

Discussion Paper Series – CRC TR 224

Discussion Paper No. 659

Project B 02

# Effective Regulation and Firm Compliance: The Case of German Privacy Policies

Jacopo Gambato<sup>1</sup>  
Bernhard Ganglmair<sup>2</sup>  
Julia Krämer<sup>3</sup>

March 2025

<sup>1</sup>University of Vienna, Email: [ja.gambato@gmail.com](mailto:ja.gambato@gmail.com)

<sup>2</sup>University of Mannheim and ZEW Mannheim, Email: [b.ganglmair@gmail.com](mailto:b.ganglmair@gmail.com)

<sup>3</sup>Erasmus University Rotterdam, Email: [j.k.kramer@law.eur.nl](mailto:j.k.kramer@law.eur.nl)

Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)  
through CRC TR 224 is gratefully acknowledged.

# Effective Regulation and Firm Compliance: The Case of German Privacy Policies\*

Jacopo Gambato<sup>†</sup>      Bernhard Ganglmair<sup>‡</sup>      Julia Krämer<sup>§</sup>

February 25, 2025

## Abstract

This paper explores the interaction between the enforcement of and compliance with difficult-to-enforce rules in the context of data regulation. We focus on the effect of the introduction of the GDPR and its transparency principle on the readability of privacy policies for a large sample of German firms. Germany has a system of state-level data protection authorities. These data regulators enforce the same rules but face diverse funding situations, allowing for an ideal setting to study the role of a regulator’s capacity in firms’ compliance decisions. We find that while, on average, the GDPR lead to less readable policies, firms active in industries that have in the past received more regulatory scrutiny and those active in jurisdictions of better-funded data regulators exhibit a stronger compliance with the GDPR’s readability requirement. These results exemplify a more general interaction between regulators’ enforcement activity and firms’ regulatory compliance.

**Keywords:** data protection, GDPR, privacy policies, readability, regulation, text-as-data

**JEL Codes:** D22; K20; L51.

---

\*This paper summarizes results from our research paper “Regulatory Compliance with Limited Enforceability: Evidence from Privacy Policies” (Ganglmair (r) Krämer (r) Gambato 2024), offers additional information on institutional features of data regulation in Germany, and provides new descriptive evidence. We thank participants at the NBER Conference “Data Privacy Protection and the Conduct of Applied Research: Methods, Approaches and their Consequences” for a lively discussion and helpful comments and suggestions. We also thank Ruobin Gong, V. Joseph Hotz, and Ian M. Schmutte for their editorial guidance. Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project B02) is gratefully acknowledged. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

<sup>†</sup>University of Vienna, [ja.gambato@gmail.com](mailto:ja.gambato@gmail.com)

<sup>‡</sup>University of Mannheim and ZEW Mannheim, [b.ganglmair@gmail.com](mailto:b.ganglmair@gmail.com)

<sup>§</sup>Erasmus University Rotterdam, [j.k.kramer@law.eur.nl](mailto:j.k.kramer@law.eur.nl)

# 1 Introduction

In this paper, we ask to what degree the firms’ compliance with the rules and requirements of the EU General Data Protection Regulation (GDPR) of 2018 is the outcome of a strategic interaction between firms and regulators. For our discussion, we focus on German firms and their compliance with the readability requirement as a central aspect of the GDPR’s transparency principle (Art. 5(1) lit. a GDPR). The requirement compels firms to make disclosures about their personal-data collection and processing in a “concise, transparent, intelligible and easily accessible form, using clear and plain language” (Art. 12(1) GDPR). A central aspect of our discussion is how the funding situation of data regulators affects firms’ compliance decisions when both GDPR enforcement and compliance are costly.

Germany provides an ideal setting to study regulators’ roles in firm-level compliance with the GDPR (German: *Datenschutz-Grundverordnung*). First, Germany is said to be one of the EU member states with the strictest data regulators<sup>1</sup> enforcing the rules and requirements of the GDPR without a national-interest or industrial-policy agenda.<sup>2</sup> Second, each of the 16 German states has an independent data protection authority with its own budget by the respective state legislature. These state-level data regulators enforce the same rules but may set different priorities because of their respective funding situations. In Section 2 of this paper, we provide further institutional details about state-level data regulation and present evidence of sizeable differences in the regulators’ budgets. Our empirical strategy builds on these differences, allowing us to document how firms adapt their compliance decisions.

Compliance with the GDPR’s transparency principle is particularly interesting from a consumer-protection perspective. The ease of access to information (or the lack thereof) in

---

<sup>1</sup>Johnson (2024) reports results from a survey of data processors asked to compare their local data regulator with others in the EU (strictest regulators in Germany and Sweden). The German business press (Anger and Neuerer 2020) quotes the president of Germany’s digital association (Bitkom), representing more than 2,200 companies of the digital economy, saying that the laxer interpretation of the rules is a disadvantage for Germany as a place for doing business.

<sup>2</sup>For instance, the Irish Data Protection Authority has been criticized for its perceived lack of commitment in addressing complaints against potential GDPR violations by big tech companies such as Meta and TikTok, which have their EU headquarters in Ireland. For an in-depth discussion, see the coverage in *Politico* (Vinocur 2019) or *The Observer* (Naughton 2020).

online terms of use, end-user license agreements, or privacy policies has been in the policy and media spotlight for several years (e.g., Hern 2015; Dwoskin 2015; Litman-Navarro 2019): easily accessible and understandable information is essential for users to make informed decisions. However, because the concept of readability is vague and ambiguous, it is a potential candidate for deprioritization by underfunded data regulators. Rational firms can be expected to exploit a possible lack of oversight, writing incomprehensible privacy policies to the disadvantage of consumers, who are easily subject to strategic (ab)use (e.g., when firms try to obfuscate by using difficult and complex language). We offer more information on the readability requirement in the GDPR’s transparency principle, where it comes from, and what it means, in Section 3 of this paper.

In Section 4, we introduce our empirical approach and present descriptive evidence of the readability of privacy policies for a larger sample of German firms from 2014 to 2021. We document that readability has not improved with the GDPR (and its stronger focus on transparency). Building on the first three sections of this paper, in Section 5, we turn to the role of regulators and how firms respond when they believe to be under higher regulatory scrutiny. This section draws from and summarizes results derived in Ganglmair <sup>Ⓘ</sup> Krämer <sup>Ⓘ</sup> Gambato (2024). These results suggest that regulation is effective despite the negative results for average readability. Firms in industries that are likely to see more stringent regulation (captured by industry-level data on past regulatory activities by the UK Information Commissioner’s Office) are also more likely to improve the readability of their privacy policies (or lower the readability less). Furthermore, firms in states with better-funded data regulators respond more by exhibiting better readability compliance.

