Privacy: NBER Tutorial

Catherine Tucker

Question

What should PhD students study about privacy?

1. What is the equation?
2. What does digitization change?
3. How can we inform policy?

Outline

What is the equation?

Digitization

What are the key policy questions?

Question

What is the dependent variable and what is the explanatory variable of interest when we study privacy?

Varian 1996

Let us think about how privacy concerns enter a basic transaction. Suppose the seller has many different kinds of apples (Jonathan, Macintosh, Red Delicious, etc.) The buyer is willing to pay at most r to purchase a Jonathan, and 0 to purchase any other kind of apple.

The buyer will in general not want the seller to know r, the maximum price that he is willing to pay for the item being sold. If this information were available to the seller, the seller would price the product at the buyer's maximum willingness to pay, and the buyer would receive no surplus from the transaction.

But how does privacy enter a utility function?

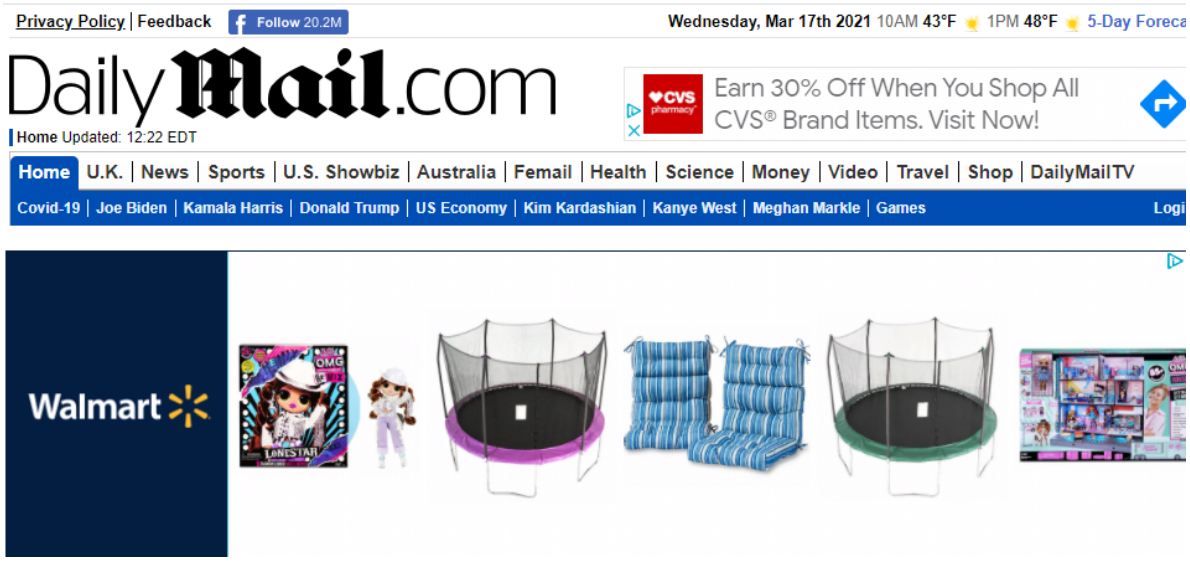


Figure: Do people mean this?

Ask yourself?

- Do you think privacy enters into a utility firm directly or indirectly?

Easiest to hardest equations to operationalize:

- Privacy regulation affects X (Miller and Tucker, 2009; Goldfarb and Tucker, 2011) + Many great GDPR papers
  $\downarrow$
- X affects how consumers value privacy or treat privacy

The Privacy Paradox

Figure: Susan Athey, Christian Catalini and Catherine Tucker (NBER 2018)

