

# Empirical Economic Research on Privacy Regulation: Lessons from the GDPR and Beyond

Garrett Johnson

Questrom School of Business  
Boston University

# Overview

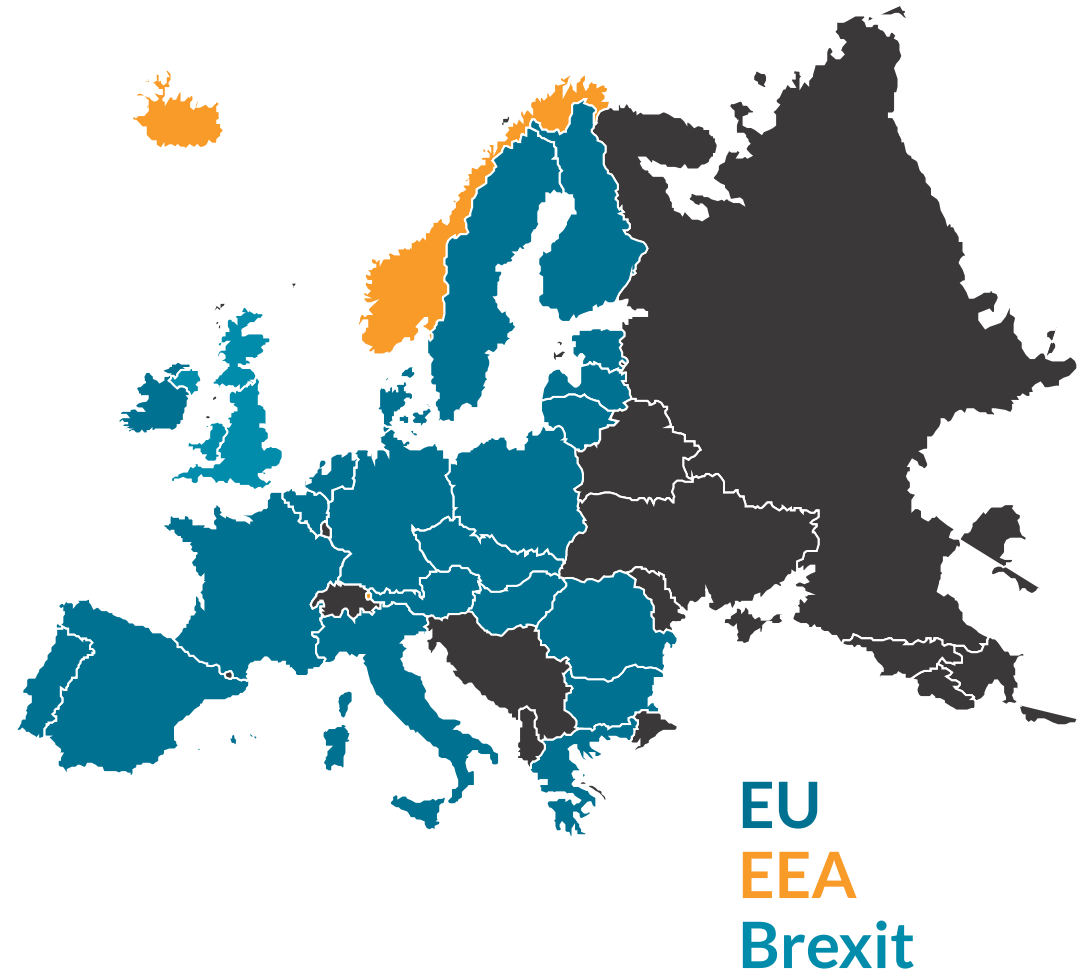
- GDPR primer
- GDPR research: challenges & solutions
- Literature survey
- Future directions



# GDPR Primer

# GDPR: Ambitious & comprehensive privacy policy

- **Personal data:** Pertains to unique individual (includes cookies IDs)
- **Enforcement deadline:** May 25, 2018
- **Geographic scope:**
  - EU firms & Non-EU firms that target EU residents
  - Enforcement split between country & EU-wide regulators
- **Global impact:** Informs global privacy regulation: US (California's CCPA), Brazil, Japan, South Korea



# Data minimization: GDPR key principle

Data minimization principle supported by many elements of GDPR

## Individual data rights

- Access data
- Correct data
- Data erasure (right to be forgotten)
- Data portability
- Object to processing
- Object to decisions based on automated processing

## Firm responsibilities

- Rights-based responsibilities
- Data audit
- Data minimization (protection by default)
- Encryption/pseudonymisation (protection by design)
- Breach notification (72 hrs)
- Data protection officer

