Privacy on the Internet

Ole Jann

Talking Economics

February 14, 2020

The Trade-offs of Privacy

- People voluntarily provide lots of information via their Smartphones, Facebook, Twitter, Google ...
- Governments systematically collect internet traffic data, metadata, read emails, can access webcams, ...
- People give up information for better prices (loyalty programs, insurance contracts)
- Some countries collect (and connect) lots of administrative data
- $\Rightarrow$ What is at stake? Is all this good for society? Should it be regulated?
- $\Rightarrow$ Can we weigh the costs and benefits of privacy? Does privacy only benefit the individual, or also society?
- (Here I mean: Privacy $\subset$ asymmetric information)

Arguments about Privacy

- The Chicago School:
  - Exemplified by George Stigler (1980, JLS) and Richard Posner (1981, AER)
  - Asymmetric information leads to inefficiencies
  - Therefore: privacy is inefficient

Arguments about Privacy

- The Chicago School:
  - Exemplified by George Stigler (1980, JLS) and Richard Posner (1981, AER)
  - Asymmetric information leads to inefficiencies
  - Therefore: privacy is inefficient
- Popular debate: the "nothing to hide" argument
  - "If you have nothing to hide, you have nothing to fear (from giving up privacy)!"
  - Eric Schmidt, CEO of Google:

    "If you do something that you don't want anyone to know, maybe you shouldn't be doing it in the first place."

What my research has done

- Developed a general theoretical model in which we can weigh the gains and losses of privacy
- Not assume “taste for privacy” – instead, we derive the individual and social informational effects of privacy
- Main result: Privacy can be efficient even when considering *informational effects only*
- We can give conditions for when privacy is efficient (and when not)

An example

- Alice thinks drugs should be legalized and wants to write on her Facebook profile about that
- She is also looking for a job
- Bob (an employer) does not want to hire drug users – but drug use is not observable
- There is a correlation between opinion on legalization and drug use

An example

- Alice thinks drugs should be legalized and wants to write on her Facebook profile about that
- She is also looking for a job
- Bob (an employer) does not want to hire drug users – but drug use is not observable
- There is a correlation between opinion on legalization and drug use
- $\Rightarrow$ Bob may look at Alice's Facebook profile and decide not to hire her (Acquisti and Fong 2015: Employers discriminate based on social media)
- $\Rightarrow$ But if Alice anticipates that and *doesn't* post on Facebook:
  Who gains and who loses? Is the world better if Bob cannot spy on Alice on Facebook?

Who wins, who loses when Alice has no privacy?

- **Alice loses:** She cannot speak her mind
- **Society loses:** Alice's opinion (and that of others) is missing from public debate (= information aggregation)
- **Bob wins:** He learns something about job applicants

Who wins, who loses when Alice has no privacy?

- **Alice loses:** She cannot speak her mind
- **Society loses:** Alice's opinion (and that of others) is missing from public debate (= information aggregation)
- **Bob wins:** He learns something about job applicants
- But *how much* does he win?

Who wins, who loses when Alice has no privacy?

- **Alice loses:** She cannot speak her mind
- **Society loses:** Alice's opinion (and that of others) is missing from public debate (= information aggregation)
- **Bob wins:** He learns something about job applicants
- But *how much* does he win?
- Consider: If Alice changes her behavior more, she is worse off ...
- ... and society is worse off, ...

