The Court Speaks, But Who Listens? Automated Compliance Review of the GDPR

Amit Zac, Pablo Wey, Stefan Bechtold,
David Rodriguez & Jose M. Del Alamo[1]

March 27, 2024

Abstract

In July 2020, the European Court of Justice invalidated the EU-US Privacy Shield with immediate effect ("Schrems II"). As a result, almost all personal data transfers from the European Union to the United States became illegal overnight. We present a unique dataset allowing us not only to observe what firms say about their behavior in privacy policies, but also how firms actually behave. Using machine-learning tools, we analyze the privacy policies of over 7,500 apps on the Spanish Google Play Store and find limited compliance with the Schrems II decision. We validate the quality of our classifier through manual inspection of privacy policies. Using tools from IT security research, we are able to observe the actual personal data traffic flows leaving apps towards the United States after Schrems II. Combining our observations on privacy policies and data traffic flows, our findings on compliance with Schrems II are sobering. A few weeks after Schrems II was decided, only 23% of the studied apps in our sample seem to comply with the decision while 77% seem to violate the GDPR. Over two years after Schrems II, the rate of compliant apps increases, yet we estimate that roughly 45% of the apps are non-compliant. We examine the implications our findings have for the design and enforcement of the GDPR, and on the notion of diffusion of EU laws ('Brussle effect') and discuss how the combination of an automated analysis of contracts and of actual data traffic flows can improve our understanding on how to regulate the digital economy at large scale.

[1] Corresponding author Zac: a.zac@uva.nl; University of Amsterdam and Center for Law & Economics, ETH Zurich; Wey & Bechtold: Center for Law & Economics, ETH Zurich; Rodriguez & Del Alamo: Universidad Politécnica de Madrid. We would like to thank participants of the Conference on Empirical Legal Studies 2023 and the University of Notre Dame Conference on Empirical, Behavioral, and Experimental Analyses of Law 2024 for very helpful feedback. Elias Landes and Simona Ramsperger provided excellent research assistance. This project was supported by the Swiss National Science Foundation via a postdoctoral fellowship of Amit Zac.

Table of Contents

1. Introduction

When the European Union enacted the General Data Protection Regulation (GDPR) in 2018, the press greeted the European Union (EU) as the "world's leading tech watchdog" (Satarino 2018). Companies were quick to point out the staggering investments that would be needed to comply with the GDPR (Smith 2018). Scholars analyzed how the European Union used its privacy law to become a leader in political agenda setting around the world (Bradford 2012, 2020).

Six years after the GDPR has been enacted, a more nuanced picture of how the GDPR impacts companies, the European Union, and global trade is emerging. Internet users have become used to lengthy privacy consent notices that pop up before they can use a website, and scholars have questioned whether these pop-ups are an effective way to protect users' privacy online (Jack et al., 2019; Warberg et al., 2023). Privacy observers have criticized the GDPR's complicated enforcement arrangement which distributes enforcement of the GDPR among national data protection authorities with a cumbersome dispute resolution system among those authorities. And empirical scholars have begun to notice that although data protection agencies may impose multi-billion Euro fines under the GDPR, compliance with the GDPR seems often low.

One of the hallmark protections that the GDPR provides to EU consumers is that it does not only protect their personal data within the European Union, but also puts in safeguards when their data leaves the EU. This is particularly relevant in the digital economy where many services and products are offered by US firms, which want to transfer the personal data of EU consumers to their US-based servers for data processing and analyzing purposes. The governance of personal data transfers from the EU to the US has a long history. In 2016, the European Union and the United States adopted the EU-US Privacy Shield, which provided a framework under which companies could safely – and legally – transfer personal data from the EU to the US. The Privacy Shield became an immediate success, with over 5,000 companies registering under the framework (Taylor 2023: 23). But it did not last long. On July 16, 2020, the European Court of Justice declared the Privacy Shield invalid (European Court of Justice 2020) as it did not provide EU citizens with adequate guarantees that their personal data would be processed in the United States with a level of privacy protection that equalled the European Union.

The Schrems II decision of the European Court of Justice came as a shock to the business world. As the decision did not provide any grace period, cross-border transfers of personal data from the EU to the US became effectively illegal overnight. The decision therefore provides an excellent opportunity to analyze the extent to which companies whose data traffic flows were affected by this decision actually reacted to the decision in an attempt to become GDPR-compliant.

In this study, we analyze how companies, which transferred personal data from the EU to the US before the Schrems II decision of the European Court of Justice, responded to this decision. As thousands of companies were affected by this decision, manually observing how these companies reacted is a daunting task. In order to scale up our analysis, we aim at integrating analytical approaches from machine learning and computer science in order to present a framework on how compliance with digital regulations can be observed and assessed at large scale.

While many existing GDPR-compliance studies focus on websites, this study focuses on smartphone apps which have emerged as a cornerstone of the digital economy. We focus on over

7,500 apps on the Spanish Google Play Store. We were able to retrieve the privacy policies these apps used before the European Court of Justice decided the Schrems II case. Using machine-learning tools, we determine whether these privacy policies state that the apps transfer personal data to the US. We then analyze these clauses to determine the legal basis they rely on to justify the transfer under the GDPR. This classification allows us to identify those apps whose "cross-border transfer clauses" violate the GDPR according to the Schrems II decision, and track whether the apps change their privacy policies in order to become GDPR-compliant after the decision.

Our study goes beyond the mere analysis of privacy policies. Using approaches from IT security research, we are able to observe the actual personal data that our apps send to servers in the US after Schrems II. Intercepting the communication between those apps and US servers allows us to analyze whether the apps transfer information about the user's identity, location, demographics or contacts to a server in the US. Combining our observations on privacy policies and data traffic flows, our findings on compliance with Schrems II are sobering. A few weeks after Schrems II was decided, only 23% of the studied apps in our sample seem to comply with Schrems II, while 77% seem to violate the GDPR. About two years after Schrems II, the level of non-complying apps is still at 45%.

As we discuss below, such numbers have to be taken with a grain of salt. Our ambition in this study is to resort to state-of-the-art techniques from machine learning and IT security research, integrate these techniques into a legal research pipeline and to assess the challenges that emerge along this road. We see the main contribution of this paper not in identifying a particular rate of non-compliance with the GDPR after the European Court of Justice decided Schrems II. Rather, we want to propose a research framework on how compliance with the GDPR and other digital regulations may be assessed at large scale in the future. In addition, our findings modify the current 'carrot and stick' theory of diffusion of EU laws globally (Bradford 2012). It is not the actual 'capacity' of EU institutions, as argued, which encourages diffusion, but the 'appetite' to go after national and global giants.

While legal scholars have become interested in empirically analyzing privacy policies over the last years, we note that such analysis can only provide an indirect proxy for compliance studies. There is a difference between whether a company states in its privacy policy that it complies with the GDPR, and whether the company actually complies with the GDPR in its data traffic flows. Our unique dataset allows us to combine both aspects. We thereby contribute to an interdisciplinary literature that combines observations on privacy policies with observations on actual firm behavior. To the best of our knowledge, this literature is still small, and several co-authors of this study have contributed to this literature before (Guáman et al. 2021, 2023; see also Peukert et al. 2022).

This paper proceeds as follows. Section 2 provides an overview of how the GDPR deals with cross-border personal data transfer to the United States, and how the Schrems II decision by the European Court of Justice interfered with typical data flows. It also discusses related literature from law and computer science. Section 3 presents the research design of our study, discusses the natural-language processing required to analyze the privacy policies of our apps, and the man-in-the-middle attacks and other techniques we use to observe the data traffic leaving our apps. We also discuss how much confidence legal researchers should have in the available natural-language processing methods to classify cross-border transfer clauses in privacy law.

Section 4 presents our results. We analyze whether the privacy policies of apps which are affected by the Schrems II decision change a few weeks and over two years after the decision. We repeat such analysis with the personal data flows of these apps. Combining our observations from privacy policies and personal data flows, we present our findings on the extent to which apps in the Spanish Google Play Store comply with the European Court of Justice's Schrems II decision. Section 5 discusses limitations of our methodological approaches. Section 6 concludes.

2. Legal Background and Related Literature

2.1 Legal Background

European privacy law and politics have grappled with personal data transfer between the EU and the US for over a quarter of a century. Before the General Data Protection Regulation became effective in 2018, European privacy law was harmonized through the EU Data Protection Directive 1995 (European Union 1995). Under that framework, the European Commission adopted the Safe Harbour agreement (European Commission 2000), which provided a set of principles governing the personal data exchange between the European Union and the United States. In 2015, the European Court of Justice declared the Safe Harbor Agreement invalid (European Court of Justice 2015). After the EU and the US had agreed on a new "Privacy Shield," the European Commission determined in July 2016 – still under the old Data Protection Directive – that US law provided an adequate level of privacy protection (European Commission 2016a).

When the General Data Protection Regulation came into effect in May 2018, it revised the structure of the legal grounds for cross-border data transfers. The GDPR allows companies to transfer personal data out of the European Union – or, more precisely, the European Economic Area – under certain conditions. In particular, the data exporter is responsible for ensuring that data transfers comply with one of the legal grounds outlined in Art. 45-49 GDPR (see Art. 44 GDPR): First, the European Commission can decide that a third country provides an adequate level of privacy protection. Following such adequacy decision, transferring personal data to this country is allowed (Art. 45 GDPR). Second, the data controller or processor[2] can transfer personal data to a third country if it uses "Standard Contractual Clauses" that are adopted or approved by the European Commission and incorporated by the controllers or processors into their contractual arrangements with the party in the third country to ensure appropriate protection (Art. 46(2)(c), (d) GDPR). Third, the data transfer may be legal under "Binding Corporate Rules" that a data protection authority approves for the group-internal data transfer within an enterprise (Art. 47 GDPR). Fourth, the data transfer can be legal in specific situations if the consumer has provided explicit consent or if such transfer is necessary for the performance of a contract between the consumer and a data controller (Art. 49 (1)(a), (b) GDPR).[3] With regard to countries covered by an adequacy decision of the European Commission, the data controller does not have to provide any further safeguards, as transfers to these countries are deemed fully equivalent to

[2] Under the GDPR, the data controller is the institution deciding why and how personal data should be, while the – potentially separate – data processor actually processes the data, see Art. 4 (7), (8) GDPR.

[3] Art. 45-49 provide some other legal grounds for personal data transfers out of the European Union, such as approved code of conducts and approved certification mechanisms, see Art. 46(2)(e), (f) GDPR. They are less relevant for our study.

EU-internal transfers. Therefore, from a company's perspective, countries covered by an adequacy decision – such as the US with the Privacy Shield decision by the European Commission – are the most attractive candidates for cross-border data transfers, as the legal risks associated with such transfer are low.

In our study, we are interested in determining the legal basis which firms use to justify their personal data transfer from the EU to the US. Importantly, firms are required to declare the legal basis for their third-country data transfers. With regards to the Privacy Shield, the Shield Framework itself required firms to declare their reliance on the Privacy Shield in their privacy policies.[4] With regards to Standard Contractual Clauses and Binding Corporate Rules, transparency obligations require data controllers to inform their users about the relevant legal basis and the safeguards taken (Art. 13(1)(c), (f), 14(1)(c), (f) GDPR). And European privacy agencies noted in 2018 that users need to be informed about the particular country to which personal data is transferred (Article 29 Data Protection Working Group 2018: 38).

In 2013, Max Schrems – an Austrian privacy activist who has emerged as an important driver in the enforcement of European privacy law – complained with the Irish Data Protection Commissioner about Facebook's transfer of his personal data from the European Union to the United States. After a complicated, decade-long procedural history, the European Court of Justice finally decided the so-called "Schrems II" case on July 16, 2020 (European Court of Justice 2020). The court held that Facebook's transfers of personal data were neither legal under Standard Contractual Clauses – which Facebook had used, based on an adoption of such clauses by the European Commission (2016b) – nor under the EU-US Privacy Shield.

As far as the Privacy Shield was concerned, the court pointed to the surveillance possibilities under US national security laws and noted that US law does not afford EU citizens a level of protection that is essentially equivalent to the protections provided under EU law. The court therefore declared the European Commission's adequacy decision with respect to the EU-US Privacy Shield invalid (European Court of Justice 2020: ¶ 201).

As far as Standard Contractual Clauses were concerned, while the court did not invalidate them as such, it stressed that it is Facebook's responsibility to determine, on a case-by-case basis, whether US law provides EU customers with a level of privacy protection that is essentially equivalent to EU law. In particular, the court noted that, in the absence of a valid adequacy decision by the European Commission, transfers to the US are only permissible if Facebook provides appropriate safeguards and consumers have enforceable rights and effective legal remedies (Art. 46 GDPR; European Court of Justice 2020: ¶¶ 91, 103). If Facebook was unable to put additional effective safeguards in place to ensure such protection, relying on Standard Contractual Clauses for the personal data transfer from the EU to the US was not sufficient according to the court (European Court of Justice 2020: ¶¶ 134, 135). Taken together, this meant that Facebook's transfer of personal data became illegal under the GDPR. Importantly, the court's

[4] US Dept. of Commerce (2016), § 6(d) ("All organizations that are placed on the Privacy Shield List by the Department must also state in their relevant published privacy policy statements that they adhere to the Privacy Shield Principles. If available online, an organization's privacy policy must include a hyperlink to the Department's Privacy Shield website and a hyperlink to the website or complaint submission form of the independent recourse mechanism that is available to investigate unresolved complaints") and § 6(f) ("An organization that withdraws from the Privacy Shield must remove from any relevant privacy policy any references to the Privacy Shield that imply that the organization continues to actively participate in the Privacy Shield and is entitled to its benefits").

decision became effective immediately, without any grace period that firms could use to adapt their data flows.[5]

The Schrems II decision created an unexpected legal turmoil with regard to personal data transfers between the EU and the US. While data transfers based on the Privacy Shield had become clearly illegal, the impact of the decision on data transfers based on Standard Contractual Clauses seemed more subtle. One day after the decision, the European Data Protection Board issued a statement stressing that firms may have to suspend personal data transfers to the US that are based on Standard Contractual Clauses and that the board will look into what additional measures Facebook and other firms could implement to make data transfers based on Standard Contractual Clauses legal again (European Data Protection Board 2020a). Eight days after the decision, the Board released a "Frequently Asked Questions" list on the decision. It pointed out that personal data transfers under the Privacy Shield have become illegal, and that transfers under Standard Contractual Clauses would only be legal if the data exporter would engage in a case-by-case analysis and then take supplementary measures ensuring that US law did not impinge on EU citizens' privacy rights. If the data exporter was unable to ensure such a level of protection, it was "required to suspend or end the transfer of personal data" (European Data Protection Board 2020b: 3). The board did not provide specific guidance on how such supplementary measures could look like ("The EDPB is looking further into what these supplementary measures could consist of and will provide more guidance," European Data Protection Board 2020b: 5).

As a result, over 5,000 companies that were registered to use the Privacy Shield (Taylor 2023: 203) faced considerable uncertainty on whether and how their personal data transfer from the EU to the US could be saved from illegality. In the United States, news commentators at that time warned that the decision "could destroy transatlantic commerce" (Stout 2020). US Secretary of Commerce Wilbur Ross stated on the day of the decision that the Department of Commerce was "deeply disappointed" with the decision, pointing to the negative consequences for the "$7.1 trillion transatlantic economic relationship" (Ross 2020; see Taylor 2023: 203-204). Stewart Baker, former Assistant Secretary for Policy at the US Department of Homeland Security, found the court's decision "gobsmacking in its mix of judicial imperialism and Eurocentric hypocrisy" and proposed to prepare sanctions against Europe (Baker 2020). Government officials on the other side of the Atlantic added their own oil to the fire. The data protection officer of the German state of Berlin, for example, advised companies under her jurisdiction to move all personal data stored in the US to Europe. She stated in a press release: "The time when personal data could be transferred to the US for convenience or cost savings is over. Now the hour for Europe's digital independence has come" (Commissioner for Data Protection and Freedom of

[5] While the court noted that, in the absence of Standard Contractual Clauses and the Privacy Shield, Article 49 of the GDPR also allows cross-border data transfers in specific situations (European Court of Justice 2020: ¶ 202), in most cases this provision would not save data transfer practices in cases that are of interest to us. While Art. 49(1)(a) GDPR allows cross-border data transfers if the consumer has consented, such consent has to be explicit, informed, specific for a particular data transfer, and can only occur on an occasional basis (European Data Protection Board 2020b: 4; see also the later decision by the Irish Data Protection Commission (2023) in the Meta case. While some of the privacy policies in our dataset mention "consent" in some capacity, we therefore do not treat this as a potential legal basis for our apps' transfer of personal data from the European Union to the United States. The same applies to contractual necessity as defined in Art. 49(1)(b) GDPR.

Information for the German State of Berlin 2020). The legal trade press noted that there was no easy solution available on how to overcome Schrems II in the short term.[6]

This uncertainty about how to structure personal data transfers from the EU to the US in a GDPR-compliant way continued in the months following the Schrems II decision. In November 2020 – about four months after the Schrems II decision – the European Data Protection Board issued recommendations on supplementary measures firms could apply when using Standard Contractual Clauses for their cross-border personal data transfers (European Data Protection Board 2020c). These recommendations were not specific to the United States, though. They provided general guidance and noted that the ultimate responsibility for the legality of personal data transfers resides with the data exporter.[7] Following a recommendation by the European Data Protection Board and the European Data Protection Supervisor in January 2021, the European Commission adopted new Standard Contractual Clauses in June 2021 (European Commission 2021). In its decision, the European Commission stressed that the data exporter should suspend the personal cross-border data transfer "if it considers that no appropriate safeguards [such as technical or organizational measures to ensure security and confidentiality] can be ensured" (European Commission 2021: 34, Recital 21).

In March 2022, the United States government and the European Commission announced their plans for a new "Transatlantic Data Privacy Framework" which would replace the defunct EU-US Privacy Shield (White House 2022), and the European Commission published a draft adequacy decision that was supposed to replace the EU-US Privacy Shield on December 12, 2022 (European Commission 2022). On July 10, 2023, the European Commission adopted the final adequacy decision, paving the way for firms to transfer personal data based on the new EU-US Data Privacy Framework (European Commission 2023).

Overall, transferring personal data from the EU to the US between July 16, 2020, and July 10, 2023, was at least highly risky, as the European Court of Justice had either invalidated or severely restricted the main legal bases for such transfer (the Privacy Shield and Standard Contractual Clauses). In its 2023 annual financial report to the US Securities and Exchange Commission, Meta mentioned the risk that it might have to pull the plug for Facebook and Instagram in Europe if no solution to this impasse was found (Meta 2023: 10). It took the European Union three years to develop a follow-up agreement to the Privacy Shield, and it is only a matter of time until the European Court of Justice will have to decide on the legality of the EU-US Data Privacy Framework.

