Incomplete Contracts and Future Data Usage

Jens Frankenreiter
Washington University in St. Louis

Talia Gillis
Columbia University

Dan Svirsky
Uber

May 2022

Abstract

Most major jurisdictions require websites to provide customers with privacy policies. While privacy policies' most important function is to provide consumers with a description of online service providers' current privacy practices, we argue that these policies also serve a second, often-overlooked function: They allocate among online services and consumers the power to decide whether a service can modify its privacy practices and use consumer data in novel ways. We furthermore argue that a central feature of the E.U.'s General Data Protection Regulation (GDPR), one of the most comprehensive and far-reaching privacy regulatory regimes, is to restrict privacy policies that allocate broad rights for future data usage to service providers. We provide a theoretical explanation for this type of regulatory intervention by adapting standard models of incomplete contracts to privacy policies. We then use the model to explain how U.S. firms reacted to the GDPR. We show that U.S. websites with E.U. exposure are more likely to change their U.S. privacy policies to have *less* stringent and more lenient modification rules. Among websites that do not have E.U. exposure, we see the opposite trend. These results suggest that websites sought to obtain residual rights over data usage after learning of the impact of the GDPR's stricter requirements.

Contents

1 Introduction

Privacy policies are ubiquitous on the internet. While jurisdictions around the world differ in their approaches to regulating data privacy, they are united in requiring websites to make privacy policies available to consumers. The most important function of these policies is to provide consumers with a description of online service providers' current privacy practices. They describe the extent to which these providers collect user's data and how they subsequently use this information. Most major jurisdictions hold online service providers accountable for breaches of their own privacy policies, rendering these policies a functional equivalent of consumer contracts governing other parts of the relationship between online service providers and the consumers they serve.[1]

In this paper, we argue that an important and often-overlooked feature of privacy policies is that they allocate among online services and consumers the power to decide whether a service can change its privacy practices and use consumer data in novel ways. This power is relevant, for example, if a provider considers rolling out new product features that require the collection or processing of consumer data in ways that are different from prior data usages. As an example, consider a social network deciding whether to introduce facial recognition technology to identify individuals in pictures uploaded by its users.[2] Will the introduction of this feature require a modification of the service provider's current privacy policy? If yes, does the consumer need to agree to this modification? The answer to both questions will, to a large extent, depend on the language of the current privacy policy.

We furthermore argue that a central feature of the E.U.'s General Data Protection Regulation (GDPR) and similar data privacy regulations is to protect consumers against privacy policies that allocate broad rights for future data usage to service providers. The GDPR requires privacy policies to explicitly describe the various ways in which service providers collect and use data. In other words, the GDPR contains a prohibition against general clauses that permit broad ranges of uses. In addition, while the GDPR allows for changes to a service provider's privacy policy, every new data collection or use requires a legal justification to be permissible. One potential justification is consumer consent. This means that a service provider cannot reserve for itself the right to unilaterally enact meaningful changes to its privacy policy.

The GDPR places few substantive limitations on the content of privacy policies and focuses primarily on what constitutes adequate notice and consumer agreement to policy terms. This is despite the growing recognition that privacy harms can occur outside the relationship between websites and users (Viljoen 2021), raising questions about whether the fact that consumers voluntarily engage in these interactions is enough to guarantee socially optimal outcomes. In this paper we consider privacy policies through the predominant regulatory lens of bilateral exchange between websites and users, although we recognize that this is only a partial consideration of privacy harms, which is complex and multi-faceted.

Yet the GDPR's approach to future data usage rights is puzzling even when privacy policies are considered through a narrow bilateral lens of an exchange between websites and users. The GDPR's limits on broad allocations of rights to future data usage to service providers are difficult to reconcile with either *rational*

1. There is an ongoing debate in the legal literature about whether courts should treat these policies as contracts. Some scholars argue that enforcing privacy polices would provide consumers with necessary protections, particularly with respect to data breaches (Matwyshyn 2013). Others are skeptical about a contractual approach to privacy polices (Karnow 2021). A similar disagreement exists over whether courts are *de facto* treating privacy policies as contracts. Some argue that courts tend to view privacy policies as "mere notices of company practice" (Solove and Schwartz 2019). By contrast, a survey of privacy disputes that reached court judgements from 2016 suggests that in most cases the courts treated privacy policies as contracts (Bar-Gill, Ben-Shahar, and Marotta-Wurgler 2017; Shahar and Strahilevitz 2016).

2. Facebook has used facial recognition technology in this way since 2010.

or *behavioral* standard accounts of consumer conduct. According to rational accounts of consumer behavior, firms and consumers have necessary information and rational beliefs, and maximize utility (or profit) accordingly. Under this view, if a social network adopts a feature that consumers consider privacy-invasive, consumers who continue to use the service do so because they value it in spite of this feature (Lenard and Rubin 2010; Westin 2000).[3] The GDPR's approach does not seem compatible with this account of consumer behavior. If consumers are able to make meaningful tradeoffs between privacy harm and the benefits of using a website, there is no need for regulation to limit the (mutually beneficial) agreements between users and websites. Consumers and firms should be able to determine the scope of data usage rights at the outset, as well as the modification rules for policy changes.

According to *behavioral* accounts of consumer conduct, consumers are unable to detect when a firm makes adopts harmful privacy practices. Consumers by and large do not read privacy policies (Obar and Oeldorf-Hirsch 2020),[4] and even if they did, it would be unlikely they would understand the policies and their implications (Strahilevitz and Kugler 2016). Therefore, they do not consider harmful privacy practices when deciding whether to use a service. If consumers do not read or understand privacy policies, and merely consent policies they are presented with, the GDPR's restrictions on broad allocations of rights to future data usage to service providers are also not likely to be meaningful. Requiring consumers to agree to future policy changes does not solve the problem that consumers are likely to agree to policies without reading or considering their content. Therefore, the approach of the GDPR, perhaps the most influential data governance regulation worldwide, is hard to justify also on behavioral grounds.

This paper makes two contributions, one **theoretical** and one **empirical**. On the **theory** front, we explain how privacy policies allocation residual data usage rights and why a regulatory model like the one adopted by the GDPR can protect consumers from harm stemming from the allocation of broad rights over future data usage to online services. For this, we develop a theoretical framework based on standard models of incomplete contracts in Hart, Shleifer, and Vishny (1997).[5] We note that the characterization of privacy policy as incomplete contracts is functional rather than legal. In particular, this modeling choice does not require that privacy policies are treated as contracts in the legal sense.

Our more nuanced account of negotiations over privacy practices assumes that there are too many future contingencies for it to be possible to contract for all future states of the world at the point of time in which parties enter the initial contract.This means that parties determine at the time of contracting whether the website has a broad right to use consumer data, even for unspecified uses ('residual data rights'). If consumers retain the residual rights to their data, yet future developments or investments require changes in consumer data usage, websites must obtain consumer consent, which we consider a 'take-it-or-leave-it'

3. This is sometimes referred to as the 'privacy calculus' (Acquisti, Taylor, and Wagman 2016). Early law and economics thinking on privacy supported the notion that privacy regulation can interfere with efficient markets, suggesting that consumers and firms benefit from minimal privacy restrictions (Posner 1981; Stigler 1980). A focus in the recent years of privacy economics is to consider the various frictions in data markets that suggest a role of regulatory interventions, relaying on the existence of transaction costs and externalities that complicate a simple story of trade-offs between consumers and firms (e.g., Romanosky and Acquisti 2009).

4. It is likely rational for consumers not to read privacy policies. A study from 2008 estimated that it would take 244 hours a year to read policies (McDonald and Cranor 2008). According to a recent estimate it would take an average person 76 working days reading privacy policies people agree to in a year. "Amazon's terms and conditions along out loud takes approximately nine hours," N.Y. Times (Feb. 2, 2019), https://www.nytimes.com/2019/02/02/opinion/internet-facebook-google-consent.html.

5. To our knowledge, the only other paper to apply the incomplete contracts theory to privacy is Dosis and Sand-Zantman (2019), who study a setting in which a monopolistic firm offers a service which generates valuable data. Similar to our setting, the use of data is profitable for the firm but generates privacy costs for consumers. However, their analysis centers on a setting of heterogeneous users with respect to their value of the service and the firm looks to extract data while potentially distorting the service usage of high types.

renegotiation. The assumption of contractual incompleteness is not hard to motivate once one recognizes that the type of data collected and how it will be used cannot be fully specified at the initial point in time that the consumer agrees to a privacy policy. When college students first signed up for Facebook in 2005, it was not obvious that the company would one day be able to track their precise geolocation at all times, nor compare their facial structures in an automated way to a vast database of biometric data. The ability to collect consumer data and the ways in which firms can profit from this information depend on technological and business innovations that might not be known at the point of initial contract. Firms are aware that consumer information is valuable in general, but they only discover over time which consumer data they need to use to generate profits and how this data needs to processed to achieve this end.

This framework also provides insights into why privacy policies use open-ended clauses that allow websites broad future discretion in how to use consumer data rather than specifying usages. In the original incomplete contracts framework, parties allocate property rights that determine who has the residual rights over an asset. Although privacy policies rarely explicitly allocate data property rights,[6] they often use open-ended language to establish whether consumer data can be used in a way not explicitly stipulated in the policy. This data right determines whether a website requires consumer consent for any changes in data usage not explicitly mentioned in the policy, which we refer to as a 'renegotiation' as it adjusts the terms of the agreement.[7] The framework also provides predictions as to when it might be beneficial for the consumer to retain residual rights to their data.

We further build on the incomplete contracts framework by developing a *behavioral incomplete contracts* framework. In this model, a behavioral agent discounts future privacy harms at the point in time in which they enter the contract. This assumption is appealing because of how it corresponds to consumer understanding in the real world. Although consumers have some, albeit limited, understanding of the possible privacy harms caused by the current use of their data, they are likely to have a particularly weak understanding of how future innovations may affect privacy. In the model, consumers in some cases give the website residual data rights at the time of contracting when it would have been beneficial for them to retain the right and renegotiate with the website when the new uses and their harms are better understood. We explicitly solve the model for the special case in which the firm has all market power and demonstrate that consumers may be indifferent, *ex ante*, between the firm or consumer holding residual data rights even when firm's holding residual data rights leads to a negative payoff. This is not the case when the consumer retains residual data rights and renegotiates only after they have rational privacy harm beliefs. Therefore, an important conclusion from the behavioral incomplete contracts model is that regulation has a role in intervening at the initial point of allocation of residual data rights. By limiting the ability for data rights to be assigned to websites at the stage of initial contracting, firms will be required to renegotiate with consumers when they wish to implement a new innovation and when the privacy harms are better known to the consumer.

