

What Do News Aggregators Do?

Evidence from Google News in Spain and Germany*

Joan Calzada[†]

Ricard Gil[‡]

Universitat de Barcelona

Johns Hopkins Carey Business School

June 20, 2017

Abstract

The impact of aggregators on news outlets is ambiguous. In particular, the existing theoretical literature highlights that although aggregators create a market expansion effect when they bring visitors to news outlets, they also generate a substitution effect if some visitors switch from the news outlets to the aggregators. Using the shutdown of the Spanish edition of Google News in December of 2014 and difference-in-differences methodology, this paper empirically examines the relevance of these two effects. We show the shutdown of Google News in Spain decreased the number of daily visits to Spanish news outlets by 14%, and that this effect was larger in outlets with less overall daily visits and a lower share of international visitors. We also find some evidence suggesting that the shutdown decreased online advertisement revenues for larger online news outlets. We then analyze the effect of the opt-in policy adopted by the German edition of Google News in October of 2014. Although such policy did not significantly affect the daily visits of all outlets that opted out, it reduced by 8% the number of visits of the outlets controlled by the publisher Axel Springer. Our results demonstrate the existence of a market-expansion effect through which news aggregators increase consumers' awareness of news outlets' contents, thereby increasing their number of visits.

*We thank Shane Greenstein, Avi Goldfarb, Hal Varian, Susan Athey, Omar Nayeem, Mike Bailey, Lisa George, Paul Seabright, Lesley Chiou, Gerard Llobet, and audiences at the Winter NBER Meeting on the Economics of Digitization, Eighth Annual Conference on Internet Commerce and Innovation at the Searle Center of Northwestern University, Johns Hopkins Carey Business School, American University, the Federal Communications Commission, Universitat de Barcelona, Universitat Oberta de Catalunya, the 10th Conference on "The Economics of Intellectual Property, Software and the Internet" at Toulouse, and the FSR Communications & Media, CMPF and ENTraNCE Annual Scientific Seminar on the Economics, Law and Policy of Communications and Media in Florence, Italy, and the First Catalan Economic Society Meetings in Barcelona, Spain, for their valuable comments and suggestions. The usual disclaimer applies.

[†]Universitat de Barcelona. Av. Diagonal 690, Barcelona 08034, Spain; calzada@ub.edu.

[‡]Johns Hopkins Carey Business School. 100 International Drive. MD 21202 Baltimore, US; ricard.gil@jhu.edu.

1 Introduction

Online platforms and aggregators have drastically changed how consumers and businesses gain access to information and interact with each other. While consumers now use aggregators and search engines to find all sorts of goods and services such as flights, accommodation, or insurance, firms have adapted their online distribution channels to the growing presence of aggregators to remain competitive. The academic community has also noticed the importance of this shift in consumption patterns and business practices; consequently, the shift has generated considerable interest among researchers in economics and management who aim to understand the role of aggregators in online markets.

A type of aggregator that has rapidly grown in importance is the one simplifying the search of news stories such as Google News, Yahoo! News, Bing News, or Summify. News aggregators offer links to news stories published by news outlets, which are usually complemented with excerpts and images. They allow consumers to save considerable time and effort in finding news. But in spite of the growing appeal of these websites to internet users, traditional news outlets around the world have been reticent to their introduction, because they are concerned about the potential effects on their audiences and on the consumers' browser activity. Even though news outlets can opt-out of aggregators by using software that blocks the links to their content, most publishers want to be indexed while receiving some economic compensation for the use of their news stories. This situation has generated multiple frictions between Google and publishers in Europe, and has led to changes in intellectual property laws in several countries.¹ This paper examines two recent events of disputes of Google News in Spain and Germany that prove useful to understand the role and impact of news aggregators on the news market in different ways. Whereas Google News completely shut down its Spanish edition, in the German case, a group of publishers decided to reduce their presence in the German edition of Google News.

The theoretical literature has identified an important trade-off in the effects that news aggregators might have on news outlets' audiences (Dellarocas et al., 2013; George and Hogendorn, 2012; Jeon and Nasr, 2015; and Calzada and Ordoñez, 2015). On the one hand, aggregators create a "market-expansion effect" because they allow consumers to discover news stories and news outlets that otherwise they would not know. Consumers also save time that they can dedicate to reading more news stories. However, the number of indirect visits received by news outlets from links to news aggregators depends on several factors such as font size of headlines, the number of words or use of images in the news excerpts. If the information released by the news aggregator is too detailed consumers may not need or be interested in clicking to the original link to read the complete news story (Dellarocas et al., 2015). On the other hand, aggregators also create a "substitution effect," because with their presence consumers may not visit the front page of the news outlets, reducing the number of visits to other pages of the news outlets and their advertisement revenues. Therefore, determining whether the indirect visits generated by the market-expansion effect compensate for

¹Google News was launched in 2002 and today is present in more than 70 countries in 35 different languages. Additionally, it also has many regional editions that are adapted to the specific needs of the audiences.

the direct visits lost by the substitution effect is a relevant empirical question. Moreover, from a policy perspective, knowing which news outlets are more likely to benefit from news aggregators and how they modify consumers’ engagement habits is also important.

This paper sheds light on these questions by first analyzing the impact of the shutdown of Google News in Spain in December of 2014. At the beginning of 2014, a reform of the Spanish intellectual property law established that firms posting links and excerpts of news stories have to pay a compulsory link fee (Google tax) to the original publishers. As a consequence, on December 16, 2014, Google News decided to shut down its Spanish edition, arguing that under the new regulation, this service would not be profitable. We exploit this quasi-natural experiment to examine the effect of news aggregators on the number of visits received by news outlets and on the consumers’ engagement metrics.

We then complement this analysis by considering a change in the linking policy of Google News adopted in Germany around the same time. In 2013, the German Parliament introduced a change in the copyright law that allowed news aggregators to link for free the news stories of news outlets if using excerpts of less than 7 words. Longer excerpts or images would require the payment of a negotiated fee to the news outlets. After a number of disputes with some German outlets, Google News finally decided to change from an “opt-out” to an “opt-in” policy. Under the new opt-in rule, those German publishers that want to be indexed by Google News must give up receiving any compensation from the aggregator. After this change in Google News’ linking policy, on October 23, 2014, more than 200 publishers decided not to opt in to protect their contents, but soon reversed their decision and accepted Google’s opt-in conditions, allegedly due to drastic traffic losses.

Our study draws from a rich data set obtained from SimilarWeb containing information for 109 newspapers in Spain, Germany, and France. This data set includes information about each domain’s daily number of visits, the monthly percentage of visits by origin (direct, from a social platform, from a search engine, or referrals), and several engagement metrics per day (page views per visit, visit duration, and bounce rate). We complement these data with information on advertisement revenues from printed and online editions of Spanish news outlets in our sample from Arce Media.

Taking into account the source of plausible exogenous variation provided by Google’s shutdown in Spain and its opt-in policy in Germany, we apply a model of difference-in-differences to assess the role of this aggregator in the news market, using French newspapers as the control group. Our results show that after Google News’ shutdown in Spain, Spanish news outlets experienced an 14% reduction in the number of visits, with a growing impact during the first six weeks. We repeat our analysis, focusing on news outlets with a high percentage of foreign visits, in an attempt to homogenize our sample of outlets in both the treated and control group, and obtain a reduction of 10% in the number of daily visits after the Google News’ shutdown. Our findings confirm previous analysis in the empirical literature showing that news aggregators have a net positive impact in outlets’ traffic (Chiou and Tucker, 2017). We also show the impact of the shutdown depends on the outlets’ specialization. Specifically, we find the shutdown affected sports outlets the most, had a lower effect on national and regional outlets, and did not significantly affect business outlets.

So far as engagement metrics are concerned, we find the shutdown of Google News reduced the

average duration of a visit by 10%, but we do not observe a long-term impact in other variables. When we decompose the effect by week, our results show that in the first few weeks, the number of pages visited decreased and the bounce rate increased, but these effects vanished in the long term. These findings suggest the shutdown did not produce a significant change in the composition of the consumers that visit the news outlets. Perhaps more surprisingly, we observe that search traffic decreased by 1.3% and direct traffic increased by 1.5%, respectively, after the shutdown. These results directly suggest the existence of a market-expansion effect and a substitution effect whereby search traffic (partly) replaces the outlets' direct traffic.

Our analysis of the effect on advertisement revenues deserves a careful description. On the one hand, our empirical analysis comparing print and online editions shows that advertisement revenues, the number of advertisers and the intensity of advertisement increased after the shutdown for the digital editions relative to print editions. This may be due to different trends in advertisements prices and the circulation/traffic for the two types of editions. In this sense, the announcers' growing interest to advertise their products in the Internet, and the increase in the consumption of digital news outlets may have overcompensated the effects of the shutdown. On the other hand, our analysis comparing larger and smaller online outlets reveals that revenues and intensity of advertisement in larger news outlets (those with above the media advertisement revenues) decreased relative to those of the rest after the shutdown.

The analysis of the German case shows Google News' opt-in policy generated a negligible effect on the visits to the news outlets that decided to stay out. Only when we restrict the analysis to the news outlets controlled by the publisher Axel Springer, do we find a 8% reduction in daily visits. We also find that during the period Axel Springer opted out, the percentage of search visits of Axel Springer's outlets decreased by 1% and the percentage of referral visits increased by 1%. Axel Springer controls (and owns stock in) a number of important media firms and news outlets in Germany, and strongly advocated for a change in the German copyright law. When Google adopted the opt-in policy, Axel Springer and other publishers promoted the opt-out option, but only the outlets in the VG Media consortium followed. Consequently, the loss of visits to these outlets can be related to their shorter excerpts (market-contraction effect) and to their lower attractiveness relative to other outlets (competition effect) that opted in. Eventually, the latter effect and the salient loss of positions in the national ranking for some of its most emblematic newspapers motivated Axel Springer to finally accept Google's opt-in conditions.

A few papers have analyzed the effects of aggregators in the news market. Athey and Mobius (2012) examine the impact that introducing local news headlines and links in Google News has on consumers' browsing activity. In 2009, the French edition of Google News enabled a local news feature that allowed those users that entered their zip code to obtain news from local outlets. Using a data set of user browsing behavior, the authors compare users who adopted the localization feature with a sample of control users that exhibited similar consumption patterns in the past. They find the addition of local news content led consumers to rely more on Google when initiating a browsing session. Moreover, after the introduction of this feature, direct navigation to local outlets increased by 5% (bypassing Google News altogether), and clicks on local outlets from the Google News page

increased by 13%. In a related paper, George and Hogendorn (2013) use a major redesign in the US edition of Google News on June 30, 2010. Similar to the previous case, the redesign placed a permanent strip of geo-targeted local news headlines onto the Google News front page. Using a sample of news visits by US households before and after the introduction of the geo-targeted links, the authors find local news visits increased by less than 1% and the likelihood of a local news visit increased between 4% and 6% from a low baseline for heavy Google News users. Interestingly, the results show no evidence of substitution away from direct outlet visits. Adding geo-targeted links increased the number of different local outlets visited per day, but not the number of unique sites visited per month. Finally, Chiou and Tucker (2017) analyze the impact of a contract dispute between Google News and The Associated Press (AP) when Google News removed all AP news articles from December 23, 2009 until sometime in February 2010. Using weekly data on the top 150 sites users navigated immediately after visiting Google News or Yahoo! News, they find Google News users were less likely to visit other news websites after visiting Google News following the removal of AP content, relative to Yahoo! News users, who did not experience such a content change.

Our paper adds to this literature by providing evidence of the change in daily visits per news outlet before and after the Spanish edition of Google News stopped operations. This natural experiment allows us to compare the treatment group (Spanish outlets) with a control group (French outlets) for which no changes to the same news aggregator occurred. We complement the analysis of this event with the examination of the opt-in policy of Google in Germany. This case allows us to measure whether the impact of news aggregators on news outlets depends on how the news aggregator exhibits information. Our data set contains information about all the daily visits that outlets receive from different sources (direct, search, referrals, and social), which allows us to identify the direct relevance of a market-expansion and a substitution effect. In addition, we use consumers' engagement metrics to examine the navigation habits of outlets' direct and search visitors.² The singularity of the two events described and the granularity of the data collected allows us to test these predictions from the theoretical literature. Moreover, our analysis of the effect of the shutdown on the news outlets' advertisement is important to understand the economic impact of the shutdown of Google News in Spain.

To the best of our knowledge, the closest paper to ours is Athey, Mobius and Pal (2017), which parallelly study the impact of the Google News shutdown in Spain using different data and identification strategy, and yet finding very similar results to ours (10% reduction in the number of visits to news publishers). While both papers implement standard difference-in-differences estimation techniques, our paper uses data on the daily number of visits to a sample of Spanish news outlets and a control group of French news outlets to, and theirs use individual-level browsing data to compare Google News users and non-Google News users in Spain. The results of both papers are in fact complementary in that our data is at the domain level but includes all visits from desktop

²Aggregators may reduce the number of loyal users and increase the number of casual consumers. This substitution between user types can reduce advertising revenues at some news outlets (Athey, Calvano, and Gans, 2011).

users, and their data is at the individual level for only users of Microsoft products (accounting for half of PC news browsing activity). Our paper also offers results on the revenue impact in Spain, as well as the effects of the opt-in Google policy in Germany, finding consistent results with the Spanish case.

