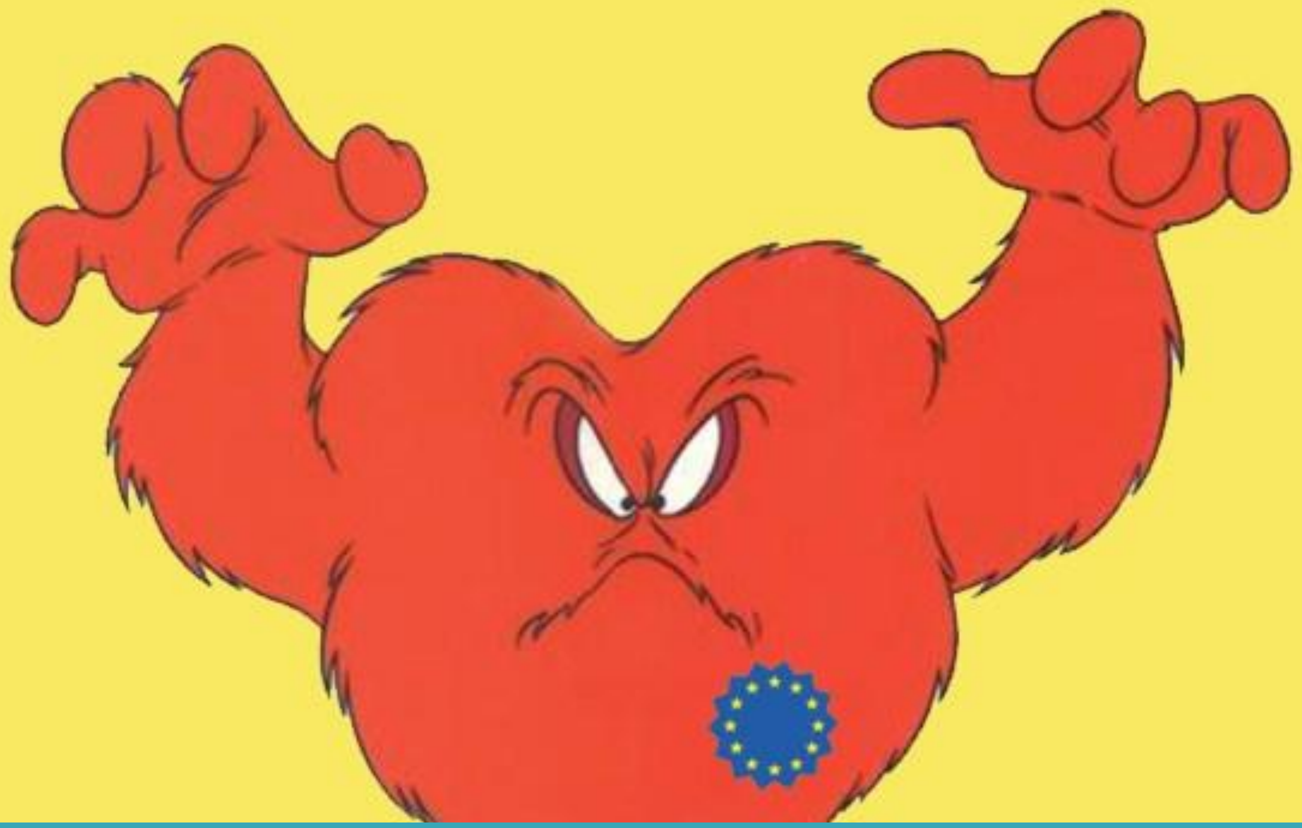
garjoh_canuck



Privacy & market concentration: Intended & unintended consequences of the GDPR

Garrett Johnson (Questrom, Boston U)
Scott Shriver (Leeds, U Colorado Boulder)
Sam Goldberg (Kellogg, Northwestern U)

Can privacy policy hurt competition?

Economies of scale

Larger firms have more resources for compliance

B2B choice of data vendors

Firms may prefer larger vendors because they have better products

Overview: GDPR & Web technology

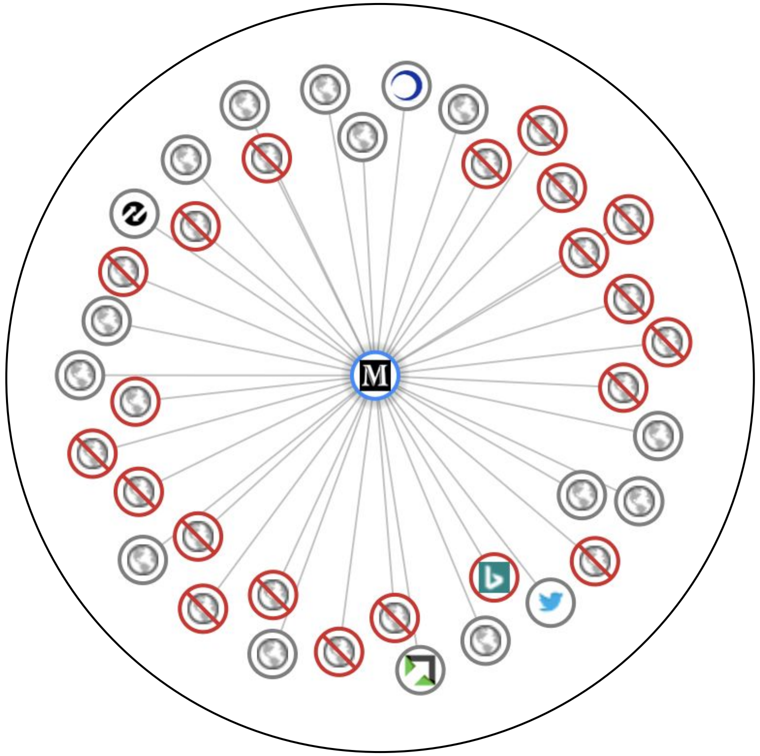
- General Data Protection Regulation (GDPR)
  - Enforced as of May 25, 2018
  - ↑ **logistical cost & legal risk of personal data processing**

- **Technology vendors** provide an **ecosystem** for the web to **thrive**
  - Vendors help websites: monetize with ads, load & share content, measure & optimize site traffic
  - Vendors often share personal data
  - Regulatory **scrutiny**, but **no fines to date**

Overview

- Data

- Results
  - GDPR effect on vendor use
  - Concentration of vendor market
  - Heterogeneity by website



Data

theguardian.com/books/2020/jul/15/great-gatsby-prequel-2021-copyright-michael-farris-smith-f-scott-fitzgerald

