The Early Impact of GDPR Compliance on Display Advertising: The Case of an Ad Publisher

Pengyuan Wang

Li Jiang

Jian Yang

Pengyuan Wang is Assistant Professor of Marketing, Department of Marketing, University of Georgia, 630 S. Lumpkin St., Athens, GA 30602, United States. Li Jiang is Assistant Professor of Marketing, Department of Marketing, George Washington University, 2201 G Street NW, Washington, D.C. 20052, United States (email: lijiang1@gwu.edu). Pengyuan Wang and Li Jiang contributed equally. The research received financial support from the Program on Economics & Privacy at Antonin Scalia Law School. The authors thank the seminar participants of the Research Roundtable on Regulating Privacy. Correspondence should be addressed to Li Jiang.

The Early Impact of GDPR Compliance on Display Advertising: The Case of an Ad Publisher

Abstract

The European Union (EU)'s General Data Protection Regulation (GDPR), with its explicit consent requirement, may restrict the use of personal data and shake the foundations of online advertising. The ad industry predicted drastic loss of revenue from GDPR compliance and has been seeking alternative ways of targeting. Taking advantage of an event created by an ad publisher's request for explicit consent from users with EU IP addresses, the authors find that for a publisher that uses a pay-per-click model, has the capacity to leverage both user behavior and webpage content information for advertising, and observes high consent rates, GDPR compliance leads to modest negative effects on ad performance, bid prices, and ad revenue. The changes in ad metrics can be explained by temporal variations in consent rates. The impact is most pronounced for travel and financial services advertisers and least pronounced for retail and consumer packaged goods advertisers. The authors further find that webpage context can compensate for the loss of access to users' personal data, as the GDPR's negative impact is less pronounced when ads are posted on webpages presenting relevant content. The results suggest that publishers and advertisers should leverage webpage-content-based targeting in a post-GDPR world.

Keywords: GDPR, privacy policy, data regulation, online advertising

Abstract Word Count: 198

Consumer data has become a valuable asset in the digital economy. As part of this data economy, the online advertising industry in the United States generated $107.5 billion in revenue in 2018, an increase of 21.8% over the previous year (IAB 2019). Online tracking and behavioral profiling play key roles in matching relevant ads to users, but they also elicit privacy concerns (Ur et al. 2012). To protect user privacy, the General Data Protection Regulation (GDPR) came into effect in 2018; it is considered a milestone for privacy legislation and a game changer for the digital ad industry (Satariano 2018). This regulation imposes restrictions on companies that collect and process personal data from users with EU IP addresses, and it mandates opt-in consent from users. Consent must be unambiguous and explicit (i.e., pre-ticked boxes or inactivity do not constitute consent) and can be revoked at any time (GDPR Art.4(11), Art.7, see Web Appendix I for more details). Fears were raised from the online advertising industry (Ghosh 2018; IHS Technology 2015), as limitations in personal data collection may lead to less-accurate matches between ads and users and incur an adverse impact on ad performance (Goldfarb and Tucker 2011a). Advertisers, faced with lower ad performance, may react by lowering their bids or leaving the publisher, which will reduce the publisher's revenue.

The advertising industry had predicted a drastic decline in ad performance and revenue, owing to the GDPR (e.g., Deloitte 2013; Ghosh 2018). Deloitte estimates that for online behavioral advertising, the GDPR could lead to direct sales losses amounting to €3.2 billion in the EU, which translates to an estimated loss of €4.2 billion in GDP and 66,000 jobs (Deloitte 2013). Given that the total online ad revenue is €19.3bn in the EU in 2017,[1] the €3.2 billion loss is around 17% of the EU's yearly ad revenue. For direct marketing, Deloitte forecasts that the GDPR could result in sales losses amounting to €62.5 billion in the EU, or an estimated loss of

[1] https://iabeurope.eu/all-news/european-digital-advertising-market-has-doubled-in-size-in-5-years/.

€85 billion in GDP and 1.3 million jobs (Deloitte 2013). The CEO of the Interactive Advertising Bureau of Europe suggested that the GDPR may "limit digital advertising's ability to continue to deliver a wide range of online content to users" (IHS Technology 2015). Such drastically pessimistic predictions may not be valid, though. For example, studies suggest that people do opt in when GDPR consent is elicited (e.g., Godinho de Matos and Adjerid 2022; Solove 2021). Different webpages and advertisers may also experience different levels of impact. Thus, despite the industry's pessimistic prediction, it is necessary to study the effect size and implications of the GDPR.

The GDPR calls for advertising models that rely less on personal data. Targeting based on user behavior usually requires users' personal data; however, placing ad creatives on webpages that have relevant content—also referred to as contextual targeting (Zhang and Katona 2012)—does not rely on user tracking or profiling and has gained renewed attention (Davies 2018; Ghosh 2018; van Bentheim 2020). This approach aligns ads to webpages that have relevant topics rather than to users. When user personal data is less available due to the GDPR, ads on webpages that have relevant topics are able to reach interested consumers and, thus, may be less affected. Many practitioners suggest that webpage-content targeting may be the future trend in a post-GDPR world; in fact, following the GDPR's rollout, ad agencies began shifting their ad budgets toward webpage-content targeting (Davies 2018). Thus, it is imperative to investigate this latter approach as an alternative or addition to personal-data-based targeting in a post-GDPR world.

In addition, previous research on online advertising tends to focus on companies using the pay-per-impression pricing model (e.g., Johnson, Shriver, and Du 2020); there is scant research on the impact of privacy regulation using companies with the pay-per-click model. The pay-per-click model is also widely used and studied (e.g., Najafi-Asadolahi and Fridgeirsdottir 2014). Hence, it is important to provide insights into the GDPR's impact on such companies.

We obtain a proprietary large-scale ad dataset directly from a large publisher headquartered in the US and has global traffic.[2] The publisher has advertisers from a variety of industries, and covers a range of webpages on various topics. It employs a pay-per-click pricing model.[3] It relies on user personal data to serve relevant ads to users, and also has extensive capacities to leverage webpage context for advertising. The focal publisher's characteristics provide opportunities to fill in research gaps and provide useful implications. First, there is a substantial amount of heterogeneity among the data in webpage content and advertisers' industries, which offers an opportunity to study the GDPR's heterogeneous effect. Second, the publisher has extensive capabilities to use webpage content for advertising in addition to user personal data. Many publishers use both behavioral and contextual information for advertising, as research shows that they work together to improve ad performance (e.g., Lu et al. 2016); with this data, we uncover the GDPR's impact on a publisher who can leverage both user-data-based targeting and webpage-based targeting. More importantly, while the GDPR limits personal data, the utilization of the webpage content may not be affected. Thus, we can examine to what extent webpage content can compensate for the loss of user personal data due to the GDPR. Third, the publisher

[2] The GDPR has a global reach and applies to all companies doing business with people located in the EU, regardless of the company location. Even if a company is not headquartered in the EU, if it offers goods or services to people located there (GDPR Art. 3(1)), it must comply with the GDPR when doing business with EU residents.

[3] The pay-per-click model is different from the pay-per-impression model that is often used in real-time bidding auctions (RTB, see Tunuguntla and Hoban 2021). In an RTB platform with a pay-per-impression model, advertisers pay for each ad impression regardless of whether the user clicks; hence, to reach users who would click, they need to leverage user information to control when and to whom their ads to be shown. In contrast, under the pay-per-click model, advertisers only pay if their ad creatives are clicked on, and thus they are guaranteed to obtain clicks with each dollar spent. Meanwhile, the number of clicks is directly related to the publisher's revenue, and hence the publisher needs to use various information to maximize clicks and ad revenue. (For more differences between the two payment models, see Asdemir, Kumar, and Jabob 2012.) Accordingly, in the focal publisher's ad system, the publisher (rather than advertisers) sets up models to use various information to serve relevant ads to users. See the "Data" section and Web Appendix III for details.

employs a pay-per-click model, in which an advertiser only pays when its ad is clicked on; our study is among the first to offer insights into the impact of privacy regulation on publishers using this payment model.

Similar to Facebook, the publisher has its own ad management system, and runs ad auctions and delivers ads using its own system. The publisher began requesting explicit consent from users with EU IP addresses on April 18, 2018 (we refer to this event as "GDPR compliance," hereafter), well before the compliance deadline (May 25, 2018), and observed high opt-in consent rates among users. Users not providing consent remain exposed to ads, but the ads are not targeted using personal data. Our data covers 3.7 billion ad impressions from around 6,000 ad creatives five weeks before and five weeks after the company's GDPR compliance, and the whole study period is before the compliance deadline. Our treatment group comprises users with EU IP addresses and our control group comprises users with non-EU IP addresses. We use a difference-in-differences (DID) model to compare the ad metrics five weeks before versus five weeks after GDPR compliance between the two groups. Our dataset covers ad performance, bid prices and ad revenue. For ad performance, we focus on two commonly used metrics: the click-through rate (CTR) and the conversion rate (CVR). Under the pay-per-click model employed by the publisher, an advertiser only pays when a user clicks on its ad. Thus, the click-through rate is critical to the publisher and the conversion rate is a key performance metric for the advertisers. The dataset also includes the advertisers' bid prices—which reflect advertisers' willingness to pay, and revenue per click—which is directly related to the publisher's advertising revenue and the advertisers' costs. In Web Appendix I, we include a conceptual framework of the changes during the study period.

In contrast to most GDPR research that uses data collected by intermediaries from multiple publishers (e.g., Aridor, Che, and Salz 2021; Goldberg, Johnson, and Shriver 2022; see Table 1 for a comparison), our data comes from a single publisher. This setting has advantages in several respects. First, industry reports suggest that responses to the GDPR vary across online publishers and that GDPR consent strategies and consent rates vary wildly across sites (ICO 2019). Some publishers apply a "GDPR-everywhere" approach and treat non-EU users similar to EU users. For example, Microsoft implemented the same data-protection mechanisms for EU and non-EU users.[4] Among publishers applying the GDPR policy only to EU users, some implemented the required consent mechanism, yet others either did not comply or they exited the EU market (ICO 2019). Therefore, aggregating across publishers that use different compliance strategies may be problematic. Focusing on a single publisher allows us to avoid the complicated nature of various consent strategies. Second, the focal publisher requests explicit consent only from users with EU IP addresses while keeping non-EU traffic intact (which is the standard practice). Therefore, we are able to clearly define users with EU IP addresses as the treatment group and users with non-EU IP addresses as the control group. Third, we examine a relatively short period during which the publisher did not implement strategic changes. Also, the publisher required GDPR compliance on April 18, 2018, well before the compliance deadline (May 25, 2018). Consequently, this greatly reduced spillover effects from other publishers. Finally, obtaining data directly from a publisher allows us to circumvent the data-recording biases caused by the GDPR's consent requirement. As required by the GDPR, intermediaries (e.g., third-party vendors such as analytics platforms) can record users' data only when users have given their consent. Thus, intermediaries may lose data from non-consenting users, which makes it difficult to isolate the true impact of the GDPR from the data-recording bias.

[4] https://blogs.microsoft.com/on-the-issues/2018/05/21/microsofts-commitment-to-gdpr-privacy-and-putting-customers-in-control-of-their-own-data.

Our analyses show that for a publisher that uses a pay-per-click model, has the capacity to use both user personal information and webpage content for advertising, and observes high consent rates, GDPR compliance leads to modest decreases in ad performance, advertisers' bid prices, and the publisher's ad revenue. The estimated decrease for revenue-per-click is 5.7%, and we show that both the decrease in the bid prices and in the number of active advertisers have contributed to a decline in the publisher's revenue. We also find evidence that the reduction in the proportion of consenting users after GDPR compliance (and, thus, the loss of user personal data and the limited ability to target users) underlies these decreases.

[INSERT TABLE 1 ABOUT HERE]

We then proceed to break down these effects using heterogeneity analyses on ads for different industries. We find that the GDPR hurts ads for travel and financial services the most and ads for retail and CPG/consumer products the least. Within the retail industry, ads for more specific products are more affected.

Finally, to examine webpage-content-targeting as an alternative to personal-data-based targeting, we find that the negative effects on the ad metrics are less pronounced for webpages on specific topics (e.g., sports) when the topic matches the advertised product, relative to webpages on general topics (e.g., assorted news). The results suggest that relevant webpage content can partially compensate for the loss of user personal data.

The GDPR is the start of a series of state- and company-level privacy policy changes (e.g., Apple's App Tracking Transparency framework, Google's termination of third-party tracking in Chrome) that are already shaking the foundations of and marking a new era for online advertising. Our study shows that the GDPR's impacts can be alleviated by relevant webpage context. In a post-GDPR world, with the cessation of third-party tracking, our findings may help companies understand the changes that come with new regulations and identify alternative solutions.

