

## **YOUR DATA RIGHTS: HOW DOES THE GDPR AFFECT THE SOCIAL MEDIA MARKET?**

Qifan Yang\*

### **Abstract**

With the development of digitalisation, personal data has gradually become a valuable resource for social media companies to extract value and obtain market dominance. Personal data processing can raise serious concerns about privacy leaks and misuse. In response, the adoption of the General Data Protection Regulation (GDPR) enhances personal data protection and market competition, but also potentially influences economic interests, the rights of data subjects, as well as market dynamics. The chapter uses the social media market to understand the complex relationship between the GDPR and market competition.

### **Table of Contents**

YOUR DATA RIGHTS: HOW DOES THE GDPR AFFECT THE SOCIAL MEDIA MARKET? .....	21
Abstract.....	21
Keywords .....	22
1. Introduction: Direct and Indirect Network Effects .....	22
2. Personal Data Regulation Practices Under the GDPR Framework.....	25

---

\* Qifan Yang is an Early-Stage Researcher in the LIDER Laboratory at Scuola Superiore Sant'Anna di Pisa and a Ph.D. student in the National Doctorate in Artificial Intelligence. This work is supported by the European Union's funded project Legality Attentive Data Scientists (LeADS) under Grant Agreement no. 956562. The author would like to thank Professor Giovanni Comandé from Scuola Superiore Sant'Anna di Pisa for his valuable contribution in reviewing this article. The author also wishes to thank the anonymous reviewers for their valuable suggestions.

3. How Personal Data Protection Affects the Market Share of Big Social Media Platforms.....	28
3.1 Concentration Ratio: an Indicator to Measure Market Competition .....	28
3.2 Market Share Dynamics in the Social Media Market Before and After the GDPR .....	30
3.3 Synthetic Control Methods to Estimate the Impact of the GDPR.....	31
4. Conclusion and Future Work.....	34
5. Selected Readings.....	35

## **Keywords**

GDPR – Personal Data – Market Concentration – Competition Regulation – Social Media Market

### **1. Introduction: Direct and Indirect Network Effects**

The technologies developed to allow the management of information and the exchange of communications have formed networks that connect people around the world and allow them to interact regardless of distance or time (Ibert et al., 2022). When using digital services - downloading software from an app shop, using Google to find the latest news, sharing a story on X or Facebook, watching a video on YouTube or TikTok, or buying an item on Amazon - one cannot help but notice the fact that digital services are dominated by a handful of well-known internet companies.

In real life, there tends to be greater flexibility of choice - one can buy coffee from a large chain such as Starbucks or homemade coffee from a local cafe operated by a neighbour. The freedom to choose between a global brand and a small local business is something we might take for granted offline. The diversity of providers allows consumers to never worry about the risk of no coffee owing to a boycott of Starbucks. But online, the providers of products or services seem to be more centralised, and

users are drawn to the services of these large companies without much thought. But why is that?

Take a platform like YouTube, for example. At first glance, it might just seem like a space where people share and watch videos. In fact, YouTube operates in a complex balance involving three key groups of users: content creators, content viewers, and advertisers. For content creators, YouTube provides a stage to distribute their work to a global audience, giving content creators fame, fun, and revenue. Millions of content viewers use YouTube to watch a wide range of videos and become potential targets for advertisers to promote their products. Unlike one-sided markets that provide products and services to consumers like retailers (e.g. Walmart, Carrefour) and legacy media (e.g. newspapers, TV), Internet companies function more as an interactive platform for communication, sharing or trading among different parties (Jullien et al., 2021; Saura et al., 2021).

The more individuals join a platform, the more opportunities to interact, connect and share content, thus improving the user experience and creating a more valuable network. The value of a platform increases with the number of individuals joining, which is known as the “direct network effect”<sup>1</sup>. In essence, it is a virtuous circle, where users attract more users, which in turn enhances the service provided to users, like a snowball rolling down a hill.

As the number of users has increased, so has the change on the other side of the platform. Sellers, recognising potential consumers, are naturally attracted by platforms with a large pool of potential buyers, like YouTube, Instagram, and Amazon (Calvano & Polo, 2021). In addition, this large user base generates huge amounts of data that are of great interest for advertisers, who are eagerly seeking to match a product or service with a target audience (Sembada & Koay, 2021). The growing user base is more like a magnet pulling in outside businesses, advertisers, and service providers. This is the “indirect network effect”<sup>2</sup> - where the value of a platform is increased by businesses, advertisers, and other external entities being attracted by the large user base (Veisdal, 2020).