Support The Guardian

Available for everyone, funded by readers

Contribute → Subscribe →

Search jobs | Sign in | Search | International edition

The Guardian

News | Opinion | Sport | Culture | Lifestyle | More

Books Music TV & radio Art & design Film Games Classical Stage

Books

The Great Gatsby prequel set for release days after copyright expires

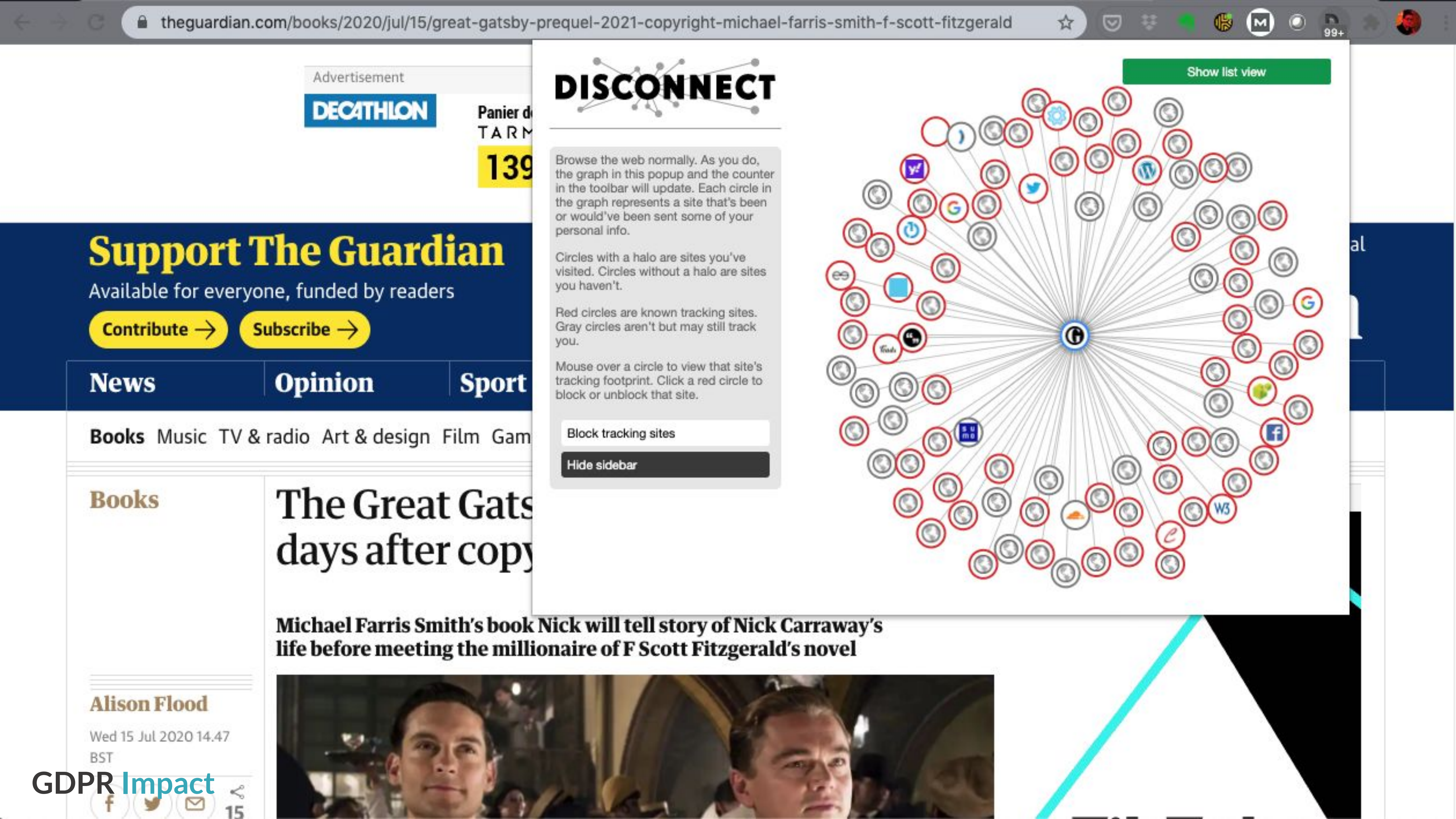
Michael Farris Smith's book Nick will tell story of Nick Carraway's life before meeting the millionaire of F Scott Fitzgerald's novel

Alison Flood

Wed 15 Jul 2020 14.47 BST

15

GDPR Impact


theguardian.com/books/2020/jul/15/great-gatsby-prequel-2021-copyright-michael-farris-smith-f-scott-fitzgerald
Advertisement
DECATHLON
DISCONNECT
Show list view
Browse the web normally. As you do, the graph in this popup and the counter in the toolbar will update. Each circle in the graph represents a site that's been or would've been sent some of your personal info.
Circles with a halo are sites you've visited. Circles without a halo are sites you haven't.
Red circles are known tracking sites. Gray circles aren't but may still track you.
Mouse over a circle to view that site's tracking footprint. Click a red circle to block or unblock that site.
Block tracking sites
Hide sidebar
Support The Guardian
Available for everyone, funded by readers
Contribute →
Subscribe →
News
Opinion
Sport
Books
Music
TV & radio
Art & design
Film
Books
The Great Gats
days after copy
Michael Farris Smith's book Nick will tell story of Nick Carraway's life before meeting the millionaire of F Scott Fitzgerald's novel
Alison Flood
Wed 15 Jul 2020 14.47 BST
15

