

# **Search Engines and News Outlet's Traffic: Evidence from Google's Core Updates**

Joan Calzada, Universitat de Barcelona

Nestor Duch-Brown, JRC European Commission

Ricard Gil, Queen's University

EEA-ESEM Barcelona Conference

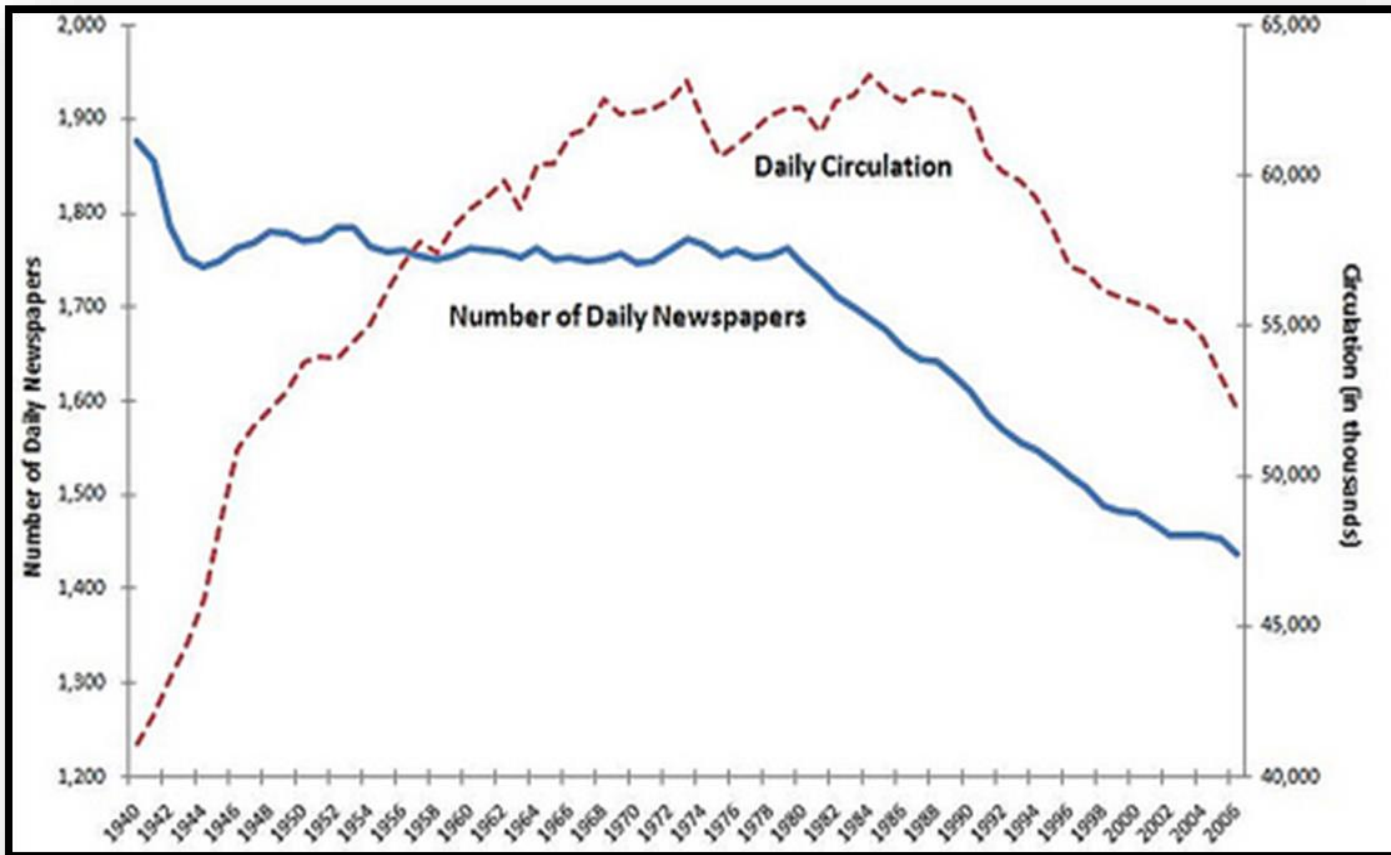
August, 2023

# Digital Technology & Media Markets

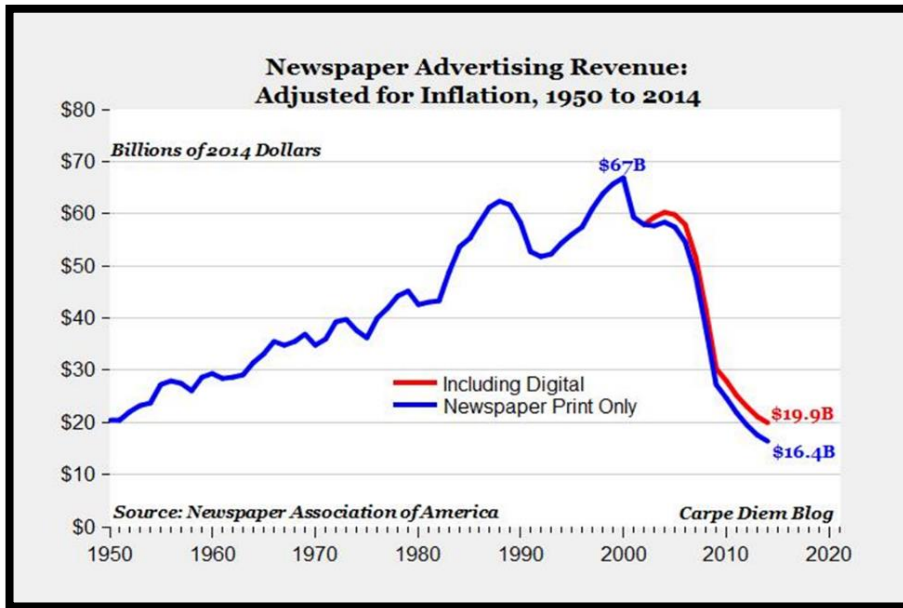
- Digital Revolution has changed the way we access and consume media content such as music, movies, and books (Waldfoegel, 2017).
  - It has increased the amount and variety of “niche” contents available to the consumers. **The long tail has become thicker.**
  - Search engines, aggregators and recommender systems are now necessary to navigate through the information clutter. **They can create “superstar” firms** that dominate the market.
- **Which effect dominates in the news markets?**