Literature Review

Although debates on privacy and online ads have been voluminous (Acquisti, Taylor, and Wagman 2016), to our knowledge there is limited research that directly examines the impact of privacy regulation on online ads. Most relevant to our current research is the work of Goldfarb and Tucker (2011a), who use self-reported purchase intentions to investigate the impact of the 2002 EU Privacy and Electronic Communications Directive (e-Privacy Directive) on ad effectiveness in the EU and find that purchase intentions fall by 65% for exposed (vs. non-exposed) consumers. However, their study is based on self-reported purchase intentions rather than on actual ad metrics and it only considers ad performance. In contrast, our study is in the context of the GDPR and we utilize a natural experiment with an array of actual ad metrics—including ad performance, bid prices, and ad revenue—which jointly provide insights to both the publisher and the advertisers. Furthermore, the focal publisher required GDPR compliance well before the compliance deadline, which has greatly reduced spillover effects from other publishers.

Using data from an anonymous intermediary, Aridor, Che, and Salz (2021) examine the impact of GDPR consent on keyword-based search ads in the travel industry. They find that GDPR consent reduces the number of clicks of keyword-based search ads and that its impact on the overall revenue for advertisers and websites is negative but not statistically significant. That study is subject to data-collection bias from third-party intermediaries due to GDPR compliance;

that is, given the GDPR's explicit consent mechanism, if a consumer does not consent to data sharing with third-party intermediaries, she would not be included in the data provided by the intermediary, as stated in Goldberg, Johnson, and Shriver (2022). In contrast, our data is directly obtained from a publisher, which avoids data-collection bias from third-party intermediaries. Also, in Aridor, Che, and Salz (2021), if a user does not consent, she is no longer exposed to ads. In our work, however, the user remains exposed to ads, but the ads are not targeted using personal data (see the "Data" section for details).

Our paper also relates to the evolving literature that analyzes the impact of the GDPR in other business domains. Regarding consent, the literature finds that EU users gave opt-in consent once GDPR consent was elicited (Godinho de Matos and Adjerid 2022). Regarding tracking technology, third-party cookies decreased by 22% (Libert et al. 2018), trackers declined by 9% (Lukic, Miller, and Skiera 2021), and third-party web technology providers fell by between 3.1% and 12.8% (Peukert et al. 2022) in the EU relative to non-EU regions after GDPR compliance. The GDPR was related to modestly lower web traffic, e-commerce orders and revenue (Goldberg, Johnson, and Shriver 2022; Schmitt, Miller, and Skiera 2022), lower venture capital investment in the EU (Jia, Jin, and Wagman 2019, 2021), and more search frictions (Zhao, Yildirim, and Chintagunta 2021), and it had no effect on publishers' content (Lefrere et al. 2022). The regulation also hurt small web technology vendors more than big vendors (Johnson, Shriver, and Goldberg 2023; Peukert et al. 2022). These studies consistently show that the impact of the GDPR is more nuanced than the industry's expectations. As Johnson (2022) points out, among the challenges to studying the GDPR's impact, a lack of a clean control group and that the GDPR impacts data observability pose big hurdles to researchers. Different from most GDPR studies, our data comes directly from a publisher, and allows us to have a clean control group and circumvent data observability issues.

Our research also contributes to the empirical literature on behavioral data and targeting. The literature suggests that behavioral data contributes to ad performance by increasing the precision of the targeting (see Bleier and Eisenbeiss 2015; Farahat and Bailey 2012; Rafieian and Yoganarasimhan 2021; Yan et al. 2009). A stream of theoretical research suggests a non-monotonic relationship between behavioral data and ad revenue, where a moderate usage of behavioral data increases ad revenue, but too much behavioral data leads to narrow targeting and a reduction in competition among advertisers—and, thus, a decrease in ad revenue (Chen and Stallaert 2014; Hummel and McAfee 2016; Levin and Milgrom 2010).

Most empirical research demonstrates that the use of behavioral data leads to an increase in ad revenue, though the magnitude varies. For example, using ad-transaction data, Beales and Eisenach (2014) find that behavioral tracking augments publisher revenue (cost per thousand impressions) by 66%. Johnson, Shriver, and Du (2020) find that when consumers opt out of behavioral tracking through AdChoices, this results in a 52% decline in ad exchange revenue. Laub, Miller, and Skiera (2022) obtain data from a large European ad exchange in 2016 and find that after controlling for differences in users, advertisers, and publishers, when user tracking is unavailable, ad price decreases by 18%. Marotta, Abhishek, and Acquisti (2019) obtain ad revenue data from a single large publisher and find that when a user's cookies are available, the publisher's revenue increases by only about 4%. Danaher (2022) applies optimal control theory to micro targeting and shows that firms' profits can be improved by more than 150%. These findings show that, when less behavioral data is used, ad revenue is likely to decrease, but the reliance on behavioral data varies and also depends on the targeting methods. We complement previous research with a natural experiment created by GDPR compliance and demonstrate the value of user personal data in this context.

In contrast to using behavioral data, another targeting strategy is through webpage content—also referred to as “contextual targeting” (Zhang and Katona 2012). Early research on webpage context is mostly conducted in the lab, which shows webpage context can increase ad performance (e.g., Moore, Stammerjohan, and Coulter 2005). An increasing number of field studies confirms this conclusion (Ada, Abou Nabout, and Feit 2022; Goldfarb and Tucker 2011a, b; Lu, Zhao, and Xue 2016; Rafieian and Yoganarasimhan 2021). Rafieian and Yoganarasimhan (2021) estimate a structural model and use counterfactuals to compare the effects of personal data and contextual information in predicting click-through rates and find that personal data is more useful than contextual information for ad performance. Lu, Zhao, and Xue (2016) find that contextual information and personal data combined are more effective in improving users’ clicks. As for ad revenue, Ada, Abou Nabout, and Feit (2022) find that advertisers value context information in addition to users’ personal data; that is, when an ad exchange provides subdomain information (i.e., ad context) to ad buyers, revenue per impression rises. Similar to Ada, Abou Nabout, and Feit (2022), our study is based on a natural experiment owing to actual policy changes of the GDPR.

Data and Metrics

Data

We obtained a large-scale display ad dataset from a large online ad publisher headquartered in the US (and the publisher is among the top 50 properties ranked by ComScore). The publisher owns a variety of content feed webpages (i.e., webpages with streams of content, such as news and articles). It collects user behaviors and interests from its own webpages, and also purchases data from external data platforms for advertising. 47% of the publisher's ad impressions are from webpages with specific topics (e.g., sports, finance), so the publisher also has extensive capacities to leverage webpage content for advertising. It inserts ad slots into the content feed; these ads are also regarded as native ads (see an example in Web Appendix II, Figure W2.1). The webpages in our dataset only include native display ads with ad texts and static images; they do not include other types of ads, such as banner ads.[5]

Native ads have experienced substantial growth in recent years. In the US market, for instance, in 2018, native ads constituted $33 billion (58% of display ad revenue) (eMarketer 2018). In 2020, native ads comprised $53 billion, accounting for 65% of display ad revenue.[6] Globally, native ads are predicted to increase by 372% between 2020 and 2025 and be worth $400 billion in 2025 (Glenday 2019). An example of a content feed webpage with multiple ads is shown in Web Appendix II, Figure W2.2. Our dataset covers 3.7 billion desktop impressions of around 6,000 ad creatives from 2,200 advertisers. Hereafter, we refer to an ad creative as an "ad."

The collaborating company's ad traffic is global, including all 28 EU countries (including the UK)[7] and 163 non-EU countries and regions. According to the GDPR, normally a company should implement the required consent, data collection and processing mechanisms for EU subjects, while leaving non-EU subjects intact. Following this requirement, the focal publisher

[5] Display ads can take the form of a banner, rich media, and in-feed native ads (IAB 2021). Because in-feed native ads target users the same way as banner ads do and account for the majority of the display ads (eMarketer 2018), we use in-feed native ads as a window through which to examine the effects of GDPR compliance.
[6] Native ads spending in 2020 is available at https://www.statista.com/statistics/369886/native-ad-spend-usa/. The proportion is calculated as the ratio of native ad spending and total display ad spending in 2020 (the total is $81.38 billion, see: https://www.statista.com/statistics/273443/online-display-advertising-revenue-in-the-united-states).
[7] The UK left the EU in 2020. However, because in 2018 the UK was part of the EU and was subject to GDPR during the study period, we include the UK in the EU group.

requests explicit consent to data collection and ad personalization from users with EU IP addresses. The publisher's ad traffic is concentrated in fifteen countries: ten EU countries (Belgium, France, Germany, Italy, the Netherlands, Poland, Romania, Spain, Sweden, and the UK, accounting for more than 95% of the EU traffic) and five non-EU countries (the US, Canada, Australia, Brazil, and Mexico, accounting for more than 95% of the non-EU traffic). The data from the collaborating company includes country labels for each of the fifteen major countries; the rest of the EU countries are aggregated into a single "other EU countries" label, and the rest of the non-EU countries and regions are aggregated into a single label titled "other non-EU countries." The focal publisher remains in the EU market; it does not decline EU webpage visitors nor does it change the number or positions of the ad slots on its webpages.

Similar to Facebook, the publisher provides an ad management system for advertisers to set up and manage their ads; it also serves ads (i.e., delivers ads to webpage viewers) using its own system. The publisher directly observes various advertising metrics, IP addresses, advertisers, and ad designs, without intermediaries. Similar to major ad publishers (Chapelle et al. 2015; Ling et al. 2017), the focal publisher builds its ad system and conducts ad targeting using various features together, such as users' personal data and online behaviors, webpage content, ad content (e.g., ad texts and images), and their interactions. The user data includes data collected from the users' behaviors on this publisher's own webpages and data bought externally such as from data management platforms (DMPs).[8] The publisher did not make strategic changes to its ad models during the study period. For example, although less user personal data might be collected after GDPR compliance, the company did not eliminate the variables that represent user characteristics from its ad system.

[8] The utilization of both data sources is subject to users' consent to ad personalization, i.e., if a user does not give consent, the publisher will not utilize data from either source for advertising.

The publisher conducts generalized second-price auction (for details, see Tunuguntla and Hoban 2021) and advertisers pay for clicks (i.e., following a pay-per-click model; see Asdemir, Kumar, and Jacob 2012 for more details of this payment model). Specifically, when a user loads a webpage that includes ad slots, the publisher's system will estimate the CTR (represented by eCTR) for each candidate ad, based on a range of features such as ad characteristics, user features, webpage content and the interactions between these factors. It then calculates the estimated cost per 1000 impressions (eCPM) as eCPM = bid×eCTR×1000. There are multiple ad positions in a stream, and the publisher allocates the ad positions according to the ranking of eCPM (i.e., it gives the top slot to the highest-ranked bidder, the second top slot to the second-highest-ranked bidder, and so on). If an ad is clicked on, the focal advertiser pays the adjusted second price (i.e., the bid of the advertiser ranked right below the focal one), capped by the focal advertiser's own bid. As we mentioned in the introduction, under a pay-per-click model, the advertisers are guaranteed to obtain clicks with their ad spend. The advertisers do not set up ad triggering rules in the focal publisher's ad system; instead, the publisher uses various information (e.g., ad characteristics, user characteristics, webpage content, and interactions between these factors) to serve relevant ads to users. Indeed, it is generally agreed that publishers have a superior capability to conduct targeting; for example, Hu, Shin, and Tang (2016) described that "publisher can automatically match the advertisement to consumers who are most likely to be interested in it by using a targeting technology based on superior knowledge of its consumers' demographics, geographical location, expressed interests, and other information". We also provide more details of the publisher's ad model in Web Appendix III.

While some demand side platforms (DSPs) with real-time bidding strategies change bid prices from impression to impression on behalf of advertisers, the focal publisher does not do so;

it simply executes the advertisers' pre-specified bid prices. The publisher does provide real-time performance monitoring; it also allows advertisers to change bid prices as often as they like. Thus, advertisers can calculate ad performance as frequently as they like and adjust their bids if low ad performance is observed. The publisher does not allow advertisers to set up different bid prices for consenting and non-consenting users, and it does not provide separate ad metrics for consenting and non-consenting users. We provide more details of how advertisers manage their bid prices in the publisher's system in Web Appendix III.