The *behavioral incomplete contracts* model provides an explanation for regulatory models like the one adopted in the GDPR. One way to understand the GDPR's requirements regarding explicit descriptions of

6. There is an extensive literature discussing whether there is data property right and whether there should be a data property right (Bergelson 2003; Litman 2000; Victor 2013). We do not directly address the question of whether data can be conceived of as property and instead focus on the ability to use data in ways that are not explicitly specified in privacy policies, which we consider a residual data right.

7. We refer to the consent to a change in terms as a renegotiation, following the language used in many incomplete contract models, although in reality there is obviously no real negotiation between firms and consumers. In some loose way, however, firms may have to offer consumers a return to their consent for changing terms. For example, if Facebook wishes consumers to continue using their platform and consent to policy changes, it must continue to maintain and invest in their platform. In this way, the adjustments made by both the firm and consumer can be referred to as a renegotiation of the initial terms.

services' data practices is that they limit the ability to assign the website the residual rights to data usage in the case of an incomplete contract. As future uses of data that depend on technological or business innovation cannot to be explicitly described at the initial point in time a consumer agrees to a policy, the GDPR blocks the possibility that consumers agree to these future uses. In other words, the consumer must retain residual data rights, requiring a renegotiation in the form of active consent if websites wish to change data collection and usage provisions. In essence, the GDPR is creating a two-part mandate for privacy policies, requiring (1) that residual rights over data usage be assigned to consumers and (2) a heightened modification rule.

The *behavioral incomplete contracts* model also sheds light on what the unconstrained choice of websites, when not exposed to the GDPR, might look like. The GDPR's requirement that consumers consent to changes to the policy, while limiting the ability to use broad language in the original policy, creates frictions and could be a meaningful constraint on websites. Therefore, when websites are unconstrained in their U.S. policies, they are likely to choose to allocate residual data rights to themselves knowing that consumers discount the full implications of allocating residual rights to the firm.

The **empirical** section of the paper informs the theory section by shedding light on whether there are frictions created by consumers retaining residual data rights (what we refer to as 'renegotiation' in the model) and whether firms seek to obtain residual data rights when unconstrained by regulation. Our empirical analysis is not a direct test of the theory but rather a close consideration of one aspect of the theory. Namely, and examination of privacy policy provisions that affect how privacy policies can be modified, and specifically, whether privacy policies stipulate that consumers must consent to policy changes. While not a formal test of our theory, our empirical results are consistent with firm behavior in our model.

In our analysis, we consider only privacy policies used in interactions with U.S. consumers. In these interactions, U.S. online services do not have to comply with the GDPR. Hence, we consider them unconstrained by European regulation: They are able to use open-ended language with respect to data usage; and they are also not required to conform with the GDPR's rules for policy changes. Within our sample of U.S. privacy policies, we compare websites that (also) cater to E.U. users, meaning that they must comply with the GDPR in some of their interactions with consumers, with websites that do not cater to E.U. users and therefore do not comply with the GDPR in any of their interactions. For every website, we compare privacy policies before and after the GDPR came into effect.

We hypothesize that the GDPR, through creating strict consent requirements, can lead websites to learn about the implications of requiring active consent for policy changes. Websites' experiences with the GDPR might translate into what they choose to do in their interactions with U.S. consumers. If websites with E.U. exposure learn that the GDPR consent requirement is costly, they may chose to limit their requirement to obtain consumer consent for changes in data usage in their interactions with U.S. consumers. By contrast, we do not expect a similar effect for website providers that were unexposed to the GDPR. While this analysis is not a direct test of the paper's theory, it can provide circumstantial evidence for the proposition that residual data usage rights matter, and that the GDPR affected service providers' allocation these rights.

We show that U.S. websites with E.U. exposure are more likely to change their U.S. privacy policies to have *less* stringent and more lenient consent requirements for policy changes. In 2018, before the GDPR came into effect, most U.S. privacy policies in our sample do not mention any consent requirement with respect to changes to the policy and continue to not mention any consent requirement in their 2020 policies. However, of the policies that chose to change their consent requirements, the E.U. exposed websites are more likely to have their later policy contain *less* stringent consent requirements, either by no longer requiring

that consumers actively consent to changes or by dropping any mention of consent. Among websites that do not have E.U. exposure we see the opposite trend. Websites that do not have any E.U. exposure are more likely to add a consent requirement for changes or *increase* the stringency of the consent requirement.

Our results support the theory that the E.U. consent requirement, a central aspect of the restriction on allocation of residual data rights to websites, can be costly for websites, and that it is primarily E.U. exposed websites that are learning about the costly requirement. Websites with E.U. exposure then choose to adjust their U.S. policies, which do not need to comply with the GDPR, to avoid these costs by providing a weak requirement for consent, if at all.

These results contribute to our understanding of how regulation in one jurisdiction may affect firm behavior in another jurisdiction. The 'Brussels Effect,' a term coined by Anu Bradford (Bradford 2020) refers to a situation in which the E.U. unilaterally regulates through voluntary firm compliance with E.U. regulations. This paper seeks to build on this framework of voluntary compliance through analyzing the response of firms to the GDPR in a particular context. Namely, consent requirements for privacy policy changes. Our results suggest that in this domain there is need for more direct regulatory intervention to shape firm practice with respect to privacy policy provisions on data usage. Understanding the theoretical underpinnings of the GDPR and its impact on website behavior help shed light on broader policy debates around restricting the substance of privacy policies and creating mandatory rules for policy modification.

Our empirical analysis also contributes to the literature on the effect of the GDPR on consumers and firms in the E.U. Some studies consider the degree to which privacy policies actually comply with the GDPR, with mixed results (Kretschmer, Pennekamp, and Wehrle 2021; Lancieri 2022). Other studies have focused on consumer behavior in response to the GDPR (Santos et al. 2021). Schmitt, Miller, and Skiera (2021) demonstrate the impact of the GDPR on the number of website users, documenting a short term and long term decline in usage that varies across websites, with more popular websites suffering less.[8] Our results are consistent with these studies, as we suggest a new avenue in which compliance can be costly for websites, namely, the requirement to actively obtain consumer consent for changes in privacy policies. Our study differs from previous literature in that we are primarily focused on websites' unconstrained responses to the GDPR when they are not required to comply with the GDPR, providing a broader perspective on the effect of the GDPR outside the E.U. and firm preferences.

The paper has the following structure. In section 2, we discuss the incomplete contract model of privacy policies on how it helps explain regulatory interventions like the GDPR and the ways in which websites respond. We build on the incomplete contract model by considering a case in which a behavioral agent mispredicts the harm resulting from the future innovations. In section 3, we discuss our empirical study of U.S. privacy policies before and after the GDPR came into effect. Our main result is that websites that are also exposed to the GDPR are more likely to select a less stringent consent requirement after the GDPR by dropping any mention of consent to policy changes.

8. Recent studies suggest a nuanced effect to some of the GDPR requirements. Aridor, Che, and Salz (2020) show how the opt-in GDPR requirement leads to a drop in observable consumers but the remaining consumers are observable for longer. Mangini, Tal, and Moldovan (2020) attempt to more directly measure the costs of compliance with GDPR, primarily in the context of the right to be forgotten.

2 An incomplete contract theory of privacy policies

To study the implications of allocating residual data usage rights, we consider a bilateral relationship between a firm and a mass of consumers. Consumers decide whether to use the service provided by the firm, while the firm decides the level of investment in an innovation that is profitable but can cause privacy harm to consumers.[9]

We start by considering a firm's profit-maximizing level of investment in an innovation when the level of investment is fully contractible. In this case, the firm can commit to the level of innovation so that the level of investment does not depend on whether the firm or the consumer has the residual data usage rights. This provides a benchmark for the efficient level of investment. We then assume that the level of investment is not contractible and investigate the economic consequences of allocating the residual data usage rights to one of the sides of the market. Finally, we consider a case in which a behavioral consumer mispredicts the harm from future investments and only learns of the true harm after the firm selects the level of innovation. We show that in certain cases, this will lead to consumers giving up on their residual rights when it would have been socially optimal for them to retain these rights. The results of this model suggest that regulation has a role in intervening through limiting consumers' ability to transfer residual data rights to websites.

2.1 Model Assumptions

We consider a bilateral relationship between a web user or consumer, C, and a firm, F, who provides the consumer with a service. The consumer, C, receives some basic benefit B_0 in return for some basic price P_0.[10] The firm has some basic cost in providing the service to the user K_0. We also assume that the consumer has an outside option of zero. For example, consider a professional networking website, like LinkedIn,[11] that allows users to create pages with personal and professional information, which is then used to connect to other users and potential employers.

Innovation.–The firm, F, can make a revenue increasing investment e that increases revenue by $r(e)$. The investment e is meant to capture a broad range of innovations, such as new technologies that allow consumer data to be collected, business innovations that create new revenue opportunities from consumer data and legal innovations that reduce the exposure of firms to litigation. For example, a professional networking website, like LinkedIn, might innovate with respect to the information it collects on users and then shares with recruiters.[12] Other innovations might include the sale of data to third party advertisers, operating on

9. Prior work on the theory of privacy focused on the economic implications of more information sharing versus more privacy (Acquisti, Taylor, and Wagman 2016). The main conclusion from surveying the literature is that the welfare implications of more or less privacy depend on the particular context of information sharing and that "both individuals and organizations face complex, often ambiguous, and sometimes intangible trade-offs" (Acquisti, Taylor, and Wagman 2016, page 462). One approach to model internet privacy is where sellers search for buyers and send solicitations, while buyers (who have heterogeneous preferences for the seller's good) dislike being solicited and invest in hiding from the seller (Hann et al. 2008). The authors argue that putting a price on solicitations increases welfare. The model does implicate some of the concerns of how new technology harms privacy, but it focuses on one specific type of privacy invasion, and ignores the type of privacy harm that we focus on and which is a bigger problem: how websites use personal data. While early writing focused narrowly on the relationship between the consumer and firm,in what Acquisti, Taylor, and Wagman (2016) characterize as the second wave of economics of privacy research, economists considered concerns created by the secondary usage of personal data, which was not considered in the first wave (see e.g. Varian 2002). Another approach is a sequential contracting model in which an upstream seller gathers information about a buyer, which it can then sell to a downstream seller (Calzolari and Pavan 2006). This model is potentially useful, especially in a dynamic setting. In this paper, we focus on a simpler, bilateral relationship that still captures the heart of the privacy debate.

