Privacy of Digital Health Information

Amalia R. Miller*

November 2022

Abstract

The widespread adoption and use of digital technologies, both within and outside of healthcare, has generated massive amounts of data on individual health that is controlled by companies. The promise of these data for improving human welfare is immense but unfettered corporate access to health information imperils personal privacy. This conflict raises fundamental questions about whether and how societies can harness the value of healthcare digitization, including big data applications and artificial intelligence, while still preserving privacy. This chapter surveys the economic and policy landscape on health data privacy in the rapidly evolving digital age. It first identifies the potential tradeoffs inherent in health privacy regulation, by examining the range of individual and social costs and benefits associated with increasing or decreasing the flow of individual health information. It then presents theoretical economic arguments for different regulatory approaches to protecting health privacy and draws insights from the empirical economics literature on the impact of health privacy rules.

* University of Virginia, IZA, and NBER. Email: armiller@virginia.edu. This chapter is based on a presentation at the Fall 2022 NBER Economics of Privacy Conference. I am grateful to the conference organizers, Avi Goldfarb and Catherine Tucker, and to other faculty and graduate student participants at the conference, for stimulating comments and feedback.

1. Introduction

Health information merits special attention within the economics of privacy because the stakes of its protection are especially high. Some of the most sensitive and revealing facts about a person pertain to their physical and mental health. Having those facts disclosed publicly can cause a person to experience both direct discomfort and indirect harms through various ways in which other people respond to the information. If patients are unable to trust medical providers to keep their information private, they may be unwilling to undergo testing or seek medical treatment or they may withhold key information about symptoms and risk factors.

Health information is also important for privacy scholars because of its special policy treatment. The United States lacks any national law that protects privacy for all types of personal data, yet federal laws addressing the privacy and security of health information have been in place for years. The most prominent of these is the 1996 Health Insurance Portability and Accountability Act (HIPAA) that produced the 2003 Privacy and Security Rules (45 CFR § 160 and 164). Further data security provisions were added in the 2009 Health IT for Economic and Clinical Health (HITECH) Act and protections for genetic information were adopted in the 2008 Genetic Information Non-discrimination Act (GINA).[1] National laws protect health privacy around the world (OECD 2022), and broad-based privacy rules typically categorize health information as particularly sensitive and require stricter protections.[2]

Health privacy policy has become increasingly important and complex as advances in computing have spurred the collection, storage, and analysis of massive amounts of personal health data. Digitization of health information makes that information easier to share and harder to protect, which increases the individual risks to health privacy. At the same, widespread digitization of health information has unique potential to increase human welfare, through improvements in healthcare delivery quality and efficiency and through data-driven innovation in medical devices and personalized medicine that can better target treatments that extend and improve lives. This dual nature of health information digitization therefore presents policymakers with a significant challenge in devising health privacy rules in a way that balances the costs and benefits of amassing and exploiting digital health data.

Economic approaches, both theoretical and empirical, can be particularly valuable for assessing these tradeoffs and for evaluating the effects of different approaches to health privacy policy. This chapter therefore offers a conceptual framework for the economics of health information

[1] Information about substance use disorders and treatments at federally funded programs are under stricter privacy protections (42 CFR § 2.11 Part 2). The US also has targeted privacy rules outside of health. The Financial Privacy Rule, created as part of the Financial Modernization Act of 1999 (the Gramm-Leach-Bliley Act, or GLBA), can also cover some health information, and the 1970 Fair Credit Reporting Act (FCRA) addresses privacy and accuracy in credit reports. Federal privacy rules also cover children (the Children's Online Privacy Protection Act, COPPA) and educational data (the Family Educational Rights and Privacy Act, FERPA). Use and dissemination of personal information by federal government agencies is regulated under the 1974 Privacy Act.

[2] These laws include the European 2016 General Data Protection Regulation (GDPR), discussed in Chapter 4 of this volume, and the 2019 Brazilian General Data Protection Law (LGPD).

privacy, surveying the existing literature, and highlighting open areas of inquiry. Section 2 delineates the various forms of harm that individuals might experience from having their health information revealed against their wishes and categorizes those harms into types. In principle, the potential harms from improper disclosure can be weighed against the benefits from of allowing unrestricted use of digital health data, discussed in Section 3, to determine the socially optimal level of privacy protection. In practice, uncertainty about, and heterogeneity in, both costs and benefits of health privacy make it impossible to find a single universally optimal level of protection. Section 4 considers economic justifications for various government interventions in health privacy, based on efficiency and fairness grounds, and links them to specific policy provisions and rules. Variation in these policies across places and time has provided valuable opportunities for empirical researchers to measure the effects of health privacy laws as implemented. Section 5 reviews the empirical economics literature on the effects of existing health privacy laws, aiming to draw insights to inform health privacy policy and shed light on privacy issues in other sectors with less history of empirical variation.

2. Costs of Health Privacy Loss

A natural starting point for assessing the economic value of protecting individual health privacy is measuring the potential harm that a person can suffer from having their personal information disclosed against their wishes. Measuring that potential harm, however, is complicated by the variety of specific harms that are commonly raised in health privacy research and advocacy (e.g., IOM 2009, Gostin 1994), listed in Table 1. We will consider these harms in turn, distinguishing first between elements of the list that reflect direct, or primary, harms that happen from the disclosure itself, regardless of whether or how the data are eventually used (items 1-3, discussed in Section 2.1), and the indirect, or secondary, harms that result from how other people react to or use the data (items 4-11, discussed in Section 2.2). Although much has been written about these harms, most of the writing has been either theoretical or anecdotal, so relatively little is known about their magnitudes or prevalence.

2.1 Direct Harm from Health Privacy Loss

People may experience direct harms from violations of their health data privacy because knowledge of the data disclosure can induce feelings of shame or embarrassment about their information being revealed to others (item 1). The extent of this harm will depend on the nature of the information, the recipients of the information, and each person's attitude toward that information.

The second direct harm is from feeling a trust has been violated by the person or organization that revealed the information. This harm is also subjective, and likely to be particularly important in the healthcare sector, where preserving the confidentiality of patient information is a longstanding professional norm.[3] Violations of this norm can erode trust in particular providers

[3] For physicians, the text of the revised Hippocratic Oath includes a promise to "respect the privacy of my patients, for their problems are not disclosed to me that the world may know."

Table 1. Potential Individual Harms from Health Privacy Loss

Direct harms	
1	Feelings of shame, embarrassment
2	Feelings of betrayal, trust violation
3	Feelings of invasion, surveillance, loss of freedom, autonomy
Indirect market harms	
4	Labor market harms (e.g., hiring, salary, promotion, termination)
5	Insurance (e.g., health, disability, life, long-term care) market harms
6	Harms in other product markets, e.g., personalized pricing, if higher
7	Targeted advertising – if manipulative, annoying, intrusive
Indirect non-market harms	
8	Social stigma, isolation
9	Harms to reputation, personal and family relationships
10	Increased risk of identity theft, other theft, impersonation, fraud
11	Increased civil or criminal legal exposure

and in the healthcare system more generally (Mechanic 1998), which can reduce healthcare seeking and treatment.[4] The salience of privacy in the healthcare provision relationship is reflected in much of the literature on health privacy discussing "patients" (rather than "consumers") and in greater legal restrictions on the use and transfer of health information by healthcare providers and health insurance plans (the focal "covered entities" in HIPAA and state health privacy laws). It is possible to maintain trust while also disclosing some private information, for example when patients are informed in advance about how their data will be used and when they give affirmative consent to those uses. However, if disclosure and consent procedures are perfunctory as a precondition for service, and offer no option to withhold consent, they can themselves be damaging to trust, particularly if the data uses extend beyond the direct functions of medical care provision.

People may also value the ability to keep certain health information private because it enhances their sense of freedom. Having one's information widely available can feel invasive, coercive, or controlling, even when there are no explicit penalties or consequences (item 3). This applies to surveillance either by private companies or by government agencies, and extensive information flows between the two sectors makes those concerns impossible to fully disentangle.[5]

These first 3 items are grouped together as direct harms because they happen within the individual and are not tied to specific responses to their private information from others. This is true despite privacy being inherently a relationship concept, about setting boundaries with other people (Nissenbaum 2004). Because they depend on circumstances and context, including the relationships among people giving and receiving data, the direct harms are unlikely to be

[4] Alsan and Wanamaker (2018) illustrates the severe and lasting consequences of trust violations related to failures of healthcare researchers to provide full disclosure and obtain informed consent.

[5] See, e.g., Qian et al. (2022), for discussion of the state's expansive use of digital surveillance tools in China.

universal in any meaningful sense. As a result of this heterogeneity and instability, the interiority and subjectivity of the direct harms makes them particularly challenging to quantify or convert to money value.

2.2 Indirect Harm from Health Data Use

The economics literature on privacy has typically focused on the indirect or secondary harms from information flows (e.g., Acquisti et al. 2016), which can be mediated by market forces (items 4-7) or not (items 8-11). Within the health privacy sphere, the areas of greatest policy attention are job (item 4) and insurance (item 5) markets, because those are markets in which health information can be especially damaging to individuals, and where individuals report feeling the greatest level of concern about potential disclosure (e.g., Forrester Research 2005).

Disclosure of personal information about physical or behavioral health conditions can make a worker less attractive to employers, which can have negative labor market effects in areas of hiring, compensation, promotion, and termination (item 4). The use of health information as a basis for differential treatment in labor markets can be considered a form of discrimination and considered using economic models developed to study discrimination by race or gender. For example, employers may use health information for statistical discrimination (Phelps 1972), because of its value in predicting worker productivity or labor supply. This use of health information could be profitable for employers, but it might also be inefficient if employers overreact to health information because they are less able to assess future productivity of workers with those conditions (Aigner and Cain 1977) or if they have biased beliefs about, and limited experience with, such workers.

Even when health information is irrelevant for productivity, it is also possible that employers have preferences related to worker health and use health information to engage in taste-based discrimination at the expense of profit-maximization (Arrow 1973, Becker 1957). This can be because employers (or customers or fellow employees) have preferences against hiring or working with people with certain health conditions, or because the health conditions are informative proxies about other non-productive characteristics, such as sexual orientation, over which they have such preferences (e.g., Badgett 2007). Discrimination by health status is also closely related to discrimination by disability status (see, e.g., Baldwin and Johnson 2006 for a survey), but not the same, for example, because not all sensitive health information is related to a current disability. Regardless of the underlying motivations, workers who anticipate adverse employer reactions to their health information will prefer to keep that information private.