The publisher updated its consent system to request explicit consent to data collection and ad personalization from users with EU IP addresses, beginning on April 18, 2018. Specifically, if a user had an EU IP address, the publisher used a pop-up window on its webpages to request consent. The pop-up window covered half of the webpage with the following message: "Click 'Agree' to allow us and our partners to use cookies and similar technologies to collect and use your data to understand your interests and provide personalized ads. Learn more about how we use your data in our Privacy Center. Once you confirm your privacy choices, you can make changes at any time by visiting your privacy dashboard." Together with the message, the publisher provided two buttons for users to indicate whether or not they wanted to give consent. The message was only shown to users who had not made a choice. Once the choice was made, regardless of whether it was consent or non-consent, the consent pop-up window would not be displayed the next time the user visited the site. The default setting was "opt-out" (i.e., if a user did not make a choice, the advertising algorithm would consider the user to be non-consenting). The consent procedure was the same for all EU countries and for registered and unregistered site users. Access to the website content was not conditional on giving consent. Upon not receiving explicit consent from users with EU IPs, the publisher stopped collecting or using personal data for ad targeting, including data collected from its own webpages and purchased externally. The publisher pre-announced the GDPR compliance date (April 18, 2018) so advertisers knew the date and could monitor their ad performance and change their ad bid prices if necessary.

Our dataset ranges five weeks before to five weeks after the publisher's GDPR compliance on April 18, 2018 (i.e., from March 14, 2018, to May 22, 2018). The provided data is aggregated for each combination of date, country, webpage, ad and ad slot position. The data is aggregated across non-consenting and consenting users. In Table 2.1, we summarize the key descriptive statistics.[9] We also obtain daily consent rates aggregated across the publisher's webpages after GDPR compliance. For a given day, the consent rate is defined as the percentage of consenting users among all visitors of the publisher's webpages *within the day*. A high proportion of website visitors give consent,[10] and there is temporal fluctuation in daily consent rates, as shown in Figure 1.

One limitation of the dataset is that, except for users' countries, the dataset does not cover user characteristics such as interests, demographics, or granular geo-location information. Also, we do not have separate data for consenting and non-consenting users. Therefore, we focus on the overall impact of GDPR compliance on the focal publisher and its advertisers. We leave the interplay among the policy, consent, and users for future research, when reliable data is available.

[INSERT FIGURE 1, TABLES 2.1 and 2.2 ABOUT HERE]

Metrics

---

[9] Our agreement with the publisher prevents us from sharing their confidential information. Hence, the descriptive statistics in the upper panel of Table 2.1 are in proportions rather than absolute scales, and the descriptive statistics in the lower panel of Table 2.1 are multiplied by a constant number to disguise these key business metrics.

[10] Our agreement with the publisher prevents us from sharing the exact consent rates, but the average consent rate is high. Also, to protect the publisher's confidential business information, in Figure 1, we report the consent rates in relative scales (i.e., relative to the average after GDPR compliance) instead of absolute scales.

We summarize the definitions of the ad metrics in Table 2.2. For ad performance, we focus on two commonly used metrics: the click-through rate and the conversion rate. The click-through rate is the likelihood of clicking on an ad upon impression, and is one of the most commonly used ad performance metrics for online advertising (e.g., Dinner, Heerde Van, and Neslin 2014). In a pay-per-click model, an ad impression only contributes to a publisher's revenue if the ad is actually clicked on. Therefore, click-through rate is a critical ad performance metric that directly affects a publisher’s revenue. The conversion rate is the number of conversions per click, which reflects the value generated by the advertisers’ paid clicks and, hence, is a critical performance metric for advertisers. The click-through and conversion rates correspond to different stages in the online purchase funnel in a web viewer’s journey from ad impression to click to conversion (Wiesel, Pauwels, and Arts 2011): the click-through rate is the transition rate at the upper funnel (i.e., from impression to click), and the conversion rate is the transition rate at the lower funnel (i.e., from click to conversion). Therefore, although both measure ad performance, they reflect ad performance for different players and at different stages along the purchase funnel.

We also investigate ad revenue, as represented by the revenue per click. The publisher conducts generalized second-price auction (Tunuguntla and Hoban 2021) with a pay-per-click model. When a user loads a webpage that includes ad slots, the publisher’s algorithm ranks the advertisers, based on the bids and other features such as ad characteristics, user features, webpage content and the interactions between these factors. If an ad is clicked on, the focal advertiser pays the adjusted second price (i.e., the bid of the advertiser ranked right below the focal one), capped by the focal advertiser’s own bid. Hence the final cost per click for the advertisers (which is also the revenue per click for the publisher) depends on the bid price and the competition among active advertisers. See details of the bidding procedure and ad cost calculation in Web Appendix III.

In this study, we also examine the advertisers' bid prices. Under the publisher's generalized second-price auction, an advertiser's bid price is the highest amount of money it would pay for a click and thus reflects the advertiser's willingness to pay. Previous literature shows that advertisers place bids according to the quality of each ad opportunity (Arnosti, Beck, and Milgrom 2016), which is usually based on the probability of conversion (Lee et al. 2012). The publisher reports the ad performance to its advertisers,[11] and the advertisers may change their bid prices accordingly, whenever needed (see Web Appendix III). The provided dataset includes the bid price for each combination of the date, country, webpage, ad and ad slot position. Because the bid price may change within a day, for each combination of the date, country, webpage, ad and ad slot position, the bid price is a weighted average, where the weight is the length of time each bid price holds within the same day.[12] The bid price can also be weighted by the number of impressions, and as a robustness check we include the "bid weighted by impressions" in Web Appendix IV.

In addition to the metrics introduced above, we also include the conversion per impression and the revenue per impression in Web Appendix IV. The former indicates the overall ad performance from impression to conversion; the latter provides further financial implications to the publisher, as this is directly related to its total revenue. These metrics together (the click-through rate, conversion rate, bid price, revenue per click, conversion per impression and revenue per impression) offer a holistic view of the impact of GDPR compliance on both the publisher and the advertisers.[13] We summarize the key statistics of the metrics in the lower panel

[11] The publisher does not report the consent rate to its advertisers, nor does it report the separated ad metrics for users with or without consent; hence, the advertisers may only observe the overall ad metrics for all users combined.
[12] When we calculate the bid price weighted by the length of time, a bid price is incorporated regardless of whether it wins impressions.
[13] Tracking of the daily ad metrics in our dataset is not affected by the GDPR. Specifically, our ad metrics are counted (i.e., aggregated) without identifying each user. For a given ad on a webpage on each day, the total number of impressions and clicks can be counted without identifying each user. The ads in our dataset specify certain actions on ad landing pages as conversions (e.g., obtaining quotes, adding items to shopping carts), and the total number of

of Table 2.1.[14]

Before we investigate the above ad metrics, we check the stability of the volume of ad impressions and the composition of these impressions before and after GDPR compliance. The results are reported in Web Appendix V. In EU and non-EU regions, we find no significant change in the ad impression volume or in the proportion of ad impressions from webpages on specific topics.

Model-free results

We present the model-free results for EU and non-EU regions (identified by IP addresses) in Figure 2. The plots show that in the EU, all the ad metrics decrease after GDPR compliance. See Web Appendix IV for additional metrics, which show similar trends.

[INSERT FIGURE 2 ABOUT HERE]

Main Effect Model and Results

Main Effect Model

We use a DID model to assess the early impact of GDPR compliance as in Model 1. The treatment group is composed of users with EU IP addresses and the control group is composed of users with non-EU IP addresses.

(1)
$$Y_{acwpt} = \alpha + \beta \, \text{Treated}_c \times \text{Post}_t + \alpha_a + \alpha_c + \alpha_w + \alpha_p + \alpha_t + \text{other control variables} + \varepsilon_{acwpt} .$$

conversions can also be counted. The publisher can thus obtain the daily-*aggregated* number of impressions, clicks, and conversions, which are not affected by the GDPR. The bid prices and revenue are recorded in the publisher's ad management system, the recording of which is not affected by the GDPR either.

[14] We also calculate the difference between the advertisers' bid prices and the payments as a proxy for the advertisers' surplus. For the analysis, see the section titled "Advertisers' Surplus."

Y_{acwpt} refers to each ad metric, collected for each combination of ad, country, webpage, and ad slot position in the feed on each day, where a, c, w and p denote the ad, country, webpage, and ad position, respectively, and t indexes the date. Y_{acwpt} is aggregated across non-consenting and consenting users. $\mathrm{Treated}_c$ is a binary variable indicating whether the data point is from the treatment group. Post_t is a binary variable that equals 1 if the time period is on or after April 18, and 0 otherwise. ε_{acwpt} is the error term. The ad, country, webpage, ad position, and date fixed effects are denoted by $\alpha_a, \alpha_c, \alpha_w, \alpha_p, \alpha_t$, respectively. The other control variables are a set of two-way interaction fixed effects, including the fixed effects for each pair of ad and country, ad and ad position, ad and webpage, country and ad position, country and webpage, and ad position and webpage.[15] We include only the interaction term $\mathrm{Treated}_c \times \mathrm{Post}_t$ in the model; the $\mathrm{Treated}_c$ and Post_t main effects are subsumed into the country and date fixed effects, so they do not need to be included in Model 1. Hence, β, the coefficient of the interaction term $\mathrm{Treated} \times \mathrm{Post}$ estimates the average treatment effect among the treatment group (i.e., EU users) and captures the impact of GDPR compliance. We obtain robust standard errors clustered at the ad and webpage level to adjust for intra-ad and intra-webpage correlation, similar to Goldfarb and Tucker (2011a).

Main Effect Model Results and Robustness

Table 3 reports the results of Model 1. The coefficient of $\mathrm{Treated} \times \mathrm{Post}$ is negative and significant for all the metrics, showing that all the metrics decreased after GDPR compliance. Relative to the pre-treatment average in the treatment group, the click-through rate decreases by

[15] The model results are robust to the existence of the interaction fixed effects. Without these, the model yields similar results for the focal variable $\mathrm{Treated} \times \mathrm{Post}$ and the conclusions stay the same.

2.1%,[16] and the conversion rate decreases by 5.4%. The results suggest that GDPR compliance leads to a modest reduction in ad performance for the focal publisher.

[INSERT TABLE 3 ABOUT HERE]

Consistent with the decrease in conversion rate (5.4%), the ad bid price decreases by 5.6%. Since conversion rate represents the business value advertisers gain from each click, this result is consistent with existing literature suggesting that the ad price depends on the overall quality of the ad opportunity (Arnosti, Beck, and Milgrom 2016). Following the 5.6% decrease in the bid price, revenue per click falls by 5.7%. The magnitude of the decrement in all of the ad metrics is modest, which may be related to the high consent rate. It is consistent with contemporary GDPR literature (e.g., Aridor, Che, and Salz 2021; Goldberg, Johnson, and Shriver 2022; Schmitt, Miller, and Skiera 2022), which also finds modest impact from the GDPR.

The decreases in conversion rate (5.4%) and revenue per click (5.7%) are of similar magnitude. This suggests that the cost per conversion (= revenue per click/conversion rate)—which is inversely related to the advertisers' return on investment (ROI)—is not impaired much by GDPR compliance. These results suggest that under the publisher's pay-per-click model, if advertisers are well calibrated, their ROI is comparable to the pre-GDPR level. Comparing the publisher's loss of revenue with the advertisers' stable ROI, our analyses suggest that the GDPR's negative impact falls primarily on the publisher rather than the advertisers.

We include a series of robustness checks in Web Appendix VI, which consistently verify our results. In Web Appendix IV, we also replicate the results for the additional outcome metrics.

DID Assumption Check

[16] We compute the percentage decrease by calculating the ratio of the coefficient of Treated × Post (-.00018) and the pre-treatment average click-through rate in the EU (.0086), so we obtain the 2.1% decrease. We compute all other metrics using the same method.

For the DID analysis, it is important to verify that the observed effects do not occur before the treatment. Following a widely used approach in the literature (Chen, Hong, and Liu 2018), we split the key covariate into a series of time indicators to show the effects over time. Specifically, for each outcome metric, we include the interaction of the `Treated` variable and the date indicators in Model 1 and we depict the coefficients in Figure 3. The plots show that for all of the metrics, the effects do not occur before the event but start to occur afterward. The results provide evidence that the DID assumption is satisfied and that the observed effects can be attributed to the publisher's GDPR compliance. Also, the daily effects demonstrate that ad performance decreases immediately after GDPR compliance and that the ad bid price and revenue per click decline after a short lag.

