PROGRAM ON ECONOMICS & PRIVACY
ANTONIN SCALIA LAW SCHOOL

THE UNIVERSITY OF ARIZONA
James E. Rogers College of Law

Responses to Written Questions

The Likely Impact of the GDPR and the CCPA on Innovation and Consumer Welfare

April 3, 2019

Jane Bambauer
Professor of Law, University of Arizona
Director, Program on Economics & Privacy at George Mason University Antonin Scalia Law School

April 3, 2019

Chair Graham and members of the Senate Judiciary Committee:

I admire the work you have done to understand the issues involved in drafting effective privacy legislation. It was an honor to think through your thoughtful and probing questions. My answers appear below.

QUESTIONS FROM SENATOR LINDSEY O. GRAHAM

1. **What are the specific areas of the CCPA that could have a negative impact on competition and innovation? What areas of the CCPA need more clarity, improvement, or removal?**

The parts of the CCPA most likely to stifle innovation and competition are the right to deletion (§1798.105), the nondiscrimination clause (§1798.125) (requiring that a business not "discriminate" against a consumer who has exercised one of their rights, including a demand to delete data or to forbid its sale), and the definition of "deidentified" information (§1798.140(h)) (setting unclear and potentially difficult-to-achieve standards for information that has been sufficiently deidentified).

Let me provide two examples to show why these clauses will be harmful to competition and consumer welfare.

A. Behavioral Advertising

First, let's revisit the question you addressed to the first panel at last month's hearing. You wondered how important tracking-based behavioral advertising was to the overall ecosystem and to a service provider's bottom line when you visit a website with golf-related content. The panel of experts seemed to agree that contextual advertising would provide roughly the same revenue as behavioral advertising in this example. This exchange left the impression that behavioral advertising could be killed off with little effect. This is not so. The aggregate amount spent on advertising is likely to be reduced somewhat. More importantly, the advertising dollars that *are* spent will be redistributed in ways that do not serve the public.

Imagine you visit two websites: one for golf enthusiasts, and one with general interest news articles. With the aid of data aggregators and advertising exchanges, both websites will profit about equally from your visit because both could deliver advertising related to golf, or related to your location, or to your other interests and preferences. On the other hand, if tracking were prohibited either as a default or because you opted out, then the prospects of the two websites are very different. The advertising dollars will flock to the golf website and away from the general interest news website. Without behavioral advertising, websites will face market pressure to produce more content that allows advertisers to match themselves to end users through context.

The implications may not seem so troubling when comparing golf to general interest websites, but imagine how these changes would affect competing news websites, or content about health and medical information. The high value content that will attract advertisers are the stories that help fragment readers into market segments by signaling their ideology or other personal characteristics based on content.

My concerns here are not theoretical; a study of the effects of the EU Data Protection Directive (the privacy framework that was in place in Europe before the GDPR took effect) found that the efficacy of advertising in Europe was reduced, but the reductions were not uniform.[1] "Websites that had general content unrelated to specific product categories (such as news and media services) experienced larger decreases in ad effectiveness after the laws passed than websites that had more specific content (such as travel or parenting websites). Customers at travel and parenting websites have already identified themselves as being in a particular target market, so it is less important for those websites to use data on previous browsing behavior to target their ads."[2]

Moreover, some of the panelists suggested that advertising revenues for a website are not highly dependent on behavioral advertising. This is not correct. Behavioral advertising has a "click-through" rate 670% higher than contextual advertising[3], and thus an ad placed with the help of behavioral data will pay the website much more than an ad placed without data (on average). The fact that the New York Times' revenues have continued to rise since ending behaviorally targeted ads in Europe may be explained by the steady increase in its readership and subscriber base since the election of President Trump; the *Times* had just over one million subscribers before the 2016 election and now has over four million.[4] Changing the privacy rules changes content and services.

Online tracking and data aggregators are popularly understood to reduce the power of consumers, but I see it differently. When ad networks can access data from aggregators and target ad dollars more effectively, it allows the consumers rather than the content-providers to control where the money goes. Each of us will reward whichever websites we like to visit, even if they don't cater to a well-defined market segment. Consumers who wish to opt out of tracking rarely understand that they are rewarding content providers who produce inherently commercial

[1] Avi Goldfarb & Catherine E. Tucker, *Privacy, Regulation, and Online Advertising,* 57 MANAGEMENT SCIENCE 57 (2011).

[2] *Id. See also* Catherine E. Tucker, *The Economics of Advertising and Privacy* ("The reduction we measured is particularly pronounced for websites offering content that is not easy to match to advertising (such as news websites and media services). It is also more pronounced for non-intrusive and smaller ads, since their appeal rests primarily on the presentation of informative rather than attention-grabbing messages. This suggests that privacy regulation might change the web landscape in unanticipated ways, with advertising becoming even more intrusive, and marketers shifting their media buys away from sites such as news providers that are difficult to match with relevant advertising.")

[3] Jun Yan et al., *How Much Can Behavioral Targeting Help Online Advertising?,* PROC. OF THE 18TH INT. CONF. ON WORLD WIDE WEB 261 (2009).

[4] Jaclyn Peiser, *New York Times Tops 4 Million Mark in Total Subscribers,* N.Y. TIMES (November 1, 2018); Rani Molla, *The 'Trump Bump' in the New York Times' Digital Subscription Growth Is Over,* RECODE (August 9, 2018). Another possibility is that the reporting in the *Times* may have changed since the election of Donald Trump as well such that the context of the articles are more ideologically skewed, and therefore send a signal for market segmentation.

or highly segmented content. The winners in this system include Facebook and Google, because these companies can match ads to users without any help from third party data aggregators.[5]

Given that the effects of behavioral advertising and other targeted services are not well understood or appreciated, I recommend holding off creating new legislation that allows consumers to delete or opt out of the sale of their data. At the very least, a privacy law that grants consumers these rights should not force companies to provide the same service to the consumers that exercise their rights. Otherwise, the law would force the people who do not opt-out to subsidize the costs of the people who do.