10. In many cases $P_0 = 0$. This is true for websites in which there is no upfront cost for using the websites services

11. https://www.linkedin.com

12. In fact, a significant source of LinkedIn's revenues comes from 'Talent Solutions' which allows recruiters access to infor-

and off the the networking platform.[13] Innovations may also include business innovations, such as investing in business relationships and developing business models that would allow the platform to profit from consumer information. We assume that the cost of each unit of innovation for the firm is constant and equal to 1.

Costs.–Another key element of out model is that while the innovation creates value for the firm, it can be costly for consumers and causes harm to the consumer's privacy $h(e)$.[14] There are at least two relevant types of costs. The first cost is the psychological harm caused by the preference for privacy. While these costs are intangible they have been documented as important for consumers (Feri, Giannetti, and Jentzsch 2016).[15] Another type of cost is related to the economic impact of data sharing, namely the possibility of price discrimination when more information is available about consumers (Acquisti, Taylor, and Wagman 2016; Richards 2021).[16]

The benefits to the consumer and profits for the firm can be summarized as:

$$B = B_0 - P_0 - h(e)$$
$$R = P_0 - K_0 + r(e) - e$$

2.2 Contractible Benchmark

As a benchmark, we consider the firm's profit maximizing level of investment by assuming that this level of investment is contractible. This assumes that all future types of investments are known at time $t = 0$ when the parties enter the contract. We also assume that consumers make rational choices that balance privacy harms and benefits.

When investments are predictable and contractible, the consumer C and firm F choose the cost saving investment e to maximize surplus, solving:

$$\max_e \{-h(e) + r(e) - e\}$$

The optimal level of cost investment, expressed as the first order condition is:

$$-h'(e^*) + r'(e^*) = 1$$

The optimal level of investment depends on the marginal benefit of spending extra effort to reduce costs, measured to take account of the marginal harm to privacy, which must equal the marginal cost of that extra effort, equalling 1.

The surplus that results from the innovation can then be split between the firm and consumer according to their relative bargaining position at time $t = 0$.

mation, ability to create targeted job posting and communicate directly with users.

13. There are other more speculative types of innovations that LinkedIn could engage in. For example, LinkedIn could develop the ability to track other online behaviors which it can then sell to potential employers. See example of discussion of LinkedIn's revenue- https://blog.waalaxy.com/en/how-linkedin-makes-money/

14. We assume that $r(e), h(e) > 0$. We also make the following assumptions in order to ensure a unique solution: $h(0) = 0; h' > 0; h'' > 0$. For the revenue function we assume that $r(0) = 0; r' > 0; r'(0) = \infty; r'(\infty) = 0; r'' < 0$. We also assume that $r' > h'$, which means that when a firm invests in an innovation, the marginal cost savings outweigh the marginal harm to the consumer.

15. Understanding and estimating people's preference has been addressed in many papers in recent years (Acquisti, Taylor, and Wagman 2016, section 3.8). One of the challenges for this literature is reconciling stated preferences for privacy and revealed choices, often known as the 'privacy paradox.'

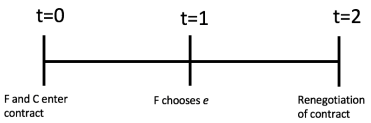
16. Another type of economic harm is the cost incurred when data is breached or misused, which may result in identity theft.

2.3 Incomplete Contract

In an incomplete contract framework, we assume that there are too many contingencies that it is impossible to anticipate and contract for all future possibilities in advance. The harms and benefits that result from the investments are all observable but not verifiable to third parties, meaning that they are not contractible. These assumptions are at the core of the incomplete contract framework as they explain why contracts are necessarily incomplete, by not being specific enough to create certainty as to the rights and obligations in all future states. These assumptions seem particularly appropriate for the privacy setting.[17] The monetizing of consumer data depends on many future innovations regarding the processing and sharing of the data, such as the ability to store and analyze large datasets. Moreover, monetizing of consumer data can depend on future business innovations and intermediaries, such as third party that aggregate consumer data for marketers. Many of these future possibilities may be unknown or accessible at date $t = 0$, when the parties enter the contract.

Given the inability to specify all future data usages, parties may instead decide at the time they enter the initial agreement whether or not the firm has the right to use the data for future usages not explicitly agreed upon by the parties. We therefore consider the firm to have the residual right to data usage when it can use consumer data for usages not specified at the time the parties initially enter the contract. Conversely, the consumer has the residual right to data usage if the firm must obtain consumer consent for data usage not specified in the initial contract. This means that the firm would need to 'renegotiate' with the consumer for further uses of the data.

As a shorthand, we sometimes refer to this residual right to data usage as a quasi-property right, or refer to situations in which the firm 'owns' the data or consumers 'own' the data. This language closely relates to the broader literature on incomplete contracts that often focuses on the key allocation question of which party has the residual control or decision right (Grossman and Hart 1986). However, it is important to note that whether data property rights can be assigned to firms or whether we should treat data as property all together, given its non-rivalrous nature, is a a topic that has been subject of much debate and is not fully addressed in this paper (Bergelson 2003; Litman 2000; Victor 2013). Instead we assume that at $t = 0$ the parties enter the contract in which they assign the residual rights to the use of consumer data, which will then determine whether the firm can decide without the consent of the consumer to use the data for purposes not specified in the contract.



Timeline.–At t=0 the website and user enter an agreement that assigns residual data right. At $t = 1$ the firm makes a decision with respect to their investment in a technology that will increase their profitability but also harm consumers. The implementation of the technology at $t = 2$ depends on who holds the residual rights to the data. If the consumer retains the residual right to the data, the firm will need consumer permission to implement the technology, requiring renegotiation with consumer. If the firm has the residual rights to the data, they are able to implement the technology without renegotiating with the consumer. We

17. We are aware of only one other paper that has used the incomplete contracts framework to study privacy. Dosis and Sand-Zantman (2019) consider the assigning of data property rights in a context of heterogeneous consumers, where the firm must incentivize consumer effort that generates monetizable data.

assume that all surplus gained from renegotiation is split in accordance to market power, where $\lambda \in [0, 1]$ notes the market power of the firm.

Renegotiation.–Referring to consumer consent to future data usages as 'renegotiation' is somewhat awkward. We use this terminology because it tracks how other incomplete contract applications refer to the stage at which the one party must obtain permission from the party with the residual datat rights as renegotiation. In reality, consumers and websites do not engage in a traditional negotiation in which consumers consent to data usage in return for a share of the surplus from the data usage. However, the take-it-or-leave offer presented by websites still requires that consumers perceive an ongoing benefit from using the website. For example, a website might know that in order to retain users in must continue to invest in the development and maintenance of it's interface and the service it provides to consumers. Absent this continuing investment, consumers may decide not to continue using a website and thereby refuse consent to data usage changes. Therefore, we can conceive of the interaction between users and websites as some loose form of renegotiation.

2.3.1 Firm has residual data rights

If firm F holds the residual rights to the data, they are able to implement the cost investment without the consumer's permission. When the firm 'owns' the data, the party payoffs are:

$$U_c = B_0 - P_0 - h(e)$$
$$U_f = P_0 - K_0 + r(e) - e$$

The firm will select the investment level that maximizes its utility, so that e_f expressed by the first order condition is:

$$r'(e_f) = 1$$

We can compare the level of innovation under a complete contract framework and the incomplete contract framework, when the firm 'owns' the data:

	cost innovation
Complete contract	$-h'(e^*) + r'(e^*) = 1$
Firm owns data	$r'(e_f) = 1$

Under the incomplete contract framework, when the firm 'owns' consumer data they over-invest in cost innovation relative to a complete contract (which we can consider as the first best), so that $e_f > e^*$. This is because the firm does not internalize the deterioration in consumer privacy that results from the innovation, and instead invests in cost innovation until the marginal cost saving equals the marginal cost (which is one).

2.3.2 Consumer has residual data rights

When the consumer 'owns' the data, the firm is unable to implement an innovation without the permission of the consumer. This means that, to implement the cost innovation, the parties must renegotiate. We

assume again that the surplus from the renegotiation is split according to the relative market power of the firm, expressed by λ.

When the consumer 'owns' the data, the parties' payoffs are:

$$U_c = B_0 - P_0 + (1-\lambda)\left[-h(e) + r(e)\right]$$
$$U_f = P_0 - K_0 + \lambda\left[-h(e) + r(e)\right] - e$$

The firm will select the investment level at $t = 1$ that maximizes its utility, so that e_c expressed as the first order condition is:

$$\lambda(-h'(e_c) + r'(e_c)) = 1$$

We can compare the cost innovation the under complete contract framework and the incomplete contract framework, when the firm or consumer own the data:

	cost innovation
Complete contract	$-h'(e^*) + r'(e^*) = 1$
Firm owns data	$r'(e_f) = 1$
Consumer owns data	$\lambda(-h'(e_c) + r'(e_c)) = 1$

We can summarize that the level of innovation is higher when the firm owns the data than when the consumer owns the data $e_f > e_c$. This is because the firm internalizes some of the harm from the cost innovation when the consumer owns the data. When the consumer owns the data, the innovation level is lower relative to its first best level in the contractible case, so that $e^* > e_c$. This is because the firm only internalizes part of the net benefit of the cost innovation, as opposed to the complete contract case in which the firm internalizes the full net benefit of the cost saving innovation. The conclusion is that allocating residual data usage rights to the firm and the consumer suffers from an inefficiency relative to the first best. When the firms owns the data they overinvest in innovation and when the consumers owns the data the firm underinvests in innovation.