Finally, the paper can also be related to other recent research that has examined the effects of news aggregators and indexing beyond the Google News case. Roos et al. (2015) investigate the influence of excerpts on consumers' decisions to consume news. They show that observing just one excerpt reduces consumers' uncertainty about their match with the excerpted site's content by about 33%. They conclude that excerpting benefits the linked site by increasing the share of traffic originating at the linking site, and benefits the linking site by making it more popular at the start of consumers' browsing sessions. The paper also finds that excerpting increases news consumption, leading consumers to browse more frequently and visit a wider range of sites. Relatedly, Cage et al. (2015) examine 84 general information media outlets in France (including newspapers, television channels, radio stations, and news agencies), and track every article these sites offered online in 2013, with the help of a plagiarism-detection algorithm that quantifies the copy rate between an article and all the articles previously published about the event. They find that half of online information production is copy-and-paste. They also explain that those outlets that produce more content receive more visits, but the rapid spillover of information occurring in the last few years has reduced the incentives of newspapers to produce original news stories. Finally, Sismeiro and Mahmood (2017) study the impact of an exogenous Facebook outage on the number of visits to a large online news website in a major Western European country. Their results are consistent with the literature in that they also find evidence of a net market-expansion effect of online social networks on news outlets visits. Their paper points out that despite these similarities, social networks differ from news aggregators in two ways. First, while aggregators rely on algorithms, online social network rely on friends to filter and recommend content. Second, because the scope of activities performed at online social networks is far richer, complementarities between news stories and recommending parties may be an important driver on clicking behavior and reading patterns.

The rest of the paper is structured as follows. Section 2 explains the main institutional details of the Spanish and German cases. Section 3 describes our data set. Section 4 analyzes the shutdown of Google News in Spain. Section 5 assesses the impact of the opt-in policy adopted by Google News in Germany. Finally, section 6 concludes.

2 Institutional Details: Google's Disputes with European Publishers

Since the release of Google News in 2002, news publishers around the world have fought against the free indexing of their content while advocating for receiving some economic compensation from Google. Even though this situation has generated several legal disputes, some European governments have recently considered creating a link fee (Google tax) that would force news aggregators

to compensate the linked outlets.³ This section first depicts the earlier history of the relationship between Google and the European news publishers and later describes the creation of a link fee in Spain and Germany, which motivate our empirical analysis.⁴

Belgium was among the first countries to regulate the activities of news aggregators. In 2006, Copiepresse (representing French- and German-language Belgian publishers) sued Google News over alleged copyright infringement. Consequently, in 2006 and 2007, two sentences forbid Google News to link the contents of Belgian publishers without their consent.⁵ In 2011, the Belgian Appeals Court ratified these decisions and established that the mere linking of newspaper websites should be considered infringement. Soon after this resolution, the Belgian publishers asked to be linked back to Google News, and on December 12, 2012, Google agreed to index Copiepresse newspapers under the condition of no future legal action for copyright infringement. The agreement also established that the two parts would partner on several business initiatives to promote both the publishers' and Google's services.⁶ Thierry Geerts, Google Belgium's managing director, clearly announced how Google aimed to address similar disputes in other countries⁷: "Instead of continuing to argue over legal interpretations, we have agreed on the need to set aside past grievances in favor of collaboration. This is the same message we would like to send to other publishers around the world - it is much more beneficial for us to work together than to fight."

Similar to the Belgian case, several publishers in France lobbied the French government in 2012 to create a link fee. Google reacted to this initiative by threatening to close its French edition if this measure were approved. By February 1, 2013, French President François Hollande and Google Executive Chairman Eric Schmidt reached an agreement such that French publishers agreed to forego the establishment of a link fee and Google agreed to create a €60 million Digital Publishing Innovation Fund that would support transformative digital publishing initiatives for French readers. Google also offered to help French publishers increase their online revenues using its advertising technology, which allow for better targeting of consumers.⁸

More generally, an overall revision of Google News' disputes in Europe shows the firm's strategy has been to lobby against the establishment of a link fee while investing numerous resources to gain the publishers' support. A clear example of such strategy is the launch by Google on April 2015 of the Digital News Initiative (DNI), partnering with Les Echos (France), FAZ and Die Zeit (Germany), the Financial Times and the Guardian (UK), NRC Group (Netherlands), La Stampa

³Similarly, in September of 2016 the European Union announced its intention to reform the copyright legislation, which might entail the creation of the so-called neighboring rights to protect the contents of press publishers. See news release here, http://europa.eu/rapid/press-release_IP-16-3010_ca.htm

⁴In 2012 the National Association of Newspapers in Brazil persuaded its 154 members to ban Google News to use its contents, arguing that Google was refusing to pay for the links and was driving traffic away from their websites. In the following years some news outlets allowed back links from the aggregator. <http://www.bbc.com/news/world-latin-america-20018221>.

⁵Today, Google offers news outlets the option to opt out of Google News if they feel harmed by the links. See, for example, the agreement between Google and the Italian anti-trust authorities in 2011. http://www.nytimes.com/2011/01/18/technology/18iht-google18.html?_r=2

⁶<http://www.theverge.com/2012/12/13/3764692/google-copyright-lawsuit-settlement-belgium>

⁷<http://googlepolicyeuropa.blogspot.de/2012/12/partnering-with-belgian-news-publishers.html>

⁸<https://googleblog.blogspot.fr/2013/02/google-creates-60m-digital-publishing.html>

(Italy), and El Pais and Grupo Godo (Spain). DNI will dedicate €150 million to projects that support innovation in digital news journalism over the next three years, and will invest in training and development resources for journalists and newsrooms across Europe.⁹

2.1 The Shutdown of Google News in Spain

The main dispute between Google News and the European publishers took place in Spain. On January 1, 2014, the Spanish Parliament passed a reform of the Law of Intellectual Property (LPI).¹⁰ The new law established that online outlets posting links and excerpts of news articles originated elsewhere must pay a link fee (*canon*) to the original publishers. The creation of the link fee was initially promoted by the publishers association AEDE (Asociación de Editores de Diarios Españoles), which lobbied the government to force news aggregators to compensate them for the use of their content.¹¹

A unique feature of the Spanish regulation is that publishers cannot refuse to receive a fee from news aggregators, in fact, the link fee must be collected by a private entity called CEDRO which will distribute back the revenues to the news outlets. This strategy tries to prevent publishers from giving away their right to receive compensation, and to enforce coordination among publishers. Note that if the fees were voluntary, some publishers could negotiate exclusivity agreements with Google and put their rivals at a competitive disadvantage.

Although the implementation of the law involved a lot of uncertainty, on December 11, 2014, Richard Gingras, world responsible of Google News, unexpectedly announced that on December 16, Google News would shut down its Spanish edition.¹² Google justified this action by claiming that the new regulation made the service unprofitable because Google News had no direct source of revenues (the firm does not show any advertising on this site).¹³ Google's decision was shortly followed by other, yet smaller, Spanish news aggregators such as Planeta Ludico, NiagaRank, Multifriki, InfoAliment, and Beeeinfo. Others tried to modify their content to avoid the effects of the law (Planet Ubuntu, Astrofisica, and Fisica).

The shutdown of Google News had an important and immediate impact on the Spanish news market. Some early reports estimated a reduction in the daily visits of the largest newspapers of more than 8%, and even bigger for smaller newspapers (NERA, 2015). As a result, the publishers in AEDE and other associations have urged the government to negotiate a solution with Google.¹⁴ Some large publishers in AEDE have even announced they would renounce any compensation payment for sharing content with news aggregators. In spite of this backlash, the solution to this

⁹See <http://googleespana.blogspot.com.es/2015/04/google-y-editores-de-medios-de-europa.html>

¹⁰<https://www.boe.es/boe/dias/2014/11/05/pdfs/BOE-A-2014-11404.pdf>

¹¹The passing of this regulation was not free of controversy. Whereas some of the biggest Spanish publishers argued in favor of it, others, such as AEPP (Asociación Española de Editoriales de Publicaciones Periódicas) and Coalición Pro-Internet, opposed it. The Spanish regulator CNMC (Comisión Nacional de los Mercados y la Competencia) also advocated for the modification of several aspects of the new regulation. See CNMC (2014) and Llobet (2015).

¹²<http://googleespana.blogspot.com.es/2014/12/novedades-acerca-de-google-noticias-en.html>

¹³<https://support.google.com/news/answer/6140047?hl=es>

¹⁴<http://www.aede.es/wp-content/uploads/2015/02/AEDEPrensa-CierreGoogleNewsDic14.pdf>

case may be delayed until the European Commission approves its new copyright legislation, which could modify the regulatory framework to protect publishers in the European Union.

2.2 The Opt-in policy of Google News in Germany

The second case that we examine in this paper is a dispute between Google News and the German news publishers. On March 1, 2013, the German Parliament passed an addendum to the copyright law that granted publishers the right to charge search engines and other online aggregators for reproducing their content, but the law also allowed the free use of text in links and brief excerpts. This addendum meant publishers can prohibit aggregators from using their news articles beyond headlines and short excerpts, and they can charge aggregators a link fee if the aggregators make a larger use of their contents. The main differences in this regulation with regards to the Spanish case are that (1) link fees have to be negotiated between the parties, and (2) it does not affect brief excerpts.

In June 2014, VG Media,¹⁵ a consortium of more than 200 publishers, including Axel Springer, sued Google and other news aggregators for displaying excerpts and preview images along with the links to their news articles. VG Media alleged that aggregators were using their content without their consent, and that according to the new law, they should receive compensation.¹⁶ Google refused to pay the publishers, and instead modified its linking policy. On October 2, 2014, the German edition of Google News announced the change from an opt-out to an opt-in system. This change implied that those German publishers that want to be indexed by Google News must explicitly grant permission and renounce any type of compensation.¹⁷

After this change in Google’s policy, publishers, and TV and radio stations associated with VG Media decided not to opt in. A leading publisher in this group was Axel Springer, which asked VG Media not to issue free licenses for its websites (welt.de, computerbild.de, sportbild.de, and autobild.de). Other publishers that followed the same course of action were Burda (bunte.de), Funke, Madsack, and M. DuMont Schaubergas. Phillip Justus, Managing Director of Google Germany, answered that Google “will not show in the future snippets and thumbnails of the publishers members of VG Media.”¹⁸

On October 23, 2014, Google News and other German news aggregators stopped showing large excerpts, video, and images from the publishers that did not opt in, in order to avoid paying them a link fee. This change allegedly significantly reduced the number of daily visits VG Media news sites received from Google and overall. Mathias Döpfner, Axel Springer Chief Executive, estimated

¹⁵In December 2015, VG Media had 332 members. <https://www.vg-media.de/de/alle-wahrnehmungsberechtigte.html>

¹⁶VG Media accused Google of “extortion” and “discrimination” against its members, saying the law passed the previous year guaranteed news providers a “reasonable compensation for the use of their content.” <http://phys.org/news/2014-10-google-news-snippets-germany-legal.html>

¹⁷Google continues to use an opt-out policy outside Germany. Publishers must take several actions to prevent indexing of their contents. For example, they must use a robot.txt file or metadata to let Google’s crawlers know what to ignore.

¹⁸<https://germany.googleblog.com/2014/10/news-zu-news-bei-google.html>

that the downgrading of search notices resulted in a loss of nearly 40% in traffic volume, and that the traffic from Google News was down by almost 80%. Moreover, welt.de dropped below its competitors in the IVM and AGOF rankings, and computerbild.de lost its Top 10 rank of all AGOF offerings in Germany.¹⁹ Shortly after, on November 5, 2014, Axel Springer and other VG Media publishers decided to opt in and gave Google a license to add excerpts to their search results for free.²⁰

3 Data

In this study, we use data at the domain level from SimilarWeb, a web measurement company providing traffic data and user-engagement statistics. This firm collects data on browsing behavior from rich and diversified panels of consumers in several countries. The data we use in this study come exclusively from desktop users. The information covers the period from June 1, 2014, to May 31, 2015, which includes the two events analyzed in the paper. Google News’ shutdown on December 16, 2014 affected Spanish news outlets. Therefore, our data cover roughly half a year before and after Google News’ shutdown in Spain. Google’s remove of excerpts and images from October 23, 2014, to November 5, 2014, affected the German news outlets belonging to the VG Media consortium.

To explore the impact of the Google News shutdown in Spain on news outlets, we chose French news outlets as control group after considering other countries such as Germany, Portugal or Italy. While discarding Germany as a control group was a no-brainer because of its own “Google News related” turbulence in October 2014, Portugal and Italy seemed very different from Spain in terms of its internet usage and demographics of internet consumers. Table shows 1 that Spain’s internet penetration rates is far higher than Portugal and Italy. Similarly, the percentage of internet users per age bracket are also fairly different between Spain, Italy and Portugal. Despite the notable difference in population size, Spain and France are more similar in internet penetration rates and internet population per age bracket than Spain is with other countries. These similarities makes France and French news outlets an adequate control group for the exercise in place. Regarding our analysis of the German market, we start using French news outlets as a comparison group to German news outlets for similar reasons. Our analysis takes us in the end to compare Axel Springer outlets to other German news outlets in our sample.