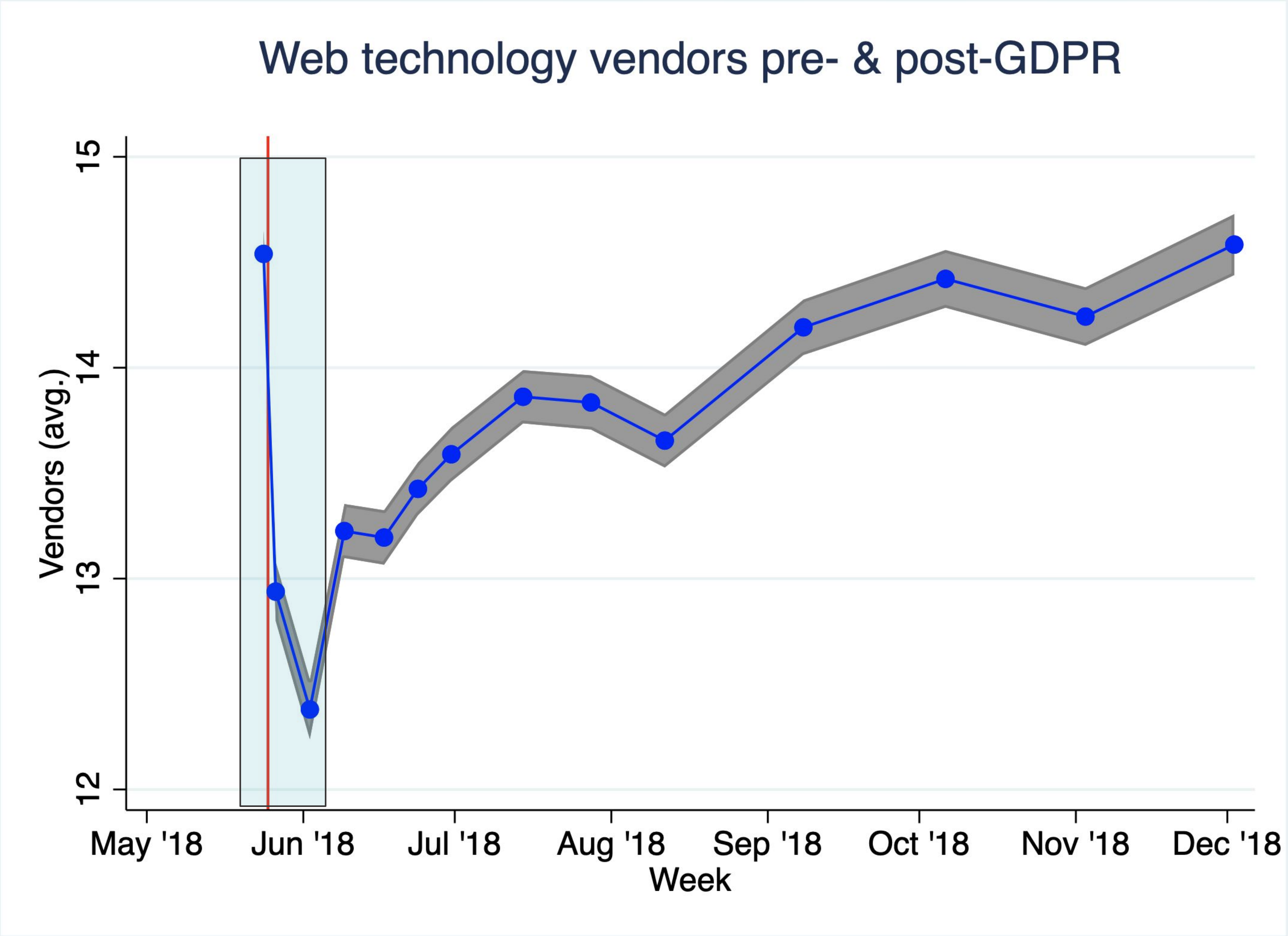
GDPR Impact

Collect data 3rd party domains on top websites

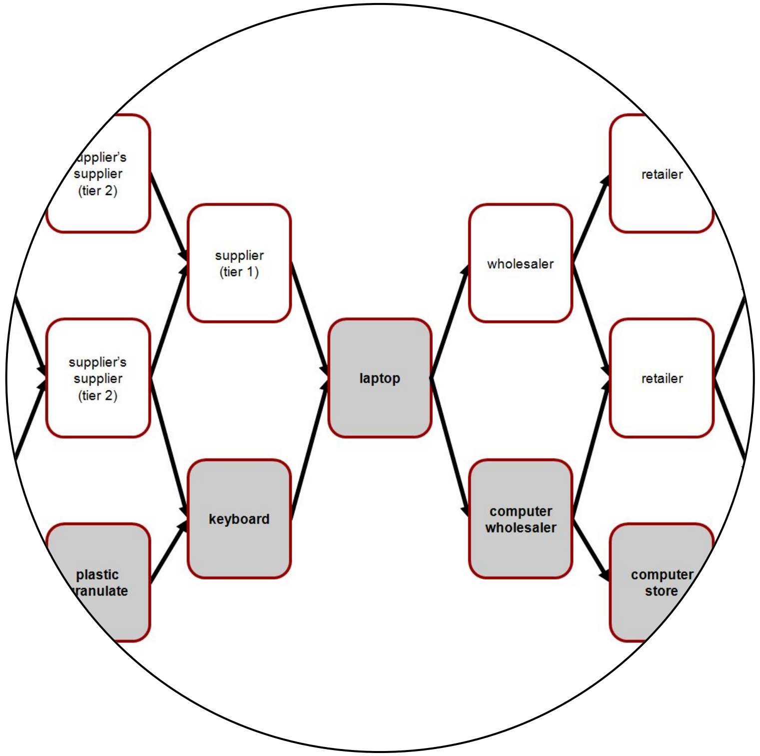
1. Use VPN service to simulate EU user (France)
2. Use "webxray" program (Libert 2015) to record all 3rd party domains when visiting a website
3. Repeat for 27K top sites regularly in 2018 (14 times)
   - Top 2,000 websites in
     - Each of 28 EU countries
     - US
     - Canada
     - Globally

Results: GDPR effect on vendor use

Short run: vendors fall 15% 1 week post GDPR



[Pre-trend]