2.3.3 Efficient allocation of residual data rights

The allocation of residual data rights to the firm is efficient if $S_f > S_c$.

$$S_f > S_c \iff -h(e_f) + r(e_f) - e_f > -h(e_c) + r(e_c) - e_c$$

Conversely, the allocation of residual data rights to the consumer is efficient if $S_f < S_c$.

In allocating residual data rights, the firm and consumer trade-off the inefficiency of over-investment (when the firm has residual rights) and under-investment (when the consumer has residual rights). Therefore, efficient allocation depends on the comparison of the harm function $h(e)$ and the revenue function $r(e)$. When the harm function is not significant allocation of residual data right to the firm is superior. This is because when the harm from the cost innovation becomes small, the cost investment under firm's holding residual data rights is closer to the first best. Conversely, if the privacy harm is not adequately offset by the benefit,

allocation of residual data rights to consumers is superior. In such a case it is more efficient to have lower incentives to invest in the cost innovation, like under consumer residual data rights.

A key prediction of the model, therefore, is that if firms and consumers allocate data rights efficiently, allocation of residual data rights to firms is more likely when the harm to privacy is less severe or when the benefits of the technology offset the harms. An example of when data innovation is very valuable to the firm but does not cause much harm to privacy, is when the firm makes use of anonymous data that can be used to train an algorithm. In that case it could be more efficient for the firm to have residual data rights.

Similarly, we would expect allocation of residual rights to consumers when the benefit to firm is more likely to be offset by the harm. For example, a dating app providing information on people's dating behavior to a third party. It is plausible that the harm is close to the benefit. More generally, the more sensitive the data that the firm is privy to, the more efficient it is for the consumer to have residual data rights.[18]

2.4 Behavioral Incomplete Contracts

In the incomplete contracts model both parties are rational but are unable to achieve the first best level of investment because they are unable to contract on the optimal level at the time of the initial contract. In this section we assume that the consumer is not fully rational. Specifically, we assume that the consumer, at the time they enter the contract, is naive about future privacy harm of innovation and so that they believe the harm function is $\delta h(e)$ rather than $h(e)$, where $0 < \delta < 1$. At $t = 0$ the consumer assumes that the firm will select the level of investment at $t = 1$ that depends on the cost function of $\delta h(e)$. Unlike the consumer, the firm assumes the correct cost function $h(e)$. At $t = 1$, after the firm has made an investment decision, the true hard function becomes known to the consumer. If the consumer 'owns' the data the parties then renegotiate when both parties accurately understand the privacy harm of innovation.

2.4.1 Firm has residual data rights

If the firm owns the data, there is no negotiation at $t = 2$. At $t = 1$ the firm selects its level of investment to maximize its utility $U_f = P_0 - K_0 + r(e) - e$.[19] Therefore its level of investment, expressed as a first order condition is:

$$r'(e_f) = 1$$

The consumer believes (incorrectly) at $t = 0$ that its utility under allocation of residual rights to the firm will be:

$$U_c^b = B_0 - P_0 - \delta h(e_f)$$

When in fact its utility at $t = 2$ will be:

$$U_c = B_0 - P_0 - h(e_f)$$

18. A high profile example of a data breach that caused a significant amount of harm due to the sensitivity of the data collected is the data breach of the website Ashley Madison. See https://en.wikipedia.org/wiki/Ashley_Madison_data_breach.

19. Although P_0 depends on the level of innovation e, as will be explained, at the time the firm makes the decision of the optimal level of innovation this has no effect on the price paid by the consumer in at $t = 1$.

$U_c^b > U_c$ and so the consumer's utility is lower than what she predicted it would be. This is because the consumer discounts the harm that will result from the firm holding residual data rights. If the basic price to access the website, P_0 is set endogenously[20] the behavioral consumer may be willing to pay a higher basic price so that $P_0^b > P_0$. This would provide a clear benefit for the website, as they are able to charge a higher price as a result of the consumer's misprediction as to future innovation levels or use the lower price to induce the consumer to agree to firm ownership when the firm prefers firm ownership.

2.4.2 Consumer has residual data rights

To understand how the behavioral consumer's decision and utility differ from the rational consumer's decision and utility, we must first understand the renegotiation at $t = 2$ and firm decision at $t = 1$.

t=2.– At $t = 2$ the parties renegotiate if the consumer retains the residual rights to the data. We assume that at $t = 2$ the consumer is aware of the **true** harm $h(e)$, and that the surplus from the negotiation is split according to the relative market power λ. This means that the party payoffs are:

$$U_c = B_0 - P_0 + (1 - \lambda)\left[-h(e) + r(e)\right]$$
$$U_f = P_0 - K_0 + \lambda\left[-h(e) + r(e)\right] - e$$

The same as in the non-behavioral case.

t=1.– Realizing that the payoffs at $t = 2$ will be determined by the true harm function, the firm will select the investment level that maximizes its utility, so that e_c expressed as the first order condition is:

$$\lambda(-h'(e_c) + r'(e_c)) = 1$$

Again, the level of investment chosen by the firm is the same as in the non-behavioral case.

t=0.– From the consumer's perspective at $t = 1$ they have incorrect beliefs as to the level of investment that the firm will select at $t = 1$. The consumer believes that the payoffs at $t = 2$ will be:

$$U_c^b = B_0 - P_0 + (1 - \lambda)\left[-\delta h(e) + r(e)\right]$$
$$U_f^b = P_0 - K_0 + \lambda\left[-\delta h(e) + r(e)\right] - e$$

So that the consumer believes that the firm at $t = 1$ will select an investment level that maximizes what the consumer expects to be the firms utility:

$$\lambda(-\delta h'(e_c^b) + r'(e_c^b)) = 1$$

Note that $r(e_c^b) > r(e_c)$, meaning that the consumer believes that the firm will chose a higher level of investment than what the firm actually ends up selecting in $t = 1$. If we compare these levels of investment to the level of investment under firm residual rights we get that $r(e_f) > r(e_c^b) > r(e_c)$, meaning that the level of investment is higher when the firm holds residual data rights than under the consumer's incorrect beliefs at $t = 0$ when consumer retains residual data rights, which is higher than the true level of investment under consumer residual rights.

20. In reality, the basic price of $P_0 = 0$ is quite sticky for many types of websites, particularly for basic access to the website.

Recall that in subsubsection 2.3.3 we discussed that whether it was optimal for the consumer or the firm to have residual rights to data depended on the relative benefit and harm from the innovation. Firm's having residual data rights is more likely when the benefit of the innovation offsets the harm from the innovation. Therefore, in the behavioral case:

1. When the firm holds residual rights, the consumer perceives the harm to be lower than the true harm of the innovation, meaning that the behavioral consumer is more likely to believe the benefit of the innovation offsets the harm of the innovation.
2. However, when consumer's retain residual rights, the consumer believes the level of investment will be higher than it really is. Therefore the consumer believes that the difference between the level of investment with firm residual rights and consumer residual rights is smaller than it truly is.

The first factor leads to the consumer discounting the harm of firm ownership while the second factors mitigates concerns over consumer ownership. From the firm's perspective, the discounted harm at $t = 0$ leads to the a higher to the consumer paying a higher price for the initial access to the service. The next section considers these factors under the assumption that the firm has all market power.

2.4.3 Special case: firm has all market power

To understand the implications of the behavioral incomplete contracts model, I consider a special case in which the firm has all market power both at the time the parties enter the initial contract, $t = 0$, and when the parties renegotiate at $t = 2$. If the firm has all market power at $t = 0$, and we assume that the consumer has a reservation price of 0, the consumer will only enter the contract if the expected payoff is non-negative. If the firm has all market power at $t = 2$ we set $\lambda = 1$.

Firm has residual data rights.–Although the consumer believes the future harm function to be $\delta h(e)$, the consumer has rational beliefs about the level of investment selected by the firm because the firm does not internalize the privacy harm under firm ownership. The firm selects e_f, the same as the non-behavioral case.

This means that the price P_0 is determined by setting the consumer payoff to 0

$$U_c = B_0 - P_0^{bf} - \delta h(e_f) = 0$$

$$P_0^{bf} = B_0 - \delta h(e_f)$$

Relative to the maximum price under rational beliefs, the consumer is willing to pay a higher price because they discount future harm- $P_0^f < P_0^{bf}$.

The firm payoff is:

$$U_f = P_0^{bf} - K_0 + r(e_f) - e_f$$

$$U_f = B_0 - K_0 + r(e_f) - e_f - \delta h(e_f)$$

Consumer has residual data rights.–When the firm has all market power at $t = 2$ it is able to fully capture the surplus from renegotiation (but still needs the consumer's consent for implementing the innovation and therefore must compensate the consumer for the privacy harm for the consumer to agree). This means that the firm maximizes:

$$\max_e\{-h(e) + r(e) - e\}$$

The optimal level of cost investment, expressed as the first order condition is:

$$-h'(e^*) + r'(e^*) = 1$$

This is the same level of investment as under the contractible benchmark.

However, because the consumer believes that the harm is onlyδ the consumer believes that the firm is maximizing:

$$\max_e\{-\delta h(e) + r(e) - e\}$$

The optimal level of cost investment, expressed as the first order condition is:

$$-\delta h'(e^{bc}) + r'(e^{bc}) = 1$$

The consumer believes that the level of investment selected by the firm is higher than e^*. This is because the harm from the innovation is lower. The consumer expects to withhold their agreement to the implementation of the innovation unless they are compensated for the privacy harm $h(e^{bc})$.

The initial price the consumer pays for the service (P_0^{bc}) is determined by setting the consumer's believed payoff to 0:

$$U_c = B_0 - P_0^{bc} - \delta h(e^{bc}) + \delta h(e^{bc}) = 0$$

$$P_0^{bc} = B_0$$

The maximum price the consumer is willing to pay for access to the service is higher than under firm ownership. The consumer beliefs about the level of investment is lower under consumer ownership ($e_f > e^{bc}$) because under firm ownership the firm does not internalize the privacy harm at all.

What actually happens is different to the consumer's beliefs. When the parties renegotiate at $t = 2$ the firm must adequately compensate the consumer for the harm of the innovation, because otherwise the consumer would not agree to the implementation of the innovation. This means that the firm actually compensates the consumer by $h(e^*)$ and not $\delta h(e^{bc})$

The firm payoff is:

$$U_f = P_0^{bc} - K_0 + r(e^*) - e^* - h(e*)$$

$$U_f = B_0 - K_0 - h(e^*) + r(e^*) - e^*$$

When the firm prefers to have the residual data rights.–The firm prefers firm ownership when its payoff under firm ownership is higher than its payoff under consumer ownership.