---

<sup>1</sup> The direct network effect means that the value or utility that users derive depends on the number of other users on the same side.

<sup>2</sup> The indirect network effect means that the utility of at least one group grows as the other groups grow.

To sum it up, the direct network effect revolves around the interplay among users, whereas the indirect network effect involves the interplay between users and other stakeholders like sellers and advertisers, the number of which will increase as the user base expands.

When companies offer better products or services, users are willing to invest more time, data, and connections on social media platforms. As users spend more time building connections, sharing content and leaving digital footprints, the cost of moving to another platform increases - not just financially but psychologically as well (Buiten, 2021). When these economic and psychological costs of switching from one alternative to another become high, social media platforms can trap individuals in their own networks, which is known as the “lock-in effect”<sup>3</sup>. Consider that a user spends years building a social network on Facebook or Instagram. Upon leaving that platform, that user not only loses friends and followers, but also data and content. Rebuilding a new social network on another platform can be time-consuming and effortful, which creates a digital trap.

With the network effect and the lock-in effect, the expanding control of user personal data from online platform companies creates an insurmountable barrier to entry for market competition in this area (Newman, 2014), since personal data have gradually become a competitive asset in the online platform market. It has raised critical concerns about the protection of personal data and the potential abuse of market dominance. In 2020, Germany’s Federal Court of Justice highlighted that Facebook’s massive user data collection exacerbated already distinct “lock-in effects” and their large user database enhances the possibilities to finance the social network by using the profits generated from advertising contracts.

The General Data Protection Regulation (GDPR) represents a significant milestone in the ongoing effort to protect personal data. The GDPR aims to empower individuals by granting them greater control over their personal data while imposing stricter rules on how companies can process personal data. With the GDPR in place, businesses based on online platforms and personal data must rethink and restructure their strategies. The GDPR was expected to help reduce market concentration, but

---

<sup>3</sup> The lock-in effect means that customers are dependent on one product from a provider and cannot use the product from another provider unless they pay significant switching costs.

what kind of impact did it actually have in terms of the company's market share<sup>4</sup>? This chapter introduces the framework and regulatory provisions of the GDPR and then selects the social media market as a use case. A statistical method for assessing the effects of interventions in comparative case studies (synthetic control method<sup>5</sup>) is employed to identify and quantify the causal effect of the GDPR reform on social media market share in the EU. The conclusion explores the impact of the GDPR along with a brief discussion of its reasons.

## **2. Personal Data Regulation Practices Under the GDPR Framework**

Since the GDPR was adopted, each EU member state has taken significant steps to align its national laws with this robust framework. This alignment includes enhancing the capabilities of domestic data protection authorities (DPAs), which play a crucial role in investigating potential violations and implementing regulatory actions to protect personal data.

Since 2018 until June 2024, EU member states have dealt with 2,141 cases related to personal data protection violations. These cases have resulted in fines that collectively exceed €4,590 million - a clear indication that non-compliance can be quite costly. The main violations that most frequently trigger GDPR regulation and fines include non-compliance with general data processing principles<sup>6</sup>, insufficient legal basis for data processing<sup>7</sup>, insufficient technical and organisational measures to ensure information security<sup>8</sup>, and others (CMS Legal Services EEIG, 2024).

Looking at the timeline from 2018 to June 2024, the number of reported cases has noticeably increased each year: just 9 cases in 2018, 143 cases in 2019, 342 cases in

---

<sup>4</sup> Generally, market share is defined as the percentage of a company's business out of the total revenue or sales in the market.