Violating the GDPR can be a costly enterprise. Infringing GDPR requirements on personal data transfers outside the EU can lead to administrative fines of up to 4% of a company's global turnover of the preceding financial year (Art. 83(5) GDPR). Recently, European data protection authorities have become more proactive in enforcing these provisions. Following a dispute

[6] Harnett et al. (2023): "Due to the continued and wide-reaching effects of the U.S's strategy on surveillance we have now entered a period of limbo, as there is no obvious methodology currently in place for cross-border transfers of EU data to the US;" Murphy (2022: 261): "[t]here is, at this point, no generalisable solution that will remedy the EU-US data transfer challenge;" see also Bradford (2023: 232).

[7] "You may ultimately find that no supplementary measure can ensure an essentially equivalent level of protection for your specific transfer. In those cases where no supplementary measure is suitable, you must avoid, suspend or terminate the transfer to avoid compromising the level of protection of the personal data," European Data Protection Board (2020c): 2.

resolution decision by the European Data Protection Board (2023), the Irish Data Protection Commission issued Meta with a fine of 1.2 billion € in May 2023 (Irish Data Protection Commission 2023). This decision was based on Meta's transfers of personal data to the US since July 16, 2020 (the day when Schrems II was decided). Meta had initially based the transfer of personal data on Standard Contractual Clauses which the European Commission had adopted – still under the auspices of the now defunct EU Data Protection Directive of 1995 – in 2010 (European Commission 2010). The Irish Data Protection Commissioner decided that Meta's transfer practices were not only illegal under the 2010 Standard Contractual Clauses, but also under the new Standard Contractual Clauses which the European Commission had adopted in 2021, as they also failed to compensate for deficiencies in US law (Irish Data Protection Commission 2023: 2, 94, 100). The Commission held that Meta could not rely on any of the derogations provided under Article 49 (1) GDPR, in particular explicit consent and contractual necessity. As a result, Meta's personal data transfer became illegal on July 16, 2020 (Irish Data Protection Commission 2023: 5, 124, 212). The Commission ordered Meta to suspend its data transfers to the United States and imposed an administrative fine of €1.2 billion on Meta (where Meta's global revenue in 2022 was $116.61 billion, Irish Data Protection Commission 2023: 211-212). So far, this has been the largest fine ever imposed under the GDPR.

As this overview shows, transferring personal data from the EU to the US is only possible if (1) there is a valid adequacy decision by the European Commission, (2) a company uses Standard Contractual Clauses that are complemented, on a case-by-case basis, by appropriate supplementary measures, or if (3) any of the other legal grounds the GDPR provides for international data transfers applies. In the period between July 16, 2020, and July 10, 2023, firms faced significant hurdles when trying to design their personal data flows from the EU to the US in a GDPR-compliant manner: The European Court of Justice invalidated the most important ground for personal data transfers to the US (the Privacy Shield). While Standard Contractual Clauses remained a potential legal basis for such transfers, the legal hurdles to become compliant on that basis were very high, and neither the court nor privacy agencies provided clear and easily implementable guidance on how to become fully compliant. Companies were therefore left alone and had to conduct their own risk assessment on the reach of US surveillance laws with regard to their data transfers. As the recent 1.2 € billion fine by the Irish Data Protection Commission against Meta demonstrates, all this could translate into a hefty bill for a company. In the aftermath of the Schrems II decision, we would therefore expect that companies put in considerable efforts to make their international personal data flows GDPR-compliant again.

2.2 Literature

Our study builds upon various strands of literature from law, economics, and computer science. On a policy level, Bradford (2012, 2020) has prominently argued that the European Union has de facto expanded some of its regulatory laws – including privacy laws – beyond its borders through a combination of market mechanisms and unilateral regulatory globalization. Peukert et al. (2022), Frankenreiter (2022) and Davis & Marotta-Wurgler (2023) investigate this effect empirically, using third-party domain requests and privacy policies of websites, respectively.

Observing firm behavior through privacy policies is an important component in analyzing compliance with privacy laws. Indeed, such data has enabled a rich literature in empirical legal scholarship that focuses on the emergence and development of privacy policies (see only

Marotta-Wurgler 2016; Davis & Marotta-Wurgler 2019; Marotta-Wurgler & Svirsky 2023; Frankenreiter 2022). But this scholarship has an important limitation. When observing firms' privacy policies, we can only learn what firms *say* about their behavior in legal documents, not how firms actually *behave*. In many industries, observing firm behavior from the outside is either impossible or very costly. But the digital economy is different. On the Internet, firm behavior can often be observed in a very detailed manner. As a result, computer scientists have explored violations of the GDPR by scraping publicly available websites. One strand of this literature has, e.g., investigated whether websites' cookie settings adhere to the GDPR. Many of these studies point to significant levels of non-compliance with the GDPR (for broader reviews, see Lancieri 2022: 65-72; Li et al. 2023). Nouwens et al. (2020), e.g., scraped the designs of the five most popular consent management platforms on the top 10,000 websites in the UK, finding that only 11.8% of the websites studied met minimal legal requirements of the GDPR for explicit consent to use such tracking technologies. Bollinger et al. (2022) indicate that an alarming 95% of 30,000 websites implementing one of three popular consent management platforms may violate European privacy law. Matte et al. (2020) scraped over 20,000 European websites and found that the cookie banners of 54% of these websites violated the GDPR or the European Union's e-Privacy Directive from 2002. Using historic data from the third-party domain requests of over 100,000 websites, Peukert et al. (2022) analyze how the GDPR has influenced market concentration in web technology markets, and how it has affected global regulatory competition. Moving beyond websites, an interdisciplinary group of computer scientists and lawyers has investigated whether websites adhere to privacy preferences with regard to newsletter registrations (Kubicek et al. 2022). And Senol et al. (2022) have shown how website registration forms collect personal information about users even before they submit the form.

Empirical research on GDPR compliance is often conducted either by analyzing privacy policies or by analyzing websites. Both approaches have the advantage that this data is relatively easily accessible. However, IT security research has gone beyond such data sources and started to investigate whether the inner operations of a program – or an app on a smartphone – can be accessed and analyzed in order to assess the program's compliance with the GDPR. Guamán et al. (2021), e.g., present a framework to assess the compliance of Android mobile apps with cross-border transfer rules established by the GDPR. They check both privacy policy disclosures and detect actual cross-border transfers in those apps at runtime. This enables them to check whether an app's actual data flows comply with the app's privacy policy. In Guamán et al. (2023), they refine their methodology and expand it to an automated method for assessing the compliance of Android apps with GDPR requirements for cross-border personal data transfers.

Our study contributes to the literature on various dimensions. First, it builds upon the computer science publications by Guamán et al. (2021, 2023) to present a framework which enables us to analyze the impact of a court decision in privacy law not only on the privacy policies of firms, but also on the actual personal data flows generated in the firms' smartphone apps. Our study goes beyond the methods presented in Guamán et al. (2021, 2023) on various dimensions. While we use their tools to analyze privacy policies and app data flows (see Sections 3.2 and 3.3), we present a novel research design. We also collect and analyze more privacy policies and app data at additional points in time. And we engage in a discussion on the validity and impact of such tools for privacy research and enforcement.

Our study contributes to the legal debate on privacy compliance. To the best of our knowledge, this is the first study directed towards a legal research community which combines observations

from privacy policies and actual data transfer practices to study compliance with privacy laws. It is also one of the few studies in privacy law scholarship that empirically analyzes the compliance of smartphone apps, as opposed to websites. We validate the machine-learning classifiers that we use for our privacy policies through human coding. And we discuss at length the extent to which the automated tools that we use in the study offer the validity and confidence that is desirable when applying such tools in the law. Finally, we analyze the implications of our proposed framework for privacy scholarship and the regulation of the digital economy more broadly.

3. Methods and Data

3.1 Research Design

As explained in Section 2.1, the European Court of Justice held on July 16, 2020, that personal data transfers from the European Union to the United States that were solely based on Standard Contractual Clauses and/or on the EU-US Privacy Shield violated European Union law. The fact that the European Court of Justice declared such transfers illegal with immediate effect and no transition period was a strong signal to the privacy world. It created a unique opportunity for empirical researchers to analyze the effect of the decision.

In a world where violations of the law are observable and the potential fines are considerable, the law should have a significant deterrent effect (Becker 1968; but see Marotta-Wurgler & Svirsky 2023: 12-14). Complying with privacy laws in the digital economy seems a textbook example of rational choice deterrence theories. The GDPR provides for considerable fines, and data protection agencies have been willing to exercise their power in case of personal data transfers, as exemplified by the recent 1.3 billion € fine against Meta (see Section 2.1).

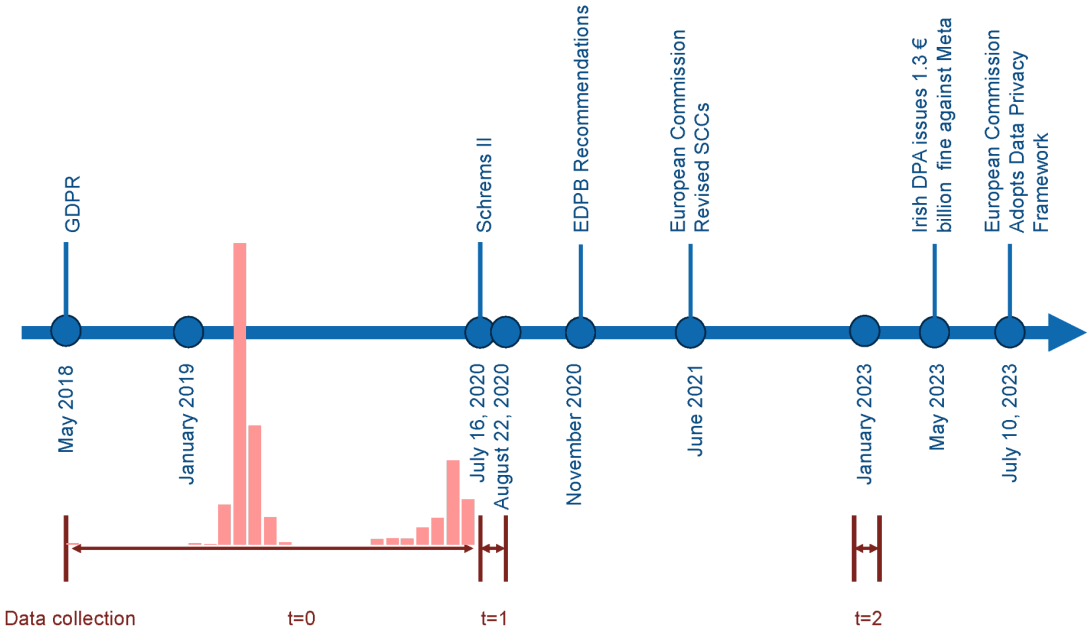
Our goal is to analyze the effects of the Schrems II decision on firm behavior (proxied by apps here). To do so, we inspect whether firms which are affected by Schrems II adapt the legal statements about their personal data transfers after the decision, and also whether these firms adapt their actual personal data transfer flows after the decision. Comparing the behavior of those firms with behavior of firms that are not affected by Schrems II allows us to approximate this to the ideal difference-in-difference research design on many dimensions.

With regard to firms' legal statements about their personal data transfers, we focus on their privacy policies. Based on data collection efforts of two of the co-authors (Guamán et al. 2021, 2023), we analyze the most popular mobile apps that are available in the Spanish Google Play store for smartphones operating under the Android operating systems. Android is the most successful smartphone operating system worldwide, currently serving over three billion devices. In order to compare the privacy policies of those apps before and after the Schrems II decision, we obtain such policies for three different points in time. Most importantly, we use the privacy policies of 7,634 apps, as obtained through a scrape (see Section 3.2) that occurred immediately after the Schrems II decision. The European Court of Justice decided Schrems II on July 16, 2020, and we obtained the privacy policies of these apps from the Spanish Google Play Store between July 20 and August 22, 2020. We refer to this period as t=1, which is the core of our analysis as it covers the time right after the judgment was delivered.

In order to check whether our apps changed their privacy policies right after Schrems II, we need access to their privacy policies before the court decided the case. The Wayback Machine archive (Internet Archive 2023) provides access to historic versions of web pages. We use the Wayback Machine to retrieve privacy policies of our apps in the period between 2018 (when the GDPR came into force) and July 15, 2020 (the day before Schrems II was decided). For 1,939 of our apps, we succeeded in retrieving a privacy policy from that period, which we refer to as period t=0.

We are not only interested in the short-term effect of Schrems II on privacy policies. It may have taken apps some time to adapt their privacy policies and data flows. Therefore, we use the scraping tools from Guamán et al. (2021, 2023) to retrieve privacy policies of our apps in January 2023. This period, which we refer to as t=2, is over two years after the Schrems II decision, which should have given firms ample time to react to Schrems II. We manage to access the privacy policies of 7,379 apps (out of the 7,634 in t=1). Figure 1 describes the timeline of the relevant legal events and of our data collection efforts.

Figure 1: Research Design Timeline



Note. The descriptions above the timeline depict legal events: On May 25, 2018, the GDPR came into effect. On July 7, 2020, the European Court of Justice decided Schrems II. On November 10, 2020, the European Data Protection Board issued its recommendation on how to deal with Schrems II (European Data Protection Board 2020). On January 14, 2021, the European Data Protection Board and the European Data Protection Supervisor issued their Implementing Decisions on Standard Contractual Clauses (European Data Protection Board & European Data Protection Supervisor 2021). The descriptions below the timeline depict our data collection efforts. The light red bars indicate the percentage of observations in t=0, to emphasize the actual dates of the privacy policies retrieved from the Wayback Machine. The overwhelming majority of policies are from after January 2019, and around 30% of policies are from 2020. The full figure is included in the Appendix as Figure A.2.

As explained in Section 2.2, observing firm behavior through privacy policies is important, but has its limits. When observing firms' privacy policies, we can only learn what firms *say* about their behavior in legal documents, not how firms actually *behave*. It may be very easy for an app

to change its privacy policy after Schrems II if it does not also change its actual data flow patterns. Given that it is much more challenging to observe app data flows, the app may not expect to get caught when it makes its privacy policy compliant with Schrems II, but continues to violate Schrems II through its data flow patterns. We therefore combine our analysis of privacy policies with observations on actual firm behavior. At t=1 (July/August 2020), we do not only have access to the privacy policies of 7,634 apps. We also analyze the extent to which these apps transfer personal data from the European Union to the United States. We use IT security methods, which two of the co-authors have developed in Guamán et al. (2021, 2023) and are described in Section 3.2, to observe such data flows while an app runs on a smartphone. We obtain the personal data flows from 10,080 apps at t=1 (July/August 2020).[8] As mentioned before, we are not only interested in the short-term effect of Schrems II on firm behavior. Therefore, as for privacy policies, we use the tools from Guamán et al. (2021, 2023) to observe personal data transfers of our apps again at t=2 (January 2023).[9] We are able to observe the personal data transfers of 3,410 apps at this point. Taken together, at t=0, we collected the privacy policies of 1,939 apps. At t=1, we collected the privacy policies of 7,634 apps and observed actual firm behavior of 10,080 apps. At t=2, we collected the privacy policies of 7,379 apps and observed actual firm behavior of 3,410 apps.

We would like not only to compare the privacy policies and actual behavior of firms that are affected by Schrems II before and after the court decision. We would also like to compare the privacy policies and behavior of these firms with privacy policies and behavior of firms that are not affected by Schrems II. If Schrems II-affected firms change their privacy policies and/or behavior after the court decision, while non-affected firms do not, this increases our confidence that these changes were triggered by the court decision.

We therefore need to split up the data on our apps depending on whether the app was affected by the Schrems II decision or not. As explained in Section 2.1, while the Schrems II decision invalidated existing Standard Contractual Clauses and the Privacy Shield with regard to personal data transfers to the US, it did not affect other apps. In particular, apps that transferred personal data to countries outside the European Economic Area but for the United States were not directly affected by the Schrems II decision. And apps that transferred personal data to the United States, but relied neither on Standard Contractual Clauses nor on the Privacy Shield were also not affected. If we are able to separate these apps from apps that were affected by Schrems II, we can compare their behavior before and after the decision.

In order to decide whether an app was affected by Schrems II, we look at their privacy policy before the decision (i.e., at t=0). In that privacy policy, the app needs to disclose whether it (1) transfers any personal data outside the European Economic Area, (2) whether such transfer goes to the United States or another country, and (3) whether a transfer to the United States is based on the Privacy Shield, on Standard Contractual Clauses, on Corporate Binding Rules, on an Approved Code of Conduct or on an Approved Certification (on the legal details, see Section 2.1). More precisely, we can distinguish between five different types of apps which are described in Table 1. This table also shows the type of cross-border transfer clauses (CBTC) in the

[8] The data collection process originated from these 10’080 apps. From this sample, we extracted sub-samples of the same apps, succeeding to collect a different number of privacy policies (e.g., only 7,634 policies in t=1 from the 10,080 apps for which we had personal flow data). These sub-samples further changed in the two other time points.

[9] Unfortunately, we are not aware of any data source that would allow us to inspect app data flows retrospectively at t=0. Therefore, we only have data about personal data transfers at t=1 and t=2.

respective privacy policy, and it lists whether an app that uses the respective clause is affected by the Schrems II decision or not.

Table 1: Different Types of Cross-Border Transfer Clauses in Privacy Policies

	Description According to the App's Privacy Policy	Type of Cross-Border Transfer Clause (CBTC)	App Is Affected by Schrems II
1	App does not transfer any personal data outside the European Economic Area	No CBTC	No
2	App transfers personal data outside the EEA, but not to the US	CBTC-NonUS	No
3	App transfers personal data to the US, but does not offer any legal basis	CBTC-US-NoLegalBasis	No (app violates the GDPR regardless of Schrems II)
4	App transfers personal data to the US without relying exclusively on the Privacy Shield and/or Standard Contractual Clauses	CBTC-US-NoShield	No
5	App transfers personal data to the US and exclusively relies on the Privacy Shield and/or Standard Contractual Clauses	CBTC-US-Shield	Yes

For our research design, we are particularly interested in apps of type 5. Apps which transfer personal data to the US and which, in t=0, base such transfer solely on either the Privacy Shield and/or Standard Contractual Clauses got into trouble on July 16, 2020, as their data transfer *de facto* became illegal with immediate effect. In an idealized world of perfect enforcement of and compliance with the GDPR, we would expect that the Schrems II decision would push these apps (of type 5) either to base their personal data traffic to the US on justifications other than the Privacy Shield and/or Standard Contractual Clauses, or to stop transferring personal data to the US at t=1, or at t=2 at the latest. And we would expect that the other types of apps would not change their privacy policies of personal data flows after Schrems II.

The research design described in this subsection requires us to collect the privacy policies of apps in the Spanish Google Play store, and it requires us to observe actual traffic flows generated by these apps, all at different points in time. So far, our description has focused on what kind of data we were able to obtain for this study. The next two subsections will explain how we got the data.

