances (cf. p. xv, 59).
- POITRAS, Marc et Daniel SUTTER (2009). « Advertiser pressure and control of the news : The decline of muckraking revisited ». *Journal of Economic Behavior & Organization* 72.3, p. 944-958 (cf. p. 83).
- PRAT, Andrea et Tommaso VALLETTI (2022). « Attention oligopoly ». *American Economic Journal : Microeconomics* 14.3, p. 530-57 (cf. p. 3, 18, 34).
- REILEY, David H, Sai-Ming LI et Randall A LEWIS (2010). « Northern exposure : A field experiment measuring externalities between search advertisements ». *Proceedings of the 11th ACM conference on Electronic commerce*, p. 297-304 (cf. p. 13).
- RICHAUD, Nicolas (2020). « Pour la première fois, le numérique s’arroe plus de la moitié du marché publicitaire en France ». *Les Échos* (cf. p. 109).
- RUTZ, Oliver J et George F WATSON (2019). « Endogeneity and marketing strategy research : An overview ». *Journal of the Academy of Marketing Science* 47.3, p. 479-498 (cf. p. xvi, 11, 42).
- SAYEDI, Amin (2018). « Real-time bidding in online display advertising ». *Marketing Science* 37.4, p. 553-568 (cf. p. 63).
- SAYEDI, Amin, Kinshuk JERATH et Kannan SRINIVASAN (2014). « Competitive poaching in sponsored search advertising and its strategic impact on traditional advertising ». *Marketing Science* 33.4, p. 586-608 (cf. p. 3, 5, 45).
- SCHIFF, Allison (2019). « Can Contextual Targeting Replace Third-Party Cookies? » *AdExchanger* (cf. p. 72).
- SEAMANS, Robert et Feng ZHU (2014). « Responses to Entry in Multi-Sided Markets : The Impact of Craigslist on Local Newspapers ». *Management Science* 60.2, p. 476-493 (cf. p. 35).
- SELDON, Barry J., R.Todd JEWELL et Daniel M. O’BRIEN (2000). « Media substitution and economies of scale in advertising ». *International Journal of Industrial Organization* 18.8, p. 1153-1180 (cf. p. 35, 39, 42, 51).
- SELDON, Barry J. et Chulho JUNG (1993). « Derived demand for advertising messages and substitutability among the media ». *The Quarterly Review of Economics and Finance* 33.1, p. 71-86 (cf. p. 35, 39, 42, 51).
- SHAMDASANI, Prem N, Andrea JS STANALAND et Juliana TAN (2001). « Location, location, location : Insights for advertising placement on the web ». *Journal of Advertising Research* 41.4, p. 7-21 (cf. p. 61).
- SHAPIRO, Bradley T, Günter J HITSCH et Anna E TUCHMAN (2021). « TV advertising effectiveness and profitability : Generalizable results from 288 brands ». *Econometrica* 89.4, p. 1855-1879 (cf. p. 12).
- SHEHU, Edlira, Nadia ABOU NABOUT et Michel CLEMENT (2021). « The risk of programmatic advertising : Effects of website quality on advertising effectiveness ». *International Journal of Research in Marketing* 38.3, p. 663-677 (cf. p. xv, 11, 61, 63, 66, 83, 97).
- SHFP (2020). *Statement from Stop Hate For Profit on Meeting with Facebook* (cf. p. 85).
- SHILLER, Robert J (2017). « Narrative economics ». *American economic review* 107.4, p. 967-1004 (cf. p. xii).
- SILK, Alvin J, Lisa R KLEIN et Ernst R BERNDT (2002). « Intermedia substitutability and market demand by national advertisers ». *Review of Industrial Organization* 20.4, p. 323-348 (cf. p. 35, 39, 51).
- SIMONOV, Andrey, Chris NOSKO et Justin M RAO (2018). « Competition and crowd-out for brand keywords in sponsored search ». *Marketing Science* 37.2, p. 200-215 (cf. p. 3).
- SRIDHAR, Shrihari et Srinivasaraghavan SRIRAM (2015). « Is online newspaper advertising cannibalizing print advertising? » *Quantitative Marketing and Economics* 13.4, p. 283-318 (cf. p. 35).
- STATISTA (2021a). *Global : travel & tourism sales channel share* (cf. p. 8, 39).