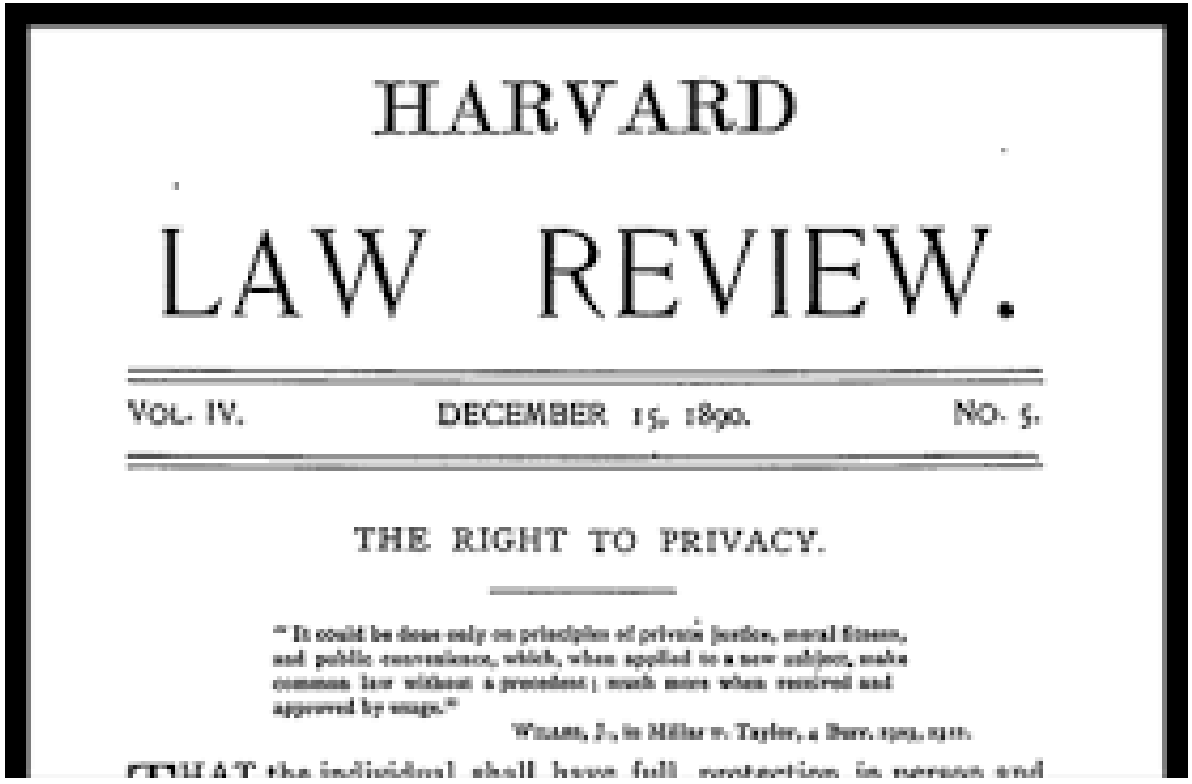
Outline

What is the equation?

Digitization

What are the key policy questions?

Initial Introduction of Privacy to Academia



HARVARD

LAW REVIEW.

VOL. IV. DECEMBER 15, 1890. NO. 5.

THE RIGHT TO PRIVACY.

"It could be done only on principles of private justice, moral fitness, and public convenience, which, when applied to a new subject, make common law without a precedent; much more when received and approved by usage."

WILLES, J., in Millar v. Taylor, 4 Burr. 2303, 2312.