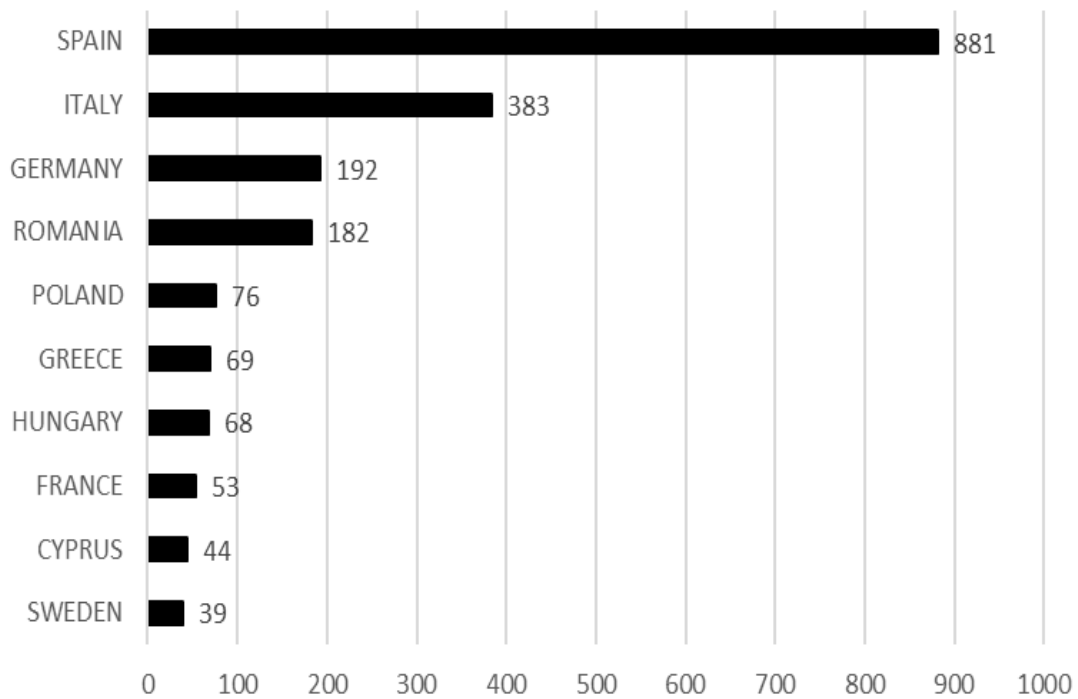
<sup>5</sup> The synthetic control method is a statistical method for assessing the effects of interventions in comparative case studies and it will be briefly described in 3.3.

<sup>6</sup> Non-compliance with general data processing principles includes failing to adhere to basic principles such as data minimisation, accuracy, and purpose limitation.

<sup>7</sup> Insufficient legal basis for data processing means that the company or the organisation lacks a sound reason for the data processing, such as consent or legitimate interest.

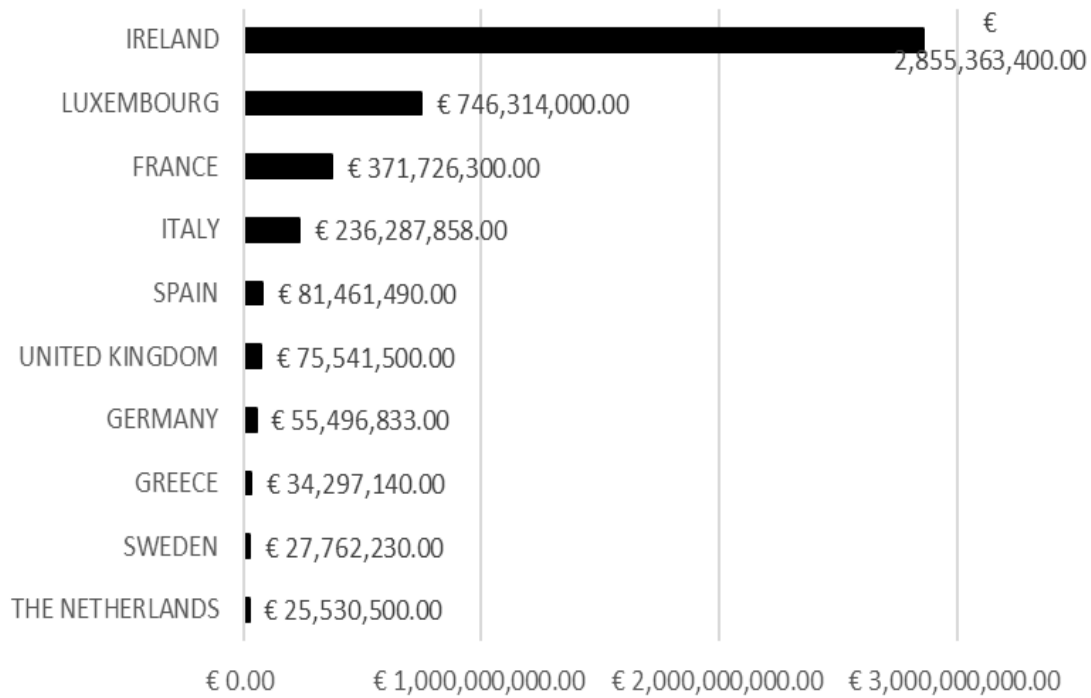
<sup>8</sup> Insufficient technical and organisational measures to ensure information security means the company or organisation does not provide appropriate and adequate technical and organisational measures to protect information security.