3.2 Obtaining and Analyzing Privacy Policies

In order to retrieve the privacy policies of our apps at t=1 and t=2 (see Figure 1), we accessed the Spanish Google Play store through an Application Programming Interface (API) at t=1 and t=2, which allowed us to fetch the apps' privacy policies (using the policy link provided in the store) for these points in time. As the Google Play Store does not provide historical app privacy policies, we reverted to the Wayback Machine of the Internet Archive in order to retrieve privacy policies of our apps at t=0 (see Section 3.1).

In order to compare apps that are affected by Schrems II with apps that are not, we need to identify apps whose privacy policies include information on cross-border data transfers (see Table 1). We are interested in personal data transfers to the United States only, as the Schrems II decision only impacts data transfers to the US. Fortunately, this simplifies our classification task. For any of our apps' privacy policies, we need a classifier to (1) determine whether the app transfers personal data outside the European Economic Area, according to the privacy policy's cross-border transfer clause, (2) determine whether this clause mentions the US as a target country for the personal data transfer, and (3) determine which legal basis the clause mentions for the data transfer (e.g., Privacy Shield, Standard Contractual Clauses, Corporate Binding Rules, etc.).

Traditionally, legal researchers would have performed this task through manual inspection. Research assistants would read privacy policies, identify the cross-border transfer clauses in the policies and determine through manual legal interpretation whether the clauses are in line with the Schrems II decision. As we are looking at over 17,000 privacy policies (from three points in time), manual inspection of these policies is beyond reach. Fortunately, the structure of cross-border transfer clauses in privacy policies is relatively straightforward. From a legal perspective, they include information on whether personal data gets transferred outside the European Economic Area, to which country it gets transferred and on what legal basis such transfer takes place. Therefore, we resort to natural-language processing techniques to extract this information and then assign our apps to one of the five types listed in Table 1.

Two of this study's co-authors developed and presented the battery of natural-language processing techniques required for our endeavor in earlier work (Guamán et al. 2021, 2023). While we refer the technically inclined reader to these computer science publications, we provide a high-level overview of our classification approach in the following. Our pipeline (see Table 2 below) starts by dividing a privacy policy into segments. Each segment corresponds to roughly one paragraph in the policy. These segments are then passed to the so-called *Cross-Border-Transfer-Classifier*. This classifier determines whether a segment discloses the intention of the app's privacy policy to transfer personal information across borders. A segment that discloses this type information is called a *cross-border transfer clause* (CBTC). Our study is not aimed at a complete classification of privacy policies, but rather focuses on these cross-border transfer clauses only. Once the Cross-Border-Transfer-Classifier has identified such a clause, the clause is passed on to a series of additional classifiers. First, the *Target-Country-Classifier* determines whether the US is mentioned as the target country of a data transfer (according to the segment). Then, the remaining classifiers (the *Shield-Classifier*, the *Standard-Contractual-Clauses-Classifier*, the *Binding-Corporate-Rules-Classifier*, the *Approved-Code-of-Conduct-Classifier* and the *Approved-Certification-Classifier*) extract the legal grounds the segment invokes (if any) to comply with the GDPR requirements described in Section 3.1. From a technical perspective, our classifiers can be categorized into two groups: The Cross-Border-Transfer-Classifier and the Shield-Classifier are supervised-learning classifiers. The other classifiers are keyword-based classifiers.

Our supervised-learning classifiers were trained using a human-annotated training dataset. Two of the study's co-authors produced a publicly available corpus (IT-100)[10] that comprises 100 manually annotated privacy policies containing a total of 281 cross-border transfer clauses among

[10] The corpus is available at https://github.com/PrivApp/IT100-Corpus.

a total of 3,715 segments. Based on this training dataset, the Cross-Border-Transfer-Classifier learned to recognize cross-border transfer clauses, while the Shield-Classifier learned to recognize references of the Privacy Shield as the legal basis of a cross-border transfer to the US. In order to successfully train a supervised-learning classifier, the training dataset must contain a minimum number of training examples as too few examples might lead to poor performance of the trained classifier. For the remaining classifiers – the Target- Country-Classifier, the Standard-Contractual-Clauses-Classifier, the Binding-Corporate- Rules-Classifier, the Approved-Code-of-Conduct-Classifier and the Approved-Certification- Classifier – the number of positive examples in the IT-100 corpus was not sufficient to train a supervised-learning classifier. We therefore resorted to keyword-based classifiers in these cases. In comparison to supervised-learning classifiers, keyword-based classifiers are relatively simple. In essence, they operate as a sophisticated word-search engine. Nonetheless, for certain tasks, they remain a viable option which can produce reliable results.

3.3 Extracting and Analyzing Personal Data Flows From Apps

As described in Section 3.1, if a study analyzing compliance with the Schrems II decision focused only on privacy policies, it would provide an incomplete picture: it may be that the app has adapted its privacy policy after Schrems II, but has not changed its actual personal data flow after the decision. This behavior may be illegal, but can be fully rational, as it is much more challenging to observe an app's data flows than to observe its privacy policy. This drastically reduces the probability of getting fined for violating the GDPR in an app's data flow.

In this study, we take advantage of recent computer science research to observe the actual data flows between an app running on a smartphone and the outside world. Two of this study's co-authors deployed methods to understand what kind of data an app handles, with whom it shares such data, and where these recipients are located. For this, Guamán et al. (2021, 2023) use a dynamic analysis method, observing the app's behavior in real time. Again, while we refer the technically inclined reader to these computer-science publications, we provide a high-level overview of our data flow extraction approach in the following.

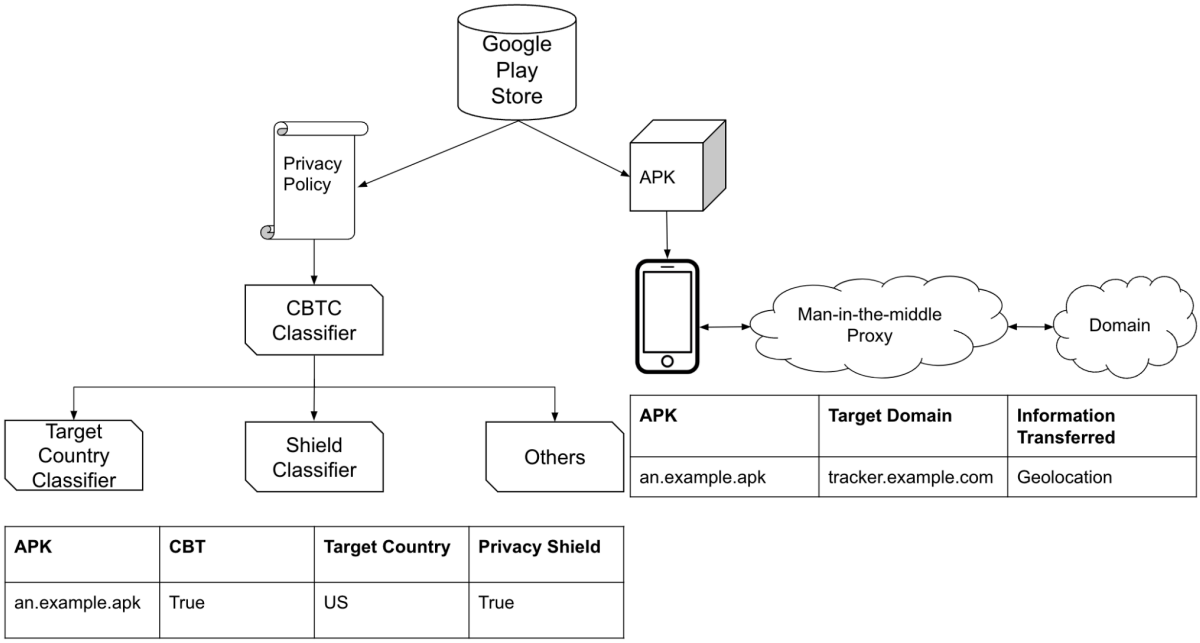
We download the relevant app (or, more precisely, the app's Android Package Kit, APK) from the Google Play Store, using an unofficial Google Play Application Programming Interface (API). We install the app on five different smartphones running under the Android operating system (Xiaomi Redmi devices). Observing data flows from the app to the outside world involves three steps:

1. *Stimulation*. We use an automated process to generate a sequence of semi-random activities that are intended to create a user-like interaction with the app (Android Studio 2023, see also Choudhary 2015). Such methods are well-known from security research stress-testing apps.
2. *Interception*. Observing the actual data traffic flows from apps is challenging, as the apps typically use various encodings and encryption techniques to protect their traffic flows from being intercepted. We use various so-called "man-in-the-middle attacks" known from IT security research that enable us to infiltrate the app's traffic flow with an intermediary through which we are able to observe all the data that an app sends from the smartphone to the outside world even when encrypted. We reuse methods developed by two of this study's co-authors in Guamán et al. (2021, 2023).

3. *Classifying and geolocating data traffic*. From the intercepted traffic flow leaving an app, we extract four different types of personal data: contact information (e.g., whether the user's email address or postal address is transmitted); demographics (e.g., whether the user's age or gender are transmitted); identifiers (e.g., whether the user's Google advertising ID, the fingerprint of the user's mobile device, or the identifier of the user's SIM card are transmitted); and location (e.g., whether the user's GPS location is transmitted). Table A.1 in the Appendix provides a full overview of all types of personal data that we are able to observe (see Guamàn 2021: 15965). In order to determine whether apps transfer personal data to the United States, we observe the IP address and domain name of the server the app communicates with. By geolocating this first-hop server, we can identify apps which transfer data to the United States (Guamàn 2021: 15965).

Figure 2 provides an overview of our extraction and classification pipeline used to create the dataset for this study.

Figure 2: Extraction and Classification Pipeline



Note. The figure describes how (1) we collect the privacy policies for apps from the Google Play Store, then use a machine-learning classifier to identify cross-border transfer clauses (CBTCs) in these privacy policies, and then classify these CBTCs in order to identify those apps which transfer personal data to the US, relying on the Privacy Shield or Standard Contractual Clauses as legal basis. The figure also describes how (2) we access the Android Package Kit of an app, interact with the app using automated means, intercept the outgoing data traffic generated by the app and determine whether this traffic contains personal data that is transferred to the United States. The names of the apps in this figure are fictional.

3.4 Validating Legal Classification of Privacy Policies

The automatic classification of privacy policies as described in Section 3.2 is a delicate task. As we are looking at over 17,000 privacy policies (from three different points in time), using

machine-learning to perform these tasks is, of course, appealing. However, as for any new method penetrating legal research, one must develop some confidence in its performance. In their original paper, two of this study's co-author had already evaluated the performance of their classifiers, as is standard practice in the machine-learning community (Guamán 2021). As the classifiers get applied to a central legal question of our study, we engage in a separate extensive manual validation of our classifiers before presenting results of this study.[11]

Machine learning models are typically evaluated on unseen data that is not part of the training dataset. Typically, for classification tasks, validation involves two main metrics: recall and precision. The recall of a classifier is the proportion of actual positives that the classifier identified correctly (or, put differently, low recall signalizes many false negatives). The precision of a classifier is the proportion of identifications that were actually correct (or, put differently, the number of true positives divided by the number of true and false positives). For our application, let us consider the following example: A classifier is tasked to identify the cross-border transfer clauses in a sample of 100 clauses. In reality, ten of those clauses are cross-border transfer clauses (positives), while 90 clauses are not (negatives). The classifier identifies a total of nine clauses as cross-border transfer clauses. However, two of those clauses are not cross-border transfer clauses. This means that the classifier had correctly classified seven clauses (true-positives), erroneously classified two clauses (false-positives) and missed three clauses that actually are cross-border transfer clauses (false-negatives). In this case, the classifier's recall is 0.7 – 7 out of 10 cross-border transfer clauses were identified correctly – and its precision is 0.78 – 7 out of 9 clauses the classifier identified as cross-border transfer clauses were in fact such clauses.

In our study, we performed two separate validations. In validation 1, the CBTC-classifier presented in Section 3.2 was tasked with classifying the clauses of 100 randomly selected privacy policies from all policies we obtained for t=1 into clauses that cover cross-border transfers and clauses that do not. We compared the results against the classification conducted manually by a research assistant with a legal background. The results are displayed in Table 2. Validation 2 focused on evaluating the performance of the transparency element classifiers – the classifiers responsible to extract relevant information such as the target country from the cross-border transfer clauses. The classifiers were tasked with extracting the relevant information from 100 randomly selected cross-border transfer clauses from t=1. Again, the results were compared against human annotation. The results are also displayed in Table 2.

[11] Our second method – the automated extraction and analysis of personal data flows – does not involve any statistical methods and therefore is close to a deterministic process. For a validation of the geolocation of these personal data flows, *see* Cozar et al., (2022).

Table 2: Comparing Machine versus Human Classification

	Total Number of Clauses	Positives	True Positives	False Positives	False Negatives	Recall	Precision	Recall (Policy)	Precision (Policy)
CBTC-Classifier	3807	78	71	26	7	0.91	0.72	0.98	0.91
Target-Country-Classifier	100	52	52	9	0	1	0.85	1	0.91
Shield-Classifier	100	25	25	3	0	1	0.89	1	1
Standard-Contractual-Clauses-Classifier	100	9	9	0	0	1	1	1	1
Binding-Corporate-Rules-Classifier	100	0	0	0	0	-	-	-	-
Approved-Code-of-Conduct-Classifier	100	0	0	0	0	-	-	-	-
Approved-Certification-Classifier	100	0	0	0	0	-	-	-	-

Note. The table reports the results of our validation. Columns two to eight report the results on the segment level, i.e., the level on which the classifiers operate. The last two columns report the results on the policy level, i.e., the level on which we base our legal reasoning. Recall and precision for the two levels are not identical since a misclassification of a single segment does not necessarily lead to a misclassification of the entire policy.

Before discussing the results, we would like to elaborate on the subtle yet important difference between what we call the *segment level* and the *policy level*. As discussed in Section 3.2, our classifiers operate on the level of individual segments (or clauses), which roughly correspond to a paragraph in a policy. For each segment, a classifier first determines whether this segment represents a cross-border transfer clause or not. If it represents a cross-border transfer clause, the classifier then extracts the relevant information from the clause (such as the target country and the safeguards invoked to comply with the GDPR requirements). To answer our research questions, however, we are not interested in the classification of single segments but in the classification of entire privacy policies. Therefore, after classifying each segment individually, we label the policy based on the aggregate of all its segments' classifications. For instance, if a policy contains a cross-border transfer clause which mentions Standard Contractual Clauses as the legal basis for a transfer, the entire policy will be labeled as relying on Standard Contractual Clauses. This distinction between segment level and policy level might seem redundant at first. However, in practice, this process of aggregation significantly increases the robustness of our classification since one can avoid that the misclassification of a single segment leads to the misclassification of the entire policy. If, for example, the Standard Contractual Clauses are mentioned in more than one segment, the fact that the classifier only identified them in one segment, while a mistake on the level of individual segments, does not influence the classification of the entire policy since the policy will nonetheless be labeled as relying on Standard Contractual Clauses.Therefore, the reported performance regarding the classification of individual segments can be seen as the lower-bound expected performance on the policy level.

Let us consider the results of our validation as reported in Table 2, starting with the CBTC-Classifier. The CBTC-Classifier scores a recall of 91% and a precision of 72% on the segment level. When aggregating the results on the policy level, recall increases to 98% and

precision to 91%. The Target-Country-Classifier scores a recall of 100% with a precision of 85%. When aggregating the results on the policy level, the precision increases to 91%. The Shield-Classifier scores a recall of 100% with a precision of 89%. On the policy level, the Shield-Classifier's precision improves considerably reaching a precision of 100%. The Standard-Contractual-Clauses-Classifier performs perfectly, scoring a precision and a recall of 100% on both the segment level and – by extension – on the policy level.[12]

Overall, our validation indicates that the deployed classifiers perform very well, especially when aggregating the classifications on the policy level. Still, they are not perfect. For our purposes, the classifiers would ideally produce a reliable lower-bound estimation of how many apps' privacy policies violate the GDPR. To respect this lower bound, our classifiers should ideally score high on precision even at the expense of a lower recall, but for a few exceptions. Such exceptions may include instances when, for example, a high recall is a step in the process (CBTC classifier) or when a positive classification could only benefit the apps (such as legal grounds which are not affected by the Schrems II decision). Our classifier which identifies the target country of a data transfer (Target-Country-Classifier) scores high in precision at the policy level, yet tends to be overinclusive thereby leaving scope for further fine-tuning. We discuss the implications and mitigating approaches in Section 5, where we also explore classifying cross-border transfer clauses through current generic large language models (LLMs).

4. Results

As described in Section 3.1, in order to analyze the impact of the European Court of Justice's Schrems II decision, we would like to observe whether apps that are affected by Schrems II change the cross-border transfer clauses in their privacy policy after the decision, and whether these apps adapt their actual personal data transfer flows after the decision. We then compare the behavior of these apps with behavior of apps that are not affected by Schrems II. In the following, we will analyze changes in apps' privacy policies first, and then combine it with an analysis of changing personal data flows.

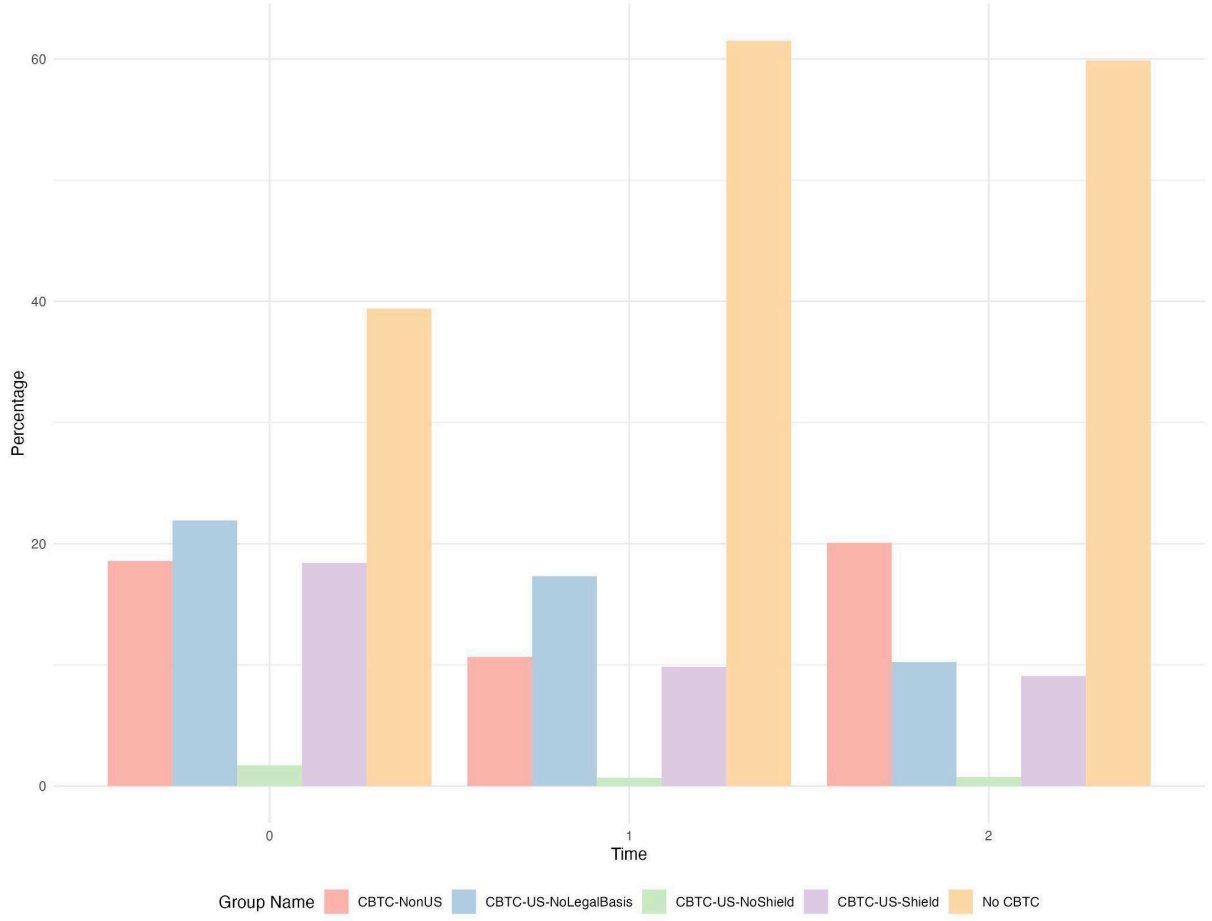
4.1 Changes in Privacy Policies

As a first cut, we are interested in the number of apps that include one of the five different cross-border transfer clauses outlined in Table 1 in their privacy policies. We track the distribution of apps over these five types before (t=0), right after (t=1) and over two years after the Schrems II decision (see Figure 1 for the timeline).