**Takeaway:** ↑ logistical cost & legal risk of personal data processing

# GDPR: Legal bases for data processing

Regulator clarified websites should obtain consent for web technology purposes (EDPB 2020 & UK ICO 2019):

- **Individual consent:** Valid consent requires
  - **Affirmative:** No pre-checked boxes
  - **Transparency:** Granular to the purpose of processing & list all third parties who process the data
  - **Freely given,** clearly visible, & uses plain language
  - **Accountability:** Firm must be able to show proof of consent

But, GDPR is not only about consent! Other legal bases:

- Contractual obligation, legal obligation, public interest, or vital interest of an individual, claim legitimate interest (controversial!)



# GDPR research: Challenges & solutions

# Three research challenges & some solutions

1. Lack of clean control group
2. Low firm compliance with regulation
3. GDPR's impact on data observability



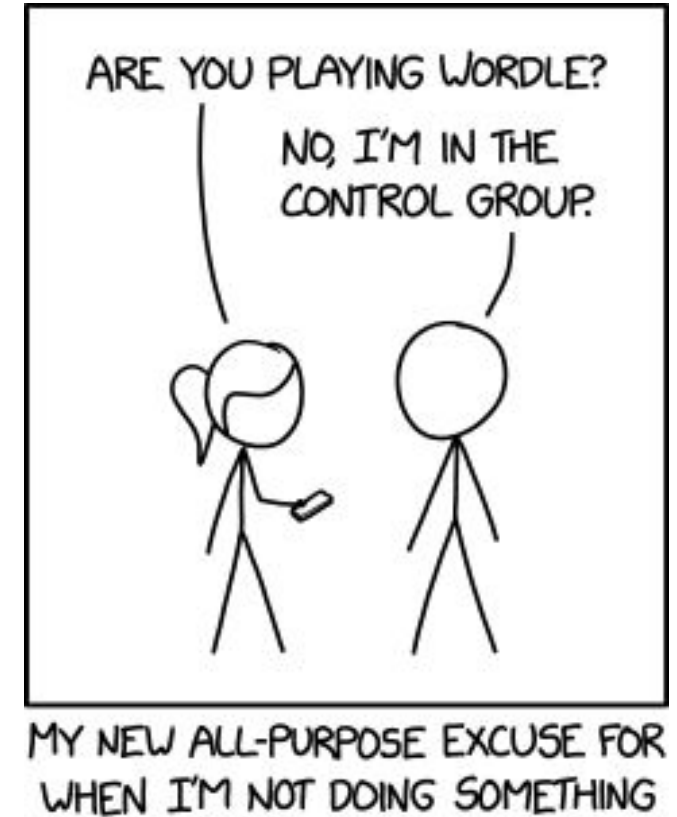


# 1) No good control: GDPR as an event study

Event studies should have a) good control group & b) clear start date

a) May lack clean control group because GDPR effects:

- 28 EU countries (+ 3 EC countries) including firms of all sizes
  - Note: Idiosyncratic economic shock to EU, post-GDPR would bias many studies
- Non-EU firms that target EU residents
- Substantial spillovers:
  - Global firms may roll out GDPR compliance globally due to cost efficiencies
  - "Brussels Effect": GDPR as an agenda-setting regulation