2020, 462 cases in 2021, 536 cases in 2022, 510 cases in 2023 and 154 cases by June 2024. This upward trend shows that more people are becoming aware of their rights under the GDPR and DPAs are getting better at enforcing them (CMS Legal Services EEIG, 2024).



**Figure 1.** Top 10 DPAs by total number of fines

(Data Source: GDPR Enforcement Tracker as of June 2024)



**Figure 2.** Top 10 DPAs by total sum of fines

(Data Source: GDPR Enforcement Tracker as of June 2024)

In terms of the number of cases and the total amount of fines, certain countries have been particularly proactive in the GDPR enforcement. As of June 2024, DPAs from Spain, Italy, Germany, and Romania have been active in dealing with data protection cases, see **Figure 1**. In terms of the total amount of non-compliance, Ireland, Luxembourg, France, and Italy stand out for imposing substantial fines, see **Figure 2**.

This further illustrates that data protection can also result in significant financial losses for those who fail to comply. Companies that use a lot of data, like tech companies, had to spend more on improving data security, training their workers, and recording their processing. Koski and Valmari (2020) used company-level data on European

and US companies from 2014-2018 to reveal that the GDPR imposes financial burdens on European companies and causes a decline of profit margins for data-intensive companies.

By imposing stringent rules on the digital world, the GDPR affects how companies compete and how they share the benefits (Jones & Tonetti, 2020; Li & Feng, 2021), just like how basic rule changes affect the play of the players in a game. Early analyses argue that the GDPR can enhance competition by lowering compliance costs by using clear rules, increasing consumer trust, and fostering the uptake of new technologies. If privacy regulation is coupled with appropriate incentives, it may positively influence the development and adoption of information exchanges (Godinho de Matos & Adjerid, 2022).

### **3. How Personal Data Protection Affects the Market Share of Big Social Media Platforms**

The present research intends to explore the actual impact of the GDPR on the EU social media market. On the social media market, platforms facilitate widespread engagement and data exchange among users (European Commission, 2021). By studying social media, we can learn how personal data protection affects market share dynamics.

#### **3.1 Concentration Ratio: an Indicator to Measure Market Competition**

The Concentration Ratio ( $CR_n$ ) is commonly employed to measure market concentration<sup>9</sup> and assess the changes of companies' market share. To put it simply, all the sales (revenue or traffic) in a market in a region (country, region, city) can be seen as a pie, and each company in that market takes its own slice of the pie based on the proportion of its sales (revenue or traffic) to the total one in the market (See

---

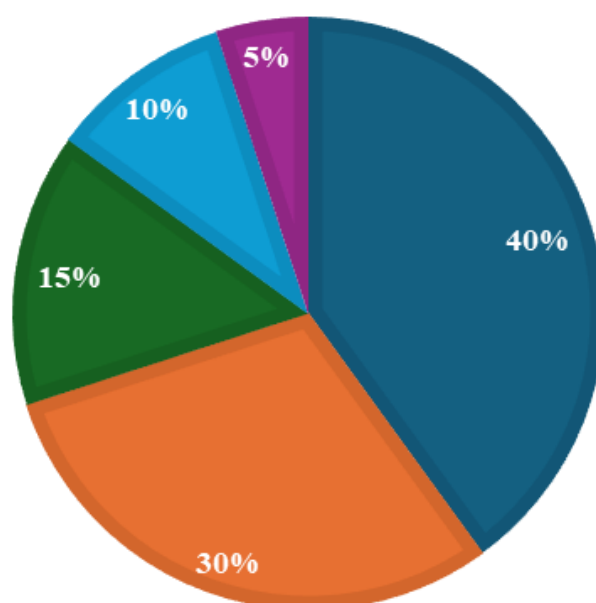
<sup>9</sup> Market concentration is the market share of a certain number of companies in a given market.



**Figure 3).**  $CR_n$  is calculated by summing the market shares of a specified number of the largest companies in a particular industry, which shows the total market share held by the  $n$  largest companies<sup>10</sup> in the market.

### MARKET SHARE IN THE MARKET

■ Largest company    ■ 2nd largest company    ■ 3rd largest company  
 ■ 4th largest company    ■ 5th largest company



**Figure 3.** Example of market share and market concentration

If there are 5 companies in the market, and the company with the largest market share occupies 40 % of the market share, the second largest company occupies 30 %, the third largest company occupies 15 %, the fourth largest company occupies 10 %, and the fifth largest company occupies 5 %. Then the  $CR_1$  is 40 %, the  $CR_2$  is the sum of the first company's market share and the second company's market share (70 %), and

<sup>10</sup>  $n$  in  $CR_n$  represents the number of companies included in the concentration ratio calculation.

the  $CR_3$  is the sum of the first company's market share, the second company's market share, and the third company's market share (85 %), and so on.  $CR_4$  is commonly used to measure the market concentration in an industry.