Figure 3 displays the distribution of our entire sample of app privacy policies (t=0: 1,939 apps; t=1: 7,634 apps; t=2: 7,379 apps) over the five types of cross-border transfer clauses. At first sight, this data seems to suggest that the percentage of apps whose privacy policies rely on the Privacy Shield or Standard Contractual Clauses ("CBTC-US-Shield") decreases from around 20 percent at t=0 to about 10 percent at t=1 and t=2.

[12] The remaining classifiers – the Binding-Corporate-Rules-Classifier, the Approved-Code-of-Conduct-Classifier and the Approved-Certification-Classifier – could not be fully validated since no segment in our validation set invoked either of these safeguards. However, since these classifiers work in the exact same way as the Standard-Contractual-Clauses-Classifier, it is fair to assume that they perform equally well.

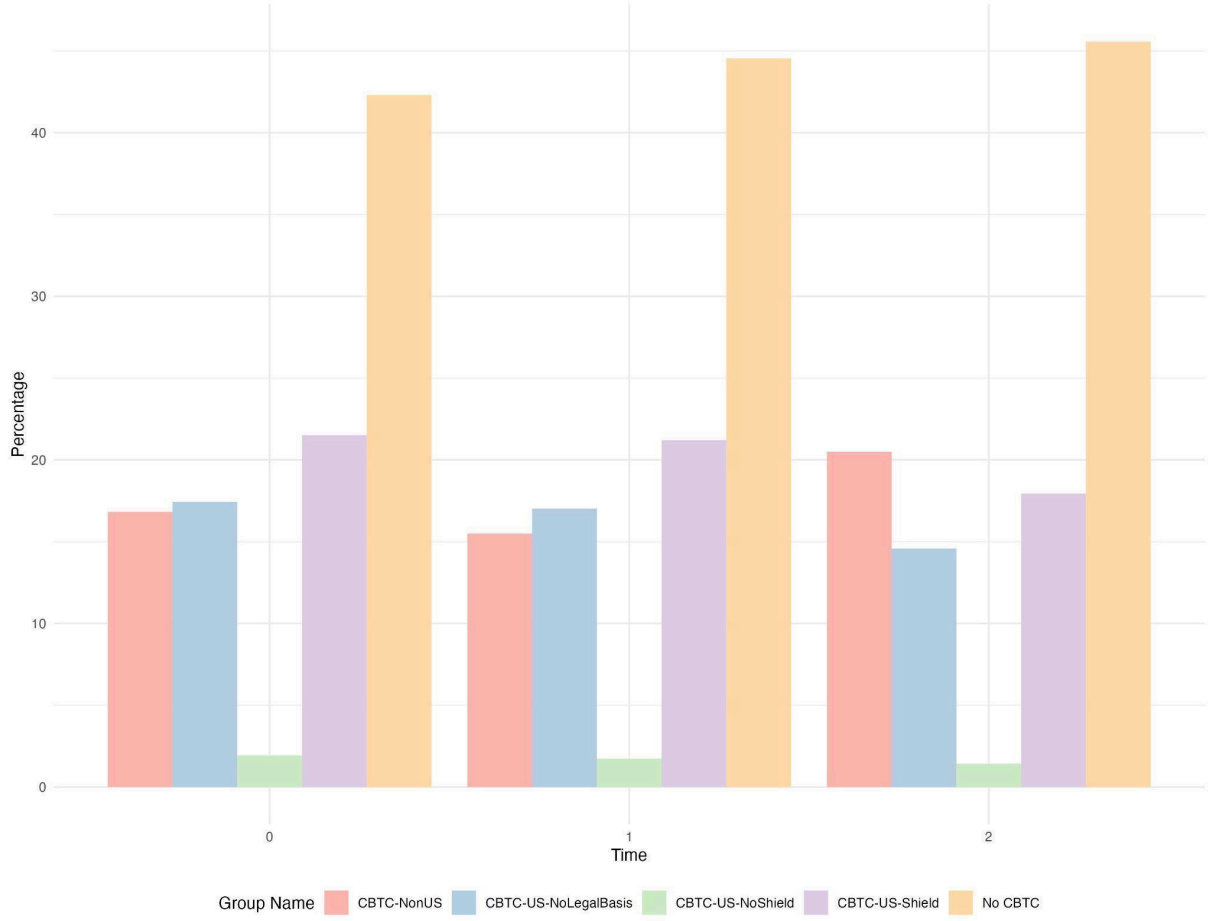
Figure 3: Distribution of all Apps Over CBTC Types



However, as the number of privacy policies that we were able to retrieve at our three points in time differed considerably (see Section 3.2), we caution against such interpretation. The composition of apps we are able to observe changes over time. Still, one already sees that the number of apps which transfer personal data to the US and provide another legal instrument than the Privacy Shield or Standard Contractual Clauses as their legal basis ("CBTC-US-NoShield") is very limited. This gives us comfort in our assessment that the two main legal instruments used by apps to justify personal data transfer to the US are the Privacy Shield and Standard Contractual Clauses (on the legal background, see Section 2.1), and that there is a considerable amount of apps (10 percent) which do not respond to the decision in both t=1 and t=2.

Matching the apps over the three time points is the simplest way to have a 'before' and 'after' picture, making sure we compare the same 'treated' apps vs the same 'control' apps in all time points. For a total of 981 of these apps, we can follow changes in their privacy policies over time until t=2. Figure 4 replicates Figure 3, but focuses on the 981 apps for which we have complete information of policy data.

Figure 4: Distribution of Apps with Complete Information Over CBTC Types



This figure seems to suggest that most apps that transfer personal data to the US based on either the Privacy Shield or Standard Contractual Clauses at t=0 ("CBTC-US-Shield") do *not* change the cross-border transfer clauses in their privacy policies right after the Schrems II decision (t=1). And even two years after the decision (t=2), only a small number of these apps seem to have removed cross-border transfer clauses relying on the Privacy Shield or Standard Contractual Clauses from their privacy policies.

By matching apps included in all time points, and specifically t=0, there is a risk of selection bias if the matched apps are different in one way or another from the full sample in t=1 (Figure 3). Note that the privacy policies from apps in t=0 are retrieved from the Wayback Machine archive. In the context of mobile apps, the t=0 sample might differ in levels of popularity or availability on the Google store (for example, new apps included in t=1 after the Schrems II decision might behave differently than the older ones). In our sample, less than ten percent of apps which are included in t=1 but not in t=0 are apps that appeared in the Google store after the European Court of Justice decided Schrems II. This suggests that the 'stickiness' of older apps is not biasing our restricted sample. Yet, we find that t=0 apps are more popular, on average, as proxied by the number of downloads, than t=1 apps (t=0 apps had an average of around 7 million downloads, versus 2 million for t=1 apps that were not included in t=0). We investigate the relationship between popularity and non-compliance in section 4.2, finding that popularity is associated with non-compliance. This might suggest that our Wayback Machine archive sample is an upper-bond

estimation, and that the number of apps which rely on the Shield after the Schrems decision is between 10 to 20 percent.

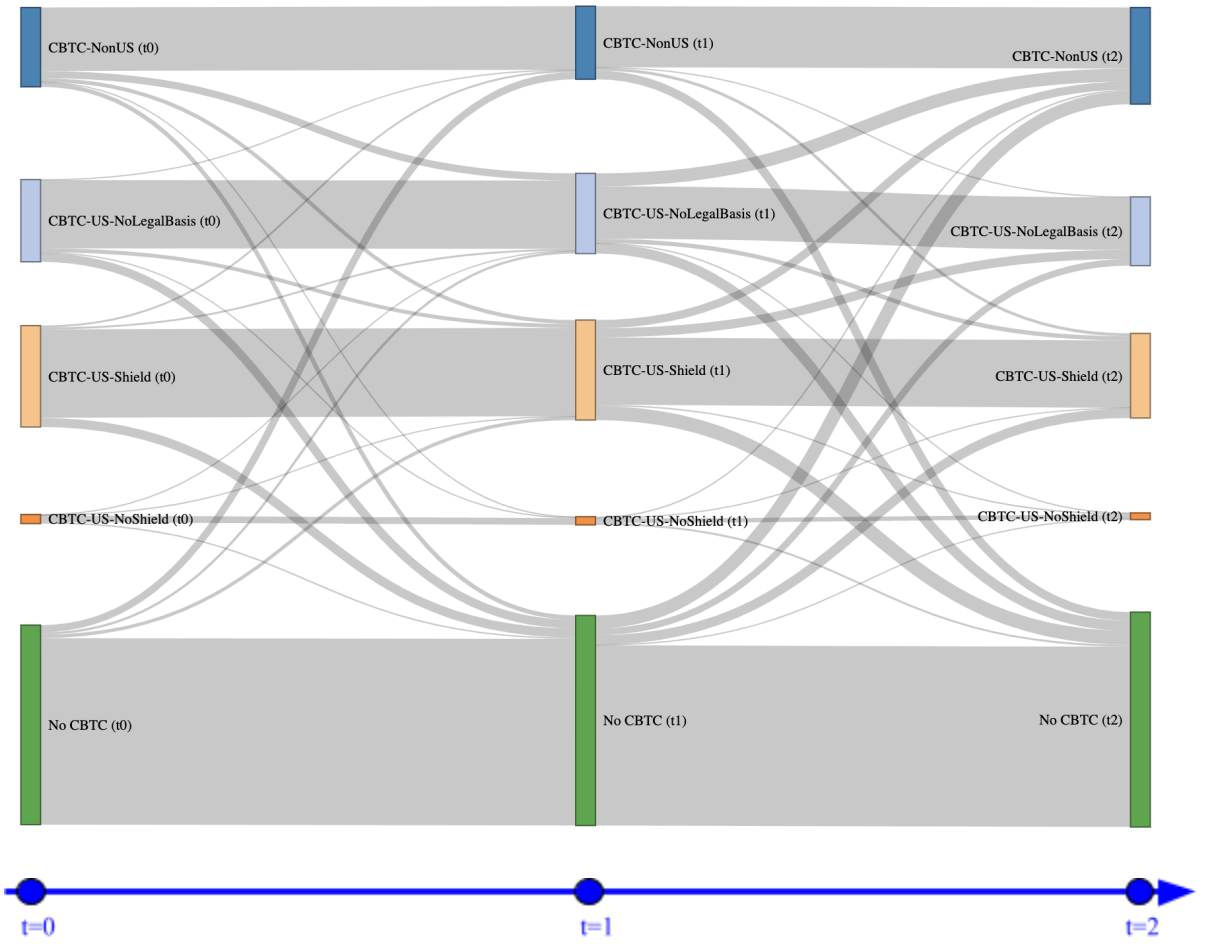
As described in Section 2.1, privacy policies relying on the Privacy Shield as the sole legal basis for third-country transfers are in clear violation of the GDPR after Schrems II, while privacy policies relying on Standard Contractual Clauses could – at least theoretically – take supplementary measures to achieve compliance. To focus on privacy policies that clearly violate Schrems II we offer more sensitivity checks in the Appendix, confirming the main findings.[13]

We further investigate how apps switch from one type of cross-border transfer clause to another type between t=0, 1 and 2. Figure 5 presents a Sankey diagram displaying how app privacy policies move from one type of cross-border transfer clause to another type over time.

For the purposes of our paper, let us focus on apps whose privacy policies state at t=0 that they transfer personal data to the US and rely on either the Privacy Shield or Standard Contractual Clauses for this transfer (CBTC-US-Shield, depicted at t=0 with the gold vertical bar in the left of the diagram). If one follows these apps to t=1 (the gold vertical bar in the middle of the diagram), one can see visually that a small fraction of apps abandon their reliance on the Privacy Shield or Standard Contractual Clauses and do not feature any cross-border transfer clause at t=1 (move to the large blue vertical bar in the middle of the diagram). However, the vast majority of apps keep their CBTC-US-Shield clause at t=1 (gold vertical bar in the middle of the diagram). If one then continues to follow the apps with whose privacy policies state at t=1 that they transfer personal data to the US and rely on either the Privacy Shield or Standard Contractual Clauses for this transfer (CBTC-US-Shield, depicted at t=2 with the green vertical bar), one can see visually that some fraction of these apps drop their cross-border transfer clause (thereby claiming that no personal data transfer outside the European Economic Area takes place); others stop the personal data transfer to the US; and others continue with personal data transfers to the U.S, but mention other (or no) legal instruments as a basis for such transfer. Similarly to t=1, however, most apps with a CBTC-US-Shield clause at t=1 keep this clause at t=2. Interestingly, there are even apps switching to a CBTC-US-Shield clause at t=1 or t=2.

[13] We replicate Figure 4 as Figure A.1 in the Appendix, but put the Privacy Shield and Standard Contractual Clauses into separate bins. It confirms the impression from Figure 4: Even two years after the decision, a large portion of the studied apps rely on the Privacy Shield as the only legal basis for their third-country transfer, and apps that also rely on Standard Contractual Clauses are slow to respond to the court decision. Figure A.2 replicates Figure 5 of the main text under the same assumption (i.e., separate bins for Standard Contractual Clauses). We also check whether the main analysis is sensitive to focusing on those apps for which we have access to their privacy policies at t=0 in 2020 (i.e, policies that we were able to pinpoint the date of collection in the Wayback Machine to between January 1 and July 15, 2020). Figure A.3 in the Appendix plots the distribution of policies per month in t=0 between May 2018 and July 2020. In total, around 30% of the observations in t=0 are from 2020, i.e. in the six months before the Schrems II decision. The results are similar: apps from the CBTC-US-Shield group remain stable from t=0 to t=1, with a small drop in t=2 as depicted in Figure A.4 in the Appendix.

Figure 5: Sankey Diagram on Movements Between CBTCs



Note. The figure visualizes how app privacy policies move from one type of cross-border transfer clause (see Table 1) to another type over time. For each type of cross-border transfer clause, the movement starts on the left with the first vertical bar depicting t=0. The second, middle vertical bar depicts movement at t=1, while the vertical bar on the right end of the figure depicts movement at t=2.

Visual inspection of the Sankey diagram seems to suggest that there is a considerable number of privacy policies moving their cross-border transfer clauses from one type to another type over time. The diagram also seems to indicate that among those apps which transferred personal data to the US based on Standard Contractual Clauses or the Privacy Shield before the Schrems II decision, the majority of these apps continue with this practice a few weeks and even over two years after the Schrems II decision. And there are apps which newly adopted Standard Contractual Clauses or the Privacy Shield after the European Court of Justice held that neither of them provide a sufficient legal basis for personal data transfer to the US. We replicate Figure 5 in the appendix as Figure A.2, but put the Privacy Shield and Standard Contractual Clauses into separate bins, based on the same logic mentioned above. The results remain similar.

To move beyond visual inspection, we quantify the odds of privacy policies changing their cross-border transfer clause over time, under the null hypothesis of no association between the type CBTC and switching tendency. If successful, we can investigate whether apps with a cross-border transfer clause of a particular type (such as CBTC-US Shield) are more or less likely than expected to switch to another cross-border transfer clause at t=1 or t=2.

First, we are interested in whether having a particular type of cross-border transfer clause (as defined in Table 1) is associated with the privacy policy switching this clause over time: Are privacy policies with a CBTC-US Shield clause more or less likely to switch the clause after Schrems II, compared to all privacy policies, or is switching behavior distributed normally across all privacy policies? The left table in Table 3 provides a statistical test of this question for the movements between t=0 and t=1, while the right table in Table 3 provides the same test for the movements between t=1 and t=2. The Chi-Square test and the Fisher's exact test indicate that there is a statistically significant relationship between privacy policies using a particular cross-border transfer clause and switching behavior. The contingency tables in Table 3 allow us to inspect which switches between cross-border transfer clauses contribute to the statistically significant relationship between type of clause and switching behavior. Zooming in on the privacy clauses with a CTBC-US Shield clause – which is our main interest – one can see that the test would predict only 38 of privacy policies to move away from the CBTC-US Shield clause at t=1. In fact, we observe in the data that only 33 privacy policies do so at t=1. At t=2, the test predicts 91 privacy policies to move away from the CBTC-US Shield clause. And we observe in the data an even higher number (152) of privacy policies moving away from using the Privacy Shield or Standard Contractual Clauses as their basis for personal data transfer to the US

The mosaic plots Figure 6 represent these findings in a visual way. Focusing on privacy policies with a CTBC-US Shield clause moving from t=0 to t=1 (left plot in Figure 6), one can see that only a small portion of privacy policies switch this clause at t=1. The CTBC-US Shield privacy policies switch less frequently, compared to other cross-border transfer clauses. When one looks at switching behavior between t=1 and t=2 (right plot in Figure 6), one can observe, however, that more privacy policies with a CTBC-US Shield clause switch to other cross-border transfer clauses at that point, and that they also switch at a frequency that is statistically higher, compared to privacy policies in the No CBTC group which becomes more stable than expected.