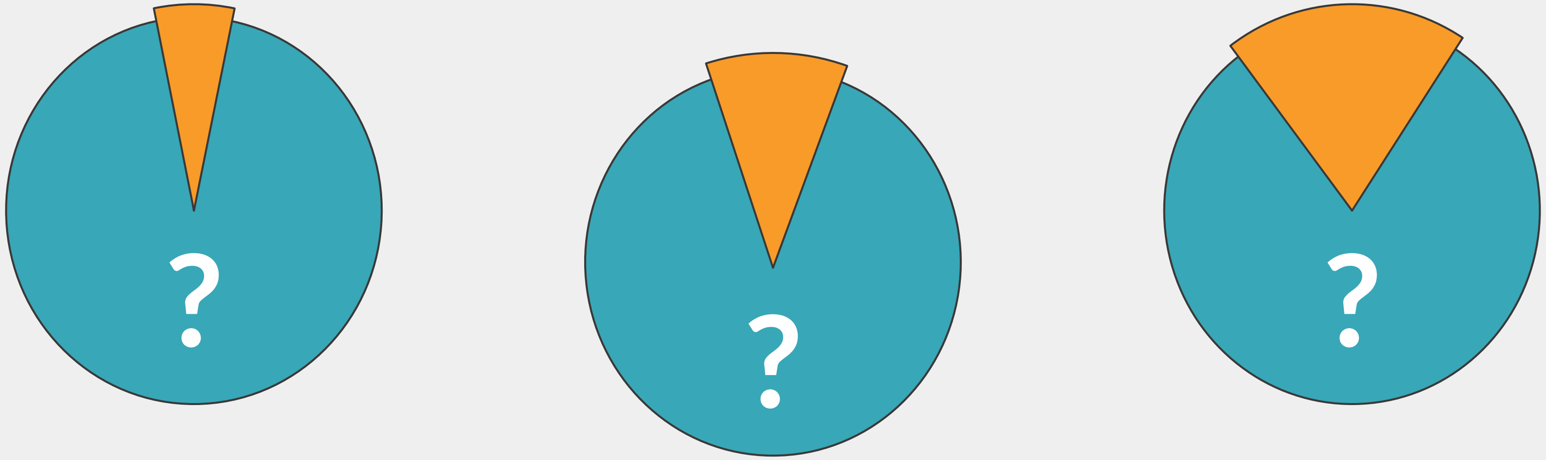


Results: Concentration of vendor market

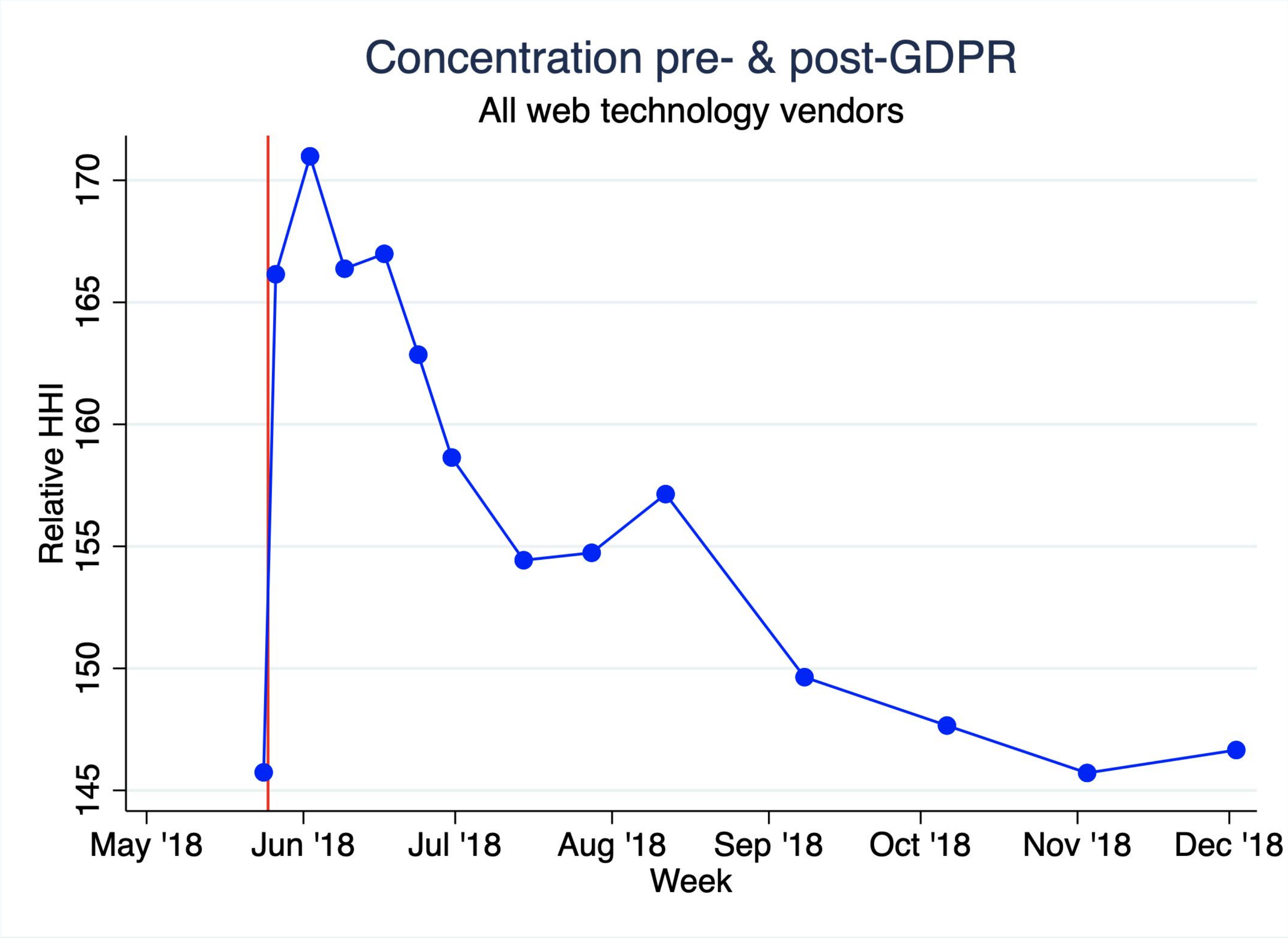
Do larger vendors get a larger share of the smaller pie?

Pre-GDPR

Post-GDPR



Short run: concentration rises 17% post-GDPR



↑ concentration in top 4 categories (94.3% of vendors)

Category	Vendor Share (Pre)	HHI Pre	HHI Post	HHI Diff. (%)
All vendors		146	171	17.3%
All categorized vendors		308	363	17.8%
Advertising	50.2%	348	436	25.3%
Hosting	20.5%	1,892	1,936	2.3%
Audience measurement	14.4%	4,116	4,355	5.8%
Social media	9.2%	4,251	4,412	3.8%