[INSERT FIGURE 3 ABOUT HERE]

Additional Analyses for the Main Effect

Relationship Between Consent Rates and Ad Metrics

To verify that the decreases in the ad metrics are related to the loss of personal data among non-consenting users, we obtain the post-GDPR daily consent rates in the EU and explore the relationship between the ad metrics and the consent rates. The daily consent rates are aggregated across all of the publisher's webpages (i.e., separate consent rates for each webpage are not provided). We then estimate Model 2 with the daily consent rates as the independent variable and the ad metrics as the dependent variable, using post-GDPR data in the EU.

(2) $$\mathrm{Y_{acwpt}} = \alpha + \gamma\,\mathrm{ConsentRate_t}$$

$$+\alpha_a + \alpha_c + \alpha_w + \alpha_p + \sum_{D=1}^{6} \omega_D \mathrm{DayOfWeek}_t^D + \text{other control variables} + \varepsilon_{acwpt}\,,$$

where $\mathrm{ConsentRate_t}$ is the proportion of consenting users among all visitors within date t, aggregated across the publisher's webpages; its coefficient γ represents the relationship between the consent rates and the ad metrics. We do not include daily fixed effects in this model because they would cause collinearity with the daily consent rates. Instead, we control for the day of the week by using a series of dummy variables dayOfWeek_t^D, where Monday is the baseline, and D ranges from 1 to 6 to represent Tuesday through Sunday; dayOfWeek_t^D equals 1 if date t is the corresponding day of the week. We include all other control variables as in Model 1.

The results of Model 2 are reported in Table 4. The coefficient of the daily consent rates is positive and significant for all the metrics (except that it is marginally significant for the advertisers' surplus; see Web Appendix IV for additional metrics), showing that more non-consenting users correspond to larger decreases in the ad metrics. We also provide a robustness check, where we estimate Model 2 using only data from webpages on general topics (which do not provide clear context for advertising); we find similar results (see Web Appendix VI, Table W6.6), which suggests that our results are not due to the publisher's specific webpage topics. These results show that the decreases in the ad metrics hinge on the share of consenting users. Thus, the reduction in the proportion of consenting users induced by GDPR compliance—and hence the subsequent loss of personal data—is likely to underlie the decreases in the ad metrics.

The above analysis also helps address an alternative explanation that the decreases in the ad metrics may be due to the mere exposure to the consent message; that is, the appearance of the consent message makes all users sensitive to ads and less likely to click. Exposure to the consent message occurs for both consenting and non-consenting users. If this is *the main driver* of the decreases in the ad metrics, then these decreases should be unrelated to the changes in daily consent rates, which are disconfirmed by the results of Model 2. Thus, it is unlikely that, merely confronting the consent message, rather than the decreased targetability among non-consenting users, is the *main driver* of the decreases in the ad metrics. We acknowledge the possibility of such an additional explanation. Future research should examine this explanation more when proper data is available.

[INSERT TABLE 4 ABOUT HERE]

What Contributes to the Changes in Ad Revenue?

Under a generalized second-price auction, the revenue per click depends on the bid prices and the competition among the active advertisers (Chen and Stallaert 2014). Hence the decrease in revenue per click may be related to the advertisers' lower bids or a reduced number of active advertisers on the publisher's webpages. As shown in the section titled "Main Effect Model Results and Robustness," we do observe a decrease in the bid price. In the following analyses, we discuss the changes in the number of active advertisers and how they contribute to the decrease in revenue per click.

We define active advertisers as those making at least one impression on the corresponding day. We obtain the number of active advertisers for each day in EU and non-EU regions. A t-test of the daily number of active advertisers in non-EU regions before and after GDPR compliance suggests no significant change ($p = .44$). A t-test of the daily number of active advertisers in EU regions, however, suggests a decrease of 2.9% after GDPR compliance ($p = .010$).[17] The reduced number of active advertisers in EU regions suggests less competition among advertisers, which may have contributed to the decrease in revenue per click.

One may argue that the decrease in the bid price may be because advertisers who used to

[17] We also use a DID model to estimate the change in the daily number of active advertisers before and after GDPR compliance in EU relative to non-EU regions. We fit the daily number of active advertisers with a binary indicator of the EU, a binary indicator of post-GDPR, and their interaction. The coefficient of the interaction term is negative and significant ($\beta = -21.7$, p-value = .024), which is consistent with the t-test results and suggests a decrease in the daily number of active advertisers in the EU (relative to the non-EU regions) after GDPR compliance.

bid higher prices (i.e., high-value advertisers) had left the publisher after GDPR compliance. We test this explanation by investigating whether there is a change in the mix of the advertisers' bid prices before and after GDPR compliance. Specifically, in EU and non-EU regions, we split the advertisers into two groups—the advertisers who left the publisher after GDPR compliance (defined as having made at least one impression before April 18 but no impression for all the 35 days on or after April 18) and those who remained—and we compare their average pre-GDPR bid prices. If high-value advertisers had left, then we would have seen a difference in the pre-GDPR bid prices between the two groups of advertisers. T-test analyses find no significant difference between the pre-GDPR bid prices of these two groups of advertisers in either EU (p = .49) or non-EU regions (p = .63).[18] Thus, the decrease in the bid price is unlikely to be caused by the departure of the advertisers that used to bid higher prices.

Taken together, our analyses suggest that both the reduced bid prices and the reduced number of active advertisers may have contributed to the decrease in revenue per click. Our results echo Chen and Stallaert's (2014) theoretical finding that a publisher's revenue depends on the advertisers' valuations and the degree of competition.

Advertisers' Surplus

We further study the impact of GDPR compliance on the advertisers' surplus. The bid price reflects the advertisers' willingness to pay, and the publisher's revenue per click (which is also the advertisers' cost) is the actual price paid, so the difference between these two variables is a

[18] We also compare the two groups of advertisers in EU and non-EU regions, using a DID model, where we fit their pre-GDPR average bid price with a binary indicator of the EU, a binary indicator of staying (indicating that the advertiser stays with the publisher after GDPR compliance), and their interaction, and the coefficient of the interaction term is not significant (β = .008, p-value = .98), which suggests no significant difference in the bid price among the two groups of advertisers in the EU relative to the non-EU regions.

proxy for the advertisers' surplus.[19] Theoretical work points out that ad revenue does not always increase with more targeting, as it is also a function of competition (e.g., Chen and Stallaert 2014; Hummel and McAfee 2016). Rafieian and Yoganarasimhan (2021) empirically show that advertisers' surplus decreases with less targeting capability and that contextual targeting alleviates the decrement. However, few studies have empirically examined the change in advertisers' surplus in the context of actual privacy regulations. Using the difference between the bid price and the revenue per click as a proxy for the advertisers' surplus, we provide preliminary insights into the change in the advertisers' surplus after GDPR compliance.

We re-estimate Model 1 using the advertisers' approximated surplus as the dependent variable (See the last column of Table 3). The coefficient of Treated × Post is negative and significant; thus, after GDPR compliance, there was a significant decline in the advertisers' surplus. This result suggests that the advertisers' surplus decreases when personal-data-based targeting is limited. These findings echo Rafieian and Yoganarasimhan (2021, Table 5).

Heterogeneous Effects Across Webpage Contexts

Models: Effect of Webpages on Specific Topics

When less personal data is available, webpage context may help reach interested users and thus alleviate the impact of the loss of personal data. Here, we examine to what extent webpage context compensates for the loss of personal data induced by the GDPR.

The focal publisher owns a variety of content feed webpages—such as homepages, assorted news, entertainment, lifestyle, sports, technology, and finance feed webpages—and

[19] Due to the data limitation, we are unable to precisely estimate impression valuation. The bid price, though non-truthful in generalized second-price auction (Edelman, Ostrovsky, and Schwarz 2007), is still the highest possible payment for a click and hence, indicates advertisers' willingness to pay.

inserts ads into these feeds. Similar to Goldfarb and Tucker (2011a), we label the homepage and assorted news as webpages on general topics, and webpages with entertainment, lifestyle, sports, technology, and finance feeds as webpages on specific topics. Webpages on specific topics provide a clearer context for advertising than webpages on general topics.

Since the webpage topic may or may not be relevant to products presented in ads (e.g., a running shoe ad on a sports webpage may be leveraging the webpage's relevant topics, but a credit card ad on a sports webpage may not), we further classify each pair of ad and webpage into three scenarios: 1) an ad is posted on a webpage on general topics (e.g., assorted news); 2) an ad is posted on a webpage that has a specific topic and the webpage topic is related to the advertised product (e.g., a running shoe ad posted on a sports webpage—and we label such a scenario a "matched ad" and represent it with a binary variable $SpecificTopicMatching$); and 3) an ad is posted on a webpage that has a specific topic but the webpage topic is unrelated to the advertised product (e.g., a credit card ad posted on a sports webpage—and we refer to such a scenario an "unmatched ad" and represent it with a binary variable $SpecificTopicNotMatching$). Two analysts of the collaborating company independently label each pair of ad and webpage with the three scenarios. They reached high inter-coder reliability with Krippendorff's α >.85. Using webpages on general topics (i.e., the first scenario) as the baseline, we obtain the differential impact of the second and third scenarios, with Model 3:

(3)
$$\mathrm{Y_{acwpt}} = \alpha + \beta_1 \mathrm{Treated_c} \times \mathrm{Post_t} \\ + \beta_2 \mathrm{Treated_c} \times \mathrm{Post_t} \times \mathrm{SpecificTopicMatching_{aw}} \\ + \beta_3 \mathrm{Treated_c} \times \mathrm{Post_t} \times \mathrm{SpecificTopicNotMatching_{aw}} \\ + \alpha_a + \alpha_c + \alpha_w + \alpha_p + \alpha_t + \text{other control variables} + \varepsilon_{acwpt} .$$

The control variables include all of the variables described in Model 1.[20] The coefficient of Treated × Post × SpecificTopicMatching captures the differential impact of ads posted on webpages that have relevant topics, and the coefficient of Treated × Post × SpecificTopicNotMatching captures the differential impact of ads posted on webpages that have irrelevant topics, compared to those on general topics.

[INSERT TABLE 5 ABOUT HERE]

Model Results: Effects of Relevant Webpage Content

The results of Model 3 are reported in Table 5. The coefficient of Treated × Post × SpecificTopicMatching is positive and significant, suggesting that ads posted on webpages that have relevant topics are less affected by the loss of personal data. The coefficient of Treated × Post × SpecificTopicNotMatching is not significant, suggesting that the GDPR has similar effects on ads posted on webpages that have irrelevant topics as on ads posted on webpages that have general topics. These results altogether suggest that webpages with clearer ad context alleviate the GDPR's impact and this alleviation effect only exists when webpage topics are relevant to the advertised products. We provide an additional analysis in Web Appendix VII, where we only use data from webpages on specific topics, and we estimate the differential impact of GDPR compliance on ads for relevant products (matched ads) versus ads for irrelevant products (unmatched ads). Consistently, we find that the negative impact of GDPR compliance is alleviated for matched ads compared to unmatched ads. The results further confirm that the *relevant* ad context is associated with a less negative impact of the GDPR.

[20] Also, to study the three-way interactions Treated × Post × SpecificTopicMatching and Treated × Post × SpecificTopicNotMatching, we include the two-way interactions SpecificTopicMatching × Post, SpecificTopicNotMatching × Post, SpecificTopicMatching × Treated, and SpecificTopicNotMatching × Treated. The SpecificTopicMatching and SpecificTopicNotMatching terms are subsumed into the interaction fixed effects of the ad and webpage and so need not be included in Model 3.

The results from Model 3 provide implications for industry players that differ in their abilities to leverage webpage context for advertising. First, in Model 3, the coefficient of $\text{Treated} \times \text{Post}$ estimates the GDPR's impact on webpages that have general topics (because we use webpages on general topics as the baseline). Such webpages do not provide clear context for advertising. The coefficients are more negative than those of $\text{Treated} \times \text{Post}$ estimated using all of the data (as listed in Table 3), suggesting that the GDPR has a more negative effect when webpage context is not leveraged. It provides insights to industry players that are not able to leverage webpage context, such as publishers who do not own webpages focusing on specific topics or niche content.

Second, when the webpages' topics match the advertised products, that is, when the webpage context is leveraged, the positive coefficient of $\text{Treated} \times \text{Post} \times \text{SpecificTopicMatching}$ in Model 3 suggests that the negative effect of GDPR compliance is alleviated. However, the net impact of the GDPR is still negative,[21] suggesting that even when the ads have leveraged the webpage context, they are not immune to the GDPR. To further investigate the impact of the GDPR when webpage context is leveraged, we re-estimate Model 1 only using the data when the advertised products match the webpage topics (in Web Appendix VI Table W6.7); indeed, the impact on the matched ads is still significant and negative.