B. AI and Healthcare

For a second example, consider emerging health technologies. Advances in Machine Learning are already changing the practice of medicine because new programs can digest and learn from vast amounts of data about other patients to make customized recommendations. These programs can also update the therapeutic recommendations in real time as it receives and learns from the health outcomes data of its users.[6] Other applications can help hospitals predict how much staffing they will need based on non-health data that might have clues about an influx of new patients. But some of the most promising applications would have to merge data from disparate sources both within and outside the health sector, and the incentives to do this are likely to depend on being able to collect, purchase, and reuse personal data.[7]

The consent and control concepts in the CCPA cannot be easily applied to these developments. What would it mean for a patient to demand data deletion, or to opt out of sale? And in what way would the program have to avoid "discriminating" against a person who exercises these rights? Would the program have to be designed to work as well as possible while removing the data from the user? Would the company have to remove that user's data from the pooled dataset used to optimize the program's recommendations to all of its users, or could it safely rely on the exemption for deidentified data? Is it sensible or fair to allow a user to benefit from the past data shared by other patients without contributing his or her own data to the system? If the company declines service or offers inferior services to a user who exercises the deletion or opt-out right, it must be prepared to show that the differential treatment is "reasonably related to the value provided to the consumer by the consumer's data"[8]—a vague standard that is sure to deter experimentation.

[5] Indeed, the privacy rules promulgated by the FCC in 2016, but which never went into effect, were likely to *increase* the dominance of Facebook and Google by preventing internet service provider companies like Verizon from entering the market for data tailoring. *See* Howard Homonoff, *Facebook and Google Win From New FCC Privacy Regulations, And Everyone Else Loses,* FORBES.COM (October 31, 2016); Roger Entner, *FCC Should Whack Privacy Rules that Favor Google and Facebook,* THE HILL (February 15, 2017).

[6] *See, e.g.,* the discussion of Ginger.io in Basel Kayyali et al., *The Big-Data Revolution in US Health Care: Accelerating Value and Innovation,* MCKINSEY REPORT (2013).

[7] Sonja Marjanovic et al., *Understanding Value in Health Data Ecosystems,* RAND RESEARCH REPORT (2017).

[8] This is particularly difficult language to interpret because the value of a consumer's data must be "provided to the consumer." It is not clear how direct the return of value to consumers must be.

The CCPA has the goal of siloing and minimizing personal data at a time when lawmakers should be facilitating more data-sharing. It is already very difficult to link databases and make them usable for a variety of purposes, and in the context of health the problem has profound consequences. For example, in the wake of the removal of the drug Vioxx from the market, Richard Platt, a professor at Harvard Medical School, showed that the fatal side effects of the drug could have been detected in just three months, rather than the five years it actually took, if health data across the country had been merged.[9] Large datasets of merged records will be vital for the success of some of the future Machine Learning applications, too. Since individual hospitals, clinics, and insurers have little incentive to give away their data, money (from sales to data aggregators) is a badly-needed lubricant.

The same issues arise outside the healthcare context, albeit with different (often lower) stakes. With better information, loans and auto insurance plans can be better matched and managed, reducing prices, defaults, and accidents.[10] Data collection, repurposing, and even its sale are likely to be necessary if the U.S. wants to continue to be the world's leader in service improvements and innovation.[11]

QUESTIONS FROM SENATOR BOOKER

1. **Big Data Discrimination**
    a. **In your view, is a private right of action critical to protecting the civil rights of individuals affected by data collection and disclosure practices?**

Yes, but such a private right of action should come from existing public accommodations and antidiscrimination laws rather than privacy laws like the California Consumer Privacy Act. (Indeed, the CCPA does very little to identify or rectify discriminatory effects of Big Data.) It is likely that the causes of action available under public accommodations laws and the Civil Rights Act will have to be updated by courts or legislatures to better fit automated scoring and decision-making, but this updating will first require sustained attention about the goals of civil rights laws. (See my discussion in the next response below.)

b. **How easy is it for seemingly non-sensitive information like a ZIP Code to become a proxy for protected class or other sensitive information? How can that information be used to discriminate?**

[9] Barbara J. Evans, *Seven Pillars of a New Evidentiary Paradigm: The Food, Drug, and Cosmetic Act Enters the Genomic Era,* 85 NOTRE DAME L. REV. 419, 455-456 (2010).

[10] Jin-Hyuk Kim & Liad Wagman, *Screening Incentives and Privacy Protection in Financial Markets: A Theoretical and Empirical Analysis,* 46 RAND J. OF ECON. 1 (2015) (home loans); Prof. Omri Ben Shahar, Testimony at the Federal Trade Commission Informational Injury Workshop, 168-69 (December 12, 2017) (discussing the benefits of privacy-invasive auto insurance plans).

[11] China's highly concentrated Internet and financial tech sector is poised to leverage its data troves to make advances in medical AI. Tom Simonite, *How Health Care Data and Lax Rules Help China Prosper in AI,* WIRED.COM (January 8, 2019).

I appreciate this question because it gets at, and helps define, the societal harms that privacy laws are meant to prevent. One principal concern about the collection and sale of large amounts of personal data is that the data may be used in ways that perpetuate or exacerbate inequalities along race, gender, and class lines. Books like *Weapons of Math Destruction*[12] and academic articles such as *Big Data's Disparate Impact*[13] help explain how this can happen. But it is important to separate the theoretically possible from the pragmatically probable. Thoughtful uses of algorithms have the potential to reduce bias—indeed they have tremendous promise for optimizing decisions to comply with a particular notion of fairness far better than any human system could. But what "fairness" means is a distinctly human issue.