Table 3: Switching CBTCs over Time, Contingency Tables

A-1 (t0->t1)

Switchers	*Group name*					***Total***
	CBTC-NonUS	CBTC-US-NoLegalBasis	CBTC-US-NoShield	CBTC-US-Shield	No CBTC	
0	188	299	21	243	556	1307
	216	310	25	238	518	1307
	14.4 %	22.9 %	1.6 %	18.6 %	42.5 %	100 %
	74.9 %	83.3 %	72.4 %	88 %	92.5 %	86.2 %
1	63	60	8	33	45	209
	35	49	4	38	83	209
	30.1 %	28.7 %	3.8 %	15.8 %	21.5 %	100 %
	25.1 %	16.7 %	27.6 %	12 %	7.5 %	13.8 %
Total	251	359	29	276	601	1516
	251	359	29	276	601	1516
	16.6 %	23.7 %	1.9 %	18.2 %	39.6 %	100 %
	100 %	100 %	100 %	100 %	100 %	100 %

χ^2=55.102 · df=4 · Cramer's V=0.191 · Fisher's p=0.000

observed values
expected values
% within Switchers
% within Group name

A-2 (t1->2)

Switchers	*Group name*					***Total***
	CBTC-NonUS	CBTC-US-NoLegalBasis	CBTC-US-NoShield	CBTC-US-Shield	No CBTC	
0	473	379	24	312	2309	3497
	448	542	28	373	2106	3497
	13.5 %	10.8 %	0.7 %	8.9 %	66 %	100 %
	84.9 %	56.2 %	68.6 %	67.2 %	88.2 %	80.4 %
1	84	295	11	152	309	851
	109	132	7	91	512	851
	9.9 %	34.7 %	1.3 %	17.9 %	36.3 %	100 %
	15.1 %	43.8 %	31.4 %	32.8 %	11.8 %	19.6 %
Total	557	674	35	464	2618	4348
	557	674	35	464	2618	4348
	12.8 %	15.5 %	0.8 %	10.7 %	60.2 %	100 %
	100 %	100 %	100 %	100 %	100 %	100 %

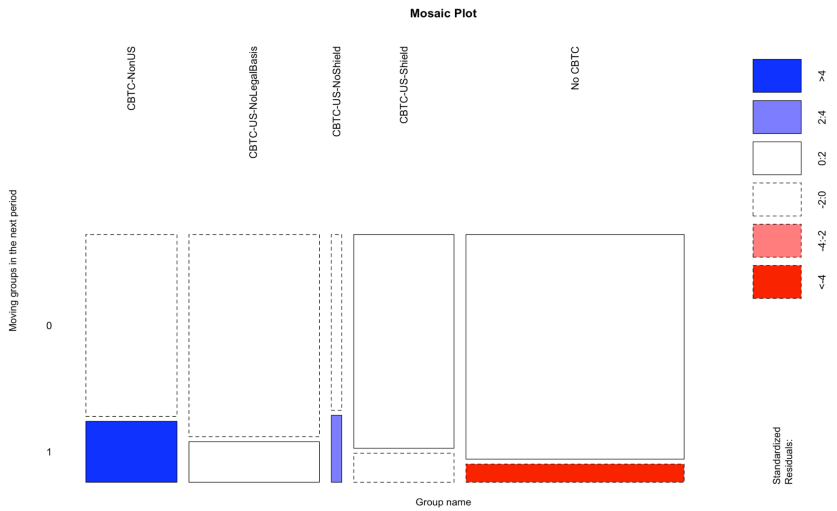
χ^2=412.583 · df=4 · Cramer's V=0.308 · p=0.000

observed values
expected values
% within Switchers
% within Group name

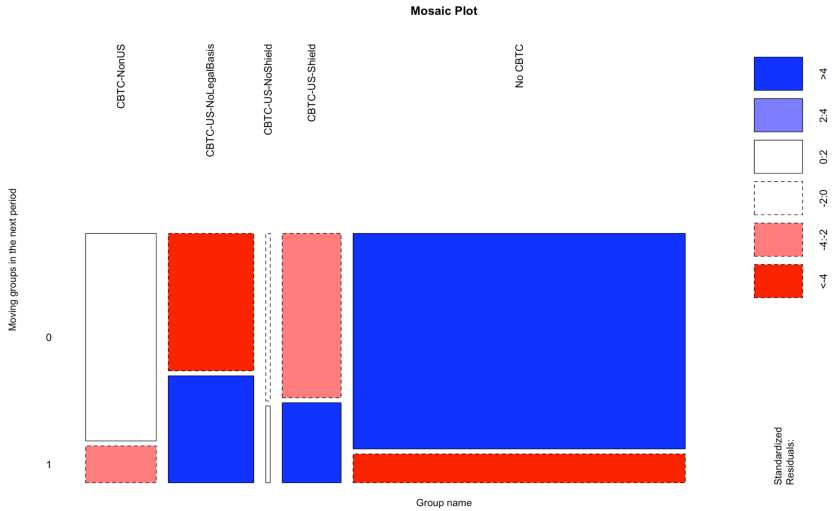
Note. In the context of the Chi-square test for independence, the null hypothesis assumes that there is no association between the categorical variables or no difference between the observed and expected frequencies, respectively. Under this null hypothesis, the test statistic follows a Chi-square distribution. When conducting a Chi-square test, the test statistic is calculated by comparing the observed frequencies with the expected frequencies (the first values in each cell of the tables – black – are the observed values and the second values – turquoise – are the expected values under the null). The test statistic follows the Chi-square distribution, and by comparing its value to the critical values (in this case 55.102 and 412.58, df=4) from the Chi-square distribution, one can determine the statistical significance of the association or difference being tested (policy group type and switching groups in the next observation). In both cases, we reject the null suggesting that the variables are associated.

Figure 6: Switching CBTCs over Time, Mosaic Plots

B-1 (t0->t1)



B-2 (t1->2)

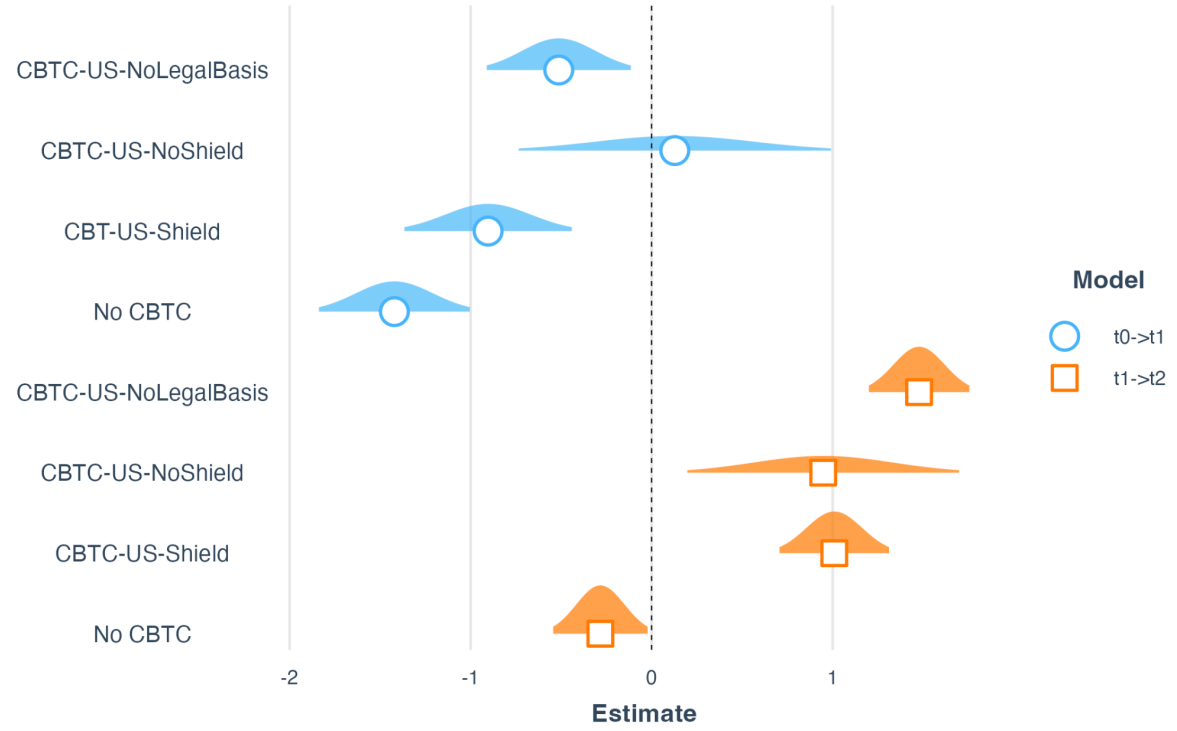


Note: The mosaic plots provide a visual representation of data in the contingency tables in Table 3. The blue color cells represent positive standardized residuals (meaning higher observed value than expected under the null) and red cells the opposite (lower observed than expected under the null). The width of each column in the plot corresponds to the proportion of that category among the groups' variables. The height of each tile within a column corresponds to the proportion of that category in the shifters variable within that specific category. On the left panel, presenting the shift from t=0 to t=1, there is little movement above the expected values from the US-Shield group. In the second period from t=1 to t=2, on the right panel, US-shield (and US without a legal basis) apps start to move (what is indicated by the blue bottom tile for y=1).

The contingency tables in Table 3 only provide a statistical test on whether there is a significant relationship between privacy policies using any particular type of cross-border transfer clauses and switching behavior, compared to a situation in which the type of cross-border transfer clause used by a privacy policy would not relate to its switching behavior. In order to identify whether this statistical relationship does not only apply to all types of cross-border transfer clauses in general, but to the CBTC-US-Shield clause in particular, and to quantify its magnitude, we use a general linear model that predicts the average likelihood of a privacy policy using a particular type of cross-border transfer clause switching to another clause at the next point in time point.

Figure 7 presents a coefficient plot of a GLM model (logistic regression), estimating whether privacy policies with a particular type of cross-border transfer clause are more or less likely to switch to another type of clause at t=1 (blue) or t=2 (orange). The output of the logistic model is on link-scale (logit). Thus, the numerical output of the model corresponds to the log-odds. For example, the coefficient of the CBTC-US-Shield variable has a numerical value of almost -1 in t=0 (moving to t=1) or 1 in t=1 (moving to t=2). Its negative sign in the first case indicates that the chance of observing a switch decreases with the association of a privacy policy with a CBTC-US-Shield clause. The magnitude of the coefficient implies that, compared to the main control group (privacy policies with a CBTC No-US clause), having a CBTC-US-Shield clause in t=0 results in -1 decreases in the log-odds ratio for the privacy policy to switch to another type of clause at t=1. Therefore, at t=1, the odds of switching to another type of clause are 63.21% (=1-exp(-1)) lower for privacy policies with a CBTC-US-Shield group, compared to the benchmark of privacy policies with a CBTC No-US clause (which are not affected by the Schrems II decision).

Figure 7: Switching CBTCs over Time, Coefficient Plot



Note. The coefficient plot summarizes the magnitude of a GLM (logit) model using each type of privacy policy cross-border transfer clause as predictors, with the switching decision at time t=1 and t=2 as the dependent variable. The benchmark (constant) is the CBTC Non-US group of privacy policies. The figure shows CBTC-US-Shield privacy policies as less likely to switch compared to the benchmark at t=1 just after the court's decision. In the long run, two years after the decision, CBTC US-Shield privacy policies are more likely to switch, though still less compared to CBTC-US-NoLegalBasis privacy policies. The shape indicates a 95% confidence interval.

Overall, the general linear model confirms our previous tests. At t=1, some privacy policies with a CBTC-US Privacy Shield clause switch to another cross-border transfer clause, but overall, these policies do not switch at a higher frequency – in a statistically significant way – than other policies that are not affected by the Schrems II decision. At t=2, more privacy policies with a CBTC-US Privacy Shield clause switch to another cross-border transfer clause, and they also do this in a statistically significant way, compared to other policies that are not affected by the Schrems II decision, such as the CBTC-US-NoShield.

These tests show that a considerable portion of apps that base their personal data transfer on either the Privacy Shield or Standard Contractual Clauses change their privacy policies after the Schrems II decision at least in the longer run. But this should not divert our attention from the fact that, even over two years after the European Court of Justice decided Schrems II, a significant portion of apps with a CBTC-US Privacy Shield clause have not responded to the court decision (see the discussion at Figure 5). On July 16, 2020, the European Court of Justice invalidated the EU-US Privacy Shield and held that, without additional safeguards, Standard Contractual Clauses cannot be used to justify personal data transfers to the US. When we observe the privacy policies of several thousand apps that are affected by this decision in July & August 2020, a large fraction of these apps do not change their privacy policies. And when we look at their privacy policies again in January 2023, non-compliance with Schrems II is still wide-spread.

4.2 Combining Observations on Privacy Policies with Personal Data Flows

As discussed in Section 3.1, checking compliance with the GDPR through observing privacy policies is a helpful proxy. But it has its limits, as one only observes that firms *state* what they are doing, rather than observing what they *actually* do. Perhaps, some apps whose privacy policies do not comply with Schrems II in t=2 have actually adapted their data traffic flows and do not transfer personal data to the US any more. In that case, the fact that their privacy policy does not comply with Schrems II may not matter.[14] Conversely, some apps which have adapted their privacy policies following Schrems II and claim they are no longer transferring personal data to the US might actually still engage in such data transfers. By merely tracking privacy policies over time, we simply do not know whether firms are misstating their actual data traffic flows in their own privacy policies or not. We therefore need a way to observe the actual personal data traffic flows leaving an app. Fortunately, as described in Section 3.3, our framework allows us to observe such traffic.

As described in Section 3.1, we are able to observe the personal data traffic flows from 10,080 apps at t=1 and from 3,410 apps at t=2. This includes contact information, demographic information, as well as information about user location and identifiers. Relevant to our analysis, the data also includes information on whether data gets transferred to third parties and to which country it gets transferred.

This information allows us to contrast the declaration in the *privacy policies* and the *actual transfers* to the US in order to observe the extent to which apps are complying with the Schrems II decision on both dimensions. Conceptually, apps may violate or comply with Schrems II both with regard to their privacy policy and their data transfers. Suppose an app transferred personal data to the US before Schrems II and used the Privacy Shield as its legal basis. Following Schrems II, this app may decide to stop any personal data transfer to the US and adapt its privacy policy accordingly. It becomes a "perfect complier." Or the app may decide to adapt its privacy policy – by, e.g., stating that no personal data transfer to the US occurs any longer – but still continue transferring personal data to the US Such app is a "covert violator" (as the statement in privacy policy can be easily detected, whereas the change in data flows is difficult to detect). The app may also decide to stop the personal data transfer to the US without indicating a valid legal basis in its privacy policy. Such app is a "lazy complier" (as the app's actually behavior does not violates the GDPR, whereas the app's statement in its privacy policy suggests that does). Finally, the app may decide to neither adapt its privacy policy nor its data traffic flows. In this case, the app is a "consistent violator." Table 4 provides an overview of these different types.

[14] On the problem of historical holdovers and obsolete contractual provisions in commercial contract production, *see* Scott, Choi & Gulati (2024); Gulati & Scott (2012).

Table 4: Types of Compliers and Violators

Privacy policy	Data Transfer	
	complies with Schrems II	violates Schrems II
complies with Schrems II	Perfect complier	Covert violator
violates Schrems II	Lazy complier	Consistent violator

In t=1, right after the European Court of Justice decided Schrems II, we find high levels of non-compliance for both privacy policies and app data flows. Among the 7,634 apps we analyze, 3,565 are covert violators. This means that out of our sample, 46.70% of the apps state in their privacy policy that no personal data transfer to the US occurs, even though the apps still continue transferring personal data to the US. Another 20.81% of the apps (1,589 out of 7,634) are consistent violators: They declare a transfer to the US in their privacy policy and actually transfer personal data to the US, without relying on a valid legal basis. 6.34% of the apps (484 out of 7,634) can be considered lazy compliers: We did not detect any actual personal data transfers to the US, while their privacy policies declare such transfers without relying on a valid legal basis. Only 26.15% of the apps (1,996 out of 7,634) are perfect compliers in that they neither state in their privacy policy to transfer personal data to the US nor actually engage in such data transfer to the US.

Two years separate our two datasets of app data flows. Some exogenous circumstances may have changed between timepoints t=1 and t=2, which may have an impact on how much and what type of traffic data we can observe. In particular, in the approximately two years between t=1 and t=2, several new Android versions were introduced. One of them – Android 10 – introduced significant privacy changes affecting both the type of personal data that an application might collect and its scope of collection.[15] Furthermore, it is likely that some apps had not been updated to support the most recent Android version running on our systems (Android 11) causing some failures to extract their data. These circumstances are evident by the mere size difference between t=1 and t=2 (10k vs 3.5k).

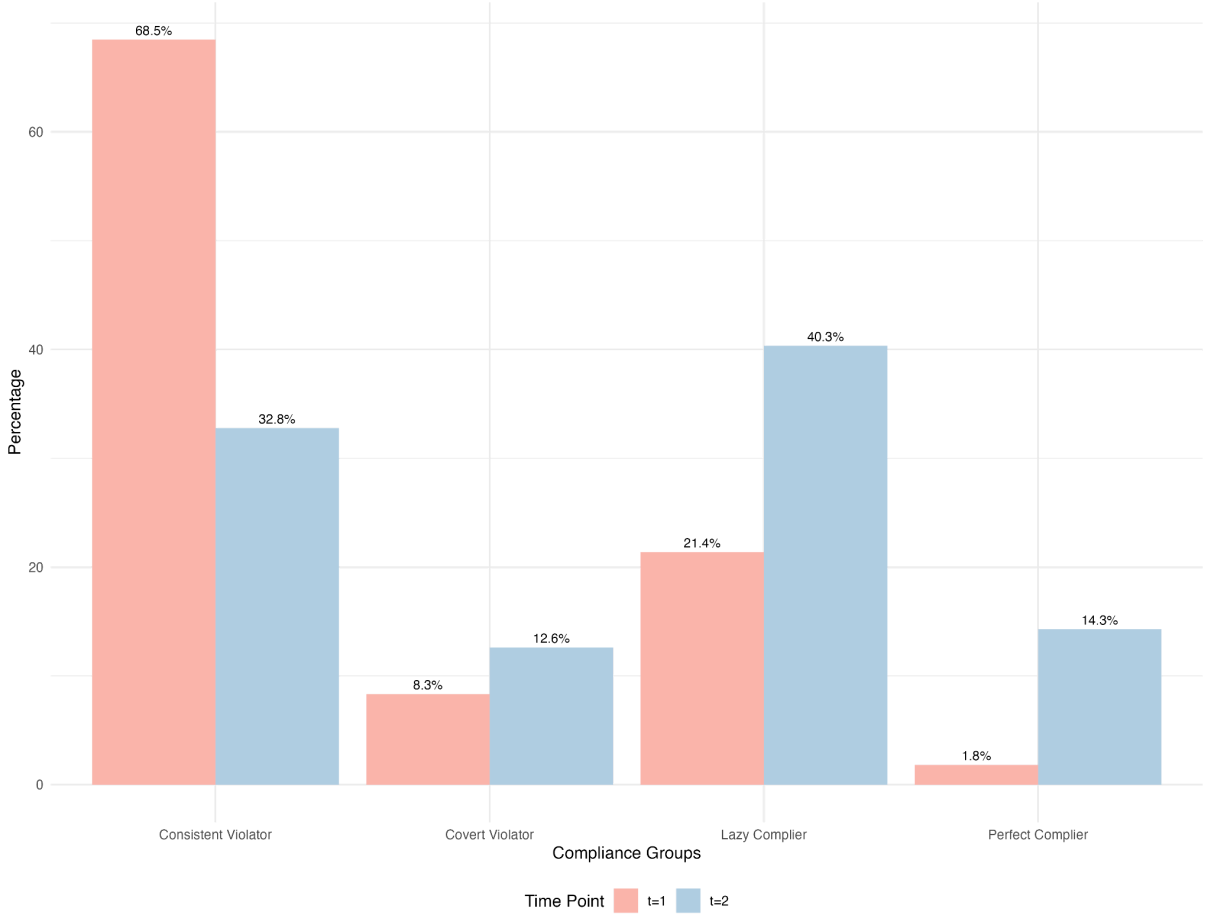
To account for these difference sampling sizes in t=1 and t=2, we focus on the intersection between the two time points, and specifically on the treated apps (i.e., apps that, according to their privacy policy, based their personal data transfer to the US on the Privacy Shield or Standard Contractual Clauses – CBTC-US-Shield).[16] We then identify their behavior when moving ahead in time, according to the four types of apps defined in Table 4 above. In Figure 8, we plot the distribution of treated apps at t=0 and observe their compliance type at t=1 (bars in red). One can observe that at t=1, the largest portion of CTBC-US-Shield apps are consistent violators, i.e. apps that adapted neither their privacy policies nor their personal data flows after Schrems II. We also

[15] For a list of privacy changes in Android 10, *see* https://developer.android.com/about/versions/10/privacy/changes#top-changes.

