

SHIFTING BATTLEFIELDS: CORPORATE POLITICAL ACTIVITY IN THE GENERAL DATA PROTECTION REGULATION

VÁCLAV OCELÍK

University of Amsterdam
Plantage Muidergracht 12, 1018 TV Amsterdam, the Netherlands

ANS KOLK

University of Amsterdam, the Netherlands

KRISTINA IRION

University of Amsterdam, the Netherlands

INTRODUCTION

Companies are increasingly lobbying the European Union (EU) in an effort to shape forthcoming policy (Coen & Vannoni, 2020), especially in the digital technology sector (Bank et al., 2021). This increased lobbying activity is a direct response to intensive regulatory activity by the European Commission (EC), which has declared its intention to “[...] manage the transformation into the digital age” (Von der Leyen, 2019: 14). Following the passage of the General Data Protection Regulation (GDPR) in 2016, legislative initiatives in the pipeline cover a whole range of areas including artificial intelligence, cyber security and high-performance computing, as well as obligations for online intermediary services through the Digital Services Act, and for so-called ‘gatekeeper’ companies via the Digital Markets Act.

What happens in the EU profoundly affects business operations both inside and outside the European single market, given the extraterritorial reach of many regulations, sometimes labeled the ‘Brussels effect’ (Bradford, 2020). The GDPR and its predecessor, for example, have spurred a wave of privacy rules in other countries and regions around the world. These and other EU regulatory activities (e.g., in the realm of the European Green Deal) have led firms to engage in corporate political activity (CPA) to actively manage their nonmarket environment (Baron, 1995), taking actions to influence governmental policies and/or decision-making processes (Getz, 1997; Hillman & Hitt, 1999).

While CPA has been studied extensively (Lenway et al., 2022; Schuler et al., 2019), we posit that the extant body of literature suffers from two significant limitations. First, CPA research pays limited attention to the *substance* of information strategies targeted at policymakers. Extant work mainly focuses on the consequences of information strategies, such as whether filing anti-dumping petitions and providing congressional testimony increase the financial performance of the firm (e.g., Ridge et al., 2019; Schuler, 1996). Second, empirical studies rarely consider the lobbying strategies of other stakeholder groups with opposing policy preferences. While non-governmental organization (NGOs), public authorities, and trade associations lobby government officials, their interests and ideal policies and those of firms are imperfectly aligned.

To help fill these gaps, this article addresses the following research question: what are the differences in informational substance of lobby documents supplied by businesses, trade associations, NGOs and public authorities? As empirical setting and case, we selected the EU and its GDPR, which constituted the first in a series of ambitious new regulations in the digital realm. Dissecting the lobbying process of the GDPR yields crucial insights into information strategies concerning other pieces of (forthcoming) legislation. Our data consists of 266 documents from 204

distinct organizations submitted to the EC in response to two GDPR consultation rounds, which we analyze using structural topic modeling (Roberts et al., 2014).

Our study makes several contributions to the CPA literature. First, building on the literature on interest group politics (Crombez, 2002; Eising, 2007), we elucidate how institutional procedures affect information strategies. Our analysis reveals that the two stages in EC consultations are characterized by distinct informational substance from stakeholder groups. In the first stage, interest groups push for the wholesale inclusion or exclusion of anticipated rules and regulation. In the second stage, the emphasis of lobbying shifts to the margins, as stakeholders attempt to limit the scope of proposed regulations. Second, our analysis reveals that, contrary to prior research (McKay & Yackee, 2007), firms do in fact monitor the lobbying activity of other stakeholder groups, and counter these attempts accordingly in the information they supply to policymakers in subsequent policy stages. This suggests that interest groups have come to appreciate that they may advance their policy objectives by observing more closely the demands of competing stakeholder groups. Hence, together these two policy phases constitute ‘shifting battlegrounds’, where firms first seek to influence what is included and excluded in the legislation, after which they need to engage the interests of other stakeholders.

DATA AND METHOD

Research setting

The EC possesses the right to initiate new legislation – giving it an important agenda-setting role within the EU – that then has to be negotiated and approved by other European institutions, namely the Council and EP (Wallace et al., 2015). Importantly, before the EC proposes a new regulation, it requests input from public and private stakeholders, broadly defined. At this point, there is no publicly available information to delineate the substance, scale, and scope of the new regulation. After the EC has taken stakeholders’ input into consideration, it publishes a communication to the other EU institution in which it clarifies what its key objectives are regarding the actual legislation. Crucially, it is only here that interest groups get a clear understanding of what other stakeholders have sought to include and exclude from the forthcoming legislation, and how the EC has internalized these demands. Following the publication of this communication, interest groups can again submit lobby documents to the EC clarifying their stance towards the different objectives and attempt to influence the EC towards certain policies. After this, the EC publishes its draft legislative proposal.