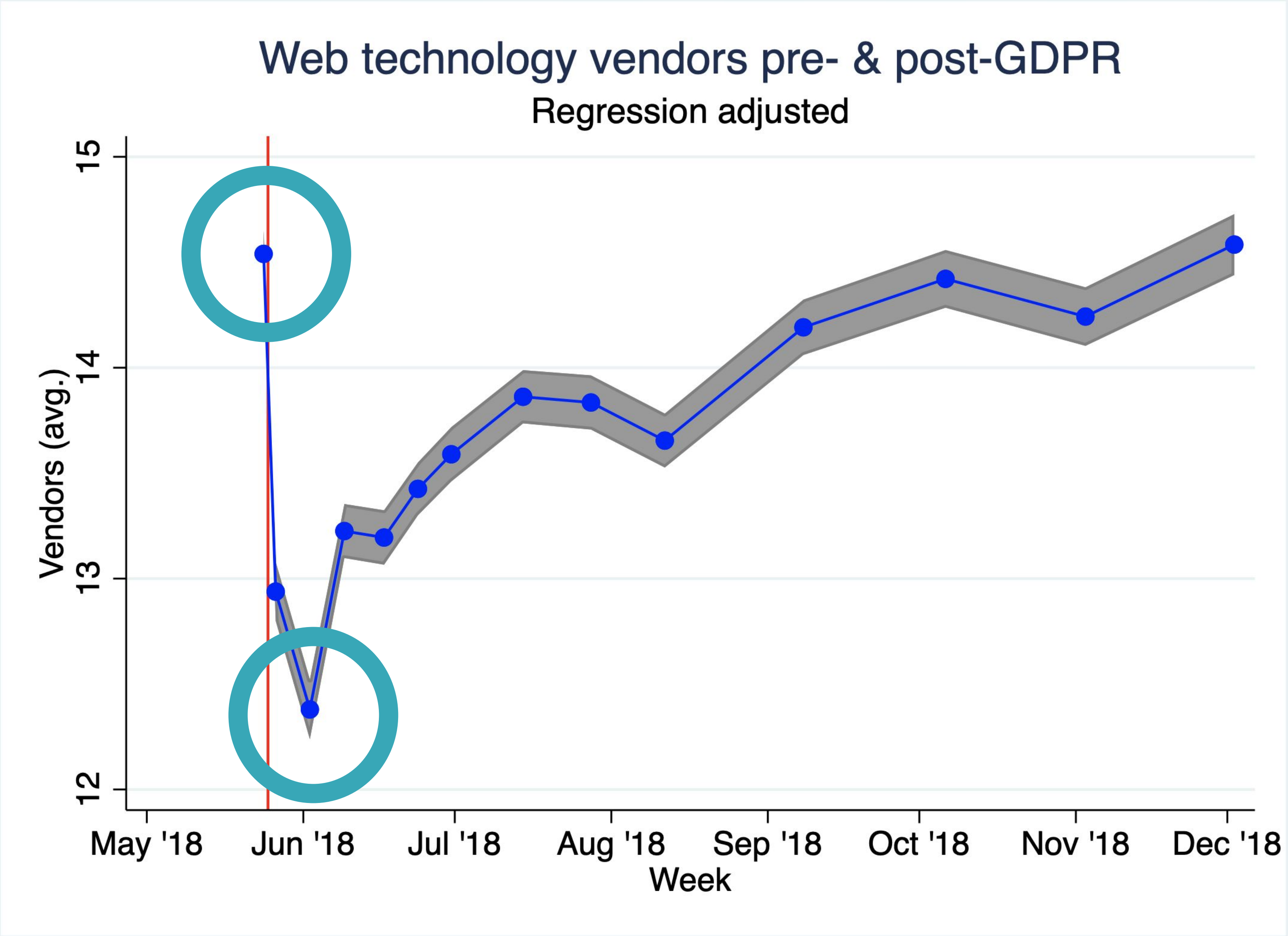
Extension: Google & Facebook drive ↑ concentration

Data samples	HHI Pre	HHI Post	Diff.	Diff. (%)
Role of top 2 companies (Google & Facebook)				
All vendors	145.7	171.0	25.2	17.3%
All but top 2 companies	46.0	43.2	-2.8	-6.2%

