DISCUSSION PAPER SERIES

DP14475

EUROPEAN PRIVACY LAW AND GLOBAL MARKETS FOR DATA

Christian Peukert, Stefan Bechtold, Michail Batikas and Tobias Kretschmer

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

EUROPEAN PRIVACY LAW AND GLOBAL MARKETS FOR DATA

Christian Peukert, Stefan Bechtold, Michail Batikas and Tobias Kretschmer

Discussion Paper DP14475
Published 08 March 2020
Submitted 06 March 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Industrial Organization

EUROPEAN PRIVACY LAW AND GLOBAL MARKETS FOR DATA

Abstract

We demonstrate how privacy law interacts with competition and trade policy in the context of the European General Data Protection Regulation (GDPR). We follow more than 110,000 websites for 18 months to show that websites reduced their connections to web technology providers after GDPR became effective, especially regarding requests involving personal data. This also holds for websites catering to non-EU audiences and therefore not bound by GDPR. We further document an increase in market concentration in web technology services after the introduction of GDPR. While most firms lose market share, the leading firm, Google, significantly increases market share.

JEL Classification: L15, L86, K21, L12

Keywords: privacy, competition policy, Antitrust, Internet regulation, compliance risk, GDPR, Brussels effect, cookies, web tracking

Christian Peukert - christian.peukert@gmail.com
ETH Zurich and UCP Lisbon

Stefan Bechtold - sbechtold@ethz.ch
ETH Zurich

Michail Batikas - michail.batikas@rennes-sb.com
Rennes School of Business

Tobias Kretschmer - t.kretschmer@lmu.de
University of Munich and CEPR

Acknowledgements
We acknowledge the support from FCT {Portuguese Foundation of Science and Technology for the project UID/GES/00407/2013, and thank Daniel Bötticher for helpful research assistance.

European Privacy Law and Global Markets for Data*

Christian Peukert[1,2], Stefan Bechtold[1], Michail Batikas[3], and Tobias Kretschmer[4,5]

[1]ETH Zurich
[2]UCP – Católica Lisbon School of Business and Economics
[3]Rennes School of Business
[4]LMU Munich
[5]CEPR London

March 6, 2020

Abstract

We demonstrate how privacy law interacts with competition and trade policy in the context of the European General Data Protection Regulation (GDPR). We follow more than 110,000 websites for 18 months to show that websites reduced their connections to web technology providers after GDPR became effective, especially regarding requests involving personal data. This also holds for websites catering to non-EU audiences and therefore not bound by GDPR. We further document an increase in market concentration in web technology services after the introduction of GDPR. While most firms lose market share, the leading firm, Google, significantly increases market share.

Keywords: Privacy, competition policy, antitrust, Internet regulation, compliance risk, GDPR, Brussels effect, cookies, web tracking

*Peukert (corresponding author): christian.peukert@gmail.com, Bechtold: sbechtold@ethz.ch, Batikas: michail.batikas@rennes-sb.com, Kretschmer: t.kretschmer@lmu.de. We acknowledge the support from FCT – Portuguese Foundation of Science and Technology for the project UID/GES/00407/2013, and thank Daniel Bötticher for helpful research assistance.

1 Introduction

The Internet has revolutionized global trade and data flows. In many aspects of our daily life, it has torn down national borders and facilitates communication and trade across the globe. Regulating such a world is challenging. As international coordination mechanisms have often proven ineffective, individual countries and regions have enacted legal regimes for the digital world, even if these regimes have extra-territorial implications. This can result in regime competition to become a leading global digital rule-maker.

In the area of privacy, examples of such regimes include the European Union (EU)'s General Data Protection Regulation of 2018 (GDPR) and the California Consumer Privacy Act of 2020 (CCPA). GDPR is the cornerstone of European privacy law and is often considered the most comprehensive, globally leading privacy regime. It has affected consumers, firms, and countries outside the EU through a variety of mechanisms.

Privacy protection in Europe has traditionally been comparably strong for historical, cultural, political and legal reasons (Bradford, 2020, 136–141), (Schwartz and Peifer, 2017, 123–127). The EU took privacy protection to a new level when it adopted GDPR, which became applicable on May 25, 2018 (European Union, 2016). It establishes common rules on data processing throughout the EU and is directly binding for companies and residents in the EU and beyond. With GDPR, the European legislator intended to harmonize privacy law and enforcement throughout the EU and increase the protection of individuals' privacy while maintaining the benefits of data processing, for example by implementing a data minimization principle (Art. 5(1)(c), 25(1) and Recitals 78, 156 GDPR).

We investigate empirically whether and how the interplay between websites, web technology providers and consumers has changed with GDPR, both within and outside the EU and explore changes in the structure and competition in markets for web technologies.

We follow more than 110,000 websites from May 2017 to November 2018 and observe which web technologies they are using. Websites may use technologies to raise advertising revenues, observe user behavior, share information through social media, or host audiovisual content. We observe whether a website uses such technologies through the HTTP requests the website makes to external servers, and map these requests to third-party firms. We also have information on the stated privacy policies of these vendors. Further, we use disaggregated consumer behavior data to investigate how the effectiveness of GDPR varies with different types of consumers.

We highlight four key findings. First, websites have reduced the number of third-parties that they request in the short-run after GDPR, but the long-run increasing trend is not changed much. However, when looking into a specific type of request – cookies – we find that websites' use of third-party technologies that interact with consumers' privacy has changed effectively with GDPR. We observe a sustained decrease in third-party cookies and a sustained increase in first-party cookies. Second, the changes have affected websites and consumers worldwide who are not directly subject to the regulation. Third, we find that privacy policies of technology vendors become more informative. Fourth, we document a substantial shift in market power. With the introduction of GDPR, the dominant firm in many markets for web technologies, Google, increases its market share whereas all other firms that supply web technology either do not see a change in market share or suffer losses.

We show that GDPR has shaped the global privacy regime beyond the EU, therefore acting as a tool of the EU to export its regulatory framework to other countries. Our results also illustrate that regulating privacy can have unintended consequences on market structure and competition.

We add to a growing empirical literature on technical and commercial implications of GDPR, and relate to theoretical work on its implications for competition policy. This literature spans across disciplines, including industry reports (WhotracksMe, 2018), and academic work in computer science and communication (Dabrowski et al., 2019; Solomos et al., 2019; Hu and Sastry, 2019; Degeling et al., 2018; Libert et al., 2018; Sørensen and Kosta, 2019; Urban et al., 2020), marketing (Johnson and Shriver, 2020; Goldberg et al., 2019; Goddard, 2017), information systems and economics (Godinho de Matos and Adjerid, 2020; Zhuo et al., 2019; Sharma et al., 2019; Economides et al., 2019; Lefrere et al., 2019; Jia et al., 2019b,a), and law (Gal and Aviv, 2019). Our contribution beyond the existing literature is manifold. First, we provide a legal analysis of implications for international privacy regulation and competition policy. Second, unique data lets us control for important pre-existing trends and therefore estimate long-run effectiveness. Third, we investigate the compliance of websites and third-party web technology providers and shed light on consumer-level heterogeneity. Fourth, we show how privacy law interacts with other policy dimensions by comparing EU-audiences and non-EU-audiences and documenting implications for market power.

2 Institutional Background

Five aspects of GDPR are of particular relevance for our paper: institutional changes, territorial application, consent, joint responsibility, and increased legal uncertainty.

2.1 Institutional Changes

Under European privacy law before GDPR, maximum fines typically varied between €12,000 and €600,000, set at the EU member state level, with considerable heterogeneity in enforcement efforts. GDPR increased the maximum fines and strengthened data protection authorities. Fines are now up to €20 million or 4% of the total worldwide annual turnover, whichever is higher (Art. 83 (5) and (6) GDPR).[1] European privacy law now resembles European antitrust law, where the EC has issued several multi-billion dollar fines over the last decade. Under the former General Data Protection Directive of 1995 (GDPD) (European Union, 1995), there were considerable differences regarding the position of data protection authorities in the member states of the EU, resulting from differences in history, case law, culture and the internal organization of the member states, as well as the GDPD itself. GDPR introduced far-reaching changes for the competence and cooperation of national data protection authorities, which are responsible for monitoring the application of GDPR.

2.2 Territorial Application

In international public law, countries are traditionally entitled to enact laws and exercise authority within their geographical boundaries. The internet has challenged the concept of rule-setting based on geography. Persons located in a country cannot only access websites operated out of that country. They may browse, communicate, and shop on any website located anywhere around the globe.

To effectively protect the privacy of persons located in the EU, GDPR can be applicable even if personal data is processed by a firm located outside the EU. It is sufficient that the firm intentionally offers goods or services to people located in the EU, or if it monitors or predicts their behavior, personal preferences or attitudes within the EU (Art. 3(1), (2) and Recital 24 GDPR, European Data Protection Board, 2019a, 15–20). GDPR protects natural persons *located* in the EU, irrespective of their citizenship, residence, or other legal status (European Data Protection Board, 2019a, 14–16). Whether a firm *targets* persons located in the EU can only be decided on a case-by-case basis. This involves the language, currency, marketing efforts (Recital 23 GDPR).

Due to this broad territorial application, GDPR can effectively apply to websites and web technology providers located anywhere in the world regardless of their business location or legal incorporation, as long as they are accessed by and used for people located in the EU. Websites and web technology

[1] As a result, for any company with a worldwide turnover higher than €15 million, GDPR led to an increased sanctioning regime.

providers that are not located in the EU but are subject to GDPR are required to designate a representative in the EU (Art. 27 GDPR). They are bound by all rules of GDPR, including its damages and fines regime. Compared to former European privacy law, this is a drastic expansion of the global reach of European privacy protection. Under Art. 4(1)(c) of the former GDPD, EU privacy law only applied if the data controller used equipment located in the EU, thereby relying on a territoriality doctrine. GDPR broadened its territorial scope by switching to an effects-based doctrine.

2.3 Consent

If a website uses web technology—such as cookies, scripts, images, web trackers or fingerprinting technologies— to identify users, it needs to gather the consent of the users whose personal data it processes (Art. 6(1)(a) GDPR).[2] GDPR is not the first European privacy law that had created a consent requirement concerning web analytics technologies. Since 2009, the revised e-Privacy Directive (European Union, 2002) had required websites to gather the consent of users for placing cookies on their devices (Article 5 (3) of the revised e-Privacy Directive).[3] However, the drastic expansion of the sanctions regime under GDPR and the broad expansion of the territorial reach of GDPR acted as a game-changer. With its broad, technology-neutral consent requirement concerning web technology, GDPR gave European privacy authorities a powerful weapon to enforce the consent requirement across the EU and beyond.

2.4 Joint Responsibility

Websites often use third-party services for their behavioral monitoring and consumer profiling. GDPR recognizes this by distinguishing between *data controllers*, which determine the purposes and means of the processing of personal data, and *data processors*, which process personal data on behalf of the controller (Art. 4 no. 7, 8 GDPR). Third-party web technology services can either be dependent data processors, or they can be data controllers themselves if they have sufficient control over the purposes and means of data processing. Under GDPR, websites cannot easily dispose of their liability for privacy violations by outsourcing web technologies to third-party providers. This is not only because the website must inform its users about any transfer of personal data to third-party providers (Art. 13

[2] GDPR also provides other legal reasons for processing, apart from user consent. In particular, if the website has a prevailing legitimate interest to process personal data— e.g., IT security measures or fraud prevention—Article 6(1)(f) GDPR allows the website to process such data without the explicit consent of the user (see also Recitals 47 and 49 GDPR). For the types of technologies that are of interest for our paper, user consent is usually required under GDPR.

[3] Note that Art. 5(3) e-Privacy Directive concerns the *storing* of a cookie on a user's device, while Art. 6(1)(a) GDPR concerns the *processing* of personal data that has been gathered through cookies or other means. On the complex relationship between GDPR and the e-Privacy Directive, see European Data Protection Board (2019b).

(1)(e) GDPR) and because the website must make sure that the third-party providers' data processing will adhere to GDPR as well (Art. 28 (1) GDPR). The website may also be jointly responsible, together with the web technology provider, for violating GDPR. In recent case law, the European Court of Justice has held that websites and web technology providers can be responsible as "joint controllers" (Art. 26 GDPR) even if the website cannot control what personal data is transmitted to or processed by the web technology provider.[4]

This legal framework of joint responsibilities can have grave consequences for websites. The websites are responsible, vis-à-vis data protection authorities, for the joint data processing with the web technology provider: it is, for example, the websites and not only Google or Facebook, which have to gather user consent to process personal data. Also, users can sue the website for damages caused by GDPR violations committed by the web technology provider within the joint responsibility framework.[5] While websites may be able to outsource tasks such as analytics to third-party providers, they are—to a considerable extent—still responsible and potentially liable for privacy violations that occur while cooperating with the technology provider.

Furthermore, if a website decides to use a web technology provider that is located outside the EU, the website must ensure that the data transfer out of the EU is in compliance with GDPR. If the personal data is transferred to a country that provides an adequate level of protection as determined by the European Commission,[6] such transfer is allowed (Art. 45 GDPR). Transfers to other countries are possible as well, but the website carries much higher compliance costs and liability risks, as it has to ensure that adequate privacy safeguards and rights are available (Art. 46 GDPR).

2.5 Increased Legal Uncertainty and Compliance Risks

GDPR has put compliance with privacy laws on the agenda of the highest echelons of corporate management not only because of its stricter enforcement regime and broader territorial scope, but also because GDPR has led to novel problems in interpreting the text of GDPR and its relationship

[4] On the contours and limits of this case law, see European Court of Justice, June 5, 2018, case C-210/16, ULD Schleswig Holstein v. Wirtschaftsakademie Schleswig-Holstein ("Fan Page"); July 29, 2019, case C-40/17, Fashion ID v. Verbraucherzentrale NRW ("Fashion ID"). In *Fashion ID*, the European Court of Justice established a joint controller responsibility between a website that used a Facebook Like button and Facebook. The Court stressed that the joint responsibility is limited to those steps of the data processing that were jointly determined.

[5] Such liability does not exist if the website can prove that it was not responsible for the violation. However, the website may often share some responsibility, or it may have difficulties proving otherwise in court (Articles 26(1), (3), 82(3), (4) GDPR). Also, the joint liability does not extend to administrative fines.

[6] The European Commission determines whether a country outside the EU offers an adequate level of data protection. Major countries that the European Commission has so far recognized include Argentina, Canada, Israel, Japan, Switzerland, and the U.S. (limited to the Privacy Shield framework).

to other bodies of EU or member state privacy laws. It is very difficult, for example, to observe a standard pattern on how European data protection authorities calculate fines or how the territorial scope of GDPR is determined in individual cases (Tobin, 2019). These problems in interpreting novel legal texts can often only be resolved by courts which develop case law within their judicial hierarchy. As this process typically takes several years at least, firms are left with increased legal uncertainty and considerable compliance risks in the meantime.[7]

2.6 Web technologies

Publishers on the internet can access a plethora of third-party services that enable personalized functionalities, help publishers understand the characteristics and behavior of their audience, as well as the services connecting publishers to advertisers. Many interactions between publisher and third-party web technology providers happen in real-time as a user visits the publisher's website. These interactions can be observed via the HTTP requests that a website makes as content is loaded. Using HTTP requests, websites can display content that is controlled by other parties, and this content can differ for each request and each user. The best example of such content is, of course, advertisements. However, a website can also embed contents from third parties that are not, or not directly, related to advertising. For example, a website may embed externally hosted video content or outsource certain interactive functionalities, such as comment or chat systems, to specialized third-party providers.