Model 3's results show that the magnitude of the decreases in ad metrics is different for webpages on general topics versus those on specific ones. If we assume that webpages on general topics do not leverage any contextual information whereas webpages on specific topics can, then by comparing the decreases in the ad metrics on the two types of webpages, we can

[21] For example, for revenue per click, the coefficient of $\text{Treated} \times \text{Post}$ is -.26, and the coefficient of $\text{Treated} \times \text{Post} \times \text{SpecificTopicMatching}$ is .11, thus the net impact is around -.26 + .11 = -.15, which is still negative.

obtain a rough estimate of the relative impact of user personal data versus contextual information on the ad metrics for the focal publisher. For example, based on the results of Model 3, the conversion rate drops by .0021 on general interest webpages after GDPR compliance (since the coefficient of $\text{Treated} \times \text{Post}$ is -.0021); the impact is alleviated by webpages on relevant topics by .00092 (since the coefficient of $\text{Treated} \times \text{Post} \times \text{SpecificTopicMatching}$ is .00092), which is 44% of .0021. Thus, relevant webpage topics roughly compensate for the loss of user personal data by 44%, which suggests the relative impact of personal data to webpage context is 56:44. A similar calculation for revenue per click suggests that relevant webpage topics roughly compensate for the loss of personal data by 42%. Although these back-of-the-envelope analyses depend on the publisher's webpages and how well they match its advertisers, and do not consider the interaction between personal data and contextual information, they provide suggestive evidence that personal data was a substantial component of targeting decision before GDPR compliance, and after GDPR compliance there may be an increasing reliance on webpage content-based targeting. These results also provide a benchmark for other publishers to estimate the effect sizes of the GDPR's impacts based on their own properties.

We also estimate the heterogeneous effects using advertisers' surplus as the dependent variable. Similar to the other metrics, when webpages present specific topics related to the advertised products, the negative impact of GDPR compliance is less pronounced. This result suggests that when personal data is limited, advertisers' surplus is higher when contextual targeting is enabled. Such results are also observed in Table 5 in Rafieian and Yoganarasimhan (2021). We also obtain similar results on the additional metrics (see Web Appendix IV).

These results suggest that relevant webpage content can, to some extent, compensate for the loss of user personal data. Thus, in a post-GDPR world, it is crucial to leverage webpage-content-based targeting and to place ads in a relevant context.

Heterogeneous Effects Across Advertisers

Next, we examine the heterogeneous effects of ads. Following eMarketer (Benes 2019), we classify advertisers (and hence their corresponding ads) into six industries, including retail, travel, automotive, financial services (e.g., loans, mortgages), healthcare and pharmaceuticals, and CPG/consumer products, plus an "others" category. The distribution of ads across industries is presented in Table 2.1. The "others" category includes advertisers that are not covered by the six categories, such as media services, telecommunications, and gaming. Using the "others" category as our baseline, we obtain the heterogeneous effects of ad industries with Model 4.

(4) $$Y_{acwpt} = \alpha + \beta_1 \, Treated_c \times Post_t + \sum_{i=1}^{6} \theta_i Treated_c \times Post_t \times Industry_a^i + \alpha_a + \alpha_c + \alpha_w + \alpha_p + \alpha_t + \text{other control variables} + \varepsilon_{acwpt} \,,$$

where i ranges from 1 to 6, indexing ads for the six industries, respectively. The variable $Industry_a^i$ is a binary indicator that equals 1 if ad a is classified in category i. All other control variables described in Model 1 are included.[22]

The results are shown in Table 6.1. The coefficient of $Treated_c \times Post_t \times Industry_a^i$ captures the differential impacts on ads for category i. The results show that GDPR compliance hurts "travel" (e.g., flights, hotels) and "financial services" (e.g., loans, mortgages) ads the most. In contrast, GDPR compliance hurts ads for retail and CPG/consumer products the least. We also replicate these results using "retail" as the baseline (see Web Appendix VIII).

[22] Also, to study the three-way interactions $Treated_c \times Post_t \times Industry_a^i$, we include the two-way interactions $Post_t \times Industry_a^i$ for each industry. The $Treated_c \times Industry_a^i$ terms are subsumed into the interaction fixed effects of the country and ad and so need not be included in Model 4.

One may argue that these results may be due to the publisher's specific webpage topics (e.g., the publisher's lifestyle webpages may have compensated for the loss of personal data on retail and CPG/consumer products ads). To rule out this alternative explanation, we conduct a robustness check using data only from webpages on general topics, which do not provide clear context for advertising, and we find similar results (see Table 6.2 and see the additional metrics in Web Appendix IV).

We also find decreases in the number of daily active advertisers in different industries: retail 2.6%, travel 3.4%, automotive 3.1%, financial services 3.3%, healthcare and pharmaceuticals 2.9%, CPG and consumer products 2.6%, and others 3.0%. The decreases in the number of daily active advertisers for travel and financial services are slightly higher than for other industries (the difference is not statistically significant though). These results suggest that it is particularly important for advertisers in travel and financial services to track and manage the impact of the GDPR.

Travel and financial services often need to reach out to specific groups of customers (e.g., business travelers visiting New York in March, or, people who need to apply for a loan); in contrast, ads for retail or CPG products (e.g., detergent) have a wide audience. To further explore the idea that ads relying more on personal data may be affected more markedly by the GDPR, we conduct another analysis, in which we focus only on one industry, retail (see Web Appendix IX for details). Retail ads cover a broad range of stores and products and, thus, allow us to examine ads that may rely on user personal data to different degrees. An analyst of the collaborating company (blind to the hypotheses) manually classified retail ad creatives into two sub-categories: retail ads about specific products (e.g., "shop for sport-specific shoes at Walmart.com"), and general retail ads that do not focus on specific products (e.g., "convenient online shopping at

Walmart.com”). Compared to the latter, the advertised products in the former are more specific to personal interests, even though the ads are both about retailing. We compare these two sub-categories using data from webpages on general topics and find that the GDPR hurts specific retail ads more than general retail ads. We also replicate the results for the CPG industry (see Web Appendix IX for details). These findings suggest that GDPR compliance may have more marked effects on ads that need to reach more specific audiences.

The heterogeneity in webpages and advertisers in the data allows us to estimate the lower and upper bounds of the effects depending on the publisher’s webpage context and advertisers’ industries. According to our results, among the three different levels of webpage context (webpages with general topics, webpages with relevant specific topics, and webpages with irrelevant specific topics), the GDPR’s impact is the smallest on webpages with relevant specific topics and biggest on webpages with irrelevant specific topics. As presented in Table 6.1, among advertisers of the six industries, retail advertisers, who have a large general audience, are affected the least, and travel advertisers, who provide specific or personal products, are affected the most. Thus, the lower bound of the effect size should occur for retail advertisers on webpages with relevant specific topics, and the upper bound should occur for travel advertisers on webpages with irrelevant specific topics. As an example, with these two scenarios as the bounds, we estimate the impact on revenue per click to range from 2.8% to 8.7%; the calculation and the estimation for other metrics are listed in Web Appendix X.[23]

[INSERT TABLES 6.1 and 6.2 ABOUT HERE]

[23] This back-of-envelope calculation is a rough estimation, as it is affected by how the focal publisher classifies product industries and how well the webpage context of the focal publisher can support relevant ads. It only intends to show the significant variation in the GDPR’s impact across different ads and webpage contexts.

Conclusions and Implications

The trade-off between economic growth and privacy protection is receiving considerable attention from academics, practitioners, and policymakers, as it is one of the most pressing problems accompanying the data economy (Acquisti, Taylor, and Wagman 2016). The GDPR is a milestone in privacy protection. It offers a role model for countries that are moving toward adopting privacy legislation as it provides a benchmark for future privacy regulations (Satariano 2018). Quantifying the effect size and implications of the GDPR and finding alternative ways to alleviate its effects are particularly important for advertisers and publishers.

This study examines the impact of a large publisher's GDPR compliance using a wide array of ad metrics. Our analysis shows that for the focal publisher (that uses a pay-per-click model, has capacities to use both user personal information and webpage context for advertising, and has relatively high consent rates), there are moderate decreases in ad performance, bid prices and ad revenue. It is below the industry's drastic predictions (Deloitte 2013) but is consistent with contemporary GDPR literature (e.g., Aridor, Che, and Salz 2021; Goldberg, Johnson, and Shriver 2022; Schmitt, Miller, and Skiera 2022). Several factors may be related to the modest effect size. First, as Solove (2021) points out, given the number of post-GDPR consent requests that European users need to click on, users may resign from this near-impossible task by simply consenting to the requests; hence, Solove questions the effectiveness of opt-in consent in privacy protection. This partly explains the high consent rate of the focal publisher and hence the GDPR's relatively modest impact. Second, the publisher provides webpages with various content; 47% of its impressions are from webpages of specific topics, which can be leveraged by various advertisers. It may have contributed to the modest effect as well. Third, the publisher has a variety of advertisers. Certain advertisers (e.g., retail) may be less affected by the GDPR if their products have large general audiences and have lower needs to target niche consumer groups. In addition, the publisher uses a pay-per-click model; advertisers are guaranteed to obtain clicks with their ad spend, and hence they may be less likely to respond drastically to privacy regulation. Importantly, our paper is not the only one that finds a modest impact of the GDPR. Contemporary GDPR studies (e.g., Aridor, Che, and Salz 2021; Goldberg, Johnson, and Shriver 2022; Schmitt, Miller, and Skiera 2022) in general demonstrate a modest effect size on the GDPR's economic impact.

We show the mechanism of the change: (1) The ad metrics move together with the daily consent rates. Thus, the loss of personal data from non-consenting users induced by GDPR compliance may underlie the decreases in the ad metrics. (2) While ad performance immediately dropped following GDPR compliance (which is expected since the loss of user personal data occurred immediately after GDPR compliance), the ad bid price dropped after a short lag. Also, the percentage decrease in the bid price was similar to that of the conversion rate. These results suggest that the advertisers monitored their ad performance and adjusted their bid prices accordingly. (3) The reductions in both the advertisers' bid prices and the number of active advertisers may have contributed to the decrease in the publisher's revenue per click, which serves as an empirical validation of Chen and Stallaert (2014).

Our findings also suggest that the decreases in the ad metrics are the most pronounced for travel and financial services ads and the least pronounced for retail and CPG products ads. Thus, the travel and financial industries may need to pay greater attention to changes in privacy policy. We further showed that within an industry (e.g., retail), the GDPR has greater impacts on ads that need to target more specific audiences.

To answer the question about how companies can serve ads to relevant users when user data is limited, we find that the GDPR's negative impact can be partially compensated by webpages with specific topics, but this alleviation effect only exists if the webpage topic is relevant to the advertised products. Thus, webpage-content-based targeting, due to its nature of not relying on personal data, may be the trend for a privacy-protective future.

Finally, we showed that there was a significant decline in advertisers' surplus after GDPR compliance, which echoes Rafieian and Yoganarasimhan (2021).

We did not have separate data for consenting and non-consenting users, nor did we have detailed information on each ad auction or each individual user's purchase journey from impression to conversion. It is possible that consenting and non-consenting users are different (e.g., consenting users may be more receptive to ads). However, our result that the GDPR has greater impacts on ads that need to target more-specific audiences partly addresses this concern, and confirms that the loss of personal data plays a role (albeit smaller than expected by industry experts) and is one of the main reasons for why GDPR impacts publishers' revenue. Our data does not allow us to examine consenting and non-consenting users separately, or whether the consent pop-up window itself may have influenced users' ad-clicking behavior. Thus, we chose to focus on the overall effect of GDPR compliance, across consenting and non-consenting users. The overall metrics we studied are widely used in practice by various players and hence the results have important implications for digital marketing. Future research may examine these issues more closely when proper data or opportunities to do experiments are available. Our dataset covered ten weeks. This helped us isolate the effects and avoid spillover effects from other publishers but, at the same time, it only allowed us to assess the early impacts of GDPR compliance. Many publishers introduced more cookies gradually after GDPR compliance

(Lefrere et al. 2022) and these publishers' strategic changes may affect the long-term outcomes of the regulation. This is something future research could examine.