The Spanish and French news outlets in our sample have been selected according to audience rankings published by Alexa (www.alexa.com) by country. We picked top rated news outlets excluding webpages from TV and radio stations, and other potential news aggregators

¹⁹ Axel Springer SE detailed that the financial damage resulting from lost marketing revenues could be in the seven-figure range per brand over the year as a whole. http://www.axelspringer.de/en/presse/Axel-Springer-concludes-its-data-documentation-Major-losses-resulting-from-downgraded-search-notices-on-Google_22070687.html

²⁰ Mathias Doepfner used this case to illustrate Google’s ability to modify competition in the media market: “As sad as it is, at least now we know precisely how enormous the consequences of discrimination are, how Google’s market power really plays out, and how Google punishes those who exercise the right to protect content.” <http://www.ibtimes.com/news-media-no-match-google-axel-springers-about-face-1720494>

such as MSN or Yahoo. While including other news aggregators would have been potentially interesting, our data sources was not able to split its main page from its news page. Thus, for example, we could not differentiate `es.yahoo.com` from `es.noticias.yahoo.com`. In order to classify news outlets, we searched for verbal descriptions in Alexa and Wikipedia (for example, https://en.wikipedia.org/wiki/List_of_newspapers_in_France). Overall, we aimed to have a well balanced sample of news outlets classified in different categories such their specialization (national regional, business, or sports), their rank at the national level and their internationalization level.

In the end, our data set contains information for 109 domains, including 46 news outlets from Spain, 33 from Germany, and 30 from France. Table 2 offers a complete listing of all 109 domains. It also includes information about the Spanish, German, and French editions of Google News (`news.google.es`, `.de`, and `.fr`) and about two additional Spanish news aggregators (`meneame.net` and `kiosko.net`). All domains are classified according to different criteria. First, we categorize them according to their specialization. They can be *National*, *Regional*, *Business*, or *Sports*. Spain also has four regional newspapers written in the *Catalan* language. Second, we divide domains according to their national rank. Specifically, we distinguish between the *Top 50%* and the *Bottom 50%* of domains of our sample. Third, we classify domains according to the number of visits they receive from other countries. *Top International* outlets are those that receive more than 25% of the visits from abroad. *Top International 50%* and *Bottom International 50%* separate the outlets of the sample into two groups according to whether their share of international visits is above or below the median in our sample. Finally, in the case of Germany, we also consider whether the domains belong to the *VG Media* consortium and whether *Axel Springer* (completely or partly) owns them. Table 2 reports the list of domains analyzed and their classification according to their specialization and whether they are in the Top or Bottom 50% of their country in our sample.

The basic measure of our analysis is the domain’s *Daily Visits*. This variable is defined as the daily entries to a web domain from a different web domain or from the beginning of an empty browsing session, and expires after 30 minutes of inactivity. We also consider several engagement metrics. *Visit Duration* is the “session” length, which is the time that elapses between the first and the last page visit, on the analyzed domain. Note that according to this definition, the visit duration is equal to zero when the visitor only visits one page within the domain. On the other hand, during the duration of the visit, all the activities such as clicking on articles and images are considered page views. *Pages per Visit* is the daily page views divided by the daily visits of the domain. Finally, *Bounce Rate* is an indicator that shows the percentage of daily single-page sessions out of all daily sessions for the domain. This variable measures how often a consumer reaches a web page and then leaves without navigating to any other page. In such instances, the visitor stays in the domain for a very short period of time.

SimilarWeb also provides information about traffic sources. Traffic coming to a domain is defined as *Direct* when the visitors type the website’s URL into their browser. Traffic can also be the result of the consumers’ search activity. We consider *Search* to be the traffic coming from search engines such as Google, Bing, and Yahoo (not an exhaustive list). SimilarWeb similarly classifies the traffic coming from news aggregators such as Google News. *Referrals* is traffic coming

from links on other websites. Finally, *Social* is the traffic coming from social networks.²¹ Note that although the domain’s daily visits can be used to measure the overall impact of the shut down or change in the design of a news aggregator, the information about traffic sources is useful to study the relevance of the market-expansion and the substitution effects for the domain.

Table 3 shows the summary statistics for all the variables obtained from SimilarWeb, and Table 4 separates the data by country. See from Table 3 that the average site in our data receives 257,000 daily visits, and 43% come from direct searches, 29% from search engines, 14% from referrals of other pages, and 13% from social media pages. On average, a visit lasts around 10 minutes and users see four pages during that time. Table 4 compares the values of these variables across sites of different countries, and we see French domains have characteristics similar to those domains in Spain in our sample. This observation validates the fact that we use French domains as a control group for Spanish sites in our empirical analysis. In fact, pages per visit, visit duration and bounce rates in Spanish and French news outlets in our sample are very similar (3.6 to 3.8, 758 to 733, and 0.51 to 0.54, respectively), and so are the proportions of national, regional, business and sports. Note, however, that Spanish domains have on average fewer daily visits than French and German domains in our sample. This is explained by the fact that our sample contains more Spanish sites with lower rankings. This may also explain the larger share of direct visits among Spanish outlets in our sample. Therefore, controlling for domain type (national, regional, or specialized such as business and sports), and whether the domain is in the top or bottom 50% of the distribution of domains across countries, is important. Outlets in all three countries have similar percentages of national visits and, if anything, our variable definition means German outlets are less likely to appear among the top international outlets in our full sample.

4 Empirical Analysis of the Shutdown of Google News Spain

This section describes the empirical methodology we use to identify the “market contraction effect” and the “substitution effect” generated by the shutdown of Google News in the Spanish news market. We then show the results of implementing our empirical strategy, and relate them to the recent changes in the Spanish market.

4.1 Empirical Methodology

We use a difference-in-differences methodology to investigate both the impact of the shutdown of Google News in Spain and the opt-in policy of Google News in Germany. Our analysis of the Spanish case mainly uses three specifications. We introduce the first one to examine the total impact of the Google News shutdown on the newspapers’ daily visits and consumers’ engagement, therefore capturing the joint net effect of the market-expansion and the substitution effects. Although we

²¹SimilarWeb also collects information about traffic originated in emails and display ads, which we do not consider in our work, because their shares of activity are far smaller than those included in the paper, and therefore less relevant.

expect the shutdown of the news aggregator to reduce the outlet’s search visits, it could also generate a change in the consumers’ navigating behavior that increases the outlet’s direct visits. We identify the overall net effect of the shutdown on the domains with the following model:

$$\ln y_{ijt} = \theta SPAIN_i * After12/16/14?_t + \gamma_i + \delta_t + \eta_0 t_{jt} + \eta_1 t_{jt}^2 + u_{ijt},$$

where y_{ijt} is the outcome and dependent variable (e.g., daily visits to site i in country j in day t), $SPAIN_i$ is a dummy variable that takes the value of 1 if site i belongs to Spain, and 0 otherwise. $After12/16/14?_t$ is another dummy variable that takes the value of 1 if day t is after December 16, 2014, and 0 otherwise. We also use site and date fixed effects γ_i and δ_t , and introduce country-specific time trends t_{jt} and t_{jt}^2 . We assume the error term u_{ijt} to be iid and normally distributed as usual. The main objective of our analysis is to identify the impact of the shutdown in the daily visits, but we also analyze other engagement visitor metrics such as average pages viewed per visit per day and site, average duration of visits per day and site, and average bounce rate per day and site. We consider that visitors of news aggregators and news outlets’ direct visitors have different navigation habits, and thus we expect that after the shutdown, news outlets experienced differences in the number of pages per visit and the duration of the average visit.

Because we use site and date fixed effects in all specifications, the dummies $SPAIN_i$ and $After12/16/14?_t$ are not separately identified. Our parameter of interest is θ (the diff-in-diff parameter), and it captures the effect of Google News’ shutdown on Spanish news outlets. Therefore, the treatment group is all Spanish outlets in our sample, and the treatment period is the days after December 16, 2014. The control group is all French news outlets during the same period.

The second specification divides the impact of the treatment by week from the first week to seventh week and beyond after the shutdown of the Spanish edition of Google News. It is as follows:

$$\ln y_{ijt} = \sum_{k=1}^7 \theta_k SPAIN_i * kWeekthAfter12/16/14?_t + \gamma_i + \delta_t + \eta_0 t_{jt} + \eta_1 t_{jt}^2 + u_{ijt},$$

where θ_k is a parameter that captures both the market-expansion and the substitution effects on Spanish newspapers k weeks after the shutdown of Google News. All other parameters and variables remain the same from the explanation above. If anything, in some specifications, we introduce a dummy for the fourth week after the shutdown in French newspapers (the control group). We do so to control for the unanticipated increase in the number of visits French news outlets received due to the Charlie Hebdo terrorist attacks in Paris. These events took place in the fourth week after the shutdown of the Spanish edition of Google News.

Finally, to complete our analysis, we introduce a third specification that aims to identify separately the aggregators’ market-expansion and substitution effects. Specifically, we want to measure the effect of the shutdown on the news outlets’ search and direct visits, although we also examine the effect on other traffic sources. This new specification differs from the first one in that our data

are available at the month level and not the day level. Therefore, we consider the following:

$$\ln y_{ijm} = \theta SPAIN_i * After12/2014?_m + \gamma_i + \delta_m + \eta_0 m_{jm} + \eta_1 m_{jm}^2 + u_{ijm},$$

where the dependent variable y_{ijm} is now monthly percentages of visits coming from direct searches, search engines (google, yahoo, etc.), social network websites, or referrals from other sites. The percentages of visits coming from emails or display ads are minor, so we choose not to study them here. The dummy variable $After12/2014?_m$ takes the value of 1 if month m is from January 2015 and on, and 0 if in 2014. We also introduce country-specific time trends m_{jm} and m_{jm}^2 , as well as site and month fixed effects γ_i and δ_m . We again assume the error term u_{ijm} to be iid and normally distributed as usual.

4.2 Results

Before showing the results of our investigation for the Spanish case, we want to confirm that the shutdown of Google News in Spain did not affect the activity of Google News in France and in Germany. Figure 1 plots the log of daily visits for all three Google News webpages. Note that although the jump downward in visits to Google News in Spain is clear after December 16, 2014, the number of visits to Google News does not change in either France nor Germany.²²

Importantly, note that the shutdown of Google News in Spain is an isolated event and did not affect other news aggregators. Figure 2 compares the log of daily visits of Google News to meneame.net and kiosko.net, which are two alternative news aggregators. Note that the number of visits to these two domains did not change around the time of the shutdown of Google News. Hence, the event we study here is not a confounder of major changes in Google News everywhere (Figure 1) or of changes to news aggregators in Spain in particular (Figure 2).

We now describe the overall effect of the Google News shutdown. Table 5 uses specification (1) to analyze the effects of the shutdown on the number of daily visits using French outlets as control group. Column 1 shows that on average Spanish news outlets saw their number of daily visits decrease by 14%. See Figure 3 for a graphical representation of the negative effect of the shutdown on the number of daily visits per outlet. Column 2 allows this effect to vary by whether the Spanish news outlet is a national, regional, business, sports or Catalan language newspaper. The results show that the heterogeneity of the effect is large, and varies from no effect (business outlets), medium size effect (national and regional news outlets), and large effect (sports and Catalan language news outlets). This heterogeneity reflects the type of news stories that a news aggregator such as Google News offers and the likelihood that these types of news outlets will be indexed. Then, columns 3 and 4 allow the effect to vary by ranking (top and bottom 50% within our sample according to number of daily visits), and by the share of international visitors,

²²Despite the shutdown of Google News in Spain, the domain kept having a small number of visitors as its webpage explains Google’s shutdown decision. See Figure A1 in the appendix for the default webpage of news.google.es still posted.

respectively. Intuitively, we find that the impact of the Google News shutdown was larger in lower-ranked domains and domains with lower proportion of international visitors. These two findings suggest that news outlets with more loyal readers (top 50% ranked outlets), and with a larger share of international visitors are less likely to be affected by the presence of news aggregators.

Table 6 investigates further whether the specialization of news outlets affects their exposition to changes in the presence and activities of news aggregators. We run separate analysis for each classification using their respective French news outlets as control group. Note there are no French news outlets in Catalan language, and therefore we cannot run that particular regression. Columns 1 and 2 show that national and regional outlets are the most affected groups, with a 21% and 18% decline in daily visits, respectively. Columns 3 and 4 find that there is no effect on business and sports specialized news outlets. This is consistent with visits to news aggregators mainly driven by consumers of national and regional news, whereas consumers of sports and business news more likely to directly visit their preferred choice of outlet according to their idiosyncratic interests. Figure 4 shows these results graphically with negative effect for national and regional outlets, a positive change in sports outlets, and no impact on business outlets.

When we analyze the impact on top- and bottom-ranked outlets, we observe in columns 5 and 6 that the impact is very similar across types (14.7% reduction vs. 13.4%). Finally, columns 7 and 8 consider the Top and Bottom International 50% outlets in our sample. The results illustrate that our findings are also robust to splitting the sample according homogenizing the treatment and control groups. The effect of the shutdown is larger for outlets with lower shares of internationalization (15.8% reduction vs. 11.8%). This is intuitive because international visitors are less likely to be affected by the shutdown of Google News Spain as the news aggregator is less likely to be part of their daily choice of news consumption.

We are also interested in exploring how the impact of the shutdown of Google News in Spain evolved over time until reaching steady state. For this purpose, we run specification (2) in Table 7 for all the sample, and for each classification separately. Note that all specifications in Table 7 include an interaction term between the dummies “Fourth Week after 12/16/2017?” and “France,” because of the Charlie Hebdo terrorist attacks which dramatically increased the number of daily visits to French news outlets. Column 1 finds that the effect across all news outlets stabilizes around 13.8% seven weeks after the Google News shutdown with very little statistically significant impact during the first four weeks. Column 2 finds an immediate reduction of 7% and a long-term reduction of 19% in national news outlets. Columns 3 and 4 show no immediate impact and a long-term reduction of 16% and 12% in regional and business outlets, respectively, whereas column 5 shows no impact on sports outlets. These findings are consistent with graphical evidence in Figure 4. Columns 6 to 9 examine the long-term impact of the Google News shutdown for the top and bottom ranked outlets, and top and bottom international visits outlets. The long-term impact across these four classifications is similar and ranges from 12.5% reduction to 15% reduction. In summary, the long-term impact found in Table 7 is consistent with those results in Table 5.