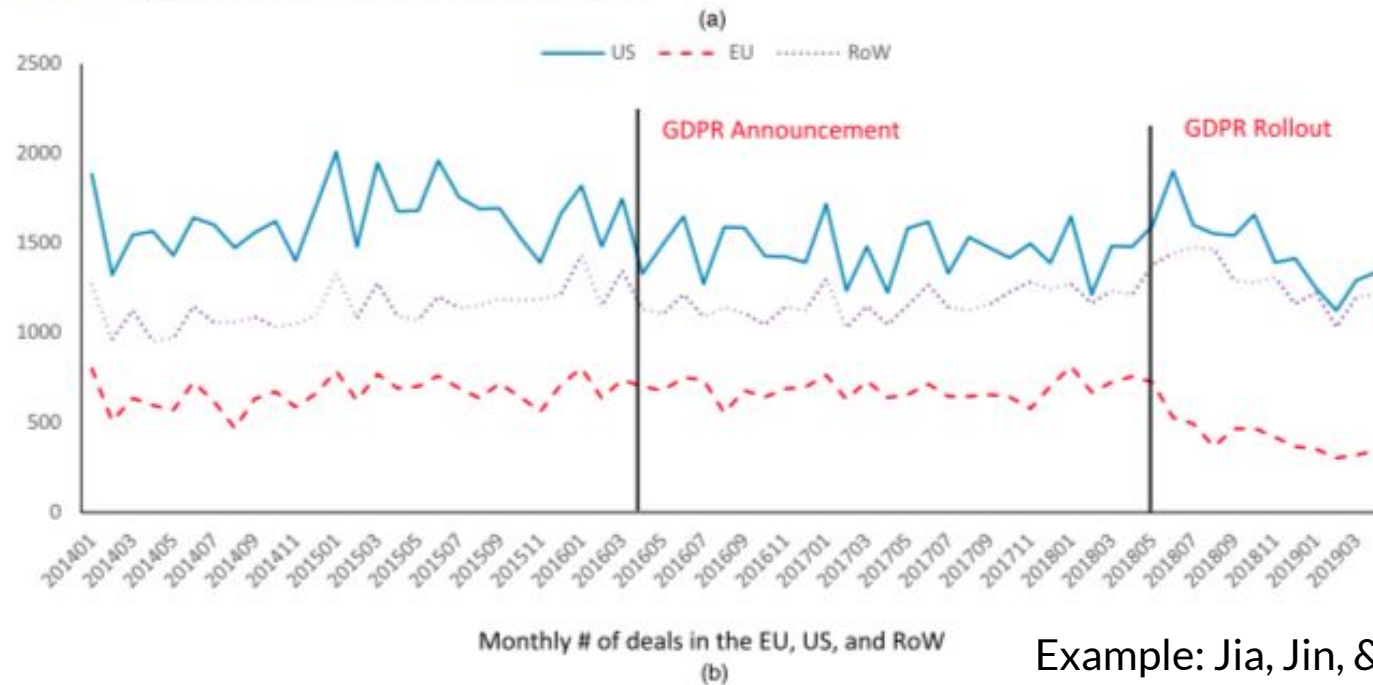


# 1) No good control: GDPR as an event study

b) Event(s): Most studies focus on enforcement deadline (May 25, 2018), but some also examine passage in April 2016

- What timing to use? E.g., compliance costs incurred before and after enforcement, but revenue effects occur after (perhaps)
- Anticipatory compliance may attenuate estimates

Figure 1. (Color online) Aggregate Monthly Trend Plots



# 1) Solutions: Identification approaches in the literature

- Difference-in-differences / synthetic controls / DiD+matching
  - Other countries as control: In some settings, GDPR spillovers may be small
- Panel differences: Diff-in-diff where previous year's units serve as control
- Before vs. after comparison
  - Large, sudden changes may evidence of GDPR's impact
  - Best suited to short-run comparisons unconfounded by long-run exogenous trends
- Exploit variation in degree of EU / GDPR exposure

## 2) Low regulatory compliance



*What do we learn about a law if compliance is low?*

## 2) Why might GDPR compliance be low?

In general, regulatory outcomes are a product of a game between firms & regulators where:

- Compliance is costly to firms
  - Small- & medium-sized firms in particular lack compliance resources
- Enforcement is costly to the regulator
  - Regulators vary in resources, strictness, tactics
  - 2019 GDPR status report emphasized under-enforcement (European Commission 2019)

## 2) GDPR compliance & enforcement are especially challenging

- Compliance may be subjective, difficult to observe & multidimensional
  - E.g., contrast with vehicle emissions standards
- In some sectors, cost of strict compliance > cost of fine
- GDPR is a "law of the whole economy"
  - Regulators necessarily must set enforcement priorities
- Debatably, firms can choose to locate in weak regulator country
- GDPR is complicated & compliance norms arise gradually
- Privacy regulators (unlike antitrust) lack enforcement experience

## 2) Solutions

- DON'T assume GDPR as written actually happens in practice
- DO examine the reality of GDPR "on the ground"
  - Do firms comply in my industry? How?
  - What are the economics of a firm's compliance decision?
  - What are related enforcement actions?
  - Do consumers make use of GDPR data rights?
- Some solutions in the literature:
  - Look where compliance most plausible and/or measurable
    - Note: Compliance can change over time
  - Look for differences by regulator strictness
  - Acknowledge limitations:
    - May be hard to generalize from short-run GDPR impact
    - Long-run GDPR impact may be confounded by exogenous trends

### 3) GDPR's impact on data observability

- GDPR limits personal data processing, which reduces the amount of data for firms and researchers
  - This poses challenges for many applied microeconomists that deliver rich insights from individual level data
  - But, no effect on accounting data or macroeconomic data
- Consent induces individual self-selection into personal data
- Identification challenge:

$$\text{Recorded outcomes} = \text{Real outcomes} * \text{Consent rate}$$

- If recorded data falls, how much is due to consent and how much is due to real economic effect?