Results: Heterogeneity by website

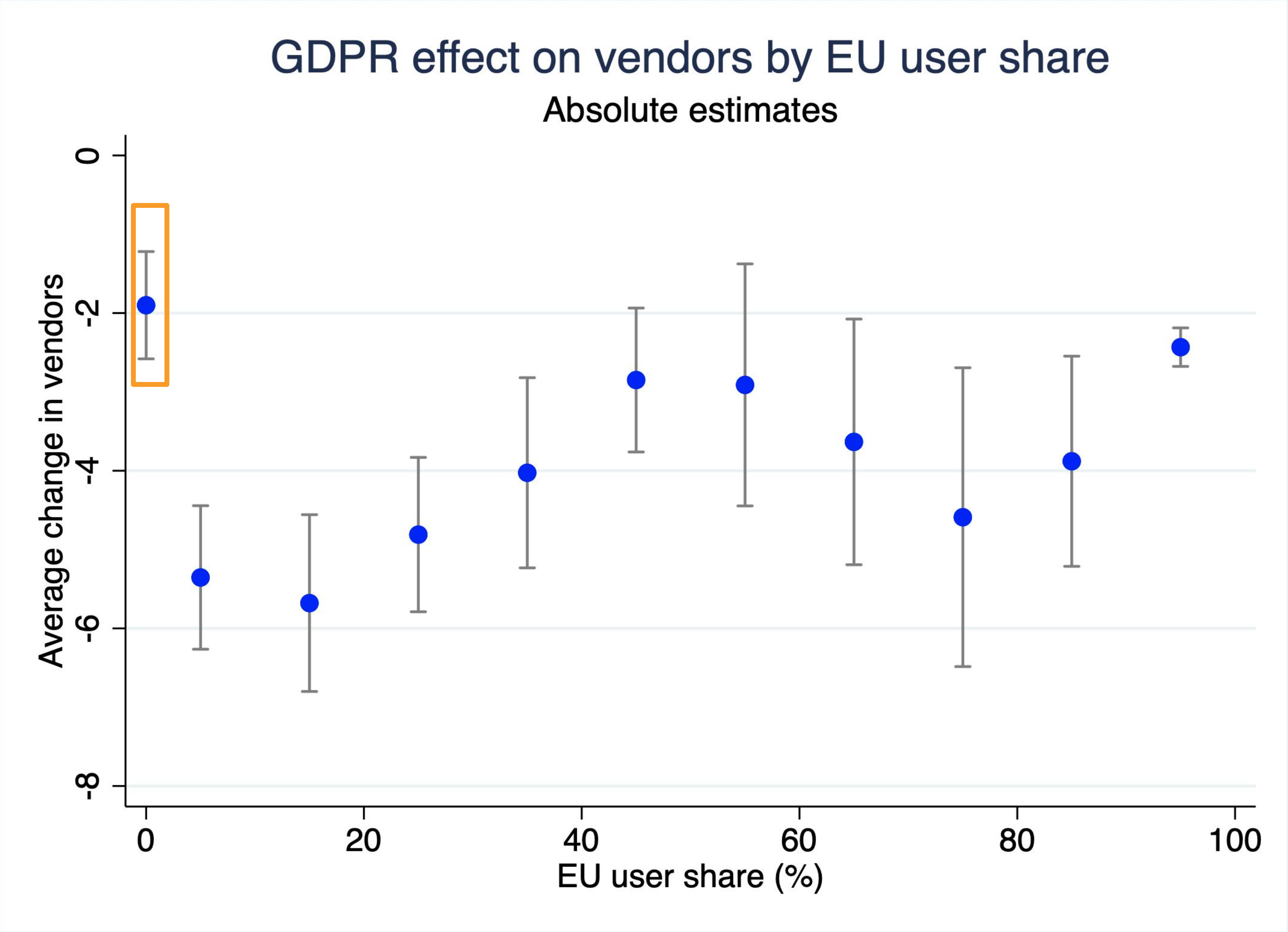
Heterogeneity in short run

Short run = full week after enforcement deadline vs pre-deadline baseline



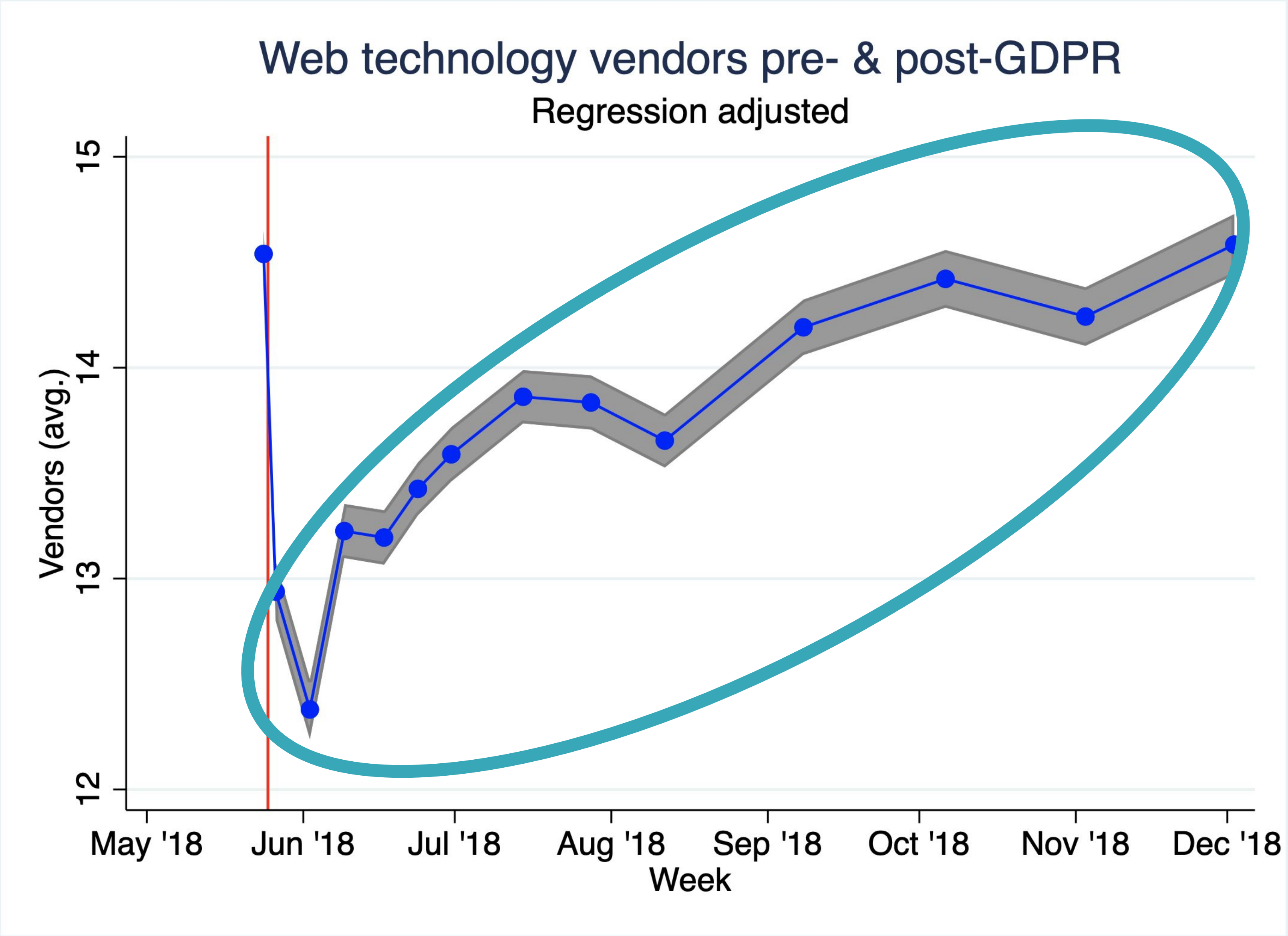
Fine on global revenues most deters non-EU sites

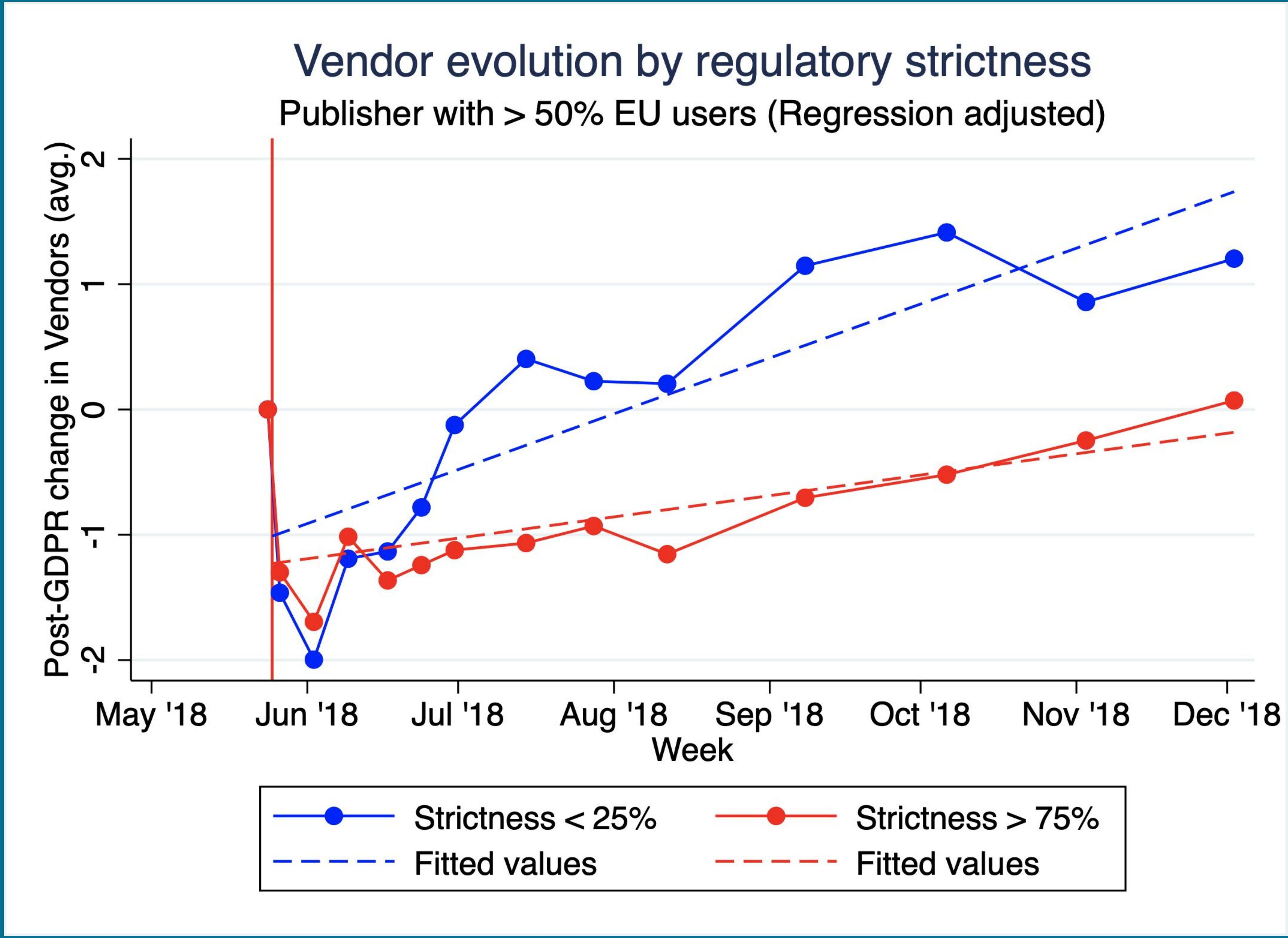
Discontinuity for sites with 0% EU users: these sites not subject to the GDPR