[16] We focus on the apps which we considered as treated in the previous time point, i.e., apps which indicated in their privacy policies at t=0 or t=1 that they justified personal data transfers to the US with the Privacy Shield or Standard Contractual Clauses as these are the apps which should respond to the Schrems II decision, i.e., treated units.

look at CBTC-US-Shield apps at t=1 and observe their compliance type at t=2 (bars in blue). Two years after the court decision, one can observe that most apps have become either lazy compliers or remain consistent violators.

Figure 8: CBTC-US-Shield Apps, Distribution of Violators & Compliers



Note. The figure visualizes how apps responded to the Schrems II decision in terms of compliance when looking at both privacy policies and personal data flows. All apps displayed are part of the CBTC-US-Shield type (see Table 1) in t=0 or t=1. They are then categorized again based on their compliance type (see Table 4) at t=1 (red bars) and t=2 (blue bars).

We started our journey with an interest in measuring the extent to which apps that transferred personal data to the US complied with the European Court of Justice's Schrems II decision. Consolidating the findings from Figure 8, we measure the ratio of apps whose privacy policies and/or data flows seem to comply with Schrems II and apps which do not. In t=1, the picture of compliance is grim. 68.5% of apps seem to violate the Schrems II decision both in their privacy policies and data transfers (consistent violator), while 8.3% of apps seem to comply with the decision in their privacy policy, yet transfer personal data to the US (covert violator). Even if one follows a lower-bound estimation approach (see our discussion at the end of Section 3.4) and treats lazy compliers as complying with the Schrems II decision,[17] only 23.2% of all apps seem to comply with Schrems II at t=1 (21.4% lazy compliers and 1.8% perfect compliers).

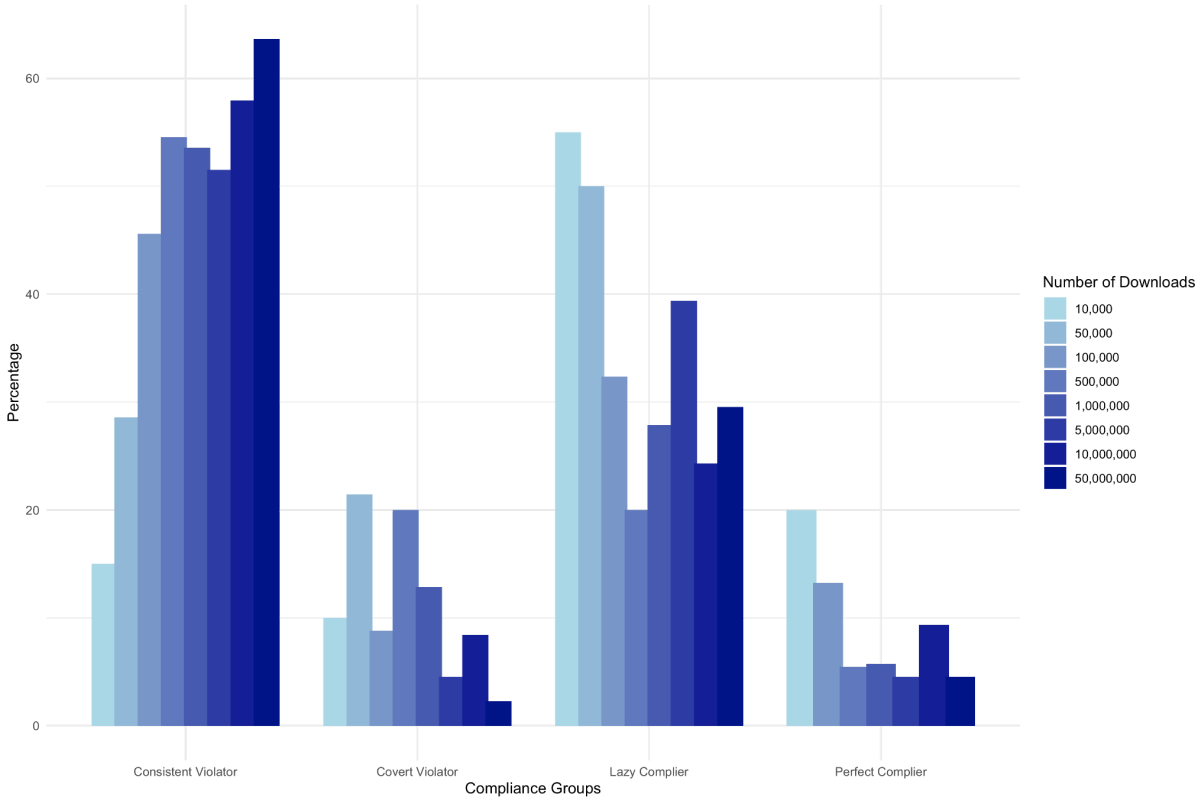
[17] In order not to overestimate non-compliance, let us assume that lazy compliers reacted to Schrems II by stopping the transfer of personal data to the U.S., but were too lazy to adapt their privacy policy accordingly. On the problem of historical holdovers and obsolete contractual provisions in commercial contract production, *see* Scott, Choi & Gulati (2024); Gulati & Scott (2012).

In t=2 (i.e., two years after the court decision), the ratio of apps apparently violating the Schrems II decision in both their privacy policies and data transfers is reduced to 32.8%, while the ratio of apps apparently complying with the decision only in their privacy policy increases to 12.6%. Following our lower-bound estimation approach, 54.6% of all apps seem to comply with Schrems II at t=2 (40.3% lazy compliers and 14.3% perfect compliers). Overall, at least 77% of apps (consistent and cover violators) seem to violate the Schrems II decision a few days after the European Court of Justice published its decision. Two years after the decision, at least 45% of apps still seem to violate the decision. Non-compliance is still high in magnitude.

Finally, in Figure 9 and Table A.2 (Appendix), we report the non-compliance distribution of the apps that were affected by the Schrems II decision, based on their popularity (as proxied by number of downloads in the Spanish Google Play Store). As popularity data is only available for a subsample of the treated units, we merge the observations from the two time points (t=1 & 2), assuming that the relationship between downloads and compliance has not changed between them. Our findings indicate that higher levels of non-compliance are associated with a higher popularity.[18] Some evidence to this interaction was found in the previous studies (Bouhoula et al., 2024). Yet, these results should be interpreted with caution, since the low number of apps in a few groups is not sufficient for statistical inference.

[18] Unlike the California Consumer Privacy Act – which applies to companies with gross revenues exceeding $ 25 million, serving over 100,000 customers and/or deriving at least 50% of their revenues from selling or sharing personal information (Cal. Civ. Code § 1798.140(d)(1) (2018)) – the GDPR has no de minimis thresholds and is also applicable to smaller companies.

Figure 9: Non-Compliance and Popularity



Note. This figure displays CBTC-US-Shield apps in t=0 and t=1 and their distribution to compliance groups (Table 4) based on their privacy policy and personal data flows, at t=1 and t=2, respectively. Light to dark colors indicate popularity (proxied by the number of downloads). The number of downloads is colored from light blue (<=10,000) to dark blue (<=50,000,000).

5. Limitations

In this study, we combine tools from machine-learning and IT security research to develop a framework that allows privacy researchers to assess compliance with the GDPR on a large scale. It should not be surprising that such a framework is subject to various limitations that our various methods have. In this section, we address some of these limitations and discuss next steps.

Focusing on the machine-learning tools to analyze the privacy policies of our apps (Section 3.2), we want to note the levels of recall and precision our classifiers achieved. We presented an extensive legal validation of our classifiers in Section 3.4. Here, we want to note some inherent challenges. Many legal documents mention key concepts throughout the document. In our case, a privacy policy may mention the "Privacy Shield" as part of a cross-border transfer clause, but it may also mention it as an example in some other part of the privacy policy. Rule-based classifiers are typically incapable of grasping the subtle differences between "invoking" and "mentioning as an example". This may pose a risk for our classifier that identifies whether a privacy policy relies on the Privacy Shield as legal basis for a personal data transfer to the US ("Shield-Classifier"). It will interpret a clause that mentions "Privacy Shield" in whatever capacity as a cross-border transfer clause that invokes the Privacy Shield as a legal basis for a cross-border transfer to the US Similar concerns might apply to our classifier which identifies the "US" as a target country

for a personal data transfer in a privacy policy: the US may not only be invoked as a target country in a cross-border transfer clause, but also mentioned as an example in other clauses.

To address some of these challenges, we explored classifying privacy policies using a large language model (LLM), which became available to the wider public as we worked on this paper. We tested GPT (4-0613) using the API offered by OpenAI and our human annotated sample (Table 2). Our focus was on our two main classifiers: the Cross-Border-Transfer-Classifier and the Target-Country-Classifier.

Table 5: Comparing OpenAI GPT versus Our Machine Classification

	Total Number of Clauses	Positives	True Positives	False Positives	False Negatives	Recall (GPT)	Precision (GPT)	Recall (ours)	Precision (ours)
CBTC-Classifier	3807	78	70	25	8	0.89	0.73	0.91	0.72
Target-Country-Classifier	100	52	46	5	6	0.88	0.9	1	0.85

As Table 5 shows, classifying cross-border transfer clauses with GPT produces results that are similar to what our tailored, segment-level classifiers achieved. For both the classification of cross-border transfer clauses and target countries, recall is lower with GPT than with our classifiers, but precision is higher. On the other tasks (i.e., shield classifier and legal basis, see Table 2), ChatGPT performed worse than our classifiers. While our paper does not intend to benchmark tailored classifiers against large language models, we want to note in passing the potential of large language models to make classification tasks more accessible, which could significantly reduce the entry barriers for researchers and data protection agencies to run analyses such as ours.

Focusing on the IT security tools we use to extract and analyze the data traffic leaving our apps (Section 3.3), these tools are prone to false negative classifications. First, while these tools unleash a wealth of data on firm behavior, we cannot be certain that they provide a full picture of the personal data traffic that is relevant for our research question. Sophisticated encryption techniques might enable apps to evade our data flow extraction attacks (Guamàn et al. 2021: 15979). And our methods only allow us to observe the traffic between an app and the first server directly reachable by the app. If an app transfers personal data first to a server located in the EU and then re-routes the data to the US, our methods would not be able to catch the data transfer to the US. We can only note that without the collaboration of the first server to which the app connects or other internet intermediaries, we are not aware of any method that would enable us to observe multiple-hop data transfers. Second, while we rely on the classification of Guamàn et al. (2021) to determine what type of data is personal data, an actual privacy investigation would more carefully consider all circumstances of the case to determine whether a particular data item is “personal data” as defined in Art. 4 (1) GDPR. All in all, we interpret the personal data we are able to observe as a minimum number of personal data flows. In this respect, our results may provide a lower-bound estimation, while the actual number of personal data transfers may be substantially higher.

Focusing on our research design (Section 3.1), while we are inspired by a difference-in-difference design, we do not want to suggest that we have the data for a clean causal identification. While

the timing of the Schrems II decision was semi-random and should not correlate with other major events, our dataset was not designed as a balanced time series which would enable a clean causal identification. The privacy policies at t=0 were collected over a longer time period and were retrieved using the Wayback Machine archive. However, the privacy policies and data flows collected at t=1 – our main point of analysis – occurred shortly after Schrems II. This allows us to study the immediate response in the days following the decision, when the risk of confounding factors is minimal. Measuring the long-term impact of a court decision is more challenging, and the best we could do is to observe privacy policies and data flows over two years after Schrems II (at t=2). Many events might have influenced privacy policies and data flows between t=1 and t=2, including, for example, changes in the Google App Store, the Android operating systems, or enforcement actions by data protection authorities or courts as we explained in Section 4.2. Our findings indicate that a substantial portion of apps still violate Schrems II and the GDPR two years after the court's decision. From a legal perspective, we find such a high level of non-compliance interesting regardless of the causal channel that led to this. While our study does not fully explore differences in compliance levels between different apps or firms, we present some suggestive evidence that more popular apps – as measured by the number of downloads from the Spanish Google Play Store – are less compliant with the Schrems II decision (see Figure 9 and Table A.2).

We would like to note that our framework will, at least in the foreseeable future, not be able to produce statements about individual cases of (non-)compliance that would hold up to legal scrutiny on their own. However, based on our empirical observation, we are rather confident that a substantial fraction of apps do not obey Schrems II even two years after the court decision. More generally, we do not foresee that our framework would enable a data protection authority to automatically generate a complaint against an app provider. Our ambition is not to fully automate the enforcement of digital regulations. Rather, we aim at a framework that will, ideally, provide reliable lower bounds of non-compliance at the aggregate level and help researchers and – ultimately – enforcement agencies to decide where to focus their attention in the digital economy.

6. Discussion

While digital technologies have enabled firms to deliver products and services that are embraced by consumers and businesses throughout the world, they also pose significant challenges for our societies. Society's track record of designing effective regulations for digital markets is mediocre at best. Regulatory frameworks for the digital economy are often designed and discussed in an environment with little empirical evidence at hand. A key reason behind these regulatory failures is that they do not appreciate the sheer scale of the markets they are trying to shape. As firms, platforms and markets grow more complex, they also grow more opaque, hindering a proper understanding of underlying business practices and their impact on society.

All these observations call for a rethinking of how to regulate the digital economy. We envision tools enabling computer scientists, economists and legal scholars to analyze firm and user behavior under digital regulations and empirically identify intended and unintended consequences of such regulations at large scale. With the increasing importance of algorithmic auditing under the European Union's Digital Markets Act, the Digital Services Act and the envisioned AI Act,

such approaches will also gain further prominence in the global debates on how to regulate the digital economy.

In many markets, it is challenging to observe whether market participants play by the rules of the game or not. Digital markets are different. As far as privacy laws are concerned, we can easily observe how firms describe their personal data flows in their privacy policies, which we then can check for compliance with applicable privacy laws. This study has introduced various techniques from IT security research enabling us, in addition, to observe how firms *actually* behave with regard to personal data flows. We hope that our framework, which combines observations on privacy policies with actual personal data traffic, will contribute to an evidence-based design and enforcement of privacy laws and beyond.

In our study, we find that a few weeks after Schrems II was decided, only 23% of the studied apps in our sample seem to comply with Schrems II, while 77% seem to violate the GDPR when contrasting their privacy policy and their actual personal data transfer. Over two years after Schrems II, the rate of compliant apps increases, yet we estimate that roughly 45% of the studied apps remain non-compliant. Several theories may provide (partial) explanations for the high level of non-compliance we observe. First, at least some of the violations we focus on are difficult to trace. Performing man-in-the-middle attacks on modified Android smartphones involve skills and tools that may not be available to the average privacy agencies these days. Thereby, firms may be less concerned about complying with the GDPR in their personal data flows, which reduces potential deterrent effects. Second, while the Irish Data Protection Commission imposed a € 1.3 billion fine on Meta in 2023, issuing such high-profile fines for violating Schrems II has remained an exception so far. This might also mitigate deterrent effects. Even under the predecessor of the Privacy Shield (the Safe Harbour Agreement), firm compliance was low (Dhont et al. 2004; Marotta-Wurgler & Svirsky 2023). Low compliance with transatlantic privacy rules has haunted EU-US trade for more than two decades (see also Lancieri 2022).

We want to note that our study is also amenable to a political economy interpretation. As described in Section 2.1, the European Court of Justice invalidated the Privacy Shield in its Schrems II decision on July 16, 2020 with immediate effect and no grace period, leaving firms in legal turmoil. While privacy authorities published some guidance in the following weeks and months, such guidance did not provide firms with a clear and easily implementable path towards compliance. When the European Court of Justice invalidated the precursor of the Privacy Shield – the Safe Harbour Agreement – on October 6, 2016 in its Schrems I decision (European Court of Justice 2015), it also invalidated the agreement with immediate effect and no grace period. However, privacy agencies were more accommodating to firms' concerns after that decision. Ten days after Schrems I was decided, the group of European privacy agencies (then called "Article 29 Working Party") issued a statement urgently calling EU institutions to create a successor to the Safe Harbor Agreement and announcing a de facto 3.5 months grace period during which privacy agencies would not enforce privacy laws against firms violating Schrems I (Article 29 Working Party 2015). After Schrems II, privacy agencies did not announce such a grace period, but urged firms to become compliant without showing a clear path on how to do so. When the European Court of Justice declared most personal data transfers to the US invalid and neither the court nor privacy agencies offered any grace period or clear guidance on how to become compliant in the short term, these institutions seemed to have accepted that a substantial portion of firms would violate the GDPR with their data transfer practices at least in the short term: it is either too complex or too costly for firms to cut their personal data transfers to the US, at least in the short

term. One interpretation why the court and privacy agencies acted in this manner is that they wanted to increase the political pressure on EU institutions and the US government to find a legally and politically stable solution to a problem that has haunted EU-US relations over the last quarter of a century. According to this interpretation, judicial actors accept massive violations of the GDPR in order to put pressure on political actors to find a political solution. We started our endeavor by asking who listens to the court when the court speaks. But we want to suggest that the real question after Schrems II may not be who listens to the court, but to whom the court speaks.