# 3) Solutions

1. Non-personal data
2. Micro data from consenting panels (Zhao, Yildirim, & Chintagunta '21)
3. Make progress with personal data (with GDPR selection effect):
  - Direct impacts on data quality are themselves interesting (Aridor, Che, & Salz '20)
    - Post-GDPR, consented data for online users is positively selected
      - Lose users who were obfuscating their data anyways (i.e., cookie blockers)
      - Better quality customers are more likely to consent
    - Data quality improves (albeit with selection), facilitating user level predictions
  - Bound contributions of GDPR's real and consent effects on recorded data (Goldberg, Johnson, & Shriver '22)



# Literature survey\*

# Overview: Literature survey

- GDPR impact on:
  - firms
  - consumers
  - competition
  - web
  - innovation
  - marketing
- How does GDPR work in practice?

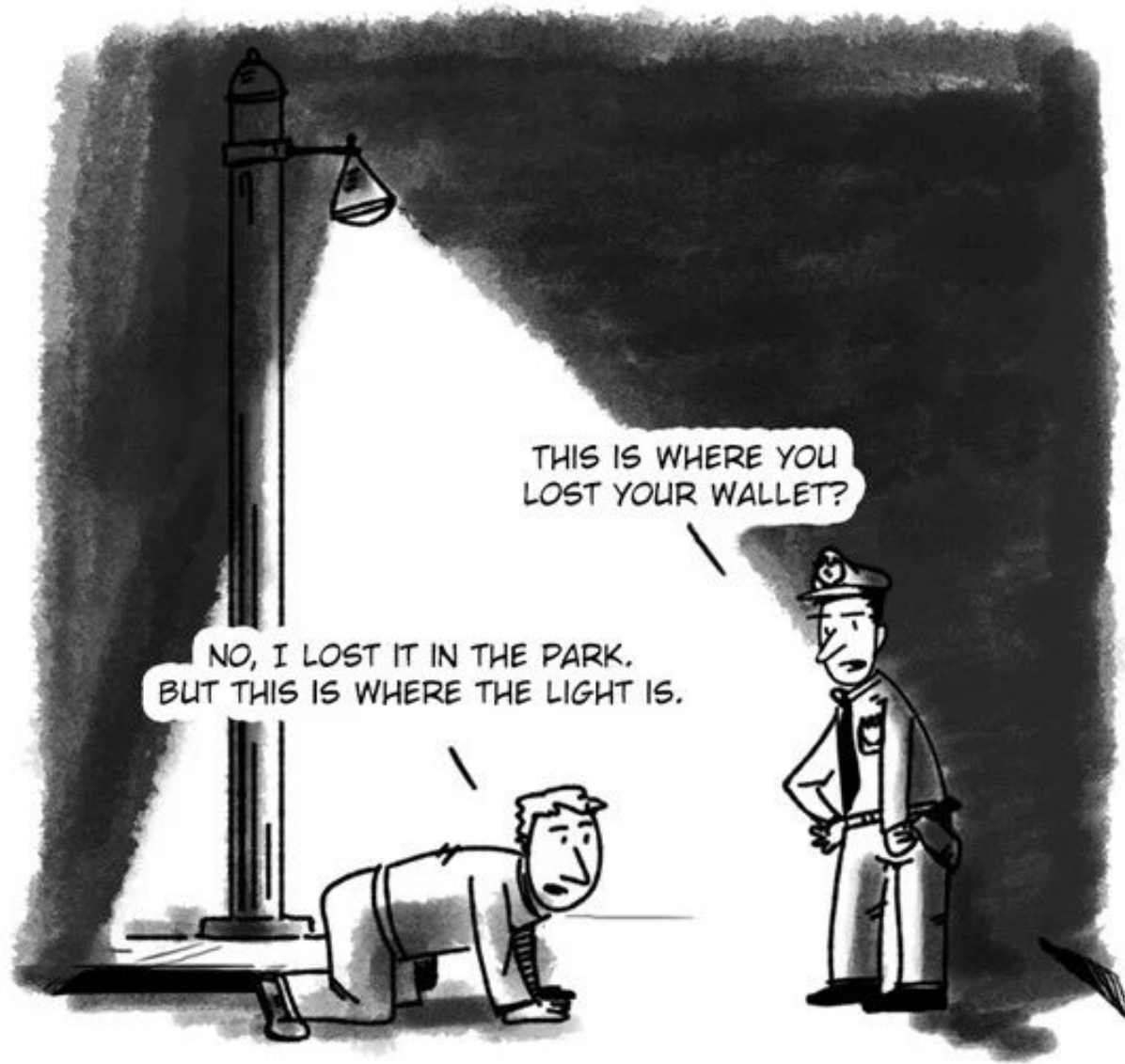


# GDPR impact on firms

- Accounting data shows lower **profit & revenue** (Yuan & Li 2019; Koski & Valmari 2020; Chen, Frey, & Presidente 2022) & higher **costs** (Bessen, et al. 2020)
  - Compliance was expensive (survey evidence):
    - Ernst & Young (2018): Cost \$7.8B for top 500 global firms
    - DataGrail (2019): 74% of small- and mid-sized organizations spent >\$100,000
- Reduced EU tech venture **investment** deals by 26% (Jia, Jin & Wagman 2021)
  - Most affected: Early-stage ventures, more data-related ventures, B2C (vs. B2B), healthcare & finance industries
  - Increase in investor **home bias** (Jia, Jin & Wagman 2019)
- Increased **exit** & reduced **entry** in app market (Janssen, Kesler, Kummer & Waldfogel 2021)

# GDPR impact on consumers

- Largely, a gap in the literature. Some exceptions:
  - GDPR hurts consumer surplus by reducing innovation in consumer products (Janssen et al. '22)
  - Theory (e.g., Ke & Sudhir '20, Wang, Xu, & Zhang '22)