Figure 1 illustrates a stylized example of third-party requests: the page hosted on `techcrunch.com` makes requests to third-party domains such as `google-analytics.com` and `wordpress.com`, embedding content that is hosted on third-party servers, such as scripts and images.

While there are many different types of third-party services that publishers can choose to interact with, as a technical principle, all third-party services that are called in real-time as a user navigates to the website can be considered a privacy-intrusive *tracking technology*. For example, it is always possible for a service to trace back the IP address that originated the HTTP request (i.e., the IP address of the user), or to respond to a request by sending a cookie to be stored on the user's machine, which would allow the service to recognize the same machine in the future, even if the IP address has changed. Depending on the business model of the third-party service, such user-level browsing information is stored and shared with others, including the first-party publisher. For the publisher, this is convenient

[7] According to industry analyses, compliance costs for large U.K. firms (FTSE 350) were $1.1 billion, and $7.8 billion for large U.S. firms (Fortune 500). See https://www.forbes.com/sites/oliversmith/2018/05/02/the-gdpr-racket-whos-making-money-from-this-9bn-business-shakedown.

as it can access a dashboard showing aggregated information on users' journeys within its website. For the third-party service, this technology is even more advantageous as it allows to track consumers across other websites that use the same service. Prominent examples of across-publisher tracking services are *Google Analytics*, but also the "like", "share" and "login" buttons operated by *Facebook* and *Twitter* (Chaabane et al., 2012; Roosendaal, 2012). These data are processed and (directly or indirectly) made available to advertisers or advertising networks, which use it to predict the "value" of a user, and which consequently affects their willingness to pay for the user's eyeballs.[8]

It is important to note that firms are often vertically integrated and assume multiple roles at the same time. For example, an advertising network can collect data by tracking consumers, use this data to enhance its prediction models, but also make information that is inferred from this data available to publishers and advertisers. These interactions are often not directly visible to the consumer, nor are the identities of the involved parties that collect and trade personal information completely transparent.[9]

Historical information on the third-party requests of websites is publicly available. It can be matched with other publicly available sources of meta-information on third-party web technology services.

Over the last 20 years, the web technology service industry has grown by a factor of 50 in terms of the number of distinct services on the market, and by a factor of 4 in terms of the median number of distinct services used per website (Lerner et al., 2016). Today, about 88% of the top 1 million most popular websites operate with at least one third-party web technology service, and out of those, the average site requests about 9 distinct third-party domains (Libert, 2015). At the same time, the industry has become considerably concentrated with the top 20 services covering about a third of the market (Gill et al., 2013; Schelter and Kunegis, 2018).

3 Data and Methods

In the following, we describe the public and proprietary data sources, which we combine to a unique dataset that enables us to study the behavior of websites, web technology providers, and consumers after GDPR was enacted.

[8] For example, Facebook states in a 2015 press release: "Last year, we introduced online interest-based advertising – ads based on people's use of other websites and apps [...]. For example, with online interest-based ads, if you visit hotel and airline websites to research an upcoming trip, you might then see ads for travel deals on Facebook.", see https://www.facebook.com/notes/facebook-and-privacy/a-new-way-to-control-the-ads-you-see-on-facebook/926372204079329/.

[9] In fact, several software tools, which come as extensions to the browser, are available to help consumers visualize and understand this data. Some tools, such as *Ghostery*, *Privacy Badger*, and *Disconnect*, offer the ability to block requests to specific third-parties.

3.1 Main dataset: HTTPArchive

We follow the data collection approach in Batikas et al. (2019) to access historical information about the HTTP requests that websites make to third parties. We use data from *HTTPArchive*, a project that periodically crawls the homepages of about half a million hosts. With a few exceptions, these data are available on a biweekly basis, stemming from a crawl at the beginning and in the middle of the month.[10] We will mostly use data for the period from May 2017 to November 2018, but we also study data starting in May 2016 as a robustness check.[11] The long time structure will be useful to control for pre-existing trends that might confound the estimate of the effect of GDPR, and to study short-run and long-run dynamics.

The level of observation is a "website-host", e.g. `subdomain.domain.com`, where the subdomain is `www` in the vast majority of cases.[12] We define a balanced sample of hosts that we can observe in each available data snapshot, which results in 110,706 hosts that we follow over 33 points in time – 21 data points before the introduction of GDPR and 12 data points after. Hence, our preferred specification includes 3,653,298 observations. We collect information about the identity of third parties that a host sends HTTP requests to, the total number of third-party requests, the number of third-party cookies, and the number of first-party cookies. In appendix A, we report the SQL code we use to extract a balanced panel of website-hosts and their HTTP requests using the publicly available database in Google BigQuery. This code includes a precise definition of "first-party," "third-party," and "cookie.".

3.2 Website-level meta-information

Our preferred measure of whether a website-host caters to an EU audience is based on the country-specific top-level domain (TLD). A second way to determine whether a website-host caters to an EU audience is to extract meta-information from the website. The idea is that the language of the website's content should be an indicator of the target audience. Excluding English, there are 23 official languages in the EU. HTTP response headers sometimes contain information on the language in which content is encoded. We use *HTTPArchive* to extract that information from the website we analyze.[13]

We add information on the country-specific demand for the websites which we observe in the *HTTPArchive*

[10] The public dataset, available in Google's BigQuery web service, does not contain information for 12-01-2016, 01-01-2017, 01-15-2017, 04-15-2018, and 05-01-2018.

[11] The earliest snapshot is 05-01-2016 and the latest snapshot is 11-15-2018.

[12] For the empirical analysis in this paper, we call `subdomain.domain.com` a *website-host* and `domain.com` a *website*.

[13] See Appendix A for a precise definition in the SQL code we use to extract this information from the publicly available database in Google BigQuery.

data from another source. From *Alexa*, we obtain the rank by country (based on page views, as of October 2019), if available, for every country in the world. Alexa does not have this information for all the websites in our sample, but we can add this information for 42,764 websites. We define a website as catering to an EU audience if it appears in Alexa's ranking in at least one EU country.

3.3 Consumer-level meta-information

Our website-level information on the popularity of websites per country is complemented by consumer-level popularity metrics. We obtain proprietary clickstream data from *Nielsen*'s Internet audience measurement service *NetView*. This service monitors the online activity of a large number of Internet users by recording all their URL visits via an application that is installed on the users' device (desktop PC or Mac) and operates in the background. Consumers are incentivized to take part and stay in the panel by a rewards program, in which they can exchange credit points for retail and travel vouchers. Consumers earn a fixed amount of credit points every month, and take part in a lottery every quarter.[14] Upon signing up, *Nielsen* requires participants to fill out a survey about basic demographics, such as income, age, and gender. This data lets us follow the top 1000 most visited domains of users in Germany and the United States, each month from May 2017 to November 2018. We have access to fine-grained aggregations of user types. We can distinguish between 84 user groups, comprised of combinations of gender (2 categories), age (6 categories), and income (7 categories). For each user group, website, and month, we observe the number of unique visitors and page views. The combined dataset, in which we match clickstream data and website-level information from *HTTPArchive*, covers 1,298 websites over 18 points in time (we use *HTTPArchive*'s first snapshot each month).

3.4 Vendor-level meta-information

For our empirical analysis, we are also interested in the location of third-party web technology providers. Since GDPR came into effect, domain registration data has not been publicly available anymore.[15] As a proxy to the web technology provider's location, we use its country-specific TLD (Batikas et al., 2019). For example, we define third-party domains to be more likely located in the EU when they use an EU-specific TLD like `de` or `fr`.

We further obtain meta-information about the domain names of third-party web technology providers in the *HTTPArchive* from *whotracks.me*. The company behind *whotracks.me* offers the browser *Clickz*

[14] See https://digitalvoice.nielsen.com/us/en/home.html.
[15] For the impact of GDPR on the Whois database, see https://www.icann.org/dataprotectionprivacy.

and browser plugins such as *Ghostery*, and is therefore able to record the HTTP requests made by websites as they are visited by users of their browser or their browser extension. Some of these users are located inside the EU, some outside the EU. The publicly available dataset is an aggregation of this individual-level data by website, month and region (Karaj et al., 2019). Most importantly for us, *whotracks.me* also provides a categorization of third-party domains into service types, e.g., advertising, social media, analytics, etc., and a match between third-party domains and the firms that operate them. The database of *whotracks.me* is extensive, but not all third-party domains that we observe in *HTTPArchive* are also included there. We consider third-party domains that do not appear in *whotracks.me* as standalone companies.

Data from *Evidon*'s industry directory additionally gives us information on the privacy policies of a subsample of trackers. *Evidon* is a company that specializes in services for the web industry, providing technology and services for collecting user consent and the monitoring and auditing of online advertising. Most importantly for us, they offer a database that includes meta-information about thousands of companies active in the web technology industry. With this database, we can observe whether a web technology provider discloses whether it collects and shares data with third parties. We can distinguish between anonymous, pseudonymous, aggregate, personally-identifiable (PII) and sensitive (relating to personal financial or health information) data. We have access to two snapshots of Evidon's data, obtained in September 2017 and July 2019.

3.5 Econometric model

We estimate a set of difference-in-differences models, beginning with:

$$\begin{aligned} Log(Number3rdParties_{it}+1) = \alpha + \beta_1 Trend_t + \beta_2 \left(Trend_t \times EU\text{-}Audience_i\right) \\ + \delta_1 Post_t + \delta_2 \left(Post_t \times EU\text{-}Audience_i\right) + \mu_i + \varepsilon_{it}, \end{aligned}$$

where $Number3rdParties_{it}$ is the number of third-party domains that website i requests at time t. $Post_t$ indicates the period after GDPR came into force on May 25, 2018. We include group-specific linear time trends and website-fixed effects μ_i. The error term ε_{it} has the standard assumptions, and we report estimates clustered on the website level. We later adapt this model along a variety of dimensions, mostly changing dependent variables, but also adding period-specific time trends that allow us to make predictions about long-run changes.

4 Findings

4.1 Interactions with third-parties and cookies

4.1.1 Requests to third-party domains

The left panel of Figure 2 shows the average log number of third-party requests per EU- and non-EU website. We see a substantial and sudden drop in the number of requested third-party domains just after the introduction of GDPR for websites with EU audiences but also for websites with non-EU audiences. In Table 1, controlling for website-host fixed effects and a linear time trend, we estimate that this reduction is -8.1% (EU) and -2.4% (non-EU).

Robustness checks and potential limitations *HTTPArchive* has introduced several changes in the technology they use to collect information on HTTP requests, and to the list of website-hosts about which they collect this information. This reduces the number of websites that we can consistently observe over a long enough period. Because four snapshots have an atypical distribution of key variables, most likely caused by measurement error, we exclude them from our sample. These snapshots are 05-01-17, 08-01-17, 08-15-17 and 09-01-17. However, as we show in Table 2, including those observations yields qualitatively similar results.

We find similar point estimates (most of them not different in a statistically significant way) when we use different ways to define websites as catering to EU versus non-EU audiences, as reported in Table 3. Specifically, we compare (1) websites that make it to Alexa's ranking in at least one EU country to those that do not, (2) websites that send at least one HTTP header encoded in one of the official languages of EU countries (excluding English) to those that do not, and (3) websites that are only visited by German users versus websites that are only visited by users in the United States according to our clickstream data.

One might be worried that unobserved confounding factors introduce a bias to our point estimates. Specifically, a concern could be that other factors that drive websites' choices to make requests to third-party domains simply coincide with the introduction of GDPR. In columns (1)–(3) of Table 4, we use data from one year before the observation period of our baseline specification. We then estimate by how much the number of requested third-party domains changes after May 25, 2017. The results suggest that websites that cater to EU audiences increased the number of third-party domains by 0.7%, and websites that cater to non-EU audiences decreased the number of third-party domains by 0.2%. The fact that these changes are in the opposite direction and very small suggests that time-

dependent unobserved factors are not likely to bias our main results to an economically sizable extent. A second placebo exercise, in column (4), looks at a specific set of websites. We observe 57 "piracy" websites that offer unlicensed content (music, movies, live sports streams) as identified by blacklists and reports of the IT security and entertainment industries (Batikas et al., 2019). For these websites, which primarily exist because it is difficult to enforce copyright law as the underlying organizations are hidden, we would expect that it is equally difficult to enforce GDPR. Hence, we should not see a reaction. It turns out that we cannot reject this hypothesis. Although the point estimate is −2.8%, the standard error is large.

A potential limitation of our data could be that *HTTPArchive* scrapes websites from servers that are located in California. Unlike in related studies (e.g., Johnson and Shriver, 2020), *HTTPArchive* does not allow to observe whether a website issues different HTTP requests depending on the geolocation from where it is accessed. We therefore also run our analysis with similar data from *whotracks.me*, which lets us observe the third-parties requested by websites when accessed by users located in different EU member states and the U.S. There are some caveats, however. First, the dataset covers only one month before the introduction of GDPR. Second, it only covers websites that are visited by the users that *whotracks.me* can track. This limits us to study a sample of 775 websites that we can observe for 14 months between April 2018 and May 2019. The results in Table 5 show that websites visited by users in the EU (U.S.) change the number of requested third-parties by −16.7% (−16.6%). As indicated by the coefficient $Post \times EU$, this difference is not statistically significant. Hence, we conclude that, at the level of our analysis, geotargeted websites should not be an issue. While websites may display geo-targeted ads to users located in different parts of the world, the *structure* of the websites should not be affected by geotargeting. Further, even if a website would trigger a different number of HTTP requests for EU than for non-EU users, our findings would be a lower bound. This is because the HTTP requests in a website version that is catered to EU users should change more strongly after the introduction of GDPR than the non-EU version of the website, which we can observe.

4.1.2 First-party and third-party cookies

The right panel of Figure 2 shows the average log number of third-party domains that respond to a website's request by sending a cookie. There is a sharp decrease after the introduction of GDPR. The opposite holds for first-party cookies, which usually do not involve sharing information with third parties. The results reported in columns (1)–(3) of Table 6, coming from a specification that

uses our preferred TLD-based definition of EU audience, suggest that the reduction in the number of third-party domains that send cookies is 12.8% (EU) and 5.5% (non-EU). This is significantly lower than the estimates we get for the overall number of requested third-party domains (in the sense that 90% confidence bands overlap). Looking at the number of first-party cookies, we find an increase for websites catering to EU audiences of 1.7% and an increase for websites catering to non-EU audiences of 2.5%. The increase in first-party cookies can perhaps partially explained by a strategic response of the web technology industry to *Apple*'s *Intelligent Tracking Prevention* policy, announced in June and introduced in November 2017, which effectively banned third-party cross-site tracking cookies from Apple devices.[16] For example, Google (in September 2017), Microsoft (in January 2018) and Facebook (in October 2018) have adjusted their tracking technology such that they can "hide" behind the domain of the first-party website when issuing cookies.[17] While the timing of these reactions seems consistent with the steady increase in first-party cookies before and after the introduction of GDPR, these strategic changes are perhaps less related to the sharp discontinuity around the introduction of GDPR that we also observe in the data.

4.2 Compliance

4.3 Short-run changes and long-run trend

The right panel of Figure 3, depicting predictions from a model with group- and period-specific time trends shows that only four months after GDPR, websites with non-EU audiences have rebounded to the level directly before GDPR. Websites with EU audiences revert to their initial level after 22 months. The middle and right panel of Figure 3 suggest that the changes in cookies, which by construction include personal data, might be more sustainable. We see both a downward shift in the number of requests with third-party cookies and a subsequent reversion of the trend. The number of first-party cookies shifts upwards and continues to grow with a steeper slope.