Heterogeneity in long run evolution

Long run evolution = post-GDPR trend in 2018





Vendor evolution by regulatory strictness
Publisher with > 50% EU users (Regression adjusted)
Post-GDPR change in Vendors (avg.)
2
1
0
-1
-2
May '18
Jun '18
Jul '18
Aug '18
Sep '18
Oct '18
Nov '18
Dec '18
Week
Strictness < 25%
Fitted values
Strictness > 75%
Fitted values

Summary: GDPR's intended & unintended consequences

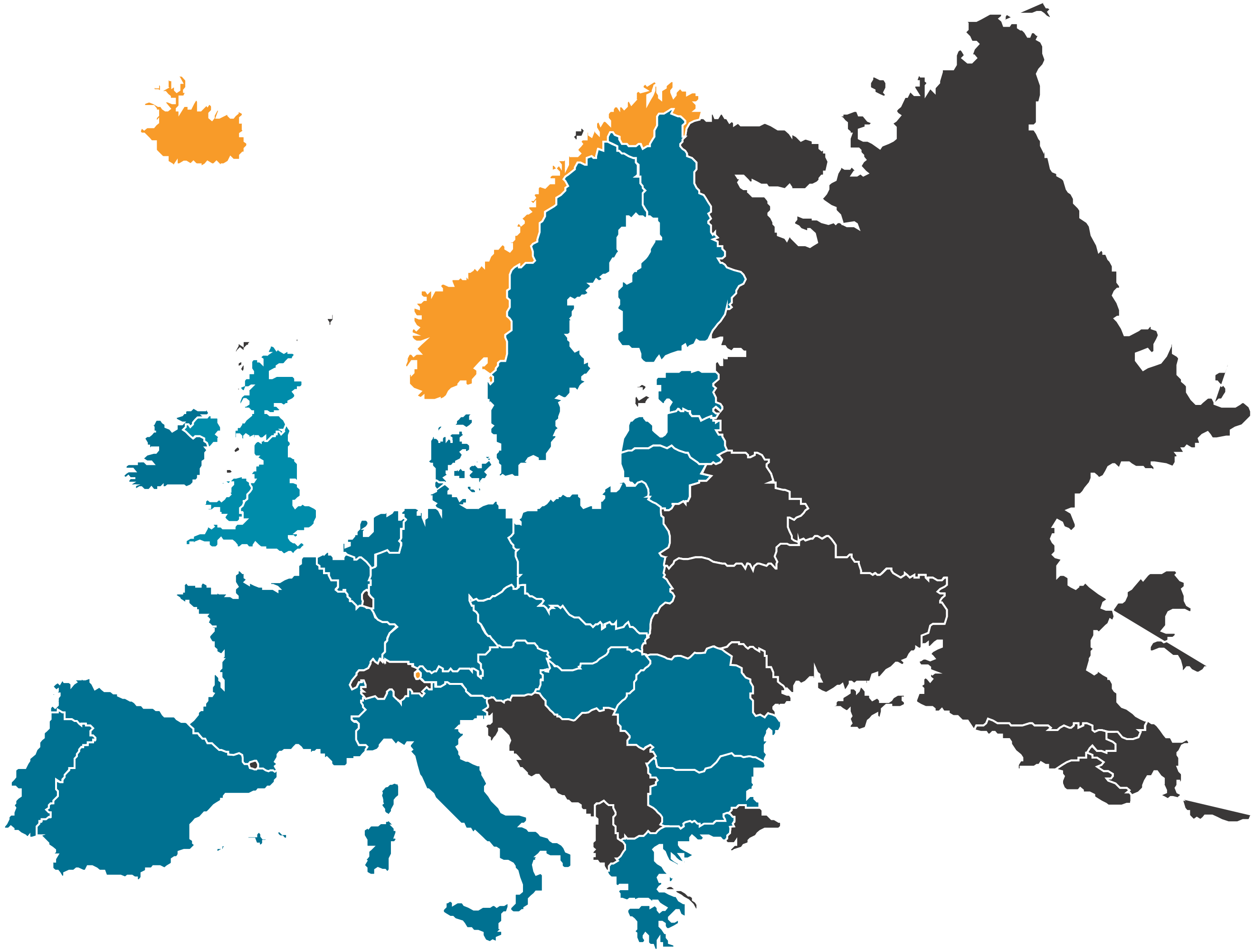
- Novel empirical evidence of **privacy-competition policy tension**
- **Intended consequence**: ↓ web tech data sharing
- **Unintended consequences**:
    - ↑ vendor concentration
    - Sites with most EU visitors reduce vendors the least