$CR_n$  is like a window into the competitive landscape of a market. Higher concentration indicates that few top companies in the industry dominate the market, i.e. the market is more monopolised and competition in the market is reduced.

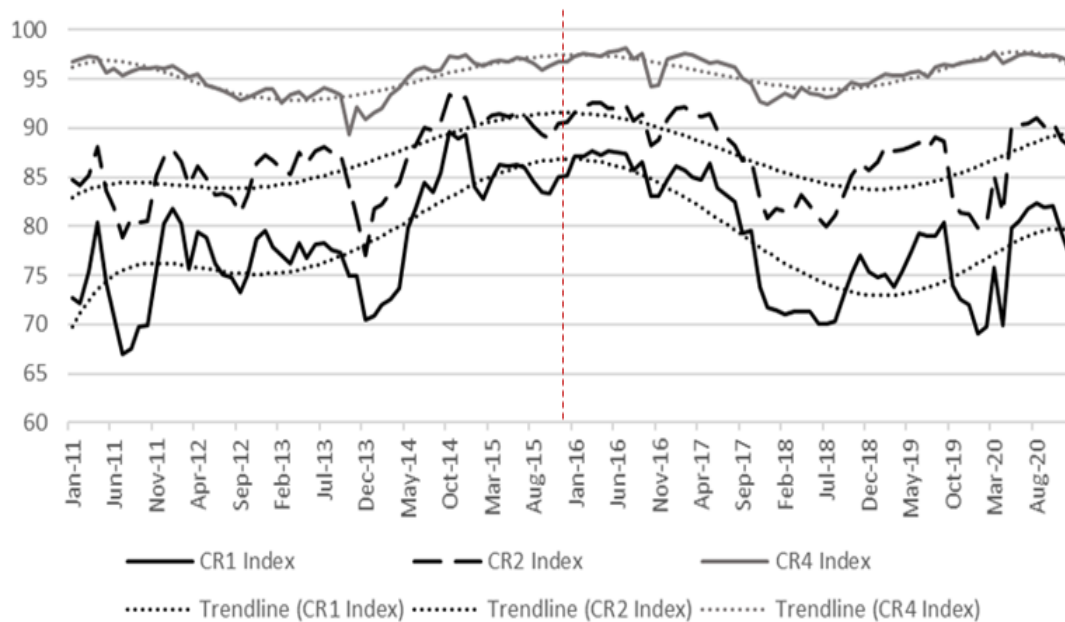
### 3.2 Market Share Dynamics in the Social Media Market Before and After the GDPR

Between 2009 and 2015, market concentration in the European social media market increased significantly before the adoption of the GDPR, see **Figure 4**:

- The  $CR_1$  index, which shows the market share of the largest company, grew from about 30% to around 85%;
- The  $CR_2$  index, representing the combined market share of the two largest companies, increased from around 55% to around 90%;
- The  $CR_4$  index, accounting for the four largest companies, went from roughly 75% to around 95%.

Before the GDPR, the growing dominance of a few major companies in the social media market could be clearly observed through the rise in the  $CR_1$  index, the  $CR_2$  index, and the  $CR_4$  index. While having a dominant market position does not automatically break antitrust laws, it is clear that a few key players are becoming more established in this market.

After the adoption of the GDPR (the right side of the red dotted line), the  $CR_1$  index, the  $CR_2$  index, and the  $CR_4$  index gradually shift from a rising trend to a declining trend until 2022. The  $CR_1$  index and the  $CR_2$  index show significant decreases, and the decrease in the  $CR_4$  index is smaller compared to that of the  $CR_1$  index and the  $CR_2$  index. Although other factors may also influence these changes, the glimpse shows that the GDPR can have a negative impact on market concentration, see **Figure 4**.



**Figure 4.** Concentration ratio of the social media market before and after the GDPR  
(Data source: StatCounter Global Stats)

### 3.3 Synthetic Control Methods to Estimate the Impact of the GDPR

When assessing the impact of the GDPR, the best way is to compare two groups: an EU group where the GDPR is in force (EU group) and a different group where the GDPR is not in force but has identical characteristics to the first group otherwise (twin EU). By comparing the two groups, it is possible to accurately measure the difference between the two groups in social media market concentration after the adoption of GDPR. Unfortunately, no such twin EU exists. This is where econometric creativity comes into play, specifically creating a synthetic twin EU through a tool known as the synthetic control method (Abadie et al., 2010). Although the synthetic control method<sup>11</sup> involves many technical parts, its main logic is very easy to understand.