Firm payoff under firm ownership is:

$$U_f = B_0 - K_0 + r(e_f) - e_f - \delta h(e_f)$$

Firm payoff under consumer ownership is:

$$U_f = B_0 - K_0 - h(e^*) + r(e^*) - e^*$$

Therefore a firm prefers firm ownership when:

$$r(e_f) - e_f - \delta h(e_f) > r(e^*) - e^* - h(e^*)$$

Whether the firm prefers firm ownership depends on the consumer's beliefs δ and the harm and benefit functions.

Conclusion.–

1. For certain values of δ, and harm function $h(e)$ and benefit function $r(e)$ the firm may prefer firm ownership because under firm ownership the consumer pays a higher price for access to the service than they would with rational beliefs.

2. When the firm has all the market power, consumers with wrong beliefs about the harm function should not have entered an agreement for firm ownership altogether, as this results in a negative payoff, $U_c = B_0 - (B_0 - \delta h(e_f)) - h(e_f) < 0$

2.5 Model implications

The goal of our model is to contribute to the conversation about privacy policies among scholars and policy makers. Specifically, our adaptation of Hart, Shleifer, and Vishny (1997) provides insight into why privacy policies use open-ended clauses that allow websites broad future discretion in how to use consumer data rather than specifying usages. The model also demonstrates how allocation of residual data rights can affect the level of innovation and privacy protections. Our behavioral extension shows that when firms are better-placed than consumers to predict future privacy harms they may seek to obtain residual data rights or make future negotiations easier at the initial time of contracting. In our special case of when the firm has all market power, this could lead to firms retaining residual data rights even when this results in a transaction that the consumer would not have entered at all with rational beliefs. This conclusion implies that regulation may have a role in intervening at the initial point of allocation of residual data rights. By limiting the ability for data rights to be assigned to websites at the stage of initial contracting, firms will be required to renegotiate with consumers when they wish to implement a new innovation and when the privacy harms are better known to the consumer.

The regulatory implications of the behavioral incomplete contracts framework may capture the essence of the limitations set by the GDPR with respect to policy changes. The GDPR's requirement that websites explicitly outline data usage in the privacy policy and the restrictions on policy modification can be seen as a method of forcing consumers to retain residual rights to their data. Under these types of regulatory interventions, firms are unable to hold residual rights to data usage and are therefore forced to renegotiate with consumers for any change in data usage. This renegotiation is typically in the form of a take-it-or-leave-it offer in which the consumer consents to the new data usage in return for the continued use of the website.

Our theory provides a framework to explain certain features of the GDPR, but it is unclear whether the assumptions of the model hold true in reality. We are therefore interested in understanding how websites respond to the GDPR and whether their response is consistent with the model. Specifically, we are interested in the unconstrained choice of websites that are exposed to the GDPR, meaning what websites choose to stipulate in the privacy policies they use for their U.S. consumers, where they are not subject to the GDPR. If the GDPR's limits on the use of broad language in privacy policies and its unalterable requirements for adopting changes to these policies create frictions and impose meaningful constraints on websites, websites might learn about the implications of consumers retaining residual data rights. Thus, in situations in which they are unconstrained by the GDPR (in interactions with U.S. consumers), they might be more likely to choose to allocate residual data rights to themselves. While this analysis is not a direct test of the paper's theory, it can provide circumstantial evidence for the proposition that residual data usage rights matter, and that the GDPR affected service providers' allocation these rights.

3 How do websites respond to the GDPR?

In this section, we study changes in privacy policies that affect how these policies can be changed. In particular, we are interested whether privacy policies stipulate that consumers must consent to future changes of the privacy policy. As we describe below, under the GDPR, websites are constrained with respect to the rules for changing privacy policies in two ways. Firstly, websites are required to explicitly lay down the types of consumer data that is collected and how this data is used, restricting website's ability to capture future changes in data collection and usage in their initial policy. Second, websites are required to justify any changes to their privacy policy that leads to more invasive privacy practices, which often requires obtaining consumer consent. The bottom line is that the GDPR lays down strict rules on consumer consent for future data rights.

In the U.S., websites do not face similar legal constraints. Therefore, we consider the unconstrained choices websites make in privacy policies used in interactions with U.S. consumers with respect to the provisions that govern changes in these policies. We hypothesize that the GDPR, through creating consent requirements, allows websites to learn about the implications of requiring active consent to policy changes. What websites learn will potentially translate into what they choose to do with their U.S. policies.

Not all U.S. websites may learn equally about the cost of consent for policies changes. If websites that are exposed to the GDPR, i.e. they must comply with the GDPR in their interactions with consumers in the E.U., they may be directly and concretely learning about these costs. On the other hand, for websites that do not cater to E.U. subjects, the costs of the consent requirement may be less salient and they might not be learning about the implications of the requirement.[21]

We consider whether websites learn about the costs of active consent to policy changes by comparing the privacy policies of GDPR-exposed websites and websites without GDPR exposure, before and after the GDPR went into effect. We show that GDPR-exposed websites are more likely to adopt less stringent rules with respect to changes in privacy policies.

21. Note that our sample of websites does not include websites run by online service providers headquartered in the E.U., which are under a legal obligation to comply with the GDPR in their interactions with consumers worldwide (GDPR Art 3(1)).

3.1 Legal Landscape

3.1.1 General Observations

In both the U.S. and Europe, regulatory efforts have focused on a 'notice and consent' approach. Under this approach, online services are required to notify consumers about their privacy practices. Insofar as they follow the requirements for notice and consent in the respective jurisdiction, there are few—if any—substantial restrictions on what online services can do with consumer data.[22] Despite their common focus on notice and consent, however, the United States and the European Union take fundamentally different approaches to regulating data privacy. In the United States, data privacy regulation is mostly limited to specific fields such as education and credit reporting. Only in recent years have several states started to enact more comprehensive rules.[23] As a result of the restraint on the part of legislators and regulators, it is largely left to businesses and consumers to specify the rules governing the use of the consumer's data by way of contracting (Davis and Marotta-Wurgler 2019).

Under E.U. law, any business that intends to collect or process personal data requires a legal justification to do so. One important way to secure this justification is by obtaining a consumer's consent (GDPR Art. 6(1)a.). Other avenues include showing that the "processing is necessary for the purposes of the [business's] legitimate interests" (GDPR Art 6(1)f.) Because of the relatively strict rules governing consumer consent, regulatory agencies and courts play a larger role in determining the scope of permissible data practices.

The E.U.'s approach to data privacy has been relatively consistent since at least the 1990s, when the first E.U.-wide data privacy law, the Data Protection Directive, entered into force. Since then, the most important milestone in the development of E.U. data privacy law was the entry into force of the GDPR in May 2018. Compared to the Data Protection Directive, the GDPR imposes substantially higher hurdles on businesses intent on collecting and processing consumer data. For example, under the GDPR, a request for consent must "be presented in a manner which is clearly distinguishable from the other matters, in an intelligible and easily accessible form, using clear and plain language" (GDPR Art. 7(2)).

A second important innovation of the GDPR is that it significantly expanded the geographic reach of E.U. data privacy law. While the GDPR's predecessor, the Data Protection Directive, applied only to businesses that were either operating out of the European Union or at least used equipment situated within the E.U. in their data processing (Data Protection Directive, Art. 4), the GDPR also applies to businesses located elsewhere, if and insofar these businesses 'target' consumers in the European Union (GDPR, Art. 3(2)).

3.1.2 Restrictions on Future Data Use

The different regulatory approaches in the United States and the E.U. also translate into different restrictions on future data usage.

In the E.U., businesses need to state ex ante in their privacy policies what type of data they collect and how they intend to use it. This was the case even before the GDPR entered into force. However, the GDPR significantly increased the level of detail with which businesses have to describe the use of consumer data. As a result, under the GDPR, it is impermissible for businesses to reserve the right to use data for purposes

22. Sloan and Warner (2014) provides a summary of the notice and consent approach and its main critiques.

23. The first of this new generation of privacy law was the California Consumer Privacy Act of 2018 (CCPA), which went into effect in January 2020.

that were not explicitly mentioned in their privacy policies at the time of data collection. Instead, in many cases, the only way to use data gathered in the past in novel ways will to be to obtain a consumer's explicit consent for the new processing techniques.

In the United States, as described above, websites are generally free to use consumer data in any way that is not in conflict with their own privacy policy. At the same time, the law imposes few constraints on the contents of these privacy policies. Against this background, websites can *de facto* reserve the right to change their processing of consumer data after its collection by using broad terms that can accommodate a range of different data practices.

The extent to which services in the United States face legal constraints when changing their privacy policies is subject to debate. Given that there exists no comprehensive data privacy law in the United States, there are a number of legal grounds that could create constraints on changing policies.

One important set of constraints on changing policies follows from the threat of enforcement actions under Section 5 of the FTC Act. According to the FTC's enforcement practice, services are generally free to change their privacy policies at any time. One exception to this rule concerns retroactive changes. Retroactive changes are instances in which a service provider changes its privacy policy to treat data that has already been collected in a way that would have been in conflict with the policy that was in place at the time. Such changes are considered a violation of Section 5 of the FTC Act, unless the service provider obtains the "informed consent of affected consumers" (FTC in re: Gateway Learning Corp. 2004). In addition, to the extent that services lay down other constraints to changing their privacy policy, such as requiring consumer consent, they are not free to act in a manner inconsistent with the provisions in the privacy policy.

The ability of services to change their privacy policies might also be constrained if courts considered privacy policies to be contracts in the legal sense. Some scholars, including the authors of the Draft Restatement of the Law, Consumer Contracts, have argued that privacy policies usually constitute 'consumer contracts' (Draft Restatement, §1, Comment 9, Bar-Gill, Ben-Shahar, and Marotta-Wurgler 2017). This classification as consumer contracts can impose restrictions on the ease with which services can change their privacy policies in a legally valid way. In particular, modifications of privacy policies may require 'assent'[24] According to the Draft Restatement, absent a contractually stipulated procedure governing modifications to a consumer contract, businesses are required to provide consumers with notice about a change as well as "a reasonable opportunity to reject the proposed modified term and continue the contractual relationship under the existing term" (Draft Restatement, §3). This requirement *de facto* implies that privacy policies that lack a clause allowing for unilateral modifications by a business cannot be changed without the consumer's consent. The Draft Restatement has yet to be adopted. At the same time, some scholars dispute the Draft Restatement's assertion that existing case law supports the notion that privacy policies are contracts (Klass 2019; Levitin et al. 2019).

