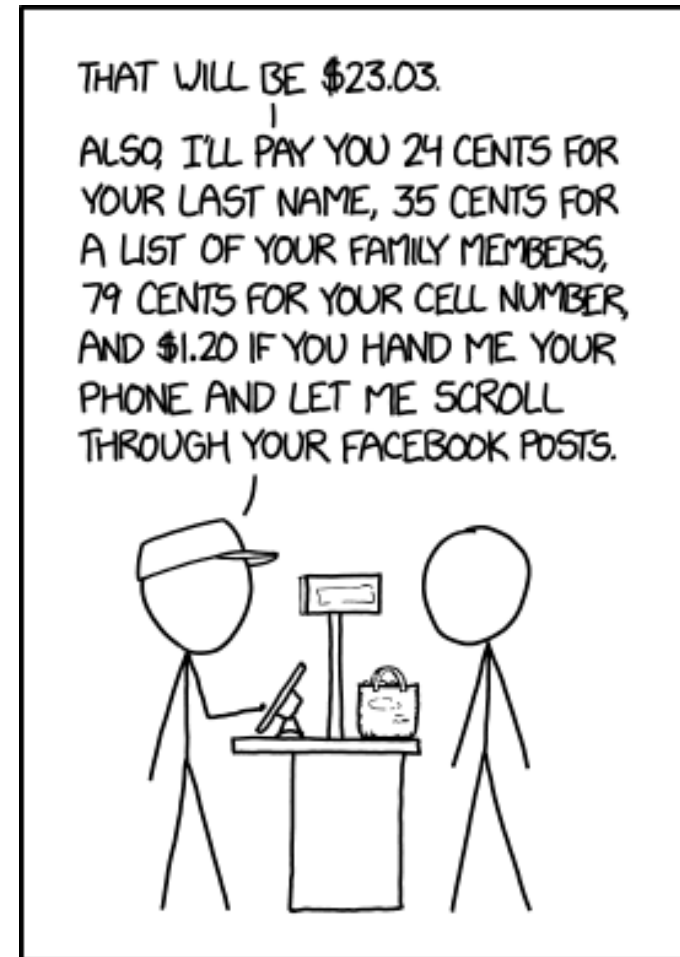


Markets and incentives for transacting on (personal) information

Katrina Ligett
Hebrew University of Jerusalem



LOYALTY CARDS AND REWARDS ACCOUNT OFFERS GET WAY WEIRDER IF YOU THINK OF THEM AS SEPARATE TRANSACTIONS.

Why study privacy?

- Increases individual/social welfare
 - reducing discrimination
 - increasing access to valuable data
 - protecting human rights
 - etc.

Economics: tools for study of welfare

For privacy intervention to maximize welfare, need to understand

- value provided to **individual** from privacy intervention
- value of **other impacts** of privacy intervention—increased/reduced access to data? changes in social norms? increased sense of control? increased saliency of privacy concerns?
- how individuals and groups will behave **in response** to privacy intervention
- ...



Why all this talk about value and money?

- Assume all types of value can be compared/exchanged
- Single dimension of comparison - simplifying
- Isn't this problematic? Makes data a commodity?
Validates/normalizes privacy losses? (Maybe. More about this later.)

Caveats

- I'm not an economist
- This is not a survey, more of a tasting menu

Talk Outline

- Information in economics
- (Private) data as a good
- Value to individuals
- Effects on society
- Other lessons from economics

Talk Outline

- **Information in economics**
- (Private) data as a good
- Value to individuals
- Effects on society
- Other lessons from economics

**excellent survey article:
“The Economics of Privacy”
Acquisti, Taylor, Wagman (2016)**

“The Economics of Information”

Stigler (1961)

- In non-centralized markets, takes effort to find out prices
- Searching more leads to finding find lower prices
- Natural that specialized traders (used car dealers) exist to reduce search costs



“The market for lemons”



Akerlof (1970)

- Information asymmetry can reduce quality of goods in a market
- Model: buyer has uncertainty about quality of good; seller does not
- Buyers thus only willing to pay average price between peach and lemon
- Sellers of peaches leave the market - “adverse selection”
- Own observation: may see related effects in elective “opt-in” privacy protections e.g., do not call lists