Markets for insurance – health, life, disability, long-term care (item 5) – are another major setting in which health information can be relevant to firm profits and firms might penalize individuals with certain health conditions or risks. In these markets the assumption is typically that health information is used solely for its predictive value, rather than for animus-based discrimination. This in no way diminishes the harm experienced by people with medical information that implies higher expected insurance claims when they are charged higher premiums or restricted in their insurance offerings. Although not directly relevant to the individual harm, the profit-motive for

using health information in insurance markets, as with labor markets, nevertheless presents challenges for health privacy regulation in these markets, for both conceptual and practical reasons. The theoretical concern is that some privacy rules might reduce overall welfare and the practical concern is that firms will be more motivated to circumvent the rules. These will be discussed further in Section 4 on regulation.

Health insurance and employment are also tightly connected in the US, where 57% of the non-elderly population is covered by an employment-based plan (KFF 2022). This means that employers are often concerned about expected medical claims. This is particularly the case for self-funded plans,[6] where employers are financially responsible for claims.[7] It can also apply to employers who buy insurance in group markets and are exposed to some degree of "experience rating" where premiums can increase based on past claims,[8] which increases the cost to the employer of providing health benefits. An implication of this connection is examined empirically, for example, in Gruber (1994), where mandated health insurance coverage of maternity affected employment outcomes for women. The close connection between employment and health insurance also entails extensive information sharing, which further connects the privacy concerns across the two settings.

Unlike the direct harms in the prior subsection, these direct harms can have significant financial impacts, which makes them potentially easier to quantify. This is certainly true for an individual who is denied a job or charged a higher price for insurance because of a specific piece of health information. However, attribution can also be challenging in assessing health privacy harms. For health conditions that affect productivity or medical costs, it is often difficult to disentangle the impact of the health information itself, separately from the observable consequences of that information. More generally, it is often difficult to identify the incremental impact of any specific piece or set of health information on labor or insurance market outcomes, relative to what could have been inferred from the rest of the available data. Going beyond individuals, even less is known about the aggregate importance of these harms at a population level, or even within specific subpopulations based on medical diagnoses.

What is known, from surveys and focus groups, is that individuals frequently cite privacy concerns about information disclosure to their employers or insurers as paramount, because of heightened fear of discrimination in those markets (IOM 2009).[9] There is also evidence that these concerns are reflected in behavior. The potential for negative predictive health information to be used against individuals in future market transactions lowers people's willingness to seek out actionable health information, such as HIV status (Vermund and Wilson 2002) and genetic testing

[6] Self-funded health plans are more commonly offered at larger employers and account for a growing majority of enrollees in employment-based health plans (Miller et al. 2013, Claxton et al. 2022).

[7] Greenhouse and Barbaro (2005) report on an internal memo at Walmart recommending hiring fewer unhealthy workers as means of reducing healthcare spending.

[8] The practice of experience rating has been largely proscribed in the individual and small group markets under the ACA.

[9] For substance use disorders, housing markets are also a key are of concern for discrimination, addressed in research and policy.

(Gostin 1991, Hellman 2003, Oster 2013). Some individuals report engaging in overt efforts to acquire the relevant information in a way that is shielded from their insurers or employers, for example by paying privately for testing (Oster et al. 2008, Miller and Tucker 2018) or testing outside of clinical settings (Figueroa et al. 2015).

Outside of health, most economics research on privacy has focused on other product markets (item 6) and the use of information about an individual's willingness to pay for a product for price discrimination or for targeted advertising or product recommendations (item 7). For examples outside of health, see Acquisti et al. (2016), Ichihashi (2020), Acemoglu et al. (2022) and references therein. Health information can plausibly be used for these purposes as well, either for marketing health services or for health-related goods, though it is unclear that health information would be especially useful. The harm in this case is also less obvious. Targeted advertising (and personalized product matching and recommendations) that is based on health information is indeed harmful if it is annoying or manipulative or if it causes further disclosure of health information to third parties. An example of manipulation is implied, for example, in the claim in Duhigg (2012) that retailers target advertisements to new parents because they are "exhausted and overwhelmed" and therefore open to trying new brands. But there can also be a positive side to personalization, even when based on health information, if it improves match quality and helps consumers find products and services most valuable to them. As with labor and insurance markets, empirical researchers face significant challenges in attempting to link any specific release or inference of health information to outcomes in these other markets, which is further complicated by the availability of similar information from other sources. Empirical work in this area could be exceptionally valuable.

Non-market factors include social stigma or isolation (item 8) and damage to personal and family relationships (item 9). Mental illness, substance abuse, and HIV status are concrete examples of health information that have been shown to disrupt family relationships, but it is also possible that relationships could be damaged by disclosure of other acute or chronic medical conditions. While responses to these disclosures can have a significant impact on individual wellbeing, they are impossible to regulate directly, other than by preventing the flow of information. For market transactions, policymakers have the added ability to regulate the use of personal information, which is an important feature of health privacy rules that generates overlap with anti-discrimination and civil rights laws (see Section for further discussion).

The two final categories of potential harm relate to potentially illegal behavior. The first category (item 10) is that disclosure of personal medical information, primarily from data breaches or involuntary loss, can increase a person's likelihood of being a victim of identity theft (medical or otherwise). This highlights the importance of addressing data security concerns, and preventing even unintentional disclosures, in maintaining health privacy. The other category (item 11) is that health information could potentially contribute to a trail of evidence used in a legal (civil or criminal) investigation or proceeding. This concern may arise because the health information provides evidence of wrongdoing (e.g., illegal drug use, child abuse or neglect, violent crimes) or because the medical treatment is itself illegal (e.g., reproductive healthcare such as abortion that

violate state level restrictions).[10] As discussed in Section 4 below, these uses of health information are typically exempted from health privacy protections, under varying conditions.[11]

2.3 Quantifying the Costs of Health Privacy Loss

Although it is relatively straightforward to list the various potential harms to individuals from lost health privacy, measuring the value of those harms presents substantial challenges. Part of the injury is subjective, and the objective parts can be hard to detect. Both subjective and objective harms are also likely to vary significantly across people and data types, and over time, and to depend on the nature and context of the disclosure. This makes it difficult to value the harm at the individual level for any specific disclosure or at the population level from overall reductions in protection.

Perhaps unsurprisingly, relatively little is known about the empirical distribution of individual or aggregate harms from health privacy loss. Major reports on health privacy rarely cite values for these harms, in total or for specific elements, and instead focus on consumer attitudes (e.g., IOM 2009, HHS 2017). For attitudes, public opinion polls typically find high fractions of respondents who report feeling concerned about their health privacy (e.g., majorities in Forrester Research 1999, 2005, California Healthcare Foundation 2005), but not universally (the 2014 Truven Health Poll cited in HHS 2017 had rates under 20%). When asked to consider hypothetical choices to protect their online information across data types, subjects in Skatova et al. (2019) consistently reported placing the highest value on protecting the privacy of medical and financial records.

Outside of health privacy, researchers have attempted to go beyond stated preferences to examine situations in which subjects make consequential choices about information sharing to infer privacy preferences. These results illustrate the difficulty of converting information on stated preferences into economic measures of value. The field experiment in Athey et al. (2017) illustrates an example of the "privacy paradox" in which individuals express strong privacy preferences yet disclose personal information for small rewards. Lin (2019) infers privacy preferences using non-response rates to personal questions in a survey, similar to the analysis of observational survey data in Goldfarb and Tucker (2012). What is unusual in Lin (2019) is that the experimental treatments are designed to separately measure two components of privacy tastes – the intrinsic value (roughly corresponding to preventing the direct harms in Section 2.1 of this chapter) and the instrumental value (indirect harms in Section 2.2). Although the mean intrinsic value is low in the sample, the paper finds substantial variation across participants. Other experiments find further evidence of heterogeneity in privacy choices, even within individuals,

[10] The latter category could take on heightened importance in the wake of the 2022 U.S. Supreme Court decision in Dobbs v. Jackson Women's Health Organization overturning prior limits on states' abilities to ban or regulate abortions. See, for example, recent news coverage in Hill (2022), Nix and Dwoskin (2022), and Kelly et al. (2022).

[11] For example, HIPPA-covered entities can provide health information in response to a court order or, after meeting notification requirements, in response to a subpoena. Data sharing with law enforcement is more strictly limited under rules for Confidentiality of Alcohol and Drug Abuse Patient Records (42 CFR § 2).

where choices vary with contextual factors and framing (e.g., Adjerid et al. 2019, Athey et al. 2017, Acquisti et al. 2015).[12]

Aggregate information on health privacy loss from data breaches is available because of mandatory reporting, but official statistics cover only the volume and type of data, and not the costs or consequences to individuals of the breach.[13] Survey responses in Ponemon (2013) indicate that medical identity theft is increasingly common in the US (affecting an estimated 0.8% of adults), with victims incurring an average out of pocket cost of $6,718 and experiencing other adverse consequences such as lost health insurance, time and effort devoted to resolving or correcting the issue, and lower trust in healthcare providers. The Ponemon survey also reveals another important feature of medical identity theft, which is that security breaches at healthcare providers and insurers are not, in fact, the primary sources of information. Instead, a significant majority of victims attribute the crime to their having knowingly shared their information (30%) or to a family member accessing their medical credentials without their consent (28%).

Despite the conceptual and practical challenges, there is significant value from empirical measures of the distribution of actual and perceived costs that individuals face from different aspects of health privacy risk. This is because (as discussed in Section 3) privacy protections are not costless. Optimal privacy policy should therefore ideally focus on preventing the most serious potential harms and addressing the areas of most widespread concern.