- Survey research? Need pre-GDPR baseline
  - No increase in general awareness of privacy or perceived control over personal data (Presthus & Sørnum '21)
- Data protection improvements?
  - E.g., GDPR plausibly reduces number of data breaches; but, GDPR breach reporting requirement leads to increase in reported breaches!
  - Short-lived reduction in site vendor use or cookies (Johnson et al. 2022; Lukic et al. 2021; Peukert, et al. 2022)



# GDPR impact on competition

- Greater GDPR harms to **smaller firms** (Bessen et al. 2020; Chen, et al. 2022; Goldberg et al. 2022; Jia, et al. 2021; Koski & Valmari 2020; Zhao et al. 2022)
  - Smaller sites may have lower consent rates (Goldberg et al. 2022)
- GDPR increased **concentration** in web vendor market (Johnson et al. 2022; Peukert, et al. 2022)
  - To reduce data sharing, sites cut some vendor partners
  - But, sites favor keeping larger partners
  - Concentration does not arise from consent mechanism (Campbell et al. 2015)
- See also law & econ. literature (Gal & Aviv 2020; Geradin, Karanikioti & Katsifis 2020)

# Impact of the GDPR on the web

- Immediate reduction in **3rd party vendor** & cookie use (Libert, Graves & Nielsen 2018; Lefrere et al. 2020; Peukert, et al. 2022; Lukic, Miller & Skiera 2022; Johnson, Shriver & Goldberg 2022)
  - Most sites track users before obtaining consent (Sanchez-Rola et al. 2019)
- Small real drop in EU site **traffic/revenue** (Schmitt et al. 2020; Lefrere et al. 2020; Goldberg, Johnson & Shriver 2021)
- Share of observed site **analytics data** shrinks due to user non-consent (Aridor, Che & Salz 2020; Goldberg, Johnson & Shriver 2021)
  - Users who provide consent are (favorably) selected
- No GDPR effect at the Internet **interconnectivity layer** (Zhuo et al. 2019)
- User **search** intensity increases post-GDPR in user panel data (Zhao, Yildirim, & Chintagunta 2022)