Table 8 analyzes the impact of the shutdown of Google News on the consumers’ engagement metrics. Interestingly, we find a long-term impact in the duration of the visits, but not in the number

of pages consumers visited or the bounce rate (see columns 1, 3, and 5). When we decompose the effect by week, we find the pages per visit and the duration of visits decreased initially, but this effect vanished after the third week for the number of pages per visit. On the other hand, the bounce rate increased in the second and third weeks after the shutdown. To interpret these results, bear in mind that the shutdown could have changed the composition of the consumers who visit news outlets. First, news outlets lost the search visitors who previously arrived to their web site via Google News. Second, after the shutdown, some consumers could substitute the news aggregator for some news outlets. In principle, these two effects should have left news outlets with users that spend more time reading the news articles and visiting more pages, but this table leads us to very different conclusions. In fact, our findings suggest that Google News users were spending longer periods of time reading news and visiting more pages, potentially because of a better match to their interests thanks to the services of Google News. In the long term (seven weeks and beyond), we find no statistically significant differences in consumers' engagement metrics. This may be driven by the fact that most consumers on average ended up visiting the same news outlets than they did before the shutdown of Google News because they also visited other news aggregators than on average indexed the same type of news than Google News would have.

Finally, we directly investigate the relevance of the market-expansion and substitution effects after the shutdown. To do so, we study the impact of the shutdown on the outlets' traffic sources in Table 9 by running specification (3) on the log of the percentage of visits per month from search engines (columns 1 to 3), direct (4 to 6), referral (7 to 9), and social networks (10 to 12). We do not find significative changes in the percentage of visits received from referrals and social networks in columns 7 and 10, respectively. In contrast, we find that the percentage of search visits decreased in column 1, and the percentage of direct visits increased in column 4. While the former result is consistent with the market-expansion effect that we have found in all results of previous tables, the latter finding provides evidence of a substitution effect mainly taking place in national and top 50% outlets. This constitutes direct evidence of the coexistence of the market-expansion and substitution effects.

In summary, the results of the Spanish case reflect that the shutdown of Google News significantly reduced the number of daily visits to news outlets. Our findings suggest the reduction in the number of daily visits concentrated around search visits, and was slightly compensated by a minor increase in direct visits. According to our results, we conclude that news aggregators play a positive role in the news market by attracting additional visitors to news outlets. Our evidence suggests that the substitution effect reduces the direct visits to news outlets, but this effect is smaller than the market-expansion effect. The shutdown of Google News in Spain did not seem to cause a long-run effect on the consumers' engagement metrics and on the percentages of traffic sources.

4.3 Robustness Checks

The analysis presented so far has relied on the use of French news outlets as a control group to analyze the impact in the Spanish news outlets. A potential concern of this approach is that the different evolution of the French and the Spanish is not completely reflected in our empirical model. To address this concern, we next present two robustness checks that test the validity of our conclusions. In addition, we examine to whether the shutdown of Google News had an effect in the outlets' advertisement revenues and number of advertisers.

Charlie Hebdo Terrorist Attack. One potential problem of our analysis is that the Charlie Hebdo attack of January 7th 2015 and the events occurred the following days (police shot on January 8th and hostage situation at a Kosher supermarket near the Porte de Vincennes on January 9th) may be driving our results. Despite the fact that these events increased news consumption everywhere, they certainly increased news consumption the most in France. Our results of applying diff-in-diff methodology may then reflect the differential effect of the Charlie Hebdo attacks in ways that day fixed effects and a France-specific fourth-week dummy are not able to control for. For this reason, in Table 10 we repeat our main analysis dropping all observations from January 7th to January 21st (two weeks).

Our findings in Table 10 show qualitatively consistent results with those reported in Table 5. We find an overall drop in daily visits of 7.4%. This is mostly explained by a reduction of 9.6% in smaller news outlets (bottom 50% ranking), and a decrease of 11.2% in outlets with lower shares of international visitors.²³

Selection of the control group. A second related concern is our choice of French news outlets as control group for Spanish news outlets. To address this issue, we perform a synthetic control group analysis with our available data. Specifically, with this method we consider a weighted average of control news outlets (synthetic control) that is as similar as possible to the treated Spanish news outlets regarding the pre-treatment outcome variable. The benefit of building this synthetic control group is that the pre-shutdown characteristics of the Spanish news outlets can be better approximated by a combination of untreated news outlets than by an unweighted group of outlets (Abadie and Gardeazabal, 2003; Abadie et al. 2015).

In order to implement our analysis, we collapse our outlet-day specific data into group-week observations where we define our groups by country of origin (Spain, France, or Germany) and news outlet type (national, regional, business, or sports). This group classification allows us to create a synthetic control group for the average Spanish news outlet and 8 potential control groups. The program creates optimal weights using the number of daily visits (for only a subset of the pre-treatment period), share of international visits, engagement metrics, and shares of visit sources.

Figures 5A and 5B show the average daily visits for Spanish news outlets and its synthetic counterpart during the period analyzed. While in the first figure we consider only French news outlets, in the second we consider French and German. Note that prior to the Google News

²³The results do not qualitatively change when dropping one week or three weeks of observations after January 7th.

shutdown, the synthetic control closely tracks the average Spanish news outlets. After the Google News shutdown in Spain, the average Spanish outlet starts to diverge from the synthetic control unit.

Advertisement revenues. An additional approach to assess the shutdown of the Spanish edition of Google News is to consider its effects on advertisement revenues. While our findings of previous sections show that the shutdown reduced the volume of daily visits received by news outlets, we now examine whether the reduction of daily visits also shaped the advertisers' willingness to pay for advertising and their advertising strategies. In order to explore this additional effect, we use data at the domain level from Arce Media, a firm specialized in the collection and analysis of advertisement information in Spain.

The advertisement data contains information for 38 online domains (29 also have a printed edition). We consider several measures for our analysis. *Daily Advertisers* is the number of brands that promote their products in the online and printed versions of the news outlets. *Advertisement Intensity* is a variable that measures the intensity of the advertisers' campaigns. Arce Media visits several times per day the pages of the online outlets, and for each advertiser it calculates a ratio of the number of times that the advertiser appears in the news outlet. The sum of the ratios for all the advertisers in all the pages of the news outlets is the Advertisement Intensity. Finally, the *Daily Revenues* measures the estimated daily advertisements revenues obtained by the online and printed versions of the outlets. Arce Media calculates the variable *Daily Revenues* taking into account the number of advertisers, the intensity of the campaigns, the number of daily visits, and the prices charged by the news outlets for each type of advertisements.

The price paid by advertisers to news outlets usually depends on the cost per thousand impressions or per page views (CPM), which is the expense incurred for every thousand potential customers who view the advertisement. Taking this into account, we expect the shutdown of Google News to reduce the outlets' advertisement revenues, although increases in CPM or advertisement intensity may compensate for the initial decrease in daily visits. We follow the diff-in-diff methodology used along the paper, and provide evidence of different treatment and control groups in Table 11.

Columns 1 to 3 show results of comparing outcomes of the online and printed version of the 29 outlets that have presence in both media. As in our previous analysis, all the estimations use time trends, and day and outlet fixed effects. Despite the decrease in the number of daily visits identified before, we find that, after the shutdown, daily revenues, number of advertisers and revenue per advertiser grows in online outlets relative to those of print media. These findings may be driven by the fact that those outlets that are getting hurt the most in their print advertising are also those investing the most in their online advertising business. Moreover, this set of results is also consistent with the general view in the sector that in the last years advertising prices and revenues in printed outlets have decreased importantly regarding those of online outlets, and makes us question the convenience of using printed outlets as a control group for our analysis.²⁴

²⁴See, for example, <http://www.pwc.es/es/sala-prensa/notas-prensa/2016/caida-ingresos-prensa-espana.html>, and <http://prnoticias.com/prensa/prensa-pr/20157747-los-diarios-ingresos-digitales>.

Because of the heterogeneity across outlets in the evolution of their online advertising business, the second half of Table 11 (columns 4 to 8) focuses on the online editions, and separates those online outlets that are above and below the median of advertising revenues. Out of the 38 online outlets in the sample, those outlets above the median are more likely to have more developed online advertising resources and those below the median are more likely to be in an upward trajectory. All the specification in columns 4 to 8 include site and date fixed effects as well as group-specific time trends. The result in column 4 shows that after the Google News shutdown the daily revenues of above median outlets decreased significantly relative to those below the median, even after accounting for differences in levels and time trends. Columns 5 to 8 aim to understand the source of this large decrease in revenue by separately considering number of advertisers (column 5), advertising intensity (column 6), revenue per advertiser (column 7), and revenue per unit of advertising intensity (column 8) as a proxy of price per advertisement. Our findings show that there is no change in the number of advertisers, but there are statistically significant decreases in advertising intensity, revenue per advertiser and revenue per unit of advertising intensity.²⁵

These results indicate that the shutdown of Google News was not innocuous for news outlets. The reduction in daily visits affected their advertising revenues through the intensive margin of the advertising business. While we find no effect on the extensive margin, namely the number of advertisers did not change, the total willingness to pay of each advertiser was significantly reduced both in the advertising intensity and the willingness to pay per advertising unit.

5 The Opt-in Policy in Germany

This section studies the impact of Google’s opt-in policy in Germany. Two major differences exist between the German case and the previously analyzed Spanish case. First, after the introduction of the opt-in policy in Germany, Google News continued to index all news outlets, but the aggregator could only complement the links with long excerpts and images from outlets that had opted in. The different amount of information portrayed in the links of outlets that opted in and outlets that opted out implies that, in addition to the market-expansion and the substitution effects, a competition effect existed that could alter the daily visits the outlets received. Second, the treatment period we examine took place for a finite amount of time from October 23 to November 5, 2014.²⁶ After this period, Axel Springer and the other news outlets that initially stayed out decided to opt in.

5.1 Empirical Methodology

We analyze the impact of VG Media’s decision to not opt in to Google’s policy, by comparing German and French news outlets during the treatment period. Our first specification compares

²⁵We are able to separate number of advertisers and advertising intensity for the front page and other pages, but we do not find different results than those reported in the table. We find no changes in number of advertisers, and we find similar reductions in the advertising intensity in both front pages and other pages.

²⁶Notice that this event took place before the Google News shutdown in Spain and the Charlie Hebdo terrorist attacks in France. But for expositional reasons we present the Spanish case first.

German news outlets (treated group) with French news outlets (control group) before, during, and after the *de facto* opt-out period from October 24 to November 5, 2014 (treatment period). It is as follows:

$$\ln y_{ijt} = \theta \text{Germany}_i * \text{Opt_Out}_t + \gamma_i + \delta_t + \eta_0 t_{jt} + \eta_1 t_{jt}^2 + u_{ijt},$$

where all dependent variables are defined as in the previous section. Germany_i is a dummy that equals 1 if online newspaper i is German, and 0 otherwise. The dummy Opt_Out_t takes the value of 1 if day t is between October 24 and November 5, 2014, and 0 otherwise. On the other hand, γ_i , δ_t , and t_{jt} and t_{jt}^2 are site and date fixed effects, and group-specific time trends. We also consider other specifications in which the treated group is the outlets that opted out. Specifically, we run separate regressions for the members of the VG Media association and for the group of news outlets controlled by Axel Springer, which was one of the more active publishers in advocating for a change in German copyright law and was the first to announce its opt-out choice in October 2014. Finally, we also break the Opt_Out_t dummy into $1stWeekOptOut_t$ and $2ndWeekOptOut_t$ that take the value of 1 if day t falls in either the first or second week, respectively, of the full opt-out period. We assume the error term u_{ijt} to be iid and normally distributed as usual.

5.2 Results

We begin the analysis of the German case by estimating the effects of the opt-out decision on the outlets' daily visits. Columns 1 and 2 in Table 12 show that after the opt-out decision, the number of visits to German outlets increased by around 7% relative to French outlets. Columns 3 and 4 show, if anything, a reduction in the daily visits to the 16 outlets belonging to the VG Media association that opted out, although the estimates are not significant.

The rest of our specifications use German outlets only and focus on the news outlets that decided not to opt in. Columns 5 and 6 still show the existence of a negative but non-significant effect of the opt-out decision on the visits to the VG Media outlets relative to all other German outlets in our sample that did not belong to VG Media. Finally, columns 7 and 8 repeat the analysis, now focusing on the 10 outlets of our data set that Axel Springer controlled during this time.²⁷ We find a negative and significant reduction in daily visits of around 8% in Axel Springer outlets relative to all other German outlets in our data. This effect was stable across weeks during the treatment period. These results suggest the change in Google's linking policy created a market-contraction effect and possibly a competitive effect that diverted some of Axel Springer's visitors to other outlets that opted in.

Next, Table 13 performs a robustness check to determine whether the Axel Springer dummy variable may be capturing the impact of shocks on the demand for news that only affected specific outlet types during the treatment period. To do so, Table 13 repeats the analysis for the Axel

²⁷ Axel Springer is one of the largest publishing houses in Europe and one of the main contributors of outlets to our sample, the largest within the conglomerate VG Media with ten outlets out of sixteen. The six VG Media's outlets not considered in this group are regional outlets with a low domestic ranking, business and sports outlets.