# Concentration in Media Markets

Long-term trend of concentration in media ... newspapers, TVs, radios either disappearing or consolidating

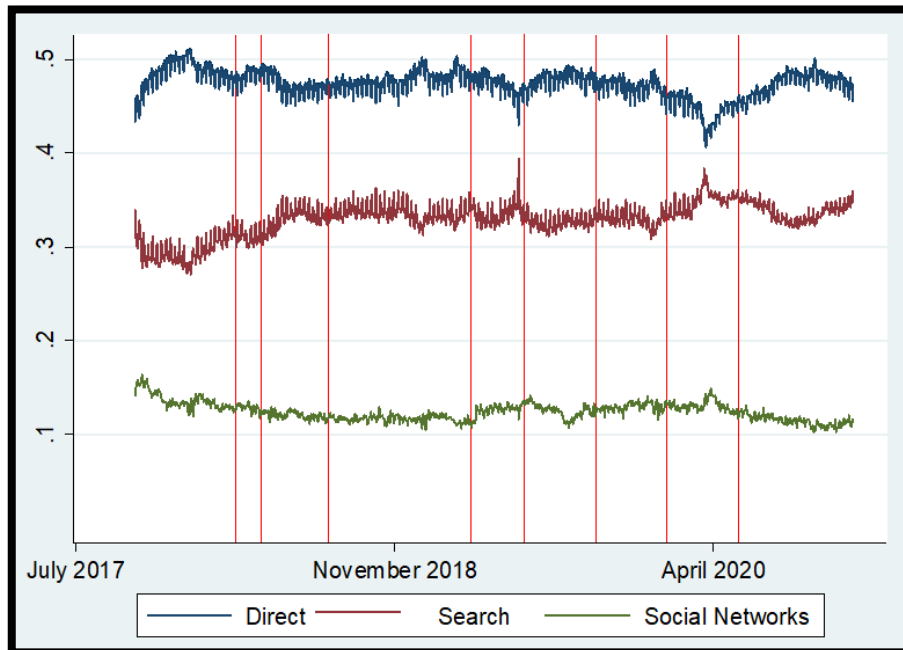


# Explanations for this consolidation?



- Digitalization:
  - Change from printed to digital newspapers
  - Change from traditional to targeted advertising

## Share of Desktop Direct, Search and Social Networks Daily Visit (Oct 2017 – Dec 2020)



- New business models:
  - News aggregators
  - 1/3 traffic from Search Engines
  - Tech platforms have a dominant position advertising

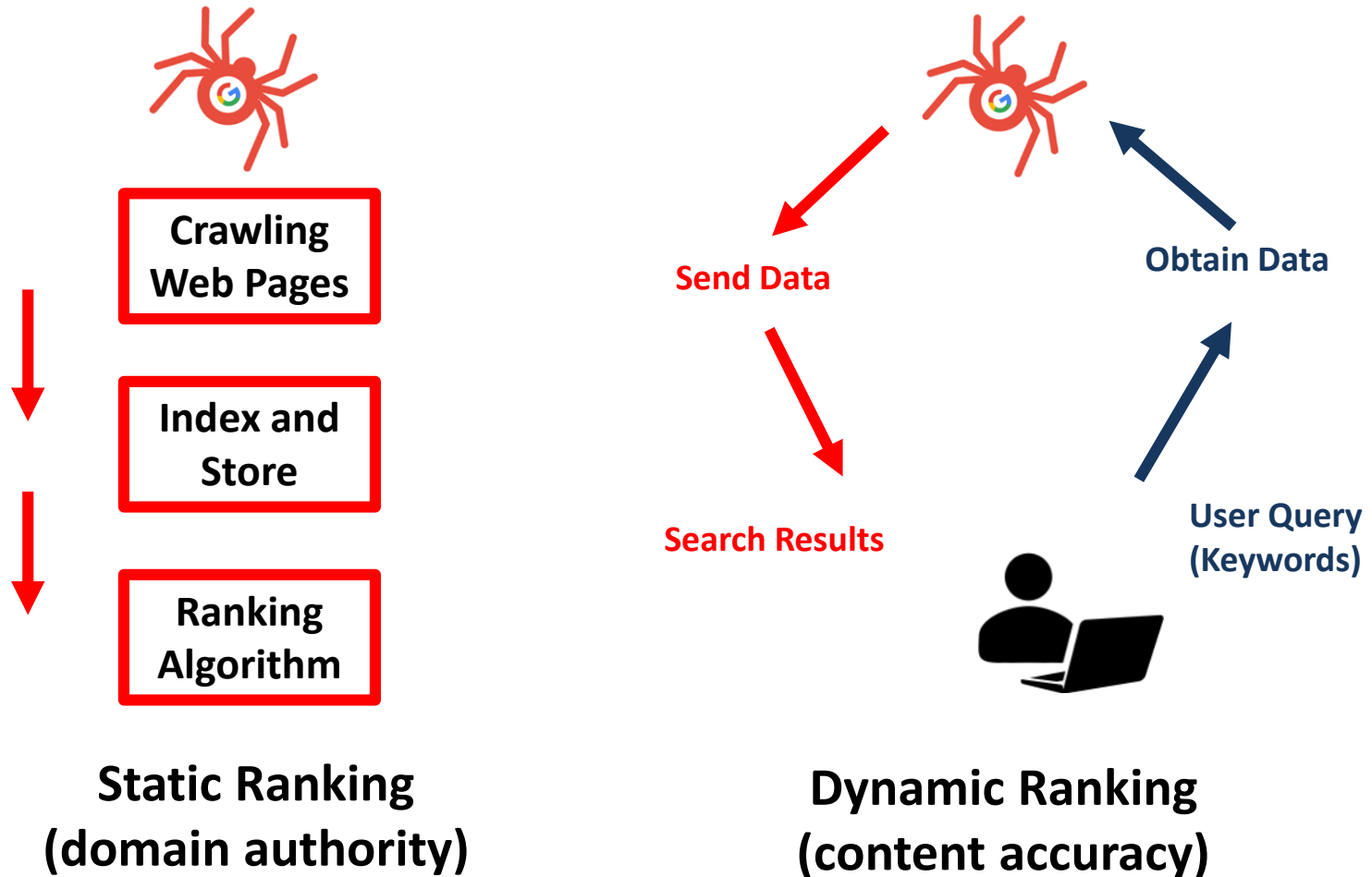
# Why should we care?