Implications

Our results are important to the advertising industry, as it suggests the industry's fears that the GDPR marks the end of ad-supported free internet may be invalid. Our results for webpage content strongly suggest that to be in an advantageous position in the post-GDPR world, publishers, especially those with niche content, should leverage webpage content for ad targeting as an alternative or an addition to personal-data-based targeting. It has been a historical practice for publishers to apply fine-grained tags to users in behavioral targeting. Instead of tagging users, in a post-GDPR world, publishers may need to tag their webpages with topics or keywords. For example, it is beneficial for publishers to tag webpages and provide at least some descriptions of the webpage content (e.g., the subdomains) in bid requests. Besides content information, publishers may also provide customized services to advertisers: for example, they can recommend to each advertiser a set of relevant webpages with high ad performance by analyzing the content of their webpages and the historical ad performance of each advertiser on each webpage.

Meanwhile, publishers may also need to be cautious about how much website content information they reveal. Too fine-grained tagging could hurt publishers' revenues as advertisers get more compartmentalized and thus lower the level of competition (Chen and Stallaert 2014; Levin and Milgrom 2010). With the reduction in behavioral data and increase in contextual data in the post-GDPR world,[24] the trade-off between providing fine-grained webpage tagging and maintaining advertiser competition level emerges as a new problem. Researchers and publishers may need to find an optimal point that balances the effectiveness of content-based targeting and

[24] https://www.forbes.com/sites/forbesagencycouncil/2021/07/22/the-new-rise-of-contextual-advertising/?sh=367b996c5e5d

the level of competition. The optimal solution would also differ depending on the relative share of opted-in site visitors. This tradeoff is especially salient if the share of opted-in visitors is small, i.e., behavioral targeting is less available for most users. In such a case, webpage content-based targeting is the dominating target method, and too much webpage content tagging may lower the competition level among advertisers.

For advertisers, it is critical that they actively investigate webpage-content-based targeting. According to our results, this is especially important for advertisers in the travel and financial services industries. Advertisers could seek publishers that provide fine-grained webpage content information, or adopt webpage analysis tools. In fact, there are a range of companies building machine learning tools (e.g., natural language processing and automatic text categorization), such as Trendii and Oracle, to help advertisers post ads in relevant contexts.

Finally, as Solove (2021) points out, when many consent requests are elicited, users may resign from this task by simply consenting; hence, we encourage future studies to keep monitoring the compliance and the effectiveness of the GDPR, and examine the long-term effects. Future research may also weigh the cost of privacy protection in ad performance and revenue against the benefits to long-term consumer welfare to evaluate the economics of privacy regulation.

References

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman (2016), "The Economics of Privacy," *Journal of Economic Literature*, 54 (2), 442–92.

Ada, Sıla, Nadia Abou Nabout, and Elea McDonnell Feit (2022), "Context Information can Increase Revenue in Online Display Advertising Auctions: Evidence from a Policy Change," *Journal of Marketing Research*, 59 (5), 1040-58.

Aridor, Guy, Yeon-Koo Che, and Tobias Salz (2021), "The Effect of Privacy Regulation on the Data Industry: Empirical Evidence from GDPR," *Proceedings of 22nd ACM Conference on Economics and Computation*, 93-94.

Arnosti, Nick, Marissa Beck, and Paul Milgrom (2016), "Adverse Selection and Auction Design for Internet Display Advertising," *American Economic Review*, 106 (10), 2852–66.

Asdemir, Kursad, Nanda Kumar, and Varghese S. Jacob (2012), "Pricing models for online advertising: CPM vs. CPC," *Information Systems Research*, 23(3): 804-822.

Beales, Howard and Jeffrey A. Eisenach (2014), "An Empirical Analysis of the Value of Information Sharing in the Market for Online Content," working paper.

Benes, Ross (2019), "Digital Ad Spending by Industry 2019," *eMarketer*, https://www.emarketer.com/content/digital-ad-spending-by-industry-2019.

Bleier, Alexander and Maik Eisenbeiss (2015), "Personalized Online Advertising Effectiveness: The Interplay of What, When, and Where," *Marketing Science*, 34 (5), 669–88.

Chapelle, Olivier (2015), "Offline Evaluation of Response Prediction in Online Advertising Auctions," *Proceedings of 24th International Conference on World Wide Web*, 919–22.

Chen, Jianqing and Jan Stallaert (2014), "An Economic Analysis of Online Advertising Using Behavioral Targeting," *Management Information Systems Quarterly*, 38 (2), 429-50.

Chen, Pei-Yu, Yili Hong, and Ying Liu (2018), "The Value of Multidimensional Rating Systems: Evidence from a Natural Experiment and Randomized Experiments," *Management Science*, 64 (10), 4629–47.

Danaher, Peter J. (2022), "Optimal Micro Targeting of Advertising," *Journal of Marketing Research*, forthcoming.

Davies, Jessica (2018), "Personalization Diminished: In the GDPR Era, Contextual Targeting is Making a Comeback," *Digiday* (June 7), https://digiday.com/media/personalization-diminished-gdpr-era-contextual-targeting-making-comeback/.

Deloitte (2013), "Economic Impact Assessment of the Proposed General Data Protection Regulat ion," Technical Report, *Deloitte*, https://www2.deloitte.com/content/dam/Deloitte/uk/Docu ments/about-deloitte/deloitte-uk-european-data-protection-tmt.pdf.

Dinner, Isaac M., Harald J. Heerde Van, and Scott A. Neslin (2014), "Driving Online and Offline Sales: The Cross-Channel Effects of Traditional, Online Display, and Paid Search Advertising," *Journal of Marketing Research*, 51 (5), 527–45.

eMarketer (2018), "Native Ad Spend Will Make Up Nearly 60% of Display Spending in 2018," *Insider Intelligence* (April 11), https://www.emarketer.com/content/native-ad-spend-will-make-up-nearly-60-of-display-spending-in-2018.

Edelman, Benjamin, Michael Ostrovsky, and Michael Schwarz (2007), "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords," *American Economic Review*, 97 (1), 242-59.

Farahat, Ayman and Michael C. Bailey (2012), "How Effective is Targeted Advertising?," *Proceedings of 21st international conference on World Wide Web*, 111–20.

Ghosh, Dipayan (2018), “How GDPR Will Transform Digital Marketing,” *Harvard Business Review*, 2–4.
Glenday, John (2019), “Native Advertising Sector Predicted to be Worth $400bn by 2025,” *The Drum* (March 6), https://www.thedrum.com/news/2019/03/06/native-advertising-sector-predicted-be-worth-400bn-2025.
Godinho de Matos, Miguel, and Idris Adjerid (2022), “Consumer Consent and Firm Targeting after GDPR: The Case of a Large Telecom Provider,” *Management Science*, 68(5): 3330-78.
Goldberg, Samuel, Garrett Johnson, and Scott Shriver (2022), “Regulating Privacy Online: An Economic Evaluation of the GDPR,” *American Economic Journal: Economic Policy*.
Goldfarb, Avi and Catherine Tucker (2011a), “Privacy Regulation and Online Advertising,” *Management Science,* 57 (1), 57–71.
Goldfarb, Avi and Catherine Tucker (2011b), “Online Display Advertising: Targeting and Obtrusiveness,” *Marketing Science,* 30 (3), 389–404.
Hu, Yu, Jiwoong Shin, and Zhulei Tang (2016), “Incentive problems in performance-based online advertising pricing: Cost per click vs. cost per action,” *Management Science*, 62(7): 2022-38.
Hummel, Patrick and R. Preston McAfee (2016), “When Does Improved Targeting Increase Revenue?” *ACM Transactions on Economics and Computation*, 5 (1), 4:1-4:29.
IAB (2019), “IAB Internet Advertising Revenue Report 2018 Full-Year Results,” Technical Report, IAB and PwC, https://www.iab.com/wp-content/uploads/2019/05/Full-Year-2018-IAB-Internet-Advertising-Revenue-Report.pdf.
IAB (2021), “Internet Advertising Revenue Report Full-Year 2020 Results,” Technical Report, IAB and PwC, https://www.iab.com/wp-content/uploads/2021/04/IAB_2020-Internet-Advertising-Revenue-Report-Webinar_4.7.21-PwC-1.pdf.
ICO (2019), “Update Report into AdTech and Real Time Bidding,” *Technical report, Information Commissioner’s Office*, https://ico.org.uk/media/about-the-ico/documents/2615156/adtech-real-time-bidding-report-201906.pdf.
IHS Technology (2015), “Paving the Way: How Online Advertising Enables the Digital Economy of the Future,” *Technical report*.
Jia, Jian, Ginger Zhe Jin, and Liad Wagman (2019), “GDPR and the Localness of Venture Investment,” working paper, *Social Science Research Network.*
Jia, Jian, Ginger Zhe Jin, and Liad Wagman (2021), “The Short-Run Effects of GDPR on Technology Venture Investment,” *Marketing Science*, 40 (4), 661-84.
Johnson, Garrett A. (2022), “Empirical Economic Research on Privacy Regulation: Lessons from the GDPR and Beyond”, working paper.
Johnson, Garrett A., Scott K. Shriver, and Shaoyin Du (2020), “Consumer Privacy Choice in Online Advertising: Who Opts Out and at What Cost to Industry?,” *Marketing Science,* 39 (1), 33–51.
Johnson, Garrett A., Scott K. Shriver, and Samuel Goldberg (2023), “Privacy and Market Concentration: Intended and Unintended Consequences of the GDPR,” *Management Science*.
Laub, René, Klaus M. Miller, and Bernd Skiera (2022), "The Economic Value of User Tracking for Publishers," SSRN (October 22). https://ssrn.com/abstract=4251233.
Lee, Kuang-chih, Burkay Orten, Ali Dasdan, and Wentong Li (2012), “Estimating Conversion Rate in Display Advertising from Past Performance Data,” *Proceedings of 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 768–76.

Lefrere, Vincent, Logan Warberg, Cristobal Cheyre, Veronica Marotta, and Alessandro Acquisti (2022), "The Impact of the GDPR on Content Providers," working paper.
Levin, Jonathan and Paul Milgrom (2010), "Online Advertising: Heterogeneity and Conflation in Market Design," *American Economic Review*, 100 (2), 603–7.
Libert, T., L. Graves, and R. K. Nielsen (2018), "Changes in Third-Party Content on European News Websites after GDPR," *Oxford University Research Archive*.
Ling, Xiaoliang, Weiwei Deng, Chen Gu, Hucheng Zhou, Cui Li, and Feng Sun (2017), "Model Ensemble for Click Prediction in Bing Search Ads," *Proceedings of the 26th International Conference on World Wide Web Companion*, 689-98.
Lu, Xianghua, Xia Zhao, and Ling Xue (2016), "Is Combining Contextual and Behavioral Targeting Strategies Effective in Online Advertising?," *ACM Transactions on Management Information Systems*, 7 (1), 1:1-1:20.
Lukic, Karlo, Klaus M. Miller, and Bernd Skiera (2021), "The Impact of the General Data Protection Regulation (GDPR) on Online Tracking," working paper.
Marotta, Veronica, Vibhanshu Abhishek, and Alessandro Acquisti (2019), "Online Tracking and Publishers' Revenues: An Empirical Analysis," working paper.
Moore, Robert S., Claire Allison Stammerjohan, and Robin A. Coulter (2005), "Banner Advertiser-Web Site Context Congruity and Color Effects on Attention and Attitudes," *Journal of Advertising*, 34 (2), 71–84.
Najafi-Asadolahi, Sami, and Kristin Fridgeirsdottir (2014), "Cost-per-click pricing for display advertising," *Manufacturing & Service Operations Management*, 16(4): 482-497.
Peukert, Christian, Stefan Bechtold, Michail Batikas, Tobias Kretschmer (2022), "Regulatory Spillovers and Data Governance: Evidence from the GDPR," *Marketing Science*, 41 (4), 318-40.
Rafieian, Omid and Hema Yoganarasimhan (2021), "Targeting and Privacy in Mobile Advertising," *Marketing Science*, 40 (2), 193–218.
Satariano, Adam (2018), "G.D.P.R., a new Privacy Law, Makes Europe World's Leading Tech Watchdog," *The New York Times* (May 24), https://www.nytimes.com/2018/05/24/technology/europe-gdpr-privacy.html.
Schmitt, Julia, Klaus Miller and Bernd Skiera (2022), "The Impact of Privacy Laws on Online U ser Behavior," SSRN (November 17), https://ssrn.com/abstract=3774110.
Solove, Daniel J. (2021), "The Myth of the Privacy Paradox," *George Washington Law Review*, 89 (1).
Tunuguntla, Srinivas and Paul R. Hoban (2021), "A Near-Optimal Bidding Strategy for Real-Time Display Advertising Auctions," *Journal of Marketing Research*, 58 (1), 1-21.
Ur, Blase, Pedro Giovanni Leon, Lorrie Faith Cranor, Richard Shay, and Yang Wang (2012), "Smart, Useful, Scary, Creepy: Perceptions of Online Behavioral Advertising," *Proceedings of 8th Symposium on Usable Privacy and Security*, 1-15.
Van Bentheim, Tom (2020), "Online Advertising 2.5 Years After the Implementation of the GDPR: What are the Lessons Learned?," *Ster* (December 8), https://www.ster.nl/online-advertising-2-5-years-after-the-implementation-of-the-gdpr-what-are-the-lessons-learned/.
Wiesel, Thorsten, Koen Pauwels, and Joep Arts (2011), "Marketing's Profit Impact: Quantifying Online and Off-line Funnel Progression," *Marketing Science*, 30 (4), 604–11.
Yan, Jun, Ning Liu, Gang Wang, Wen Zhang, Yun Jiang, and Zheng Chen (2009), "How Much Can Behavioral Targeting Help Online Advertising?," *Proceedings of 18th International Conference on World Wide Web*, 261–70.