3. Digitization and Costs of Health Privacy Protection

Although privacy risks are present with any form of medical record keeping, they are substantially higher for digital records than for paper files. Electronic records are much cheaper and easier to store, access, and transfer. This greater portability of digital records threatens data confidentiality, by making intentional disclosures less expensive, as well as data security, by potentially enabling massive data breaches carried out by distant attackers. Electronic health information is also easier to combine with other data sources, to manipulate, and to analyze, which increases the risks of indirect harms from how the information is used after disclosure. It is not surprising, therefore, that the increased impetus for health privacy in the late 1990s was closely tied to the diffusion of electronic medical records and health information exchange, particularly among medical providers and payors, and that the HIPAA Privacy and Security Rules focused on entities that transfer information in electronic form.

If protecting privacy rights entails allowing individuals to decide for themselves what information to conceal from others (Posner 1981), then stronger protections of health privacy will require restrictions on the volume, flows, and usage of digital health data, and will reduce the amount of

[12] This heterogeneity in privacy concerns is also found in focus group discussions. For example, attitudes about privacy and security in mobile health applications are highly variable across people, information types, and context (Atienza et al. 2015). Goldfarb and Tucker (2012) also find significant heterogeneity in privacy-preserving behavior across demographic groups and over time.

[13] The Department of Health and Human Services (HHS) maintains a public listing at https://www.hhs.gov/hipaa/for-professionals/breach-notification/breach-reporting/index.html

health data available to companies. Data elements and uses that are not directly and sufficiently beneficial to consumers will become more expensive to access or will no longer be available. For that reason, the primary cost of strong privacy protection comes from reducing the gains that would otherwise be generated by greater use of digital health data.

What are the benefits of digital health information? Advances in information technology and computing have significantly lowered costs of data collection and use and driven a shift to digital record keeping and processing across the economy (Goldfarb and Tucker 2019). As an industry, healthcare has been slow to transition away from paper records, despite arguments that electronic medical records (EMRs) have potential to both improve healthcare quality, by reducing errors due to inaccurate or incomplete information about patients (IOM 2000), and to lower administrative costs (Hillestad et al. 2005). Several reasons have been posited for the slow diffusion of EMRs, including privacy concerns, as well as positive externalities from EMR adoption from quality improvements and information sharing across organizations (Miller and Tucker 2009). The 2009 HITECH Act allocated over $25 billion of federal government funding to provider incentives for health IT adoption. At the time of its passage, while only 2 percent of US hospitals had an EMR system in place that met the government's "meaningful use" criteria (Jha et al. 2010).

A substantial literature examines the impact of adoption of digital health records by medical providers in the periods before and after the HITECH Act. Studies have found significant improvements in quality, particularly for the most vulnerable patients and complex cases (Gresenz et al. 2017, Miller and Tucker 2011a, Derksen et al. 2022, McCullough et al. 2016, Freedman et al. 2015), though the gains have not been universal across providers or patient groups (e.g., Spetz et al. 2014, Agha 2014, Hitt and Tambe 2016).[14] Lin et al. (2019) sheds some light on the heterogeneous effects of EMRs across hospitals. The study finds no effects of technology alone, but significant quality improvements from achieving "meaningful use" criteria of the HITECH Act. Lin et al. (2019) also finds larger quality improvements at small and rural hospitals, suggesting an important role for health IT in reducing health disparities. The estimated effects of EMRs on hospital operating costs are also heterogeneous across hospitals, with the benefits from adoption favoring hospitals located in areas with a stronger labor market presence of IT workers (Dranove et al. (2014).

In addition to the stand-alone benefits that accrue to EMR adopters and their patients, there can also be benefits from participating in health data exchanges with other providers (Walker et al. 2005). Indeed the "meaningful use" criteria for system interoperability were aimed at promoting network benefits from data exchange. These benefits cross the boundaries of individual firms, but still accrue to the patients whose data are shared. Empirical studies have found evidence supporting these spillover gains from health information exchange in both quality and costs. Janakiraman et al. (2022) finds quality improvements at emergency departments in the form of shorter inpatient stays and lower patient readmission rates, while Lammers et al (2014) finds a

[14] Also see Atasoy et al. (2019) and Bronsoler et al. (2022) for reviews of the literature on the effects of health IT on clinical quality, productivity, and healthcare utilization.

reduction in duplicate testing for patients who visit multiple hospitals. Despite these gains, data exchange can be particularly hampered by privacy concerns (McGraw et al. 2009).

While information exchange among medical providers can improve healthcare operations, the social gains are from health data use extend beyond the healthcare system and the data subjects themselves. These uses are well-illustrated in the 12 "national priority purposes" for which the HIPAA Privacy Rule permits disclosure and use of personal health information without express permission.[15] These purposes include compliance with legal requirements and regulations, operation of government insurance programs, regulatory oversight and enforcement, and for law enforcement and crime prevention purposes. Of these, public health activities and research uses are likely to have the most significant economic impacts.

The need for timely and extensive data on health conditions for effective public health operations stands in conflict with absolute individual rights to privacy. Health privacy regulations, including HIPAA and state laws, typically relax disclosure rules for public health uses such as disease surveillance and contact tracing programs to monitor and contain outbreaks of infectious diseases. These uses, and the vital importance of free-flowing health data, were especially salient in the government response to the COVID-19 pandemic (e.g., Halpern 2020, Buckman et al. forthcoming).[16] Even outside of epidemic control, privacy protections often need to be relaxed to promote public health, such as registries for non-communicable disease. Other examples including limiting patients' control over their prescription information to curb opioid abuse through drug monitoring programs (Maclean et al. 2020) and limiting parental control over child health information to protect abused and neglected children.

Perhaps the largest social benefit from digital health data comes from its use as an input into health research and development. Digital health data, whether generated from clinical encounters and insurance claims, from sources outside of the healthcare system, or by merging existing public and private sources, can advance health research by significantly lowering the costs of conducting large-scale studies that study novel treatments and measures. Massive datasets with health information can be especially valuable for research into rare conditions and for assessing heterogeneous effects across detailed sub-groups, making it a key input into the development and deployment of personalized medicine, where disease prevention, diagnosis and treatment are all tailored to the patient's individual genetic, social and environmental characteristics (Miller and Tucker 2017, 2018).[17]

[15] See, e.g., https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html, which describes the aim as "striking the balance between the individual privacy interest and the public interest need for this information."

[16] Lawmakers reacted to the threat to health privacy from these expanded public health data uses by introducing a specific COVID-19 Public Health Emergency Privacy Act. The bill was introduced in the senate in January 2021 (see text at https://www.congress.gov/bill/117th-congress/senate-bill/81) and in the house in February 2021 (text at https://www.congress.gov/bill/117th-congress/house-bill/651/text).

[17] A prominent public investment in this area is the Precision Medicine Initiative at the US National Institutes of Health (NIH) and the resulting *All of Us* research program (All of Us Research Program Investigators 2019).

Large quantities of health data are also needed to exploit novel information processing tools, such as machine learning and artificial intelligence, for healthcare applications (Price and Cohen 2019, Smalley 2017, Sanders et al. 2019, Yu et al. 2018, Shilo et al. 2020, Goldfarb et al. 2020, Bates and Syrowatka 2022). Better use of health information can also lead to process improvements within healthcare delivery systems, through internal quality improvement program evaluations and utilization reviews.[18] One conception of this is the idea of the "learning health system" (IOM 2007, 2013, Friedman et al. 2015), whose activities can fall outside of formal research, but nevertheless contribute to improving performance and health outcomes.

Although the precise economic value of the resulting medical innovation is hard to measure, and even harder to predict for future discoveries, its potential is enormous, because of the immense economic value of extending lives and improving health (e.g., Murphy and Topel 2003, 2006). HIPAA's Privacy Rule includes provisions aimed at reducing the barriers to using previously collected health data for research purposes. These include the possibility to waive consent with approval from an Institutional Review Board (IRB) or Privacy Board,[19] to disclose a limited data set with an data use agreement, or to exclude the health information from protected status by rendering it anonymous and stripped of its personal identifiers (or de-identified).[20] Despite these allowances, evidence suggests that even the limited burdens imposed by the HIPAA Privacy Rule have been detrimental to health research (IOM 2009).[21]

An increasing volume of health data is now being generated outside of clinical settings from a growing number of mobile health devices and applications used by consumers directly to manage their physical and mental health conditions or to invest in general wellness and disease prevention (e.g., HHS 2017, EDPS 2015). Some of these applications are used to exchange data with healthcare providers, or under their supervision, which could help expand access to

[18] It is also important to note concerns that big data and algorithms can potentially exacerbate existing inequalities, by race or other protected groups, including in healthcare (HHS 2017). Alternatively, increased use of computer algorithms and decision support in healthcare could improve outcomes more for traditionally disadvantaged groups, by reducing the impact of human biases (as found in Bartlett et al. 2022 in financial services). As noted elsewhere in this chapter, digitization in healthcare, by increasing standardization and reducing error rates, has been found to produce greater gains for disadvantaged populations, such as larger improvements in survival rates of Black infants in Miller and Tucker (2011a), lower amputation rates for Black patients in Ganju et al. (2020), and improved quality measures at smaller and rural hospitals in Lin et al. (2019). Further empirical evidence is needed to understand whether and how more advanced data applications affect health disparities.

[19] The regulations governing IRB's and ethical guidelines for protecting the privacy of research subjects are addressed in the Federal Policy for the Protection of Human Subjects, known as the Common Rule, adopted by 20 federal agencies and departments. See https://www.hhs.gov/ohrp/compliance-and-reporting/common-rule-agencies-contacts/index.html

[20] This de-identification can be accomplished by an expert or through the Safe Harbor method of removing 18 types of personal identifiers. Heightened privacy concerns around the use of government data, and greater awareness of re-identification risk from previously anonymized data (Komarova et al. 2018), have led to renewed debate about the adequacy of HIPAA's provisions for health research (e.g., IOM 2009). Outside of health, revised disclosure methods to improve privacy protection have been implemented in producing Census data products for the public, possibly at the expense of statistical accuracy (Abowd and Schmutte 2019, Hotz et al. 2022).

[21] Outside of health, Miller et al. (2021) discusses costs (and advantages) of using informed consent to collect individual internet browsing histories for research.

healthcare for people in rural and under-served areas, while others are used with no connection to formal healthcare providers. Although these applications can significantly increase the amount of health data at risk of disclosure, even consumers with strong preferences for health data privacy and security express a willingness to sacrifice on those dimensions to benefit from the convenience and quality improvements (Atienza et al. 2015).