The GDPR – the focus of this paper – also went through this legislative procedure. The GDPR is a refinement and extension of the 1995 Data Protection Directive, which suffered from poor compliance and enforcement (Hoofnagle et al., 2019). The GDPR essentially constitutes a data governance framework that compels companies operating in the EU to take privacy and data protection issues as seriously as they do anti-trust and foreign corrupt practices law. Whereas previously corporations enjoyed the *de facto* freedom to cast a wide net in cyberspace and harvest all data they deemed remotely relevant for current or future business purposes, the GDPR incentivizes companies to take a more deliberate, responsible, and parsimonious approach to data collection and mining. Most importantly, the GDPR was the most lobbied law in EU history: the entire process took over two years; witnessed around 4000 amendments; and took another two years to be finalized and adopted (D’Cunha, 2020). This demonstrates the significance of the GDPR for the nonmarket environment of businesses.

Data collection

The documents we used as input for the topic modeling algorithm were accessed through a freedom of information request submitted to the EC. We requested to receive all documents the Commission had received related to the legislative procedure of the GDPR from the period of January 2009 until June 2011. These documents consist of replies to two open consultations, namely the consultation on the commission's 'Comprehensive Approach on Personal Data Protection in the European Union' (CAPDP), and the consultation on 'the Legal Framework for the Fundamental Right to the Protection of Personal Data (FRPDP). The CAPDP documents date around the period of January 2010, while the FRPDP document are timestamped around January 2011. Together these two consultations received a total of 491 documents. The documents were authored by companies, non-governmental organizations (NGOs), (national) public authorities, and trade associations. We categorized the different author types into one of four options: companies (e.g., Facebook), NGOs (e.g., European NGO Alliance for Child Safety Online), trade associations (e.g., American Chamber of Commerce), and public authorities (European Medical Agency).

Structural Topic Modeling

Topic modeling constitutes a form of unsupervised machine learning based on the Bayesian statistical technique of latent Dirichlet Allocation (Blei et al., 2003). Topic modeling is increasingly utilized by management scholars to reveal phenomenon-based constructs and grounded conceptual relationships in textual data (Hannigan et al., 2019). The key innovation of structural topic models is that it permits researchers to incorporate arbitrary document metadata into the topic model and estimate relationships between this metadata and topical prevalence and topical content (Roberts et al., 2014). Here, topical prevalence refers to the degree to which a document is associated with a topic, while topical content refers to the words used within topics. By metadata, we refer to information about the document itself, such as the author and publication date. In simpler terms, it allows researchers to estimate, for example, how often a particular topic, such as technological neutrality, occurs within a document written by a trade association. The incorporation of metadata into the topic model allows for a more structured interpretation of the underlying themes in the corpus. For our analysis, we incorporate the type of author of each document – i.e., whether it was authored by a company, NGO, public authority, or trade association.

DISCUSSION

The literature on CPA distinguishes three types of political strategies: financial incentives strategies, information strategies and constituency-building strategies (Hillman & Hitt, 1999). Information strategies are central to achieving lobbying success within the EU (Coen et al., 2021). Yet CPA scholars have little to say regarding the substance of these information strategies (Lenway et al., 2022), except insofar as they aim to influence or manipulate public discourse (Nyberg & Murray, 2020). Furthermore, CPA scholars have generally eschewed the EU as their empirical context in favor of the US (Chalmers & Macedo, 2021). Given the fact that the EU often acts as a global norm-setter in areas such as digitalization, consumer health and the natural environment (Bradford, 2020), it is a highly relevant context for CPA scholars to expand our knowledge.

Responding to calls for such research (e.g., Lenway et al., 2022), we have explored the substance of information strategies within the context of the GDPR. While ours is not the first

empirical study on the GDPR (e.g., Andrew & Baker, 2021), it is the first to systematically and comprehensively explore the entire set of lobby documents submitted to the EC in response to two GDPR consultation rounds. Aided by recent advances in the field of unsupervised machine learning (Roberts et al., 2014), we were able to uncover the underlying themes in GDPR lobby documents and link these themes to different author types.

Our analysis reveals that the consultation rounds differ significantly in terms of underlying themes. During the first consultation round, companies, NGOs, public authorities, and trade associations lobbied for the wholesale inclusion or exclusion of policy proposals. During the second consultation, where there were a number of regulatory ideas articulated and consulted, the lobbying strategy shifted towards limiting or expanding the proposed set of regulatory ideas. The key difference is that the manifestation of regulatory proposals and ideas for updating data protection rules served as an institution that constrained the demands stakeholder groups could make towards the EC.