Zhang, Kaifu and Zsolt Katona (2012), "Contextual Advertising," *Marketing Science*, 31 (6), 980–94.
Zhao, Yu, Pinar Yildirim, and Pradeep K. Chintagunta (2021), "Privacy Regulations and Online Search Friction: Evidence from GDPR," *Available at SSRN 3903599*.

Table 1. Impact of Privacy Regulation on the Online Economy.

Study	Data	Privacy Regulation	Metrics and Results (After vs. Before the Regulation)
Goldfarb & Tucker (2011a)	9596 ad campaigns	e-Privacy Directive	Self-reported user purchase intent (surveyed) decreases by 65% after e-Privacy Directive in the EU (vs. US).
Jia, Jin, & Wagman (2019, 2021)	Crunchbase (CB) and VentureXpert (VX) venture capital investments datasets	GDPR	Investment in new and emerging technology firms decreases in the EU (vs. US) after the GDPR.
Aridor, Che, & Salz (2021)	An intermediary in online travel industry	GDPR	12.5% reduction in intermediary-observed consumers. Advertising revenue decreased but not statistically significantly.
Lukic, Miller, & Skiera (2021)	Data from WhoTracks.me	GDPR	The number of online trackers decreases by 9% in the EU after the GDPR.
Zhao, Yildirim, & Chintagunta (2021)	Panel data	GDPR	21.6% more search terms to access information and 16.3% more pages browsed after the GDPR in the EU (vs. non-EU regions)
Godinho de Matos & Adjerid (2022)	Data from a large telecommunication provider	GDPR	Opt-in for different data types increased once GDPR consent was elicited
Goldberg, Johnson, & Shriver (2022)	A third-party intermediary (Adobe Analytics)	GDPR	12% reduction in both web traffic and e-commerce sales after the GDPR.
Lefrere et al. (2022)	Various content providers	GDPR	GDPR has reduced the number of third-party cookies and tracking responses, more evidently for EU IP addresses. The amount of new content posted does not change much.
Peukert et al. (2022)	Various public and proprietary data sources	GDPR	The number of third-party HTTP requests first decreases and then increases; the number of cookies decreases; and market concentration among technology vendors increases.
Schmitt, Miller, & Skiera (2022)	Data from SimilarWeb for the Top 1,000 websites	GDPR	The number of total visits to a website decreases by 4.9% short-term and 10% long-term after the GDPR.
Johnson, Shriver, & Goldberg (2023)	Panel data on connection between web technology vendors and top websites	GDPR	Market concentration among technology vendors increases by 17% after the GDPR in the EU.
The current study	Ad metrics directly from a single publisher	GDPR	Ad performance, bid prices, and ad revenue directly recorded by the publisher experience a modest decrease in the EU (vs. non-EU) after the GDPR.

Table 2.1. Descriptive Statistics.

Ad Traffic Proportion	
EU IP Address	23%
Non-EU IP Address	77%
Webpages on General Topics	53%
Webpages on Specific Topics	47%
Ad Slot Position 1	62%
Ad Slot Position 2	25%
Ad Slot Position 3	9%
Ad Slot Position 4 and below	4%
Retail Ads	19%
Automotive Ads	7%
Travel Ads	11%
Financial Services Ads	14%
Healthcare and Pharma Ads	10%
CPG and Consumer Product Ads	15%
Other Ads	24%

Summary of Ad Metrics

Variable	Mean	Std. dev.	Min	Max	Observations
Click-through rate	.0085	.0057	.0016	.098	3920212
Conversion rate	.04	.029	.0013	.22	3920212
Revenue per click ($)	.51	.38	.083	4.4	3920212
Bid prices (weighted by time, $)	.57	.42	.086	5.6	3920212
Revenue per impression ($)	.0038	.0036	.00022	.031	3920212
Conversion per impression	.00029	.00030	.0000072	.011	3920212
Bid weighted by impressions ($)	.62	.46	.14	6.1	3920212
Advertisers' surplus ($)	.12	.11	.045	2.2	3920212

Note: Per the request of the publisher, the statistics are multiplied by a constant number to protect confidential business information. Ad revenue (revenue per click and revenue per impression) and bid prices were initially recorded in their original currency (e.g., Euros) and were converted into US dollars using the average exchange rate of 2018; hence, they are in US dollars in the final dataset. The data is aggregated to the level of date, country, webpage, ad and ad slot position. Thus, the number 3920212 represents 3920212 combinations of date, country, webpage, ad and ad slot position.

Table 2.2. Descriptions of Ad Metrics.

Ad Metrics	
Click-through rate	Click-through rate = clicks/impressions, the number of clicks per impression. Primary ad performance metrics for publishers.
Conversion rate	Conversion rate = conversions/clicks, the proportion of clicks that yield conversions. Primary ad performance metrics for advertisers.
Revenue per click	Publishers' ad revenue generated by each click, which also equals advertisers' cost for each click.
Bid prices (weighted by	Weighted average of bid prices within the same 24-hour day, where

time)	the weight is the length of time each bid price holds.
Revenue per impression	Revenue per impression = revenue/impressions.
Conversion per impression	Conversion per impression = conversions/impressions, the proportion of impressions that yield conversions.
Bid weighted by impressions	The bid price weighted by impressions, calculated as a robustness check.
Advertisers' surplus	For each click an advertiser obtains, its surplus is approximated by the difference between the advertiser's bid price and cost.

Table 3. Overall Effect of GDPR Compliance, Represented by the Two-Way Interaction Treated × Post.

	Revenue per click	Click-through rate	Bid price	Conversion rate	Advertisers' surplus
Treated × Post	-.022***	-.00018***	-.025***	-.0017***	-.0036***
	(.00049)	(.0000081)	(.00051)	(.000042)	(.00014)
Ad FE	Yes	Yes	Yes	Yes	Yes
Ad slot position FE	Yes	Yes	Yes	Yes	Yes
Webpage FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes
Other control variables†	Yes	Yes	Yes	Yes	Yes
Observations	3920212	3920212	3920212	3920212	3920212
R^2	.82	.77	.84	.75	.71

Note: Three asterisks indicate 1% statistical significance (two-tailed).
† Other control variables include the set of two-way interaction fixed effects as described in Model 1, i.e., the fixed effects for each pair of ad and country, ad and ad position, ad and webpage, country and ad position, country and webpage, and ad position and webpage.

Table 4. Effect of Temporal Variations in Daily Consent Rates.

	Revenue per click	Click-through rate	Bid price	Conversion rate	Advertisers' surplus
Daily Consent Rate	.054***	.00061***	.056***	.0055***	.0066*
	(.014)	(.00022)	(.015)	(.0011)	(.0038)
Ad FE	Yes	Yes	Yes	Yes	Yes
Ad slot position FE	Yes	Yes	Yes	Yes	Yes
Webpage FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Weekday FE	Yes	Yes	Yes	Yes	Yes
Other control variables†	Yes	Yes	Yes	Yes	Yes
Observations	1093215	1093215	1093215	1093215	1093215
R^2	.77	.74	.80	.71	.66

Note: Three asterisks and one asterisk indicate 1% and 10% statistical significance (two-tailed) respectively.
† Other control variables include the set of two-way interaction fixed effects as described in Model 2, i.e., the fixed effects for each pair of ad and country, ad and ad position, ad and webpage, country and ad position, country and webpage, and ad position and webpage.

Table 5. Differential Effect for Ads on Webpages that Have Relevant Topics (Represented by Treated × Post × SpecificTopicMatch) or Irrelevant Topics (Represented by Treated × Post × SpecificTopicNotMatch), Compared to Ads on Webpages on General Topics.

	Revenue per click	Click-through rate	Bid price	Conversion rate	Advertisers' surplus
Treated × Post × SpecificTopicMatch	.011*** (.0010)	.00011*** (.000017)	.013*** (.0011)	.00092*** (.000090)	.0032*** (.00028)
Treated × Post × SpecificTopic NotMatch	-.00091 (.0011)	-.000021 (.000019)	-.00094 (.0012)	-.00010 (.000099)	-.00024 (.00031)
Treated × Post	-.026*** (.00085)	-.00022*** (.000014)	-.030*** (.00088)	-.0021*** (.000075)	-.0044*** (.00024)
Ad FE	Yes	Yes	Yes	Yes	Yes
Ad slot position FE	Yes	Yes	Yes	Yes	Yes
Webpage FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes
Other control variables†	Yes	Yes	Yes	Yes	Yes
Observations	3920212	3920212	3920212	3920212	3920212
R^2	.83	.78	.84	.75	.72

Note: Three asterisks indicate 1% statistical significance (two-tailed).
† Other control variables include the set of terms as described in Model 3, i.e., the fixed effects for each pair of ad and country, ad and ad position, ad and webpage, country and ad position, country and webpage, and ad position and webpage; also, to study the three-way interactions Treated × Post × SpecificTopicMatching and Treated × Post × SpecificTopicNotMatching, we include the two-way interactions SpecificTopicMatching × Post, SpecificTopicNotMatching × Post, SpecificTopicMatching × Treated, and SpecificTopicNotMatching × Treated. The SpecificTopicMatching and SpecificTopicNotMatching terms are subsumed into the interaction fixed effects of the ad and webpage and so need not be included.

Table 6.1. Differential Effect for Different Industries (with "Others" Category as the Baseline) Using the Full Dataset, Represented by Three-Way Interactions.

	Revenue per click	Click-through rate	Bid price	Conversion rate	Advertisers' surplus
Treated × Post × Retail	.0034** (.0013)	.000048** (.000022)	.0039*** (.0013)	.00038*** (.00011)	.00067* (.00037)
Treated × Post × Automotive	.00066 (.0019)	-.000014 (.000032)	.00075 (.0020)	-.00012 (.00016)	.00016 (.00053)
Treated × Post × Travel	-.0045*** (.0016)	-.000056** (.000025)	-.0051*** (.0016)	-.00047*** (.00014)	-.00084* (.00045)
Treated × Post × Financial Services	-.0029** (.0015)	-.000044* (.000024)	-.0031** (.0015)	-.00030** (.00012)	-.00048 (.00042)
Treated × Post × Healthcare and Pharma	-.00047 (.0016)	-.000010 (.000026)	-.00049 (.0016)	.000046 (.00015)	-.000082 (.00047)
Treated × Post × CPG and Consumer Product	.0024* (.0014)	.000043* (.000023)	.0024* (.0014)	.00028** (.00012)	.00039 (.00041)

Treated × Post	-.023*** (.00082)	-.00019*** (.000014)	-.025*** (.00084)	-.0018*** (.000070)	-.0038*** (.00024)
Ad FE	Yes	Yes	Yes	Yes	Yes
Ad slot position FE	Yes	Yes	Yes	Yes	Yes
Webpage FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes
Other control variables†	Yes	Yes	Yes	Yes	Yes
Observations	3920212	3920212	3920212	3920212	3920212
R^2	.83	.78	.85	.75	.72

Table 6.2. Differential Effect for Different Industries (with "Others" Category as the Baseline), Represented by Three-Way Interactions, Using Data from Webpages on General Topics.