The re-use of personal health information for marketing purposes is more controversial. While it is true that health data used for personalized advertising, pricing, or product recommendations can be unfavorable to consumers (as discussed in Section 2.2), that is not universally the case. Consumers can also benefit from improved match quality in seeing more relevant advertisements and learning about products they are more likely to want to purchase. Using health data as an input to better predictions of product matches is similar in spirit to personalized medicine, though in a different context, and can similarly have a public good component from more data sharing from one person improving the quality of matches for others (e.g., Loertscher and Marx 2020). However, unlike medical applications that improve health, better predictions in other markets can sometimes be used in ways that benefit firms at the expense of some consumers or to infer hidden health information. For those uses, the spillovers across people from increased data sharing, and improved predictions, would be negative. Whether consumers benefit or suffer harm from personalized marketing depends on the nature of the marketing they receive and on their subjective preferences about the underlying health information and the promoted products.

Consumers can also benefit indirectly from the value their data provides to advertisers if a robust market for consumer health data supplies revenue to developers to create new digital health-related products and offer them to consumers for free or at low cost. In that case, consumers are effectively compensated for their data through useful digital products (e.g., Kummer and Schulte 2019). A recent analysis of over 15,000 free mobile health (medical or health and fitness) apps available through Google Play found that 88% included programming code that could collect user data, with 88% of data collection operations sending information to external third parties for tracking, analytics or advertising (Tangari et al. 2021).[22] While it is unclear if these benefits are particularly high for sensitive medical information, the increasing depth and scope of personal information used in digital advertising means that health-related information is increasingly being amassed from sources unrelated to healthcare, including internet browsing and mobile device location services.[23] It can also be increasingly possible to infer health information from non-health data, as was done in Merchant et al. (2019), using textual analysis of social media

[22] An earlier study of 211 diabetes apps on Google Play found that fewer than 1 in 5 had a privacy policy and only 4 said they would obtain user permission for data sharing (Blenner et al. 2016). A smaller study of apps for smoking cessation and depression found that 80% transmitted data to Facebook or Google services, but fewer than half of those notified users in their privacy policies (Huckvale et al. 2019).

[23] See, for example the August 2022 FTC case against Kochava, in which the data broker was alleged to have sold geolocation information that could be used to trace individual visits to "sensitive" locations, including abortion clinics. Also note the "Socioeconomic Health Risk" product offered by LexisNexis Risk Solution, described at https://www.lexisnexis.com/risk/downloads/literature/health-care/Socioeconomic-Health-Risk-Score-br.pdf as a way to "predict health risk more precisely - without medical claims data."

posts to predict health conditions and treatments found in EMR records. Because health information is so tightly enmeshed with other types of information, strict privacy rules that cover any health information, regardless of source or confidence level, could operate effectively as broad-based restrictions and raise the costs of collecting and using of all forms of personal data.

4. Economic Foundations for Health Privacy Regulations

As the previous two sections illustrate, decisions about when to keep health information private and when to disclose it, produce a wide variety of private and social effects. This variety is reflected in the various concepts of privacy employed in the theoretical literature in economics on privacy (discussed in Chapter 2 of this volume) as well as in the various approaches to privacy protection employed by policymakers. This section connects theory to policy by presenting economic foundations for different types of privacy rules.

Decisions about disclosing health information are made by consumers, who are the subjects of the information, and by firms, who control and manage it. Although there can be significant overlap in their interests, the alignment is imperfect. Individuals typically bear the main costs from improper disclosure and misuse of their information, making them prefer a higher level of privacy and security than firms do. A natural starting point for resolving disagreements between firms and individuals about information privacy is through private contracts and bargaining, as in Coase (1960). In the ideal case, when the Coase Theorem holds, bargaining delivers the efficient level of privacy, balancing the costs and benefits between the parties, regardless of the initial allocation of property rights. In that case, government intervention would be unnecessary (if it leaves the final allocation unchanged) or harmful (if it were binding). Unfortunately, it is unlikely that the Coase Theorem will hold.

Bargaining over health privacy is impossible if the initial allocation of ownership rights is ambiguous. Companies that create, collect, and maintain health data have plausible claims to property rights, as do individuals who are its subjects. Without specific information to the contrary, people might assume a higher level of protection than is being offered. One function of privacy laws is therefore resolving the ambiguity. This entails setting a default initial allocation of ownership rights, either to consumers or to firms. It can also entail establishing notification and consent rules to ensure that consumers are explicitly informed about relevant privacy terms before they make decisions about sharing their health information or using products or services that will generate a trail of personal health data. An even weaker form of this policy is the Federal Trade Commission (FTC) guidance on best practices for mobile health app developers that recommends clear communication with consumers about privacy data collection and privacy policy but is otherwise focused on data security.[24]

While providing detailed information to consumers about the different potential uses of their data is essential, it may not be enough to enable them to make optimal decisions about their own desired levels of privacy. This is because privacy choices and risks can be complex and

[24] https://www.ftc.gov/business-guidance/resources/mobile-health-app-developers-ftc-best-practices

uncertain, involving hypothetical outcomes that are hard for people to assess. Furthermore, health data can be persistently informative over a lifetime and the disclosure risks can increase as science and technology advance (see Miller and Tucker 2018 on the evolving privacy risk from genetic data). It seems inevitable that some people would struggle to understand and optimize over these risks. Outside of healthcare, researchers find empirical support for concerns about non-standard decision-making in privacy choices in their sensitivity to framing and provision of extraneous reassuring information (e.g., Athey et al. 2017, Adjerid et al. 2019). This suggests a potential role for some paternalistic government interventions to protect individuals from mistakenly giving up too much privacy at the time of initial information disclosure. This could be accomplished, for example, by nudging people toward more privacy-preserving choices through default options (Bhargava and Loewenstein 2015), or possibly by mandating or supplying education to citizens about privacy issues and access to counseling and decision support services.

Even when consumers are well-informed and fully comprehend their privacy choices, consent requirements at the initial contracting stage may still not be sufficient to ensure efficient privacy levels if the range of privacy options offered to them are severely restricted. This can happen when consumers need to accept a take-it-or-leave-it lists of allowed disclosures as a precondition to obtaining medical treatment or using data-intensive products. When it comes to online activity, Acquisti et al. (2020) argue that consumers face prohibitive costs to constraining digital tracking. To the extent that the limited set of privacy options comes from coordination among competitors or the exercise of market power, it could be addressed under antitrust policy. However, this is unlikely on its own to ensure privacy-preserving options for consumers who value them. Simply assigning property rights to consumers could also be ineffectual if they agree to transfer the rights as part of the terms of trade. Instead, if there are categories of re-disclosures or uses that are harmful to a significant majority of consumers, there could be a role for government in limiting the allowed uses, i.e., preventing certain trades from happening, even when both sides are willing. This could increase overall welfare even if there is a loss from the forgone trade. Health privacy rules that require explicit patient authorization for every re-disclosure of their data have that form, as do rules that set time limits on authorizations, in that they preclude agreements that provide blanket or perpetual authorization. These provisions are notably absent from the HIPAA Privacy Rule, for example, which has been criticized for its leniency because it only requires notification of privacy practices, but not that patients are given specific options to limit disclosures (Rothstein 2007).

Asymmetric information about what happens to data after the initial disclosure presents another challenge to private bargaining. Once a consumer discloses information to a company, they are unable to monitor how the company handles the data in their possession, what information they re-disclose to third parties, or how the information is further transferred and used after it moves beyond the boundaries of the initial company that directly interacted with them. The data may be sold or transferred for some in-kind benefit or as part of a merger or acquisition of the original company. Violations of contractual restrictions on data use could easily go undetected, which limits consumers' ability to obtain recourse through legal or reputational channels. Here, again, the government has an important role in providing the necessary structure to enforce property

rights over privacy, which includes both detecting and penalizing violations.[25] This motivates rules that increase consumer information about harmful data loss (through breach notification requirements), as well as standards and mandates for information security requirements that apply to companies that collect or use health data as well as ones that develop and manufacture devices and products that do the same.

In areas of both privacy and security, there could also be role for government in imposing tighter restrictions than the market would provide to address negative spillovers from data sharing across people. These can arise among close family members, as in the cases of genetic information in Miller and Tucker (2018) and children's HIV status in Derksen et al. (2022), but they are not limited to those cases. In the model of data markets studied in Acemoglu et al. (2022), data spillovers come from information about one person revealing information about others, which can lead to inefficiently low levels of privacy.

Even when people have an interest in sharing personal health information with healthcare or other service providers, they may still want to keep it secret from other firms or people, where it can be used against them. Health privacy rules that limit all data collection or re-disclosure may be too broad because they prevent even positive data uses. At the same time, by focusing on the party making the disclosure (and sending the data), they may not do enough to align incentives and prevent the most severe financial risks associated with health data loss. Another approach to privacy protection is therefore to regulate how companies, primarily employers and insurance providers, can use personal health information, by restricting the types of personal information they can acquire or consider.

Health privacy rules that focus on data use by employers or insurers operate as anti-discrimination rules that treat health characteristics as protected categories. A prime example is the 2008 Genetic Information Non-discrimination Act (GINA) that treats genetic information as a protected category. Other federal laws that ban discrimination based on health-related information include the 1978 Pregnancy Discrimination Act, the 1990 Americans with Disabilities Act (ADA), the 2010 Patient Protection and Affordable Care Act (ACA)'s ban on individual insurance market consideration of information on pre-existing health conditions. The provisions in HIPAA (outside the Privacy and Security Rules) restricting the treatment of pre-existing conditions in group health plans would also fit in this category. Restrictions on data use are not focused on the informational aspect, but they can function similarly. Interestingly, the equal protection requirements of these rules could increase the amount of health information workers disclosure to their employers, particularly if they seek workplace accommodations under the ADA.[26]

[25] There may also be a role for public enforcement of privacy rules. For the general literature on public versus private enforcement, see, for example, Landes and Posner (1975) and Polinsky and Shavell (2000).

[26] Because of this, title I of the ADA also features requirements that employers preserve the privacy of information they receive about employee's health conditions. The requirements are not limited to materials produced by healthcare providers or that contain information on medical diagnoses or treatments and include information such as accommodation requests.