“Job market signaling”

Spence (1973)

- University degree is a costly signal
- Signal could be useful to employers not because makes you more productive worker, but because positively correlated with greater ability
- Aggregate cost of signaling activity may outweigh benefits



“The economics of privacy”

Posner (1981)

- Concealment of personal characteristics makes markets (employment, marriage, etc.) less efficient. Predicts that, e.g., laws making credit history private will result in increased interest rates. Is pretty non-PC at best
- Argues some forms of privacy protections have economic benefit— increasing value of information, preventing false information, making communication more effective without someone eavesdropping
- Counterarguments: sometimes society benefits when individuals can select optimal level of an embarrassing activity (e.g., drug treatment)
- Recent related point (Bushway (2004) and Strahilevitz (2008)): when employers without access to information may rely more on statistical discrimination strategies

“Economic aspects of personal privacy”

Varian (1997)

- Consumer may not wish for willingness-to-pay to be known
- But may wish for some information to be known (e.g., interest in offers for particular product category)
- Raises concerns about secondary uses, resale



Property rights for information

- Discussed by Laudon (1997), Litman (2000), Samuelson (2000), Schwartz (2004)
- Could benefit both individuals and users of data—proper market-based compensation
- Under simple model, a monopolist offering personal vs. fixed pricing may get all users' data for free, because marginal anonymous consumer makes no surplus

Price discrimination

- Large literature, e.g. Fudenberg and Tirole (1998)
- Some models (Taylor (2004), Acquisti and Varian (2005)) suggest that consumers “only” need regulatory protection if they are naive about how information about them can be used
- Council of Economic Advisers (2015) concludes bad for consumers

Who benefits from information?



- Gehrig and Stenbacka (2007): information sharing between lenders reduces switching costs (sounds good for consumers!)
- Reduces need to compete for initial customer choice, reducing the welfare of borrowers.

Information-sharing between competing firms

- Raith (1996) summarizes literature on incentives of firms to share information
- Bergemann and Morris (2013) study information design in this setting

Who benefits from information?

- Board and Lu (2015): when consumers are anonymous, sellers may present all goods
- When consumers are tracked, they are steered to the most profitable products (implicit collusion between sellers)

“Markets for Information”

Bergemann and Bonatti, Annual Review of Economics, 2019

- With single firm purchasing data, get price discrimination, known (Schmalensee (1981)) to lower consumer welfare and total welfare
- Demand uncertainty and informativeness of the information dictate how many consumers are needed to make data intermediary profitable
- “The optimal information policy for a data intermediary remains a wide open question.”
- Noise lowers value of information, but also cost of obtaining it; may benefit intermediary
- Admati and Pfleiderer (1986): possible dilution in the value of information due to its leakage through informative prices

Auctions and mechanism design for information

- Sponsored search as an example (Edelman, Ostrovsky, and Schwartz (2007) and Varian (2007))
- Information design for info used to reduce risk (Bergemann, Bonatti, and Smolin (2018))

Additional issues

- Sale of consumer scores as a form of aggregated data
- Use of selective release of data to manipulate recipient into “exploring” in explore-exploit settings (Kremer, Mansour, and Perry (2014))

Talk Outline

- Information in economics
- **(Private) data as a good**
- Value to individuals
- Effects on society
- Other lessons from economics

(Private) data as a good

- How is data different from sneakers?
- What are the goods of data?
- How are other goods different from data?

A good in economics is any object or product that is useful. - Wikipedia

A service is a non-material good.

Data vs. Sneakers

nearly free to make additional copies	production costs
difficult to control resale	limits to resale
difficult to find right buyer; markets not very functional	effective markets
degrade with use (statistical concerns)	degrade with use
value depends on who else has it	(less so)
curve of marginal value not well understood?	decreasing marginal value
value may be sensitive [c.f. Frauke's talk]	value presumed not sensitive
can be re-purposed in surprising ways	use pretty clear

Privacy: What is the good?