These findings also challenge "carrot and stick" theories of the diffusion of EU laws (Bradford 2012). One of the conditions for the successful diffusion of EU laws is the capacity of EU institutions to enforce laws. As Anu Bradford argues, a country can only exert authority within or outside its jurisdiction if it has sufficient regulatory expertise and capacities to enforce its laws. The authority and ability to impose sanctions in case of noncompliance is an important component of regulatory capacities, as it enables the country to exclude non-complying firms from its market (Bradford 2012, 12-13). The level of non-compliance we observe in this paper, as well as other findings concerning the GDPR and EU privacy laws (Li, et al. 2023), suggests that many apps and websites perceive the risk of enforcement by EU institutions to be low and continue their business as usual. This challenges the notion that the EU signals sufficient regulatory capacity to enforce its law and yet diffusion of EU privacy laws is well documented. While European privacy authorities are increasingly imposing large fines against big tech companies, this might signal selective regulatory capacity at best, hinting that diffusion can occur despite low enforcement capacity, maybe with a strong facade of enforcement against global giants. We leave the empirical investigation of such selective signals to future work.

Lastly, we hope that frameworks such as ours will help legal researchers and regulatory agencies in the long run to understand the intricate relationship between digital regulations and the digital economy. Employing methods from machine-learning and IT security research, we anticipate that the legal and regulatory discourse will develop a better understanding of how digital regulations interact with the millions of firms and consumers operating under these regulations. And we hope that this will help to prepare the discourse when the complexity of digital regulations increases. As Chander & Schwartz (2023) point out, the European Union is not the only regulator that governs cross-border transfers of personal data. Nowadays, more than sixty countries evaluate whether foreign countries provide for an adequate privacy protection. In this study, we have explored the extent to which techniques from machine learning and IT security research can help us understand compliance with one single regulation. We leave the question on how to scale this to 60 regulating entities, which interact and sometimes overlap with each other, to future research.

References

Android Studio (2023). UI/Application Exerciser Monkey. https://developer.android.com/studio/test/other-testing-tools/monkey.

Article 29 Data Protection Working Party (2015). Statement of the Article 29 Working Party, Oct. 16, 2015. https://ec.europa.eu/justice/article-29/press-material/press-release/art29_press_material/2015/20151016_wp29_statement_on_schrems_judgement.pdf.

Article 29 Data Protection Working Party (2018). Guidelines on Transparency Under Regulation 2016/679. https://edpb.europa.eu/system/files/2023-09/wp260rev01_en.pdf.

Baker, S. (2020). How Can the US Respond to Schrems II? Lawfare, July 21, 2020. https://www.lawfaremedia.org/article/how-can-us-respond-schrems-ii.

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. Journal of Political Economy 76, 169-217.

Bollinger, D., Kubicek, K., Cotrini, C., & Basin, D. (2022). Automating Cookie Consent and GDPR Violation Detection. 31st USENIX Security Symposium (USENIX Security 2022), 2893-2910.

Bouhoula, A., Kubicek, K., Zac, A., Cotrini, C., & Basin, D. (2024). Automated, Large-Scale Analysis of Cookie Notice Compliance. Forthcoming In USENIX Security Symposium.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A Training Algorithm for Optimal Margin Classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, 144-152.

Bradford, A. (2012). The Brussels Effect. Northwestern University Law Review 107, 1-68.

Bradford, A. (2020). The Brussels Effect: How the European Union Rules the World. Oxford University Press.

Bradford, A. (2023). Digital Empires: The Global Battle to Regulate Technology. Oxford University Press.

Chander, A., & Schwartz, P. (2022). Privacy and/or Trade. University of Chicago Law Review 90, 94-136.

Choudhary, S. R., Gorla, A., & Orso, A. (2015). Automated Test Input Generation for Android: Are We There Yet? Proceedings of the 30th IEEE/ACM International Conference on Automated Software Engineering (ASE), 429-440.

Commissioner for Data Protection and Freedom of Information for the German State of Berlin (2020). Nach “Schrems II”: Europa braucht digitale Eigenständigkeit. Press release 711.424 (July

17, 2020). https://www.datenschutz-berlin.de/fileadmin/user_upload/pdf/pressemitteilungen/2020/20200717-PM-Nach_SchremsII_Digitale_Eigenstaendigkeit.pdf.

Cozar, M., Rodriguez, D., Del Alamo, J. M., & Guaman, D. (2022, June). Reliability of IP geolocation services for assessing the compliance of international data transfers. In 2022 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW) (pp. 181-185). IEEE.

Davis, K. E., & Marotta-Wurgler, F. (2019). Contracting for Personal Data. New York University Law Review 94, 662-705.

Davis, K. E., & Marotta-Wurgler, F. (2023). Filling the Void: How E.U. Privacy Law Spills Over to the US Manuscript.

Dhont, J. Pérez Asinari, M. V., Poullet, Y. (2004). Safe Harbour Decision Implementation Study. https://www.europarl.europa.eu/meetdocs/2009_2014/documents/libe/dv07_etude_safe-harbour-2004_/07_etude_safe-harbour-2004_en.pdf.

European Commission (2000). Commission Decision of 26 July 2000 Pursuant to Directive 95/46/EC of the European Parliament and of the Council on the Adequacy of the Protection Provided by the Safe Harbour Privacy Principles and Related Frequently Asked Questions Issued by the US Department of Commerce. Official Journal of the European UNion L 215, August 25, 2000, 7.

European Commission (2010). Commission Decision of 5 February 2010 on Standard Contractual Clauses for the Transfer of Personal Data to Processors Established in Third Countries under Directive 95/46/EC of the European Parliament and of the Council. Official Journal of the European Union L 39, February 12, 2010, 5.

European Commission (2016a). Commission Implementing Decision (EU) 2016/1250 of 12 July 2016 pursuant to Directive 95/46/EC of the European Parliament and of the Council on the Adequacy of the Protection Provided by the EU-US Privacy Shield. Official Journal of the European Union L 207, August 8, 2016, 1.

European Commission (2016b). Commission Decision 2010/87/EU of 5 February 2010 on standard contractual clauses for the transfer of personal data to processors established in third countries under Directive 95/46 (Official Journal of the European Union 2010 L 39, 5), as amended by Commission Implementing Decision (EU) 2016/2297 of 16 December 2016. Official Journal of the European Union 2016 L 344, 100.

European Commission (2021). Commission Implementing Decision (EU) 2021/914 of 4 June 2021 on Standard Contractual Clauses for the Transfer of Personal Data to Third Countries Pursuant to Regulation (EU) 2016/679 of the European Parliament and of the Council. Official Journal of the European Union 2021 L 199, 31.

European Commission (2022). Commission Implementing Decision of XXX Pursuant to Regulation (EU) 2016/679 of the European Parliament and of the Council on the Adequate Level of Protection of Personal Data Under the EU-US Data Privacy Framework.

https://commission.europa.eu/system/files/2022-12/Draft%20adequacy%20decision%20on%20EU-US%20Data%20Privacy%20Framework_0.pdf.

European Commission (2023). Commission Implementing Decision of 10.7.2023 Pursuant to Regulation (EU) 2016/679 of the European Parliament and of the Council on the Adequate Level of Protection of Personal Data under the EU-US Data Privacy Framework. C(2023) 4745 final.

European Court of Justice (2015). Maximilian Schrems vs. Data Protection Commissioner. Case C-362/14, ECLI:EU:C:2015:650.

European Court of Justice (2020). Data Protection Commissioner v. Facebook Ireland Ltd. Case C-311/18, ECLI:EU:C:2020:559.

European Data Protection Board (2020a). Statement on the Court of Justice of the European Union Judgment in Case C-311/18 – Data Protection Commissioner v Facebook Ireland and Maximillian Schrems. https://edpb.europa.eu/news/news/2020/statement-court-justice-european-union-judgment-case-c-31118-data-protection_en.

European Data Protection Board (2020b). Frequently Asked Questions on the Judgment of the Court of Justice of the European Union in Case C-311/18 – Data Protection COmmissioner v. Facebook IReland Ltd. and Maximilian Schrems. https://edpb.europa.eu/our-work-tools/our-documents/other/frequently-asked-questions-judgment-court-justice-european-union_en.

European Data Protection Board (2020c). Recommendations 01/2020 on Measures that Supplement Transfer Tools to Ensure Compliance with the EU Level of Protection of Personal Data. https://edpb.europa.eu/our-work-tools/our-documents/recommendations/recommendations-012020-measures-supplement-transfer_en.

European Data Protection Board & European Data Protection Supervisor (2021). EDPB – EDPS Joint Opinion 2/2021 on the European Commission's Implementing Decision on Standard Contractual Clauses for the Transfer of Personal Data to Third Countries for the Matters Referred to in Article 46(2)(c) of Regulation (EU) 2016/679. https://edpb.europa.eu/our-work-tools/our-documents/edpbedps-joint-opinion/edpb-edps-joint-opinion-22021-standard_en.

European Data Protection Board (2023). Binding Decision 1/2023 on the Dispute Submitted by the Irish SA on Data Transfers by Meta Platforms Ireland Limited for its Facebook Service (Art. 65 GDPR) https://edpb.europa.eu/our-work-tools/our-documents/binding-decision-board-art-65/binding-decision-12023-dispute-submitted_en.

European Union (1995). Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals With Regard to the Processing of Personal Data and on the Free Movement of Such Data. Official Journal of the European Communities L 281, November 23, 1995, 31.

Frankenreiter, J. (2022). Cost-Based California Effects. Yale Journal on Regulation 39, 1155-1217.

Guamán, D. S., del Alamo, J. M., & Caiza, J. C. (2021). GDPR Compliance Assessment for Cross-Border Personal Data Transfers in Android Apps. IEEE Access 9, 15961-15982.

Guamán, D. S., Rodriguez, D., del Alamo, J. M., & Such, J. (2023). Automated GDPR Compliance Assessment for Cross-border Personal Data Transfers in Android Applications. Computers & Security 130, 103262.

Gulati, M., & Scott, R. (2012). The Three and a Half Minute Transaction: Boilerplate and the Limits of Contract Design. University of Chicago Press.

Harnett, L., Rodgers, D., & Newton, C. (2023). Largest GDPR Fine to Date: DPC Penalises Meta Once More, But Did They Ever Have a Choice? https://www.lexology.com/library/detail.aspx?g=d21b5a94-bed6-4dc4-88ee-f8d443acea7c.

Internet Archive (2023). Wayback Machine. https://archive.org/web.

Irish Data Protection Commission (2023). In the Matter of Meta Platforms Ireland Limited. IN-20-8-1. https://edpb.europa.eu/system/files/2023-05/final_for_issue_ov_transfers_decision_12-05-23.pdf.

Jack, M. C., Sovannaroth, P., & Dell, N. (2019). ”Privacy is Not a Concept, But a Way of Dealing With Life”: Localization of Transnational Technology Platforms and Liminal Privacy Practices in Cambodia. Proceedings of the ACM on Human-Computer Interaction, 3 (CSCW), 1-19.

Kampanos, G., & Shahandashti, S. F. (2021). Accept All: The Landscape of Cookie Banners in Greece and the UK. In A. Jøsang, L. Futcher, & J. Hagen (eds.), ICT Systems Security and Privacy Protection, 36th IFIP TC 11 International Conference (SEC 2021), Springer, 213-227.

Kubíček, K., Merane, J., Cotrine, C., Stremitzer, A., Bechtold, S., & Basin, D. (2022). Checking Websites’ GDPR Consent Compliance for Marketing Emails. Proceedings on Privacy Enhancing Technologies, 2, 282-303.

Kuner, C. (2023). Protecting EU Data Outside EU Borders Under the GDPR. Common Market Law Review 66, 77-106.

Lancieri, F. (2022). Narrowing Data Protection's Enforcement Gap. Maine Law Review 74, 15-72.

Li, W., Li, Z., Li. W., Zhang, Y. & Li, A. (2023). Mapping the Empirical Evidence of the GDPR’s In-Effectiveness: A Systematic Review. http://ssrn.com/abstract=4615186.

Libert, T. (2018). An Automated Approach to Auditing Disclosure of Third-party Data Collection in Website Privacy Policies. In Proceedings of the 2018 World Wide Web Conference, 207-216.

Linden, T., Khandelwal, R., Harkous, H., & Fawaz, K. (2020). The Privacy Policy Landscape After the GDPR. Proceedings on Privacy Enhancing Technologies, 1, 47-64.

Maas, M., Stöver, A., Pridöhl, H., Bretthauer, S., Herrmann, D., Hollick, M., & Spiecker, I. (2021). Effective Notification Campaigns on the Web: A Matter of Trust, Framing, and Support. Proceedings of the 30th USENIX Security Symposium (USENIX Security 2021), 2489-2506.

Marotta-Wurgler, F. (2016). Self-Regulation and Competition in Privacy Policies. Journal of Legal Studies 45, 13-39.

Marotta-Wurgler, F., & Svirsky, D. (2023). Do FTC Privacy Enforcement Actions Matter? Compliance Before and After US-EU Safe Harbor Agreement Actions. Manuscript.

Matte, C., Bielova, N., & Santos, C. (2020). Do Cookie Banners Respect my Choice? Measuring Legal Compliance of Banners from IAB Europe's Transparency and Consent Framework. Proceedings of the IEEE Symposium on Security and Privacy 2020, 791-809.

Meta Platforms, Inc. (2023). Annual Report (Form 10-K) for the fiscal year ended Dec. 31, 2022. https://www.sec.gov/Archives/edgar/data/1326801/000132680123000013/meta- 20221231. htm.

Murphy, M. H. (2022). Assessing the Implications of Schrems II for EU-US Data Flow. International & Comparative Law Quarterly Vol. 71, 245-262.

Naef, T. (2023). Data Protection without Data Protectionism: The Right of Protection of Personal Data and Data Transfers in EU Law and International Trade Law. Springer.

Nouwens, M., Liccardi, I., Veale, M., Karger, D., & Kagal, L. (2020). Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence. Proceedings of the Conference on Human Factors in Computing Systems, 194.

Peukert, C., Bechtold, S., Batikas, M., & Kretschmer, T. (2022). Regulatory Spillovers and Data Governance: Evidence from the GDPR. Marketing Science 41, 318-340.

Ross. W. (2020). US Secretary of Commerce Wilbur Ross Statement on Schrems II Ruling and the Importance of EU–US Data Flows, July 16, 2020. https://useu.usmission.gov/u-s-secretary-of-commerce-wilbur-ross-statement-on-schrems-ii-ruling-and-the-importance-of-eu-u-s-data-flows.

Sanchez-Rola, I., Dell'Amico, M., Kotzias, P., Balzarotti, D., Bilge, L., Vervier, P.-A., & Santos, I. (2019). Can I Opt Out Yet? GDPR and the Global Illusion of Cookie Control. Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security, 340-351.

Satarino, A. (2018). G.D.P.R., a New Privacy Law, Makes Europe World's Leading Tech Watchdog. New York Times, May 25, 2018, Section A, p. 1. https://www.nytimes.com/2018/05/24/technology/europe-gdpr-privacy.html.

Scott, R., Choi, S., & Gulati, M. (2024). Commercial Boilerplate: A Review and Research Agenda, forthcoming in the Annual Review of Law and Social Science. http://ssrn.com/abstract=4697536.

Senol, A., Acar, G., Humbert, M., & Zuiderveen Borgesius, F. (2022). Leaky Forms: A Study of Email and Password Exfiltration Before Form Submission. Proceedings of the 31st USENIX Security Symposium (USENIX Security 22), Boston, MA, 1813-1830.

Smith, O. (2018). The GDPR Racket: Who's Making Money From This $9bn Business Shakedown? Forbes, May 2, 2018, https://www.forbes.com/sites/oliversmith/2018/05/02/the-gdpr-racket-whos-making-money-from-this-9bn-business-shakedown.

Stout, K. (2020). EU Data Transfer Laws Might Destroy Transatlantic Commerce. The Hill, Oct. 26, 2020. https://thehill.com/blogs/congress-blog/technology/522805-eu-data-transfer-laws-might-destroy-transatlantic-commerce.

Taylor, M. (2023). Transatlantic Jurisdictional Conflicts in Data Protection Law: Fundamental Rights, Privacy and Extraterritoriality. Cambridge University Press.

US Department of Commerce (2016). EU-US Privacy Shield Framework Principles. Published as Annex II to European Commission (2016a).

Warberg, L., Lefrere, V., Cheyre, C., & Acquisti, A. (2023, November). Trends in Privacy Dialog Design after the GDPR: The Impact of Industry and Government Actions. In Proceedings of the 22nd Workshop on Privacy in the Electronic Society, 107-121.

White House (2022). United States and European Commission Announce Trans-Atlantic Data Privacy Framework. https://www.whitehouse.gov/briefing-room/statements-releases/2022/03/25/fact-sheet-united-states-and-european-commission-announce-trans-atlantic-data-privacy-framework/.