Our findings are relevant beyond the European Union. In the United States, data privacy regulation is becoming increasingly fragmented, with many states passing their own data privacy acts. Many of these use language similar to that used for the readability requirement in the GDPR. For instance, the California Consumer Privacy Act (CCPA) requires that information be made available in a “format that is easily understandable to the average

consumer” (1798.130. (B) (iii) CCPA), and the very recent Colorado Privacy Act (CPA) mandates that a privacy notice should be “reasonably accessible, clear, and meaningful” (§6-1-1308 (1)(a) CPA).<sup>3</sup> Researchers studying the GDPR and (data) regulation more generally need to understand that regulation may not happen as written, but firms will optimally respond given their circumstances (see, e.g., Johnson 2024). We expect our results (in Ganglmair (r) Krämer (r) Gambato (2024) and this paper) to inform the debate on how best to regulate firms in an environment of potentially underfunded regulators that have to set priorities and may neglect one set of rules over another.

As such, this paper relates to a broader literature on privacy policies (e.g., Milne et al. 2006; Degeling et al. 2019; Linden et al. 2020; Becher and Benoliel 2021; Amos et al. 2021; Frankenreiter 2022; Wagner 2023), the effects of the GDPR (e.g., Yuan and Li 2019; Koski and Valmari 2020; Peukert et al. 2022; Johnson et al. 2023; Goldberg et al. 2024), and the performance of vague and ambiguous laws and regulations or regulation under capacity constraints (e.g., Stern 2000; Laffont 2005; Armstrong and Sappington 2006; Giommoni et al. 2023). We argue that these intertwined topics highlight a more general interaction between firms’ incentives to comply with regulation and regulators’ ability to enforce said compliance.

## 2 Privacy Regulation in Germany

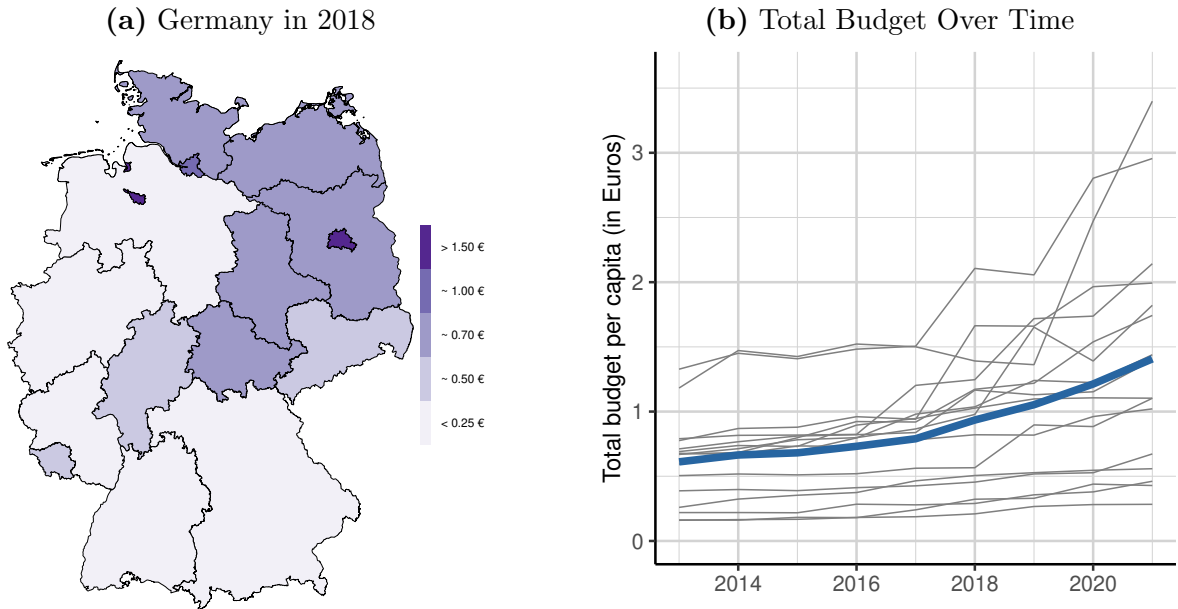
In Germany, each state has its own independent *supervisory authority* for data protection, with the exception of Bavaria, which has two separate authorities: one for public entities and another for private ones.<sup>4</sup> In addition to these 17 state authorities, there is the Federal Commissioner for Data Protection (*Bundesbeauftragter für Datenschutz*), who oversees the data protection activities of federal public authorities (§8 BDSG, Federal Data Protection Act) and serves as a member of the European Data Protection Board (EDPB). To ensure

---

<sup>3</sup>The European Commission, too, continues to use the same type of language, for instance, in the Platform-to-Business (P2B) Regulation or the Digital Services Act.

<sup>4</sup>The GDPR uses the term “supervisory authority” for what we refer to as the data protection authority (i.e., the “regulator”).

Figure 1: German GDPR Regulators Vary in Budget



Notes: In this figure, we depict the total budget per capita (in Euros) of the 16 state data protection authorities in Germany (for Bavaria, with two authorities, we restrict our attention to the one governing private actors). In panel (a), we show a map of Germany with the budget in 2018; in panel (b), we plot the time series of the total budget for the individual data protection authorities (in gray) and the nationwide average (in blue). *Source:* Ganglmair (P) Krämer (P) Gambato (2024) (for budget values) and Statistisches Bundesamt (Destatis) (2024) (for population numbers).

consistent legal interpretation and application (of the same set of rules set out in the GDPR), all state data protection authorities and the Federal Commissioner regularly convene in the *Datenschutzkonferenz*, a forum dedicated to maintaining uniform legal enforcement.

The individual data protection authorities are responsible for overseeing their subjects (e.g., firms) in their respective jurisdictions. The relevant jurisdiction of a given firm is determined based on the location of the firm’s central administration (Art. 4(16) GDPR). This central administration refers to the establishment where the primary management activities (e.g., headquarters) occur, regardless of whether data processing actually takes place at this location.<sup>5</sup>

<sup>5</sup>See Recital 36 GDPR. In cases where the relevant authority in Germany is unclear (e.g., when a firm has multiple establishments in different states), §40 of the German Federal Data Protection Act outlines a procedure to identify the responsible data protection authority. Once identified, the designated data protection authority will assume control of the case.

A firm’s enforcing data protection authority is determined by location. The capacity of that authority—and its capability to implement the uniform legal rules—is determined by its budgetary situation. The authorities are independent institutions (Art. 52(1) GDPR) but are subject to funding by the respective member states (here, the German states). The GDPR compels the member states to provide “the human, technical and financial resources” that are “necessary for the effective performance of its tasks” (Art. 52(4) GDPR). The result of this funding mandate, potentially leaving data protection authorities exposed to budget politics, is considerable variation of the authorities’ budgets across states (both within Germany and across EU member states), as shown in Figure 1.