4.3.1 Geolocation of third-party vendors

We are now interested in whether websites' reduction in requests sent to third-party domains varies with the jurisdiction the third-party operates in. If a website cares about GDPR compliance, it may prefer to use an EU-based web technology provider, given the joint responsibility risks between data

[16] See https://webkit.org/blog/7675/intelligent-tracking-prevention/ for a description of ITP

[17] See for technical details regarding Google https://searchengineland.com/google-analytics-adwords-response-apple-intelligent-tracking-prevention-282233, Microsoft https://about.ads.microsoft.com/en-us/blog/post/january-2018/conversion-tracking-update-on-bing-ads, and Facebook https://www.adweek.com/programmatic/the-facebook-pixel-will-add-a-first-party-cookie-option/.

controllers and processors described in section 2.4. We approximate the location of a third-party web technology provider by country-specific TLDs. Columns (1) of Table 7 shows that, after the enactment of GDPR, the share of requests to third-party domains with EU-TLDs increases significantly, not only for websites that cater to an EU audience but also for websites that do not. The point estimates are an increase of 0.34 percentage points (EU websites) and 0.03 percentage points (non-EU websites). In column (2), we repeat this exercise but look at how the share of requests to third parties with TLDs of countries that the EU considers to provide adequate privacy protection changes with the introduction of GDPR. The results suggest that websites decrease the share of trackers with TLDs from these countries by 0.02 percentage points (no significant difference between EU and non-EU websites). Finally, in column (3), we look at the share of requests to third parties with TLDs from all other countries. Here we see a much stronger decrease for EU websites (0.33 percentage points) than for non-EU websites (0.02 percentage points). Hence, these results suggest that especially websites that cater to EU audiences increase their requests to third-party services in the EU, and reduce their requests to third-party services outside the EU, but much less so regarding services that operate in countries that the EU considers to provide adequate privacy protection.

4.3.2 Privacy policies of third-party vendors

We now turn to an analysis of how behavior of third-party web technology providers and the market for these technologies as a whole changes with the introduction of GDPR.

We analyze whether firms in the market for web technologies change their privacy policies after GDPR became effective. We have access to detailed information on the data collection and sharing policies of 3,993 firms via the industry database provided by *Evidon*, before and after GDPR comes into force. On the extensive margin, column (1) of Tables 8 and 9 show that web technology providers are more likely after GDPR to disclose publicly whether they collect any data at all, and whether they are sharing any data with third-parties. On the intensive margin, we show in Table 8 that web technology providers that have already disclosed their data collection policy before GDPR (column 2) are less likely to state that they collect personal data after GDPR. Web technology providers that only disclose their data collection policy after GDPR (column 3), however, are more likely to state that they collect personal data. In Table 9, we show that web technology providers are more likely to state that they share personal data with third-parties, independent of whether they have already disclosed their data sharing policy before GDPR (column 2) or whether they start to do so after GDPR (column 4). In

columns (3) and (5), we additionally show that web technology providers that state that they collect personal data are also more likely to state that they share personal data with third parties. Before GDPR, we do not find a significant relationship between collection and sharing of personal data. In column (1) of Table 10, we show that web technology providers are more likely to disclose any data retention policy after GDPR. In column (2), we find that those that only disclose a data retention policy after GDPR are more likely to state that theu store data as long as necessary to fulfill a business need or a required by law. In column (3), we do not find a significant increase in the stated retention period in months, conditional on disclosing any data retention policy.

As a next step, we are interested in whether web technology providers that collect or share personal data have lost market share with the introduction of GDPR. We calculate market shares as $s_i = n_i / \sum_{k=1}^{N} n_k$, where n_i is the number of EU/non-EU websites that send requests to third-party domain/firm i (reach), distinguishing between third-party domains that state to collect/share personal data. The left-hand panels of Figure 4, show how EU and non-EU market shares of third-party services that collect (upper panel) and share (lower panel) personal data evolve. There are trends visible, but no clear discontinuities that coincide with the introduction of GDPR. The difference between EU and non-EU markets appears to be relatively stable, although there seems to be a change after the enactment of GDPR in that EU market shares decrease more than non-EU market shares. The right-hand panels of Figure 4 show how the EU and non-EU market shares of third-party services that do not collect (upper panel) and do not share (lower panel) personal data evolve. The patterns are not clear, but it appears that the gap between EU and non-EU market shares has widened as far as third-party services that do not collect personal data are concerned. In contrast, there are no substantial changes visible regarding those that do not share personal data.

Table 11, showing results from regressions models that control for pre-existing trends, confirms that there is no statistically significant difference in market shares between third-party web technology providers with privacy policies that include collection or sharing of personal data and those that have different privacy policies. Much like in the upper right panel of Figure 4, we find a significant increase in the EU market share of services that collect personal data. The point estimate is 0.0006 percentage points, which is a 1.6% increase relative to the sample average.

4.3.3 Website popularity

So far, we have shown that websites decrease their connections to third-party web technology providers after GDPR. We are also interested in whether the reactions of websites depend on how popular the websites are. For example, the average reduction in the number of requested third-party domains could hide heterogeneity based on the popularity of websites. If not all websites reduce their requests to third-party domains by the same amount, and, for example, more frequently-visited websites reduce their requests to third-party domains to a much larger extent, this would mean that we would have so far underestimated the overall change.

To shed some light on this question, we turn to an analysis at the consumer level. Figure 5 depicts the average likelihood of encountering at least one third-party cookie when browsing the top 1000 most visited websites in a given month (the popularity being measured by using *Nielsen*, the third-party cookies being measured by using *HTTPArchive* data). We distinguish between consumers in the United States and consumers in Germany. The left-hand panel shows the unweighted average, where we again observe a sharp decrease as GDPR comes into force. This is true for either US and German users, but the difference between US and German users seems to get smaller as well, perhaps suggesting that the change is more pronounced for German users. The right-hand panel shows an average weighted by websites' total audience. Here we do not observe this sharp discontinuity and a perhaps more general downward trend. The difference between US and German users seems to remain stable.

We follow up on these observations more carefully in Table 12. Controlling for group-specific linear trends and holding time-invariant unobserved user-group-specific variation fixed in column (1), we estimate that the share of visited websites that send requests to third-parties with cookies decreases by 1.9% (German users) and 0.9% (US users). We now turn to a linear probability model, which additionally holds time-invariant website-specific variation fixed. In column (2), we find a reduction in the likelihood of receiving at least one third-party cookie in a given month of 1.7% (German users) and 0.8% (US users). In columns (3) and (4), we add interaction terms, measuring the size of the website's audience. We find that consumers visiting more popular websites have a significantly lower likelihood to received third-party cookies after GDPR comes into force. However, this is only true for German users. The point estimates suggest that the change for German users varies between -0.4% and -4.2% when we consider the websites with the least and most pageviews in our sample. The

variation in the change for US users is much smaller, ranging between −0.8% (least page views) and −1.0% (most page views).

4.3.4 Privacy-sensitive and young consumers

In addition to website popularity, websites with specific demographic target audiences could have also changed their third-party requests differently after GDPR came into force.

Two groups seem particularly interesting in the context of GDPR's consent requirement: consumers with higher preferences for privacy and consumers who are too young to give legally binding consent. We follow Avi Goldfarb and Catherine Tucker (2012) and categorize consumers as more privacy-sensitive if they choose the option "Prefer not to say" when asked about income in Nielsen's survey. Since we observe users in the age group 2–17, we can directly distinguish between adults and children/adolescents.

In Table 13, we report results for the share of websites that are connected to third-party services sending cookies (columns 1–3), and for the likelihood of receiving at least one third-party cookie when navigating to the top 1000 most visited websites in a given month (columns 4–6). We find strong evidence that young German users experience a three times larger decrease in the share of websites with third-party requests (−4.3% versus −1.5% for Germans of age $\geq$ 18), and a 1.5 times larger decrease in the likelihood of receiving a third-party cookie. Young users in the U.S. are not affected differently than older users in the U.S. Our results do not suggest that users with stronger preferences for privacy are affected differently. We only see differences concerning the share of websites that send third-party cookies in column (3). The share has decreased by 1.6% for privacy-sensitive users in the United States, which is almost two times larger than the decrease of other US users (−0.9%). In contrast, the share of websites sending third-party requests has decreased three times less for privacy-sensitive users in Germany than for other users in Germany (−0.5% versus −1.5%).

These results are, of course, a reflection of the types of websites that different user groups tend to visit. In Figure 6, we plot a concentration index for young, privacy-sensitive, and all other consumers. This shows that demand is much more concentrated for young users, less so for privacy-sensitive, and least so for all other consumers. Hence, young consumers visit the most popular websites (e.g., YouTube) relatively more often than privacy-sensitive and other consumers. Together with the results in Table 13, where we show that the change coming with GDPR is stronger for popular websites, this can explain our findings.

4.4 Market structure and competition

4.4.1 Concentration in EU and non-EU markets

We now turn to investigate whether the market structure in the web technology markets has changed after GDPR. Using metadata from *whotracks.me*, we observe the ownership structure of third-party domains. For example, *Google* operates a variety of different domains, including less obvious domains such as `doubleclick.net`, `invitemedia.com`, and `2mdn.net`. In the left-hand panel of Figure 7, we show how the Herfindahl-Hirschmann Index (the sum of squared market shares of all firms) develops over time. We define markets as according to the market of websites that cater to EU audiences and the market of websites that cater to other audiences. The HHI for the non-EU market follows an increasing trend over time, with no clearly visible change as GDPR comes into force. In contrast, there is no trend visible in the HHI for the EU-market before GDPR, but we see a large increase after the introduction of GDPR. This is particularly true in the short run, but there seems to be an increasing trend. At the end of our observation period, the EU market has caught up with the non-EU market.

In the right-side panel of Figure 7, we plot the hypothetical market structure when we exclude the firm with the largest market share. In both EU and non-EU markets, the largest firm is *Google*. The dominance of Google in this market becomes apparent when we compare the scales of the left-hand and right-hand panel. The HHI including Google varies roughly between 900 and 1000, whereas the variation in the hypothetical HHI excluding Google is only roughly between 60 and 80. We see a downward trend in both markets, with no clearly visible change in the non-EU market, but a steep decrease in the EU market after GDPR comes in to force. Following GDPR, we find that the market for web technologies becomes more concentrated, with the HHI index increasing by 3%. In Table 14 we show that this result is driven by changes in Google's market shares (columns 1–2), not by those of other firms (columns 3–4). With the caveat of a small sample size that naturally comes with aggregation to the market level, these regressions suggest that the HHI in the EU-market has significantly increased by 26.3 (30.4) points from a base of 911.3 (914.5) points or 2.9% (3.3%). We do not find that the concentration in the non-EU market has significantly increased. The point estimates of *Post* are also much smaller than the point estimates of *Post* × *EU-based Websites*. Turning to the hypothetical HHI in columns (3) and (4), we find a decrease of 5.5 (5.3) points from a base of 74.1 (74.2) points or 7.4% (7.1%). This suggests that the increase in overall market concentration seems

to be explained solely by an increase in the market of the dominant firm.

4.4.2 Market shares

To provide a more detailed picture of the results we see on the aggregated market-level, we now turn to a disaggregated analysis of market shares. Figure 8 plots changes in market shares of the ten firms with the largest increases ("winners") and of the ten firms with the largest decreases ("losers") in the EU and non-EU market, two weeks before and two weeks after GDPR comes into force. The same exercise is carried out in Figure 9, but here we look at the entire observation period of six months before and six months after the introduction of GDPR. While the firms that lose the most market share vary across time horizons and markets, Google sticks out as the clear winner in the short run and the long run, both in the EU and the non-EU market.

These insight carry over in a regression analysis. In column (1) of Table 15, we show that EU and non-EU market shares of third-party domains do not change on average after GDPR comes into force. In columns (2)–(4), we aggregate the data to combine all trackers of a web technology provider in our calculation of market shares. Again, we find that EU and non-EU market shares of web technology providers do not change on average after GDPR comes into force. When we specifically compare the two web technology providers with the highest market share (Facebook and Google) to all other firms in column (3), we see that point estimates dramatically increase. However, the estimated coefficients are not significantly different from zero. In column (4), we compare the firm with the largest market share (Google) to all other firms and see that Google's market share has increased significantly. Regarding the non-EU market, our estimates suggest an increase of 0.1 percentage points, and regarding the EU market, we find an increase of 0.5 percentage points. Relative to Google's pre-GDPR market share with non-EU websites, and their pre-GDPR market share with EU websites, this is an increase of 0.3% and 2.2%, respectively. It is difficult to attach a monetary value to these market share numbers because detailed enough data on revenues is not publicly available. For an illustrative back-of-the-envelope calculation, however, we collect information from Google's parent company's and Facebook's quarterly filings with the U.S. Securities and Exchange Commission. Alphabet reports Google advertising revenues and disaggregated total revenues (of which Google ad revenues are about 86% in the third quarter of 2018) by world regions. Europe, Middle East and Africa are treated as one region (EMEA). Assuming that EMEA's share of Google ad revenues is the same as EMEA's share of Alphabet's overall revenues, we extrapolate Google ad revenues in EMEA over time. Facebook

reports advertising revenues separately for Europe and other regions of the world. In a regression where we look at the log of advertising revenues and where we compare the intercept of the before to the intercept of the after period, controlling for quarter fixed effects, and with the caveat of a very small sample size, we do not see a significant change for Google (the coefficient is -3% in EMEA and outside). Facebook, however, loses significantly (-12% outside and -14% in Europe). This analysis is of course not as rigorous as it should be, but the patterns are at least consistent with the findings in our paper.[18]

Finally, we further differentiate between submarkets in which Google is active in Table 16. This is helpful to understand in which service categories Google is mostly gaining market share after GDPR comes into force. Using a slight variation of the classification from Whotracks.me, we distinguish between *Video/Audio Players* (column 1), *Advertising* (column 2), *Tagmanager* column 3), *Analytics* (column 4), *CDN/API* (column 5), and *Other/Unknown* services in column 6. Reinforcing the results from Table 15, we only see significant changes in the market shares of Google, and not for other firms. First, looking at changes in Google's market share concerning non-EU websites, the largest increase is 1.9% concerning analytics, whereas the *Other/Unknown* category is suffering the largest decrease with −16.6%. Note, however, that Google's market share in the latter category was already less than 0.25% before GDPR. Concerning EU websites, we find substantial decreases in the audio/video player market (−7.5%), the tag manager market (−8.0%), and the CDN/API market (−1.4%). We find the largest increases in the analytics market (7.2%) and the advertising market (5.4%). Hence, we conclude that the average effect of 0.3% increase in Google's overall market share with non-EU websites and the average effect of 2.2% increase in Google's overall market share with EU websites in Table 15 hides significant heterogeneity. At least in the market for EU websites, Google mainly becomes stronger in the markets in which they had been strongest before GDPR (advertising and analytics) with market shares of 25.8% and 38.4%, respectively.

5 Discussion

5.1 Privacy law and dynamic compliance risks

Given the risks created by GDPR's joint responsibility regime, if websites cannot accurately assess the privacy risks at their web technology provider, the best way to reduce their exposure is to reduce their use of web technology providers, i.e. by insourcing behavioral monitoring and consumer profil-

[18] Due to their speculative nature, we do not report these results in more detail here, but they are available on request.

ing. Similarly, moving from third-party to first-party cookies reduces compliance risks for a website, especially for websites catering to particularly privacy-sensitive consumer groups.