Anti-discrimination rules can increase market efficiency if the source of the discrimination was based on bias or animus against people with certain genetic or health conditions, but they can be source of inefficiency if they block companies from using economically relevant information. When insurers are not able to incorporate information on health conditions in setting premiums or coverage levels, that could significantly improve access to insurance for individuals with higher expected health costs, but it could worsen access for people without those conditions if their premiums are raised. At a market level, making health information invisible (or non-actionable) to insurers effectively creates an information asymmetry, because the information is known to consumers. This can cause problems of adverse selection if people with fewer medical risks opt to reduce their insurance purchases or if companies attempt to steer consumers using non-price features to different contracts based on health status (Handel et al. 2015, Oster et al. 2010, Einav et al. 2010, Akerlof 1970, Rothschild and Stiglitz, 1976). This has a potential efficiency cost in reducing the size of the market.

Similar distortions can occur in labor markets if the antidiscrimination rule causes average expected productivity, overall or for identifiable groups, to drop so much that hiring is curtailed (e.g., Herman and Katz 2006). Companies may try to circumvent anti-discrimination rules that focus on specific types of health information by increasing attention to non-protected data elements that function as proxies. The use of proxies can have its own perverse effects, as seen in the findings of increased racial disparities in labor markets from regulations delaying employer access to criminal background information (Doleac and Hanen 2020, Agan and Starr 2018) and of labor market effects of the Pregnancy Discrimination Act affecting all women of childbearing ages (Gruber 1994). It is notable that in both cases the affected groups are themselves protected classes under labor market antidiscrimination rules; this points to the general challenge of enforcing these rules and of detecting violations.

Policy debates around restricting the use of health information in insurance markets therefore tend to center on the tradeoff between the inefficiency produced by increasing the potential role for adverse selection and the distributional aim of providing financial support to people with adverse health shocks (Posner 1981).[27] Yet even when motivated by fairness concerns, the distributional aims of health privacy rules have been argued to serve efficiency goals. One basis for this is similar to the idea behind social insurance more generally, namely, that the *ex post* redistribution that happens after negative health shocks are realized provides *ex ante* insurance to the population from the financial risk of experiencing those shocks (a form of insurance that private markets are ill-equipped to supply). A second argument for the efficiency of redistribution based on negative health status is if people have altruistic preferences and care about the wellbeing of people in poor health, wanting them to have access to healthcare and gainful employments. While those preferences can be expressed through voluntary contributions to private charities, the wellbeing of the disadvantaged group takes on the characteristics of a public good – non-excludable and non-rivalrous. In that case, private charity will under-provide, and

[27] There can also be moral hazard concern from rules that prevent insurers from pricing based on health status, because they lower individual incentives to make "self-protective" investments in preventative care, a healthy diet, and regular exercise (Ehrlich and Becker 1972).

there can be a social gain from increasing the level of support. A counterpoint to these arguments is that health privacy rules, or indeed health status, may not be the best way to target the neediest populations, as was found in the analysis of the federal disability insurance program in Deshpande and Lockwood (2022).

Health privacy rules typically include special restrictions on the use of personal health information for marketing purposes. Nevertheless, in recognition of the potential value to consumers of some targeted marketing of health and insurance products based on their health data, these uses are not typically banned. Even the HIPAA Privacy Rule allows for certain limited marketing uses of personal health data without prior authorization, such as communication about products and services at their current provider or relevant to their course of treatment or disease management.

Outside of the specific exceptions, however, permission from patients is needed before their HIPAA-protected health information can be used, transferred, or sold for marketing purposes. This includes data from certain medical and wellness tracking devices and apps that are used under physician direction, but it leaves most consumer health products outside of the scope of federal privacy rules.[28] Some mobile medical apps are regulated by the Food and Drug Administration, but the requirements are for security and not confidentiality. The FTC acts against companies that violate the terms of their privacy or security policies, or that make false claims about their data practices, under general consumer protection rules against unfair or deceptive practices.[29] Taken together, these rules leave a substantial amount of health information outside of federal privacy rules. Extending the full set of HIPAA protections to these other sources of health information would impose significant costs on companies that collect, disseminate, or share health data, and could reduce the quality and variety of health products that rely on consumer data.

Finally, it is important to note that health privacy laws often include provisions aimed at increasing the flow of health information. One way they do this is by granting to consumers greater rights of access to information about their own health that is held by companies. These provisions fit with the idea of health privacy laws as providing a form of consumer protection and are discussed more in Section 5.3 below. Another important set of provisions are the exemptions and carve-outs from other privacy rules that are aimed at allowing for health data application that serve public good (like the "national priority purposes" in HIPAA), discussed in Section 3. In addition to these relaxations of existing privacy rules, the government also provides significant public subsidies to support voluntary participation and health data sharing for research purposes and makes data sharing compulsory for public health registries and administrative oversight.

[28] State health privacy rules sometimes have broader scope of covered entities or stricter provisions, but they also tend to focus on healthcare providers rather than general health information.

[29] Recent health privacy cases, listed on the FTC webpage at https://www.ftc.gov/legal-library/browse/cases-proceedings, include the 2022 action against Kochava, Inc., and the settlements with Flo Health in 2021 and with SkyMed in 2020.

Furthermore, because the initial provision of health information, or contracting decision, involves a voluntary choice by consumers, it is theoretically possible that stronger privacy protections, such as those that limit data re-disclosure or discriminatory uses, could serve to increase the initial supply of information. This can happen if the legal rules provide reassurance and structure that makes people willing to provide information and use data-demanding products. Absent this reassurance, people may avoid seeking care at all, or invest in costly efforts to mask or obscure their identities when seeking treatment or provide limited or misleading information for their records (as discussed in Section 2.2).

5. Insights from Empirical Studies of Health Privacy Regulation

Theoretical predictions about the effects of privacy laws depend fundamentally on factors that are hard to measure in advance, creating a pressing need for empirical analysis to guide ongoing policy debates. This section discusses four key insights from the empirical literature on health privacy regulation.

5.1 Privacy Rules Can Inhibit Digitization

The first insight from the economics literature on health privacy policy is empirical confirmation of the prediction that privacy laws can inhibit digitization of health information. This is the main finding in Miller and Tucker's (2009) analysis of new adoption of electronic medical records (EMRs) in US hospitals between 1996 and 2005. Controlling for hospital and year fixed effects, as well as a variety of time-varying factors related to the hospital and local area, strict health privacy rules governing disclosure of patient data are found to have significantly slowed the diffusion of EMRs, reducing annual adoption at hospitals by over 24%.

The mechanism through which privacy laws depressed EMR adoption is from the elimination of the positive network effects that would otherwise cause EMR adoption by one hospital in a health services area to increase the likelihood of adoption by other hospitals in the same area. This channel is consistent with the privacy laws reducing the net benefit of EMRs to hospitals by increasing the costs of sharing data, effectively creating a regulatory barrier preventing hospitals from realizing the value of the technological innovation in EMRs that reduces the cost of exchanging patient data. This result does not imply that privacy regulation did not also serve a role in reassuring some patients and increasing their comfort with digital health data collection and sharing, only that the positive effect for consumers was small relative to the negative effect on firms.

The empirical variation in privacy rules comes from primarily from states adopting laws that preceded the HIPAA Privacy Rule. These rules exceeded the relatively weak provisions in HIPAA related to data re-disclosure (Rothstein 2007). Because they were not preempted by HIPAA, they continued to be operative after its enactment, unless modified at the state level (Pritts 2001, Pritts et al. 2009). As shown in Figures 1 and 2 of the paper, these laws were geographical dispersed across the country, but more common in larger states with higher average income levels. The paper therefore also addresses the concern that their presence may be endogenous,

by repeating the analysis using state political representation to instrument for privacy laws, confirming the significant negative effect of privacy laws on EMR adoption, operating through spillovers from other local adoption.

This core finding is again replicated and then extended in Miller and Tucker (2011a), which further examines the welfare impact of delayed EMR adoption by studying its effects on neonatal mortality. The paper focuses on newborn health because it is a key measure of health system performance on which the US routinely underperforms relative to other high-income countries. It is also a setting in which EMRs can also be particularly valuable by helping medical specialists monitor and access patient data needed to track and manage the progress of high-risk pregnancies and births.

Using 11 years of panel data derived from vital statistics records of every live birth and infant death in the nation, the paper finds that hospital adoption of EMRs is associated with significant reductions in neonatal mortality, after controlling for a wide range of maternal, pregnancy, hospital, and county level controls, as well as location and year fixed effects. A 10% increase in basic EMR adoption in a county is associated with a reduction of neonatal deaths of 16 per 100,000 live births, with larger reductions coming when EMR adoption is coupled with adoption of an obstetric-specific IT system. Consistent with predictions, the improvements in neonatal survival were largest for pregnancies with perinatal complications and premature births and not present for pregnancies with no prenatal care (and therefore no prior data to access). EMRs also had no effects on mortality from causes that are not affected by information flows, including congenital defects, sudden infant death syndrome, and accidents.

To address concerns about the potential endogeneity of adoption decisions by hospitals, the study uses state-level health privacy laws as a source of instrumental variables for EMR adoption. The instruments include indicators for having in a place a variety of health privacy provisions (as discussed in Section 4.1), as well as interactions between re-disclosure rules and the latent value of health data exchange (using the size of the hospital, its membership in a system of hospitals and the number of other hospitals in the area; each of these is also included as controls in the main equation). The first stage estimates reveal that disclosure rules are the most important provisions that slow EMR adoption overall, with effect sizes that are larger for hospitals with more local opportunity for data exchange (with more hospitals in the area) and smaller at large hospitals that may have less need for data sharing. The IV estimates confirm the OLS finding that health IT adoption improves neonatal survival rates.

These results, together with other studies of health IT, discussed in Section 3, suggest significant social costs from delayed EMR adoption on hospital quality.[30] Although these studies have focused on the adoption and use of existing technologies, the effect of privacy laws is unlikely to

[30] Derksen et al. (2022) further illustrates the conflict between respecting privacy preferences and harnessing the value of IT in healthcare delivery. EMR adoption in Malawian clinics lowered AIDS mortality through improved tracing of HIV-positive patients for follow-up care; the largest benefits were among patients who asked not to be traced.

be limited to those outcomes. To the extent that privacy laws affect expected adoption rates for health IT, they can affect investments in innovation and the future evolution of technology (similar to the dynamic effects from vaccine policy in Finkelstein 2004).