- Industrial Organization side:  
Raising concentration (i) increases cost of advertising which are passed through to the consumers; and (ii) affects the quality of the journalistic contents.
- Political Economy side:  
Raising concentration (i) diminishes plurality of media outlets; and (ii) affects development of democratic institutions.

# Our Goal in this Paper

- Our research questions:
  - How do Google's indexation affect news outlets' visits?
  - Does Google modify the concentration of media markets?
- Empirical strategy:
  - Daily desktop and mobile traffic for 606 outlets in 15 EU countries
  - Use Google core updates as shifter of keyword indexation:
    - Effect of # keywords in top 10 positions in search visits
    - Effect of core updates on market concentration
- Main takeaway: Google core updates decrease outlets' visits and heterogenous impact on market concentration.

# Google Search: crawling, indexing, ranking



The traffic of news sites depends on how Google's weights the two things

## Top of the First Search Results Page

Google search results for 'energy prices'. The search bar shows 'energy prices' and the results are categorized under 'Crude Oil & Natural Gas'. A table lists various energy indices and their prices.

Index	Units	Price
CL1:COM WTI Crude Oil (Nymex)	USD/bbl.	103.38
CO1:COM Brent Crude (ICE)	USD/bbl.	108.07
CP1:COM Crude Oil (Tokyo)	JPY/kl	70,730.00
NG1:COM Natural Gas (Nymex)	USD/MMBTu	5.78

Below the table, there are several search results with snippets:

- <https://www.bloomberg.com> › energy  
**Energy - Bloomberg**
- <https://www.iea.org> › reports › ener...  
**Energy Prices: Overview – Analysis - IEA**  
This database includes annual energy prices data for more than 100 non-OECD countries, for gasoline, automotive diesel, natural gas, electricity and other ...
- <https://www.ofgem.gov.uk> › check...  
**Check if the energy price cap affects you | Ofgem**  
hace 11 horas — The price cap limits the rates a supplier can charge for their default tariffs . These include the standing charge and price for each kWh of electricity and gas ...
- <https://ec.europa.eu> › index.php  
**Electricity price statistics - European Commission**  
28 ene 2022 — EU statistics on electricity prices for households and non-households analyse their evolution and the differences between countries.

## Bottom of the First Search Results Page

### Otras preguntas de los usuarios

Will energy prices go up UK?

What is the energy price cap UK per kWh?

Who is the cheapest energy supplier per kWh?

What is the cheapest electricity tariff in the UK?

Enviar comentarios

<https://www.independent.co.uk> › c...  
Traducir esta página

Customers urged to check meters before 54% jump in energy ...

<https://www.bbc.co.uk> › news › bu...  
Traducir esta página

Why are energy bills so high and when should I take a meter ...

<https://www.bbc.co.uk> › news › uk...  
Traducir esta página

Energy prices: Why are power bills going up?

<https://www.britishgas.co.uk> › energy  
Traducir esta página

Gas and electricity | Energy prices and tariffs

Energy price cap update. After the recent announcement from Ofgem, our tariff rates will be rising in line with the price cap to take effect from ...

### Videos



Why are energy prices rising? - BBC News

YouTube · BBC News  
30 ene 2022

It is important to know the number of keywords that news outlets have in the first results page



# Google's core algorithm updates

- Google Search introduces many changes in its algorithms and systems every day aiming to improve its service to users.
- A few times per year Google makes large “**core updates**” that have significant effects in the search market.
  - Announced by Google (the same day, or one day before...).
  - Global: they affect all Google search regions and languages; not focused on specific types of search queries or web characteristics.
  - Web sites are heterogeneous: They are affected in different ways by the updates.