To illustrate, consider the much maligned[14] recidivism risk scores (such as the COMPAS scores) that are intended to measure the chance that a person in the criminal justice system will be arrested for another crime if released.[15] These scores could be biased by virtue of the chosen outcome variable—likelihood of arrest. There is good reason to believe that police are often more likely to detect and arrest Latinos and African-Americans than whites who commit the same crimes, in part because low SES and minority neighborhoods are patrolled more often.[16] So even before predictors like zip code have the chance to create racial disparities, the choice of predicted outcome variable can cement inequities. (What we'd really like to predict is likelihood of *committing a crime*, but it is not surprising that arrests are used as the closest available alternative.)

If the chosen outcome is sufficiently legitimate, an algorithm can still have a disparate impact on minorities even when race is not included as a predictor in the model. However, it is not obvious which disparate impacts should be actionable, particularly when the variables that correlate with race also correlate with the outcome that is being predicted. In human systems, experts often look for differences in the success or failure rates for members of different groups who are scored or treated the same way. (Becker 1957). For example, in Floyd v. NYPD, the case challenging the practices of the NYPD stop and frisk program, the fact that frisks of African-Americans were *less likely* to produce a weapon than frisks of whites showed that police were using a different standard for minorities. This problem, though, will usually not occur with machine learning algorithms because they work backwards from the predicted outcome. Every variable that is included in a machine learning model will be used in service of making the prediction score match the outcomes that are observed later. However, there are other measures of fairness that can be considered, such as whether a member of a disadvantaged group is more likely to produce a false positive, or less likely to get the benefits of a false negative, than a member of the majority group. These different versions of "fairness" are in tension with one another, so making improvements for one can make others worse.

[12] CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION (2016)

[13] Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact,* 104 CAL. L. REV. 671 (2016).

[14] Julia Angwin et al., *Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks,* PROPUBLICA (2016).

[15] These scoring systems do not use the vast amount of data and sophisticated processing that AI/machine learning is capable of, but they are valuable for showing that problems with bias are real, and also misunderstood.

[16] *See, e.g.,* Samuel Gross & Katherine Barnes, *Road Work: Racial Profiling and Drug Interdiction on the Highway,* 101 MICH. L. REV. 651 (2002).

At some point, it may be necessary to have public commitments to a specific definition of discrimination and fairness in the digital era. That definition would have to work out compromises between competing interests in accuracy, equitable distribution of errors, and race-neutral treatment.[17] But it would be premature to do so now because the *human* decision-making that digital tools can augment or replace are at least as flawed along the same metrics (often more so, and with less transparency).[18] The few studies that compare the effects of using recidivism scores to bail and sentencing decisions made in their absence find that jailing is reduced for members of *every* race, and that pretrial detentions could be even further reduced if judges are removed from the decision-making.[19]

These findings are at odds with the way COMPAS scores are portrayed in the popular media, but they are consistent with studies in other areas finding that machine algorithms, as actually implemented, tend to improve race and gender disparities rather than exacerbate them. One study found that a hiring algorithm had a positive effect on racial minority applicants compared to the status quo recruiting process[20], and another found that a machine learning algorithm could be used to select corporate directors who were more likely to be female *and* more likely to outperform the directors actually selected by the boards.[21] Home mortgages will tend to have lower interest rates and lower default rates when banks are able to make use of Big Data profiles that go beyond the income and credit score information that is typically collected, suggesting that machine learning has promise for helping low income applicants prove that they are more creditworthy than loan officers have historically thought.[22] And a study of risk scores used by child protective service centers found that when human decision-makers deviated from the

[17] Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan, and Cass Sunstein provide a good start. Jon Kleinberg et al., *Discrimination in the Age of Algorithms,* NBER WORKING PAPER NO. 25548 (2019) ("The existence of disparate impact is clear; the data will prove it, and also show its magnitude. We can also attempt to specify the practice that gives rise to those impacts. Then the question is the standard one: Can the disparate impact be justified, given the relevant standard? That is the same question that would be asked if an algorithm were not involved. The presence of the algorithm goes further – it makes it possible to quantify the tradeoffs that are relevant to determining whether there is 'business necessity' (or some other justification for disparate impact). Algorithms by construction produce not just a single ranking of applicants. They can produce a set of rankings that vary based on one's tolerance for disparate impact. For each of these rankings, they additionally quantify their effect on the objective, such as sales. This allows us to answer exactly the question, 'What is the magnitude of the disparate impact, and what would be the cost of eliminating or reducing it?'").

[18] Policy analysis should compare existing and future problems of AI to the counterfactual by asking what sort of distributional outcomes we could expect in the absence of the technology. *See* Bo Cowgill & Catherine Tucker, *Algorithmic Bias: A Counterfactual Perspective,* NSF WORKING PAPER: TRUSTWORTHY ALGORITHMS (2017).

[19] John Kleinberg et al., *Human Decisions and Machine Predictions,* NBER WORKING PAPER NO. 23180 (2017). *See also* Megan Stevenson, *Assessing Risk Assessment in Action,* 103 MINN. L. REV. 303 (2018) (finding that a law requiring judges to at least consider risk assessment scores caused a short-term reduction in pretrial detension, but that the reduction faded over time as judges returned to their previous habits. Stevenson also found that pretrial arrests increased when the scores were influencing judge's decisions, but pretrial arrests for violent crimes went down slightly. Stevenson did not find promising reductions in the race gap, though.)

[20] Bo Cowgill, *Productivity in Humans and Algorithms: Theory and Evidence from Resume Screening* (working paper).

[21] Isil Erel et al., *Selecting Directors Using Machine Learning,* NBER WORKING PAPER NO. 24435 (2018).