THAT the individual shall have full protection in person and

The economics of digital intrusion has changed

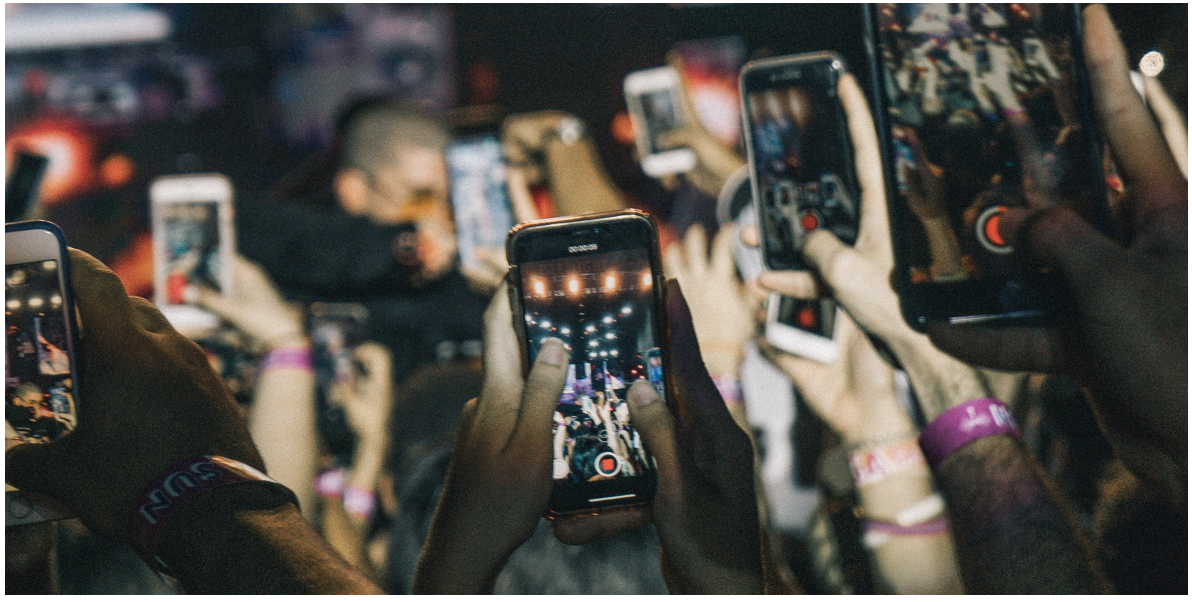


Figure:

What economic implications can we think about?

1. data persistence
2. data spillovers
3. data repurposing

Data persistence: But our privacy preferences change over time leading to dynamics: (Goldfarb and Tucker, 2012)

Shifts in Privacy Concerns

Avi Goldfarb

Catherine Tucker

AMERICAN ECONOMIC REVIEW
VOL. 102, NO. 3, MAY 2012
(pp. 349-53)

Download Full Text PDF

Article Information

Abstract

This paper explores how digitization and the associated use of customer data have affected the evolution of consumer privacy concerns. We measure privacy concerns by reluctance to disclose income in an online marketing research survey. Using over three million responses over eight years, our data show: (1) Refusals to reveal information have risen over time, (2) Older people are less likely to reveal information, and (3) The difference between older and younger people has increased over time. Our results suggest that the trends over time are partly due to broadening perceptions of the contexts in which privacy is relevant.

Citation

Goldfarb, Avi, and Catherine Tucker. 2012. "Shifts in Privacy Concerns." *American Economic Review*, 102 (3): 349-53.

DOI: 10.1257/aer.102.3.349

Fig

Data spillovers: Genetic (Miller and Tucker, 2017) + Visual Data (Augmented Reality, Facial Recognition):

ome > Innovation Policy and the Economy > Volume 17

PREVIOUS ARTICLE NEXT ARTICLE

FREE

Frontiers of Health Policy: Digital Data and Personalized Medicine

Amalia R. Miller and Catherine Tucker

University of Virginia, IZA, and NBERMIT Sloan School of Management and NBER

Abstract | Full Text | PDF

SECTIONS

Abstract

This paper argues that due to two unstoppable mechanisms, some of the most pressing future questions in health policy will relate to the use of digital technologies to analyze data concerning patient health. The first mechanism is the shift away from a system where patient data was essentially temporary and not intended to be reused or easily accessed again, to a new digital world where patient data is easily transferred and accessed repeatedly. The second mechanism is a fundamental deepening of the nature of patient data that enables increased personalization of health care for each individual patient, based on not only their detailed medical history, but also their likely future medical history that can be projected for their genetic makeup. We summarize our research investigating the potential consequences of policies in this new world where patient data is virtually costless to store, share, and individualize. We emphasize that issues of data management and privacy are now at the forefront of health policy considerations.

Digital data and digital technologies have the potential to transform medicine through two mechanisms. First, digital patient data is far easier to share and access than traditional paper records. This has many potential upsides, but also raises the question of how the potential benefits of sharing patient data are moderated by privacy concerns. Second, the advent of digital storage has now made it possible to store, virtually costlessly, vast swathes of data about any one individual patient. Such individualized data also enables a patient-centric approach to medicine, often referred to as "personalized" or "precision" medicine, based on that individual patient's genetic makeup.