Springer outlets in columns 7 and 8 of Table 12 while including as independent variables the interaction between the *OptOut* dummy variable and the specialization categories of outlets. Columns 1 and 2 take into account the classification of the news outlets according to their ranking (bottom 50%), and columns 3 and 4 according to their content (national, regional, business, and sports). Results in columns 3 and 4 show the effect of the opt-out decision was only significant in the second week of the treatment. However, the estimates in columns 1 and 2 offer similar insights to those in Table 12.

Table 14 examines the effects of Axel Springer’s opt-out decision on the engagement metrics, but we do not observe statistically significant changes in the engagement metrics of consumers during the opt-out period in Axel Springer outlets. Finally, we use the monthly data in traffic sources to analyze the relevance of the market-expansion and substitution effects for Axel Springer outlets. Table 15 considers the percentages of visits from search engines (columns 1 and 2), direct traffic (3 and 4), referrals (5 and 6) and social platforms (columns 7 and 8). Whereas column 1 shows an overall statistically non-significant 1% decrease in search traffic share, the results in column 2 indicate this effect mainly took place in the first week of the treatment period during October. Table 15 finds no evidence of substitution effect through change in the share of direct visits, but we do find evidence of an increase in 1% in the share of referral visits during the month of October. Note that our data and empirical methodology cannot disentangle whether this finding is due to a market-expansion effect (fewer visitors coming from Google News) or a competition effect (Axel Springer news stories were less attractive in Google News for consumers relative to other outlets because they exhibited shorter excerpts). Hence, similar to the Spanish case, our findings suggest news aggregators may increase the overall number of daily visits through a market-expansion effect.

6 Concluding Remarks

Amidst the growing importance of online platforms and aggregators, news aggregators are one of the most successful new players in the internet’s new era, quickly rising to occupy top positions in audience rankings. Yet, since their introduction, they have faced the opposition of news publishers, who consider aggregators as free-riders that resell their content. This situation has motivated the amendment of copyright laws in several countries, which have limited the use aggregators can make of the publishers’ content. Google News’ strategy in this new environment has been to avoid paying any link fee for the indexation of news stories. In Spain, after the government created a compulsory link fee, Google shut down its Spanish edition, sending a clear message to the publishers and governments of other countries that it would not accept paying for indexing news stories. In Germany, where the linking fee could be negotiated, Google has adopted an opt-in policy that, in practice, forces news outlets to waive any linking fee. Google is complementing this strategy with other actions, such as the creation of the Digital News Initiative, which gives support to European publishers for developing products that increase their revenue and traffic, stimulate innovation in digital news journalism, and promote training and academic research into journalism. These solutions do not seem to compensate traditional publishers, who in the last decade have seen how

the significant increase in their online visits and advertising revenues have not compensated for the reduction in advertisement revenues of printed newspapers. Consequently, the debate about regulating this market continues, and in the last year, European publishers have managed to move the discussion from the national arena to the European Union level. The European Union will have to decide whether to include this issue on its revision of the copyright legislation.

The goal of this research has been to examine the role of online news aggregators and their effect on the number of visits and the advertising revenues of news outlets. The economics literature has identified two potential types of effects news aggregators may have on news outlets. Whereas the market-expansion effect increases the search visits they receive by bringing new visitors that otherwise would not be aware of their presence, the substitution effect reduces their direct visits because they compete with news aggregators for the same audience. Our main contribution to the existing literature is to estimate the relative strength of these effects and shed light on how aggregators help news outlets expand their audience.

Our analysis of the shutdown of the Spanish edition of Google News shows a significant reduction in the audience of news outlets, driven by the reduction in search visitors. This result confirms that news aggregators are an important channel for attracting visitors to news outlets. Moreover, we observe the existence of a substitution effect that may have increased the number of direct visits to news outlets upon the shutdown of Google News Spain. Our findings suggest aggregators may overall increase the total number of daily visits to news outlets by allowing consumers to “rediscover” new sources of information. However, their impact on news outlets’ advertisement revenues remains an open question. While our analysis has not find an effect of the shutdown on the number of advertisers, it could have modified the advertisers’ total willingness to pay, reducing their advertising intensity and the price they offer per advertisement.

The examination of the German case is also useful because it shows that changes in the size of the excerpts the aggregators release modify the traffic news outlets receive. The decision to opt out significantly reduced the number of daily visits to Axel Springer outlets, and similar to the Spanish case, the reduction in direct and search visits drove this effect. Note, however, that in the German case, the reduction in search visits was also motivated by a competition effect, because Google only exhibited short excerpts for Axel Springer’s links, whereas it showed regular excerpts and images for the outlets that had “opted in” since the beginning of the policy.

Although our research answers the question of whether news aggregators are predominantly expanding the market audience of news outlets, we believe future research may expand our findings by further examining the impact of news aggregators on consumers’ engagement metrics. A full understanding of the effect of news aggregators on news consumption is essential for copyright policy design that benefits consumers and societies overall.

References

- [1] Abadie, A. and J. Gardeazabal (2003), The Economic Costs of Conflict: A Case Study of the Basque Country, *American Economic Review*, vol. 93(1), 113-132.
- [2] Abadie, A., A. Diamond and J. Hainmueller (2015) Comparative Politics and the Synthetic Control Method, *American Journal of Political Science*, vol. 59(2), 495-510.
- [3] Athey, S., E. Calvano and J. Gans (2011). “The Impact of Internet on Advertising Markets for News Media,” Discussion paper, Harvard University.
- [4] Athey, S. and Mobius, M. (2012). “The Impact of News Aggregators on Internet News Consumption: The Case of Localization,” Working Paper, Microsoft Research.
- [5] Athey, S., M. Mobius and J. Pal (2017). “The Impact of Aggregators on Internet News Consumption,” Working Paper, Microsoft Research.
- [6] Cage, J., N. Herve, and M.-L. Viaud. (2015). “The Production of Information in an Online World,” NET Institute Working Paper #15-05.
- [7] Calzada, J. and Ordóñez, G. (2012). “Competition in the News Industry: Fighting Aggregators with Versions and Links,” NET Institute Working Paper No. 12-22.
- [8] Chiou, L. and Tucker, C. (2017). “Content Aggregation by Platforms: The Case of the News Media,” NET Institute Working Paper No. 11-18. Forthcoming in *Journal of Economics and Management Strategy*.
- [9] CNMC (2014). “Propuesta referente a la modificacion del articulo 32.2 del proyecto de ley que modifica el texto refundido de la Ley de Propiedad Intelectual,” Comisión Nacional de los Mercados y la Competencia; PRO/CNMC/0002/14.
- [10] Dellarocas, C., Katona, Z. y Rand, W. M. (2013). “Media, Aggregators, and the Link Economy: Strategic Hyperlink Formation in Content Networks,” *Management Science*, 59(10): 2360-2379.
- [11] Dellarocas, C. J. Sutanto, M. Calin, E. Palme. (2015). “Attention Allocation in Information-Rich Environment: the case of Aggregator,” *Management Science*, forthcoming.
- [12] George, L. and Hogendorn, C. (2012). “Aggregators, search and the economics of new media institutions,” *Information Economics and Policy*, 24, issue 1: 40-51.
- [13] George, L. and Hogendorn, C. (2013) “Local News Online: Aggregators, Geo-Targeting and the Market for Local News,” Working paper.
- [14] Isbell, K. (2010). “The Rise of the News Aggregator: Legal Implications and Best Practices,” Berkman Center Research Publication No. 2010-10.

- [15] Jeon, D.S. and N. Nasr (2016). “News Aggregators and Competition Among Newspapers in the Internet,” *American Economic Journal: Microeconomics*, forthcoming.
- [16] Llobet, G. (2015). “Los agregadores de noticias y la prensa escrita en la era de Internet,” *Papeles de Economía Española*, N. 145, 172-179.
- [17] NERA (2015). “Impacto del Nuevo Artículo 32.2 de la Ley de Propiedad Intelectual Informe para la Asociación Española de Editoriales de Publicaciones Periódicas (AEEPP),” <http://www.aepp.com/pdf/InformeNera.pdf>
- [18] Roos, J., C.F. Mela and R. Shachar. (2015). “The Effect of Links and Excerpts on Internet News Consumption,” mimeo.
- [19] Rutt, J. (2011). “Aggregators and the News Industry: Charging for Access to Content,” Working Papers 11-19, NET Institute.
- [20] Sismeiro, C. and A. Mahmood (2017). “Competitive Versus Complementary Effects in Online Social Networks and News Consumption: A Natural Experiment,” Imperial College Working Paper.