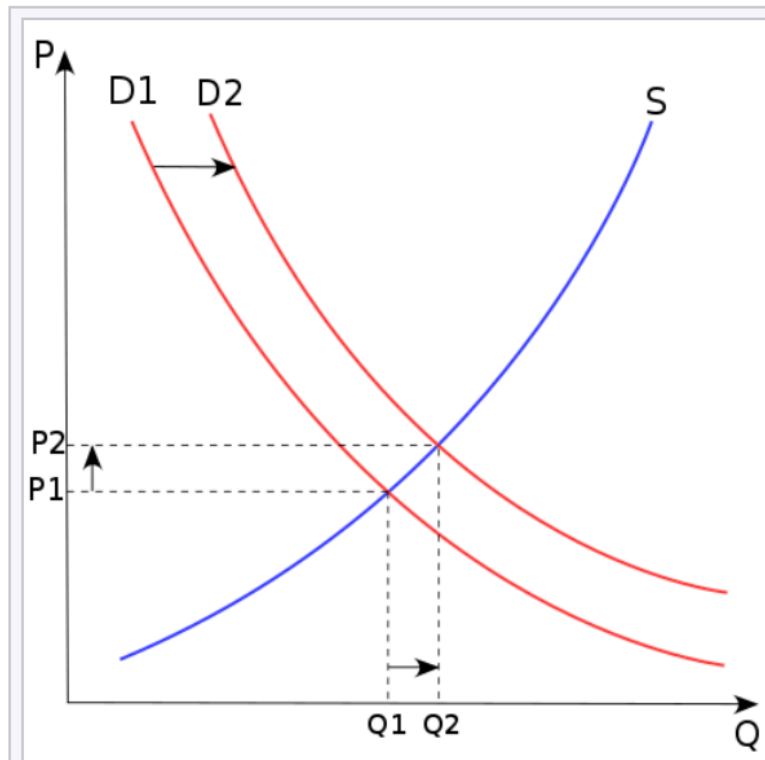
- A record?
- An aggregate statistic?
- A synthetic dataset?
- A privacy “service”?
- The ability to serve information to a targeted audience?
- ...



What should a data price cover?

- Value of its current and future **uses** (and resale?)?
- Compensation for loss of **intrinsic right**?
- **Insurance** against future harms that might result?

Supply and demand



The price P of a product is determined by a balance between production at each price (supply S) and the desires of those with purchasing power at each price (demand D). The diagram shows a positive shift in demand from D_1 to D_2 , resulting in an increase in price (P) and quantity sold (Q) of the product.

Talk Outline

- Information in economics
- (Private) data as a good
- **Value to individuals**
- Effects on society
- Other lessons from economics

Revealed preferences



Artist Risa Puno collected photos, addresses, driver's license numbers, phone numbers, mother's maiden name, fingerprints, social security numbers in exchange for a cookie.

Experimental and data-based evidence

- Many studies. E.g., Savage and Waldman (2013) willingness to pay for privacy:

\$2.28	browser history
\$4.05	contacts list
\$1.19	location
\$3.58	text message contents
\$2.12	no ads

“Privacy and Mechanism Design”

Roth and Pai

SigEcomExchanges 2013

Definition *A mechanism $M : \mathcal{T}^n \rightarrow \mathcal{O}$ is ϵ -differentially private if for all pairs of neighboring type vectors $t, t' \in \mathcal{T}^n$, and for all functions $u : \mathcal{O} \rightarrow \mathbb{R}^+$:*

$$\mathbb{E}_{o \sim M(t)}[u(o)] \leq \exp(\epsilon) \mathbb{E}_{o \sim M(t')}[u(o)].$$

Definition $M : \mathcal{T}^n \rightarrow \mathcal{O}$ is ϵ -approximately dominant strategy truthful if for every player i , for every $t_{-i} \in \mathcal{T}^{n-1}$, and for every $t' \in \mathcal{T}$:

$$\mathbb{E}_{o \sim M(t_i, t_{-i})}[u_i(o)] \geq \mathbb{E}_{o \sim M(t'_i, t_{-i})}[u_i(o)] - \epsilon$$