If a website still needs to use an outside provider, it can reduce its compliance risk by switching to an EU-based provider. A website may, on average, have higher confidence that an EU-based web technology provider is GDPR-compliant compared to non-EU-based providers. Further, a website may choose large web technology providers over small ones because these may have more resources to weather legal challenges created by GDPR. By choosing a large web technology provider, a website may reduce its own compliance risk as well.

We find that the number of requested third-parties rebounds after 4–24 months. This is in line with a compliance risk interpretation: After GDPR initially significantly increased legal exposure due to legal uncertainty, websites had an interest in adopting a cautious approach when interpreting GDPR. However, over time this uncertainty decreased as new policy guidelines (e.g., by the European Data Protection Board[19]) were released. Websites may have learned to adapt by interpreting such guidelines, seeking legal counsel and observing competitors. As a result, the initial caution regarding GDPR decreased over time, and websites have become more active in generating third-party requests again. In contrast, we find a persistant reduction in third-parties with cookies. Given that third-party cookies typically include personal data at least by identifying a browser and device, this seems in line with the goal of data minimization. Further, we estimate that the reduction in third-parties with cookies is about 2–7 times larger in percentage terms (4–15 times larger in absolute terms) than the increase in first-party cookies. Over a longer run, we predict that the rate of decrease in third-parties with cookies is higher than the rate of increase in first-party cookies—but only for websites targeted at EU audiences.

5.2 Privacy law and trade policy

We observe that even websites catering to a non-EU audience reduce their use of third-party web technology providers after GDPR. Non-EU websites are moving from third-party cookies to first-party cookies, and they increasingly rely on web technology providers in the EU.

These changes, while rational for websites subject to GDPR, seem puzzling behavior by websites not operating under GDPR. They become understandable, however, if one considers the broad territorial application of GDPR. Under general principles of international public law, the EU cannot regulate

[19] See https://edpb.europa.eu/our-work-tools/general-guidance/gdpr-guidelines-recommendations-best-practices_en.

the processing of personal data that takes place outside of and is not related to the EU. Yet, the EU has expanded the *de facto* territorial reach of European privacy laws well beyond the geographical boundaries of the EU. As complying with GDPR is costly, some global technology companies have decided to apply GDPR to all their consumers worldwide, even though GDPR does not require them to do so. These companies save costs by not having to offer two versions of their products and services (for persons in and outside the EU).[20] In addition, GDPR has served as a role model for other countries to adopt similar privacy regimes (Bradford, 2020, 143–144).

The literature on international regulatory competition has developed a general theory that the EU has *de facto* externalized several of its strict regulatory laws outside its border through a combination of market mechanisms and unilateral regulatory globalization. European privacy law is one example of the *Brussels effect*, according to Bradford (2012, 2020), in addition to European antitrust, health and environmental law (see also Goldstein and Wu, 2006 and Schwartz, 2019). Our study offers one of the first empirical explorations of this argument in European privacy law.

The EU is not the only jurisdiction whose privacy laws may have extraterritorial reach (Rustad and Koenig, 2019). The California Consumer Privacy Act (CCPA)—which became effective January 1, 2020, and incorporates several GDPR concepts such as the rights of access, portability, and data deletion—applies to all for-profit entities doing considerable business in California. It protects California residents that are either in California for other than a temporary or transitory purpose or are domiciled in California but are currently outside the state for a temporary or transitory purpose (Cal. Civ. Code §1798.140(c), (g); Cal. Code Regs. tit. 18, §17014). Consequently, the CCPA can apply to websites and web technology providers located outside California or even outside the United States, as long as they cater to California residents. Several other U.S. states are considering introducing similar or even stricter privacy laws than the CCPA. As with the *Brussels effect* in the context of GDPR, we may observe a *California effect* (Vogel, 1995) in future U.S. privacy law, and state privacy laws converge towards the jurisdiction with the strictest regulatory standard.

5.3 Privacy law and antitrust policy

Increasing concentration and an increasing market share of the dominant firm is most likely not what the European legislature had in mind when designing GDPR. Indeed, the European Commission (EC)

[20] In May 2018, Microsoft announced, e.g., that it would apply GDPR to all consumer customers worldwide, see https://blogs.microsoft.com/on-the-issues/2018/05/21/microsofts-commitment-to-gdpr-privacy-and-putting-customers-in-control-of-their-own-data; See also (Bradford, 2020, 143–144);(Rustad and Koenig, 2019, 391–96).

stressed in 2012 how the pro-competitive effects of the future GDPR would increase the attractiveness of Europe as a location to do business (European Commission, 2012, 148–149).

GDPR implemented and enforced the consent requirement for websites on a large scale, which disproportionately benefits larger firms offering a broader range of services (Campbell et al., 2015). The concentration in the markets for web technologies may additionally be explained by economies of scale. Why would only the firm with largest pre-GDPR market share increase its market share? In Table 17, looking at the advertising network submarket, we find that Google and Facebook both increased market share in EU-targeted websites, but Google's increase is four times larger. This may be due to network effects that make it more appealing for websites to use larger advertising networks since Google's pre-GDPR market share in the advertising technology market was already 3.5 times larger than Facebook's. Also, Facebook's reputation concerning privacy issues has suffered over the last years. Facebook has lost 9.8% market share with websites catering to EU audiences in their strongest market before GDPR (market share of 59.5%): the market for like/share/login buttons, which additionally allow Facebook to track users across websites (Roosendaal, 2012). This is consistent with the issues of joint responsibility, and recent European case law that involved websites and their usage of Facebook's like buttons.[21]

The increased concentration in web technology markets may have been an unintended but unavoidable consequence of GDPR. This raises the question of how privacy law and antitrust policy are related. Under European law, antitrust and privacy laws have traditionally been distinct. Enforcement of European antitrust laws was well-developed and done by both public antitrust authorities at the EU and member state levels; enforcement of privacy laws was traditionally weak and left to member state authorities, which sometimes delegated this task further in their federal structures. As recently as 2014, the EC noted in its approval of the merger between Facebook and WhatsApp that "any privacy-related concerns flowing from the increased concentration of data within the control of Facebook as a result of the transaction do not fall within the scope of the EU competition law rules but within the scope of the EU data protection rules."[22]

Our results indicate that it is increasingly difficult to conceptualize antitrust and privacy law as two

[21] See European Court of Justice, July 29, 2019, case C-40/17, Fashion ID v. Verbraucherzentrale NRW ("Fashion ID").

[22] European Commission, Decision in Case COMP/M.7217 (Facebook/WhatsApp), Oct. 3, 2014, C(2014) 7239 final, note 164. In its antitrust merger approval of the Google/DoubleClick merger in 2006, the EC had adopted a similar hands-off approach, leaving potential privacy problems to national privacy authorities once the merger had been implemented, European Commission, Decision in Case COMP/M.4731 (Google/DoubleClick), March 11, 2008, C(2008) 927 final, note 368.

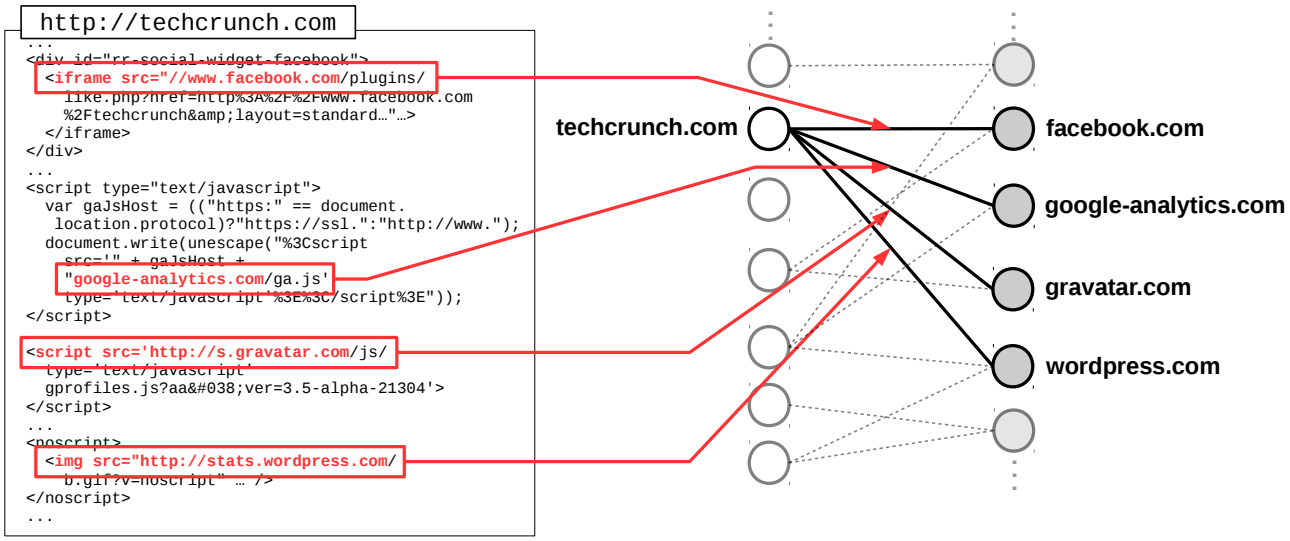
distinct areas of the law with different goals, remedies, and enforcement mechanisms (Economides and Lianos, 2019). On the one hand, network effects, lacking competition on terms of service and privacy policies as well as the limited effectiveness of user consent in privacy law (Acquisti and Grossklags, 2005; Barth and de Jong, 2017) may enable firms to increase their dominant position by violating privacy laws.[23] On the other hand, we find that laws aimed at increasing privacy protection may, at the same time, decrease competition in related technology markets. In world where processing personal data, analyzing user-profiles and predicting consumer behavior are cornerstones of highly concentrated Internet markets, designing privacy laws that do not have immediate implications for antitrust policy (or vice versa) is near-impossible.

6 Conclusion

We provide robust large-scale evidence on the changes coming with GDPR in the context of websites and web technology providers. We demonstrate how firms are reducing their compliance risks after GDPR: websites reduce the number of third-party trackers, and they move from third-party to first-party cookies, from non-EU-based to EU-based web technology providers, and from smaller to larger providers. We document a decrease in websites' use of third-party web technologies in the short-run, but we do not see a change in the long-run increasing trend. Decreases in first-party and third-party cookies, however, seem more sustained. We provide empirical evidence of the *Brussels effect* in European privacy law: Websites and web technology providers catering to non-EU audiences and therefore not subject to GDPR still comply with it. Finally, we demonstrate that, following GDPR's enactment, markets for web technologies became more concentrated, with the dominant firm—Google—increasing its market share. Our findings suggest that some of the most important effects of GDPR may not relate to privacy, but to antitrust and trade policy. They also indicate how closely linked these areas of the law have become. We leave the implications for the theoretical relationship between privacy and antitrust laws to future research.

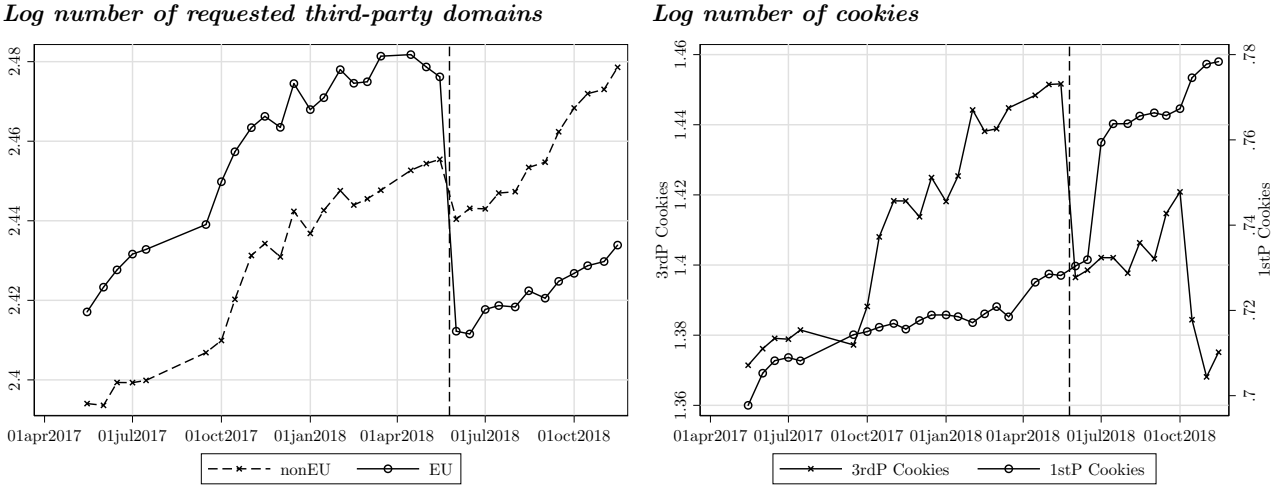
[23] Along similar lines, the German antitrust authority (*Bundeskartellamt*) has recently attempted to block Facebook from combining user data from different sources, including WhatsApp and Instagram; see *Bundeskartellamt Prohibits Facebook from Combining User Data from Different Sources*, Feb. 7, 2019, available at https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2019/07_02_2019_Facebook.html; (Sideri, 2019; Scharf, 2019). While the decision by the *Bundeskartellamt* was reversed by an appeals court (Oberlandesgericht Düsseldorf, case VI-Kart 1/19, August 26, 2019) and is still pending in the German court system, it indicates how antitrust authorities are exploring the connections between antitrust and privacy law in digital platform markets.

Figure 1: Example of third-party requests



Note: Similar to Figure 1 in Schelter and Kunegis (2018).

Figure 2: Requested third-party domains and cookies



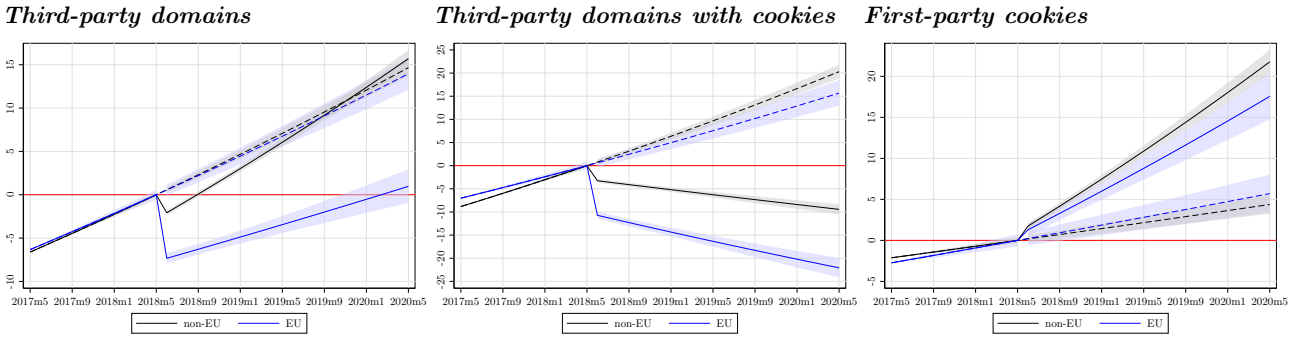
Note: Left-hand panel shows average log number of third-party domains to which website-hosts with(out) EU country-specific top-level domains send requests. Right-hand panel shows average log number of third-party domains that respond with a cookie, and number first-party requests (same domain as website-host) that respond with a cookie. Vertical line indicates the implementation of GDPR on May 25, 2018.