**We consider 9 core updates in 2018-2020 (moz.com)**



Google SearchLiaison   
@searchliaison



Later today, we are releasing a broad core algorithm update, as we do several times per year. It is called the May 2020 Core Update. Our guidance about such updates remains as we've covered before. Please see this blog post for more about that:

	What webmasters should know about Google's core up... Official news on crawling and indexing sites for the Google index <a href="https://webmasters.googleblog.com">webmasters.googleblog.com</a>
--	---

8:29 PM · May 4, 2020 · Twitter Web App

1,660 Retweets 908 Quote Tweets 2,337 Likes



Google SearchLiaison   
@searchliaison · May 4, 2020



Replying to @searchliaison

The May 2020 Core Update is now rolling out live. As is typical with these updates, it will typically take about one to two weeks to fully roll out.



61



414



506



Google SearchLiaison   
@searchliaison · May 18, 2020



The May 2020 Core Update rollout is complete.



131



526



516



Niyazi @Niyazi\_dortkol · May 5, 2020



Replying to @searchliaison

@dannysullivan I think this core update is being done to further increase the authority of major news sites. 😊

Because small and medium-sized sites that do quality publications are no longer able to stand out.



2



1



11



GOOGLE AMP BLOWS @CecilBMoor\_Gucc · May 7, 2020



That maybe the outcome as smaller sites typically do not use AMP.

This was just a money grab to keep pushing AMP, of which you

**Table 1A. Google's confirmed core updates**

<p><b>December 2020 Core Update</b> (December 3, 2020)</p>	<p><b>Google's Confirmation:</b> <a href="https://twitter.com/searchliaison/status/1334521448074006530">https://twitter.com/searchliaison/status/1334521448074006530</a></p> <p>Some industry experts explain that this was of the more impactful algorithm adjustments to hit the SERP over the past year or so.</p>
<p><b>May 2020 Core Update</b> (May 4, 2020)</p>	<p><b>Google's Confirmation:</b> <a href="https://twitter.com/searchliaison/status/1257376879172038656">https://twitter.com/searchliaison/status/1257376879172038656</a></p> <p>According to <a href="#">Moz</a>, this update was the second-highest Core Update after the August 2018 "Medic" update.<sup>44</sup></p>
<p><b>January 2020 Core Update</b> (January 13, 2020)</p>	<p><b>Google's Confirmation:</b> <a href="https://twitter.com/searchliaison/status/1216752087515586560">https://twitter.com/searchliaison/status/1216752087515586560</a></p> <p><a href="#">Moz</a> considers that the effects of this core update were considered smaller than the August 2018 "Medic" core update.</p>
<p><b>September 2019 Core Update</b> (September 24, 2019)</p>	<p><b>Google's Confirmation:</b> <a href="https://twitter.com/searchliaison/status/117647392383225221">https://twitter.com/searchliaison/status/117647392383225221</a></p> <p>This update focused on improvements in the content quality in the SERPs. For the second time, Google pre-announced a core algorithm update "in advance".</p>
<p><b>June 2019 Core Update</b> (June 3, 2019)</p>	<p><b>Google's Confirmation:</b> <a href="https://twitter.com/searchliaison/status/1135275028834947073">https://twitter.com/searchliaison/status/1135275028834947073</a></p> <p>This is considered as one of the Google's most important core updates. Moreover, for the first time in the history of core updates, Google announced this update 24 hours ahead of time on Google Search Liaison Twitter channel. According to <a href="#">Moz</a>, the impact was smaller than the August "Medic" update.<sup>45</sup></p>
<p><b>March 2019 Core Update</b> (March 12, 2019)</p>	<p><b>Google's Confirmation:</b> <a href="https://twitter.com/searchliaison/status/1105842166788587520">https://twitter.com/searchliaison/status/1105842166788587520</a></p> <p>Google stated that this was the third major core update since they began using that label. The update generated ranking shifts for keywords related to health and other sensitive topics. The update affected search queries that are covered by the acronym E-A-T (Expertise, Authoritativeness, and Trust).</p>
<p><b>Medic Core Update</b> (August 1, 2018)</p>	<p><b>Google's Confirmation:</b> <a href="https://twitter.com/searchliaison/status/1024691872025833472">https://twitter.com/searchliaison/status/1024691872025833472</a></p> <p>Expert report large impact in search results, <a href="#">specially</a> for health and wellness.</p>
<p><b>Unnamed Core Update</b> (April 17, 2018)</p>	<p><b>Google's Confirmation:</b> <a href="https://twitter.com/searchliaison/status/987397051997663232">https://twitter.com/searchliaison/status/987397051997663232</a></p> <p>According to experts, a heavy algorithm flux that peaked on April 17 and continued for over a week. Google later confirmed a "core" update</p>
<p><b>Brackets Core Update</b> (March 8, 2018)</p>	<p><b>Google's confirmation:</b> <a href="https://twitter.com/searchliaison/status/973241540486164480">https://twitter.com/searchliaison/status/973241540486164480</a></p> <p>Google confirmed a "core" update on March 7th, but volatility spiked as early as March 4th, with a second spike on March 8th, and continued for almost two weeks. The "Brackets" name was coined by Glenn Gabe.</p>

**BIG**

**BIG**

**BIG**

# Data Description

- Data at the domain level from SimilarWeb (web measurement company)
  - From October 1, 2017 to December 31, 2020.
  - Information for 606 news outlets, from 15 countries.
  - Main analysis with (1) **Daily Search Desktop** visits; (2) Daily Total Desktop visits; and (3) Daily Total Mobile visits.
  - Also data on direct traffic, social networks, display ads, and paid search.
- News outlets are categorized according to:
  - Specialization (National, Regional, Business, Sports, TV/Radio)
  - National rank (Top 50% and Bottom 50%)
  - Percentage of visits from other countries (Top & Bottom International 50%)
  - Percentage of search visits from GS (Top & Bottom Google 50%)
- Information on # keywords and on domain rank from Ahrefs (SEO software firm)