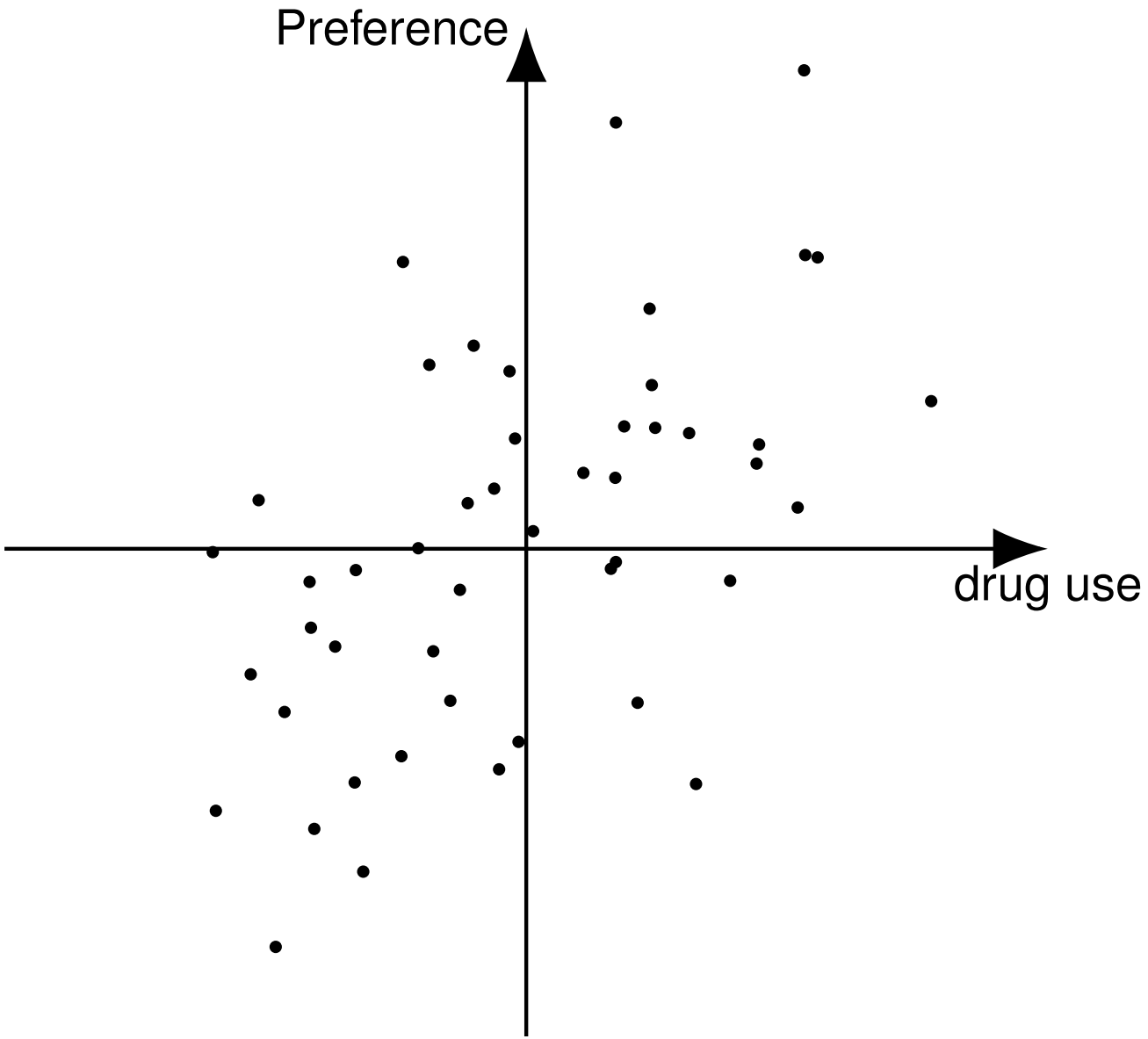
Who wins, who loses when Alice has no privacy?

- **Alice loses:** She cannot speak her mind
- **Society loses:** Alice's opinion (and that of others) is missing from public debate (= information aggregation)
- **Bob wins:** He learns something about job applicants
- But *how much* does he win?
- Consider: If Alice changes her behavior more, she is worse off ...
- ... and society is worse off, ...
- ... and Bob learns less as well!
- $\Rightarrow$ Sometimes, all welfare effects point in the same direction

A theoretical model

- *n* individuals
- Each individual *i* has a preference and a hidden type (“drug use”)
- the first is correlated with the second
- There is an opposing player (“employer”) who gains from treating people differently based on drug use

Correlation between preference and drug use



Correlation and machine learning

Hello [redacted],

We thought you might be interested in knowing that customers who bought "200g*0.01g Mini Digital Pocket Scale for Jewelry Kitchen Gram Oz Ct" also bought these items.

Customers Also Bought...

100 Plastic Resealable Grip Seal Bags...

Sold by: AGOODBUYFROMME

100 Grip Seal Bags 2.25 x 3 Inch...

Sold by: Express Goods UK

250 Raw Filter TIPS card booklets...

Sold by: Martins Deals

S AND S Â£20 Pound Note Design...

Sold by: BARGAIN BASE

100 Grip Seal Bags 1.5" x 2.5"

Sold by: Swoosh Supplies

JUICY Juicy Jays Kingsize...