[22] Jin-Hyuk Kim & Liad Wagman, *Screening Incentives and Privacy Protection in Financial Markets: A Theoretical and Empirical Analysis,* 46 RAND J. OF ECON. 1 (2015). *See also* Will Dobie et al., *Measuring Bias in Consumer Lending*, NBER WORKING PAPER NO. 24953 (2018). By contrast, studies of human-driven mortgage lending decisions continue to find racial bias even after controlling for credit history and income. *Id.*

recommendation of a scoring system, they tended to screen more black families into the high risk treatment.[23]

So, it is very likely that some factors used by Big Data algorithms will correlate with race. But this is true of the factors used by human and "little data" decision-makers, too, and it does not mean on its own that something untoward or legally discriminatory has occurred. With guidance, antidiscrimination law can adapt so that notions of discriminatory treatment or disparate impact have even more coherence and clarity than they do today.

c. Significant amounts of data about us are gathered by companies most people have never heard of. Do we need a registry of data brokers, similar to what Vermont established last year?

I do not believe we need a national registry until there is better understanding and consensus about the risks and harms that data brokers pose.

2. Living without the Big 5

a. How would you respond to the following argument? "If people are uncomfortable with the data practices of certain tech companies, they simply shouldn't use their services."

Avoiding the data collection of the major tech firms able to collect information is highly impractical, and impossible for anybody who wishes to live with modern amenities. Yet this does not prove on its own that the companies are harming consumers or behaving anticompetitively. By way of analogy, it is nearly impossible to avoid using an Intel microchip.[24]

b. What does providing consent mean in a world where it's extremely difficult to avoid certain companies?

Kashmir Hill's article raises important questions about what value consent is meant to serve. Assuming no significant market failures, the largest tech companies evidently provide services of exceedingly high value not only directly to consumers, but also to the websites, companies, and friends with whom the consumers interact. Consent requirements enshrined in law could have the opposite effect from what is intended. They will favor the companies that have already made it—that have already been able to prove their value to consumers or to their preferred service-providers during an era when novel data practices could be employed without explicit consent. By contrast, young companies will have many more hurdles to persuade consumers to try out a new data-intensive service. This will be particularly true when those companies are challenging

[23] Alesandra Chouldechova, ***A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions,*** **81 PROC. MACHINE LEARNING RESEARCH 1 (2018)** (citing Alan J. Dettlaff et al., *Disentangling Substantiation: The Influence of Race, Income, and Risk on the Substantiation Decision in Child Welfare,* 33 CHILDREN AND YOUTH SERVICES REVIEW 1630 (2011).)

[24] Christopher Mims, *Just Four Companies Will Produce the Microchips on Which the Global Economy Depends,* BUSINESS INSIDER (April 26, 2013).

Google's and Facebook's backend ad delivery services and have only an indirect relationship with consumers.

3. The Do-Not-Track Experience

a. What purpose does a notice-and-consent regime serve if the most prominent consent mechanism is only regarded as a suggestion at best?

I regard Do-Not-Track as a failed experiment in part because it wrongly gave web users the illusion that they were not being tracked. Just as importantly, it also wrongly gave web users the illusion that privacy is costless, and that the services they enjoy would be the same even if DNT were respected. The experiment would have been less of a failure if user-facing websites and service providers warned individuals who signaled the DNT request that the request cannot or will not be honored because data-tracking is a necessary component for their economic or technical viability. This response would have further fueled resentment that many services operate on a "take it or leave it" basis when it comes to data privacy, but in the long run it may have educated consumers that data collection and repurposing is critical to much of what works and funds the modern Internet.

b. How much faith should the failure of Do-Not-Track give us in the ability of the industry stakeholders to regulate themselves?

In my writing, I rarely refer to "industry self-regulation." There is little reason to believe that a corporation has the right incentives to purposefully raise the public's interest above its own self-interest. But tech companies are greatly constrained (and in this sense, regulated[25]) by market discipline. As long as competitors have the ability to undercut a company by offering better service, or the same service at a lower "price" (and in the context of online services, a lower price might mean the company collects less data, or inspires more trust), the company will have all the incentive it needs to give the consumer as good a deal as they can. These pressures have induced organizations like the Digital Advertising Alliance to develop and enforce industry standards that websites and apps can subscribe to on a voluntary basis.

c. In your view, should this approach be abandoned, or would federal legislation requiring companies to respect the Do-Not-Track signal breathe new life into the mechanism?

Federal legislation should not require companies to respect the DNT signal while also offering the same services to those who use it. If legislation helps foster a program like Do-Not-Track in order for consumers to easily express their preferences, it should allow companies to decline service to those who opt out of data collection.

[25] The market is one of the forms of regulation, in addition to law, norms, and architecture, in Lawrence Lessig's famous taxonomy. LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE (1999).

To understand why I recommend a legal framework that permits a "take it or leave it" response, please see my lengthy discussion of behavioral advertising in my response to Senator Graham's question (pages 2-4). Changing the privacy rules will change online content and services in ways that are not compatible with most people's values. Moreover, consumers who turn on the Do-Not-Track signal unwittingly reward Facebook and Google, because these companies can match ads to users without any help from third party data aggregators.[26]

Given that the effects of behavioral advertising and other targeted services are not well understood or appreciated, I do not recommend bolstering Do-Not-Track mechanisms through federal regulation. If Do-Not-Track is resuscitated, lawmakers should allow companies to offer inferior services or decline service altogether to those who turn it on. Otherwise, the law would force the 75% of people who do not turn DNT on to subsidize the costs of the 25% of people who do.

4. Preemption of the California Consumer Privacy Act

a. In your view, should a federal data privacy law preempt state data privacy laws? Why?

Yes, federal data privacy law should preempt state data privacy law for at least three reasons. First, privacy law has a complicated relationship to First Amendment. Privacy rules must be crafted within the bounds of free speech precedent that have routinely recognized that state-imposed restrictions on information flows must be justified by, and tailored to, concrete risks of harm.[27] Striking the right balance between privacy and free speech requires much more care for omnibus privacy laws that are expected to be implemented across industries than for narrower, industry-specific privacy rules.