# Empirical methodology

- We could analyze the relationship between the **search visits** of news outlets and the **#keywords** they have in top 10 search results. Potential endogeneity.
- We consider an IV estimation such that the exogenous variation (instrument) is **Google's core updates** driving #keywords, which drives search visits
- We estimate the following IV model in first-differences of t and t-4:

$$\Delta \ln[visits_{it}] = \alpha + \beta \Delta \ln[Words Top_{it}] + \gamma \Delta X_{it} + \Delta \delta_t + \Delta u_{it}$$

where the first stage is

$$\Delta \ln[Words Top_{it}] = \theta_0 + \theta_1 CoreUpdatePlus7_{it} + \theta_2 \Delta X_{it} + \Delta \omega_{it}$$

Regressions control for first differences in direct traffic, day-of-week, week, year fixed effects. Standard errors are clustered at the news outlet level.

Table 3. First Differences OLS Regressions of Search Visits, Total Visits and Mobile Visits on the Number of Key Words and Google Core Updates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	OLS	OLS	1st Stage - IV	2nd Stage -IV	2nd Stage -IV	2nd Stage -IV
Dependent Variable	$\Delta \ln(\text{Search Visits})$ t-4	$\Delta \ln(\text{Desktop Visits})$ t-4	$\Delta \ln(\text{Mobile Visits})$ t-4	$\Delta \ln(\text{Words top 10})$ t-4	$\Delta \ln(\text{Search Visits})$ t-4	$\Delta \ln(\text{Desktop Visits})$ t-4	$\Delta \ln(\text{Mobile Visits})$ t-4
$\Delta \ln(\text{Words top 10})$ t-4	0.0869*** (0.0181)	0.0453*** (0.0122)	0.1270*** (0.0195)		6.6376*** (1.1211)	3.7405*** (0.7467)	4.1166*** (1.4006)
"Big" Google Core Update t+7				-0.0008* (0.0004)			
"Not Big" Google Core Update t+7				-0.0025*** (0.0003)			
$\Delta \ln(\text{Desktop Direct Visits})$ t-4	0.3444*** (0.0257)	0.6021*** (0.0242)	0.3914*** (0.0263)	0.0001 (0.0001)	0.3436*** (0.0258)	0.6017*** (0.0242)	0.3910*** (0.0264)
Constant	0.0332*** (0.0072)	-0.0001 (0.0050)	0.0997*** (0.0081)	0.0075*** (0.0008)	-0.0156 (0.0126)	-0.0277*** (0.0079)	0.0724*** (0.0138)
Week FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
Day of Week FE	YES	YES	YES	YES	YES	YES	YES
Observations	644469	645589	597962	645589	644463	645589	597962
R-squared	0,26	0,64	0,16	0.04		0.55	0.08

Robust standard errors clustered at the domain level are in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

**Table 7. Impact of Google Core Updates on number of search, desktop and mobile visits per domain type**

Dependent Variable	FIRST STAGE		IV		
	$\Delta \ln(\text{KeyWords TOP10}) t-4$		$\Delta \ln(\text{Search Visits}) t-4$	$\Delta \ln(\text{Desktop Visits}) t-4$	$\Delta \ln(\text{Mobile Visits}) t-4$
Coefficients of Interest	$\beta$ "Big" Core Update	$\beta$ "Small" Core Update	$\beta \Delta \ln(\text{KeyWords TOP10}) t-4$	$\beta \Delta \ln(\text{KeyWords TOP10}) t-4$	$\beta \Delta \ln(\text{KeyWords TOP10}) t-4$
<u>TOP RANK</u>	0.0001 (0.0006)	-0.0026*** (0.0005)	5.0845*** (1.4868)	2.9693*** (0.8491)	2.9184* (1.6065)
<u>BOT RANK</u>	-0.0015*** (0.0006)	-0.0024*** (0.0004)	5.4388*** (1.3257)	2.8425*** (1.0117)	2.8790 (1.8926)
<u>TOP DOM%</u>	-0.0013** (0.0005)	-0.0028*** (0.0004)	5.7484*** (1.2311)	3.4065*** (0.8461)	5.9952*** (1.7329)
<u>BOT DOM %</u>	-0.0002 (0.0006)	-0.0022*** (0.0006)	7.1389*** (2.1115)	4.0163*** (1.4005)	0.4720 (2.1136)
<u>TOP GOOGLE %</u>	-0.0020 (0.0032)	-0.0016 (0.0021)	-7.0346 (6.3987)	5.4355 (10.5390)	-6.6891 (13.8042)
<u>BOT GOOGLE %</u>	-0.0007* (0.0004)	-0.0025*** (0.0003)	6.8018*** (1.1433)	3.8415*** (0.7589)	4.3679*** (1.4259)