<sup>11</sup> Synthetic control method (SCM) a statistical method to evaluate treatment effects in comparative case studies, which allows the construction of a counterfactual by selecting a weighted average of the outcome variable from a group of units similar to the treated unit.

The logic of the synthetic control method can be illustrated with a simple example. Now there is a glass of mixed juice consisting of 30% apple juice, 30% orange juice, and 40% grape juice, and a test is required to examine the effect of adding mango juice to this glass of juice. It is impossible to obtain another juice with the same composition, but a similar (synthetic) juice can be created by mixing apple, orange and grape juices in the same proportions, and the effect of adding mango juice can be observed by using the synthetic juice as the twin group.

In our case,  $CR_4$  is collected as the outcome variable measuring changes in market concentration,<sup>12</sup> so the  $CR_4$  of the EU group is the “mixed juice”. The different types of juices are the countries or regions that have not adopted the GDPR but are very similar to the EU in many ways, like GDP, population, technology, size of internet users, government efficiency, and level of regulation. We used 24 countries or regions<sup>13</sup> with some similar characteristics to the EU as a donor pool, like a pool of different juices.

Based on this pool, the statistical methods are used to select countries or regions that contribute to the  $CR_4$  of the EU group (type of juice contained in the “mixed juice”) and to find a combination of the selected countries (proportion of different juices in the “mixed juice”) that matched the various characteristics and market trends of the EU before the adoption of the GDPR. Combining these selected countries<sup>14</sup> with the given weights will create a synthetic EU group (twin EU). After building the synthetic group, the real EU market with the GDPR is compared with the synthetic EU market without the GDPR to find the differences, and the gap between the real market and the synthetic market is the effect of the GDPR.

**Figure 5** illustrates that the trend of the  $CR_4$  index in the synthetic EU and the actual EU matches closely before the GDPR was introduced. This indicates that the

---

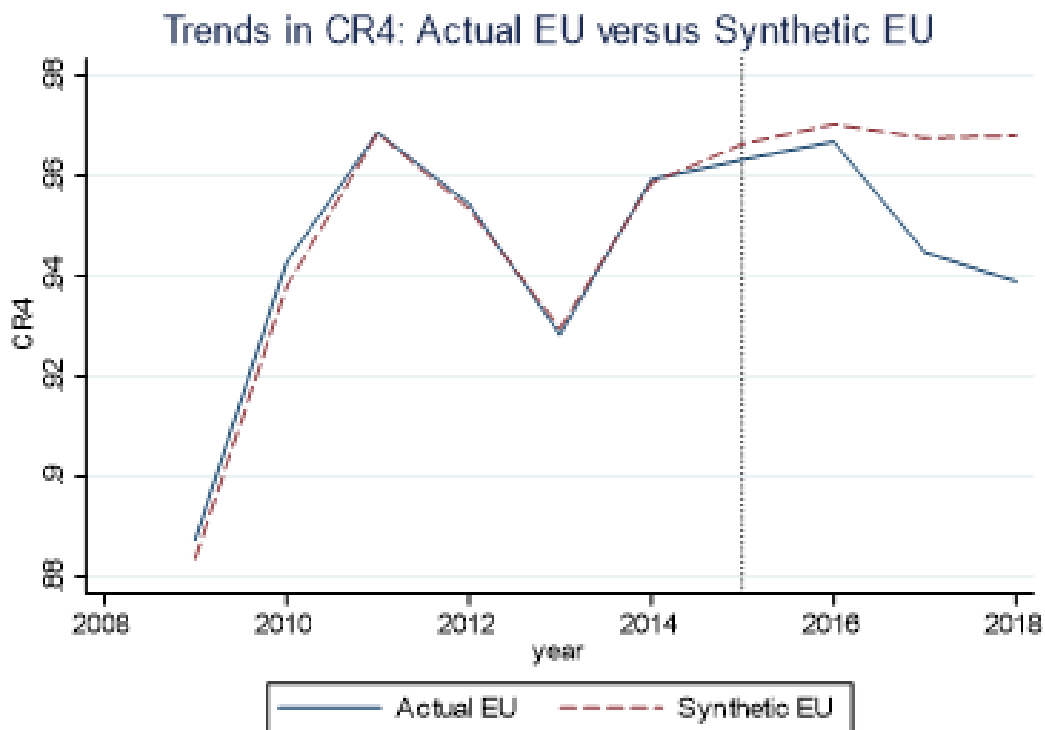
<sup>12</sup> To ensure the similarities between the treated unit and the units in the donor pool, the parallel trends of  $CR_1$ ,  $CR_2$ ,  $CR_3$ , and  $CR_4$  in the donor pool and the treated unit are tested.  $CR_4$  is closer to the treated units than the other indexes and is less affected by a single event targeting a particular player, so it is more appropriate as an outcome variable.