Table 1: Change in number of requested third-party domains

	(1) EU	(2) NonEU	(3) All
Trend	0.0173***	0.0194***	0.0194***
	(0.0008)	(0.0004)	(0.0004)
Post	-0.0840***	-0.0245***	-0.0245***
	(0.0027)	(0.0011)	(0.0011)
Trend × EU			-0.0021**
			(0.0009)
Post × EU			-0.0596***
			(0.0029)
Observations	728112	2925186	3653298
$\overline{R^2}$	0.8903	0.9033	0.9007
Mean DV	2.4453	2.4385	2.4399

Note: Dependent variable is the log number of third-party domains that the website-host requests. *Trend* is a linear time trend in days, where May 1, 2017 is 0.01. *Post* indicates the period after May 25, 2018. *EU* indicates whether a website-host has a top-level domain that is specific to a country in the EU. Website-host fixed-effects in all specifications. Standard errors in parentheses, clustered on the website-host-level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Figure 3: Change in number of requested domains and cookies, long-run prediction



Note: Predictions based on regression models which allow for different trends across EU/non-EU websites and before/after GDPR, respectively. Normalized to the month before the introduction of GDPR and in percent ($exp(coeff) - 1 * 100$). Dashed lines indicate counterfactual trends, shaded areas indicate 90% confidence bands.

Table 2: Change in numnber of requested third-party domains (not excluding outliers)

	(1) EU	(2) NonEU	(3) All
Trend	0.0227***	0.0219***	0.0219***
	(0.0008)	(0.0004)	(0.0004)
Post	-0.0929***	-0.0282***	-0.0282***
	(0.0027)	(0.0011)	(0.0011)
Trend × EU			0.0008
			(0.0009)
Post × EU			-0.0646***
			(0.0029)
Observations	816368	3279754	4096122
$\overline{R^2}$	0.8881	0.9008	0.8982
Mean DV	2.4397	2.4324	2.4339

Note: Dependent variable is the log number of requests that the website-host sends to third parties. *Trend* is a linear time trend in days, where May 1, 2017 is 0.01. *Post* indicates the period after May 25, 2018. *EU* indicates whether a website-host has a top-level domain that is specific to a country in the EU. Website-host fixed-effects in all specifications. Standard errors in parentheses, clustered on the website-host-level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Table 3: Change in number of requested third-party domains, different measures for EU

	EU			non-EU					
	(1) Rank	(2) Language	(3) DE Visits	(4) not Rank	(5) not Language	(6) US Visits	(7) Rank	(8) Language	(9) DE Visits
Trend	0.0185***	0.0172***	0.0151***	0.0192***	0.0277***	0.0295***	0.0192***	0.0277***	0.0295***
	(0.0006)	(0.0017)	(0.0049)	(0.0004)	(0.0011)	(0.0047)	(0.0004)	(0.0011)	(0.0047)
Post	-0.0705***	-0.0823***	-0.1153***	-0.0218***	-0.0355***	-0.0608***	-0.0218***	-0.0355***	-0.0608***
	(0.0021)	(0.0054)	(0.0153)	(0.0011)	(0.0030)	(0.0130)	(0.0011)	(0.0030)	(0.0130)
Trend × EU							-0.0006	-0.0104***	-0.0143**
							(0.0008)	(0.0020)	(0.0067)
Post × EU							-0.0487***	-0.0468***	-0.0546***
							(0.0024)	(0.0062)	(0.0201)
Observations	1091475	173877	21318	2561823	385011	22143	3653298	558888	43461
$\overline{R^2}$	0.8928	0.8968	0.9098	0.9041	0.9123	0.9025	0.9006	0.9080	0.9125
Mean DV	2.4829	2.3641	2.7489	2.4216	2.5433	3.2671	2.4399	2.4876	3.0129

Note: Dependent variable is the log number of requests that the website-host sends to third parties. *Trend* is a linear time trend in days, where 1 is May 1, 2017. *Post* indicates the period after May 25, 2018. *EU* in columns (1), (4) and (7) indicates whether the website appears on Alexa's rank for any country in the EU. *EU* in columns (2), (5) and (8) indicates whether the website-host returns content in any of the official languages of member countries of the EU (except English). *EU* in columns (3), (6) and (9) indicates whether the website-domain is visited by user in Germany but not users in the US in the Nielsen clickstream data. Website-host fixed-effects in all specifications. Standard errors in parentheses, clustered on the website-host-level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Table 4: Placebo exercises

	One Year Before			Piracy Websites
	(1) EU	(2) nonEU	(3) All	(4) All
Trend	0.029238***	0.025183***	0.025183***	-0.027374
	(0.000803)	(0.000397)	(0.000397)	(0.021992)
Post	0.007204***	-0.001806*	-0.001806*	-0.034253
	(0.001791)	(0.000922)	(0.000922)	(0.064958)
Trend × EU			0.004055***	
			(0.000895)	
Post × EU			0.009010***	
			(0.002014)	
Observations	871012	3406530	4277542	2109
$\overline{R^2}$	0.875597	0.877641	0.877232	0.848776
Mean DV	2.421103	2.404705	2.408044	1.858701

Note: Dependent variable is the log number of requests that the website-host sends to third parties. *Trend* is a linear time trend in days, where May 1, 2017 is 0.01. *Post* indicates the period after May 25, 2017 in column (1) and after May 25, 2018 in column (2). Website-host fixed-effects in all specifications. Standard errors in parentheses, clustered on the website-host-level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Table 5: Change in number of requested third-party domains, data from whotracks.me

	(1) EU	(2) US	(3) EU+US
Trend	7.1832***	7.7033***	7.7033***
	(0.2183)	(0.2097)	(0.2098)
Post	-0.1823***	-0.1815***	-0.1815***
	(0.0156)	(0.0126)	(0.0126)
Trend × EU			-0.5201***
			(0.1677)
Post × EU			-0.0008
			(0.0140)
Observations	10850	10850	21700
$\overline{R^2}$	0.7787	0.8181	0.7398
Mean DV	2.6547	2.5855	2.6201

Note: Dependent variable is the log number of third-parties that a website requests. *Trend* is a linear time trend in days, where April 1, 2018 is 0.01. *Post* indicates the period after May 25, 2018. *EU* indicates whether a website was visited by a user in the EU versus a user in the U.S. Website fixed-effects in all specifications. Standard errors in parentheses, clustered on the website-level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Table 6: Change in number of cookies

	3rdP Cookies			1stP Cookies		
	(1) EU	(2) NonEU	(3) All	(4) EU	(5) NonEU	(6) All
Trend	0.0154***	0.0214***	0.0214***	0.0091***	0.0081***	0.0081***
	(0.0011)	(0.0006)	(0.0006)	(0.0010)	(0.0005)	(0.0005)
Post	-0.1369***	-0.0563***	-0.0563***	0.0169***	0.0246***	0.0246***
	(0.0037)	(0.0015)	(0.0015)	(0.0028)	(0.0014)	(0.0014)
Trend × EU			-0.0060***			0.0010
			(0.0013)			(0.0011)
Post × EU			-0.0806***			-0.0076**
			(0.0041)			(0.0031)
Observations	728112	2925186	3653298	728112	2925186	3653298
$\overline{R^2}$	0.8490	0.8614	0.8589	0.8334	0.8044	0.8114
Mean DV	1.4445	1.3990	1.4080	0.8099	0.7137	0.7329

Note: Dependent variable is the log number of requests that the website-host sends to third parties that respond with cookies in columns (1)–(3). In columns (4)–(6) the dependent variable is the log number of cookies that originate from the same domain as the website-host. *Trend* is a linear time trend in days, where May 1, 2017 is 0.01. *Post* indicates the period after May 25, 2018. *EU* indicates whether a website-host has a top-level domain that is specific to a country in the EU. Website-host fixed-effects in all specifications. Standard errors in parentheses, clustered on the website-host-level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Table 7: Change in geographic share of requested third-party domains

	(1) EU 3rd Parties	(2) APP 3rd Parties	(3) Other 3rd Parties
Trend	-0.0004***	-0.0002***	0.0006***
	(0.0000)	(0.0000)	(0.0000)
Trend × EU	-0.0020***	0.0001***	0.0019***
	(0.0001)	(0.0000)	(0.0001)
Post	0.0003***	-0.0002***	-0.0002**
	(0.0001)	(0.0001)	(0.0001)
Post × EU	0.0031***	-0.0000	-0.0031***
	(0.0004)	(0.0001)	(0.0004)
Observations	3570592	3570592	3570592
$\overline{R^2}$	0.8721	0.8995	0.8765
Mean DV	0.0172	0.0069	0.9759

Note: Dependent variable is the share of requests that the website-host sends to third-parties with EU top level domains in columns (1), to third-parties with top levels domains of countries that the EU considers to provide adequate privacy protection in column (2) and to third-parties with all other top levels domains in column (3). *Trend* is a linear time trend in days, where May 1, 2017 is 0.01. *Post* indicates the period after May 25, 2018. *EU* indicates whether a website-host has a top-level domain that is specific to a country in the EU. Website-host fixed-effects in all specifications. Standard errors in parentheses, clustered on the website-host-level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Table 8: Change in privacy policies: collection of personal data

	Any collection	Collection of personal data	
	(1) Disclosure	(2) Always disclosed	(3) Disclosed after
Post	0.055096***	-0.190277***	0.539634***
	(0.007293)	(0.015631)	(0.040880)
Observations	7986	3826	656
$\overline{R^2}$	0.786694	0.305165	0.447233
Mean DV	0.533684	0.786200	0.553354

Note: Dependent variable is an indicator of whether a tracker discloses to collect personally-identifiable or sensitive data (based on data from Evidon). In columns (2)–(5) we distinguish between trackers that always disclosed their data collection policy (of any kind of data), and trackers that start to disclosed their data collection policy (of any kind of data) after GDPR. *Post* indicates the period after May 25, 2018. Because we only observe information on the privacy policy in two points in time (September 20, 2017 and July 19, 2019) we cannot estimate a trend. Standard errors clustered on the tracker-level in parentheses. All specifications include tracker-fixed effects. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Table 9: Change in privacy policies: sharing of personal data

	Any sharing	Sharing of personal data			
	(1) Disclosure	(2) Always disclosed	(3) Always disclosed	(4) Disclosed after	(5) Disclosed after
Post	0.096419***	0.131719***	0.057571**	0.431193***	0.231244***
	(0.007514)	(0.012664)	(0.026741)	(0.033600)	(0.067452)
Collection			0.056764		0.038425
			(0.035539)		(0.069459)
Post × Collection			0.112629***		0.225236**
			(0.032330)		(0.091699)
Observations	7986	3386	3386	872	872
$\overline{R^2}$	0.774434	0.678621	0.690196	0.300459	0.325565
Mean DV	0.484974	0.302717	0.302717	0.227064	0.227064

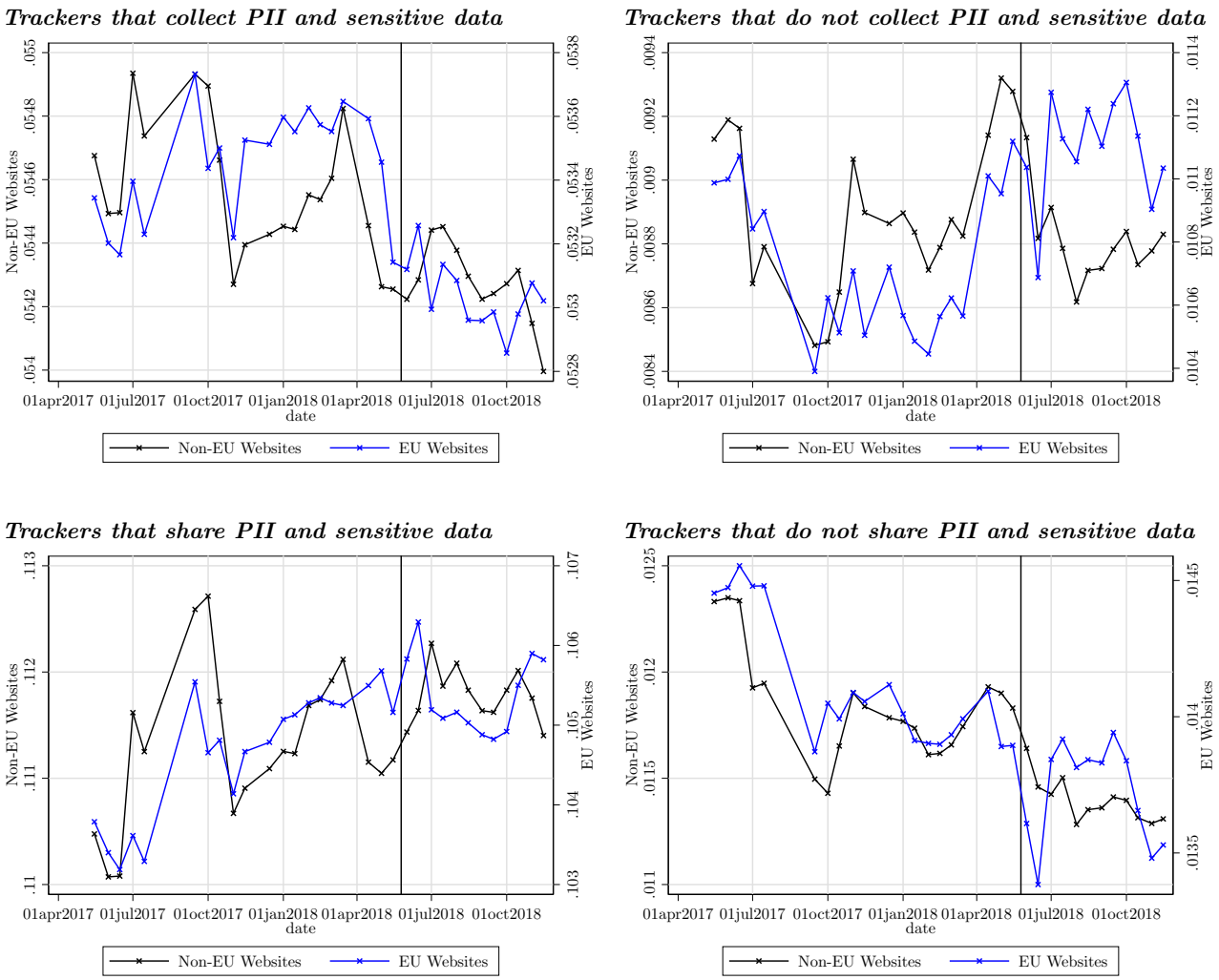
Note: Dependent variable is an indicator of whether a tracker discloses to share personally-identifiable or sensitive data with third parties (based on data from Evidon). In columns (2)–(5) we distinguish between trackers that always disclosed their data sharing policy (of any kind of data), and trackers that start to disclosed their data sharing policy (of any kind of data) after GDPR. *Post* indicates the period after May 25, 2018. Because we only observe information on the privacy policy in two points in time (September 20, 2017 and July 19, 2019) we cannot estimate a trend. Standard errors clustered on the tracker-level in parentheses. All specifications include tracker-fixed effects. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Table 10: Change in privacy policies: data retention policies

	Disclosure of Data Retention		Data Retention
	(1) Any policy	(2) As required by law	(3) In Months
Post	0.052842***	-0.001394	1.852941
	(0.005509)	(0.001539)	(2.963954)
Disclosed after		0.802320***	
		(0.038776)	
Observations	7986	7729	346
$\overline{R^2}$	0.732381	0.914148	0.954623
Mean DV	0.130103	0.082546	7.2e+01