Panel (a) of the figure depicts the total budgets per capita of the German data protection authorities for the year 2018. The states of Berlin, Bremen, and Hamburg (all densely populated) have the best-funded authorities. Among the other states, those in the north and northeast are better funded than the authorities in the southwest and west. Comparing the situation with other EU member states (in 2013), Schütz (2018) observes that Germany takes an average position with its aggregated budget of around one-half Euro per capita. For reference, the data protection authorities in the UK, Italy (both EUR 0.39), Spain (EUR 0.29), or France (EUR 0.26) face budget situations of a similar order of magnitude as Germany.

In panel (b) of Figure 1, we depict the budget situation over time for the years 2014 to 2021. All data protection authorities have seen an improvement in their funding situation—with some more than others—that has accelerated in 2018 with the enforcement of the GDPR. Overall, the average per-capita budget has almost tripled in the nine years depicted in the figure. Despite some of these increases in annual budgets, a common theme in the state authorities’ annual reports is the discussion of a tight budgetary situation given the variety of enforcement tasks imposed on the authorities by the GDPR.

One of these tasks is the investigation of individuals’ complaints. Individuals who believe their rights have been infringed can file a complaint against a firm (i.e., the “data controller”)

with the relevant data protection authority. Such a complaint (and subsequent investigation) can eventually result in an authority decision identifying an infringement by a firm.<sup>6</sup> Within the enforcement framework of the GDPR, mechanisms for both public law sanctions (through the data regulator) and private enforcement of civil claims (through courts) co-exist. In addition to filing complaints with the data protection authorities, individuals also have the right to pursue GDPR claims in court against private entities. This course of action is especially significant for seeking damages under Art. 82 GDPR. The two pathways complement each other: damages serve a compensatory purpose for individuals, while public fines imposed by data protection authorities are intended to have a preventive effect (Chamberlain and Reichel 2020, 668).

### 3 The Transparency Principle in the GDPR

The GDPR has fundamentally transformed the way any data controller—any entity that determines the means and purposes of collecting and processing personal data (Art. 4(7) GDPR)—is required to collect, process, and store data and communicate these details to users and the public. One of the principles contained in the GDPR is *transparency* (Art. 5(1) lit. a GDPR). It requires any information concerning the processing of personal data to be easily accessible and understandable, which enables users to make informed decisions about who is allowed to process their data and under what conditions.

More specifically, Art. 13 and 14 GDPR require a firm to provide consumers with details about the data processing (e.g., what data is collected, how, and by whom),<sup>7</sup> and Art. 12 GDPR specifies the procedural and technical aspects of that information provision, compelling firms to make the required disclosures in “concise, transparent, intelligible and easily accessible form, using clear and plain language” (Art. 12(1) GDPR)—a *readability requirement*.

---

<sup>6</sup>Firms can contest these infringement decisions (and associated fines) in national courts (Art. 78 GDPR).

<sup>7</sup>Elements that have to be disclosed include, for instance, the contact details of a firm, the legal basis the data processing is based on, and the duration of the storage of personal data.





Readability, however, is a vague and ambiguous concept, and the GDPR provides little to no guidance on how this particular aspect of the GDPR should be interpreted and enforced. The Art. 29 Working Party (2018), a former advisory body within the EU’s data protection framework, has tried to address these issues by providing non-binding guidelines to facilitate a consistent application of the law. It gave terms such as “concise and transparent,” “intelligible,” and “clear and plain language” a more precise definition. It also emphasized the needs of the “average member of the intended audience” and how that average user ought to be able to easily access information expressed in “as simple a manner as possible.” As a standard to assess compliance with Art. 12(1), the Working Party proposed mechanisms such as “readability testing.” However, because the Working Party did not provide guidance on *which* measures should be considered suitable for privacy policies (and linguists have over the years developed dozens of candidate indices and scores), and because the European Court of Justice has not yet clarified how to assess compliance with Art. 12(1), the Working Party’s suggestions serve, at best, as loose guidance for both firms and data protection authorities.

As we discuss in greater detail below, for our analysis, we use a readability score with regulatory history, one that has been used in the U.S. to regulate the language of insurance contracts. We believe it is an obvious candidate for data regulators to use—and, equally important, for firms to anticipate.

## 4 Firm-Level Compliance

In this section, we present descriptive evidence of German firms’ responses to the GDPR, zooming in on the readability requirement that compels firms to disclose information about the nature of their data collection, processing, and use (Art. 13–14 GDPR) in accessible and readable language (Art. 12(1) GDPR).

## 4.1 Data and Measurement



For our analyses, we use the texts of privacy policies of German firms web-scraped from the Internet Archive’s Wayback Machine (Ganglmair  Krämer  Gambato 2024). We restrict our sample to firms for which we have at least one policy before the enforcement of the GDPR in May 2018 and one policy after, allowing us to track each firm over time. With this sample restriction, we have 585,329 quarterly observations (i.e., privacy policies) posted by 75,683 German firms between Q1 2014 and Q2 2021.<sup>8</sup> The average number of policies per firm is 4.4 pre-GDPR enforcement (May 2018) and 3.3 post-GDPR enforcement.

Our measure of compliance is based on a firm’s communication to the public (consumers, suppliers, competitors, and regulators) through its privacy policy. A key aspect of the transparency principle is making communication accessible in “concise, transparent, intelligible and easily accessible form, using clear and plain language” (Art. 12(1) GDPR).

To assess compliance with the readability requirement, we rely on the work of linguists who have developed readability indices for many decades to measure the ease (or difficulty) of written texts. One of the most popular readability indices (developed for English texts) is the Flesch Reading Ease Score (FRE) (Flesch 1948), a weighted average of the average sentence length and the average word length of a text. The FRE is widely used in research<sup>9</sup>. Moreover, it has a relevant regulatory history in the United States. For instance, in Michigan and Massachusetts, an insurance contract must have an FRE score of at least 50; in Texas,

---

<sup>8</sup>The average firm in our sample has 36 employees and sales of EUR 15 million. We further classify 61.6% as micro firms (less than 10 employees), 36.3% as small and medium-sized enterprises (between 10 and 250 employees), and 2% as large firms (more than 250 employees). Comparing our sample to the 2017 distribution of firm sizes in the Mannheim Enterprise Panel (*Mannheimer Unternehmenspanel, MUP*), the most comprehensive micro-data base of companies in Germany beside official administrative data (Bersch et al. 2014), we find that micro and large firms are somewhat underrepresented. The largest sector in our sample is the services sector (58.6% of all firms in 2017), followed by trade (22.3%), manufacturing (9.6%), construction (7.0%), utilities (1.5%), and agriculture/mining (1%). Services, manufacturing, and utilities are over-represented (relative to the MUP), whereas trade, construction, and agriculture/mining are underrepresented.