Second, unlike the industry-specific privacy laws (like HIPAA and FERPA) which tend to govern companies with local clientele, general data privacy laws will force companies that operate in a national market to comply with a patchwork of regulations. Thus, it will be the most restrictive state law that guides the development of most data-driven companies. That means that the benefits of the most restrictive state law will spill over into other states, but so too will the drawbacks. For all-industry privacy laws like the CCPA, the drawbacks are likely to severely dampen service quality and innovation.

Third, and relatedly, state laws that provide greater privacy rights for its constituents are not necessarily more protective of consumers' interests as a whole. For example, imagine if the Fair Credit Reporting Act did not preempt state law, and a state passed a law providing greater privacy protections by prohibiting banks and other creditors from accessing credit reports or

[26] Indeed, the privacy rules promulgated by the FCC in 2016, but which never went into effect, were likely to *increase* the dominance of Facebook and Google by preventing internet service provider companies like Verizon from entering the market for data tailoring. *See* Howard Homonoff, *Facebook and Google Win From New FCC Privacy Regulations, And Everyone Else Loses,* FORBES.COM (October 31, 2016); Roger Entner, *FCC Should Whack Privacy Rules that Favor Google and Facebook,* THE HILL (February 15, 2017).

[27] Sorrell v. IMS Health Inc., 564 U.S. 552 (2011).

income information about loan applicants. Such a law would be more protective of consumers' privacy, but it would not protect consumers in the most important senses. It would cause consumers in that state to lose access to affordable credit. Thus, more stringent state privacy protections should not be equated with greater consumer protection.

b. **In your view, should a federal data privacy law implement the requirements of the California Consumer Privacy Act as a floor? If not, please explain the most significant change you would suggest.**

No. Federal law should facilitate a means to identify and incisively reduce specific data uses and practices that are likely to cause harm to consumers. This approach is orthogonal to the California Consumer Privacy Act because it does not center around user consent. Consent rules expect each individual consumer to develop expertise on the data practices of the moment and guard their interests themselves.

There is some risk that consumers will fail to guard their interests when risk arises. The greater concern, though, is that consumers and companies will be reluctant to permit data to be gathered or sold or reused in novel ways, impoverishing the advances we've already seen and will continue to see in the digital economy.

To see why this is so, please see my lengthy discussion of emerging health technology in my response to Senator Graham's question (pages 4-5).

c. **The specific wording of a proposed preemption provision will invite considerable debate in Congress and, ultimately, will still require courts to interpret and clarify the provision's scope. Should the Federal Trade Commission have notice-and-comment rulemaking authority to aid in the statute's interpretation and to clarify which types of state laws are preempted? Or, alternatively, is case-by-case adjudication of multiple state privacy laws preferable? Would rulemaking authority obviate the need for Congress to solve each and every preemption issue in drafting the text?**

I do not anticipate that the drafting of a preemption clause would be unusually difficult in this context. A federal law that carefully balances the competing interests related to data privacy and security should inhabit the field for legal redress related solely to the collection, storage, or dissemination of personal information. Exceptions to preemption should be limited to issues that would not affect the foundations of the digital economy.[28] The policy considerations related to preemption are not at the core of the Federal Trade Commission's expertise, so it would make more sense for federal courts to address any ambiguities that statutory drafting has not resolved.

[28] Examples might include mandated disclosure rules requiring companies to explain their data practices, or traditional privacy torts like Public Disclosure of Private Facts that address single, more egregious acts of privacy violation rather than routine business practices.

d. **The preemption language in, for example, the amendments to the Fair Credit Reporting Act was included as part of a heavily negotiated process in which consumers received a package of new rights in exchange for certain preemption provisions. Rather than centering the federal privacy bill debate on the existence of a preemption provision, shouldn't our starting point be: "Preemption in exchange for what?" In other words, what basic consumer protections should industry stakeholders be willing to provide in exchange for preemption? Do the requirements of the California Consumer Privacy Act represent a good floor for negotiating preemption?**

For all the reasons explained above, the California Consumer Privacy Act should not be the floor for a federal statute. But consumers want and deserve some reassurance, through enactment of new law, that regulators are not asleep at the wheel. FTC rulemaking authority that helps focus the agency on privacy and data security risks would provide that assurance. At the same time, it is important to insulate the FTC from popular pressure to punish large tech companies when the theories of consumer harm are poorly defined or unsupported by evidence. Thus, a new federal statute should authorize FTC rulemaking power with specific benchmarks for cognizable consumer harm.

5. **At the hearing, several witnesses indicated that opt-out requirements that permit users to tell companies not to process and sell their data are more protective of data privacy and more conducive to the user experience, since they do not impose the "take it or leave it" dynamic that opt-ins tend to create. In your view, are opt-outs preferable to opt-ins in terms of both data privacy and user experience? Why?**

I agree that an opt-in rule tends to frustrate consumers who can become fatigued from all of the pop-up windows and banners that clutter their experience.[29] Also, outside the World Wide Web experience or devices with a screen, opt-in mechanisms are harder to design. (This would be relevant for the Internet of Things.)

It is also true that if the law permits companies to offer their services on a take-it-or-leave-it basis, companies may counterintuitively wind up with more data about their users (although presumably at least a few end users will occasionally choose to not proceed.) However, if a company found that more than a trivial number of its users were opting out of data collection that is important for the firm's business model, the firm could force a take-it-or-leave it at the point of opt-out. Thus, while I recommend an opt-out rule over an opt-in one, my reasons for doing so stem from the evidence that on balance, consumers benefit from modern data practices even while they distrust them.

6. **At the hearing, several witnesses also indicated that the Federal Trade Commission, and perhaps state attorneys general, should have primary enforcement authority**

[29] Pop-up windows and overlays are experienced as annoying by website visitors in the context of advertising. Steven M. Edwards et al., *Forced Exposure and Psycological Reactance: Antecedents and Consequences of the Perceived Intrusiveness of Pop-Up Ads,* 31 J. ADVERTISING 83 (2013).

for data privacy violations. In your view, what additional authority and/or resources would the FTC need to perform this function effectively?