# GDPR impact on innovation & marketing

## Innovation

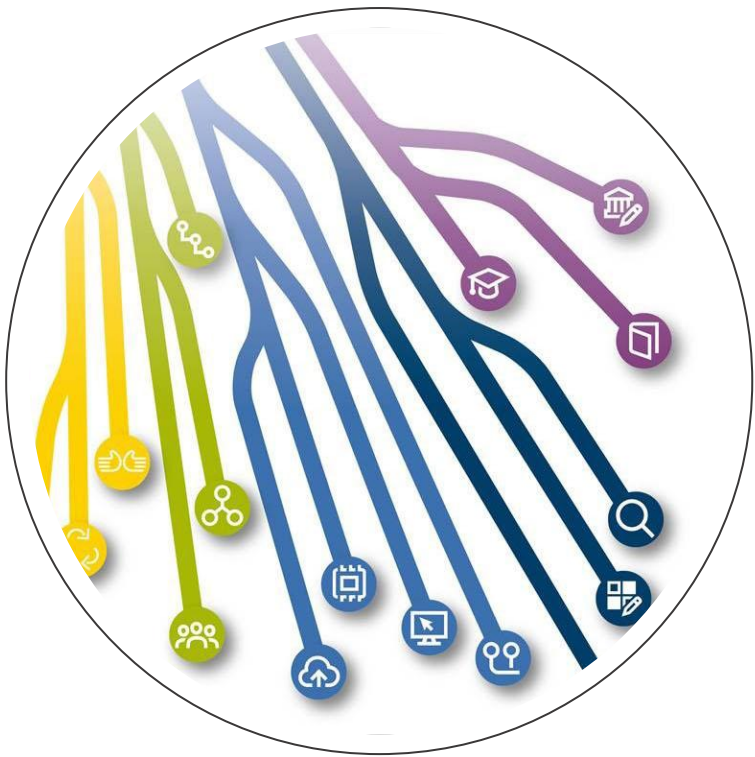
- Bessen et al. (2020) survey AI startups:
  - GDPR imposes costs by adding new position, reallocating resources, deleting data
  - Use of data protection means does not differ by whether firms have customers in Europe
- Reduced app entry (Janssen et al. 2021)

## Marketing

- Firms can run permissioning campaigns to gain consent to market to consumers and thereby increase sales (Godinho de Matos & Adjerid (2022))
- Larger recorded reductions in site traffic originating from email & display ad clicks (Goldberg et al. 2022)

# How does GDPR work in practice?

- GDPR effects moderated by **regulatory strictness** (Goldberg et al. 2022; Jia et al. 2019;2021; Johnson et al. 2022)
  - Limited regulator resources vary by country (EU DPB 2020)
- **International spillovers** (Peukert et al. '22)
  - Non-EU users see fewer vendors on EU sites
  - Non-EU users see fewer vendors on non-EU sites ("Brussels effect")
- **Penalty design:** For EU users, foreign sites with small share of EU users cut more vendors than EU sites (Johnson et al. 2022)
  - Foreign firms serving EU users may fear 4% fine on *global* revenue
- Differences by greater **reliance on data & sensitive data** (Jia et al. 2019; Yuan & Li 2019)
  - Firms with experience dealing with sensitive data may have lower GDPR compliance costs (Koski & Valmari '20)



# Future research

# Privacy regulations and changes on the horizon

- **GDPR:** enforcement deadlines, major cases, self-regulatory adjustments, private actions (e.g. NOYB warning sites)
- **EU:** Digital Services & Markets Acts, future ePrivacy Regulation
- **Privacy regulation** in other countries & US state laws
  - FTC proposed rulemaking on "commercial surveillance"
  - UK proposal to revisit GDPR
- **Tech platform changes** like Apple's App Tracking Transparency

*But, beware of "Lucy pulling the football"!!*



# Privacy regulation: Benefits & improvements

- To date, have we seen more research on the *unintended* consequences than the *intended* consequences of GDPR?
  - What are privacy benefits to consumers and how do we quantify them?
  - What are compliance gains in improving data protection?
- How to enforce GDPR effectively?
  - We have seen that compliance game between regulators & firms is complex & interesting!
  - What are effective enforcement strategies (e.g., fines, notices, choice of targets, establishing precedent)?
- How do you design better privacy regulation? How do you limit unintended consequences?

# Privacy-enhancing technologies (PETs)

PETs are "technologies that embody fundamental data protection principles by minimising personal data use, maximising data security, and empowering individuals." (ICO '22)

- Example PETs: Differential privacy, federated learning, on-device computation, zero-knowledge proof, secure multi-party computation
- Practical examples: US Census, Google's Privacy Sandbox
- So far, privacy regulations ignore these developments



# PETs research needs

## Fundamental research:

- PET design
- Mapping out privacy versus value frontier
- Challenges for inference (Nekipelov & Komarova '21)

## Key economic questions:

- PET adoption as innovation
- Competitive consequences
- Economic impact assessment
  - Firm & market adaptation
  - Check out: "Privacy-Centric Digital Advertising: Implications for Research" (Johnson, Runge, Seufert '22)