<sup>9</sup>The article introducing the FRE has more than 6,000 Google Scholar citations; the FRE generates more than 25,000 Google Scholar search hits. The runner-ups for citations are the Flesch-Kincaid Readability Score and Simple Measure of Gobbledygook (SMOG). The Gunning’s Fog Index ranks fourth with the second-highest number of Google Scholar search results per March 31, 2023 (Ganglmair  Krämer  Gambato 2024).

the minimum score of the FRE is 40; and similar guidelines (with a minimum score of 45) exist in Florida.<sup>10</sup>

Because our text corpus comprises privacy policies in German, we use the German version of the FRE (we refer to it as the *German FRE*), which was developed by Amstad (1978).<sup>11</sup> By applying different weights to the average sentence length, *ASL*, and word length, *AWL* (in syllables), it takes into account the specific features of the German writing style, for which long words and long sentences are common.<sup>12</sup> The German version of the FRE is defined as:

$$\text{German FRE} = 180 - \text{ASL} - 58.5 \times \text{AWL}.$$

Higher values of the German FRE imply higher readability (presuming that texts with longer sentences and longer, more complex words are more difficult to read). Values below 50 are considered difficult texts; texts with values between 60 and 70 are appropriate for 7–8th graders; and texts with values of 90 or more are appropriate for 5th graders (Immel 2014, 17–19).

## 4.2 Firms Did Not Improve the Readability of Their Policies

Based on the German FRE, we find that the readability of German firms' privacy policies has *not* improved in response to the GDPR in 2018. In fact, the average readability score of post-GDPR privacy policies is 3–4% lower than that of those policies posted before the GDPR. We plot the evidence for our raw data in panel (a) of Figure 2.

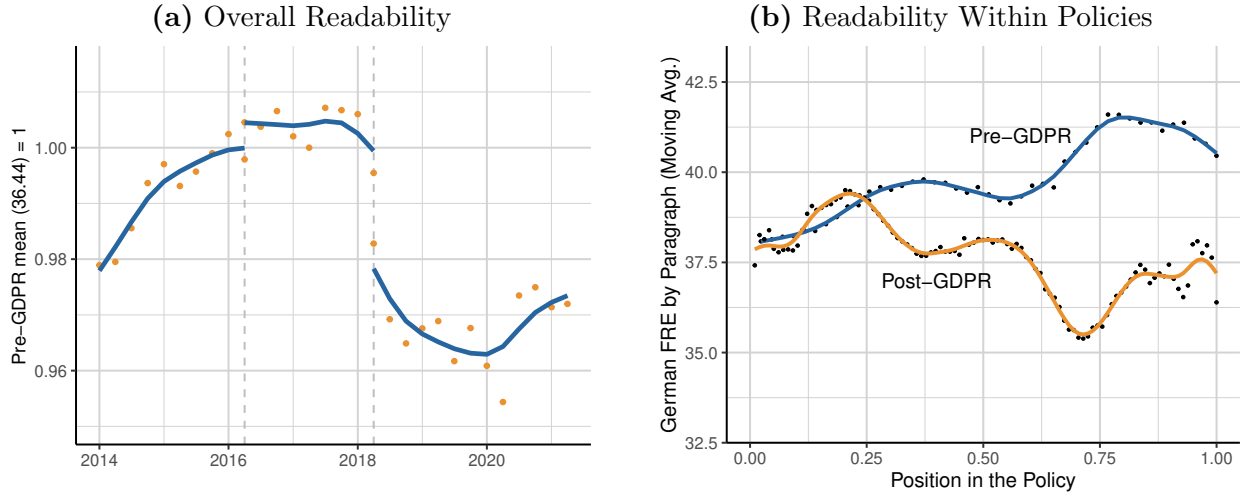
---

<sup>10</sup>Michigan Compiled Laws, Section 500.2236 (2020); General Laws of Massachusetts, Title XXII, Chapter 175 Section 2B. (2014); Texas Insurance Code, Section 2301.053 (2019); Florida Statute §627.4145, Readable language in insurance policies; available at <https://flsenate.gov/Laws/Statutes/2021/0627.4145> (accessed August 13, 2024).

<sup>11</sup>Several other readability measures explicitly calibrated for German texts have been proposed over the years. See, for instance, Bamberger and Vanecek (1984). The German FRE in Amstad (1978) has the benefit of its proximity to existing regulatory practices.

<sup>12</sup>Legal writings (such as privacy policies) have that in common. They have unusually long sentences, with an average sentence containing twice as many words as in other categories of texts (Gustafsson 1984).

Figure 2: Readability of Privacy Policies Before and After the GDPR



Notes: This figure presents the evolution of the readability of privacy policies, measured as the German version of the Flesch Reading Ease Score (German FRE) by Amstad (1978). In panel (a), we plot the German FRE’s quarterly averages (dots) relative to the mean value of 36.44 for all pre-GDPR observations. The curves (in blue) are fitted to the data (spline). The vertical dashed lines indicate the GDPR passage in Q2 2016 and GDPR enforcement in Q2 2018. In panel (b), we plot the moving average (10-paragraph wide window) of the paragraph-level readabilities for all pre-GDPR privacy policies (in blue) and post-GDPR policies (in orange). Length of policies is normalized to one; values on the horizontal axis represent the relative position within a policy. *Source:* Ganglmair (r) Krämer (r) Gambato (2024) (for the privacy policy panel), Benoit et al. (2018) (for the construction of readability scores), own calculations.

The figure depicts the quarterly averages of the German FRE, split into three different phases: (i) the time before the adoption of the GDPR in Q2 2016, (ii) the time after the adoption but before the GDPR went into effect in Q2 2018, and (iii) the time after the GDPR went into effect. We plot the values relative to the mean of all pre-GDPR observations (the mean pre-GDPR German FRE is 36.44).<sup>13</sup> Relative to the pre-GDPR mean, the readability of policies posted after the GDPR went into effect is about 3–4% lower (phase (iii)). We do not observe a sizeable announcement effect (between phases (i) and (ii)), suggesting that firms were not eager to update their privacy policies before they were forced to.<sup>14</sup>

In panel (b) of Figure 2, we plot the readability as it varies by position within the

<sup>13</sup>Ironically (yet in line with the literature on the European Commission’s communications (Rauh 2023)), the German text of the GDPR itself (the *Datenschutz-Grundverordnung*, a seven-syllable word) is highly unreadable, with a German FRE score of 8.83.

<sup>14</sup>The increase in readability in the early years of our sample during phase (i) might be the result of firms responding to the draft proposal (European Commission 2012) and early responses (European Parliament 2013).

policy, from the beginning to the end. The blue curve depicts the readability of all pre-GDPR policies (phases (i) and (ii)), and the orange line represents the readability of all post-GDPR policies. For both sets of policies, the first quarter of an average policy exhibits increased readability (higher values). For pre-GDPR policies, this trend continues: privacy policies become easier to read further into the text. For post-GDPR policies, we observe a different pattern: the latter parts of policies become more difficult to read (with a partial reversal in the last quarter). This pattern is consistent with a drafting strategy focusing on (relative) readability in the early portions of a policy (presumably with a larger readership) while neglecting that aspect of transparency in later parts in which more of the required disclosures are made. It does, of course, not explain why we do not observe this strategy for pre-GDPR policies. A possible explanation is that the “transparency paradox” (Nissenbaum 2011) became binding only after firms started incorporating the new disclosure requirements imposed by the GDPR.