Sold by: Premier Life Store

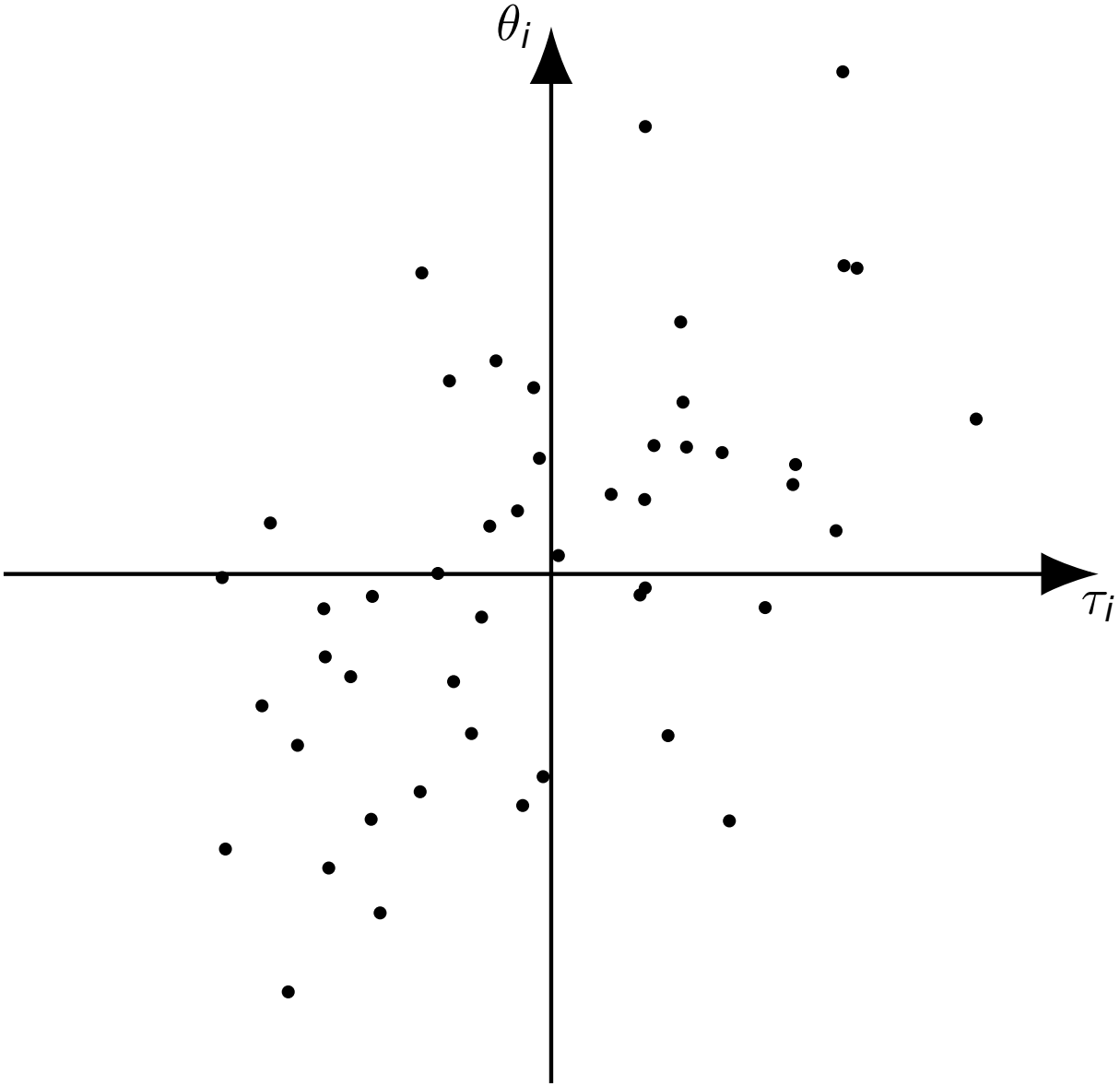
S AND S 750 Roaches Roach Filter...