5.2 Different Privacy Rules Produce Different Effects

A second insight from the empirical literature is that different approaches to health privacy policy can produce different effects. This result is perhaps best illustrated in Miller and Tucker (2018), which studies the effects of different types of state laws addressing genetic privacy. Although health privacy laws often cover genetic information, there is additional policy and advocacy focus on genetic privacy because of the heightened privacy risks from genetic information disclosure (Hellman 2003, Oster et al. 2010). Genetic data can reveal a significant amount of information about a person and their biological relatives. Unlike internet browsing or phone location histories, genetic information is persistently informative over a person's lifetime, and the meanings and uses of the information are likely to expand in unpredictable ways as biological science advances. At the same time, genetic information is increasingly valuable in healthcare for disease prevention and treatment, and it is a key input in the development of more personalized medicine (e.g., All of Us Research Program Investigators 2019).

Miller and Tucker (2018) studies individual decisions to undergo genetic testing for cancer risks, using data from over 80,000 people surveyed across the 2000, 2005 and 2010 waves of the National Health Interview Surveys (NHIS) Cancer Control Modules. Patients with known genetic markers for cancer risk (such as BRCA1 or BRACA2 mutations) can received tailored care, such as more frequent screenings or preventative medication (such as raloxifene or tamoxifen) or surgery (such as prophylactic mastectomy). Yet testing rates are very low in the population (under 1%), and even among populations with elevated risk factors from family history who have discussed genetic testing with their physician (only 20%). Availability of genetic testing services at hospitals is also limited, covering only about 11% of hospitals in the American Hospital Association (AHA) survey.

Concerns about increased privacy risks and potential discriminatory uses of genetic information have been proposed as possible reasons for the low testing rates and as a motivation for specific privacy rules that address genetic data to reassure patients and increase their willingness to seek testing. At the same time, the results from Miller and Tucker (2009) and (2011a) suggest that privacy rules may lower availability of testing at hospitals if they face increased compliance costs or perceive the data to be less valuable.

The analysis in Miller and Tucker (2018) therefore empirically examines the separate effects of 3 different dimensions of genetic privacy laws: 1) explicit notification requirement on privacy risks as part of informed consent; 2) requirements that companies obtain individual consent before data re-disclosure, effectively assigning ownership rights to individuals over their data; and 3) restrictions on downstream uses of data through anti-discrimination rules. State laws typically include one or more of these protections, while GINA (which comes into effect at the end of the sample period) focuses on the third related to discriminatory uses.

Consistent with the countervailing mechanisms inherent in privacy rules, the paper finds different empirical effects of the different dimensions of privacy policy. The policy of requiring clear and detailed notification of privacy risks is associated with lower testing rates, but policies that strengthen patients' ownership and control over their data are associated with increased adoption. Restrictions on third-party discriminatory uses, at the state level or federal level from GINA, are not found to have any detectable impact on testing rates. While the finding that ownership right increase testing is promising for the potential role of privacy rules to reassure the public, the implications from the other results are more concerning. The importance of notification requirements appears to come in large part from lower supply of testing at hospitals, but it may also come from the greater salience of privacy at the time of testing decisions or from better guiding consumers in making informed choices about their information.

Similarly, the lack of an effect of anti-discrimination laws on testing could reflect the difficulties that consumers anticipate in detecting illegal discrimination and enforcing of future claims. This null effect differs from findings in the literature on the effects of the ADA, which prohibits discrimination by disability status (DeLeire 2000, Acemoglu and Angrist 2001, Jolls 2004) and of the pregnancy discrimination act (Gruber 1994). One potential source for that difference is that those studies focused on labor market effects rather than on data generation and disclosure.

5.3 Privacy Rules Can Sometimes Increase Data Flows

In contrast to the theoretical concerns that privacy rules increase the costs of information exchange and lower the value of health IT, supported by the empirical findings in Miller and Tucker (2009 and 2011a) that privacy rules lower EMR adoption, another insight from the literature is that privacy rules can increase IT adoption and data use in some circumstances.

One mechanism for this is that privacy laws provide reassurance to patients about data security and limits on re-use that makes them more willing to undertake medical testing or seek treatments that will create sensitive records. This mechanism was supported empirically for genetic testing in Miller and Tucker (2018). It was also argued as a reason why privacy rules could help promote health information exchange (McGraw et al. 2009), as found in Adjerid et al. (2016) when combined with and financial incentives for adoption. The findings in Buckman et al. (2022) for COVID-19 vaccination rates similarly support the idea that privacy concerns can reduce healthcare seeking for some patients, and that legal health privacy protections (in their case the right to remove identifying information from the vaccine registry) can provide the needed reassurance.

A second mechanism is that privacy laws that strengthen consumers' ownership rights over their health information can increase the production and use of health data by making it easier to extract data from the control of organizations that generate and control it. Giving patients more control over their data can increase data flaws because patients sometimes want their data to be transferred and shared more easily than providers do. When patient records belong to healthcare providers, it becomes a form of proprietary information that some firms want to keep siloed away from competitors. One such motivation for hospitals is to "lock-in" existing patients

for follow-on care. Miller and Tucker (2014) find empirical support for this in the lower rates of external health data exchange among hospitals that are part of larger systems. This pattern is present despite the fact that those hospitals tend to have greater technological capacity for health IT and engage in significantly more internal data exchange within their systems. [31] The effect is larger for hospitals with patients who are otherwise more mobile (having non-HMO insurance coverage), with higher paid staff, and with specialty services (such as cariology or oncology), suggesting that the reduction in external data exchanges come from a strategic motivation to retain patients.

Notwithstanding these considerations, the "meaningful use" requirements for health IT data compatibility and exchange in the original HITECH Act were focused solely on the technological capacity and ability to exchange data. This proved insufficient when providers and vendors had financial incentives to block data flows (Pear 2015). This was addressed in part in the 21st Century Cures Act of 2016 (Cures Act), which updated the HITECH Act (effective April 2021) to prohibit data blocking by technology vendors or healthcare providers.

These results point to an important aspect of health privacy rules that is sometimes neglected in the literature. Although the rules often aim to restrict information flows and prevent unwanted disclosures, those are not their only goals. Privacy rules often originate from a consumer protection perspective and aim at bolstering individual property rights over personal data. This is reflected in US federal laws described above and in the GDPR provisions related to individual rights of data access and erasure and to data portability.[32]

Advancing the goals of increasing consumer control over their data can increase information flows and improve efficiency if it reduces inefficient data hoarding by companies. This idea is central to the theory Jones and Tonetti (2020), who focus on the non-rivalrous aspect of data use. The idea also receives empirical support in the finding in Baker et al. (2015) that rules that increase consumer access to their health data, in their case, state laws that capped the charges that healthcare providers could impose for copies of paper medical records, led to increased adoption of EMRs at hospitals.

5.4 Regulating Technology Is Not Enough

A final theme from the empirical literature is that a focus on technological solutions alone is not enough to protect privacy. This was seen above when "meaningful use" rules for technological capability were not enough to ensure meaningful flow of patient data across providers. It is further illustrated in Miller and Tucker's (2011b) study of health data security.

The technology of interest in that paper is data encryption, which can preserve privacy in the event of a breach, by rendering the information unreadable to anyone without the key. Data

[31] Concerns about corporate data control being used to foreclose competition arises outside of healthcare as well, for example, in debates about car repairs (Magliozzi 2022).

[32] Data access rights are also included in state-level broad-based privacy laws in California, Virginia,

security rules in several states recognize this feature by exempting breached data from mandatory disclosure requirements if encryption was in place. Miller and Tucker (2011b) first confirms the motivation for devoting special attention to the security of electronic health records, by showing that digitization of hospital records indeed increases the loss of patient data through publicized data breaches, and then studies the role of data encryption in preventing data loss. State policies that exempt encrypted data from breach notification rules are found to have their intended effect of increasing the adoption of data encryption at hospitals. However, the paper also finds that encryption alone does not reduce the amount of data lost. Instead, public reports of lost data at higher at hospitals with encryption. To address concerns that this relationship could come from hospitals with higher value data being both more likely to adopt encryption and more likely to be targeted, the paper also estimates effects of encryption using the legal variation as a source of exogenous variation. Again, the results confirm that encryption increases data loss. The reason for this surprising effect is that encryption increases data loss from internal fraud and from lost equipment.

This result highlights a key challenge in regulating data privacy and security. Technical solutions can be effective for firms but focusing policy on them can be counterproductive if it draws attention away from human factors that also contribute to data protection. This is analogous to the multi-task principal-agent problem in employment contracts (Holmstrom and Milgrom 1991) and an example of the general problem of unintended consequences of regulation. There is also a lesson specific to privacy regulation. Technological change is the key source of the increased risks of privacy loss, so understanding and addressing technology is essential to managing the risk. However, the solutions will not come from technology alone. Privacy problems are inherently about human behaviors; effective privacy policy needs to keep human factors, such as cost and incentives, at its center. This is true even for data security, where firms and consumers share some common interest in data protection. Incentives must play an even larger role in addressing confidentiality and intentional disclosures.

6. Conclusion

The economic approach to digital health privacy, presented in this chapter, is a complement to approaches from other fields centered on legal rights and principles or on technological challenges and solutions. The approach is characterized by its consideration of costs and benefits of different data uses and restrictions and it is grounded in both theory models and empirical evidence of how firms and individuals make decisions and how market operate. These features make the economic literature on health privacy particularly promising for providing a foundation for assessing the impacts of existing health privacy rules, and for predicting effects of new rules. Although the focus of this chapter is on health privacy, the increasing prevalence of health-related information outside of traditional medical and insurance settings presents new challenges for policymakers, raising questions about the desirability of expanding the scope of existing health privacy rules or enacting broad-based privacy rules. Economic research in health privacy can therefore serve to inform pressing policy debates and to advance scientific understanding of the fundamental tradeoffs between preserving privacy and harnessing the value of IT and data-driven innovation in healthcare.

References

Abowd, J. M., & Schmutte, I. M. (2019). An economic analysis of privacy protection and statistical accuracy as social choices. *American Economic Review*, *109*(1), 171-202.