Note: Dependent variable in column (1) is an indicator of whether a tracker discloses their data retention policy (based on data from Evidon). In columns (2)–(3) it is an indicator of whether a tracker discloses that they retain data as long as necessary to fulfill a business need or as required by law. In column (4) the dependent variable is the disclosed retention period in months. *Post* indicates the period after May 25, 2018. Because we only observe information on the privacy policy in two points in time (September 20, 2017 and July 19, 2019) we cannot estimate a trend. Standard errors clustered on the tracker-level in parentheses. All specifications include tracker-fixed effects. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Figure 4: Average market shares by privacy policy



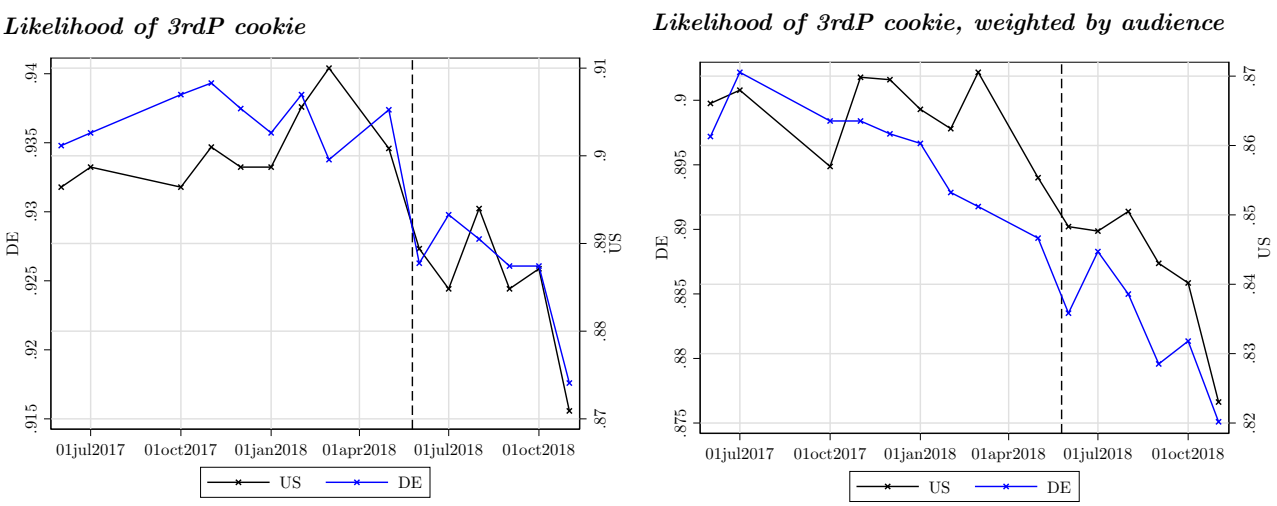
Note: Subsample of trackers for which we observe a privacy policy in data from Evidon before the implementation of GDPR. Market shares defined as number of website-hosts with (non-)EU country-specific top level domains that send requests to any of a firm's domains divided by the number of website-hosts with (non-)EU country-specific that send requests to third-party domains. Vertical lines indicate the implementation of GDPR on May 25, 2018.

Table 11: Change in market shares by privacy policy

	(1) Collection	(2) Sharing
EU × Personal Data	-0.003289*	-0.009673***
	(0.001829)	(0.002970)
Post	-0.000160	-0.000123
	(0.000269)	(0.000280)
Post × EU	0.000592**	0.000158
	(0.000292)	(0.000208)
Post × Personal Data	0.000114	0.000133
	(0.000512)	(0.000828)
Post × EU × Personal Data	-0.001086	-0.000920
	(0.000689)	(0.001473)
Observations	137024	137024
$\overline{R^2}$	0.998289	0.998296
Mean DV	0.036835	0.036835

Note: Dependent variable is market share defined as $s_i = n_i / \sum_{k=1}^{N} n_k$ where n_i is the number of websites that send requests to tracker i owned by Google. *Post* indicates the period after May 25, 2018. *Trend* is a linear time trend in days, where May 1, 2017 is 0.01. *Google* indicates the joint market share of all tracker domains of the firm in the category denoted by the title of the respective column. Category information comes from *Whotracks.me*. Standard errors clustered on the firm-level in parentheses. All specifications include firm-fixed effects. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Figure 5: Likelihood of 3rdP cookie in top 1000 visited websites



Note: Average of a variable indicating whether a website-domain sends requests to third-parties which respond with cookies. Underlying sample includes the top 1000 most visited websites by users in Germany and users in the U.S. as measured by Nielsen.
Left panel: Unweighted average.
Right panel: Average weighted by website-domain's audience (Nielsen's estimate of the number of users in the respective population). Vertical line indicates the implementation of GDPR on May 25, 2018.

Table 12: Consumer exposure to third-party cookies

	Share	Likelihood of 3rdP cookie		
	(1)	(2)	(3)	(4)
Post	-0.0095***	-0.0085***	0.0019	-0.0069***
	(0.0010)	(0.0002)	(0.0017)	(0.0017)
Post × UserGermany	-0.0097***	-0.0084***		0.0231***
	(0.0018)	(0.0003)		(0.0024)
Post × Log(Pageviews)			-0.0013***	-0.0001
			(0.0002)	(0.0001)
Post × Log(Pageviews) × UserGermany				-0.0030***
				(0.0002)
Observations	2520	1253661	1253661	1253661
$\overline{R^2}$	0.7444	0.8214	0.8214	0.8214
Mean DV	0.9123	0.9171	0.9171	0.9171

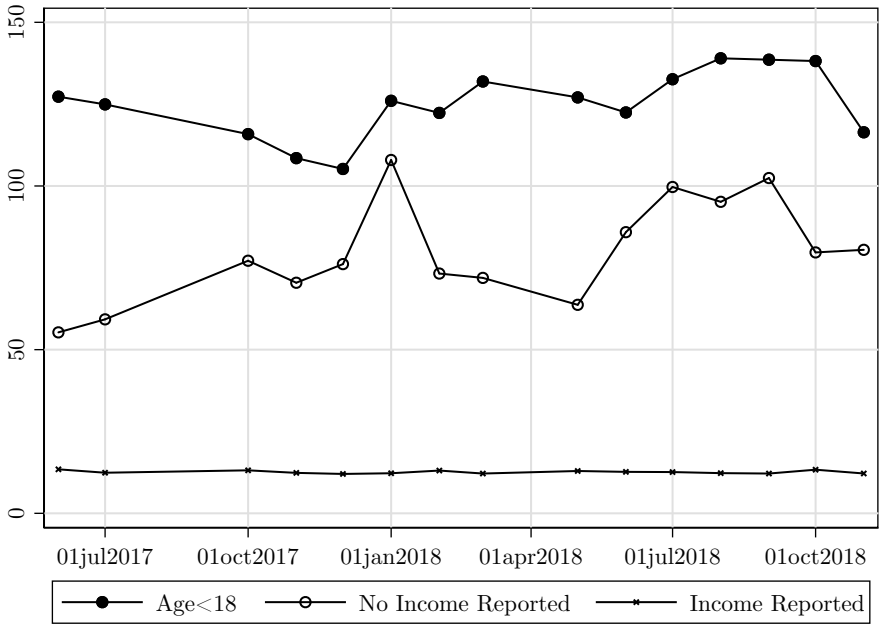
Note: The dependent variable in column (1) is the share of visited domains that send requests to third-parties which respond with cookies. In columns (2)–(4) the dependent variable is an indicator whether a visited domain sends requests to third-parties which respond with cookies. *Post* indicates the period after May 25, 2018. Usergroup-fixed effects in all models, additional website-fixed effects in columns (2)–(4). Standard errors in parentheses, clustered on the usergroup-level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Table 13: Consumer exposure to third-party cookies, by consumer types

	Share websites 3rdP cookies			Likelihood of 3rdP cookie		
	(1) DE	(2) US	(3) All	(4) DE	(5) US	(6) All
Post	-0.0151***	-0.0088***	-0.0089***	-0.0164***	-0.0085***	-0.0084***
	(0.0005)	(0.0004)	(0.0008)	(0.0003)	(0.0002)	(0.0002)
Post × Age<18	-0.0232**	-0.0021	0.0020	-0.0110***	-0.0002	-0.0008
	(0.0089)	(0.0057)	(0.0054)	(0.0029)	(0.0008)	(0.0010)
Post × NoInc	-0.0021	-0.0026	-0.0076**	-0.0016	-0.0003	-0.0005
	(0.0033)	(0.0027)	(0.0032)	(0.0014)	(0.0009)	(0.0007)
Post × UserDE			-0.0061***			-0.0081***
			(0.0015)			(0.0003)
Post × Age<18 × UserDE			-0.0293***			-0.0083***
			(0.0063)			(0.0022)
Post × NoInc × UserDE			0.0104*			-0.0010
			(0.0054)			(0.0012)
Observations	1260	1260	2520	594818	658843	1253661
$\overline{R^2}$	0.6260	0.2471	0.7716	0.7971	0.8521	0.8214
Mean DV	0.8914	0.9332	0.9123	0.8985	0.9340	0.9171

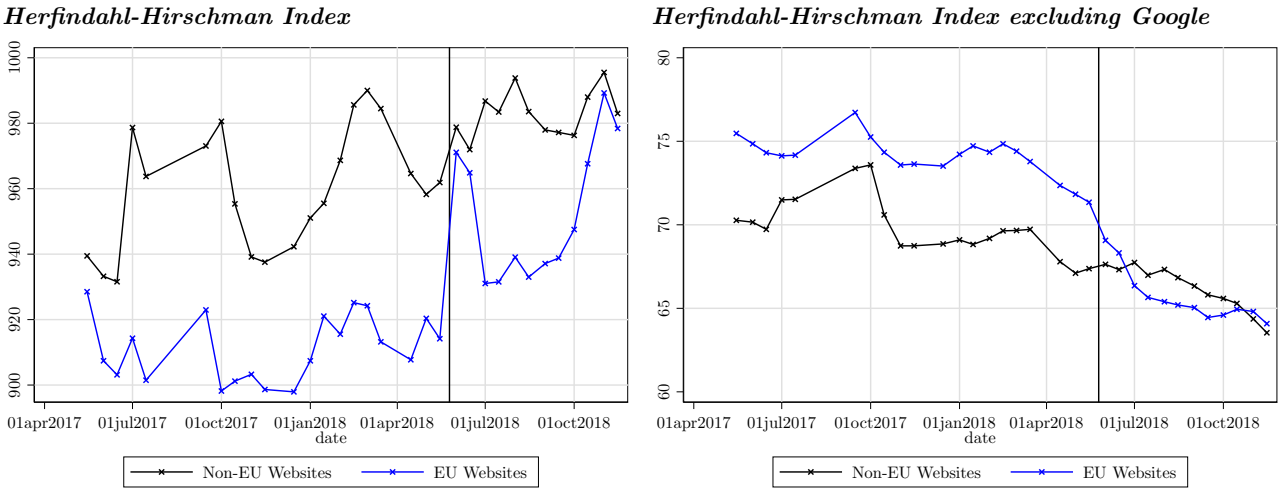
Note: The dependent variable in columns (1)–(3) is the share of visited domains that send requests to third-parties which respond with cookies. In columns (4)–(6) the dependent variable is an indicator whether a visited domain sends requests to third-parties which respond with cookies. *Post* indicates the period after May 25, 2018. Usergroup-fixed effects in all models, additional website-fixed effects in columns (4)–(6). Standard errors in parentheses, clustered on the usergroup-level. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Figure 6: Demand concentration, by user types



Note: Sum of squared market shares (HHI) in the top 1000 most visited website-domains of different user types.
Market shares are defined as $s_i = v_i / \sum_{k=1}^{N} v_k$ where v_i is the number of pageviews that a website domain receives from user type i.

Figure 7: Change in market structure of the web technology industry



Note: Market shares defined as $s_i = n_i / \sum_{k=1}^{N} n_k$ where n_i is the number of websites that send requests to third-party domains of firm i (reach). HHI is the sum of squared market shares. Markets are defined based on the audiences that websites cater to. *EU Websites* and *non-EU Websites* are defined based on country-specific top-level domains. Vertical lines indicate the implementation of GDPR on May 25, 2018.

Table 14: Change in market structure of the web technology industry

	All		w/o Google	
	(1)	(2)	(3)	(4)
Trend	6.3468**	3.3995	-1.0873***	-1.2842***
	(2.8259)	(2.7583)	(0.2307)	(0.2167)
EU-based Websites	-42.5656***	-37.6137***	3.8699***	3.9458***
	(9.2568)	(11.3974)	(0.7978)	(0.8206)
Post	5.7988	11.8158	-0.5481	0.0669
	(8.1989)	(8.3108)	(0.5472)	(0.7122)
Trend × EU-based Websites	-3.0778	-4.8635	0.2351	0.1575
	(3.7099)	(4.5255)	(0.3074)	(0.3047)
Post × EU-based Websites	26.3429**	30.3990**	-5.5375***	-5.2732***
	(11.4205)	(12.9623)	(0.9579)	(1.1626)
Observations	64	74	64	74
$\overline{R^2}$	0.7732	0.7009	0.9044	0.7977
Mean DV	947.5992	947.5509	69.6881	70.0832

Note: Dependent variable is HHI, calculated as squared sum of market shares defined as $s_i = n_i / \sum_{k=1}^{N} n_k$ where n_i is the number of websites that send requests to domains of firm i (reach). *Post* indicates the period after May 25, 2018. *Trend* is a linear time trend in days, where May 1, 2017 is 0.01. We exclude snapshots with outlier observations in columns (1) and (3). White-robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Table 15: Change in market shares of web technology providers

	Domains	Firms		
	(1)	(2)	(3)	(4)
Post	0.000000	-0.000000	-0.000001	-0.000002
	(0.000011)	(0.000011)	(0.000011)	(0.000011)
Post × EU	-0.000000	0.000000	-0.000004	-0.000009
	(0.000012)	(0.000015)	(0.000011)	(0.000012)
EU × Top2			-0.215112	
			(0.408395)	
Post × Top2			0.035368	
			(0.048553)	
Post × EU × Top2			0.141851	
			(0.285693)	
EU × Google				-0.792670***
				(0.000034)
Post × Google				0.104032***
				(0.000011)
Post × EU × Google				0.545883***
				(0.000012)
Observations	4083968	4050688	4050688	4050688
$\overline{R^2}$	0.988631	0.997941	0.998001	0.998215
Mean DV	0.001567	0.001580	0.001580	0.001580
% Google-Effect nonEU				0.346784
% Google-Effect EU				2.221910