Sold by: BargainShop_London

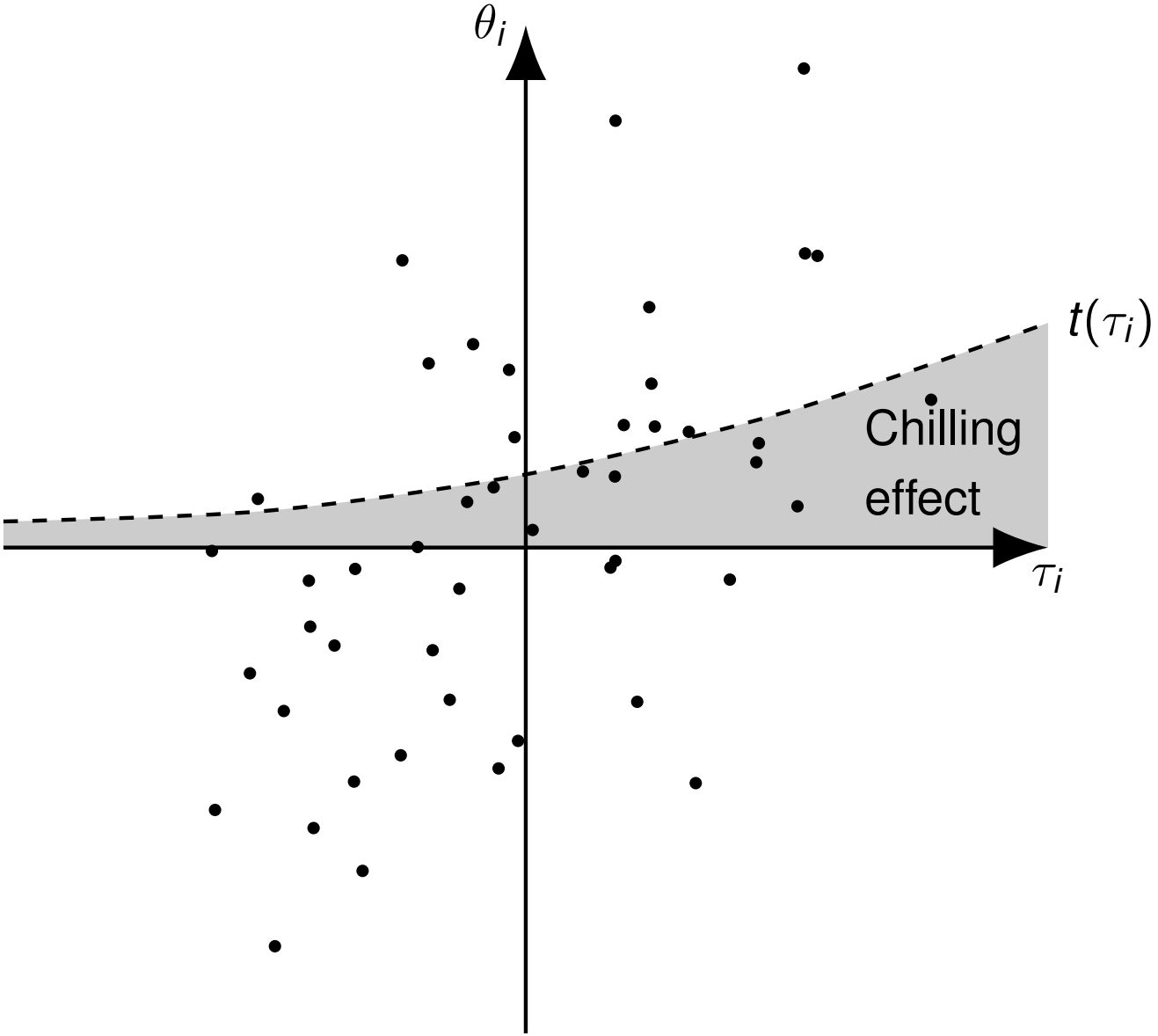
Timing

- Two stages:
  1. Information aggregation stage: Each individual chooses a policy based on their preference. A policy is implemented if the majority support it
  2. Interaction stage: An employer chooses, for each individual, whether to hire them or not
- Privacy: Can employer see policy choice when deciding how to treat individual i?

The Chilling effect: Graph



The Chilling effect: Graph



Welfare Results

- Welfare = Welfare of individuals + payoff of employer
  ... with arbitrary weights!

Result I:

Privacy is ex ante Pareto superior if either

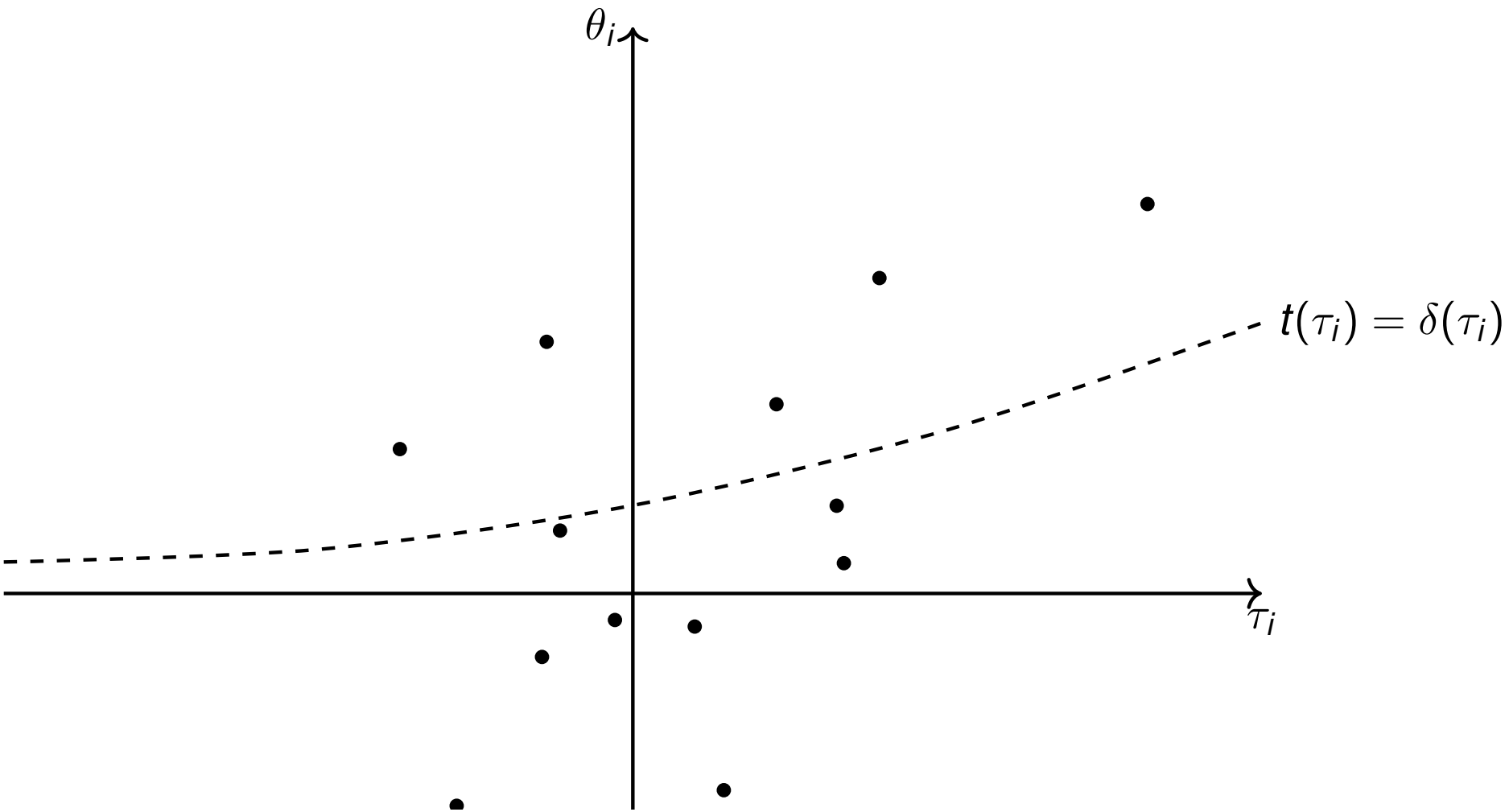
- there are many individuals (each individual has little influence),
- the cost of not being hired is large (loss of privacy induces large behavior change).