**Table 8. Impact of Google Core Updates on number of search, desktop and mobile visits per domain**

Dependent Variable	FIRST STAGE		IV		
	$\Delta \ln(\text{KeyWords TOP10}) t-4$		$\Delta \ln(\text{Search Visits}) t-4$	$\Delta \ln(\text{Desktop Visits}) t-4$	$\Delta \ln(\text{Mobile Visits}) t-4$
Coefficients of Interest	$\beta$ "Big" Core Update	$\beta$ "Small" Core Update	$\beta \Delta \ln(\text{KeyWords TOP10}) t-4$	$\beta \Delta \ln(\text{KeyWords TOP10}) t-4$	$\beta \Delta \ln(\text{KeyWords TOP10}) t-4$
<u>NATIONAL</u>	-0.0016** (0.0007)	-0.0028*** (0.0004)	7.9725*** (1.6319)	5.4994*** (1.0500)	6.2104*** (1.9389)
<u>REGIONAL</u>	-0.0018*** (0.0006)	-0.0037*** (0.0005)	5.3915*** (1.0139)	3.4021*** (0.7939)	5.6872*** (1.4540)
<u>BUSINESS</u>	0.0005 (0.0013)	-0.0017** (0.0007)	12.4077** (5.8336)	3.9189 (3.3803)	10.8587 (8.3448)
<u>SPORTS</u>	0.0017 (0.0014)	-0.0017** (0.0008)	-6.8526 (4.3187)	-3.8886 (2.5973)	-11.2317* (6.4024)
<u>TV/RADIO</u>	0.0012 (0.0012)	-0.0018** (0.0007)	10.9817** (5.0367)	5.2162** (2.5761)	7.8365 (6.8450)



**Table 6. Impact of Google Core Updates on number of search, desktop and mobile visits per country**

Dependent Variable	FIRST STAGE		IV		
	$\Delta \ln(\text{KeyWords TOP10})$ t-4		$\Delta \ln(\text{Search Visits})$ t-4	$\Delta \ln(\text{Desktop Visits})$ t-4	$\Delta \ln(\text{Mobile Visits})$ t-4
Coefficients of Interest	$\beta$ "Big" Core Update	$\beta$ "Small" Core Update	$\beta \Delta \ln(\text{KeyWords TOP10})$ t-4	$\beta \Delta \ln(\text{KeyWords TOP10})$ t-4	$\beta \Delta \ln(\text{KeyWords TOP10})$ t-4
<b>Austria</b>	0.0004 (0.0012)	-0.0006 (0.0013)	18.5679 (25.6457)	-0.0515 (10.2516)	1.1297 (23.5141)
<b>Belgium</b>	-0.0006 (0.0021)	-0.0046*** (0.0011)	2.3189 (1.6237)	2.7453** (1.0280)	5.1430* (2.5972)
<b>Denmark</b>	-0.0030* (0.0016)	-0.0042*** (0.0009)	7.0638** (2.7906)	1.4451 (1.8946)	
<b>Finland</b>	0.0014 (0.0012)	-0.0027** (0.0011)	-3.2330 (1.9556)	-4.1269** (2.0029)	-0.3457 (5.0669)
<b>France</b>	-0.0034*** (0.0012)	0.0019 (0.0032)	-7.8110 (5.6929)	-4.0528 (3.0022)	-5.6614 (4.8236)
<b>Germany</b>	-0.0005 (0.0012)	-0.0014* (0.0008)	6.8452 (6.0689)	5.7789 (3.8744)	-1.4815 (6.1417)
<b>Greece</b>	0.0085*** (0.0020)	-0.0030*** (0.0011)	-3.0628* (1.8023)	-2.3780** (1.0637)	-2.1888* (1.1690)
<b>Ireland</b>	-0.0074*** (0.0021)	-0.0066*** (0.0016)	-0.0312 (0.5429)	-0.3740 (0.8609)	1.1909 (1.2613)
<b>Italy</b>	0.0002 (0.0012)	-0.0001 (0.0009)	46.9613 (204.4139)	16.7190 (74.5345)	-66.1106 (293.4514)
<b>Netherlands</b>	-0.0034*** (0.0012)	-0.0032*** (0.0008)	0.7125 (1.3368)	1.9045** (0.8993)	-3.0322 (2.5607)
<b>Poland</b>	-0.0032*** (0.0010)	-0.0034*** (0.0009)	10.3908*** (2.1928)	6.0095*** (1.5756)	7.8332*** (2.2501)
<b>Portugal</b>	0.0019 (0.0018)	-0.0041*** (0.0013)	2.6860 (2.4840)	1.8328 (1.5616)	6.1433 (4.2944)
<b>Spain</b>	0.0013 (0.0009)	-0.0047*** (0.0006)	6.9571*** (1.0232)	3.2958*** (0.7012)	2.3239 (1.4320)
<b>Sweden</b>	-0.0045*** (0.0015)	-0.0008 (0.0008)	0.5642 (1.3304)	0.7596 (1.5677)	-1.9122 (2.3194)
<b>United Kingdom</b>	-0.0030** (0.0014)	-0.0024** (0.0009)	-4.4395* (2.3095)	-3.4895* (1.8522)	1.3677 (3.2620)

# Are news outlets reacting to GS core updates?

- The change in visits due to Google's algorithm core updates seems important enough to grant a reaction from news outlets.
- We do not observe internal records of each (or any) outlet, but instead we examine behavior in three ways that would be consistent with outlet-level reaction
  - Follow keywords in top10 two weeks after algorithm's update.
  - Changes in visits through paid search
  - Changes in visits through display ads
- We find no evidence of counter-reaction, yet this does not mean news outlets do not react at all to GS algorithm's core updates.