Acemoglu, D., & Angrist, J. D. (2001). Consequences of employment protection? The case of the Americans with Disabilities Act. *Journal of Political Economy*, *109*(5), 915-957.

Acemoglu, D., Makhdoumi, A., Malekian, A., & Ozdaglar, A. (2022). Too much data: Prices and inefficiencies in data markets. *American Economic Journal: Microeconomics*, *14*(4), 218-56.

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and Human Behavior in the Age of Information. *Science*, *347*(6221), 509-514.

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2020). Secrets and likes: The drive for privacy and the difficulty of achieving it in the digital age. *Journal of Consumer Psychology*, *30*(4), 736-758.

Acquisti, A., Taylor, C., & Wagman, L. (2016). The Economics of Privacy. *Journal of Economic Literature*, *54*(2), 442-92.

Adjerid, I., Acquisti, A., & Loewenstein, G. (2019). Choice architecture, framing, and cascaded privacy choices. *Management Science*, *65*(5), 2267-2290.

Adjerid, I., Acquisti, A., Telang, R., Padman, R., & Adler-Milstein, J. (2016). The impact of privacy regulation and technology incentives: The case of health information exchanges. *Management Science*, *62*(4), 1042-1063.

Agan, A., & Starr, S. (2018). Ban the box, criminal records, and racial discrimination: A field experiment. *Quarterly Journal of Economics*, *133*(1), 191-235.

Agha, L. (2014). The Effects of Health Information Technology on the Costs and Quality of Medical Care. *Journal of Health Economics* 34:19–30.

Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *ILR Review*, *30*(2), 175-187.

All of Us Research Program Investigators. (2019). The "All of Us" research program. *New England Journal of Medicine*, *381*(7), 668-676.

Alsan, M., & Wanamaker, M. (2018). Tuskegee and the health of Black men. *Quarterly Journal of Economics*, *133*(1), 407-455.

Arrow, K. (1973). The Theory of Discrimination, in Orley Ashenfelter and Albert Rees, eds., Discrimination in Labor Markets, Princeton University Press.

Atasoy, H., Greenwood, B. N., & McCullough, J. S. (2019). The digitization of patient care: a review of the effects of electronic health records on health care quality and utilization. *Annual Review of Public Health*, *40*, 487-500.

Athey, S., Catalini, C., & Tucker, C. (2017). The digital privacy paradox: Small money, small costs, small talk. NBER Working Paper #23488.

Atienza, A., Zarcadoolas, C., Vaughon, W., Hughes, P., Patel, V., Chou, W-Y. S. & Pritts, J. (2015). Consumer Attitudes and Perceptions on mHealth Privacy and Security: Findings from a Mixed-Methods Study, *Journal of Health Communication*, 20:6, 673-679

Badgett, M. L. (2007). *Sexual Orientation Discrimination: An International Perspective*, Routledge.

Baldwin, M. L., & Johnson, W. G. (2006). A critical review of studies of discrimination against workers with disabilities. In ed. Rodgers, W. M. *Handbook on the Economics of Discrimination*. Edward Elgar Publishing.

Baker, L. C., Bundorf, M. K., & Kessler, D. P. (2015). Expanding patients' property rights in their medical records. *American Journal of Health Economics*, *1*(1), 82-100.

Bartlett, R., Morse, A., Stanton, R., & Wallace, N. (2022). Consumer-lending discrimination in the FinTech era. *Journal of Financial Economics*, *143*(1), 30-56.

Bates, D. W., & Syrowatka, A. (2022). Harnessing AI in sepsis care. *Nature Medicine*, *28*(7), 1351-1352.

Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago Press.

Bhargava, S., & Loewenstein, G. (2015). Behavioral economics and public policy 102: Beyond nudging. *American Economic Review*, *105*(5), 396-401.

Blenner, S. R., Köllmer, M., Rouse, A. J., Daneshvar, N., Williams, C., & Andrews, L. B. (2016). Privacy policies of android diabetes apps and sharing of health information. *JAMA*, *315*(10), 1051-1052.

Bronsoler, A., Doyle, J., & Van Reenen, J. (2022). The Impact of Health Information and Communication Technology on Clinical Quality, Productivity, and Workers. *Annual Review of Economics*, 14.

Buckman, J., Adjerid, I., & Tucker, C. E. (forthcoming). Privacy Regulation and Barriers to Public Health. *Management Science*.

Claxton, G., Rae, M., Damico, A., Wager, E., Young, G., & Whitmore, H. (2022). Health Benefits In 2022: Premiums Remain Steady, Many Employers Report Limited Provider Networks For Behavioral Health: Study examines employer-sponsored health benefits in 2022. *Health Affairs*, 10-1377.

Coase, R.H. (1960). The Problem of Social Cost. *Journal of Law and Economics*, 3, 1-44.

DeLeire, T. (2001). Changes in wage discrimination against people with disabilities: 1984-93. *Journal of Human Resources*, 144-158.

Derksen, L., McGahan, A., & Pongeluppe, L. (2022). Privacy at What Cost? Using Electronic Medical Records to Recover Lapsed Patients Into HIV Care. Working Paper.

Deshpande, M., & Lockwood, L. M. (2022). Beyond health: Non-health risk and the value of disability insurance. *Econometrica*, *90*(4), 1781-1810.

Doleac, J. L., & Hansen, B. (2020). The unintended consequences of "ban the box": Statistical discrimination and employment outcomes when criminal histories are hidden. *Journal of Labor Economics*, *38*(2), 321-374.

Dranove, D., Forman, C., Goldfarb, A., & Greenstein, S. (2014). The Trillion Dollar Conundrum: Complementarities and Health Information Technology. *American Economic Journal: Economic Policy* 6 (4):239–70.

Duhigg, C. (Feb. 6, 2012). See How Companies Learn Your Secrets. *New York Times*.

Ehrlich, I., & Becker, G. S. (1972). Market Insurance, Self-Insurance, and Self-Protection. *Journal of Political Economy*, *80*(4), 623-648.

Einav, L., Finkelstein, A., & Cullen, M. R. (2010). Estimating welfare in insurance markets using variation in prices. *Quarterly Journal of Economics*, *125*(3), 877-921.

European Data Protection Supervisor (EDPS). (May 21, 2015). Mobile health: Reconciling technological innovation with data protection. Available at:

https://edps.europa.eu/sites/edp/files/publication/15-05-21_mhealth_en_0.pdf
Figueroa, C., Johnson, C., Verster, A., & Baggaley, R. (2015). Attitudes and acceptability on HIV self-testing among key populations: a literature review. *AIDS and Behavior*, *19*(11), 1949-1965.
Finkelstein, A. (2004). Static and dynamic effects of health policy: Evidence from the vaccine industry. *The Quarterly Journal of Economics*, *119*(2), 527-564.
Freedman, S., Lin, H., & Prince, J. (2018). Information technology and patient health: analyzing outcomes, populations, and mechanisms. *American Journal of Health Economics*, *4*(1), 51-79.
Friedman, C., Rubin, J., Brown, J., Buntin, M., Corn, M., Etheredge, L., Gunter, C., Musen, M., Platt, R., Stead, W. and Sullivan, K. (2015). Toward a science of learning systems: a research agenda for the high-functioning Learning Health System. *Journal of the American Medical Informatics Association*, *22*(1), 43-50.
Ganju, K. K., Atasoy, H., McCullough, J., & Greenwood, B. (2020). The role of decision support systems in attenuating racial biases in healthcare delivery. *Management science*, *66*(11), 5171-5181.
Goldfarb, A., Taska, B., & Teodoridis, F. (2020). Artificial intelligence in health care? Evidence from online job postings. AEA Papers and Proceedings (Vol. 110, pp. 400-404).
Goldfarb, A., & Tucker, C. (2012). Shifts in Privacy Concerns. *American Economic Review*, *102*(3), 349-53.
Goldfarb, A., & Tucker, C. (2019). Digital economics. *Journal of Economic Literature*, *57*(1), 3-43.
Gostin, L. (1991). Genetic Discrimination: The Use of Genetically Based Diagnostic and Prognostic Tests by Employers and Insurers. *American Journal of Law & Medicine* 17:109.
Gostin, L. O. (1994). Health information privacy. *Cornell L. Rev.*, *80*, 451.
Gresenz, C. R., Laughery, S., Miller, A. R., & Tucker, C. E. (2017). Health IT and ambulatory care quality. SSRN Working Paper #2665664.
Gruber, J. (1994). The incidence of mandated maternity benefits. *American Economic Review*, 622-641.
Halpern, S. (2020, April 27). Can we track COVID-19 and protect privacy at the same time? *New Yorker Magazine*.
Handel, B., Hendel, I., & Whinston, M. D. (2015). Equilibria in health exchanges: Adverse selection versus reclassification risk. *Econometrica*, *83*(4), 1261-1313.
Hellman, D. (2003). What makes genetic discrimination exceptional? *American Journal of Law & Medicine*, *29*(1), 77-116.
Hermalin, B. E., & Katz, M. L. (2006). Privacy, property rights and efficiency: The economics of privacy as secrecy. *Quantitative marketing and economics*, *4*(3), 209-239.
Hill, K. (June 30, 2022). Deleting Your Period Tracker Won't Protect You. *New York Times*.
Hillestad, R., J. Bigelow, A. Bower, F. Girosi, R. Meili, R. Scoville, and R. Taylor. 2005. "Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, and Costs. *Health Affairs* 24 (5): 1103-17.
Hitt, L. M., & Tambe, P. (2016). Health care information technology, work organization, and nursing home performance. *ILR Review*, *69*(4), 834-859.
Holmstrom, B., & Milgrom, P. (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, & Organization*, *7*, 24-52.

Hotz, V.J., Bollinger, C.R., Komarova, T., Manski, C.F., Moffitt, R.A., Nekipelov, D., Sojourner, A. and Spencer, B.D., 2022. Balancing data privacy and usability in the federal statistical system. *Proceedings of the National Academy of Sciences*, 119(31), p.e2104906119.

Huckvale, K., Torous, J., & Larsen, M. E. (2019). Assessment of the data sharing and privacy practices of smartphone apps for depression and smoking cessation. *JAMA Network Open*, *2*(4), e192542-e192542.