Figure 1. Log Daily Visits Google News

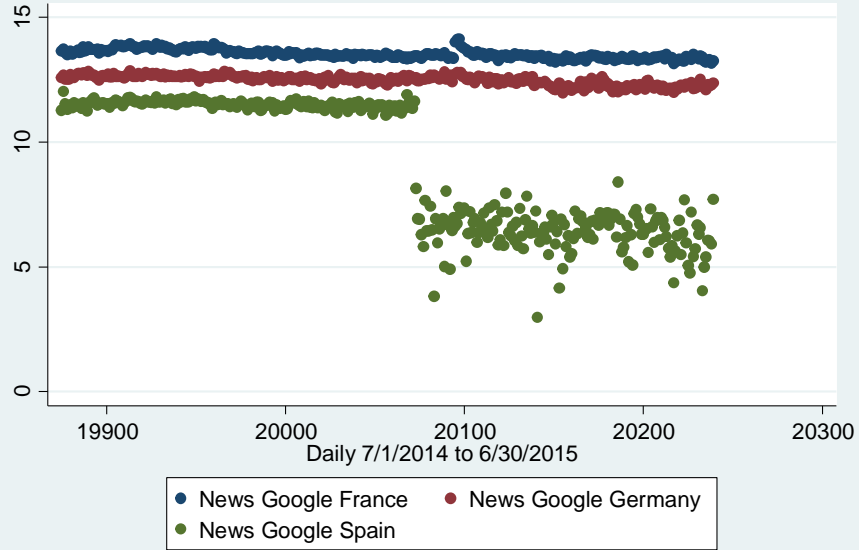


Figure 2. Log Daily Visits News Aggregators in Spain

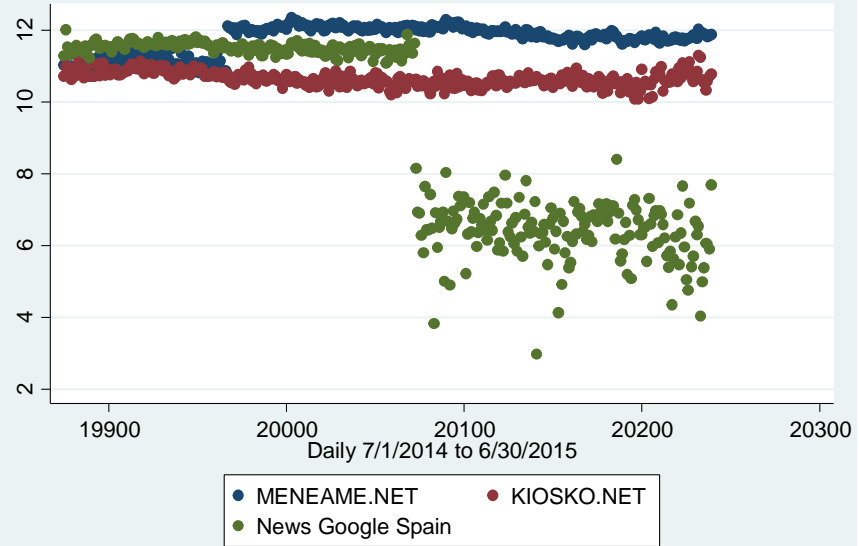


Figure 3. Full Sample

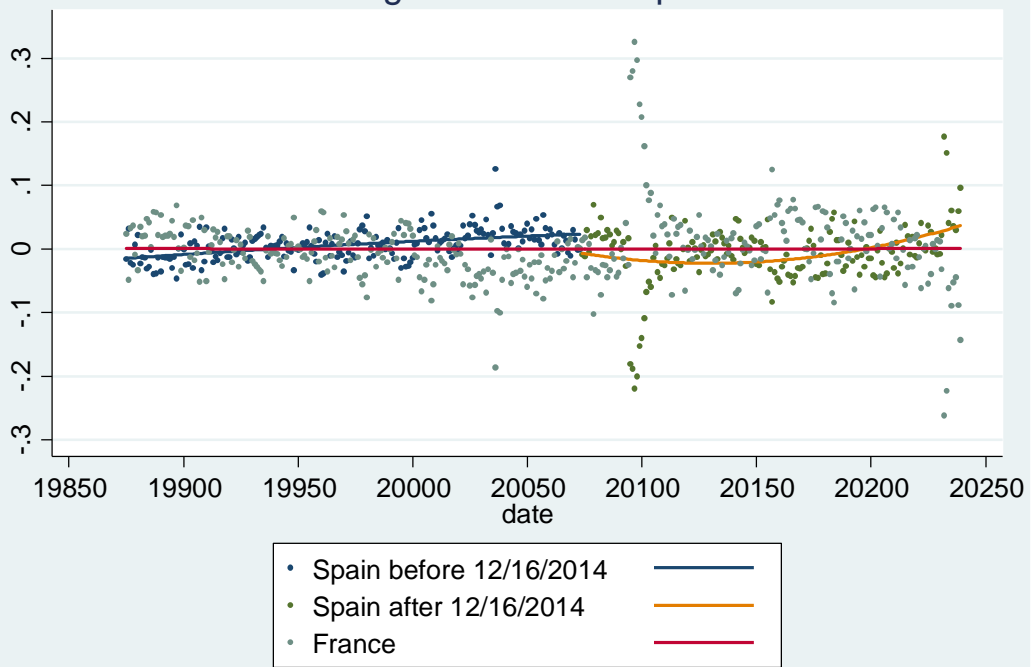


Figure 4. Diff-in-Diff by Newspaper Type

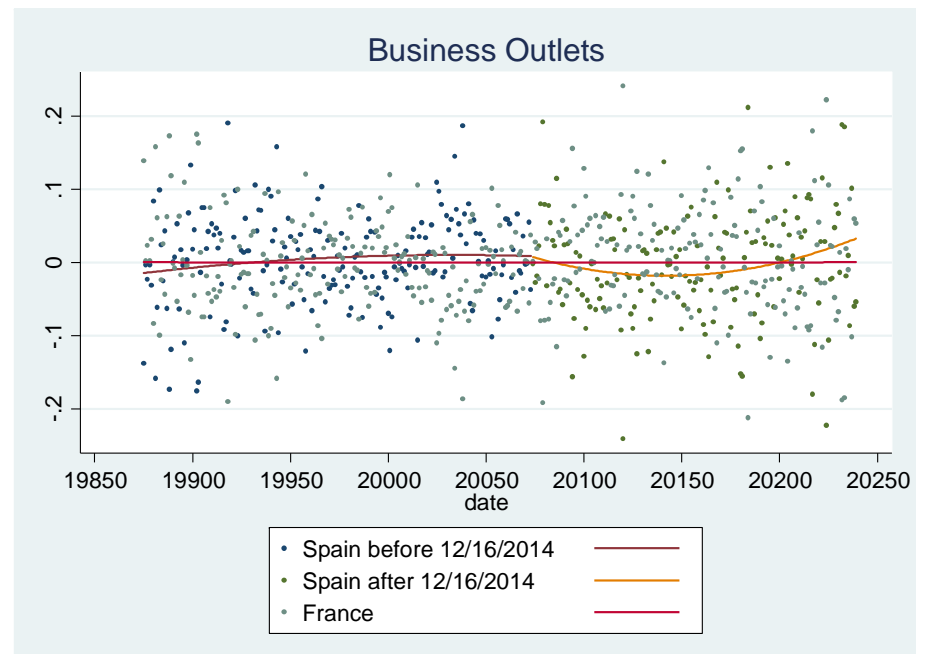
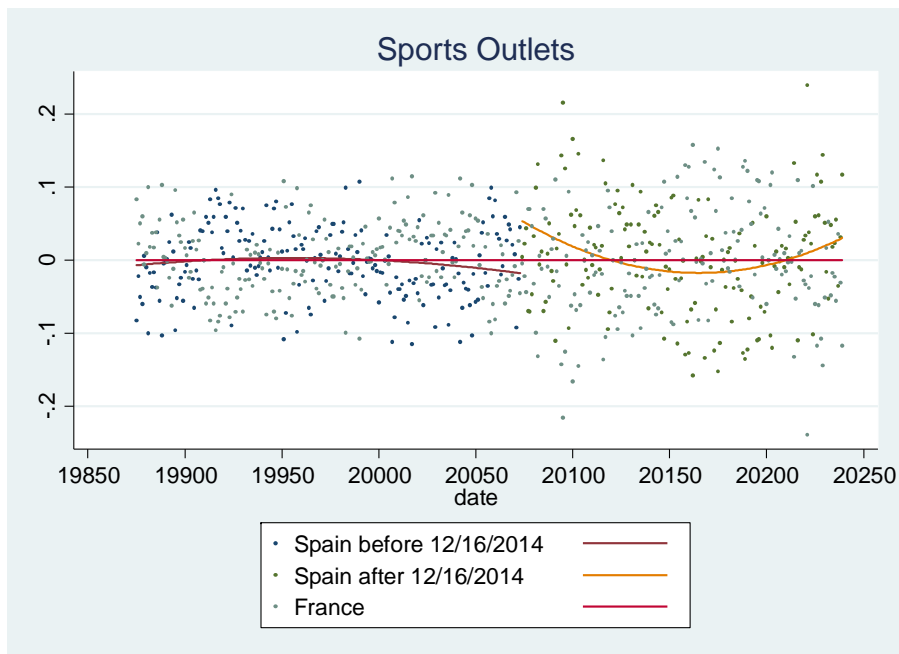
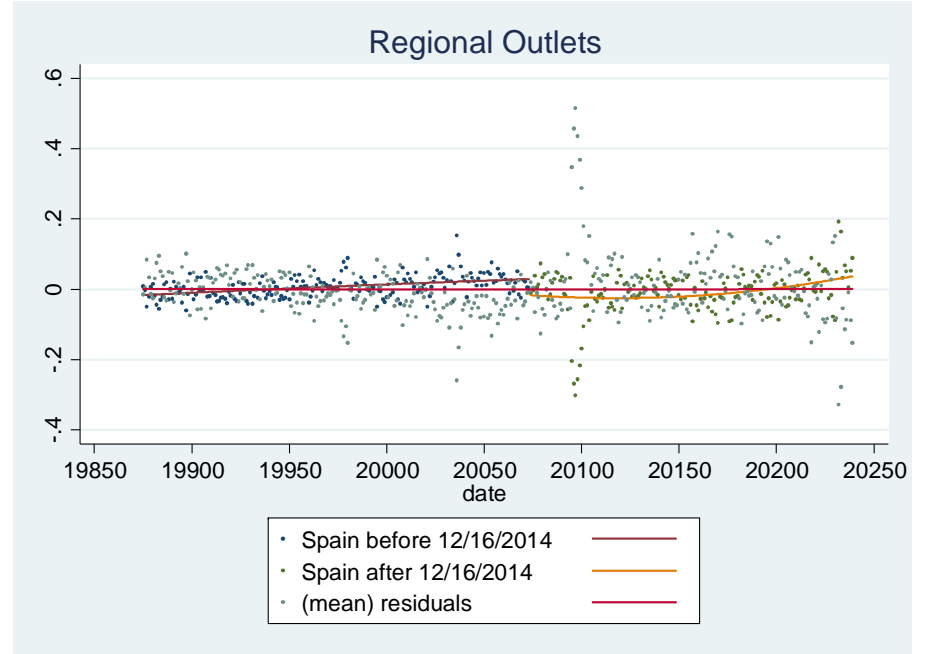
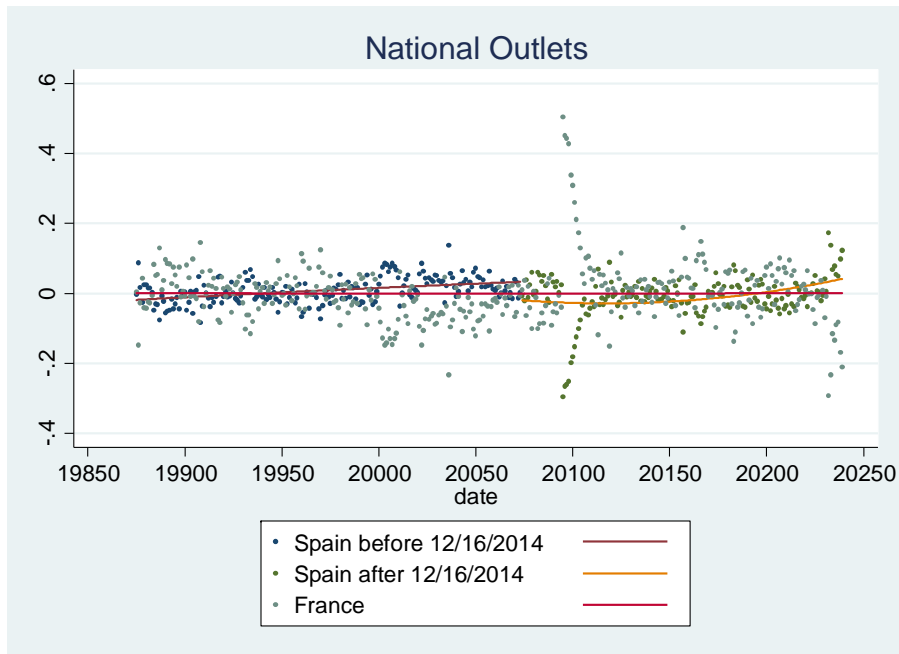


Figure 5A. Synthetic Control Unit Analysis with Germany and France News Outlets

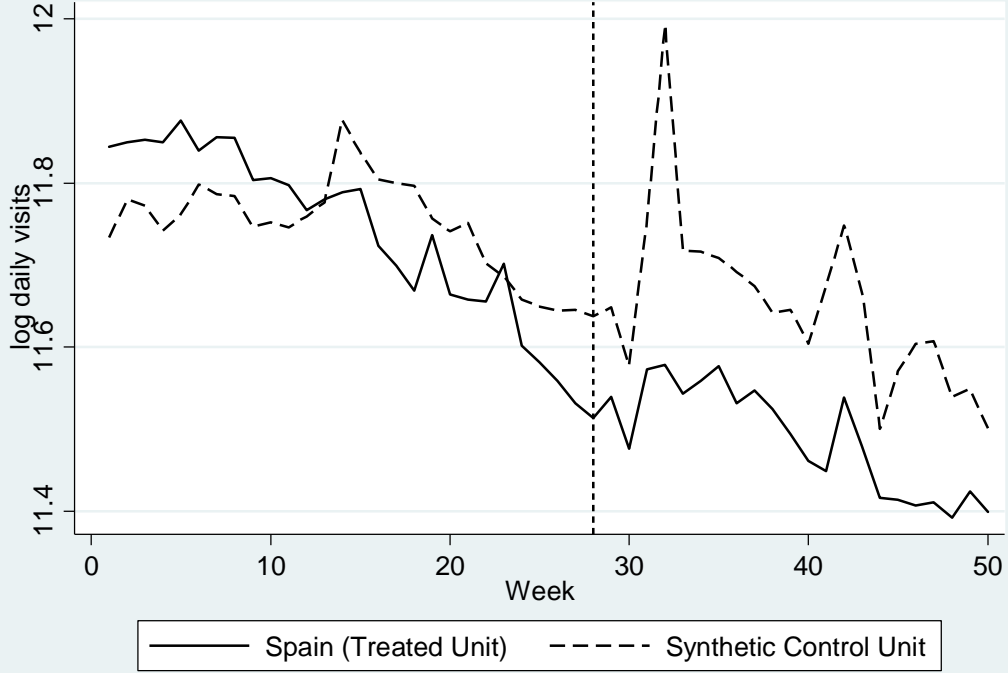


Figure 5B. Synthetic Control Unit Analysis with Only France News Outlets

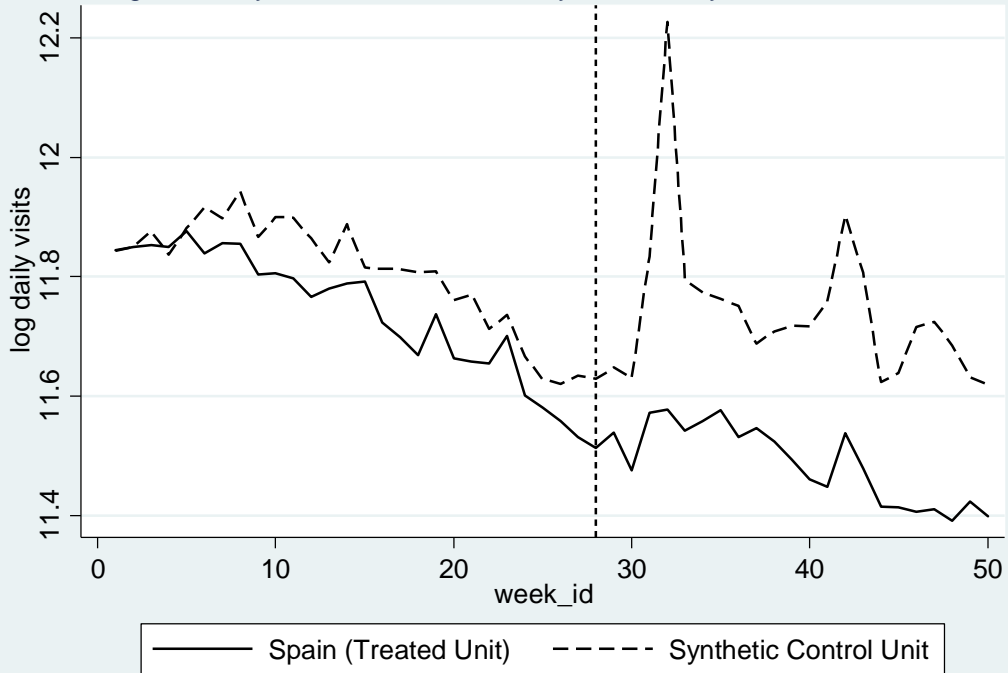


Table 1. Internet User Characteristics in European Union Countries

	Population 2016	Internet Users 2016	% Penetration	% Users EU	FB Members	% Users 16-24	% Users 25-34	% Users 35-44	% Users 45-54	% Users 55+
Spain	46,439,864	35,705,960	0.77	0.09	22,000,000	97%	96%	93%	85%	81%
France	66,132,169	55,429,382	0.84	0.14	32,000,000	98%	97%	91%	81%	77%
Germany	81,174,000	71,727,551	0.88	0.18	29,000,000	99%	94%	91%	77%	69%
Portugal	10,374,822	7,015,519	0.68	0.02	5,600,000	98%	91%	85%	77%	67%
Italy	60,795,612	37,668,961	0.62	0.09	28,000,000	97%	96%	87%	82%	76%

Data available from <http://www.internetworldstats.com/stats9.htm> and www.statista.com.