<sup>13</sup> The 30 countries or regions are the EU, United Arab Emirates, Argentina, Australia, Brazil, Canada, Switzerland, Chile, China, Egypt, Arab Rep., United Kingdom, Hong Kong SAR(China), Indonesia, India, Israel, Japan, Korea, Rep., Mexico, Malaysia, New Zealand, Philippines, Russian Federation, Saudi Arabia, Singapore, Thailand, Turkey, United States, Viet Nam, South Africa.

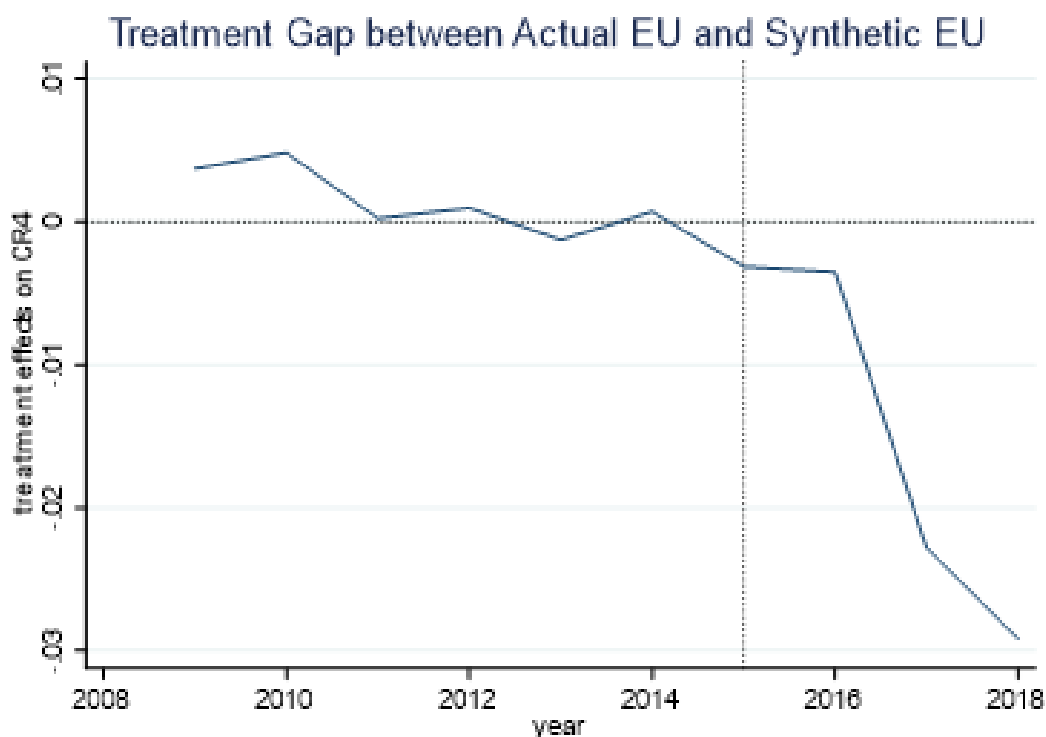
<sup>14</sup> In our case, the  $CR_4$  index in the EU social media market is best reproduced by the combination of Korea, Rep.(0.514), Chile(0.338), Hong Kong SAR, China (0.094) and Russian Federation (0.055).

synthetic EU without GDPR regulations is a good copy of the real EU's market concentration (the  $CR_4$  index). The disparity between the real EU  $CR_4$  index and its synthetic unit emerged before the time of the GDPR adoption, which may be likely influenced by the European Parliament's vote for the GDPR in 2014 and the agreement on the GDPR by the European Parliament, the Council and the European Commission in 2015.

**Figure 5** also reveals the  $CR_4$  index in the synthetic EU had a very small decline, indicating a natural trend in market concentration over time. However, in the real EU, the  $CR_4$  index dropped sharply after GDPR adoption, indicating that the GDPR had a significant negative impact on market concentration in the EU social media market.