Result II

- Privacy is efficient unless the correlation between preference and drug use is sufficiently high

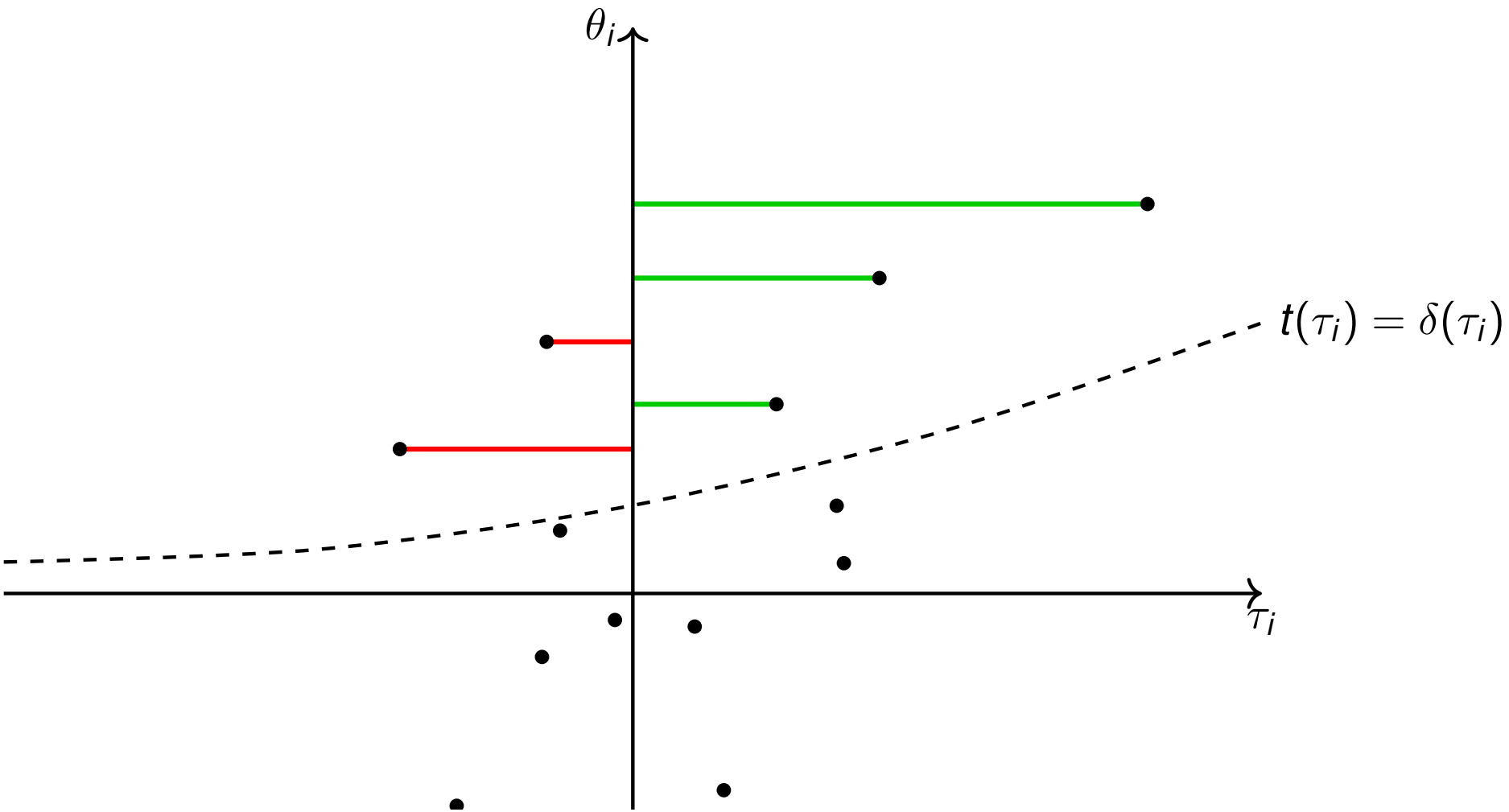
Gains and losses from privacy (intuition for $n = 1$)

- Gains (green) and losses (red) from taking away privacy
- horizontal = employer, vertical = individuals