Table 2. List of Domains per Country

Spain		Germany		France	
Domain	Classification	Domain	Classification	Domain	Classification
20minutos.es	N, TOP	abendblatt.de	N,BOT	20minutes.fr	N,TOP
abc.es	N,TOP	autobild.de	B,BOT **	boursorama.com	B,TOP
ara.cat	N,C,BOT	berliner-zeitung.de	R,BOT **	capital.fr	B,BOT
as.com	N,TOP	bild.de	N,TOP **	challenges.fr	B,BOT
bolsamania.com	B,BOT	computerbild.de	B,TOP **	eurosport.fr	S,TOP
cincodias.com	B,TOP	derwesten.de	R,BOT **	footmercato.net	S,BOT
diaridegirona.cat	R,C,BOT	deutsche-wirtschafts-n.	B, BOT	huffingtonpost.fr	N,TOP
diariosur.es	R,BOT	express.de	R,BOT *	journaldesfemmes	N,TOP
diariovasco.com	R,BOT	faz.net	N,TOP	journaldunet.com	B,TOP
e-noticies.cat	R,BOT	finanzen.net	B,TOP **	ladepeche.fr	R,TOP
elcomercio.es	R,BOT	focus.de	N,TOP	laprovence.com	R,BOT
elconfidencial.com	N,TOP	fr-online.de	R,BOT	latribune.fr	N,BOT
elconfidencialdigital.com	N,BOT	handelsblatt.com	B,BOT *	lavoixdunord.fr	R,BOT
elcorreo.com	R,BOT	hna.de	R,BOT *	ledauphine.com	R,BOT
eldiario.es	N,TOP	huffingtonpost.de	N,TOP	lefigaro.fr	N,TOP
eldiariomontanes.es	R,BOT	kicker.de	S,TOP	lemonde.fr	N,TOP
eleconomista.es	B,TOP	ksta.de	R,BOT *	leparisien.fr	R,TOP
elmundo.es	N,TOP	manager-magazin.de	B,BOT	lepoint.fr	R,TOP
elpais.com	N,TOP	n24.de	N,TOP **	leprogres.fr	R,BOT
elperiodico.com	N,TOP	news.de	N,BOT	lequipe.fr	R,TOP
europapress.es	N,TOP	news.google.de	A,TOP	lesechos.fr	N,TOP
expansion.com	B,TOP	rp-online.de	R,BOT *	letelegramme.fr	R,BOT
heraldo.es	R,BOT	spiegel.de	N,TOP	liberation.fr	R,TOP
huffingtonpost.es	N,TOP	sport1.de	S,TOP *	metronews.fr	R,TOP
ideal.es	R,BOT	sportbild.bild.de	S,TOP **	midilibre.fr	R,BOT
invertia.com	B,BOT	stern.de	N,TOP	news.google.fr	R,TOP
kiosko.net	A,BOT	sueddeutsche.de	R,TOP	ouest-france.fr	R,TOP
lainformacion.com	N,BOT	tagesschau.de	N,TOP	rugbyrama.fr	R,BOT
larazon.es	N,BOT	tagesspiegel.de	N,BOT	sudouest.fr	R,BOT
lasprovincias.es	R,BOT	taz.de	N,BOT	tempsreel.nouvelobs	R,TOP
lavanguardia.com	N,TOP	transfermarkt.de	S,TOP **		
laverdad.es	R,BOT	welt.de	N,TOP **		
lavozdegacia.es	R,TOP	zeit.de	N,TOP		
levante-emv.com	R,BOT				
libertaddigital.com	N,TOP				
lne.es	R,BOT				
marca.com	S,TOP				
meneame.net	A,TOP				
mundodeportivo.com	S,TOP				
naciodigital.cat	R,C,BOT				
news.google.es	A,TOP				
periodistadigital.com	N,BOT				
publico.es	N,BOT				
sport.es	S,TOP				
vilaweb.cat	R,C,BOT				
vozapopuli.com	N,BOT				

Note: List of Domains in our sample per country.

Outlets classification: N= National; R= Regional; B= Business; S= Sports; A=Aggregator; C=Catalan language outlets; TOP=Top ranked outlets; BOT=Bottom ranked outlets. In the case of German outlets, * VG Media and ** Axel Springer.

Note Axel Springer outlets also belong to VG Media.

Table 3. Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Daily Visits	40,150	257867.00	331105.50	19.54	3314330
Pages Per Visit	40,150	3.61	2.30	1	130.97
Visit Duration	40,150	606.68	596.98	0	23641.75
Bounce Rate	40,150	0.53	0.11	0	1
Spain	40,150	0.43	0.49	0	1
Germany	40,150	0.30	0.46	0	1
France	40,150	0.27	0.45	0	1
After 12/16/2014?	40,150	0.45	0.50	0	1
National	40,150	0.36	0.48	0	1
Regional	40,150	0.34	0.47	0	1
Business	40,150	0.15	0.35	0	1
Sports	40,150	0.11	0.31	0	1
Catalan Language	40,150	0.05	0.23	0	1
News Aggregator	40,150	0.05	0.21	0	1
National Rank	39,055	273.38	377.13	11	3134
Top 50% Rank	39,055	0.50	0.50	0	1
% Search Visits	1,320	0.29	0.14	0	0.80
% Direct Visits	1,320	0.43	0.13	0.13	0.77
% Social Visits	1,320	0.13	0.10	0.0003	0.53
% Referrals Visits	1,320	0.14	0.09	0.04	0.72

Note that the variables % are monthly while other variables are daily.
Google News sites in France, Germany and Spain are not ranked here
because google.com is #1 everywhere.

Table 4. Summary Statistics per Country

Variable	<u>Spain (46 outlets, 17155 obs)</u>				<u>France (30 outlets, 10950 obs)</u>				<u>Germany (33 outlets, 12045 obs)</u>			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Daily Visits	204950.80	306133.10	19.54	2616957	312705	295028.50	18324.36	3314330	283379.90	381413.20	9413.13	2552446
Pages per Visit	3.60	2.11	1	50.42	3.82	1.59	1.43	18.46	3.45	2.98	1.17	130.97
Visit Duration	758.13	563.33	0	23641.75	733.79	646.94	75.18	4183.07	275.43	445.50	32.48	4376.13
Bounce Rate	0.51	0.11	0	1	0.54	0.10	0.19	0.80	0.54	0.10	0.25	0.82
After 12/16/2014?	0.45	0.50	0	1	0.45	0.50	0	1	0.45	0.50	0	1
National	0.36	0.48	0	1	0.33	0.47	0	1	0.39	0.49	0	1
Regional	0.38	0.49	0	1	0.33	0.47	0	1	0.27	0.45	0	1
Business	0.11	0.31	0	1	0.17	0.37	0	1	0.18	0.39	0	1
Sports	0.09	0.28	0	1	0.13	0.34	0	1	0.12	0.33	0	1
Catalan Language	0.13	0.33	0	1	0	0	0	0	0	0	0	0
News Aggregator	0.06	0.24	0	1	0.03	0.18	0	1	0.03	0.17	0	1
National Rank	321.33	510.77	11	3134	222.28	196.21	19	810	250.78	243.90	12	783
Top 50% Rank	0.43	0.50	0	1	0.59	0.49	0	1	0.50	0.50	0	1
National Visits	80.82	10.01	58.4	97.5	80.72	8.18	65	94.5	82.52	5.26	72.8	91.9
Top International	0.16	0.36	0	1	0.24	0.42	0	1	0.06	0.25	0	1
Top International 50%	0.47	0.50	0	1	0.58	0.49	0	1	0.35	0.48	0	1
% Search Visits	0.27	0.12	0	0.61	0.34	0.14	0.01	0.80	0.26	0.15	0.01	0.64
% Direct Visits	0.46	0.13	0.17	0.77	0.36	0.11	0.13	0.60	0.46	0.13	0.21	0.75
% Social Visits	0.13	0.10	0.00	0.53	0.16	0.09	0.004	0.47	0.12	0.09	0.002	0.51
% Referrals Visits	0.14	0.10	0.05	0.72	0.13	0.10	0.04	0.60	0.15	0.08	0.04	0.51

Note: The variables in % are monthly, whereas other variables are daily. Google News sites are not ranked here, because google.com is #1 everywhere.

Table 5. Diff-in-diff of Shutting Down Google News in Spain after 12/16/2014

	(1)	(2)	(3)	(4)
Dep Variable	ln(Daily Visits)			
After 12/16/2014?*Spain	-0.142*** (-4.617)			
After 12/16?*Spain*National		-0.111*** (-3.326)		
After 12/16?*Spain*Regional		-0.114*** (-3.072)		
After 12/16?*Spain*Business		-0.0387 (-0.765)		
After 12/16?*Spain*Sports		-0.196*** (-4.943)		
After 12/16?*Spain*Catalan		-0.259** (-2.625)		
After 12/16?*Spain*Top 50%			-0.113*** (-3.435)	
After 12/16?*Spain*Bottom 50%			-0.165*** (-4.083)	
After 12/16?*Spain*Top Int 50%				-0.0965*** (-3.212)
After 12/16?*Spain*Bottom Int 50%				-0.181*** (-4.351)
Constant	13.27*** (4.395)	13.27*** (4.394)	13.27*** (4.395)	13.27*** (4.395)
Site FE	YES	YES	YES	YES
Day FE	YES	YES	YES	YES
Time Trends	YES	YES	YES	YES
Control Group	France	France	France	France
Observations	26,280	26,280	26,280	26,280
R-squared	0.96	0.96	0.96	0.96

Robust t-statistics in parentheses clustered at the site level.

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

Table 6. Diff-in-diff by Online Newspaper Type, National Ranking and International Position

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep Variable	ln(Daily Visits)							
After 12/16?*Spain	-0.210*** (-6.217)	-0.183*** (-3.392)	0.0445 -0.748	-0.0691 (-1.228)	-0.147*** (-3.952)	-0.134*** (-2.838)	-0.118*** (-3.187)	-0.158*** (-3.412)
Constant	57,547*** (7.15)	933.9** (2.16)	46.87*** (5.08)	-3,409** (-2.455)	6.452 (1.53)	21.23*** (7.01)	4,165*** (3.83)	-12,327** (-2.299)
Site FE	YES	YES	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES	YES	YES
Time Trends	YES	YES	YES	YES	YES	YES	YES	YES
Control Group	France	France	France	France	France	France	France	France
Sample	National	Regional	Business	Sports	TOP 50%	BOTTOM 50%	TOP INT 50%	BOTTOM INT 50%
Observations	9,855	9,855	2,920	3,650	13,140	13,140	13,505	12,775
R-squared	0.97	0.95	0.97	0.93	0.93	0.89	0.95	0.93

Robust t-statistics in parentheses and clustered at the site level. *** p<0.01, ** p<0.05, * p<0.1.

Table 7. Diff-in-diff of Evolution of Effect by Online Newspaper Type and National Ranking Position

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep Variable	ln(Daily Visits)								
First Week After*Spain	-0.0291 (-1.084)	-0.0687* (-1.785)	-0.0281 (-0.552)	0.0483 (0.796)	-0.0126 (-0.216)	-0.0291 (-0.829)	-0.0317 (-0.827)	-0.0192 (-0.556)	-0.0272 (-0.680)
Second Week After*Spain	-0.0154 (-0.422)	-0.0264 (-0.560)	-0.054 (-0.898)	0.016 (0.144)	0.0473 (0.399)	0.0238 (0.509)	-0.0697 (-1.414)	0.00807 (0.169)	-0.0379 (-0.727)
Third Week After*Spain	-0.0563 (-1.476)	-0.0581 (-1.455)	-0.0772 (-1.000)	-0.105 (-1.000)	0.0411 (0.392)	-0.0311 (-0.831)	-0.0968 (-1.441)	-0.0129 (-0.328)	-0.097 (-1.514)
Fourth Week After*Spain	-0.0493 (-1.059)	-811.8*** (-5.903)	-0.159 (-1.590)	0.893*** (3.959)	-0.272** (-3.369)	0.505*** (3.635)	0.0527 (0.644)	11.81*** (3.880)	-26.95* (-1.832)
Fifth Week After*Spain	-0.206*** (-4.508)	-0.340*** (-5.171)	-0.215** (-2.758)	-0.143** (-3.084)	0.0988 (1.789)	-0.255*** (-4.423)	-0.134** (-2.171)	-0.199*** (-3.417)	-0.192*** (-2.987)
Sixth Week After*Spain	-0.123*** (-4.033)	-0.171*** (-3.969)	-0.150*** (-2.790)	-0.0675 (-1.100)	0.0207 (0.283)	-0.152*** (-4.770)	-0.0895* (-1.705)	-0.122*** (-3.290)	-0.107** (-2.286)
Seventh Week + After*Spain	-0.138*** (-4.254)	-0.189*** (-3.411)	-0.161** (-2.748)	-0.119** (-2.324)	0.0547 (0.729)	-0.150*** (-4.854)	-0.130** (-2.082)	-0.145*** (-3.657)	-0.125** (-2.455)
Fourth Week After*France	0.482*** (6.406)	-811.1*** (-5.903)	0.580*** (5.227)	1.037*** (4.032)	-0.447** (-3.405)	1.064*** (5.429)	0.511*** (6.423)	12.23*** (3.984)	-26.30* (-1.785)
Constant	13.36*** (4.431)	55,254*** (5.897)	753.9 (1.585)	-3,767** (-2.569)	45.85*** (5.681)	6.572 (1.556)	21.28*** (7.092)	4,470*** (3.882)	-10,141* (-1.826)
Site FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time Trends	YES	YES	YES	YES	YES	YES	YES	YES	YES
Control Group	France	France	France	France	France	France	France	France	France
Sample	All	National	Regional	Business	Sports	TOP 50%	BOTTOM 50%	TOP INT 50%	BOTTOM INT 50%
Observations	26,280	9,855	9,855	3,650	2,920	13,140	13,140	13,505	12,775
R-squared	0.96	0.968	0.949	0.927	0.967	0.932	0.886	0.954	0.933

Robust t-statistics in parentheses and clustered at the site level. *** p<0.01, ** p<0.05, * p<0.1.