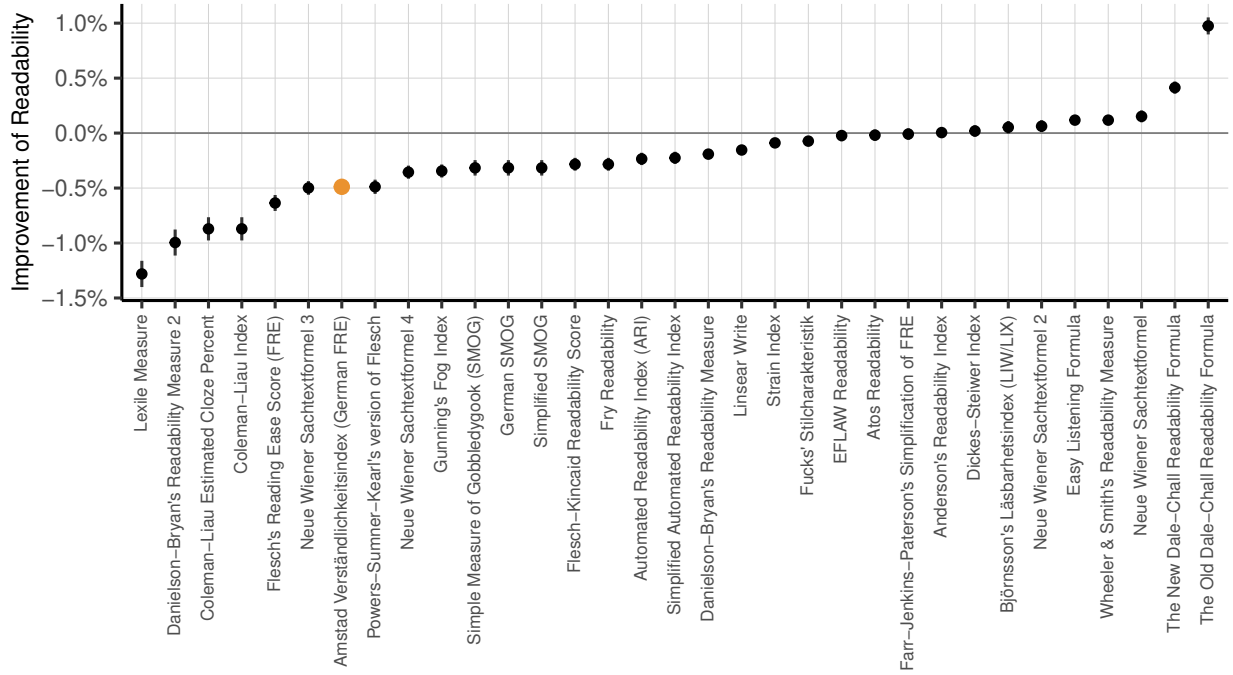
The average readability based on the German FRE has worsened in response to the GDPR, as documented in Figure 2. This result, however, is not limited to the use of the German FRE as our measure for readability. We repeat our exercise for 33 readability scores and use them one-by-one as dependent variables in simple regression models with a Post-GDPR indicator variable equal to one if a given policy was posted after the GDPR went into effect and zero if posted before.<sup>15</sup> We further control for firm size (using the number of firm employees) and industry concentration (using the Herfindahl-Hirschman Index); we also use firm fixed effects and year fixed effects to control for unobserved heterogeneity. The Post-GDPR indicator variable captures the average change in the readability of privacy policies once other observable and unobservable factors have been accounted for.

In Figure 3, we plot the estimation coefficients (and the 99% confidence intervals) of this indicator variable for all 33 readability scores, ranked from the largest GDPR-induced

---

<sup>15</sup>We use Benoit et al. (2018) to calculate the readability scores for our privacy policy panel. The list represents a small fraction of the more than 200 scores and indices developed by researchers over the years (Immel 2014, 17–19); it does include the most popular ones.

Figure 3: Consensus Among Readability Scores



Notes: This figure presents the coefficients (and 99th confidence intervals) of the post-GDPR indicator variable (equal to one for all post-GDPR observations, zero otherwise) from OLS regressions with log values of various readability scores (and number of firm employees, industry HHI at the 4-digit NACE industry-level as additional control variables; firm fixed effects, and year fixed effects). Positive values on the vertical axis imply an improvement in the readability of privacy policies following the enforcement of the GDPR in Q2 2018. *Source:* Ganglmair (R) Krämer (R) Gambato (2024) (for the privacy policy panel), Benoit et al. (2018) (for the construction of readability scores), own calculations.

decrease to the largest increase in readability. We find a statistically negative change in the readability for 22 of the 33 scores (including the German FRI) and a positive effect for only seven of the scores (from LIW to Dale-Chall).<sup>16</sup> Our main findings are, therefore, not just an artifact of our readability score of choice but highlight a broader pattern: two out of three readability scores indicate a decrease in the readability of privacy policies.

<sup>16</sup>For the remaining five scores (from EFLAW to Dickes-Steiner), we find a precisely estimated null effect.

### 4.3 Hypothetical Compliance Rates

Aggregate statistics for firm-level compliance with the GDPR do not exist.<sup>17</sup> In the case of readability, such statistics would be difficult to compile because, for an assessment of compliance, one needs a minimum level of readability (i.e., a threshold value for the readability score of choice) against which a firm’s policy is compared.

Neither the GDPR nor any commentary provides such a readability threshold. However, we can resort to existing readability thresholds in insurance-contract regulation to determine the level of compliance if privacy regulators in Germany were to use the same or similar thresholds. In the United States, state regulators often require a minimum value of the Flesch Reading Ease score for insurance contracts. For instance, in Texas, an insurance contract must have an FRE score of at least 40 (a “lenient” requirement); in Michigan and Massachusetts, the minimum score of the FRE is 50 (a “strict” requirement).