Table 7. Impact of Google Core Updates Two Weeks After

Dependent Variable	(1)	(2)	(3)	(4)
	$\Delta \ln(\text{Words top 10})$ t-4	$\Delta \ln(\text{Search Visits})$ t-4	$\Delta \ln(\text{Desktop Visits})$ t-4	$\Delta \ln(\text{Mobile Visits})$ t-4
$\Delta \ln(\text{Words top 10})$ t-4		2.5431*** (0.4055)	1.4547*** (0.2859)	3.7283*** (0.6376)
Google Core Update t to t+7	-0.0021*** (0.0003)			
Google Core Update t+8 to t+14	-0.0035*** (0.0003)			
$\Delta \ln(\text{Desktop Direct Visits})$ t-4	0.0001 (0.0001)	0.3441*** (0.0258)	0.6019*** (0.0242)	0.3910*** (0.0264)
Constant	0.0074*** (0.0008)	0.0149* (0.0084)	-0.0106* (0.0057)	0.0751*** (0.0102)
Week FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Day of Week FE	YES	YES	YES	YES
Observations	645,589	644,463	645,589	597,962
R-squared	0.04	0.22	0.63	0.10

Robust standard errors clustered at the domain level are in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

**Table 8. Paid Search Visits and Display Ads Visits**

---

	(1)	(2)
<b>Dependent Variable</b>	<b><math>\Delta \ln(\text{Paid Search Visits})</math> t-4</b>	<b><math>\Delta \ln(\text{Display Ads Visits})</math> t-4</b>
<b><math>\Delta \ln(\text{KeyWords TOP10})</math> t-4</b>	0,0375 (0,0893)	0,0723 (0,5421)
<b><math>\Delta \ln(\text{Desktop Direct Visits})</math> t-4</b>	0.0008* (0,0004)	0,0107 (0,0089)
<b>Constant</b>	-0,0004 (0,0010)	0,0036 (0,0049)
<b>Observations</b>	644.463	405.805

---

Standard errors clustered at the domain level in parentheses.

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

# Impact on Country Level HHI

- We compute the *daily* HHI in each country for search, desktop and mobile visits.
- We run first differences of HHI t to t-4 on dummy whether Core Update has taken place in the last 7 days such that,

$$\Delta \ln[HHI_{ct}] = \varphi_0 + \varphi_1 CoreUpdatePlus7_{ct} + \varphi_2 \Delta X_{ct} + \Delta \varepsilon_{ct}$$

- We control for first differences in aggregate direct visits at the market level, as well as day-of-week, month and year FE. Joint regressions clustered at the country level, robust standard errors in per country regressions.

**Table 9. Impact of Core Updates on HHI of Search, Desktop and Mobile Visits**

Dependent Variable	$\Delta \ln(\text{HHI Search Visits}) \text{ t-4}$		$\Delta \ln(\text{HHI Desktop Visits}) \text{ t-4}$		$\Delta \ln(\text{HHI Mobile Visits}) \text{ t-4}$	
Coefficient	$\beta$ "Big" Core Update	$\beta$ "Small" Core Update	$\beta$ "Big" Core Update	$\beta$ "Small" Core Update	$\beta$ "Big" Core Update	$\beta$ "Small" Core Update
<b>All</b>	-0.0110* (0.0054)	0.0086** (0.0029)	-0.0014 (0.0035)	0.0019 (0.0023)	-0.0081 (0.0052)	0.0021 (0.0026)
<b>Austria</b>	-0.0105 (0.0174)	0.0150 (0.0164)	-0.0126 (0.0101)	0.0036 (0.0109)	-0.0180 (0.0165)	-0.0022 (0.0126)
<b>Belgium</b>	0.0170 (0.0178)	0.0080 (0.0100)	0.0081 (0.0094)	0.0021 (0.0077)	-0.0018 (0.0081)	0.0043 (0.0058)
<b>Denmark</b>	-0.0024 (0.0294)	-0.0003 (0.0106)	0.0146 (0.0176)	-0.0049 (0.0044)	-	-
<b>Finland</b>	-0.0186* (0.0108)	0.0142* (0.0082)	-0.0042 (0.0051)	-0.0038 (0.0038)	-0.0116* (0.0069)	-0.0074 (0.0056)
<b>France</b>	-0.0115 (0.0077)	0.0025 (0.0075)	-0.0054 (0.0051)	-0.0041 (0.0053)	-0.0146 (0.0113)	0.0130 (0.0105)
<b>Germany</b>	-0.0477*** (0.0131)	0.0103 (0.0082)	-0.0040 (0.0057)	0.0041 (0.0045)	-0.0345** (0.0138)	0.0122 (0.0108)
<b>Greece</b>	-0.0433*** (0.0147)	-0.0091 (0.0134)	0.0096 (0.0078)	0.0065 (0.0093)	-0.0269*** (0.0101)	0.0109 (0.0076)
<b>Ireland</b>	-0.0208 (0.0222)	0.0096 (0.0172)	-0.0223 (0.0154)	-0.0045 (0.0121)	-0.0015 (0.0159)	-0.0076 (0.0120)
<b>Italy</b>	0.0043 (0.0126)	0.0038 (0.0128)	0.0237 (0.0218)	0.0410** (0.0199)	0.0246 (0.0160)	-0.0055 (0.0101)
<b>Netherlands</b>	0.0052 (0.0146)	0.0287* (0.0172)	0.0059 (0.0069)	0.0061 (0.0063)	-0.0007 (0.0084)	-0.0090 (0.0081)
<b>Poland</b>	-0.0041 (0.0124)	0.0042 (0.0098)	0.0044 (0.0075)	0.0058 (0.0057)	0.0235** (0.0115)	0.0074 (0.0114)
<b>Portugal</b>	0.0243* (0.0143)	-0.0060 (0.0129)	0.0035 (0.0103)	0.0074 (0.0074)	0.0026 (0.0087)	0.0146 (0.0098)
<b>Spain</b>	-0.0001 (0.0111)	0.0164 (0.0124)	-0.0009 (0.0119)	-0.0084 (0.0082)	-0.0164 (0.0133)	0.0017 (0.0079)
<b>Sweden</b>	-0.0069 (0.0148)	0.0264 (0.0163)	0.0046 (0.0100)	-0.0026 (0.0073)	0.0111 (0.0089)	0.0007 (0.0079)
<b>United Kingdom</b>	-0.0078 (0.0094)	0.0169 (0.0110)	-0.0056 (0.0079)	0.0133 (0.0091)	-0.0071 (0.0095)	0.0175* (0.0102)