24. The Draft Restatement on Consumer Contracts and Bar-Gill, Ben-Shahar, and Marotta-Wurgler (2017) provide a further discussion of the differences between consent and assent. In this paper we refer loosely to both standards as 'consent' primarily because all notions of consent in this paper refer to the procedure of agreeing to the terms of a consumer contract.

3.2 Data

In the analysis, we use a dataset consisting of the texts of the privacy policies of 238 websites. For each website, we obtain one observation from April 2018 and one observation from April 2020.[25] To create this dataset, we begin with all websites that appeared in the Top 500 ranking in Alexa's Top Site service in late 2017.[26] We exclude websites operated by online services that are not headquartered in the United States.[27] The resulting dataset consists of 325 websites and their privacy policies. Notably, as these websites are operated by services located in the United States, they were not legally obliged to treat U.S. consumers in line with the requirements stipulated in the GDPR (GDPR, Art. 3(1), Frankenreiter 2021).

We construct our treatment and control groups from the remaining 325 websites. Our treatment group consists of the U.S. versions of privacy policies of websites that either cater to E.U. consumers alongside consumers in other jurisdictions or offer one or more separate 'E.U. versions' of their websites.[28] Importantly, because of their substantial exposure to E.U. consumers, these services have to comply with the GDPR insofar as they serve their customers in the European Union (GDPR, Art. 3(2)). Yet we focus on privacy policies used in interactions between these services and consumers in the United States, where the GDPR does not apply. The treatment sample consists of 130 websites, corresponding to 260 privacy policies.

Our control group contains of the privacy policies of websites that we consider not to have been 'exposed' to the GDPR, meaning that they do not have to comply with the GDPR in any of their interactions with consumers. The reason for the lack of exposure to the GDPR is that these websites are not 'targeting' consumers in the E.U. in the way described above.[29] We exclude from our control sample all websites of services in industries that are covered by sector-specific privacy regulations in the U.S.[30] We furthermore exclude all websites that, although they do not explicitly target E.U. consumers, receive more than 10% of their traffic from the E.U., or for which the ratio of E.U. versus U.S. traffic is above 0.11.[31] We are left with a sample of 108 'control' websites and 216 privacy policies. The final dataset therefore consists of 476 policies from 238 websites. In addition to the texts of the privacy policies, we further obtain a range of background variables for all websites.[32]

Our main analysis focuses on a binary classification of E.U. exposure. In practice, however, exposure to E.U. regulation could be a matter of degree. A firm that caters primarily to E.U. subjects may have greater exposure to the GDPR than firms that caters primarily to U.S. consumers. In our Online Appendix, we therefore replicate the analysis presented here using a measure of E.U. exposure that is not binary, but measures the degree of E.U. exposure. For this, we obtain information on the overall percentage of traffic to a website comes from E.U. countries.[33]

25. Frankenreiter (2021) and Frankenreiter and Hermstruewer (2021) provide more details on how the privacy policies were obtained.

26. https://www.alexa.com/topsites.

27. We further restrict the sample to websites that have policies in English and do not use privacy policy as another website in the sample.

28. One example in the former category is facebook.com, which is available in many languages primarily spoken in the European Union. The latter category includes Google, which makes its services available to E.U. residents mainly through country-specific versions of its website such as google.de.

29. Examples of websites in this category include large retailers without operations in the E.U. such as Lowes and Bed, Bath & Beyond.

30. The industries we exclude are financial services, academia, government and medical.

31. To make this determination, we use traffic data obtained from Alexa Web services in the summer of 2020. Frankenreiter (2021) describes this data in more detail. We exclude websites from our sample for which such traffic data is not available.

32. This includes information such as the website category and website entry data.

33. The estimates of the percentage of website traffic from E.U. countries are from October 2020.

We coded the policies in this dataset according to a coding scheme developed specifically for this project. The coding scheme captures various features of a website's privacy policy, including (1) the types of data that a service is allowed to gather, (2) the different ways in which a service is allowed to use data under the policy, and (3) rules related to policy changes. With regard to (1) and (2), the coding scheme contains lists of different types of information, and different information uses whose legal status under the policy we ascertain through our coding. Examples of items in the list for (1) are personally identifiable information and geolocation data. The list for (2) included data sharing with third parties for advertisement and marketing and data sharing for other activities external to the website. For each item, we asked our coder to ascertain whether the collection of this type of data or the information use was (i) explicitly allowed under the policy, (ii) implicitly permitted under the policy (because, for example, the policy contained open-ended language that did not restrict the practice), or (iii) not allowed under the policy. With regard to (3), the coding scheme requires coders to identify whether the rules fell into one of the following categories: (i) No mention of consent, (ii) policy stipulates that continued use of the website is treated as consent, (iii) users need to actively consent to any changes. The coding was done in the summer of 2021 by a research assistant from Columbia Law School.

3.3 Results

We begin by reporting how U.S. privacy policies change after the GDPR came into effect, comparing GDPR exposed websites to websites not exposed to the GDPR. We then consider whether policies of GDPR exposed websites differ from non-exposed websites with respect to the type of information they collect and the ways they use this data. It is important to consider this second dimension of data collection and usage because to some extent broad data rights and consent requirements can be viewed as substitutes. A policy that sets down very broad data rights may not require material change in the future making strict consent requirements less consequential. We therefore consider differences in data rights articulated in the privacy policies before and after the GDPR comes into effect.

3.3.1 Consent

As described above, we distinguish between three types of consent provisions in policies. Most policies in our sample contain no mention of whether consumers must consent to changes to the privacy policy ('no mention').[34] For the policies that do require consent, they either stipulate that consumers consent to the updated privacy policy when they continue to use the website ('continued use as consent')[35] or they require consumers to actively consent to the new policy ('active consent')[36].

34. For example, in the 2018 zappos.com privacy policy there is mention of the possibility that the privacy policy can change but no provisions related to consumer consent: 'If we change or update this Privacy Policy, we will post changes and updates on the Site so that you will always be aware of what information we collect, use and disclose. We encourage you to review this Privacy Policy from time to time so you will know if the Privacy Policy has been changed or updated. If you have any questions about the Privacy Policy, please contact us at 1-800-943-1633 or cs@zappos.com.'

35. For example, in Twitter's 2018 it says with respect to changes in the policy: 'We may revise this Privacy Policy from time to time. The most current version of the policy will govern our use of your information and will always be at https://twitter.com/privacy. If we make a change to this policy that, in our sole discretion, is material, we will notify you via an @Twitter update or email to the email address associated with your account. By continuing to access or use the Services after those changes become effective, you agree to be bound by the revised Privacy Policy.'

36. For example, in Verizon's 2018 policy it says with respect to changes in the policy: 'We reserve the right to make changes to this privacy policy, so please check back periodically for changes. You will be able to see that changes have been made by checking to see the effective date posted at the end of the policy. If Verizon elects to use or disclose information that identifies you as an individual in a manner that is materially different from that stated in our policy at the time we collected that

Table 1 shows the consent provisions in our sample, separately for 2018 and for 2020. Over half of all policies in the sample do not mention any consent requirement for changes to the policy policy (56.4% in 2018 and 59.4% in 2020). The vast majority of policies that do set down rules for consenting to changes in policies state that continued use of the websites constitutes consent to the policy changes ('continued use as consent'), rather than requiring that consumers actively consent to changes.

Table 1: Consent provisions

	2018 consent			
	Active consent	Continued use as consent	No mention of consent	Total
	(1)	(2)	(3)	
No GDPR Exposure	1 (1%)	44 (41%)	63 (58%)	108
GDPR Exposure	1 (1%)	57 (44%)	72 (55%)	130
Total	2 (1%)	101 (42%)	135 (57%)	238
	2020 consent			
	Active consent	Continued use as consent	No mention of consent	Total
	(1)	(2)	(3)	
No GDPR Exposure	2 (2%)	50 (46%)	56 (52%)	108
GDPR Exposure	0	44 (34%)	86 (66%)	130
Total	2 (1%)	94 (39%)	142 (60%)	238

3.3.2 Changes in consent stringency

We are primarily interested in whether privacy policies are more or less stringent in their 2020 policy, following the GDPR coming into effect, relative to their 2018 policy. Most policies are unchanged with respect to their consent provision,[37] with 83.7% of websites with no GDPR exposure remaining with the same consent provision type in their 2018 and 2020 policies and 78.5% of websites with GDPR exposure remaining with the same consent provision type in their 2018 and 2020 policies.

Defining stringency.– We define policies as **less stringent** if in 2018 the policy required 'active consent' and in 2020 the policy either required 'continued use as consent' or there is 'no mention' of consent in the policy. A policy also becomes less stringent with respect to its consent requirement if in 2018 it required 'continued use as consent' and in 2020 it no longer mentions consent.[38] Conversely, a policy is **more stringent** with respect to its consent provision if the 2018 policy did not mention consent and the 2020 requires 'continued use as consent' or 'active consent'. The policy is also more stringent if in 2018 it required 'continued use as consent' and in 2020 required 'active consent.' It is important to note that although at first 'continued use as consent' could be viewed as close to having not consent requirement at all, in practice 'continued use as consent' could create significant barriers for websites in obtaining containing. For example,

information from you, we will give you a choice regarding such use or disclosure by appropriate means, which may include use of an opt out mechanism.' Note that this provision only requires active consent in a narrow set of circumstances. We nonetheless code this as 'active consent' because it does provide for circumstances under which active consent it required.

37. This does not necessarily mean that the exact language in the policy remained the same from 2018 to 2020 but that in essence the type of consent required has not changed even if the language has changed.

38. As noted above, FTC jurisprudence suggests that that retroactive changes to policies as requiring informed consent it remains unclear what change would be considered retroactive. Moreover, websites themselves do not distinguish between prospective and retroactive changes to policies.