	Revenue per click	Click-through rate	Bid price	Conversion rate	Advertisers' surplus
Treated × Post × Retail	.0035** (.0017)	.000053* (.000029)	.0042** (.0017)	.00039*** (.00015)	.00063 (.00049)
Treated × Post × Automotive	-.0022 (.0026)	-.000045 (.000044)	-.0025 (.0028)	-.00038* (.00022)	-.00030 (.00074)
Treated × Post × Travel	-.0081*** (.0022)	-.000091** (.000036)	-.0095*** (.0023)	-.00077*** (.00019)	-.0012* (.00065)
Treated × Post × Financial Services	-.0051** (.0020)	-.000065* (.000033)	-.0058*** (.0021)	-.00048*** (.00017)	-.0010* (.00059)
Treated × Post × Healthcare and Pharma	-.0015 (.0024)	-.000026 (.000039)	-.0017 (.0024)	-.000043 (.00020)	-.00024 (.00068)
Treated × Post × CPG and Consumer Product	.0025 (.0018)	.000051* (.000031)	.0026 (.0019)	.00031* (.00016)	.00016 (.00055)
Treated × Post	-.025*** (.0011)	-.00020*** (.000019)	-.027*** (.0012)	-.0019*** (.000098)	-.0043*** (.00033)
Ad FE	Yes	Yes	Yes	Yes	Yes
Ad slot position FE	Yes	Yes	Yes	Yes	Yes
Webpage FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes
Other control variables†	Yes	Yes	Yes	Yes	Yes
Observations	2058628	2058628	2058628	2058628	2058628
R^2	.84	.80	.85	.77	.74

Note: Three and two asterisks, and one asterisk, indicate 1%, 5%, and 10% statistical significance (two-tailed), respectively.

† Other control variables include the set of terms as described in Model 4, i.e., the fixed effects for each pair of ad and country, ad and ad position, ad and webpage, country and ad position, country and webpage, and ad position and webpage; also, to study the three-way interactions Treated × Post × Industry, we include the two-way interactions Post × Industry for each industry. The Treated × Industry terms are subsumed into the interaction fixed effects of the country and ad and so need not be included.

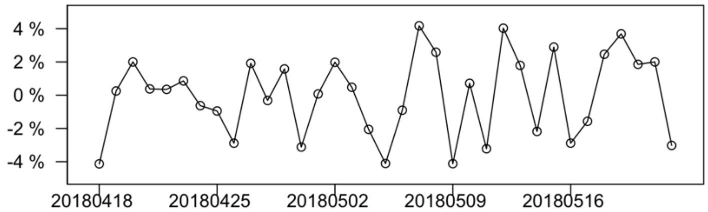
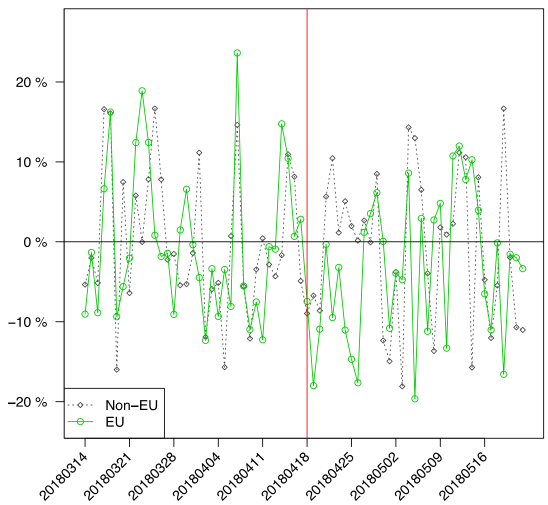
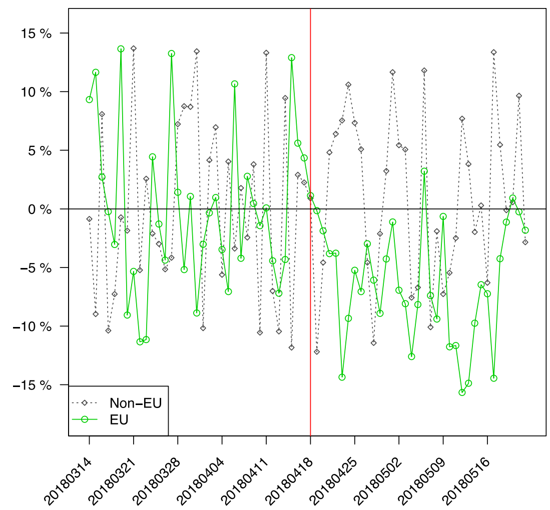


Figure 1. Daily consent rates relative to the average after GDPR compliance.
Note: The x-axis indicates the dates after GDPR compliance (35 days in total). The y-axis indicates the daily consent rates, relative to the average. We are unable to plot the absolute scale, per the request of the data provider to protect its confidential business information.

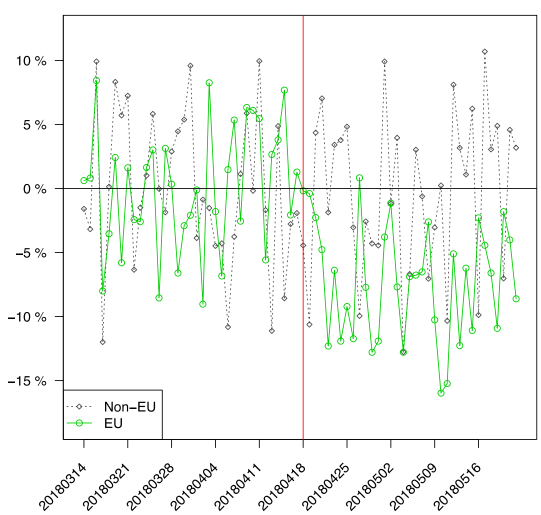
Click-through rate



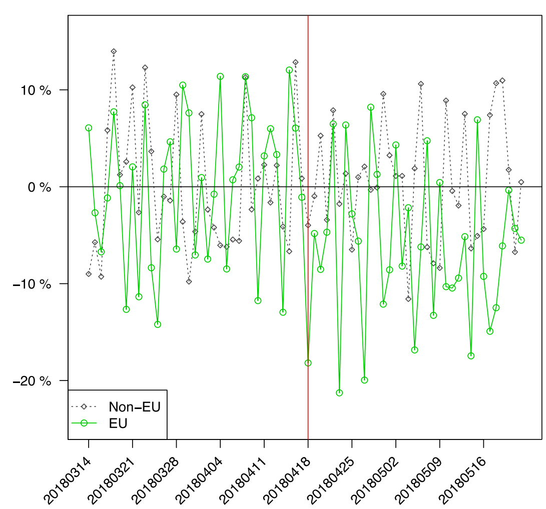
Revenue per click



Bid price



Conversion rate



Advertisers' surplus

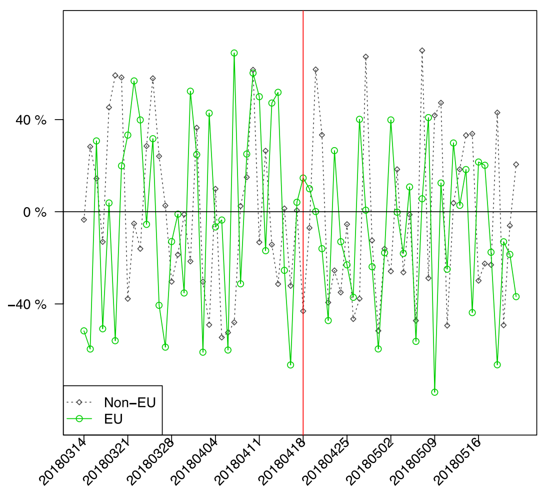
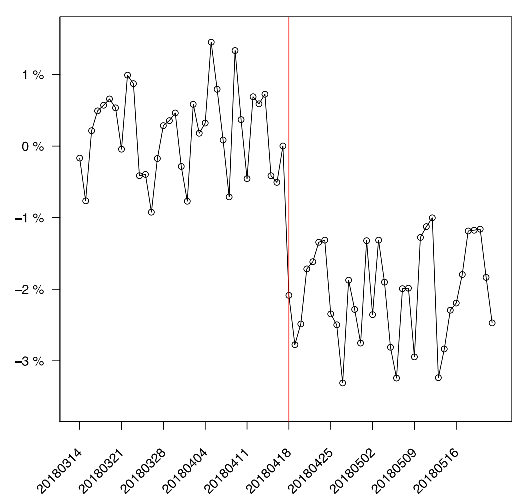
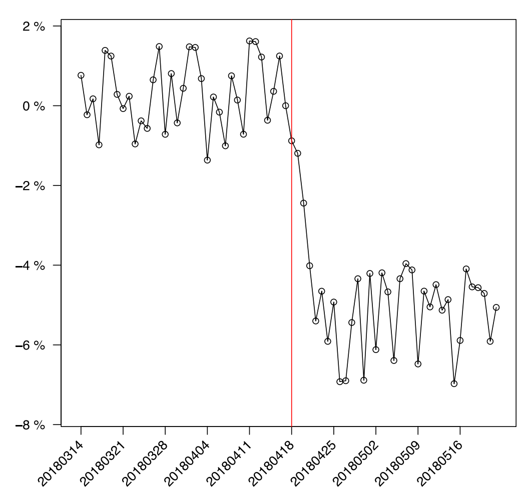


Figure 2. Model-free results. In each graph, the plot depicts the corresponding ad metric in EU (solid line with circles) and non-EU (dotted line with squares) regions, relative to the pre-treatment average of the corresponding metric. (Specifically, the numbers are first subtracted and then divided by the pre-treatment average of the corresponding metric in the corresponding region to obtain the relative scale. We are not able to plot the absolute scale per the request of the data provider to protect its confidential business information.) The vertical line in the middle indicates the date of GDPR compliance. Traffic from EU and non-EU regions is identified by IP addresses.

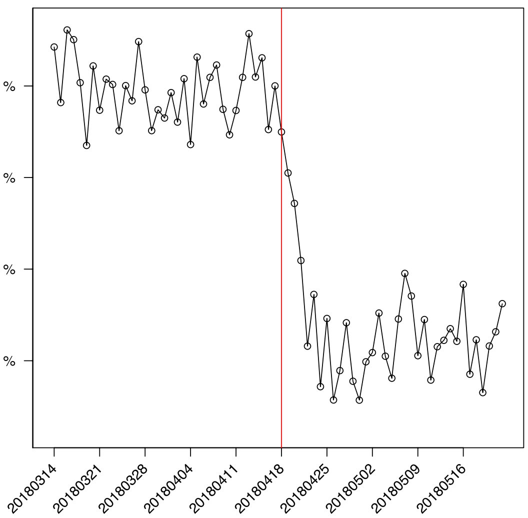
Click-through rate



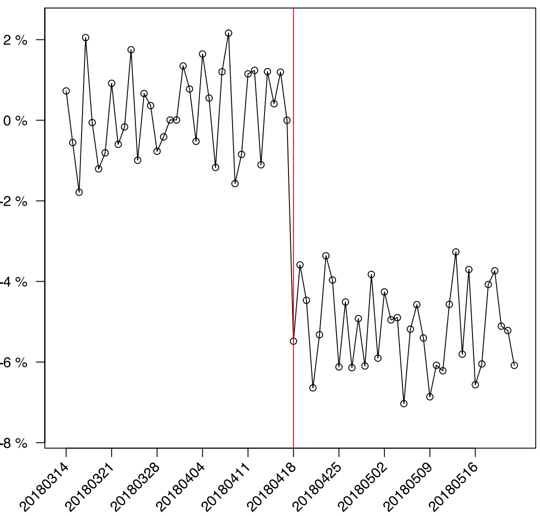
Revenue per click



Bid price



Conversion rate



Advertisers' surplus

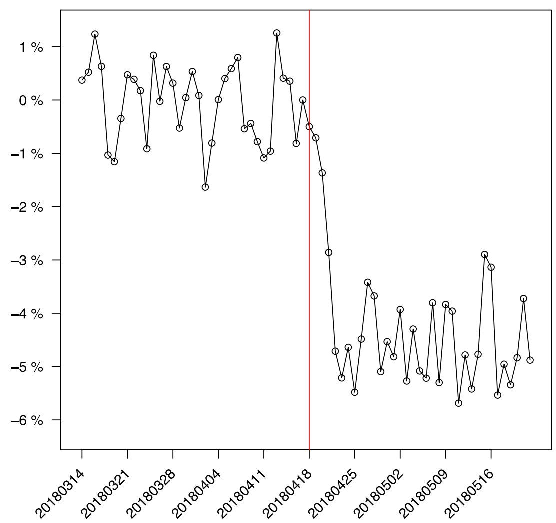


Figure 3. Daily effects over time. In each graph, the plot depicts the coefficient of the impact of GDPR compliance on each day, relative to the pre-treatment average of the corresponding metric. (That is, the coefficients are divided by the pre-treatment average of the corresponding metric to obtain the relative scale. We are not able to plot the absolute scale per the request of the data provider to protect its confidential business information.) The vertical line in the middle indicates the date of GDPR compliance.