# From market level HHI to keyword level HHI ...

- Detailed data of traffic generated from all keywords to each outlet
  - *Monthly data*
  - *Focus in Spain*: 65 news outlets; easy to analyze for us
  - Drop keywords that show up in our data in only one month or one outlet.
- Compute monthly HHI for all remaining keywords across outlets in Spain.

$$\ln[HHI_{kt}] = \varphi_0 + \varphi_1 CoreUpdate_t + \varphi_2 CoreUpdatePlus1_t + \varphi_3 CoreUpdatePlus2_t + \theta FE_{ct(m)} + \varepsilon_{kt},$$

Regressions contain keyword fixed effects plus year\*month and length run fixed effects. Standard errors clustered at the keyword level.

# Impact of Google Core Updates on HHI per keyword in Spain

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<b>Dep Var: ln(HHI traffic)</b>								
<b>Google Core Update</b>	-0.0342*** (0.0067)	0.0064 (0.0066)	0.0042 (0.0174)	0.0064 (0.0050)	0.0346** (0.0148)	0.0034 (0.0072)	-0.0106 (0.0080)	0.0185** (0.0090)
<b>Google Core Update + 1 month</b>	-0.0057 (0.0065)	0.0023 (0.0067)	0.0269 (0.0173)	-0.0001 (0.0051)	0.0334** (0.0150)	-0.0021 (0.0072)	-0.0164** (0.0081)	0.0267*** (0.0092)
<b>Google Core Update + 2 months</b>	0.0310*** (0.0065)	-0.0079 (0.0067)	-0.0061 (0.0159)	0.0095* (0.0050)	0.0301** (0.0149)	-0.0140* (0.0072)	0.0143* (0.0081)	-0.0023 (0.0090)
<b>"Big" Google Core Update</b>		-0.0406*** (0.0066)	-0.0780*** (0.0172)	-0.0070 (0.0049)	-0.0190 (0.0139)	-0.0502*** (0.0075)	0.0664*** (0.0078)	-0.0911*** (0.0090)
<b>"Big" Google Core Update + 1 month</b>		-0.0080 (0.0065)	-0.0276 (0.0171)	0.0009 (0.0049)	-0.0140 (0.0140)	-0.0102 (0.0072)	0.0868*** (0.0078)	-0.0587*** (0.0089)
<b>"Big" Google Core Update + 2 months</b>		0.0389*** (0.0063)	0.0819*** (0.0158)	0.0050 (0.0049)	0.0005 (0.0138)	0.0486*** (0.0070)	0.0833*** (0.0078)	0.0062 (0.0084)
<b>Constant</b>	-0.7630*** (0.0035)	-0.7630*** (0.0035)	-1.2252*** (0.0078)	-0.5945*** (0.0025)	-0.8710*** (0.0064)	-0.7302*** (0.0041)	-0.6002*** (0.0045)	-0.4102*** (0.0046)
<b>Month Run Length FE</b>	YES	YES	YES	YES	YES	YES	YES	YES
<b>KeyWord FE</b>	YES	YES	YES	YES	YES	YES	YES	YES
<b>Month-Year FE</b>	YES	YES	YES	YES	YES	YES	YES	YES
<b>Sample</b>	All	All	> 4 sites	< 5 sites	> 8 months	< 9 months	Top 10 Kword	Not Top 10 Kword
<b>Observations</b>	3,644,630	3,644,630	1,003,707	2,640,923	945,559	2,699,071	1,863,610	1,781,020
<b>R-squared</b>	0.61	0.61	0.49	0.49	0.68	0.54	0.56	0.63



# Summary of results

- We examine how search engines affect distribution of search visits across news outlets ... through Google algorithm's core updates.
- We unveil a mechanism:
  - Core updates reduced the # of Keywords in top 10 search results.
  - A 1% decrease in the # of keywords in top 10 search results generates a 6.6% decrease in search visits, and around 4% decrease in total desktop and mobile visits.
- Core updates affect market concentration
  - “Big” updates reduce market concentration by 1%, while “non-big” updates increased market concentration by 0.08% at the market level.
  - “Big” updates reduce market concentration by 4%, relative to “non-big” updates and no updates at the keyword-market level.

**Thanks!**

# Contribution

- Recent studies analyze the effects of **digitalization** on the media market.
  - Competition (Athey et al. 2017, Chiou & Tucker, 2017; Calzada & Gil, 2021);
  - Quality of journalism (Cagé et al, 2020; Bandy & Diakopoulos, 2020);
  - Democratic institutions (Gentzkow & Shapiro, 2011; Boxell, Gentzkow, & Shapiro, 2017; Peterson, Goel & Iyengar, 2019).
- Theoretical literature analyze the existence of **bias**
  - **in search engines** (Cornière and Taylor, 2014; Burguet, Caminal, and Ellman, 2015; Belleflamme and Peitz, 2018);
  - **in recommendation systems** (Bourreau and Gaudin, 2018; Bourreau et al. 2021; Drugov and Jeon, 2017).
- Little is known about the effects of **search engines** in the development of media markets (Sismeiro & Mahmood, 2018; Cagé, Hervé & Mazoyer, 2020).