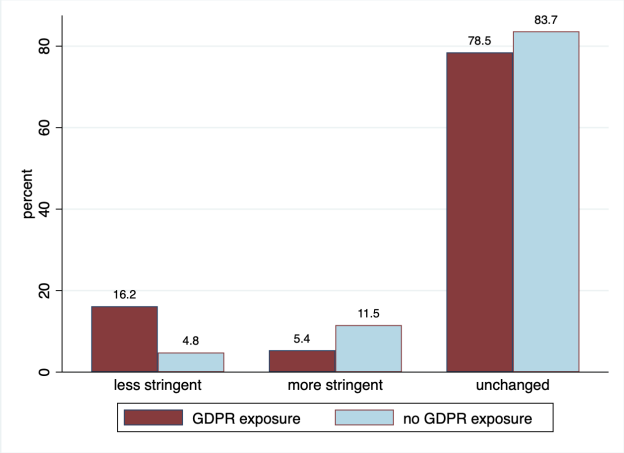
if a website uses data of a user who is no longer active on the website.

Our definitions of consent stringency rely on our understanding that firms that drop any mention of consent seek to lower their consent requirement, for example, by allowing the mere posting of a new policy to be sufficient for the new policy to apply. If privacy policies are considered contracts,[39] it is possible that a privacy policy that does not lay down a mechanism for policy changes would actually require active consent for policy changes, the default for consent for privacy policies. This could mean that from a legal perspective, dropping any mention of consent would create a **more stringent** consent requirement.

Despite ongoing debates on whether privacy policies are considered contracts, we adopt the interpretation that websites that drop any mention of consent intend to adopt a **less stringent** consent requirement. First, it continues to be disputed whether privacy policies are indeed contracts and in any event there is little case law on the exact implications of policy silence on required consent for policy changes. Second, we are primarily interested in website's subjective beliefs. Based on conversations with industry practitioners, we consider it unlikely that websites believe that by omitting mention of consent they are opting in to the most stringent consent requirement.

Consent stringency results.– Figure 1 shows the percentage of privacy policies that became **less stringent**, **more stringent**, or **unchanged** in their 2020 policy, reporting percentages separately for websites that are exposed to the GDPR and websites that are not exposed to the GDPR. Figure 1 shows that websites that are exposed to the GDPR are more likely to make their policies **less stringent** with respect to consent requirements while websites that are not exposed to the GDPR are more likely to make their consent requirements **more stringent**. While most policies do not make an essential change to their consent provisions, of the websites that do change their policy's consent provision there is a clear difference between GDPR exposed websites, who are more likely to choose to have less stringent policies in 2020 and websites with no GDPR exposure who are more likely to make their policies more stringent. Table 2 reports these results along with their statistical significance. In out Online Appendix we report the results using a logistical regression with the percentage of E.U. users' as the independent variable.

Figure 1: Provision changes



For GDPR-exposed websites, most policies that do not change their consent provisions offer the lowest level of consent stringency. Figure 2 shows the consent provisions for policies that did not change from 2018

39. See discussion in subsection 3.1. Classifying privacy policies as contracts may also have implications for whether explicit consent provisions are enforceable (Bar-Gill and Davis 2010).

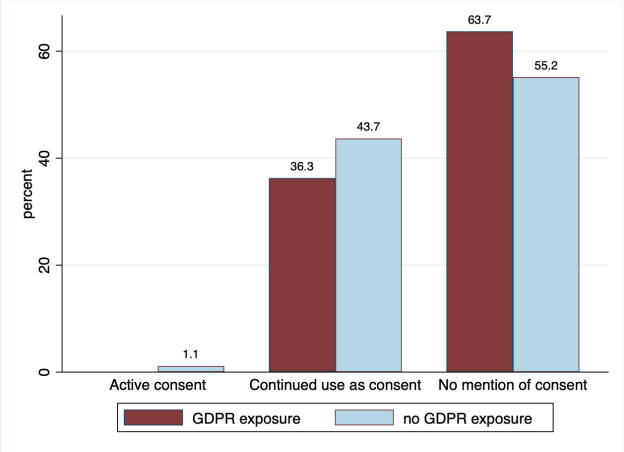
Table 2: Consent provision changes

	Consent Change			
	Less Stringent	More Stringent	Unchanged	Total
	(1)	(2)	(3)	
GDPR Exposure	21 (16.15%)	7 (5.38%)	102 (78.46%)	130
	14.4	10.6	105.0	
No GDPR Exposure	5 (4.81%)	12 (11.54%)	87 (83.65%)	104
	11.6	8.4	84.0	
Total	26 (11.11%)	19 (8.12%)	189 (80.77%)	234

Pearson Chi-square = 9.5818, P-value = 0.008, Fisher's exact = 0.007

to 2020. The graph shows that 63.7% of unchanged GDPR exposed policies do no mention consent at all. This is important because 'no mention' is a lower bound for consent provisions, and therefore the results could understate a website's preference for less stringent consent provisions post GDPR, as many policies already had the lowest consent stringency possible.

Figure 2: Provisions for policies that are unchanged



A key challenge to any comparison between GDPR exposed websites and non-exposed websites is that whether a website is exposed to the GDPR is not random. As we document in our Online Appendix, the treatment and control groups vary with respect to their industry or category type. Moreover, the treatment and control group are not identical with respect to their user numbers even though both control and treatment groups contain high traffic websites.

We build on our analysis of changes in consent provisions of the full sample by considering the differences in a matched sample. To create this matched sample we identify all treatment websites that have an Alexa category of which there is at least one control website with the same category. We end up with a sample of 96 treatment firms. When there is more than one website in the control sample with the same category as a treatment website we select a match, with replacement, based on similarity in user numbers. Table 3 shows that also within this matched sample GDPR exposed websites are more likely to adopt a less stringent consent requirement for policy changes.

Table 3: Consent provision changes: Matched sample

	Consent Change			
	Less Stringent	More Stringent	Unchanged	Total
	(1)	(2)	(3)	
GDPR Exposure	16 (16.7%)	5 (5.2%)	75 (78.1%)	96
	11.0	4.0	81.0	
No GDPR Exposure	6 (6.3%)	3 (3.1%)	87 (90.6%)	96
	11.0	4.0	81.0	
Total	22 (11.5%)	8 (4.2%)	162 (84.4%)	192

Pearson Chi-square = 5.9343, P-value = 0.051

3.3.3 Data collection and usage

On the dimension of the scope of permissible data gathering, we see a move towards policies that explicitly allow for the gathering of certain sensitive types of data. This is true for both websites of services that were exposed to the GDPR and for websites of services that were not. As an example, Table 4 displays the results for whether privacy policies allow for the gathering of geolocation data, separated by the year in which the policies were gathered. Geolocation data is a particularly sensitive type of data because services (especially those that users access frequently using their mobile devices) can compile comprehensive movement profiles of their users.

We can see that only 1 out of 238 privacy policies adopted a provision restricting the collection of geolocation data prior to when the GDPR entered into force. In 2020, only 2 out of 238 services had provisions that outlawed the use of such data. At the same time, a substantial number of services went from implicitly allowing for the geolocation data in 2018 to explicitly allowing for the collection of geolocation data in 2020. In 2018, most services already allowed for the collection of such data, but 24% of websites had privacy policies in place that included broad, open-ended clauses that arguably covered the collection of many types of data, including geolocation data. By 2020, the number of services without a clause explicitly allowing for such data collection had dropped to less than half.

Importantly, this trend can be observed both for websites exposed to the GDPR and those that were not. A larger share of exposed services adopted explicit provisions allowing for the collection of geolocation data. However, the difference is slight (and not statistically significant).

Similar trends can be observed for privacy policy provisions addressing the scope of permissible data uses. To illustrate these changes, Table 5 displays results for whether privacy policies allow for the sharing of data with third-party advertisement services. Similar to the provisions regarding geolocation data described above, the number of privacy policies with provisions explicitly allowing for this sharing of data grew over the two years covered in our analysis. We again see this trend for privacy policies of services exposed to the GDPR and those not. Similar trends can be observed for all other types of information and data uses for which information is available in our dataset.

Table 4: Geolocation data collection provisions

	2018 geodata			
	Not allowed	Implicitly allowed	Explicitly allowed	Total
	(1)	(2)	(3)	
No GDPR Exposure	1 (1%)	24 (22%)	83 (77%)	108
GDPR Exposure	0	33 (25%)	97 (75%)	130
Total	1 (0%)	57 (24%)	180 (76%)	238

	2020 geodata			
	Not allowed	Implicitly allowed	Explicitly allowed	Total
	(1)	(2)	(3)	
No GDPR Exposure	2 (2%)	12 (11%)	94 (87%)	108
GDPR Exposure	0	12 (9%)	118 (91%)	130
Total	2 (1%)	24 (10%)	212 (89%)	238

Table 5: Third-party advertisers provisions

	2018 third-party advertising			
	Not allowed	Implicitly allowed	Explicitly allowed	Total
	(1)	(2)	(3)	
No GDPR Exposure	2 (2%)	19 (18%)	87 (81%)	108
GDPR Exposure	5 (5%)	18 (14%)	107 (82%)	130
Total	7 (3%)	37 (16%)	194 (82%)	238

	2020 third-party advertising			
	Not allowed	Implicitly allowed	Explicitly allowed	Total
	(1)	(2)	(3)	
No GDPR Exposure	0	10 (9%)	98 (91%)	108
GDPR Exposure	3 (3%)	6 (5%)	121 (93%)	130
Total	3 (1%)	16 (7%)	219 (92%)	238

3.4 Discussion

Our results show different trends in consent provisions for websites exposed to the GDPR and websites that are not exposed to the GDPR. Websites that are exposed to the GDPR are more likely to adjust their policies post-GDPR to contain a less stringent consent provision. Conversely, websites that are not exposed to the GDPR are more likely to adjust their policies post-GDPR to contain more stringent consent provisions. This is in striking contrast to the dimensions related to data collection collection in usage in which both GDPR exposed and non-exposed websites exhibit a similar trend of moving from implicit data rights to explicit data rights.