Thus, information strategies take one of two forms. In the absence of any proposed legislation, companies and other stakeholder groups engage in a ‘categorical information’ lobbying strategy. We define this strategy as providing information to policymakers to sway them to either include or exclude wholesale elements from proposed legislation. This information strategy will dominate early lobbying efforts in the EU, where it has become apparent to relevant stakeholders that legislation is forthcoming, but there is uncertainty as to its actual substance. This was borne out in our data, where during the first consultation rounds discussions focused on broad topics such as the international transfer of personal data (topic 17); principles and definitions, role of self-regulation/co-regulation (topic 16); and defining data controllers and processors in a distributed environment (topic 6). These topics, their top 5 terms and example documents can be found in Table 1.

Table 1 about here

Second, after the EC communication had been published, companies and other stakeholders engage in a ‘conditional information’ lobbying strategy. We define this strategy as providing information to policymakers to sway them to introduce exceptions or limitations to legislation. One example we saw in our data was how during the second consultation lobbying efforts turned towards more specific issue areas, such as the organizational appointment of a data protection officer (topic 23); the categories of sensitive data (topic 2); and accountability and privacy-by-design (topic 22). These topics, their top 5 terms and example document can be found in Table 2.

Table 2 about here

This article makes several contributions. First, we expand the CPA literature by distinguishing the two following types of information strategies: categorical information strategy and conditional information strategy. We believe this distinction will provide CPA scholars with a meaningful tool to differentiate between the substance of information strategies, an area hitherto considered a blind sport in CPA research (Lenway et al., 2022).

Second, we expand the concept of ‘institutions’ in nonmarket strategy studies to include

institutional procedures. Contemporary CPA research has been heavily influenced by the work of Baron (1995, 1997), who mainly posited institutions as governmental agencies, such as Congress and the Federal Trade Commission. However, we have demonstrated that firms deal with other institutions in their attempt to influence government policy as well. Namely, during the policymaking process, their opportunities to influence public policy are more extensive before regulatory bodies communicate their legislative intentions. These institutional procedures are equally important for firms to recognize and incorporate in their nonmarket strategies.

In addition to these main contributions, we also demonstrate that stakeholder groups consider the lobbying activity of others and engage with this lobbying activity in subsequent rounds. This contradicts earlier research that found that interest groups do not in fact engage in counter lobbying (McKay & Yackee, 2007). Our finding seems to suggest that interest groups have come to the realization that they can achieve more from their lobbying activities by monitoring the behavior of their opponents. Finally, in methodological terms we show, following the suggestion of Lenway et al. (2022), how newer approaches can help to study CPA topics.

REFERENCES AVAILABLE FROM THE AUTHORS

Table 1. Three example topics from the ‘comprehensive approach on personal data protection in the European Union’ corpus.

Topic	Top 5 terms	Label	Example document (truncated)
17	transfer, country, standard, international, rule	International transfers of personal data	The rules on model clauses for international data transfers are a good example of this phenomenon. Model clauses concern the existence – on paper – of rules that may not be observed in practice.
6	controller, processor, service, provider, party	Defining data controllers & processors in a distributed environment	For instance, it is effectively the case when companies outsource the processing of data or in cloud computing scenario. [...] Therefore rules of liability of the data controller should be made more flexible.
16	principle, practice, base, approach, regulatory	Principles and definitions, role of self-/co-regulation	Directive 95/46/EC has stood well the influx of these technological developments because it holds principles and uses concepts that are not only sound but also technologically neutral.

Table 2. Three example topics from the ‘legal framework for the fundamental right to the protection of personal data’ corpus.

Topic	Top 5 terms	Label	Example document (truncated)
2	sensitive, category, context, type, definition	Categories of sensitive data	A broadening of the special categories of sensitive data to unnecessary types of data could lower the perceived threshold where data subjects are granted additional safeguards, and thereby potentially jeopardise

			the protection of genuinely sensitive data.
23	company, organisation, business, market, internal	Organizations appointing a data protection officer	As regards making the appointment of an independent Data Protection Officer mandatory, we have no objections to any company choosing voluntarily to appoint a Data Protection Officer (DPO).
22	principle, design, accountability, measure, concept	Accountability and Privacy-by-design	Public authorities have failed to provide examples of best practice in the EU. From health databases to smart metering and publicly-funded transport ticketing systems, there has been a fundamental failure to ensure proportionality, privacy by design and data minimisation.