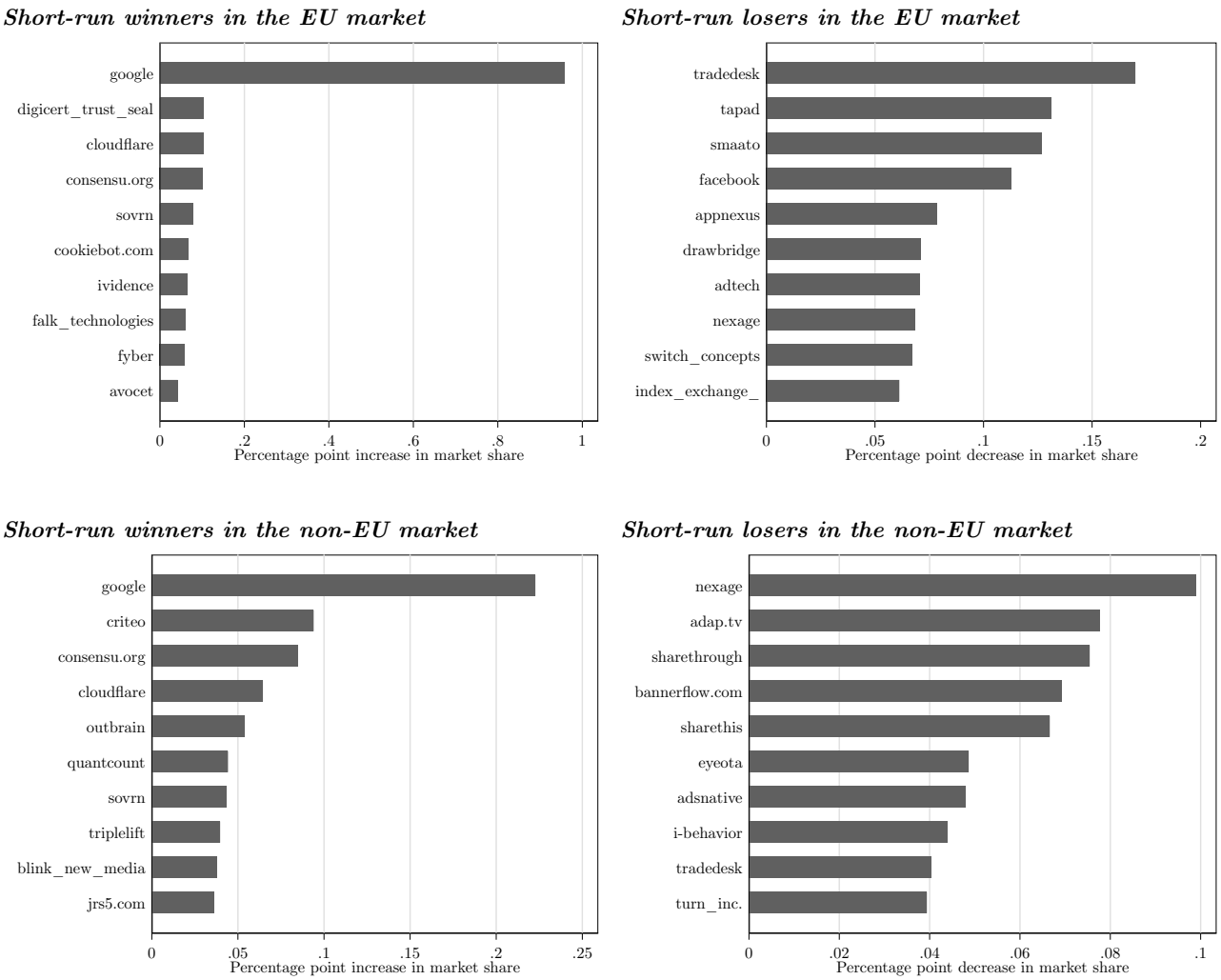
Note: Dependent variable is market share defined as $s_i = n_i / \sum_{k=1}^{N} n_k$ where n_i is the number of websites that send requests to service i (reach). In column (1) we consider each third-party domain as an independent service, in columns (2)–(4) we aggregate all domains of a firm using information from *Whotracks.me*. *Post* indicates the period after May 25, 2018. *Trend* is a linear time trend in days, where May 1, 2017 is 0.01. *Top2* indicates the joint market share of all tracker domains of the two firms with the highest marketshare (Google and Facebook). *Google* indicates the joint market share of all tracker domains of the firm. Standard errors in parentheses. Fixed effects and standard errors clustered on the tracker-domain-level in columns (1)–(2), on the firm-level in columns (3)–(6). * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Table 16: Change in market shares, by submarket, Google vs. all other firms

	(1) Video	(2) Advertising	(3) Tagmanager	(4) Analytics	(5) CDN/API	(6) Other
EU	1.5e-11	-4.2e-12	-2.8e-12	-3.5e-12	5.3e-11	-4.5e-12
	(.000134)	(.000061)	(.000165)	(.000087)	(.000052)	(.000071)
Post	-.000019	7.0e-06	-.000012	-.000012	-.000015	9.2e-07
	(.000044)	(.000023)	(.000058)	(.000024)	(.000019)	(.000032)
Post × Google	1.20566***	-.441783***	.731151***	.78715***	.971121***	-.058424***
	(.000031)	(.00003)	(.000035)	(.000031)	(.000022)	(.000038)
Post × EU	.000113	-.000029	.000093	-.000031	.000031	-1.5e-06
	(.000085)	(.000029)	(.000102)	(.000035)	(.000031)	(.000032)
Post × EU × Google	-7.15514***	1.83314***	-5.91109***	1.9904***	-1.97809***	.095707***
	(.00006)	(.000042)	(.000087)	(.000072)	(.000028)	(.000072)
Observations	4050688	4050688	4050688	4050688	4050688	4050688
$\overline{R^2}$	.998331	.995211	.99491	.993765	.999512	.90875
Mean DV	.00158	.00158	.00158	.00158	.00158	.00158
% Effect nonEU	1.59465	-1.58394	1.19014	1.95054	1.39394	-16.5879
% Effect EU	-7.46261	5.40282	-7.96724	7.23468	-1.39309	15.722

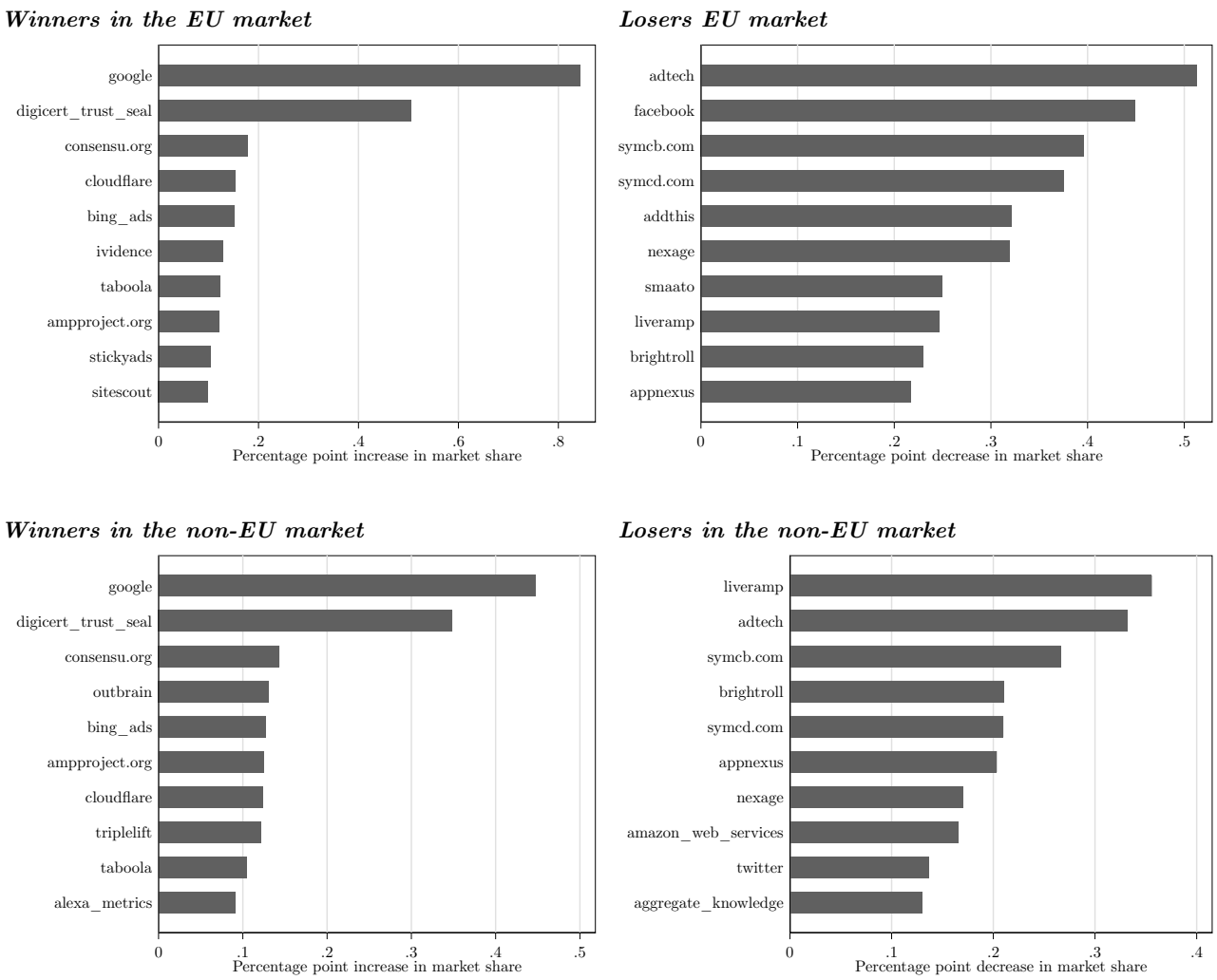
Note: Dependent variable is market share defined as $s_i = n_i / \sum_{k=1}^{N} n_k$ where n_i is the number of websites that send requests to tracker i owned by Google. *Post* indicates the period after May 25, 2018. *Google* indicates the joint market share of all tracker domains of the firm in the category denoted by the title of the respective column. Category information comes from *Whotracks.me*. Standard errors clustered on the firm-level in parentheses. All specifications include group-specific linear trends (non-EU, EU, Google) and firm-fixed effects. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

Figure 8: Winners and losers: 2 weeks before/after



Note: Ten firms with largest increase/decrease in average marketshare 2 weeks before and 2 weeks after the implementation of GDPR on May 25, 2018. Market shares defined as number of website-hosts with (non-)EU country-specific top level domains that send requests to any of a firm's domains over number of website-hosts with (non-)EU country-specific that send requests to third-party domains.

Figure 9: Winners and losers: 6 months before/after



Note: Ten firms with largest increase/decrease in average marketshare 6 months before and 6 months after the implementation of GDPR on May 25, 2018. Market shares defined as number of website-hosts with (non-)EU country-specific top level domains that send requests to any of a firm's domains divided by the number of website-hosts with (non-)EU country-specific that send requests to third-party domains.

Table 17: Change in market shares, by submarket, Google and Facebook vs. all other firms

	(1) Advertising	(2) Social Media
EU	-4.2e-12	-2.7e-12
	(.000061)	(.000263)
Post	5.3e-06	-.00007
	(.000023)	(.000089)
Post × Facebook	.103305***	4.42552***
	(.00003)	(.00003)
Post × Google	-.441782***	
	(.00003)	
Post × EU	-.000028	.000174
	(.000029)	(.000174)
Post × EU × Facebook	-.054528***	-10.9839***
	(.000042)	(.000088)
Post × EU × Google	1.83314***	
	(.000042)	
Observations	4050688	4050688
$\overline{R^2}$	.99528	.994232
Mean DV	.00158	.00158
% Effect nonEU Facebook	1.34712	8.43895
% Effect EU Facebook	.674025	-9.78537
% Effect nonEU Google	-1.58394	
% Effect EU Google	5.40281	

Note: Dependent variable is market share defined as $s_i = n_i / \sum_{k=1}^{N} n_k$ where n_i is the number of websites that send requests to tracker i owned by Google. *Post* indicates the period after May 25, 2018. *Google* indicates the joint market share of all tracker domains of the firm in the category denoted by the title of the respective column. Category information comes from *Whotracks.me*. Standard errors clustered on the firm-level in parentheses. All specifications include group-specific linear trends (non-EU, EU, Google) and firm-fixed effects. * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$

A Obtaining HTTPArchive data from Google BigQuery

Listing 1: "BigQuery Code - Websites"

```
#standardSQL
CREATE TABLE myhttparchive.common_hosts_balanced AS
SELECT p.pageid, NET.HOST(p.url) as host
FROM `httparchive.summary_pages.*` p,
(
WITH dlist AS
(
SELECT DISTINCT host, date
FROM
(
SELECT SUBSTR(_TABLE_SUFFIX, 0, 10) AS date, NET.HOST(p.url) as host
FROM `httparchive.summary_pages.*` p,
(
SELECT NET.HOST(t1.url) as host
  FROM `httparchive.summary_pages.2018_06_01_desktop` t1
  INNER JOIN `httparchive.summary_pages.2018_07_01_desktop` t2
  ON NET.HOST(t1.url)=NET.HOST(t2.url)
) l
WHERE  _TABLE_SUFFIX like '%desktop' AND (_TABLE_SUFFIX like '%2017_05%'
        OR _TABLE_SUFFIX like '%2017_06%'
        OR _TABLE_SUFFIX like '%2017_07%'
        OR _TABLE_SUFFIX like '%2017_08%'
        OR _TABLE_SUFFIX like '%2017_09%'
        OR _TABLE_SUFFIX like '%2017_10%'
        OR _TABLE_SUFFIX like '%2017_11%'
        OR _TABLE_SUFFIX like '%2017_12%'
        OR _TABLE_SUFFIX like '%2018%'
  AND _TABLE_SUFFIX NOT like '%2018_12%' )
AND NET.HOST(p.url)=l.host
)
),
times AS
(
SELECT host, COUNT(date) AS nooc FROM dlist
```

```
GROUP BY host
),
max AS
(
SELECT MAX(nooc) AS maxnooc FROM times
)
SELECT times.host
FROM times, max
WHERE times.nooc=max.maxnooc
) l
WHERE  _TABLE_SUFFIX like '%desktop' AND (_TABLE_SUFFIX like '%2017_05%'
        OR _TABLE_SUFFIX like '%2017_06%'
        OR _TABLE_SUFFIX like '%2017_07%'
        OR _TABLE_SUFFIX like '%2017_08%'
        OR _TABLE_SUFFIX like '%2017_09%'
        OR _TABLE_SUFFIX like '%2017_10%'
        OR _TABLE_SUFFIX like '%2017_11%'
        OR _TABLE_SUFFIX like '%2017_12%'
        OR _TABLE_SUFFIX like '%2018%'
  AND _TABLE_SUFFIX NOT like '%2018_12%')
AND NET.HOST(p.url)=l.host;

#standardSQL
CREATE TABLE myhttparchive.common_hosts_balanced_meta AS
SELECT NET.HOST(p.url) as host, NET.PUBLIC_SUFFIX(p.url) AS tld,  NET.REG_DOMAIN(p.
    url) AS domain
FROM `httparchive.summary_pages.*` p, `myhttparchive.common_hosts_balanced` l
WHERE  _TABLE_SUFFIX like '%desktop' AND (_TABLE_SUFFIX like '%2017_05%'
        OR _TABLE_SUFFIX like '%2017_06%'
        OR _TABLE_SUFFIX like '%2017_07%'
        OR _TABLE_SUFFIX like '%2017_08%'
        OR _TABLE_SUFFIX like '%2017_09%'
        OR _TABLE_SUFFIX like '%2017_10%'
        OR _TABLE_SUFFIX like '%2017_11%'
        OR _TABLE_SUFFIX like '%2017_12%'
        OR _TABLE_SUFFIX like '%2018%'
  AND _TABLE_SUFFIX NOT like '%2018_12%' )
```

```
AND NET.HOST(p.url)=l.host
GROUP BY host, tld, domain;
```