The draft bill that I circulated with my written testimony shows how the Federal Trade Commission could be empowered to make informed interventions on the public's behalf. I recommend rulemaking authority related to data security and privacy-related harm. The privacy-related harms include: (1) a data practice that produces unjustified risk of substantial injury; (2) a data practice that was not effectively disclosed and for which a majority of consumers (or a majority of a particular group of consumers) would have chosen another option available to them at the time; or (3) failing to contractually restrict and reasonably monitor business partners with whom the company shares personal information.

To do this work, especially to create a rigorous evidence base for wise regulatory interventions, the FTC would benefit from additional resources to support an expansion of its Bureau of Economics.

QUESTIONS FROM SENATOR CHUCK GRASSLEY OF IOWA

1. **Please briefly explain the importance of transparency and ensuring that consumers can make informed decisions about the information they share.**

Transparency can be valuable for giving consumers, consumer watch dog organizations, and regulators insight into how data is used, but it is very difficult even for experts to understand how a particular data practice will ultimately affect a consumer, and whether that effect will be positive or negative. Understanding will be all the more difficult as AI and Machine Learning take over a greater number of increasingly complex tasks. For this reason, companies and regulators, rather than individual consumers, should be charged with the task of anticipating or responding to unjustified risks and harms.

2. **Often times, comprehensive regulations end up just benefiting the large, entrenched entities that have teams of lawyers to ensure compliance. Should small businesses be treated differently in any federal data privacy framework? And if so, how?**

The theoretical and empirical literature indeed suggests that data privacy laws have adverse effects on smaller and newer firms.[30] An approach to privacy that centers around harmful uses of data rather than user control will be less burdensome for all businesses (including startups) because they will not have to respond to individual consumers' requests, so long as their business model is a fair one. If a new federal statute does include the mechanisms of user control that make compliance more difficult, the law could exempt firms that have annual revenues below the U.S. Small Business Administration's standard for its industry. The drawback to this approach is that the smaller, fly-by-night operations may be more likely to do something reckless with consumer data. However, since fraudulent and illegitimate businesses are unlikely to be

[30] James Campbell et al., *Privacy Regulation and Market Structure*; Jian Jia..

deterred by federal privacy law anyway, an exemption for the legitimate small businesses that really are adding value to the economy would be sensible.

3. **If Congress enacts federal data privacy legislation, how do we ensure that companies are still incentivized to innovate in their privacy and data protections, rather than just 'check the box' of regulatory compliance?**

A virtue of the harm- or risk-centered approach, rather than a consent and control frameworks, is that the standard of unjustified risk rises over time. The FTC's unfairness standard is one that operates at the margins: a practice that is reasonable today will not be tomorrow if a better, less risky option is developed in the interim. By contrast, a consumer who agrees to data collection today is unlikely to revisit that decision if circumstances change later.

4. **How do we best craft a federal data privacy law that keeps pace with our ever-evolving tech and data landscape? And can we do that without giving unfettered discretion to the regulators?**

This is an insightful question. So often, calls for GDPR-style regulation assume that user consent and control are the best way to allow privacy rights to keep pace with technological change, but users are at an even greater disadvantage than regulators when trying to assess the ultimate impact of a new data practice. On the other hand, if an agency like the FTC is given too much flexibility to respond as it sees fit, the formal and informal rules can change rapidly based on the idiosyncrasies of the regulators or the zeitgeist of the times. For this reason, I recommend giving the FTC rulemaking authority, but statutorily tying that authority to specific definitions of consumer risk and harm. For example, in my draft privacy bill (appended to my pre-hearing written testimony), I recommended FTC rulemaking authority (1) to define reasonable data security, (2) to identify when a data practice produces unjustified risk of substantial injury; (3) to identify a data practice that was not effectively disclosed and for which a majority of consumers (or a majority of a particular group of consumers) would have chosen another option available to them at the time; and (4) to identify when a firm has failed to contractually restrict or reasonably monitor business partners with whom the company shares personal information.

<u>QUESTIONS FROM SENATOR MAZIE K. HIRONO</u>

1. In your testimony, you called the CCPA's anti-discrimination provisions "vague concepts."

   What do you consider vague about the anti-discrimination provisions? How would you suggest making the provisions more clear?

Section 1798.125 of the CCPA prohibits businesses from discriminating against a consumer when the consumer exercises one of the privacy rights (including the right to deletion or to opt out of data sale). The business cannot deny service or offer the service at a lower quality or higher price unless the difference in treatment is "reasonably related to the value provided to the consumer by the consumer's data."

Consider a website that earns its revenues from behaviorally targeted advertising. Some of the panelists at last month's hearing suggested that contextual advertising commands roughly the same price to websites as behavioral advertising. This is not correct. Behavioral advertising has a "click-through" rate 670% higher than contextual advertising[31], and so an ad placed with the help of behavioral data will pay the website much more than an ad placed without data (on average).[32] Now, suppose that 25% of the website's visitors exercise their rights under the CCPA to opt out of the data exchanges that permit behavioral ad targeting. What can the website do in response?

Suppose compliance with the CCPA antidiscrimination provision requires the website to treat everyone the same. Given the lost advertising revenue, the website may have to reduce the quantity or quality of their content, or increase the number of ads each viewer has to see. Either way, consumers who have not exercised their privacy rights will be forced to subsidize the services of those who have.

Alternatively, suppose the antidiscrimination provision allows the website to provide different services based on the different revenues they can expect to draw for the two types of consumers. In this case, the website could display more ads, untargeted as they must be, to the consumers who exercised their privacy rights. Or they could limit these consumers' access to content that had higher production costs. If this is permitted, the law would no longer force cross-subsidization, but it would also largely render the antidiscrimination provision moot.

These issues will become more complex with AI and machine learning services. If the CCPA is going to be a model for federal legislation, I recommend removing the antidiscrimination clause and allowing businesses to offer services on a take-it-or-leave-it basis.