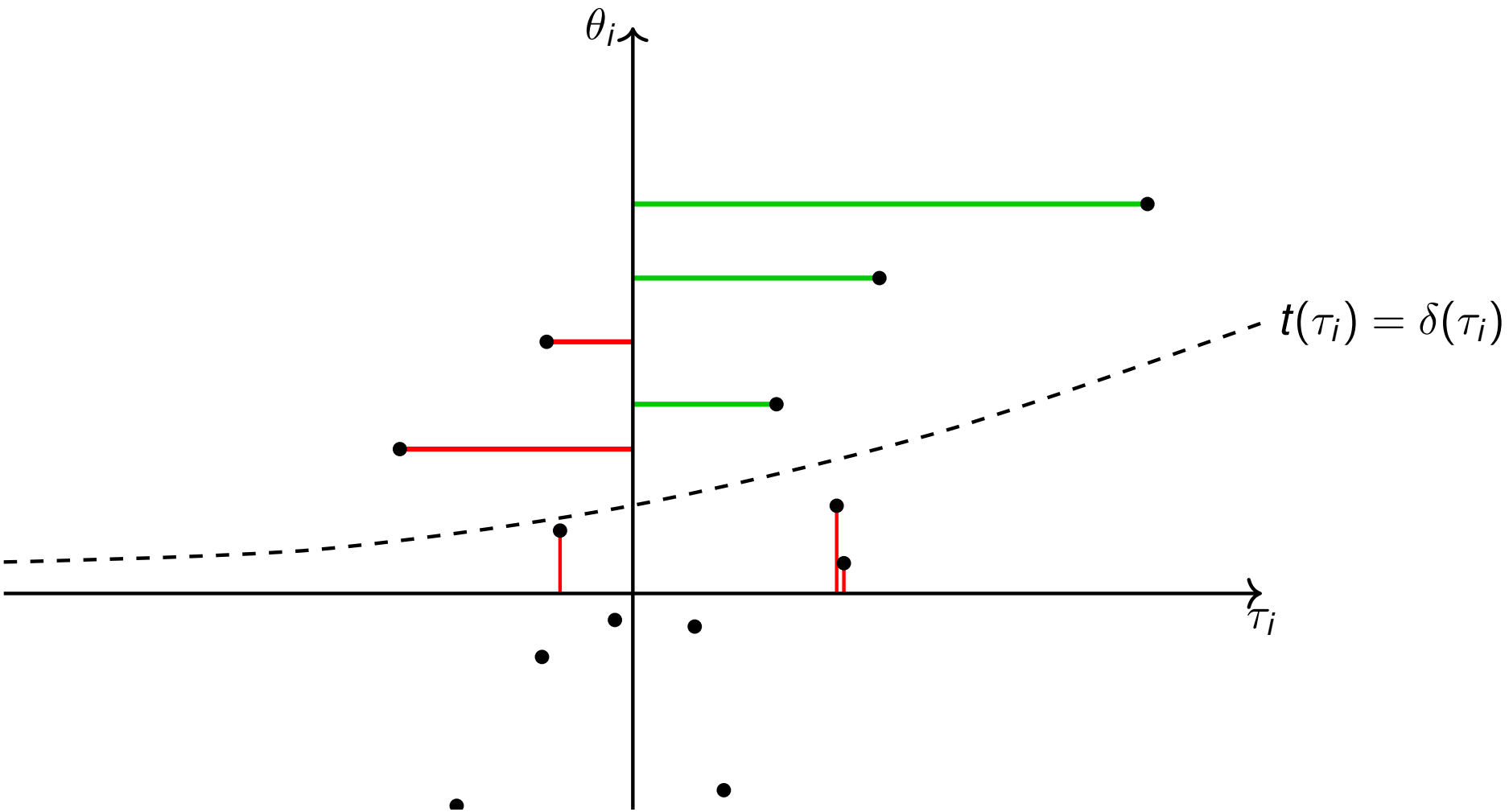
Gains and losses from privacy (intuition for $n = 1$)

- Gains (green) and losses (red) from taking away privacy
- horizontal = employer, vertical = individuals