-
- (2021b). *Net digital display ad revenues of Facebook as percentage of total digital display ad revenues in the United States from 2016 to 2019* (cf. p. 18).
- (2022a). *Dépenses publicitaires en Europe de l'Ouest de 2015 à 2021 (en millions de dollars des États-Unis), par média* (cf. p. 35).
- (2022b). *Number of monthly active Facebook users worldwide as of 3rd quarter 2022* (cf. p. 87).
- (2022c). « Soft Drinks - Worldwide (Statista Outlook) ». *Statista* (cf. p. 39).
- STERN, David I (2011). « Elasticities of substitution and complementarity ». *Journal of Productivity Analysis* 36.1, p. 79-89 (cf. p. 37).
- STIGLER, George J (1961). « The economics of information ». *Journal of political economy* 69.3, p. 213-225 (cf. p. xi).
- STIGLER, George J et Gary S BECKER (1977). « De gustibus non est disputandum ». *The american economic review* 67.2, p. 76-90 (cf. p. xii).
- STIPP, Horst (2018). « How Context Can Make Advertising More Effective ». *Journal of Advertising Research* 58.2, p. 138-145 (cf. p. 59, 82, 83).
- STOCK, James et Motohiro YOGO (2005). « Testing for Weak Instruments in Linear IV Regression ». *Identification and Inference for Econometric Models*. New York : Cambridge University Press, p. 80-108 (cf. p. 51).
- TUCKER, Catherine (2013). « The Implications of Improved Attribution and Measurability for Antitrust and Privacy in Online Advertising Markets ». *Georges Mason Law Rev* 203, p. 1025-1054 (cf. p. 5).
- (2014). « Social networks, personalized advertising, and privacy controls ». *Journal of marketing research* 51.5, p. 546-562 (cf. p. 11, 66).
- UHL, Christina, Nadia Abou NABOUT et Klaus MILLER (2020). « How much ad viewability is enough ? The effect of display ad viewability on advertising effectiveness ». *Working Paper* (cf. p. 62).
- VAN HEERDE, Harald J, Maarten J GIJSENBURG, Marnik G DEKIMPE et Jan-Benedict EM STEENKAMP (2013). « Price and advertising effectiveness over the business cycle ». *Journal of Marketing Research* 50.2, p. 177-193 (cf. p. 12).
- VILLAGRA, Nuria (2021). « Firm value impact of corporate activism : Facebook and the stop hate for profit campaign ». *Journal of Business Research*, p. 8 (cf. p. 84).
- WILBUR, Kenneth C. et Yi ZHU (2009). « Click fraud ». *Marketing Science* 28.2, p. 293-308 (cf. p. xv).
- XU, Yiqing (2017). « Generalized synthetic control method : Causal inference with interactive fixed effects models ». *Political Analysis* 25.1, p. 57-76 (cf. p. 91, 92, 101, 102).
- ZENETTI, German, Tammo HA BIJMOLT, Peter SH LEEFLANG et Daniel KLAPPER (2014). « Search engine advertising effectiveness in a multimedia campaign ». *International Journal of Electronic Commerce* 18.3, p. 7-38 (cf. p. 19).
- ZENITH (2021). *Global internet ad spend by format 2023* (cf. p. 59).
- ZHANG, Kaifu et Zsolt KATONA (2012). « Contextual advertising ». *Marketing Science* 31.6, p. 980-994 (cf. p. 61).
- ZIGMOND, Dan et Horst STIPP (2010). « Assessing a New Advertising Effect : Measurement of the Impact of Television Commercials on Internet Search Queries ». *Journal of Advertising Research* 50.2, p. 162-168 (cf. p. 4).