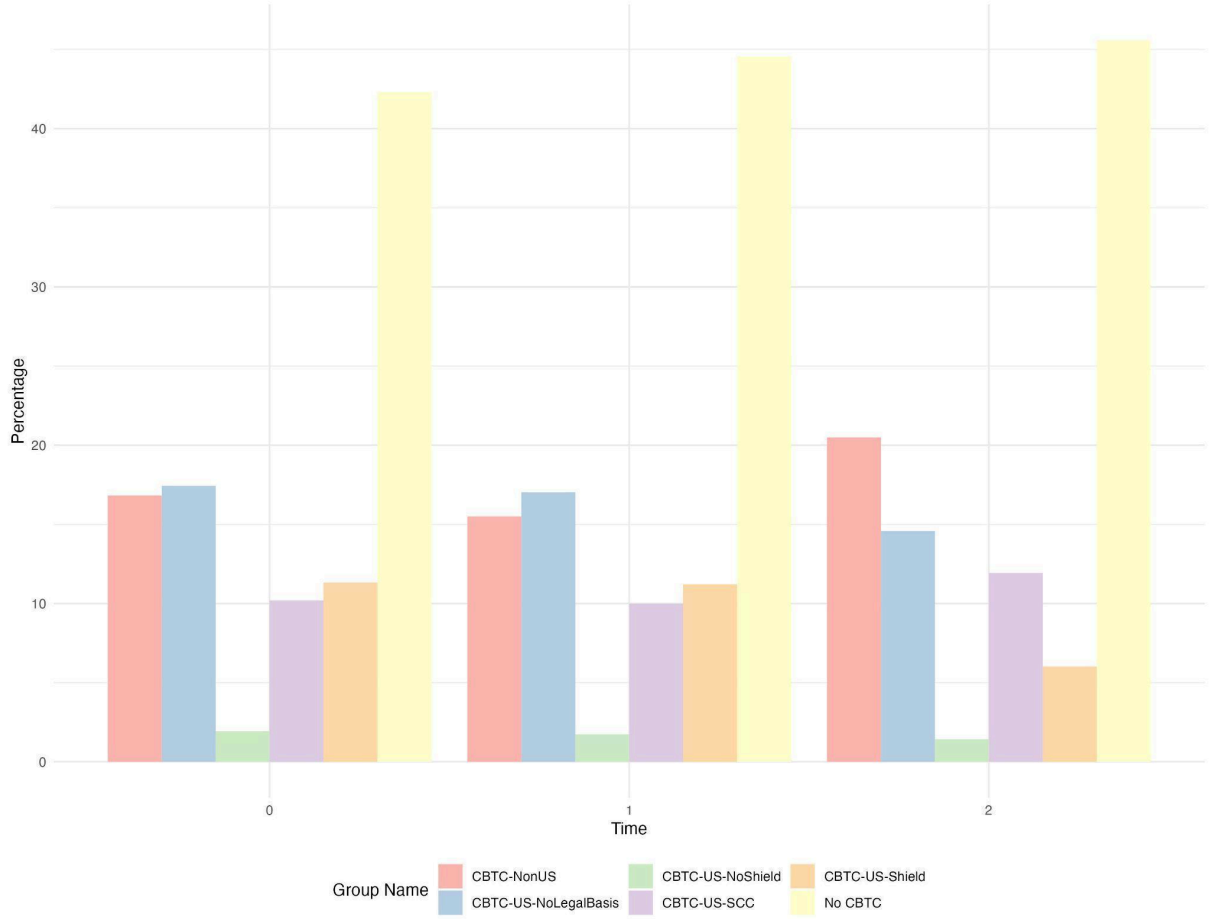
Appendix

Table A.1 Types of Personal Data Observed

Category	Data Type	Description
Contact	Contact_Address_Book	Contact data from the data subject's address book
	Contact_Name	Data subject's name
	Contact_City	Data subject's city
	Contact_Phone_Number	Data subject's phone number
	Contact_Postal_Address	Data subject's postal address
	Contact_ZIP	Data subject's ZIP code
Demographic	Demographic_Age	Data subject's age (including age range)
	Demographic_Gender	Data subject's gender
Identifiers	Identifier_Ad_ID	Google Advertising ID
	Identifier_Cookie_or_similar_Tech	Mobile device fingerprint
	Identifier_Device_ID	Android ID
	Identifier_IMEI	IMEI (International Mobile Equipment Identity)
	Identifier_IMSI	IMSI (International Mobile Subscriber Identity)
	Identifier_MAC	MAC address
	Identifier_SIM_Serial	SIM serial number
	Identifier_SSID_BSSID	Wi-Fi SSID or BSSID
Location	Location_Bluetooh	Bluetooth-based location data
	Location_GPS	GPS location data
	Location_WiFi	Wi-Fi-based location data

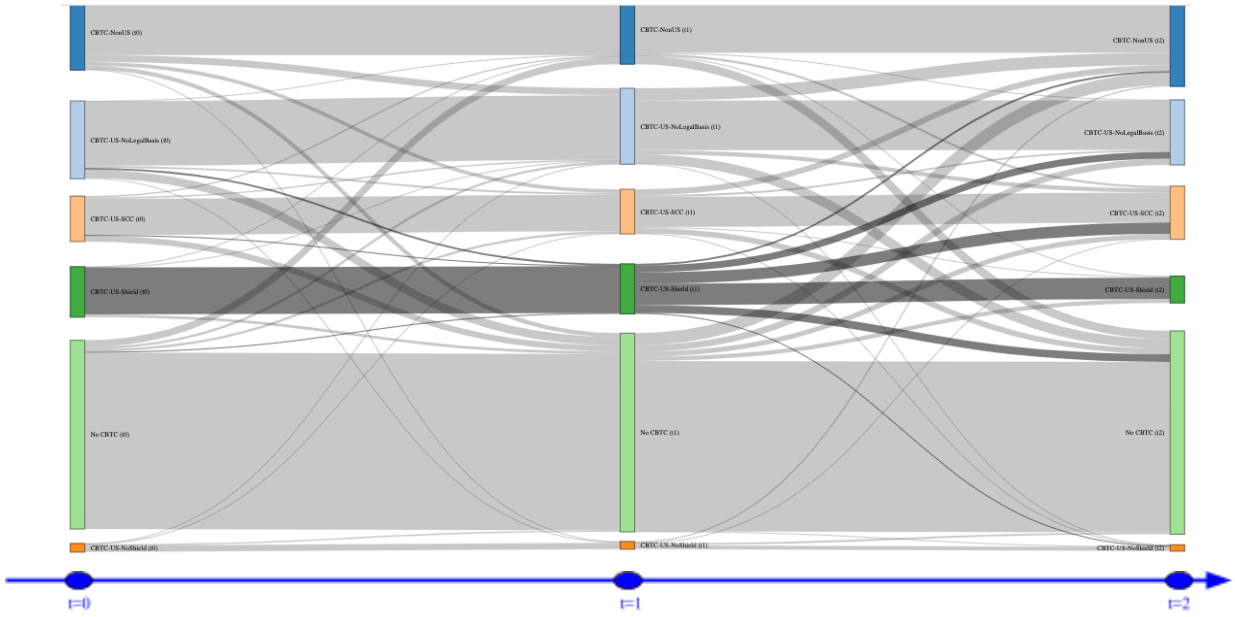
Note: This table presents the type of personal data captured by our traffic data flow analysis. The table is reproduced from Guamàn et al. (2021): 15965.

Figure A.1: Distribution of Apps with Complete Information Over CBTC Types, Distinguishing Between Privacy Shield and Standard Contractual Clauses



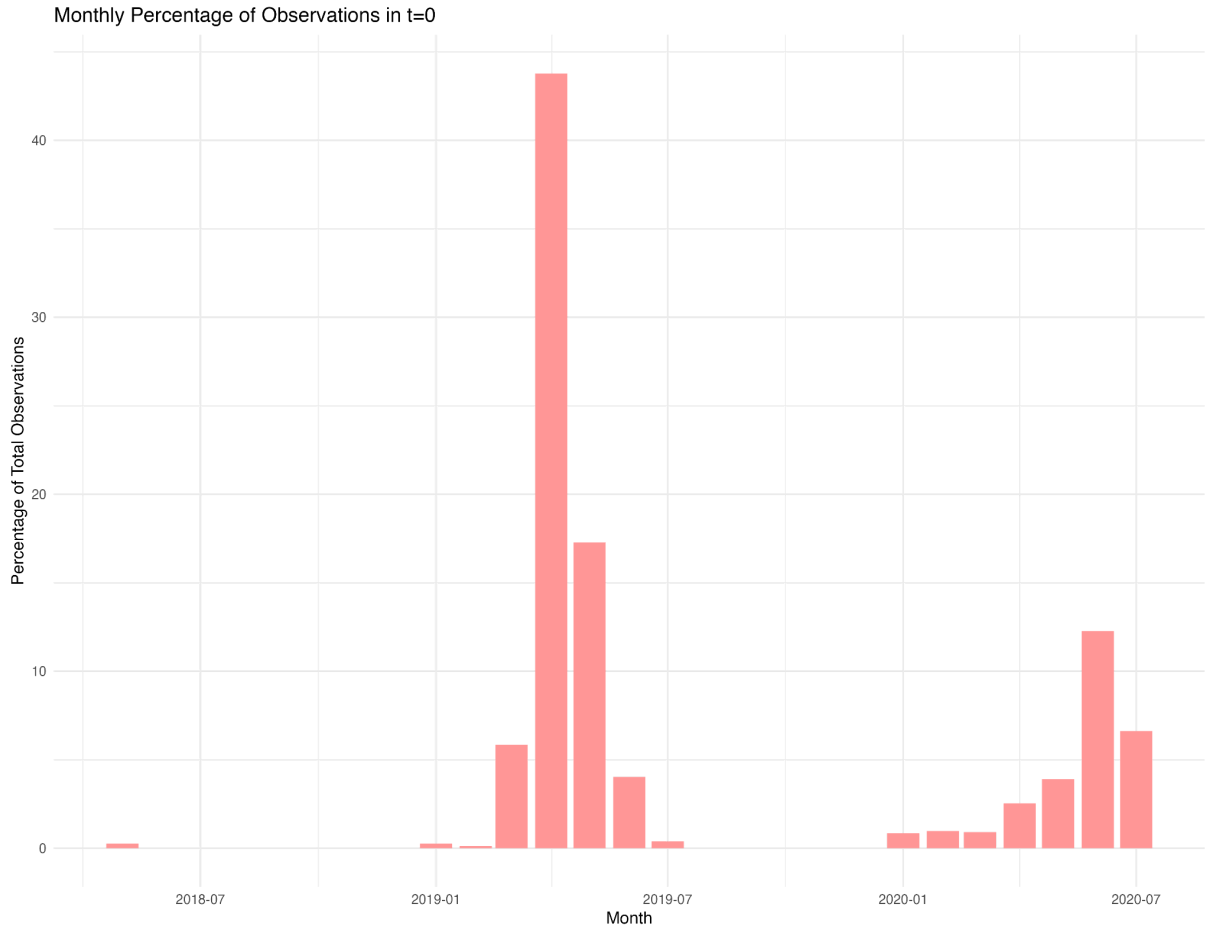
Note. This figure replicates Figure 4 in the main text, but splits the CBTC-US-Shield group into two groups: one group (CBTC-US-Shield) includes privacy policies that rely on the Privacy Shield as the sole legal basis for third-country transfers, while the other group (CBTC-US-SCC) includes privacy policies that rely on Standard Contractual Clauses as the sole legal basis or on Standard Contractual Clauses and the Privacy Shield as the legal basis for third-country transfers.

Figure A.2: Sankey Diagram on Movements Between CBTCs
SCC As a Separate Group



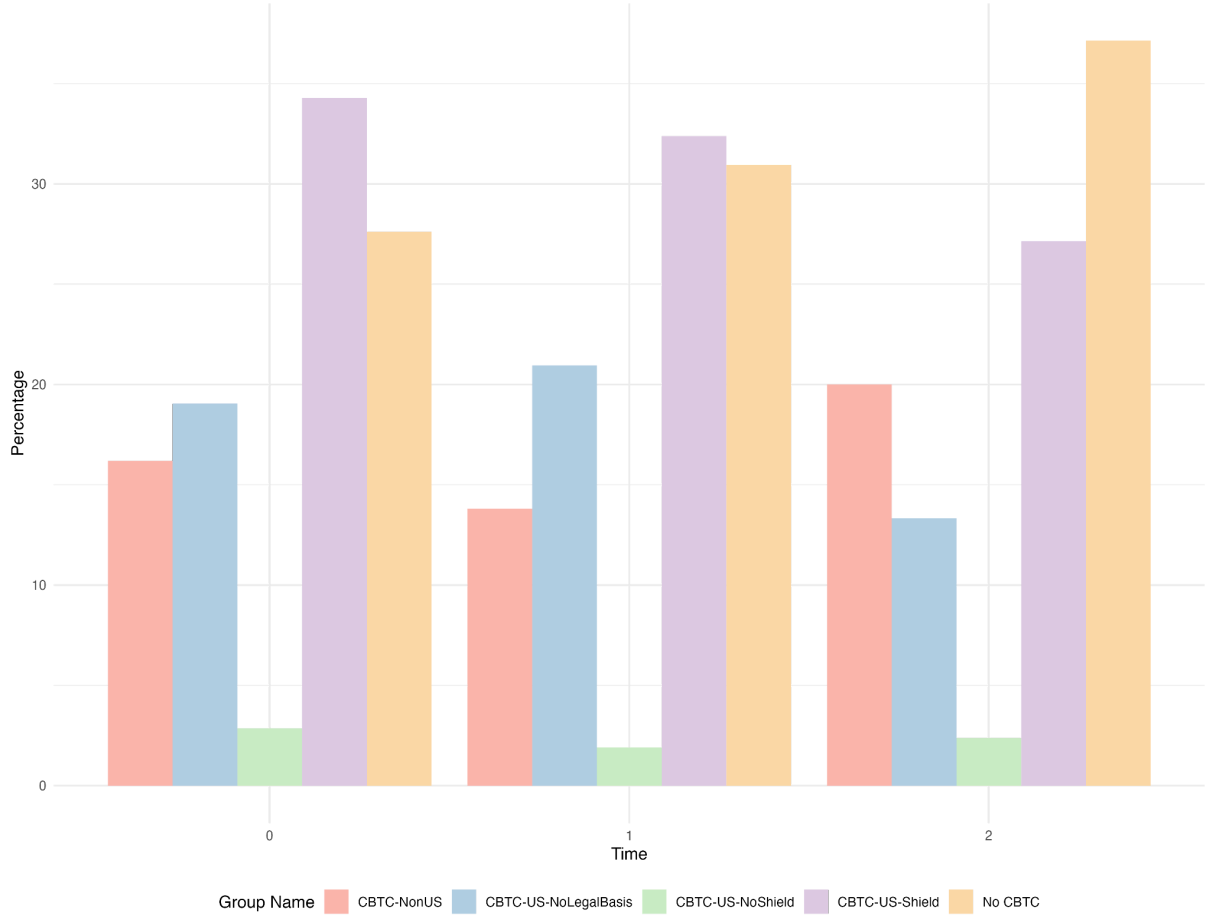
Note. The figure visualizes how app privacy policies move from one type of cross-border transfer clause (see Table 1) to another type over time. For each type of cross-border transfer clause, the movement starts on the left with the first vertical bar depicting t=0. The second, middle vertical bar depicts movement at t=1, while the vertical bar on the right end of the figure depicts movement at t=2. This figure replicates Figure 5 in the main analysis separating the CBTC-US-Shield group to two groups, one which only includes the privacy shield as a legal basis, and CBTC-US-SCC, which includes policies that rely on Standard Contractual Clauses (with or without the privacy shield). It confirms the main findings. Privacy policies using the CBTC-US-Shield only respond in the later stages, and still a large portion of the apps continue to include the shield as the only legal basis.

Figure A.3: Distribution of Privacy Policies Collected in t=0



Note. The figure visualizes the percentage of privacy policies in t=0 we retrieved from the Wayback Machine, aggregated per month, between May 2018 and July 2020. Around 30% of the observations in our t=0 dataset come from 2020 (i.e.,in the six months before the Schrems decision) and the overwhelming majority of the policies are from 2019 and 2020.

Figure A.4: Distribution of Apps with Complete Information Over CBTC Types, Between January and July 2020

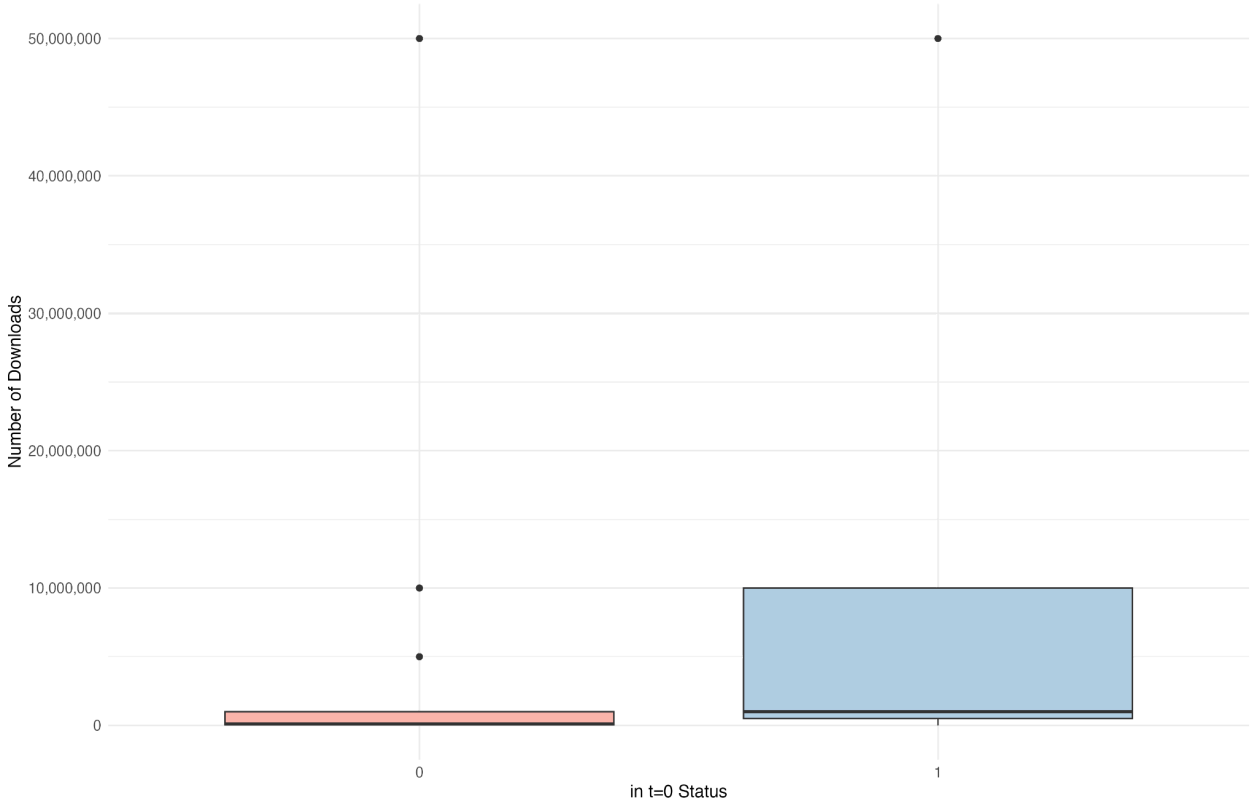


Note. This Figure replicates Figure 4 in the main text, but limits the observation in t=0 to the period between January 1 and July 15,2020. The total size of the intersection between all three time points is 210.

Table A.2: Non-Compliance and Popularity

Number of Downloads	Consistent Violators	Covert Violators	Lazy Compliers	Perfect Compliers	Total Number of Apps
10,000	3 (15%)	2 (10%)	11 (55%)	4 (20%)	20
50,000	4 (28.57%)	3 (21.43%)	7 (50%)	0 (0%)	14
100,000	31 (45.59%)	6 (8.82%)	22 (32.35%)	9 (13.24%)	68
500,000	30 (54.55%)	11 (20%)	11 (20%)	3 (5.45%)	55
1,000,000	75 (53.57%)	18 (12.56%)	39 (27.56%)	8 (5.71%)	140
5,000,000	34 (51.52%)	3 (4.55%)	26 (39.39%)	3 (4.55%)	66
10,000,000	62 (57.94%)	9 (8.41%)	26 (24.30%)	10 (9.35%)	107
50,000,000	28 (63.64%)	1 (2.27%)	13 (29.55%)	2 (4.55%)	44

Figure A.5: Distribution of Apps's Downloads in t=1 by the Status in t=0



Note. The boxplot visualizes the distribution of the number of downloads for apps in t=1, based on their status in t=0. Apss that are included in t=0 (retrieved from the Wayback Machine) on the blue bar (right) show a higher number of downloads on average than the red bar (left). The boxplot in figure represent the 25, 50 (median in bold) and 75 quartiles.

.