Ichihashi, S. (2020). Online privacy and information disclosure by consumers. *American Economic Review*, *110*(2), 569-95.

Institute of Medicine (IOM). (2000). *To err is human: Building a safer health system.* Washington, DC: National Academy Press.

Institute of Medicine (IOM). (2007). *The Learning Healthcare System: Workshop Summary*. Washington, DC: National Academies Press.

Institute of Medicine (IOM) (2009). *Beyond the HIPAA Privacy Rule: Enhancing Privacy, Improving Health Through Research*. Washington, DC: National Academies Press.

Institute of Medicine (IOM) (2013). *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America*. Washington, DC: National Academies Press.

Janakiraman, R., Park, E., M. Demirezen, E., & Kumar, S. (2022). The effects of health information exchange access on healthcare quality and efficiency: An empirical investigation. *Management Science*.

Jha, A. K., DesRoches, C. M., Kralovec, P. D., & Joshi, M. S. (2010). A progress report on electronic health records in US hospitals. *Health Affairs*, *29*(10), 1951-1957.

Jolls, C. (2004). Identifying the effects of the Americans with Disabilities Act using state-law variation: preliminary evidence on educational participation effects. *American Economic Review*, *94*(2), 447-453.

Jones, C. I., & Tonetti, C. (2020). Nonrivalry and the Economics of Data. *American Economic Review*, *110*(9), 2819-58.

Kelly, H., Hunter, T., & Abril, D. (June 26, 2022). Seeking an abortion? Here's how to avoid leaving a digital trail. *Washington Post*.

Komarova, T., Nekipelov, N. & Yakovlev, E. (2018). Identification, Data combination, and the Risk of Disclosure. *Quantitative Economics* 9, no. 1: 395-440.

Kummer, M., & Schulte, P. (2019). When private information settles the bill: Money and privacy in Google's market for smartphone applications. *Management Science*, *65*(8), 3470-3494.

Lammers, E. J., Adler-Milstein, J., & Kocher, K. (2014). Does Health Information Exchange Reduce Redundant Imaging? Evidence from Emergency Departments. *Medical Care* 52 (3):227–34.

Landes, W. M., & Posner, R. A. (1975). The private enforcement of law. *The Journal of Legal Studies*, *4*(1), 1-46.

Lin, T. (2022). Valuing intrinsic and instrumental preferences for privacy. *Marketing Science*.

Lin, Y. K., Lin, M., & Chen, H. (2019). Do electronic health records affect quality of care? Evidence from the HITECH Act. *Information Systems Research*, *30*(1), 306-318.

Loertscher, S., & Marx, L. M. (2020). Digital monopolies: Privacy protection or price regulation? *International Journal of Industrial Organization*, *71*, 102623.

Maclean, J. C., Mallatt, J., Ruhm, C. J., & Simon, K. (2020). Economic studies on the opioid crisis: A review. NBER Working Paper #28067.

Magliozzi, R. (2022). 'Car Talk' host: Independent auto shops deserve the right to repair your car. *Washington Post*.

McCullough, J. S., Parente, S. T., & Town, R. (2016). Health information technology and patient outcomes: the role of information and labor coordination. *The RAND Journal of Economics*, *47*(1), 207-236.

McGraw, D., Dempsey, J. X., Harris, L., & Goldman, J. (2009). Privacy as an enabler, not an impediment: building trust into health information exchange. *Health Affairs*, *28*(2), 416-427.

Mechanic, D. (1998). The functions and limitations of trust in the provision of medical care. *Journal of Health Politics, Policy and Law*, *23*(4), 661-686.

Merchant, R.M., Asch, D.A., Crutchley, P., Ungar, L.H., Guntuku, S.C., Eichstaedt, J.C., Hill, S., Padrez, K., Smith, R.J. and Schwartz, H.A., 2019. Evaluating the predictability of medical conditions from social media posts. *PloS One*, *14*(6), p.e0215476.

Miller, A. R., Eibner, C., & Gresenz, C. R. (2013). Financing of employer sponsored health insurance plans before and after health reform: What consumers don't know won't hurt them? *International Review of Law and Economics*, *36*, 36-47.

Miller, A. R., Ramdas, K., & Sungu, A. (2021). Browsers Don't Lie? Gender Differences in the Effects of the Indian COVID-19 Lockdown on Digital Activity and Time Use. SSRN Working Paper #3930079.

Miller, A. R., & Tucker, C. (2009). Privacy Protection and Technology Diffusion: The Case of Electronic Medical Records. *Management Science*, *55*(7), 1077-1093.

Miller, A. R., & Tucker, C. (2011). Can Health Care Information Technology Save Babies? *Journal of Political Economy*, *119*(2), 289-324.

Miller, A. R., & Tucker, C. (2011). Encryption and the Loss of Patient Data. *Journal of Policy Analysis and Management*, *30*(3), 534-556.

Miller, A. R., & Tucker, C. (2014). Health Information Exchange, System Size and Information Silos. *Journal of Health Economics*, *33*, 28-42.

Miller, A. R., & Tucker, C. (2017). Frontiers of health policy: Digital data and personalized medicine. *Innovation Policy and the Economy*, *17*(1), 49-75.

Miller, A. R., & Tucker, C. (2018). Privacy Protection, Personalized Medicine, and Genetic Testing. *Management Science*, *64*(10), 4648-4668.

Murphy, K. M., & Topel, R. H. (2003). Measuring the gains from medical research: An economic approach. University of Chicago Press.

Murphy, K. M., & Topel, R. H. (2006). The value of health and longevity. *Journal of Political Economy*, *114*(5), 871-904.

Nissenbaum, H. (2004). Privacy as Contextual Integrity. *Washington Law Review*, *79*, 119.

Nix, N., & Dwoskin, E. (August 12, 2022). Search warrants for abortion data leave tech companies few options. *Washington Post*.

Organization for Economic Co-operation and Development (OECD). (2022). Health Data Governance for the Digital Age: Implementing the OECD Recommendation on Health Data Governance, OECD Publishing, Paris, https://doi.org/10.1787/68b60796-en.

Oster, E., Dorsey, E. R., Bausch, J., Shinaman, A., Kayson, E., Oakes, D., Shoulson, I. & Quaid, K. (2008). Fear of health insurance loss among individuals at risk for Huntington disease. *American Journal of Medical Genetics Part A*, *146*(16), 2070-2077.

Oster, E., Shoulson, I., & Dorsey, E. (2013). Optimal expectations and limited medical testing: Evidence from Huntington disease. *American Economic Review*, *103*(2), 804-30.

Oster, E., Shoulson, I., Quaid, K., & Dorsey, E. R. (2010). Genetic adverse selection: Evidence from long-term care insurance and Huntington disease. *Journal of Public Economics*, *94*(11-12), 1041-1050.

Pear, R. (May 26, 2015). Tech Rivalries Impede Digital Medical Record Sharing. *New York Times*.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review*, *62*(4), 659-661.

Polinsky, A. M., & Shavell, S. (2000). The economic theory of public enforcement of law. *Journal of Economic Literature*, *38*(1), 45-76.

Ponemon Institute (2013). *Survey on Medical Identity Theft.* Available at: https://www.ponemon.org/research/ponemon-library/security/2013-survey-on-medical-identity-thef-2.html

Posner, R. A. (1981). The Economics of Privacy. *American Economic Review*, *71*(2), 405-409.

Price, W. N., & Cohen, I. G. (2019). Privacy in the age of medical big data. *Nature Medicine*, *25*(1), 37-43.

Pritts, J. L. (2001). Altered states: state health privacy laws and the impact of the Federal Health Privacy Rule.

Pritts, J.L. Lewis, S., Jacobson, R., Lucia, K., & Kayne, K. (2009). Privacy and Security Solutions for Interoperable Health Information Exchange: Report on State Law Requirements for Patient Permission to Disclose Health Information. Available at https://www.healthit.gov/sites/default/files/disclosure-report-1.pdf

Qian, I., Xiao, M., Mozur, P., & Cardia, A. (2022). Four Takeaways From a Times Investigation Into China's Expanding Surveillance State. *New York Times*.

Rothstein, M. A. (2007). Health privacy in the electronic age. *Journal of Legal Medicine*, *28*(4), 487-501.

Sanders, S. F., Terwiesch, M., Gordon, W. J., & Stern, A. D. (2019). How artificial intelligence is changing health care delivery. *NEJM Catalyst*, 5(5).

Shilo, S., Rossman, H., & Segal, E. (2020). Axes of a revolution: challenges and promises of big data in healthcare. *Nature Medicine*, *26*(1), 29-38.

Skatova, A., McDonald, R. L., Ma, S., & Maple, C. (2019). Unpacking Privacy: Willingness to pay to protect personal data. Working Paper. DOI: 10.31234/osf.io/ahwe4.

Smalley, E. (2017). AI-powered drug discovery captures pharma interest. *Nature Biotechnology*, *35*(7), 604-606.

Spetz, J., Burgess, J. F. & Phibbs, C. S. (2014). The Effect of Health Information Technology Implementation in Veterans Health Administration Hospitals on Patient Outcomes. *Healthcare* 2 (1):40–47.

Tangari, G., Ikram, M., Ijaz, K., Kaafar, M. A., & Berkovsky, S. (2021). Mobile health and privacy: cross sectional study. *BMJ*, 373:n1248.

U.S. Department of Health and Human Services (HHS). (December 13, 2017). Health Information Privacy Beyond HIPAA: A 2018 Environmental Scan of Major Trends and Challenges. Available online at https://ncvhs.hhs.gov/wp-content/uploads/2018/05/NCVHS-Beyond-HIPAA_Report-Final-02-08-18.pdf.

Vermund, S. H., & Wilson, C. M. (2002). Barriers to HIV testing-where next? *Lancet*, *360*(9341), 1186-1187.

Walker, J., Pan, E., Johnston, D., Adler-Milstein, J., Bates, D. W., & Middleton, B. (2005). The Value of Health Care Information Exchange and Interoperability: There is a business case to be made for spending money on a fully standardized nationwide system. *Health Affairs*, *24*(Suppl1), W5-10.

Yu, K. H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature Biomedical Engineering*, 2(10), 719-731.