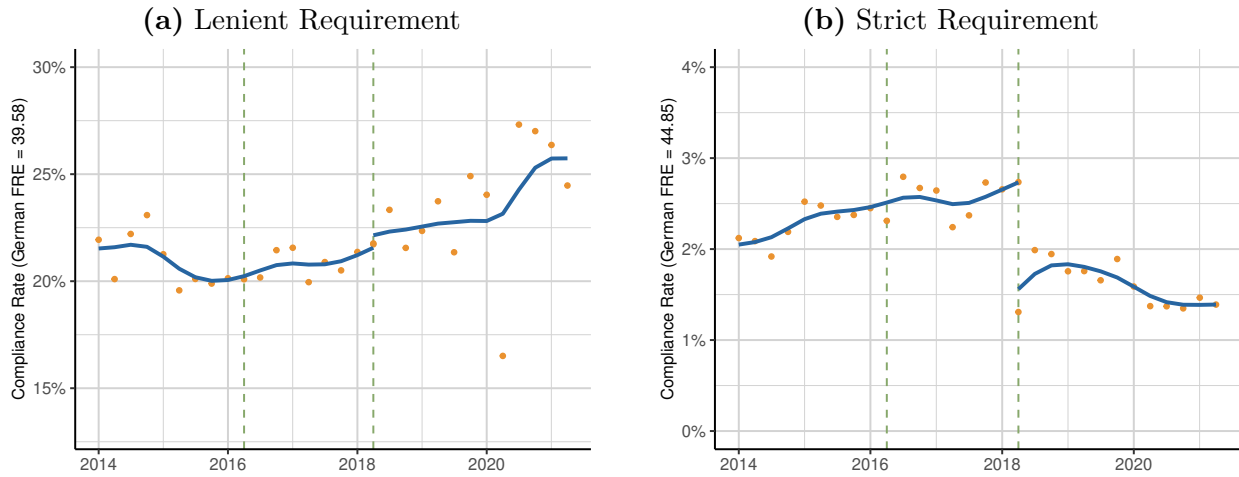
For the construction of *hypothetical compliance rates* of German privacy policies, we use German equivalents of these regulation-tested thresholds (for the English FRE).<sup>18</sup> In Figure 4, we plot the quarterly compliance rates for the lenient requirement with the lower threshold (in panel (a)) and the strict requirement with the higher threshold (in panel (b)). For the lenient requirement, average compliance rates are well below 30%: only three out of ten firms write privacy policies that pass our hypothetical readability test. For the strict requirement, average compliance rates are in the low single digits and decreased even further with the enforcement of the GDPR in 2018.

---

<sup>17</sup>The German data protection authorities do not publish such numbers, nor are we aware of any scholarly work that provides large-scale analysis of GDPR compliance.

<sup>18</sup>While the German FRE is a recalibrated English FRE to fit German texts, we acknowledge that legal texts (such as privacy policies) might introduce additional language-specific variation (see, e.g., Gustafsson 1984) that the recalibration cannot capture. We, therefore, first determine the German equivalents of the U.S. thresholds by using the distribution of the FRE scores of a sample of more than 100,000 English-language privacy policies posted between 2014 and 2017 (Amos et al. 2021). For the German equivalent of the U.S. threshold, we find the position (i.e., percentile) of a given FRE score in the distribution of the FRE scores of English-language policies and then determine the value of the German FRE at that very position of the distribution of the German FRE scores of our privacy policies. For the U.S. insurance-contract threshold of 40, the German equivalent is 39.58; for the threshold of 50, the equivalent is 44.85.

Figure 4: Hypothetical Compliance Rates for the Readability Requirement



Notes: This figure presents hypothetical compliance rates for privacy policies if German regulators used the same thresholds as those used in Texas (lenient requirement in panel (a), FRE of 40) and Massachusetts and Michigan (strict requirement in panel (b), FRE of 50) for the regulation of insurance contracts (see Wagner 2023). To determine the corresponding thresholds for German texts, we use a sample of English-language privacy policies (posted pre-GDPR, between 2014 and 2017) compiled by Amos et al. (2021), calculate their respective FRE scores, and determine the percentages of policies with scores below 40 and 50. Assuming that the distributions for German and English FRE scores are the same but for a shift parameter, we can then calculate the percentiles of these percentage thresholds. For the “lenient requirement,” this threshold is 39.58; for the “strict requirement,” it is 44.85. *Source:* Ganglmair (r) Krämer (r) Gambato (2024) (for the privacy policy panel), own calculations.

## 5 The Role of Regulators

The evidence we present in Figures 2–4 paints a sobering picture: on average, the readability of privacy policies has worsened with the GDPR, and compliance with (hypothetical) readability testing is strikingly low. Some of this may be the consequence of the “transparency paradox” (Nissenbaum 2011) or what the Art. 29 Working Party (2018, para. 34) referred to as an “inherent tension” between providing detailed information in a “concise, transparent, intelligible and easily accessible” form. In Ganglmair (r) Krämer (r) Gambato (2024), we show that the average length of privacy policies has doubled to tripled in length, their scope has doubled, and the volume of required disclosures has tripled to quadrupled in response to the GDPR. This stark increase in information may have simply come at the cost of readability.



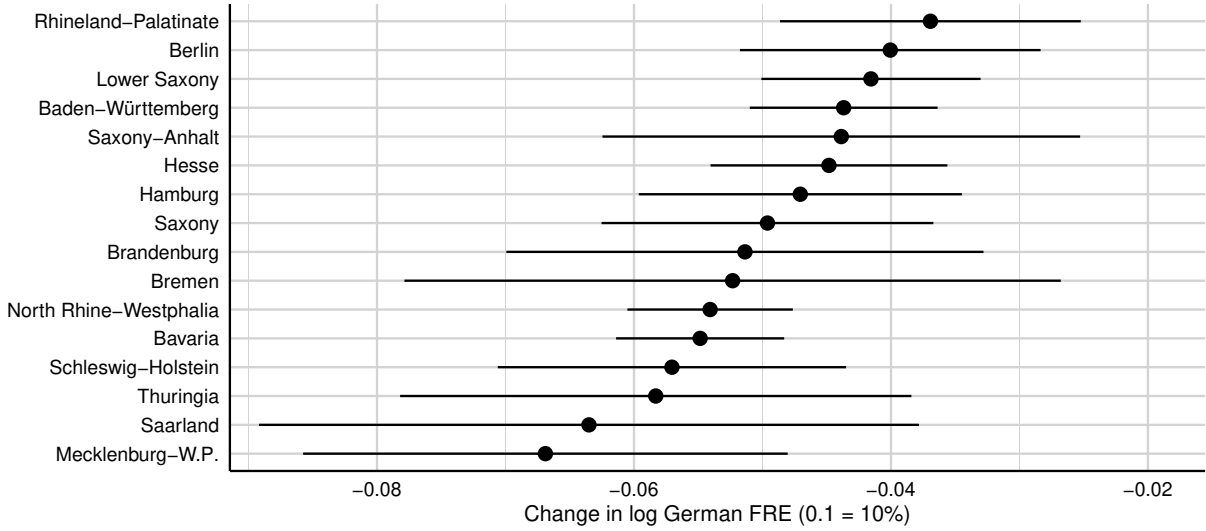
Our findings in Ganglmair [\(r\)](#) Krämer [\(r\)](#) Gambato (2024), however, also suggest that the effects on readability are not purely mechanical. In that paper, we argue that firms do indeed respond to the readability requirement of the GDPR—but not in a one-size-fits-all fashion. For instance, firms with pre-GDPR privacy policies of low readability improved the readability of their policies. In contrast, firms with high-readability policies before the GDPR saw a decline in their readability after the GDPR—they experienced a weaker GDPR treatment intensity. We do not find evidence suggesting that these patterns are the result of a general convergence but, indeed, the response to the GDPR.