QUESTIONS FROM SENATOR BEN SASSE

1. **Aside from situations in which compliance costs lead to higher product prices and foregone spending on research and development, in what ways is CCPA affecting Americans outside of California?**

The CCPA is likely to cause companies operating in a national market to develop their practices in a way that complies with the CCPA even if the end user is in a different state. This will be particularly true for machine learning and AI systems that need to pool as much data as possible to have the power for precision predictions. Thus, the principal risk to residents in other states is not that compliance is a costly process, but that the substance of the CCPA will constrain how innovation can develop. Please see my lengthy discussion of AI and healthcare in my response to

[31] Jun Yan et al., *How Much Can Behavioral Targeting Help Online Advertising?,* PROC. OF THE 18TH INT. CONF. ON WORLD WIDE WEB 261 (2009).

[32] The fact that the New York Times' revenues have continued to rise since ending behaviorally targeted ads in Europe may be explained by the steady increase in its readership and subscriber base since the election of President Trump. Jaclyn Peiser, *New York Times Tops 4 Million Mark in Total Subscribers,* N.Y. TIMES (November 1, 2018).

Senator Graham's question (pages 4-5) for an example of the lost innovation that residents of all states are likely to experience.

2. **Which types of sites, apps, and platforms are able to provide different user experiences between California and the rest of the country in a manner that is technologically feasible and cost-effective? Which are not?**

This is outside my area of expertise. My limited understanding is that the geo-restriction and geo-blocking that media firms routinely do in the international market is more difficult to replicate within the United States unless the service has access to GPS or other more precise geolocation data. Businesses could also require consumers to register and log in before using their services, and part of the registration could be designed to separate California residents from others by asking for the user's location. Some of these will not fit well with a business's service.

3. **What is a principled way we can think about the possibility of federal preemption in the data privacy context? When is not appropriate to let states regulate as they wish, even if we disagree with their policy choices? In what situations should we be comfortable with letting one state drive nationwide policy as a practical matter?**

A non-exhaustive list of factors should include (a) whether a state law targets an industry where firms develop policies and practices locally or nationally; (b) whether the risks and benefits of the relevant area of policy are hard to assess, and should leverage the expertise of federal agencies; (c) whether there is a substantial risk that the state law violates the U.S. Constitution; and (d) whether the state law is bad policy.

4. **To what extent has GDPR deprived European users from accessing particular types of content on the internet?**

Immediately after the passage of GDPR, about one-third of the top 100 U.S. news websites blocked traffic from the EU (including the *L.A. Times* and the *Chicago Tribune*).[33] And to date, the European Right to be Forgotten (which predated the GDPR but has now been codified by it) has led to the delisting of 1.1 million websites in Europe.[34] Although I do not have it, data from WordPress may be able to reveal the number of small companies selling goods and services on the web who have decided to block sales for EU customers using a popular plug in for EU-compliant Internet sales.[35]

European privacy law also changes the type of content that is developed in the first place and rewarded with advertising dollars. A study of the effects of the EU Data Protection Directive (the privacy framework that was in place in Europe before the GDPR took effect) found that the

[33] Steve Dent, *Major Us News ites Are Still Blocking Europeans Due to GDPR,* ENGADGET (August 9, 2018).
[34] GOOGLE TRANSPARENCY REPORT, SEARCH REMOVALS UNDER EUROPEAN PRIVACY LAW, https://transparencyreport.google.com/eu-privacy/overview
[35] WooCommerce EU VAT Compliance Assistant, WordPress.org

efficacy of advertising in Europe was reduced, and the reductions were not uniform.[36] "Websites that had general content unrelated to specific product categories (such as news and media services) experienced larger decreases in ad effectiveness after the laws passed than websites that had more specific content (such as travel or parenting websites). Customers at travel and parenting websites have already identified themselves as being in a particular target market, so it is less important for those websites to use data on previous browsing behavior to target their ads."[37] So it is difficult to know the full effects of the GDPR on the information ecosystem.

5. Which aspects of GDPR, CCPA, and proposed aspects of potential federal legislation most worry you in terms of protecting incumbents in particular markets and creating major barriers to entry for new firms?

The GDPR, the CCPA, and many of the proposals for federal legislation focus on user control rather than significant risks and harm. For reasons I stated in my written testimony last month, this shift would be unfortunate for businesses and consumers alike because compliance is more difficult for new firms that have fewer resources and less name recognition.

Assuming that a consent-based system is adopted in a federal privacy statute, firms should be permitted to decline service, or offer lower quality service, to a person who exercises their privacy rights. (The GDPR and CCPA forbid firms from creating two tiers of service unless the difference can be justified by a direct connection to the consumer's value in the transaction.)

To understand why I recommend a legal framework that permits a "take it or leave it" response, let's revisit the question that Senator Lindsey Graham asked the panel of tech company representatives at last month's hearing. Senator Graham wondered how important tracking-based behavioral advertising was to a service provider's bottom line when he visits a website with golf-related content. The panel of experts seemed to agree that contextual advertising would provide roughly the same revenue as behavioral advertising, leaving the impression that behavioral advertising could be killed off with little effect on end users. This is not so. Imagine Senator Graham were to visit two websites: one for golf enthusiasts, and one with general interest articles. With the aid of data aggregators and advertising exchanges, both websites will profit about equally from Senator Graham's visit because both could deliver advertising related to golf, to his location, or to his other interests and preferences. On the other hand, if tracking were prohibited either as a default or because Senator Graham opted out, then the prospects of the two websites are dramatically different. Throughout the economy there may be roughly the same amount spent on advertising, but the advertising dollars will flock to the golf website and away from the general interest website.

[36] Avi Goldfarb & Catherine E. Tucker, *Privacy, Regulation, and Online Advertising,* 57 MANAGEMENT SCIENCE 57 (2011).