Proposition *If a mechanism M is ϵ -differentially private, then M is also 2ϵ -approximately dominant strategy truthful.*

DP as a tool for mechanism design

- Good news:
 - Composition! (Usual strategy-proof mechanisms need not compose)
 - Mechanism design without money!
- Bad news:
 - Approximate truthfulness
 - *Any* report is approximately dominant strategy [NissimSmorodinskyTennenholtz12]

DP for mechanism design

- Digital goods auctions [McSherryTalwar07]
- Equilibrium selection mechanisms
[KearnsPaiRothUllman12,CummingsKearnsRothWu15]
- Joint differential privacy
- Design of exactly truthful mechanisms
[NissimSmorodinskyTennenholtz12]

Eliciting private data: Buying Private Data With Verification

- [GhoshRoth11] introduced problem of buying private data
- [GR11]: truthful auctions to get accurate statistics when individuals don't care about privacy of their costs
- [GR11, NissimVadhanXiao14]: strong impossibility result for individually rational mechanisms when the costs themselves are private

Responding to impossibility

- [FleischerLyu12]: c_i drawn from known prior given b_i ; relies on knowing prior exactly
- [LigettRoth12]: take-it-or-leave-it offers (lose individual rationality); revised model of privacy costs
- [NVX14]: monotonicity of correlation between bits and costs; known bound on how many players' costs exceed a given threshold
- [GhoshLigettRothSchoenebeck14] Bayesian setting, but privacy not reliant on prior's correctness

Measuring the costs of privacy

- [GR11]: Linear function of epsilon?
- [NOS12]: Linear function of epsilon as upper bound
- [LR12]: Any privacy function that's a deterministic function of epsilon leads to problematic predictions
- More sophisticated proposal of [Chen et al. 13] model losses from realized outcome rather than worst-case
- [LiLiMiklauSuciu12] Model for pricing private data; consideration of arbitrage opportunities

Impacts of privacy concerns

- [CummingsIoannidisLigett 15]: What if sensitive data is input to a computation? Example: linear regression. How to elicit participation despite biased private estimator?
- [CummingsPennockWortmanVaughan16]: Privacy in one-shot vs. dynamic prediction markets

Design of a data market

- [CummingsLigettRothWuZiani15]: How to aggregate a simple statistic under a variance constraint
- So much left open!

Two econ lessons on how to design experiments

- Don't lie to your subjects
- Incentivized choice

Privacy Decision Making

[Cummings, Dekel, Heffetz, Ligett]

- Key idea: Create sensitive data in the lab
- Setup: groups of people play public good game

Therefore, your earnings from this task will be:

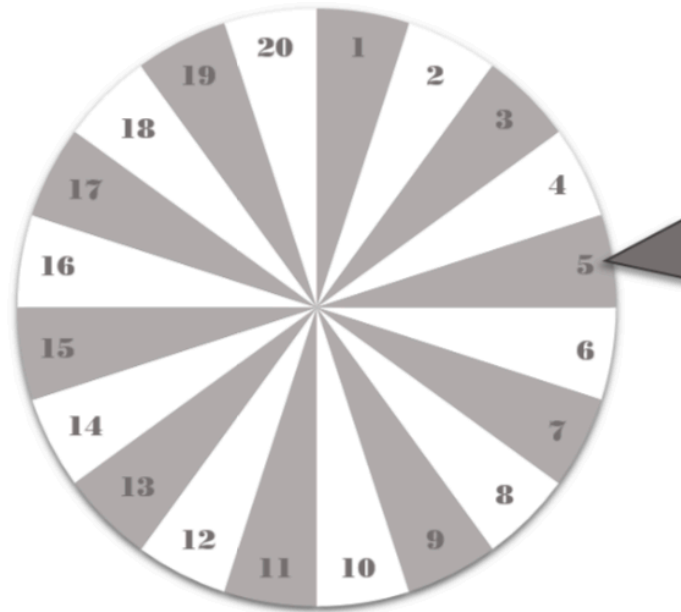
The number of dollars you allocate to the personal account

+ y times the number of dollars you allocate to the group account

+ r times the number of dollars all other $n - 1$ group members allocate to the group account.