Our interpretation of what constitutes a more or less stringent consent requirement is not obvious. Most websites that change their consent provisions either drop any mention of consent from their privacy policies (in most cases because they drop their 'continued use as consent' provision) or by adding a 'continued use as consent' provision (when previously the policy did not mention consent at all). Viewing privacy policies

as contracts, a position taken by the Draft Restatement on Consumer Contracts and some in the literature (Bar-Gill, Ben-Shahar, and Marotta-Wurgler 2017), suggests that dropping a mechanism for policy changes means consumers must then **actively consent** to a policy change. A policy that drops any mention of consent is possibly opting into a stricter consent regime of active consent. Such a result would lead to the opposite conclusion to what we state above – firms exposed to the GDPR move closer to the GDPR in their U.S. policies by dropping mention of consent.

We believe it is unlikely that U.S. websites that drop any mention of consent intend to create a more stringent consent requirement. Firstly, it continues to be a topic of contention whether privacy policies should be considered contracts, with some scholars arguing that case law does not indicate a clear trend of recognizing policies as contracts (Klass 2019). Moreover, there are no cases that directly consider this question of policy updating and FTC guidance suggests that only retroactive changes require active consent. Lastly, even if a court would determine that a silent privacy policy requires active consent for policy changes, we are primarily interested in firm responses to the GDPR which rely on the beliefs of firms. Based on our conversations with industry practitioners, we consider it unlikely that websites that drop any mention of consent intend to increase consent stringency.

We interpret these results as suggestive that websites through their exposure to the GDPR learn about the impact of the GDPR's consent requirement. These websites may learn that requiring 'active consent' for privacy policy changes could be particularly costly if a critical number of consumers do not consent to the changes, preventing websites from using and sharing consumer data in ways not originally stipulated in the policy. Even a more lenient consent requirement of 'continued use as consent' could be costly if a website has data of consumers that are no longer active on the website. Websites that experience these costs may then choose to avoid consent requirements in policies that are not subject to the GDPR, and therefore adjust their U.S. policies to have less stringent consent provisions.

Our findings are consistent with prior research, which suggests that the GDPR's consent requirement has had implications for consumer behavior and website traffic. Schmitt, Miller, and Skiera (2021) show that, in the 18 months after the GDPR came into effect, the average number of unique website visitors declined for the $6,286$ most popular websites in E.U. countries. This result suggests that privacy features, such as requesting consumers explicit consent for data collection, effects consumer behavior. Similarly, Aridor, Che, and Salz (2020) show that the ability to opt-out of cookie tracking lead to a reduction of cookies collected by websites. These results suggest that the GDPR's consent requirement may have real implications for website's ability to obtain consent.

A learning story is not the only way to understand the move of GDPR exposed firms to less stringent consent requirements. For example, the need to comply with the GDPR may have prompted GDPR exposed firms to revise and update their privacy policies, where non-exposed firms may opt to continue with their existing policy. This explanation is consistent with the slightly higher number of policies with consent provisions changes for GDPR exposed firms (21.53% for GDPR exposed firms versus 16.35% for non-GDPR exposed firms) although this would not explain why firms not exposed to the GDPR opt into stricter consent regimes.

These results also reflect a specific context in which the GDPR has not raised privacy policy standards for policies not directly subject to the GDPR. The theory of unilateral regulatory globalization, commonly known as the 'Brussels Effect' (Bradford 2020), refers to a situation in which firms comply with E.U. law even when not directly subject to it. If it is costly for businesses to treat consumers in different jurisdictions

differently, this could lead to firms with consumers in the E.U. to treat everyone in accordance with GDPR standards. However, in this context we see that websites with GDPR exposure choose to be less stringent with respect to consent.

Acknowledgments

We are grateful for useful comments and suggestions provided by Anu Bradford, Sophie Calder-Wang, Oliver Hart, Greg Klass, Sarath Sanga, Rory Van Loo, Florencia Marotta-Wurgler and Ari Waldman. We thank Nuri Albayrak, Jiyoung Kim, and Sahana Ragunathan for superb research assistance.

Appendix

Our online appendix can be found here.

References

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman. 2016. "The economics of privacy." *Journal of economic Literature* 54 (2): 442–92.

Aridor, Guy, Yeon-Koo Che, and Tobias Salz. 2020. "The economic consequences of data privacy regulation: Empirical evidence from gdpr." *NBER working paper,* no. w26900.

Bar-Gill, Oren, Omri Ben-Shahar, and Florencia Marotta-Wurgler. 2017. "Searching for the common law: The quantitative approach of the restatement of consumer contracts." *The University of Chicago Law Review,* 7–35.

Bar-Gill, Oren, and Kevin Davis. 2010. "Empty promises." *S. Cal. L. Rev.* 84:1.

Bergelson, Vera. 2003. "It's personal but is it mine-toward property rights in personal information." *UC Davis L. Rev.* 37:379.

Bradford, Anu. 2020. *The Brussels effect: How the European Union rules the world.* Oxford University Press, USA.

Calzolari, Giacomo, and Alessandro Pavan. 2006. "On the optimality of privacy in sequential contracting." *Journal of Economic theory* 130 (1): 168–204.

Davis, Kevin, and Florencia Marotta-Wurgler. 2019. "Contracting for Personal Data." *New York University Law Review* 94:662–705.

Dosis, Anastasios, and Wilfried Sand-Zantman. 2019. "The ownership of data." *Available at SSRN 3420680.*

Feri, Francesco, Caterina Giannetti, and Nicola Jentzsch. 2016. "Disclosure of personal information under risk of privacy shocks." *Journal of Economic Behavior & Organization* 123:138–148.

Frankenreiter, Jens. 2021. "The Missing 'California Effect' in Data Privacy Law." *Yale Journal on Regulation, forthcoming.*

Frankenreiter, Jens, and Yoan Hermstruewer. 2021. "Privacy's Great Shock."

Grossman, Sanford J, and Oliver D Hart. 1986. "The costs and benefits of ownership: A theory of vertical and lateral integration." *Journal of political economy* 94 (4): 691–719.

Hann, Il-Horn, Kai-Lung Hui, Sang-Yong T Lee, and Ivan PL Png. 2008. "Consumer privacy and marketing avoidance: A static model." *Management science* 54 (6): 1094–1103.

Hart, Oliver, Andrei Shleifer, and Robert W Vishny. 1997. "The proper scope of government: theory and an application to prisons." *The Quarterly Journal of Economics* 112 (4): 1127–1161.

Karnow, Curtis EA. 2021. "The Internet and Contract Formation." *Berkeley Bus. LJ* 18:135.

Klass, Gregory. 2019. "Empiricism and privacy policies in the restatement of consumer contract law." *Yale J. on Reg.* 36:45.

Kretschmer, Michael, Jan Pennekamp, and Klaus Wehrle. 2021. "Cookie Banners and Privacy Policies: Measuring the Impact of the GDPR on the Web." *ACM Transactions on the Web (TWEB)* 15 (4): 1–42.

Lancieri, Filippo. 2022. "Narrowing Data Protection's Enforcement Gap." *Main Law Review* 74 (1): forthcoming.

Lenard, Thomas M, and Paul H Rubin. 2010. "In defense of data: Information and the costs of privacy." *Policy & Internet* 2 (1): 149–183.

Levitin, Adam J, Nancy S Kim, Christina L Kunz, Peter Linzer, Patricia A McCoy, Juliet M Moringiello, Elizabeth A Renuart, and Lauren E Willis. 2019. "The Faulty Foundation of the Draft Restatement of Consumer Contracts." *Yale J. on Reg.* 36:447.

Litman, Jessica. 2000. "Information privacy/information property." *Stanford Law Review,* 1283–1313.

Mangini, Vincenzo, Irina Tal, and Arghir-Nicolae Moldovan. 2020. "An empirical study on the impact of GDPR and right to be forgotten-organisations and users perspective." In *Proceedings of the 15th International Conference on Availability, Reliability and Security,* 1–9.

Matwyshyn, Andrea M. 2013. "Privacy, the Hacker Way." *S. Cal. L. Rev.* 87:1.

McDonald, Aleecia M, and Lorrie Faith Cranor. 2008. "The cost of reading privacy policies." *Isjlp* 4:543.

Obar, Jonathan A, and Anne Oeldorf-Hirsch. 2020. "The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services." *Information, Communication & Society* 23 (1): 128–147.

Posner, Richard A. 1981. "The economics of privacy." *The American economic review* 71 (2): 405–409.

Richards, Neil. 2021. *Why Privacy Matters.* Oxford University Press.

Romanosky, Sasha, and Alessandro Acquisti. 2009. "Privacy costs and personal data protection: Economic and legal perspectives." *Berkeley Tech. LJ* 24:1061.

Santos, Cristiana, Midas Nouwens, Michael Toth, Nataliia Bielova, and Vincent Roca. 2021. "Consent Management Platforms under the GDPR: processors and/or controllers?" In *Annual Privacy Forum,* 47–69. Springer.

Schmitt, Julia, Klaus M Miller, and Bernd Skiera. 2021. "The Impact of Privacy Laws on Online User Behavior." *arXiv preprint arXiv:2101.11366.*

Shahar, Omri Ben, and Lior Jacob Strahilevitz. 2016. "Contracting over Privacy: Introduction." *Journal of Legal Studies* 45:S7.

Sloan, Robert H, and Richard Warner. 2014. "Beyond notice and choice: Privacy, norms, and consent." *J. High Tech. L.* 14:370.

Solove, Daniel, and Paul M. Schwartz. 2019. *Privacy Law Fundamentals 2019.* IAPP.

Stigler, George J. 1980. "An introduction to privacy in economics and politics." *The Journal of Legal Studies* 9 (4): 623–644.

Strahilevitz, Lior Jacob, and Matthew B Kugler. 2016. "Is privacy policy language irrelevant to consumers?" *The Journal of Legal Studies* 45 (S2): S69–S95.

Varian, Hal R. 2002. "Economic aspects of personal privacy." In *Cyber Policy and Economics in an Internet Age,* 127–137. Springer.

Victor, Jacob M. 2013. "The EU general data protection regulation: Toward a property regime for protecting data privacy." *Yale LJ* 123:513.

Viljoen, Salome. 2021. "A relational theory of data governance." *Yale LJ* 131:573.

Westin, Alan F. 2000. "Intrusions." *Public Perspective* 11 (6): 8–11.