[37] Id. *See also* Catherin E. Tucker, *The Economics of Advertising and Privacy* ("The reduction we measured is particularly pronounced for websites offering content that is not easy to match to advertising (such as news websites and media services). It is also more pronounced for non-intrusive and smaller ads, since their appeal rests primarily on the presentation of informative rather than attention-grabbing messages. This suggests that privacy regulation might change the web landscape in unanticipated ways, with advertising becoming even more intrusive, and marketers shifting their media buys away from sites such as news providers that are difficult to match with relevant advertising.")

With contextual advertising, websites will face market pressure to produce more content that allows advertisers to match themselves to end users via context. The implications may not seem so troubling when comparing golf to general interest websites, but imagine how these changes would affect news websites or content about health and medical information. The high value content that will attract advertisers are the stories that help fragment readers in a way that signals their ideology or other personal characteristics based on content. The winners in this system counterintuitively include Facebook and Google, because these companies can match ads to users without any help from third party data aggregators thanks to the long-term direct relationship with the consumer and to the highly specific searches and newsfeeds that the consumer is viewing at any given time.

6. **Which aspects of GDPR, CCPA, and proposed aspects of potential federal legislation most worry you in terms of harming innovation? How much should we worry about regulation hampering innovation in the West and giving China a competitive advantage in the development of new technology such as artificial intelligence?**

Restrictions on the retention and sale of personal data and vague standards on the requirements for deidentification are the most troubling aspects of the GDPR, CCPA, and consent-centered federal proposals when it comes to innovation. To understand why I focus on these, please see my lengthy discussion of AI and healthcare in my response to Senator Graham (pages 4-5).

7. **In terms of the different proposals for giving the Federal Trade Commission new rulemaking authority, how should we think about balancing between ensuring flexibility to adapt a regulatory framework to fit emerging technologies and avoiding delegation of what should be lawmaking authority properly exercised by Congress to a "fourth branch" of government?**

The Federal Trade Commission is in a much better position to assess the ultimate impact of a new data practice on consumers and competition. But as your question suggests, if the agency is given too much flexibility, the formal and informal rules can change rapidly, resulting in unpredictability and scope creep. For this reason, I recommend giving the FTC rulemaking authority, but statutorily tying the authority to specific definitions of redressable risk and harm. For example, FTC rulemaking authority could be restricted to four purposes: (1) to define reasonable data security, (2) to identify when a data practice produces unjustified risk of substantial injury; (3) to identify a data practice that was not effectively disclosed and for which a majority of consumers (or a majority of a particular group of consumers) would have chosen another option available to them at the time; and (4) to identify when a firm has failed to contractually restrict or reasonably monitor business partners with whom the company shares personal information.

8. **Do you foresee any situations in which data portability requirements actually enhance some firms' abilities to build more data-rich profiles of individual users?**

Yes, but this is not necessarily a problem. A data portability requirement may be a boon to existing data aggregator companies or to future startups who want to offer services directly to the consumer. For example, I could imagine a system of behaviorally targeted ads that pays consumers a small cut of its revenue in exchange for porting their Facebook and other data into a locker and permitting advertising across their devices and across the web to make use of that data. This may not be the most likely application to develop using ported data (particularly since behavioral advertising is most effective using detailed but short-term data), but it serves to illustrate that data porting could inspire existing and new firms to expand the profiles they would otherwise have on their users by inducing the users to port and store additional data with them. These developments could alleviate some of the competition concerns that loom over Facebook and Google, too.

That said, I am not sure a regulatory intervention is necessary to spur this sort of development. In theory companies could already offer consumers an incentive to allow their devices to be tracked across all of the websites and applications they use by installing a program onto their computers (a program that would look like malware if it weren't specifically sanctioned.)

9. **Do you foresee any situations in which opt-in requirements actually increase the amount and types of data that firms collect from individual users?**

It's certainly possible that opt-in requirements would result in more data collection either because pop-up windows or take-it-or-leave-it deals convince users to opt in and never look back, or because the psychological experience of having control inspires the user to trust the service provider more.[38] But these counterintuitive effects should not be overstated. Opt-in rules are likely to alter the decisions made by firms to attempt to collect data in the first place. And for most companies most of the time, only a small fraction of users will bother to opt out of data collection.[39]

10. **To what extent do you think privacy policies and user agreements are drafted deliberately to dissuade users from closely reading them?**

The FTC has brought many cases against companies under their deceptive practices authority when the companies' privacy policies have given the impression that a consumer's privacy is better protected than it actually is. Some of these cases have arguably gone too far, holding companies responsible for highly technical or improbable misreadings of their promises. (For example, the enforcement action against Snapchat penalized the company for, among other things, assuring its users that messages sent through its app will "disappear" which, while true for the app's design, was technically not true if the recipient of the message quickly took a screen shot.[40]) To steer clear of FTC enforcement, privacy policies must explain their practices in rather

[38] *See* Catherine Tucker, *Social Networks, Personalized Advertising and Privacy Controls,* 51 J. MARKETING RESEARCH 546 (2014).

[39] For example, the New York Times ran a front page story about the data practices of the startup Nomi Technologies, which tracked shopper's phones in physical retail spaces. In response to the article, Nomi's website (where shoppers could opt out) received 3,840 new visits, but only 3.8% of the visitors opted out of the data collection. Dissenting Statement of Commissioner Joshua D. Wright In the Matter of Nomi Technologies, Inc. (2015).

[40] Federal Trade Commission Complaint, In the Matter of Snapchat, Inc. (2014).

stark terms, and to bury them in a long end user agreement. The FTC's high standards for avoiding deception have impeded the Commission's goal to give users meaningful and easy-to-understand information.

That said, as long as consumers find aggressive data-collection and repurposing creepy, companies are unlikely to be fully forthcoming about their practices, and consumers will continue to be surprised and dispirited if they lack faith that regulators are poised to step in to protect them from downstream harm.