In our paper, we further document that firms’ compliance with the readability requirement is (partially) driven by the level of scrutiny to which they can reasonably expect to be exposed—in their respective industry or by their state’s data regulator.

First, we find that firms respond to how much scrutiny they anticipate from data regulators, which means that firms that expect regulators to be more active display better compliance. More specifically, firms in industries that are likely to see more stringent regulation are also more likely to improve the readability of their privacy policies (or lower the readability to a smaller degree). For this analysis, we use data on past regulatory activities by the UK Information Commissioner’s Office (provided by Koutroumpis et al. (2022)). The UK ICO enforces data privacy laws, and industry-level case counts (scaled by the number of firms in the respective industries) measure this data regulator’s enforcement activities across different industries. For this approach, we assume that industries that a regulator scrutinized before the GDPR were also primary targets after the GDPR. We believe it is reasonable to assume that this enforcement variation is also true for Germany.

Second, firms not only show differences in readability compliance across industries (with enforcement history a key factor) but also across states, or rather, regulatory bodies. In Figure 5, we plot the estimated coefficients from our regression models (see the description for Figure 3) separately for each state and rank the states by the respective size of the coefficient. The GDPR-induced change in readability is negative in all states (as is the

Figure 5: The GDPR-Effect on Readability by State



Notes: This figure presents the coefficients (and 95th confidence intervals) of the post-GDPR indicator variable (equal to one for all post-GDPR observations, zero otherwise), conditional on the firm’s state, from OLS regressions with log values of the German FRE (and number of firm employees, industry HHI at the 4-digit NACE industry-level as additional control variables; and firm fixed effects). Negative values on the horizontal axis imply a worsening of the readability of privacy policies following the adoption of the GDPR. *Source:* Ganglmair (R) Krämer (R) Gambato (2024) (for the privacy policy panel), own calculations.

average effect), but it exhibits considerable variation. For instance, in the northeastern state of Mecklenburg-Western Pomerania, the negative effect of the GDPR on readability is almost twice as strong as in the southwest state of Rhineland-Palatinate.

In Ganglmair (R) Krämer (R) Gambato (2024), we link these state-level differences to the budget situation of the states’ data protection authorities. We find that firms in states with data protection authorities with higher budgets and more staff show better readability compliance than those with data protection authorities under more binding budget constraints. We interpret these results as evidence of firms responding to the readability requirement of the GDPR (and ultimately evidence of the conditional effectiveness of that requirement). Readability is a vague concept and difficult to enforce; constrained regulators may simply neglect the readability requirement and focus on “easier” aspects of the GDPR. Rational firms, of course, anticipate this enforcement focus and respond with little to no compliance. On the other hand, regulators with a higher budget can afford to pay attention to what is

difficult to enforce, in return, achieving better compliance by firms.

## 6 Concluding Remarks

In this paper, we explore the interaction between regulation activity and compliance with ambiguous and, therefore, difficult-to-enforce rules in the context of data regulation. We show that, while the average effect of the GDPR (and its transparency principle) on the readability of privacy policies in Germany is weak and strongly relies on how we measure readability, the areas in which it was effective in changing the drafting practices of firms indicate that there is a strategic dimension to regulatory compliance. Firms in industries more likely to see stringent regulation, and those in states with better-funded regulators respond more to the GDPR, displaying a stronger degree of readability compliance. This suggests that regulation is effective at inducing compliance as long as firms can reasonably expect regulators to enforce the rules actively.

Our findings have several broader implications. First, the result reinforces the notion that ambiguous regulatory rules are inherently costly because only well-funded regulators can credibly commit to enforcing them. Second, our results speak to the effectiveness (or lack thereof) of regulatory tools based on difficult-to-verify information. This is relevant for the European Union, where recent legislation uses language similar to that of the GDPR defining its transparency standards, and in the United States, where we see state-level privacy laws mushrooming that include provisions targeting readability.

## References

**Amos, Ryan, Gunes Acar, Eli Lucherini, Mihir Kshirsagar, Arvind Narayanan, and Jonathan Mayer.** 2021. “Privacy Policies over Time: Curation and Analysis of a Million-Document Dataset.” In *Proceedings of The Web Conference 2021*, WWW ’21 22, Association for Computing Machinery.

- Amstad, Toni.** 1978. *Wie verständlich sind unsere Zeitungen?*. Zurich, Switzerland: Studenten-Schreib-Service.
- Anger, Heike, and Dietmar Neuerer.** 2020. “Datenschutz-Verstöße: Zahl der Bußgelder ist drastisch gestiegen.” *Handelsblatt*, January 1, 2020, available at <https://www.handelsblatt.com/politik/deutschland/dsgvo-datenschutz-verstoesse-zahl-der-bussgelder-ist-drastisch-gestiegen/25364576.html> (accessed August 8, 2024).
- Armstrong, Mark, and David E. M. Sappington.** 2006. “Regulation, Competition, and Liberalization.” *Journal of Economic Literature* 44 (2): 325–366.
- Art. 29 Working Party.** 2018. “Article 29 Working Party: Guidelines on Transparency Under Regulation 2016/679 (wp260rev.01).” The Working Party on the Protection of Individuals with Regard to the Processing of Personal Data, available at <https://ec.europa.eu/newsroom/article29/items/622227> (accessed August 15, 2024).
- Bamberger, Richard, and Erich Vanecek.** 1984. *Lesen-Verstehen-Lernen-Schreiben*. Vienna: Jugend und Volk.
- Becher, Shmuel, and Uri Benoliel.** 2021. “Law in Books and Law in Action: The Readability of Privacy Policies and GDPR.” In *Consumer Law and Economics*, edited by Mathis, Klaus, and Avishalom Tor 179–204, Cham, Switzerland: Springer.
- Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, and Akitaka Matsuo.** 2018. “quanteda: An R Package for the Quantitative Analysis of Textual Data.” *Journal of Open Source Software* 3 (30): 774–777.
- Bersch, Johannes, Sandra Gottschalk, Bettina Müller, and Michaela Niefert.** 2014. “The Mannheim Enterprise Panel (MUP) and Firm Statistics for Germany.” ZEW Discussion Paper 14-104, ZEW – Leibniz Centre for European Economic Research, Mannheim, Germany.
- Chamberlain, Johanna, and Jane Reichel.** 2020. “The Relationship Between Damages and Administrative Fines in the EU General Data Protection Regulation.” *Mississippi Law Journal* 89 (4): 667–696.
- Degeling, Martin, Christine Utz, Christopher Lentzsch, Henry Hosseini, Florian Schaub, and Thorsten Holz.** 2019. “We Value Your Privacy . . . Now Take Some Cookies: Measuring the GDPR’s Impact on Web Privacy.” *Proceedings of the 26th Network and Distributed System Security Symposium (NDSS’19)*.
- Dwoskin, Elizabeth.** 2015. “Privacy Policies More Readable, But Still Hard to Understand.” *The Wall Street Journal*, December 30, 2015, available at <https://www.wsj.com/>

[articles/BL-DGB-44469](#) (accessed August 15, 2024).