Listing 2: "BigQuery Code - Requests Cookies Language"

```
#standardSQL
CREATE TABLE myhttparchive.requests3 AS
SELECT SUBSTR(_TABLE_SUFFIX, 0, 10) AS date, l.host, COUNT(DISTINCT NET.REG_DOMAIN(r.
    url)) as requests
FROM `httparchive.summary_requests.*` r, `myhttparchive.common_hosts_balanced` l
WHERE  _TABLE_SUFFIX like '%desktop%' AND
(_TABLE_SUFFIX like '%2017_05%'
  OR _TABLE_SUFFIX like '%2017_06%'
  OR _TABLE_SUFFIX like '%2017_07%'
        OR _TABLE_SUFFIX like '%2017_08%'
        OR _TABLE_SUFFIX like '%2017_09%'
        OR _TABLE_SUFFIX like '%2017_10%'
        OR _TABLE_SUFFIX like '%2017_11%'
        OR _TABLE_SUFFIX like '%2017_12%'
        OR _TABLE_SUFFIX like '%2018%'
  AND _TABLE_SUFFIX NOT like '%2018_12%' )
AND r.pageid = l.pageid AND NET.REG_DOMAIN(r.url)!=NET.REG_DOMAIN(l.host) AND (status
    =200 OR status=302)
GROUP BY date, host;

#standardSQL
CREATE TABLE myhttparchive.cookies3 AS
SELECT SUBSTR(_TABLE_SUFFIX, 0, 10) AS date, l.host, COUNT(DISTINCT NET.REG_DOMAIN(r.
    url)) as requests
FROM `httparchive.summary_requests.*` r, `myhttparchive.common_hosts_balanced` l
WHERE  _TABLE_SUFFIX like '%desktop%' AND
(_TABLE_SUFFIX like '%2017_05%'
  OR _TABLE_SUFFIX like '%2017_06%'
  OR _TABLE_SUFFIX like '%2017_07%'
        OR _TABLE_SUFFIX like '%2017_08%'
        OR _TABLE_SUFFIX like '%2017_09%'
        OR _TABLE_SUFFIX like '%2017_10%'
        OR _TABLE_SUFFIX like '%2017_11%'
```

```
        OR _TABLE_SUFFIX like '%2017_12%'
        OR _TABLE_SUFFIX like '%2018%'
  AND _TABLE_SUFFIX NOT like '%2018_12%' )
AND r.pageid = l.pageid AND NET.REG_DOMAIN(r.url)!=NET.REG_DOMAIN(l.host) AND
    respCookieLen>0 AND (status=200 OR status=302)
GROUP BY date, host;

#standardSQL
CREATE TABLE myhttparchive.cookies1 AS
SELECT SUBSTR(_TABLE_SUFFIX, 0, 10) AS date, l.host, COUNT(DISTINCT r.requestid) as
    requests
FROM `httparchive.summary_requests.*` r, `myhttparchive.common_hosts_balanced` l
WHERE  _TABLE_SUFFIX like '%desktop%' AND
(_TABLE_SUFFIX like '%2017_05%'
  OR _TABLE_SUFFIX like '%2017_06%'
  OR _TABLE_SUFFIX like '%2017_07%'
        OR _TABLE_SUFFIX like '%2017_08%'
        OR _TABLE_SUFFIX like '%2017_09%'
        OR _TABLE_SUFFIX like '%2017_10%'
        OR _TABLE_SUFFIX like '%2017_11%'
        OR _TABLE_SUFFIX like '%2017_12%'
        OR _TABLE_SUFFIX like '%2018%'
  AND _TABLE_SUFFIX NOT like '%2018_12%' )
AND r.pageid = l.pageid AND NET.REG_DOMAIN(r.url)=NET.REG_DOMAIN(l.host) AND
    respCookieLen>0 AND status=200
GROUP BY date, host;

#standardSQL
CREATE TABLE myhttparchive.language1 AS
SELECT SUBSTR(_TABLE_SUFFIX, 0, 10) AS date, l.host, COUNT(DISTINCT r.requestid) as
    requests, resp_content_language as lang
FROM `httparchive.summary_requests.*` r, `myhttparchive.common_hosts_balanced` l
WHERE  _TABLE_SUFFIX like '%desktop%' AND
(_TABLE_SUFFIX like '%2017_05%'
  OR _TABLE_SUFFIX like '%2017_06%'
  OR _TABLE_SUFFIX like '%2017_07%'
        OR _TABLE_SUFFIX like '%2017_08%'
```

```
        OR _TABLE_SUFFIX like '%2017_09%'
        OR _TABLE_SUFFIX like '%2017_10%'
        OR _TABLE_SUFFIX like '%2017_11%'
        OR _TABLE_SUFFIX like '%2017_12%'
        OR _TABLE_SUFFIX like '%2018%'
  AND _TABLE_SUFFIX NOT like '%2018_12%' )
AND r.pageid = l.pageid AND NET.REG_DOMAIN(r.url)=NET.REG_DOMAIN(l.host)
GROUP BY date, host, lang;
```

Listing 3: "BigQuery Code - Trackers"

```
#standardSQL
CREATE TABLE myhttparchive.trackers AS
SELECT SUBSTR(_TABLE_SUFFIX, 0, 10) AS date, NET.REG_DOMAIN(r.url) as tracker, COUNT(
    r.requestid) as requests, COUNT(DISTINCT NET.REG_DOMAIN(l.host)) as websites
FROM `httparchive.summary_requests.*` r, `myhttparchive.common_hosts_balanced` l
WHERE  _TABLE_SUFFIX like '%desktop%' AND
(_TABLE_SUFFIX like '%2017_05%'
  OR _TABLE_SUFFIX like '%2017_06%'
  OR _TABLE_SUFFIX like '%2017_07%'
        OR _TABLE_SUFFIX like '%2017_08%'
        OR _TABLE_SUFFIX like '%2017_09%'
        OR _TABLE_SUFFIX like '%2017_10%'
        OR _TABLE_SUFFIX like '%2017_11%'
        OR _TABLE_SUFFIX like '%2017_12%'
        OR _TABLE_SUFFIX like '%2018%'
  AND _TABLE_SUFFIX NOT like '%2018_12%' )
AND r.pageid = l.pageid AND NET.REG_DOMAIN(r.url)!=NET.REG_DOMAIN(l.host) AND (status
    =200 OR status=302)
GROUP BY date, tracker;

#standardSQL
CREATE TABLE myhttparchive.trackers_eu AS
SELECT SUBSTR(_TABLE_SUFFIX, 0, 10) AS date, NET.REG_DOMAIN(r.url) as tracker, COUNT(
    r.requestid) as requests, eu.eu as website_eu, COUNT(DISTINCT NET.REG_DOMAIN(l.
    host)) as websites
FROM `httparchive.summary_requests.*` r, `myhttparchive.common_hosts_balanced` l, `
    myhttparchive.common_hosts_balanced_meta_eu` eu
```

```
WHERE  _TABLE_SUFFIX like '%desktop%' AND
(_TABLE_SUFFIX like '%2017_05%'
  OR _TABLE_SUFFIX like '%2017_06%'
  OR _TABLE_SUFFIX like '%2017_07%'
        OR _TABLE_SUFFIX like '%2017_08%'
        OR _TABLE_SUFFIX like '%2017_09%'
        OR _TABLE_SUFFIX like '%2017_10%'
        OR _TABLE_SUFFIX like '%2017_11%'
        OR _TABLE_SUFFIX like '%2017_12%'
        OR _TABLE_SUFFIX like '%2018%'
  AND _TABLE_SUFFIX NOT like '%2018_12%' )
AND r.pageid = l.pageid AND NET.REG_DOMAIN(r.url)!=NET.REG_DOMAIN(l.host) AND (status
    =200 OR status=302) AND l.host=eu.host
GROUP BY date, website_eu, tracker;
```

References

Acquisti, A., and Grossklags, J. (2005). "Privacy and Rationality in Individual Decision Making." *IEEE Security & Privacy*, *3*, 26–33.

Avi Goldfarb and Catherine Tucker (2012). "Shifts in Privacy Concerns." *American Economic Review: Papers & Proceedings*, *102*(3), 349–353.

Barth, S., and de Jong, M. D. (2017). "The Privacy Paradox: Investigating Discrepancies between Expressed Privacy Concerns and Actual Online Behavior—A Systematic Literature Review." *Telematics and Informatics*, *34*, 1038–1058.

Batikas, M., Claussen, J., and Peukert, C. (2019). "Follow the Money: Online Piracy and Self-regulation in the Advertising Industry." *International Journal of Industrial Organization*, *65*, 121–151.

Bradford, A. (2012). "The Brussels Effect." *Northwestern Unviersity Law Review*, *107*, 1–68.

Bradford, A. (2020). *The Brussels Effect*. Oxford University Press.

Campbell, J., Goldfarb, A., and Tucker, C. (2015). "Privacy Regulation and Market Structure." *Journal of Economics & Management Strategy*, *24*, 47–73.

Chaabane, A., Kaafar, M., and Borelli, R. (2012). "Big Friend is Watching You: Analyzing Online Social Network Tracking Capabilities." *WOSN 12 Proceedings*.

Dabrowski, A., Merzdovnik, G., Ullrich, J., Sendera, G., and Weippl, E. (2019). "Measuring Cookies and Web Privacy in a Post-GDPR World." In *International Conference on Passive and Active Network Measurement*, 258–270, Springer.

Degeling, M., Utz, C., Lentzsch, C., Hosseini, H., Schaub, F., and Holz, T. (2018). "We Value Your Privacy... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy." *Working Paper*.

Economides, N., and Lianos, I. (2019). "Restrictions on Privacy and Exploitation in the Digital Economy: A Competition Law Perspective." *Working Paper*, SSRN–ID 3474099.

Economides, N., Lianos, I., et al. (2019). "Privacy Restrictions as Restrictions of Competition." *Working Paper*.

European Commission (2012). "Impact Assessment Accompanying the document Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation) and Directive of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data by Competent Authorities for the Purposes of Prevention, Investigation, Detection or Prosecution of Criminal Offences or the Execution of Criminal Penalties, and the Free Movement of Such Data." SEC(2012) 72 final.

European Data Protection Board (2019a). "Guidelines 3/2018 on the Territorial Scope of the GDPR (Article 3)."

European Data Protection Board (2019b). "Opinion 5/2019 on the Interplay Between the ePrivacy Directive and the GDPR, in Particular Regarding the Competence, Tasks and Powers of Data Protection Authorities."

European Union (1995). "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data." *Official Journal, L 281*, 31–50.

European Union (2002). "Direcitve 2002/58/EC of the European Parliament and of the Council of 12 July 2002 Concerning the Processing of Personal Data and the Protection of Privacy in the Electronic Communications Sector (Directive on Privacy and Electronic Communications, lasted amended by Directive 2009/136/EC." *Official Journal, L 201*, 37–47.

European Union (2016). "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation)." *Official Journal, L 119*, 1–88.

Gal, M. S., and Aviv, O. (2019). "The Competitive Effects of the GDPR." *Working Paper.*

Gill, P., Erramili, V., Chaintreau, A., Krishnamurthy, B., and Papagiannaki, D. (2013). "Follow the Money ? Understanding Economics of Online Aggregation and Advertising." *Proceedings of the 2013 ACM Internet Measurement Conference.*

Goddard, M. (2017). "The EU General Data Protection Regulation (GDPR): European Regulation that Has a Global Impact." *International Journal of Market Research, 59*(6), 703–705.

Godinho de Matos, M., and Adjerid, I. (2020). "Consumer Consent and Firm Targeting after GDPR: The Case of a Large Telecom Provider." *Working Paper.*

Goldberg, S., Johnson, G., and Shriver, S. (2019). "Regulating Privacy Online: The Early Impact of the GDPR on European Web Traffic & E-Commerce Outcomes." *Working Paper.*

Goldstein, J., and Wu, T. (2006). *Who Controls the Internet? Illusions of a Borderless World.* Oxford University Press.

Hu, X., and Sastry, N. (2019). "Characterising Third Party Cookie Usage in the EU after GDPR." *arXiv preprint arXiv:1905.01267.*

Jia, J., Jin, G. Z., and Wagman, L. (2019a). "GDPR and the Localness of Venture Investment." *Working Paper.*

Jia, J., Jin, G. Z., and Wagman, L. (2019b). "The Short-run Effects of GDPR on Technology Venture Investment." *Working Paper.*

Johnson, G., and Shriver, S. (2020). "Privacy & Market Concentration: Intended & Unintended Consequences of the GDPR." *Working Paper.*

Karaj, A., Macbeth, S., Berson, R., and Pujol, J. M. (2019). "WhoTracks.Me: Shedding light on the opaque world of online tracking." *Working Paper, arXiv:1804.08959v2.*

Lefrere, V., Warberg, L., Cheyre, C., Marotta, V., and Acquisti, A. (2019). "The Impact of GDPR on Ad-Supported Content Providers." *Working Paper presented at WISE 2019.*

Lerner, A., Kornfeld Simpson, A., Kohno, T., and Roesner, F. (2016). "Internet Jones and the Raiders of the Lost Tracker: An Archaeological Study of Web Tracking from 1996 to 2016." *Proceedings of the 25th USENIX Security Symposium.*

Libert, T. (2015). "Exposing the Hidden Web: An Analysis of Third-Party HTTP Requests on One Million Websites." *International Journal of Communication*, *9*, 3544–3561.

Libert, T., Graves, L., and Nielsen, R. K. (2018). "Changes in Third-party Content on European News Websites after GDPR." Tech. rep.

Roosendaal, A. (2012). "We Are All Connected to Facebook... by Facebook!" In S. Gutwirth, R. Leenes, P. DeHert, and Y. Poullet (Eds.), *European Data Protection: In Good Health?*, 3–19, Springer, New York.

Rustad, M. L., and Koenig, T. H. (2019). "Torwards a Global Data Privacy Standard." *Florida Law Review*, *71*, 365–453.

Scharf, A. (2019). "Exploitative Business Terms in the Era of Big Data: the Bundeskartellamt's Facebook Decision." *European Competition Law Review*, 332–339.

Schelter, S., and Kunegis, J. (2018). "On the Ubiquity of Web Tracking: Insights from a Billion-Page Web Crawl." *Journal of Web Science*, *4*(4), 53–66.

Schwartz, P. M. (2019). "Global Data Privacy: The EU Way." *New York University Law Review*, *94*, 771–818.

Schwartz, P. M., and Peifer, K.-N. (2017). "Transatlantic Data Privacy Law." *Georgetown Law Journal*, *106*, 115–179.

Sharma, P., Sun, Y., and Wagman, L. (2019). "The Differential Effects of New Privacy Protections on Publisher and Advertiser Profitability." *Working Paper*.

Sideri, K. (2019). "German Facebook Decision: the Interplay of Competition and Data Protection Law." *European Competition Law Review*, 354–357.

Solomos, K., Ilia, P., Ioannidis, S., and Kourtellis, N. (2019). "Clash of the Trackers: Measuring the Evolution of the Online Tracking Ecosystem." *Working Paper*.

Sørensen, J., and Kosta, S. (2019). "Before and After GDPR: The Changes in Third Party Presence at Public and Private European Websites." In *The World Wide Web Conference*, WWW '19, 1590—-1600, New York, NY, USA: Association for Computing Machinery.

Tobin, O. (2019). "GDPR: 3 Areas of Ambiguity." *Privacy & Data Protection*, *20*(2), 15–16.

Urban, T., Tatang, D., Degeling, M., Holz, T., and Pohlmann, N. (2020). "Measuring the Impact of the GDPR on Data Sharing in Ad Networks." In *Proceedings of the 15th ACM ASIA Conference on Computer and Communications Security*, ASIA CCS '20, New York, NY, USA: Association for Computing Machinery.

Vogel, D. (1995). *Trading Up: Consumer and Environmental Regulation in a Global Economy*. Harvard University Press.

WhotracksMe (2018). "GDPR – What Happened?" Tech. rep.

Zhuo, R., Huffaker, B., claffy, k., and Greenstein, S. (2019). "The Impact of the General Data Protection Regulation on Internet Interconnection." *Working Paper*.