GDPR

Appendix

GDPR

General Data Protection Regulation

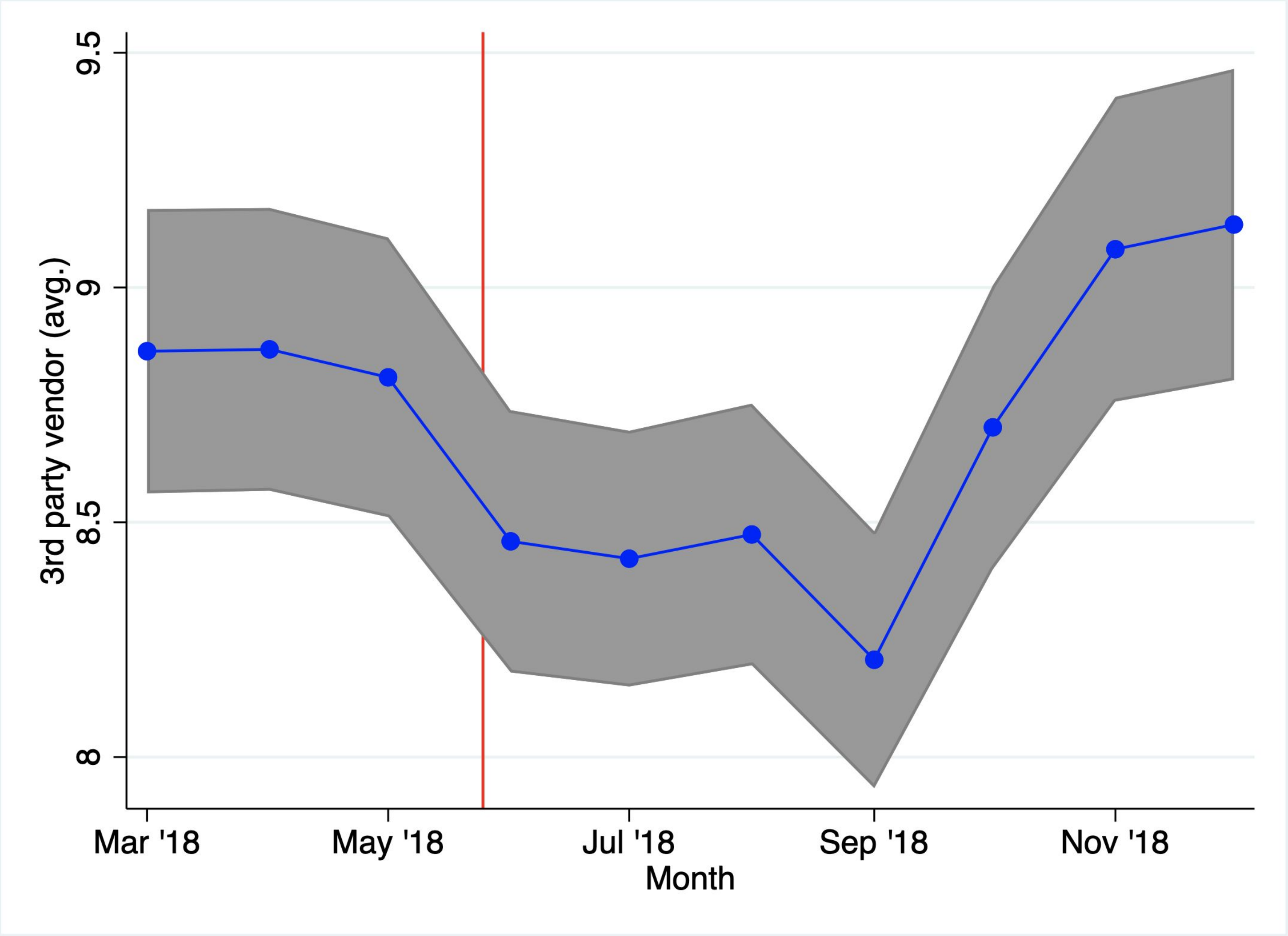


EU
EEA
Brexit

GDPR Impact

No pre-GDPR trend in vendors

1,452 site sample overlapping with ours from WhoTracksMe dataset



Fewer vendors in all categories but compliance

Short run estimates: 1 week post-GDPR

	Avg. Vendors		
Category	Pre	Post	Diff. (%)
All vendors	14.44	12.35	-14.5%
All categorized vendors	8.40	6.91	-17.7%
Advertising	4.35	3.29	-24.3%
Hosting	1.78	1.61	-9.7%
Audience measurement	1.25	1.11	-10.9%
Social media	0.79	0.70	-11.5%
Design optimization	0.22	0.20	-10.5%
Security	0.15	0.12	-17.8%
Native ads	0.078	0.066	-14.8%
CRM	0.022	0.019	-9.6%
Privacy compliance	0.017	0.021	23.2%

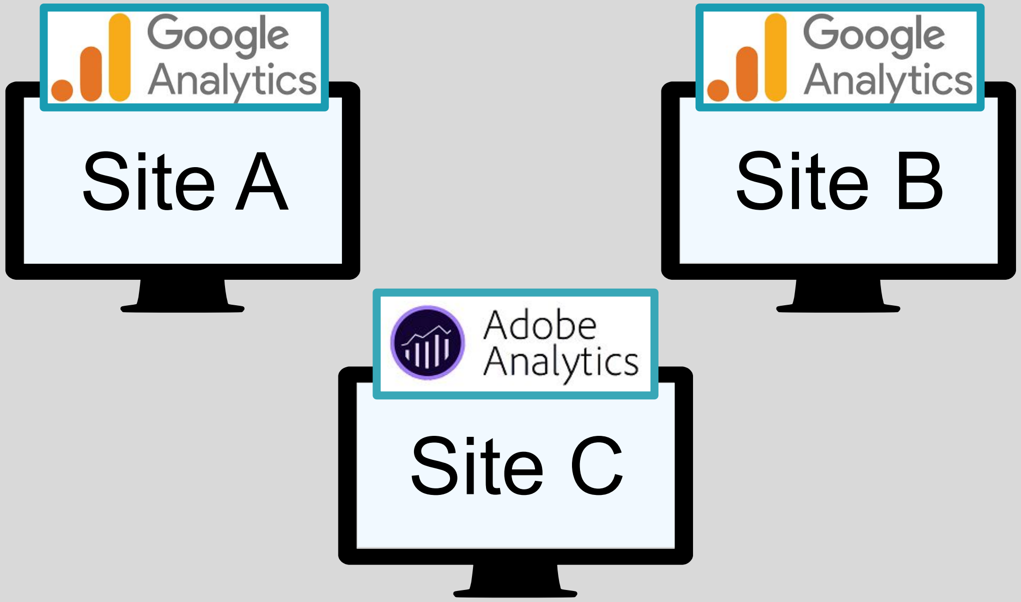
Defining relative market concentration

- Reach: # of websites using vendor
- **Vendor's relative market share**:

  market share=own-reach / ∑reach

- **Herfindahl–Hirschman Index (HHI)**

  $\sum \textit{market share}^2$



	Reach	Share
Google Analytics	2	2/3
Adobe Analytics	1	1/3

↑ concentration in top 4 categories (94.3% of vendors)

Category	Vendor Share (Pre)	HHI		
		Pre	Post	Diff. (%)
All vendors		146	171	17.3%
All categorized vendors		308	363	17.8%
Advertising	50.2%	348	436	25.3%
Hosting	20.5%	1,892	1,936	2.3%
Audience measurement	14.4%	4,116	4,355	5.8%
Social media	9.2%	4,251	4,412	3.8%
Design optimization	2.6%	2,874	2,861	-0.5%
Security	1.8%	8,926	9,722	8.9%
Native ads	0.9%	4,229	4,024	-4.8%
CRM	0.2%	6,408	6,119	-4.5%
Privacy compliance	0.2%	3,925	4,116	4.9%

Robustness of concentration effect

- Defining markets by vendor category
  - Categories include: advertising, social media, site analytics
  - Two categorizations
    - *Webxray*: ↑ Concentration in top 4 categories (94% of total)
    - *WhoTracksMe*: ↑ Concentration in top 3 categories (85% of total)
- Alternate concentration metrics
  - Concentration ratios: Total share of top *N* firms
  - "Head-to-head competition": Probability site retains dominant firm given drops only one of two vendors

Webxray categorization