Details | Figures | References | Cited by

Innovation Policy and the Economy
Volume 17
2017

Sponsored by the National Bureau of Economic Research (NBER)

Article DOI
https://doi.org/10.1086/688844

Views: 354

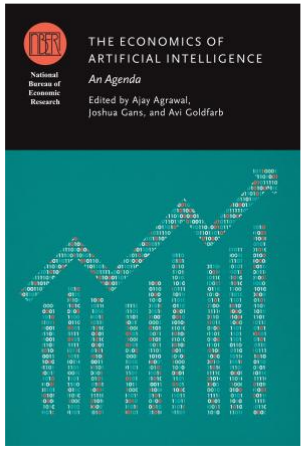
Figure:

Data Repurposing (Miller and Tucker, 2018)

Privacy, Algorithms, and Artificial Intelligence

Catherine Tucker

PUBLISHED DATE May 2019 **COPYRIGHT** 2019 **ISBN** 978-0-226-61333-8



CONFERENCE HELD SEPTEMBER 13-14, 2017
BOOK: THE ECONOMICS OF ARTIFICIAL INTELLIGENCE: AN AGENDA
BOOK EDITORS: AJAY AGRAWAL, JOSHUA GANS & AVI GOLDFARB
PUBLISHER: UNIVERSITY OF CHICAGO PRESS

Artificial intelligence can use an individual's data to make predictions about what they might desire, be influenced by, or do. The use of an individual's data in this process raises privacy concerns. This article focuses on what is novel about the world of artificial intelligence and privacy, arguing that the chief novelty lies in the potential for data persistence, data repurposing and data spillovers.

Figure:

Provocative viewpoint: Given this why is the debate on digital privacy so focused on advertising?

1. data persistence
2. data spillovers
3. data repurposing

Big Picture

Traditional models of the economics of privacy need to also reflect

- Data persistence and dynamics implied
- Data repurposing and uncertainty implied
- Data spillovers and lack of control or choice implied

Outline

What is the equation?

Digitization

What are the key policy questions?

Congressional testimony (which I made you read)

- A piece of personal advice....

Some notes on the framing of the issue

- How the privacy debate has moved from privacy to data-based discrimination to algorithmic bias

Why might algorithms be biased?

- Biased Programmers
- Biased training data
- Bias is learned from humans interacting with the algorithm

What can economists do to inform the algorithmic bias debate

- Explore areas where we can understand the mechanism which might explain algorithmic bias (Lambrecht and Tucker, 2018)
- Point out counterfactual thinking and the existing economics literature (Cowgill and Tucker, 2017)

Your task: Think up your ideal paper that you would write now to inform your congressional testimony in 2030

- Points for imagination
- The privacy debate was reframed as algorithmic bias in 2019. Back to privacy in 2021. You need to think about how to inform the framing?
- What do you think the big policy issue about data will be? What will be sensitive and private data in 2030?

Some of my ideas

1. Are there ways of measuring privacy preferences?
2. Do consumers value personalize advertising?
3. Do consumers distinguish between data privacy and data security?
4. Do consumers have hyperbolic discounting when it comes to future data use? How can we characterize uncertainty over data reuse?
5. Is there any win on the idea of ‘privacy competition’ Or is there always a tradeoff between privacy and competition?
6. Do sectoral or unified approaches to privacy competition work better?
7. What about privacy-protective technologies? Transaction costs or worthwhile?
8. Please study government surveillance

https://www.allourideas.org/privacy_paper_congress

Returning to the idea of what is different: Punchline

1. data persistence
2. data spillovers
3. data repurposing

Thank you! Time for questions and for you to tell me about your work.

cetucker@mit.edu

Cowgill, B. and C. Tucker (2017). Algorithmic bias: A counterfactual perspective. *NSF Trustworthy Algorithms*.

Goldfarb, A. and C. Tucker (2012). Shifts in privacy concerns. *American Economic Review: Papers and Proceedings 102*(3), 349–53.

Goldfarb, A. and C. E. Tucker (2011, January). Privacy regulation and online advertising. *Management Science 57*(1), 57–71.

Lambrecht, A. and C. Tucker (2018). Algorithmic Discrimination? Apparent Algorithmic Bias in the Serving of STEM Ads. *Management Science*.

Miller, A. and C. Tucker (2017). Privacy protection, personalized medicine and genetic testing. *Management Science*.

Miller, A. and C. Tucker (2018). Historic patterns of racial oppression and algorithms. *Mimeo, MIT*.

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