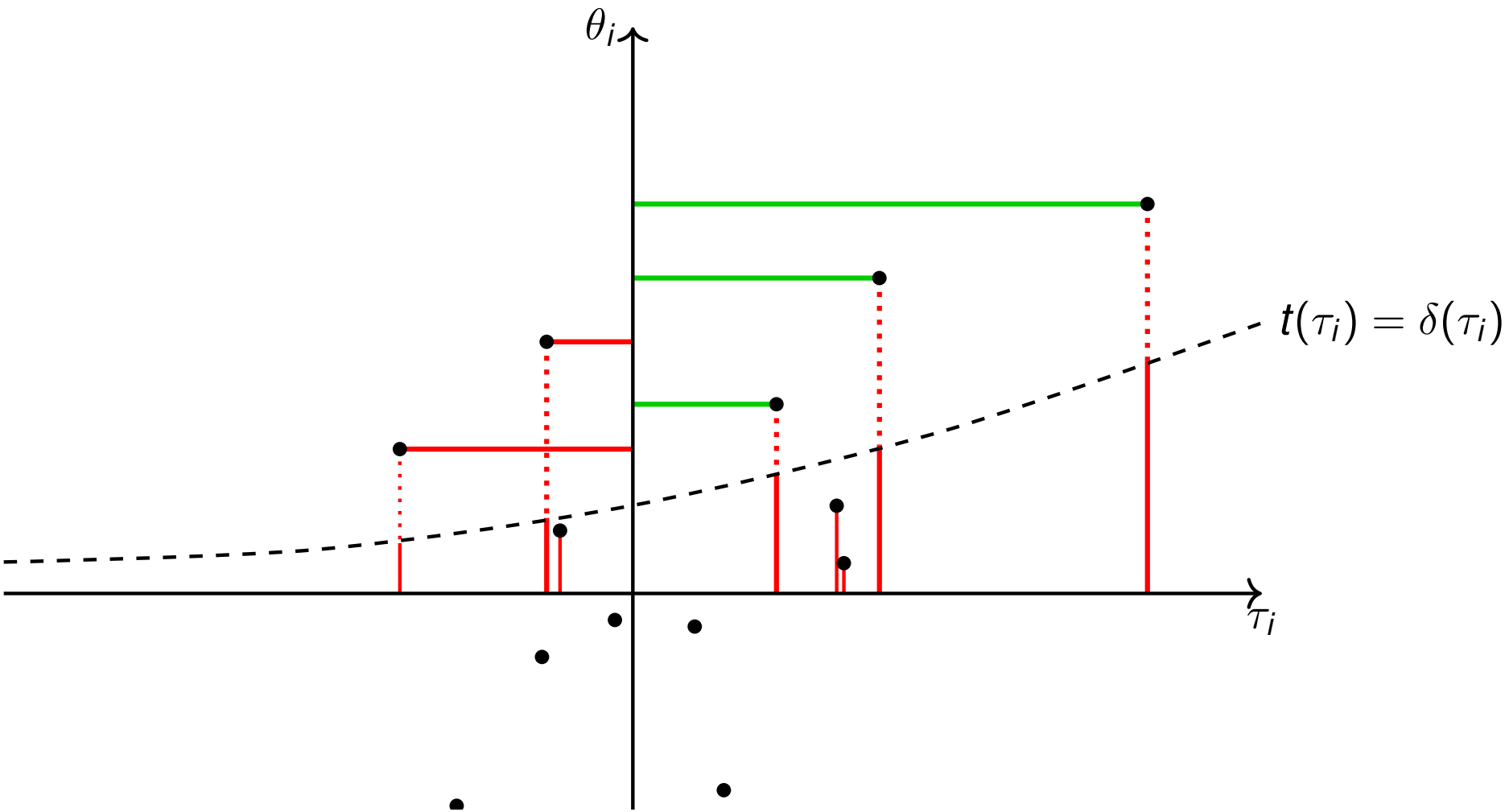
Gains and losses from privacy (intuition for $n = 1$)

- Gains (green) and losses (red) from taking away privacy
- horizontal = employer, vertical = individuals