- European Commission.** 2012. “Proposal for a Regulation on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation).” Document COM/2012/010 final – 2012/0010 (COD), European Commission.
- European Parliament.** 2013. “Report on the Proposal for a Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation).” available at [https://www.europarl.europa.eu/doceo/document/A-7-2013-0402\\_EN.pdf](https://www.europarl.europa.eu/doceo/document/A-7-2013-0402_EN.pdf) (accessed August 13, 2024).
- Flesch, Rudolf.** 1948. “A New Readability Yardstick.” *Journal of Applied Psychology* 32 (3): 221–233.
- Frankenreiter, Jens.** 2022. “Cost-Based California Effects.” *Yale Journal on Regulation* 39 (3): 1155–1217.
- Ganglmair, Bernhard** (r) **Julia Krämer** (r) **Jacopo Gambato.** 2024. “Regulatory Compliance with Limited Enforceability: Evidence from Privacy Policies.” CRC Discussion Paper 547, Collaborative Research Center Transregio 224, Bonn and Mannheim.
- Giommoni, Tommaso, Luigi Guiso, Claudio Michelacci, and Massimo Morelli.** 2023. “The Economic Costs of Ambiguous Laws.” unpublished manuscript, Einaudi Institute for Economics and Finance.
- Goldberg, Samuel G., Garrett A. Johnson, and Scott K. Shriver.** 2024. “Regulating Privacy Online: An Economic Evaluation of the GDPR.” *American Economic Journal: Economic Policy* 16 (1): 325–58.
- Gustafsson, Marita.** 1984. “The Syntactic Features of Binomial Expressions in Legal English.” *Text: An Interdisciplinary Journal for the Study of Discourse* 4 (1-3): 123–142.
- Hern, Alex.** 2015. “I Read All the Small Print on the Internet and it Made Me Want to Die.” *The Guardian*, June 15, 2015, available at <https://www.theguardian.com/technology/2015/jun/15/i-read-all-the-small-print-on-the-internet> (accessed August 15, 2024).
- Immel, Karl-Albrecht.** 2014. *Regionalnachrichten im Hörfunk: Verständlich schreiben für Radiohörer*. Wiesbaden: Springer Fachmedien Wiesbaden.
- Johnson, Garrett.** 2024. “Economic Research on Privacy Regulation: Lessons from the GDPR and Beyond.” In *The Economics of Privacy*, edited by Goldfarb, Avi, and Catherine Tucker, Chicago, Ill.: University of Chicago Press.

- Johnson, Garrett, Scott Shriver, and Samuel Goldberg.** 2023. “Privacy and Market Concentration: Intended and Unintended Consequences of the GDPR.” *Management Science* 69 (10): 5695–5721.
- Koski, Heli, and Nelli Valmari.** 2020. “Short-Term Impacts of the GDPR on Firm Performance.” ETLA Working Papers 77, ETLA Economic Research.
- Koutroumpis, Pantelis, Farshad Ravasan, and Taheya Tarannum.** 2022. “Under Investment in Cyber Skills and Data Protection Enforcement: Evidence from Activity Logs of the UK Information Commissioner’s Office.” unpublished manuscript, available at <https://ssrn.com/abstract=4179601>.
- Laffont, Jean-Jacques.** 2005. *Regulation and Development*. Cambridge, UK: Cambridge University Press.
- Linden, Thomas, Rishabh Khandelwal, Hamza Harkous, and Kassem Fawaz.** 2020. “The Privacy Policy Landscape After the GDPR.” *Proceedings on Privacy Enhancing Technologies* 2020 (1): 47–64.
- Litman-Navarro, Kevin.** 2019. “We Read 150 Privacy Policies. They Were an Incomprehensible Disaster.” *The New York Times*, June 12, 2019, available at <https://www.nytimes.com/interactive/2019/06/12/opinion/facebook-google-privacy-policies.html> (accessed August 15, 2024).
- Milne, George R., Mary J. Culnan, and Henry Greene.** 2006. “A Longitudinal Assessment of Online Privacy Notice Readability.” *Journal of Public Policy & Marketing* 25 (2): 238–249.
- Naughton, John.** 2020. “Data Protection Laws are Great. Shame They are not Being Enforced.” *The Observer*, May 2, 2020, available at <https://www.theguardian.com/commentisfree/2020/may/02/data-protection-laws-are-great-shame-they-are-not-being-enforced> (accessed August 8, 2024).
- Nissenbaum, Helen.** 2011. “A Contextual Approach to Privacy Online.” *Daedalus (Journal of the American Academy of Arts & Sciences)* 140 (4): 32–48.
- Peukert, Christian, Stefan Bechtold, Michail Batikas, and Tobias Kretschmer.** 2022. “Regulatory Spillovers and Data Governance: Evidence from the GDPR.” *Marketing Science* 41 (4): 318–340.
- Rauh, Christian.** 2023. “Clear Messages to the European Public? The Language of European Commission Press Releases 1985–2020.” *Journal of European Integration* 45 (4): 683–701.

- Schütz, Philip.** 2018. “Zum Leben zu wenig, zum Sterben zu viel? Die finanzielle und personelle Ausstattung deutscher Datenschutzbehörden im Vergleich.” In *Die Fortentwicklung des Datenschutzes: Zwischen Systemgestaltung und Selbstregulierung*, edited by Roßnagel, Alexander, Michael Friedewald, and Marit Hansen 251–268, Wiesbaden: Springer Vieweg.
- Statistisches Bundesamt (Destatis).** 2024. “12411-0010: Bevölkerung: Bundesländer, Stichtag (1958–2023) [Dataset].” Federal Statistical Office, Germany, data table available at <https://www-genesis.destatis.de/genesis//online?operation=table&code=12411-0010> (last accessed: August 6, 2024).
- Stern, Jon.** 2000. “Electricity and Telecommunications Regulatory Institutions in Small and Developing Countries.” *Utilities Policy* 9 (3): 131–157.
- Vinocur, Nicholas.** 2019. “How One Country Blocks the World on Data Privacy.” *Politico*, April 29, 2019, available at <https://www.politico.com/story/2019/04/24/ireland-data-privacy-1270123> (accessed August 8, 2024).
- Wagner, Isabel.** 2023. “Privacy Policies Across the Ages: Content of Privacy Policies 1996–2021.” *ACM Transactions on Privacy and Security* 26 (3): 1–32.
- Yuan, Bocong, and Jiannan Li.** 2019. “The Policy Effect of the General Data Protection Regulation (GDPR) on the Digital Public Health Sector in the European Union: An Empirical Investigation.” *International Journal of Environmental Research and Public Health* 16 (1070): 1–15.