Table 8. Diff-in-diff using Engagement Metrics as Dependent Variables

	(1)	(2)	(3)	(4)	(5)	(6)
Dep Variable	ln(1+pages per visit)		ln(1+visit duration)		ln(1+bounce rate)	
After 12/16/2014?*Spain	-0.035 (-0.852)		-0.101* (-1.884)		0.002 (0.441)	
First Week After*Spain		-0.074 (-1.629)		-0.127** (-2.475)		0.000 (-0.0602)
Second Week After*Spain		-0.088** (-2.022)		-0.160** (-2.537)		0.0097* -1.748
Third Week After*Spain		-0.099** (-2.329)		-0.216*** (-3.062)		0.0094* (1.723)
Fourth Week After*Spain		-0.027 (-0.511)		-0.184*** (-3.342)		-0.002 (-0.335)
Fifth Week After*Spain		0.020 (0.472)		-0.049 (-0.821)		-0.001 (-0.215)
Sixth Week After*Spain		-0.049 (-1.079)		-0.052 (-0.734)		0.002 (0.446)
Seventh Week + After*Spain		0.002 (0.034)		-0.053 (-0.693)		0.000 (0.059)
Fourth Week After*France		-0.093*** (-3.034)		-0.160** (-2.444)		0.006 (0.942)
Constant	-1.856 (-0.728)	-2.015 (-0.779)	-12.03*** (-2.924)	-12.24*** (-2.954)	-0.160 (-0.455)	-0.153 (-0.438)
Site FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Time Trends	YES	YES	YES	YES	YES	YES
Control Group	France	France	France	France	France	France
Observations	26,280	26,280	26,280	26,280	26,280	26,280
R-squared	0.65	0.65	0.86	0.86	0.80	0.80

Robust t-statistics in parentheses and clustered at the site level. *** p<0.01, ** p<0.05, * p<0.1.

Table 9. Diff-in-diff using % of visits per Source and Month

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dep Variable	ln(1 + % Search)			ln(1 + % Direct)			ln(1 + % Referral)			ln(1 + % Social)		
After 12/16/2014?*Spain	-0.0133** (-2.541)			0.0155*** (2.903)			-0.00119 (-0.304)			-0.00355 (-0.558)		
After 12/16?*Spain*National		-0.00843 (-1.151)			0.0159** (2.257)			0.00179 (0.357)			-0.0154** (-2.183)	
After 12/16?*Spain*Regional		-0.00953 (-1.043)			0.00759 (0.957)			0.00541 (1.114)			0.000166 (0.021)	
After 12/16?*Spain*Business		-0.0222*** (-3.117)			0.01 (0.889)			0.0150** (2.311)			0.00 (0.00)	
After 12/16?*Spain*Sports		-0.00547 (-0.421)			-0.0001 (-0.00591)			-0.0200** (-2.019)			0.00265 (0.339)	
After 12/16?*Spain*Catalan		0.00823 (0.601)			0.000162 (0.011)			-0.00418 (-0.813)			-0.0101 (-0.761)	
After 12/16?*Spain*Top 50%			-0.00691 (-0.987)			0.0135* (1.968)			0.00164 (0.305)			-0.0156** (-2.253)
After 12/16?*Spain*Bottom 50%			-0.0104 (-1.549)			0.0071 (1.005)			0.00126 (0.297)			-0.000637 (-0.103)
Constant	0.291*** (5.969)	0.291*** (5.960)	0.291*** (5.967)	0.218*** (5.215)	0.218*** (5.209)	0.218*** (5.218)	0.206*** (6.384)	0.206*** (6.387)	0.206*** (6.380)	0.103*** (3.107)	0.103*** (3.093)	0.103*** (3.096)
Site FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year/Month FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time Trends	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Control Group	France	France	France	France	France	France	France	France	France	France	France	France
Observations	864	864	864	864	864	864	864	864	864	864	864	864
R-squared	0.96	0.96	0.96	0.97	0.97	0.97	0.91	0.91	0.91	0.95	0.95	0.95

Note: Observations are monthly. Robust t-statistics in parentheses and clustered at the site level. *** p<0.01, ** p<0.05, * p<0.1.

Table 10. Shutting Down Google News Spain Deleting Charlie Hebdo Observations

	(1)	(2)	(3)	(4)
Dep Variable	ln(Daily Visits)			
After 12/16/2014?*Spain	-0.0748*** (-2.822)			
After 12/16?*Spain*National		-0.0463 (-1.527)		
After 12/16?*Spain*Regional		-0.047 (-1.377)		
After 12/16?*Spain*Business		0.0291 (0.589)		
After 12/16?*Spain*Sports		-0.134*** (-3.472)		
After 12/16?*Spain*Catalan		-0.248** (-2.445)		
After 12/16?*Spain*Top 50%			-0.0476 (-1.571)	
After 12/16?*Spain*Bottom 50%			-0.0963** (-2.633)	
After 12/16?*Spain*Top Int 50%				-0.032 (-1.182)
After 12/16?*Spain*Bottom Int 50%				-0.112*** (-2.923)
Constant	13.10*** (4.340)	13.10*** (4.339)	13.10*** (4.340)	13.10*** (4.340)
Site FE	YES	YES	YES	YES
Day FE	YES	YES	YES	YES
Time Trends	YES	YES	YES	YES
Control Group	France	France	France	France
Observations	25,200	25,200	25,200	25,200
R-squared	0.96	0.96	0.96	0.96

This table repeats the exercise in Table 5 dropping observations from January 7th to January 21st 2015 (both included).

Robust t-statistics in parentheses clustered at the site level.

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

Table 11. Advertising Effects of Google News Shutdown on Online News Outlets

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	ln(Daily Revenue)	ln(Daily Advertisers)	ln(Revenue/Advertiser)	ln(Daily Revenue)	ln(Daily Advertisers)	ln(Advertising Intensity)	ln(Revenue/Advertiser)	ln(Revenue/Advertising Intensity)
Online*Post 12/16/14	0.459*** (4.99)	0.139*** (3.39)	0.321*** (3.46)					
Above Median*Post 12/16/14				-0.477** (-2.39)	-0.024 (-0.37)	-0.102* (-1.90)	-0.453** (-2.03)	-0.376* (-1.87)
Treatment Group	Online News Outlets			Online News Outlets Above Median Advertising Revenue				
Control Group	Print Newspapers			Online News Outlets Below Median Advertising Revenue				
Site FE	YES	YES	YES	YES	YES	YES	YES	YES
Date FE	YES	YES	YES	YES	YES	YES	YES	YES
Group-Specific Time Trends	YES	YES	YES	YES	YES	YES	YES	YES
Observations	20,696	20,698	20,696	13,752	13,755	13,729	13,752	13,726
R-squared	0.78	0.77	0.88	0.83	0.80	0.91	0.82	0.86

This table shows two different set of diff-in-diff results using as dependent variables the log of daily advertising per site, the number of advertisers per site, the average revenues per advertiser, the average advertising intensity, and the revenues per advertising intensity. Advertising intensity is a measure of number of advertising per day in a site.

Columns (1) to (3) compare online newspapers to their own print newspapers. Note that online advertising was increasing and print advertising was decreasing during this time beyond what trends and square of trends are able to capture. Columns (4) to (8) compare online outlets with well-established advertising units to small and growing advertising outlets. We also control for different time trends across groups. All specifications include site, and date fixed effects.

Robust t-statistics in parentheses, clustered at the site level. *** p<0.01, ** p<0.05, * p<0.1.

Table 12. Diff-in-Diffs of Opting-out Period in Germany, October and November 2014

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep Variable	ln(Daily Visits)							
Opt Out*Germany	0.0710*** (3.038)		0.0829*** (2.721)	0.0869*** (2.933)				
1st Week Opt Out*Germany		0.0609** (2.298)						
2nd Week Opt Out*Germany		0.0721** (2.410)						
Opt Out*VGM			-0.02 (-0.698)		-0.0238 (-0.688)			
1st Week Opt Out*VGM				-0.0345 (-0.889)		-0.0266 (-0.653)		
2nd Week Opt Out*VGM				-0.0342 (-0.880)		-0.0369 (-0.864)		
Opt Out*Axel Springer							-0.0763** (-2.459)	
1st Week Opt Out*Axel Springer								-0.0814** (-2.074)
2nd Week Opt Out*Axel Springer								-0.0899** (-2.177)
Site FE	YES	YES	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES	YES	YES
Trends	YES	YES	YES	YES	YES	YES	YES	YES
Sample	France & Germany	France & Germany	France & Germany	France & Germany	Germany	Germany	Germany	Germany
Observations	22,265	22,265	22,265	22,265	11,680	11,680	11,680	11,680
R-squared	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95

Robust t-statistics are in parentheses and clustered at the site level. *** p<0.01, ** p<0.05, * p<0.1.

Table 13. Diff-in-Diffs Germany Opt-out Robustness Check

	(1)	(2)	(3)	(4)
Dep Variable	ln(Daily Visits)			
Opt Out*Axel Springer	-0.0830*** (-2.934)		-0.0645* (-1.964)	
1st Week Opt Out*Axel Springer		-0.0881** (-2.495)		-0.070 (-1.528)
2nd Week Opt Out*Axel Springer		-0.0966** (-2.388)		-0.0784** (-2.117)
Opt Out*BOT 50%	-0.046 (-1.329)	-0.046 (-1.335)		
Opt Out*National			0.068 (1.354)	0.066 (1.322)
Opt Out*Regional			0.114** (2.399)	0.113** (2.368)
Opt Out*Sports			0.111 (1.486)	0.111 (1.496)
Constant	11.08*** (174.8)	11.08*** (174.8)	11.08*** (176.2)	11.08*** (176.2)
Site FE	YES	YES	YES	YES
Day FE	YES	YES	YES	YES
Time Trends	YES	YES	YES	YES
Observations	11,680	11,680	11,680	11,680
R-squared	0.95	0.95	0.95	0.95

Robust t-statistics are in parentheses and clustered at the site level.

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

Table 14. Diff-in-Diffs using Engagement Metrics Germany Opt-out

	(1)	(2)	(3)	(4)	(5)	(6)
Dep Variable	ln(1+pages per visit)		ln(1+visit duration)		ln(1+bounce rate)	
Opt Out*Axel Springer	0.016 (0.938)		-0.033 (-0.984)		0.002 (0.347)	
1st Week Opt Out*Axel Springer		-0.006 (-0.231)		-0.052 (-1.510)		0.005 (0.986)
2nd Week Opt Out*Axel Springer		0.033 (1.374)		-0.023 (-0.572)		-0.001 (-0.189)
Constant	1.110*** (37.130)	1.110*** (37.140)	4.846*** (159.400)	4.846*** (159.400)	0.519*** (87.660)	0.519*** (87.640)
Site FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Time Trends	YES	YES	YES	YES	YES	YES
Observations	11,680	11,680	11,680	11,680	11,680	11,680
R-squared	0.74	0.74	0.79	0.79	0.87	0.87

Robust t-statistics are in parentheses and clustered at the site level. *** p<0.01, ** p<0.05, * p<0.1.

Table 15. Diff-in-Diffs of Monthly Shares per Visit Source Germany Opt-out: Search, Direct, Referral or Social

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep Variable	ln(1 + Search %)		ln(1 + Direct %)		ln(1 + Referral %)		ln(1 + Social %)	
Opt Out*Axel Springer	-0.008 (-1.630)		0.001 (0.453)		0.005 (1.634)		0.002 (0.354)	
October*Axel Springer		-0.0118* (-1.909)		0.005 (1.272)		0.0107* (1.719)		0.002 (0.382)
November*Axel Springer		-0.003 (-0.661)		-0.002 (-0.341)		-0.001 (-0.247)		0.002 (0.239)
Constant	0.407*** (61.060)	0.406*** (61.090)	0.311*** (54.610)	0.311*** (54.540)	0.0754*** (7.284)	0.0756*** (7.344)	0.0632*** (8.838)	0.0632*** (8.889)
Site FE	YES	YES	YES	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES	YES	YES	YES
Time Trends	YES	YES	YES	YES	YES	YES	YES	YES
Observations	384	384	384	384	384	384	384	384
R-squared	0.98	0.98	0.97	0.97	0.88	0.88	0.94	0.94

Robust t-statistics are in parentheses and clustered at the site level. *** p<0.01, ** p<0.05, * p<0.1