**Figure 5.** Trends of  $CR_4$  between actual EU and synthetic EU



**Figure 6.** Treatment gap of CR<sub>4</sub> between actual EU and synthetic EU<sup>15</sup>

#### 4. Conclusion and Future Work

Based on the above analyses and empirical evidence, we can easily find that the enhancement of personal data protection not only benefits the data subjects, but also plays a role in slowing down the market concentration, at least in the social media market. The negative effect of GDPR on social media market concentration may stem from the following reasons:

- The GDPR's transparency and accountability requirements limit social media platforms' power to misuse user data, empowering users with

<sup>15</sup> To test the significance of the result, I test each country and region with similarities in the sample by applying the same synthetic control method. For the distribution of the post/pre-GDPR ratios, the difference between before and after GDPR in the EU unit is about 48.165 times, a much larger difference than any of the 23 control regions. The probability of this difference happening by random chance is very low, 1/24 (about 0.0416), making the conclusion reliable.

greater control and compelling platforms to weigh the costs of extensive data collection.

- Data portability provisions enable users to transfer their data between platforms, reducing data exclusivity and promoting market competition.
- Strict data protection regulations impose compliance costs on dominant platforms, leveling the competitive landscape by restricting their ability to exploit data collection advantages.
- By regulating dominant platforms' data processing, new entrants can compete more effectively without facing exclusive data constraints. Apparently, the implications of the regulation of personal data spill over into the market sphere.

This chapter attempts to provide a new perspective on the impact of the GDPR on social media market concentration in the EU, but also has some limitations. Future research could delve into the divergence of personal data regulations across different jurisdictions. Understanding these differences and their implications can shed light on the feasibility and challenges of harmonising global data governance standards. Additionally, the synthetic control evaluation has scope for further refinement in this study.

## 5. Selected Readings

- (1) Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105(490), 493–505.
- (2) Buiten, M. C. (2021). Exploitative abuses in digital markets: Between competition law and data protection law. *Journal of Antitrust Enforcement*, 9(2), 270–288.
- (3) Calvano, E., & Polo, M. (2021). Market power, competition and innovation in digital markets: A survey. *Information Economics and Policy*, 54, 100853.

- (4) CMS Legal Services EEIG. (2024, August 7). *List and Overview of Fines and Penalties under the EU General Data Protection Regulation (GDPR, DSGVO)*.
- (5) European Commission. (2021). *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS 2030 Digital Compass: The European way for the Digital Decade*.
- (6) Godinho de Matos, M., & Adjerid, I. (2022). Consumer Consent and Firm Targeting After GDPR: The Case of a Large Telecom Provider. *Management Science*, 68(5), 3330–3378.
- (7) Ibert, O., Oechslen, A., Repenning, A., & Schmidt, S. (2022). Platform ecology: A user-centric and relational conceptualization of online platforms. *Global Networks*, 22(3), 564–579.
- (8) Jones, C. I., & Tonetti, C. (2020). Nonrivalry and the Economics of Data. *American Economic Review*, 110(9), 2819–2858.
- (9) Jullien, B., Pavan, A., & Rysman, M. (2021). Two-sided markets, pricing, and network effects☆. In K. Ho, A. Hortaçsu, & A. Lizzeri (Eds.), *Handbook of Industrial Organization* (Vol. 4, pp. 485–592). Elsevier.
- (10) Koski, H., & Valmari, N. (2020). *Short-term Impacts of the GDPR on Firm Performance* (Working Paper No. 77). ETLA Working Papers.
- (11) Li, S., & Feng, J. (2021). Is Data Ownership Empowerment Welfare-Enhancing? *Hawaii International Conference on System Sciences 2021 (HICSS-54)*.
- (12) Newman, N. (2014). Search, Antitrust, and the Economics of the Control of User Data. *Yale Journal on Regulation*, 31(2), 401–454.
- (13) Saura, J. R., Ribeiro-Soriano, D., & Palacios-Marqués, D. (2021). From user-generated data to data-driven innovation: A research agenda to understand user privacy in digital markets. *International Journal of Information Management*, 60, 102331.
- (14) Sembada, A. Y., & Koay, K. Y. (2021). How perceived behavioral control affects trust to purchase in social media stores. *Journal of Business Research*, 130, 574–582.



- (15) Veisdal, J. (2020). The dynamics of entry for digital platforms in two-sided markets: A multi-case study. *Electronic Markets*, 30(3), 539–556.