RÉSUMÉ

La publicité en ligne abaisse considérablement les coûts de ciblage du consommateur. Cette thèse étudie les apports et limites de la publicité en ligne au travers de 4 études empiriques basées sur des données d'annonceurs. Le **Chapitre 1** montre que les liens sponsorisés (publicités *search*) bénéficient largement des publicités hors-ligne. Je montre qu'en augmentant son activité publicitaire hors-ligne d'1%, une marque génère jusqu'à 0.95% de clics supplémentaires sur ses liens sponsorisés. Le **Chapitre 2** porte sur la substituabilité des publicités hors-ligne et numériques. À l'aide d'un modèle translog, je conclus que les publicités hors-ligne et numériques sont des substituts limités. Les formats publicitaires numériques (*display* et *search*) sont en revanche fortement substituables. Dans le **Chapitre 3**, je m'intéresse aux asymétries d'information quant au placement des publicités en ligne. Je montre que les contrats en coût-par-impression (CPM) incitent moins les régies à rendre la publicité visible que les contrats de coût-par-vue. L'achat programmatique — basé sur des intermédiaires publicitaires — expose l'annonceur à une visibilité et une audience de moins bonne qualité que l'achat en direct. En outre, l'adéquation des publicités aux contenus des sites engendre des taux de clic 69% supérieurs au ciblage des consommateurs indépendamment du contexte. Enfin, si les inventaires premium ne sont pas plus cliqués à court terme, ils semblent chasser les mauvais espaces du marché standard. Les effets de contexte sont également abordés dans le **Chapitre 4**. À partir de modèles en double-différences et d'une estimation de contrefactuels, je montre que la circulation des contenus controversés et la dégradation de l'image de Facebook durant le boycott de juillet 2020 ont altéré la valeur des publicités sur la plateforme. Les publicités Facebook ont enregistré 5000 à 10 000 clics de moins, comparées aux autres formats *display* de la marque. Leur prix a également baissé. Cette thèse conclut que la publicité en ligne est davantage un complément de la publicité traditionnelle qu'un substitue. Enfin, elle plaide pour une meilleure contextualisation de la publicité. Celle-ci apparaît aujourd'hui essentielle à l'heure où la régulation limite l'usage des données personnelles à des fins publicitaire.

MOTS CLÉS

Publicité, Économétrie Appliquée, Ciblage, Externalités

ABSTRACT

Online advertising significantly lowers the costs of targeting individuals. This thesis studies the contributions and limitations of online advertising through 4 empirical studies based on advertiser data. **Chapter 1** shows that sponsored links (*search* ads) benefit greatly from offline ads. I show that when increasing its offline advertising activity by 1%, a brand generates up to 0.95% additional clicks on its sponsored links. **Chapter 2** focuses on the substitutability between offline and digital ads. Using a translog model, I find that offline and digital ads are limited substitutes. Digital ad formats (*display* and *search*) are highly substitutable. **Chapter 3** focuses on information asymmetries in the placement of online ads. I show that cost-per-impression (CPM) contracts do not provide incentives for advertisers to make ads visible compared to cost-per-view contracts. Programmatic buying - based on advertising intermediaries - exposes the advertiser to a lower visibility and audience quality compared to direct buying. In addition, matching ads with website content results in 69% higher click-through rates than ads that only target consumers regardless of context. Finally, while ads bought from premium inventories are not more clicked, it seems to be driving out low-quality ad spaces from the standard inventories. Context effects are also discussed in **Chapter 4**. Using differences-in-difference and counterfactual estimations, I show that the circulation of controversial content and the degradation of Facebook's credibility during the July 2020 boycott altered the value of ads on the platform. From June to July 2020, Facebook ads recorded 5,000 to 10,000 fewer clicks compared to the brand's other display campaigns. Their price also dropped. This thesis concludes that online advertising is more a complement to traditional advertising than a substitute. I also advocate for a better contextualization of advertising. This appears to be essential as regulation limits the use of personal data for advertising purposes.

KEYWORDS

Advertising, Applied Econometrics, Targeting, Externalities