You will make your allocation decision in private, and will receive no feedback until the very end of the experiment. However, if this task is selected at the end of the experiment, an announcement will be made about each group member's chosen allocation. This announcement may or may not be the same as the group member's true allocation, and will be determined as follows:

Each of you will be asked to spin a virtual roulette wheel like this:



- If your spin result is one of the following:

① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩

then your **true** allocation will be announced.

- However, if your spin result is one of the following:

⑪ ⑫ ⑬ ⑭ ⑮ ⑯ ⑰ ⑱ ⑲ ⑳

then a **random** allocation will be announced instead of you your true allocation. This random allocation will be determined by asking you to roll a virtual 11-sided die numbered 0-10. The result of this die roll will be your announced allocation to the group account. For example, if the result of the die roll is 5, then your announced allocation to the group account will be \$5.

Privacy Decision Making

[Cummings, Dekel, Heffetz, Ligett]

- Participants have the option to trade money for having less-embarrassing data
 - Could instead try to design experiment where people pay for privacy, but seems difficult if don't want to lie to subjects
- What will the contribution vs. epsilon curve look like?
- What does theory predict? (Depends how you model what goes into behavior—how do others' beliefs about me affect my actions? How do others form beliefs about me? What's the role of epsilon?)

Talk Outline

- Information in economics
- (Private) data as a good
- Value to individuals
- **Effects on society**
- Other lessons from economics

More privacy => need more samples

- More samples can be prohibitively expensive and time-consuming [c.f. Frauke's talk]
- Relatively easy to estimate cost—we should do this!

The privacy regulation hurdle

- Privacy experts are expensive (?)
- Takes time and money to get privacy right (e.g., Rubin and Lenard (2001))
- Favors large companies with
 - big legal departments
 - big budgets for privacy tech

Philosophical/legal/moral benefits of privacy

- Seem difficult to quantify.
- What might an economist do?
 - Survey to elicit the importance of living in a privacy-just world for, e.g., happiness
- Hope privacy “wins” even without this factor.

Privacy protections prevent harms

- Databases that don't exist can't be hacked
- Trying to protect yourself as an individual is costly
- Insurance against privacy harms (identity theft)

More data, better data, more access

- More and better data captured?
- Broader (industry, society) access to data?
- Greater competition, interoperability can drive economic growth
- Privacy reduces complexity in interactions (cf. Milberg et al., 2000)

Privacy => Statistical validity

- Enable more data sharing, reuse
- Fewer resources wasted as a result of wrong science, statistics
- Increase *trust* in science, government statistics
- Can we ever argue that the cost of additional samples is offset by protections of statistical validity?

CummingsLigettPaiRoth16



Privacy policy:

flip w.p. 1-q



NOT BUY



offer A

Posted price

p

BUY!

Main Question:

How does equilibrium change as we vary privacy policy?

Not obvious!



offer B

Equilibrium analysis

As accuracy of the signal is decreased, we can see:

- more information about the consumer is revealed
- consumer utility decrease
- lender utility increase
- player utilities non-monotone, discontinuous in privacy parameter
- multiplicity/non-existence of equilibria

Talk Outline

- Information in economics
- (Private) data as a good
- Value to individuals
- Effects on society
- **Other lessons from economics**

Other lessons from economics

- Brandimarte et al. (2013) : greater perceived control can result in greater privacy risk-taking
- The way you present information has substantial effects on behavior (Adjerid et al., 2013)
- Salience bias (Kahneman et al. 1982)—do privacy protections have a cost by making us “feel” risks more?
- Gneezy and Rustichini (2000) “A fine is a price” - introducing late fine increased lateness
- Adverse selection with compensated surveys: may induce non-truthful reporting behavior or select an unfavorable segments of the population

“Evite.com may sell lists of consumers attending a party in a given location” - Bergemann and Bonatti 2018

.

only fifty cents, huh? well,
i guess the real profit's
made when you COLLECT
MY DATA AND SELL IT
TO THE GOVERNMENT!