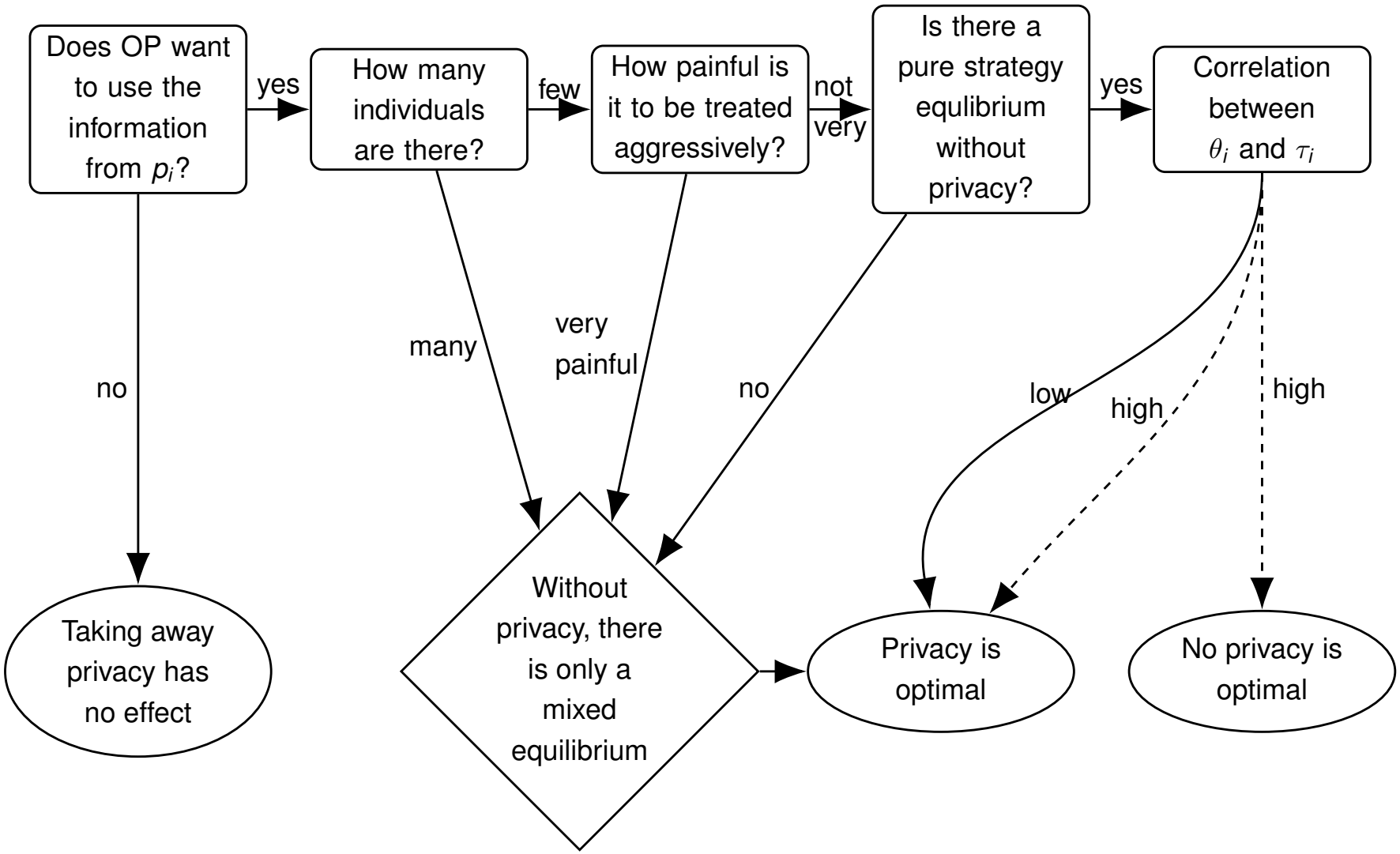


Gains and losses from privacy (intuition for $n = 1$)

- Gains (green) and losses (red) from taking away privacy
- horizontal = employer, vertical = individuals



When is privacy optimal: Sufficient conditions



Can markets help?

- Can privacy be voluntary?
  - No: Some people have no incentive to choose privacy, so that choosing privacy becomes informative in itself
  - Consider the example of the secret ballot (Schelling 1960)
- Can prices (or property rights) for data help?
  - If Alice decides to sell her data, that tells the employer something about others who didn't $\Rightarrow$ but they didn't get compensated!

Example: Credit scoring

- A bank has to decide to whom to lend; repayment probability is not directly observable
- Consider two preferences that are predictive of repayment prob: Taste for education ($\Rightarrow$ education level) and music taste
- Low education and a preference for rap music predict low repayment probability
- (There are "social scoring" companies who collect data about music taste etc from social networking sites, Facebook has a patent on a method!)

Example: Credit scoring

- A bank has to decide to whom to lend; repayment probability is not directly observable
- Consider two preferences that are predictive of repayment prob: Taste for education ($\Rightarrow$ education level) and music taste
- Low education and a preference for rap music predict low repayment probability
- (There are "social scoring" companies who collect data about music taste etc from social networking sites, Facebook has a patent on a method!)
- There is a chilling effect in both cases (but in the first we might consider it desirable!)

Example: Credit scoring

- A bank has to decide to whom to lend; repayment probability is not directly observable
- Consider two preferences that are predictive of repayment prob: Taste for education ($\Rightarrow$ education level) and music taste
- Low education and a preference for rap music predict low repayment probability
- (There are "social scoring" companies who collect data about music taste etc from social networking sites, Facebook has a patent on a method!)
- There is a chilling effect in both cases (but in the first we might consider it desirable!)
- Should the bank be allowed to use data on music taste? ("Equal Credit Opportunity Act" outlaws "redlining" in the US, but does so by blacklisting!)

Example: Work in committees

- A committee is debating two policies (e.g. raise interest rates or not)
- The debate and vote can either be in secret or in public
- There is a correlation between policy preference and competence
- Members worry about being perceived as incompetent; advocate less radical positions

Example: Work in committees

- A committee is debating two policies (e.g. raise interest rates or not)
- The debate and vote can either be in secret or in public
- There is a correlation between policy preference and competence
- Members worry about being perceived as incompetent; advocate less radical positions
- The Fed is forced to publish minutes of FOMC meetings since 1993; studies show an increase in conformity and a decrease of disagreement with the chairman (Meade and Stasavage 2008)
- Thomas Hoenig, President of the Kansas City Fed: “The tape has had some chilling effect on our discussions. I see a lot more people reading their statements.”

What can we learn from this?

- It is rational to care about your privacy
- It also makes sense to care about society's approach to privacy $\Rightarrow$ sometimes privacy can make *everyone* better off
- Markets may not provide optimal privacy

What can we learn from this?

- It is rational to care about your privacy
- It also makes sense to care about society's approach to privacy $\Rightarrow$ sometimes privacy can make *everyone* better off
- Markets may not provide optimal privacy
- Paper recently published:
  Jann, O., & Schottmüller, C. (2020). An informational theory of privacy. *Economic Journal*, 130(625), 93-124.

Questions for the coming years:

- I was awared a PRIMUS project from Charles University (€200k for 3 years) to do more research on “Information Revelation and Privacy in the Information Age”

Questions for the coming years:

- I was awared a PRIMUS project from Charles University (€200k for 3 years) to do more research on “Information Revelation and Privacy in the Information Age”
- What I will be working on in the coming years:
  - how can privacy concerns inhibit the revelation of information in conversations and debates?
  - how can we build mechanisms that make the most out of the Internet’s possibilities to exchange and collect information?
  - when can we rely on markets, when do we need rules/law to protect privacy but reveal information?
- I will be working theoretically, with (big) data and with experiments