Competitive Markets for Personal Data

Simone Galperti Jacopo Perego
UCSD Columbia

May 2023

Competitive Markets for Personal Data

Simone Galperti Jacopo Perego
UCSD Columbia

May 2023

Preliminary

Motivation

Personal data fuels multi-billion dollar industries and is essential input to the digital economy

Yet, consumers are **imperfectly compensated** for their data, and have **limited control** over their use

Expropriation and barter, common practice in the industry (FTC '15)

Motivation

Personal data fuels multi-billion dollar industries and is essential input to the digital economy

Yet, consumers are **imperfectly compensated** for their data, and have **limited control** over their use

Expropriation and barter, common practice in the industry (FTC '15)

Status quo may be source of "market failures:" (Seim et al. '22)

- Lack of compensation may affect incentives and equity
- Lack of control may facilitate market power (no trade)

Motivation

Personal data fuels multi-billion dollar industries and is essential input to the digital economy

Yet, consumers are **imperfectly compensated** for their data, and have **limited control** over their use

Expropriation and barter, common practice in the industry (FTC '15)

Status quo may be source of "market failures:" (Seim et al. '22)

- Lack of compensation may affect incentives and equity
- Lack of control may facilitate market power (no trade)

A "well-functioning" "market for data" would correct these problems

Motivation

Personal data fuels multi-billion dollar industries and is essential input to the digital economy

Yet, consumers are **imperfectly compensated** for their data, and have **limited control** over their use

Expropriation and barter, common practice in the industry (FTC '15)

Status quo may be source of "market failures:" (Seim et al. '22)

- Lack of compensation may affect incentives and equity
- Lack of control may facilitate market power (no trade)

A "well-functioning" "market for data" would correct these problems **?**

This Paper

In this paper,

1. We study **competitive markets for data** and characterize their equilibria

   Extreme yet important benchmark we should know about

This Paper

In this paper,

1. We study **competitive markets for data** and characterize their equilibria

   Extreme yet important benchmark we should know about

2. Show that **pooling externalities** (GLP22) can lead to market failure

   So severe that consumers can be better off under expropriation

This Paper

In this paper,

1. We study **competitive markets for data** and characterize their equilibria

   Extreme yet important benchmark we should know about

2. Show that **pooling externalities** (GLP22) can lead to market failure

   So severe that consumers can be better off under expropriation

3. A solution *a la* Lindhal: more complete markets to restore efficiency

   Price of data needs to depend on its intended use

This Paper

In this paper,

1. We study **competitive markets for data** and characterize their equilibria

   Extreme yet important benchmark we should know about

2. Show that **pooling externalities** (GLP22) can lead to market failure

   So severe that consumers can be better off under expropriation

3. A solution *a la* Lindhal: more complete markets to restore efficiency

   Price of data needs to depend on its intended use

Pooling Externalities

The source of the inefficiency

We model a platform as an information intermediary:

- It uses personal data to intermediate consumers and merchants
- In many cases, platform optimally **withholds** some information

Pooling Externalities

The source of the inefficiency

We model a platform as an information intermediary:

- It uses personal data to intermediate consumers and merchants
- In many cases, platform optimally **withholds** some information

When info is withheld, data of different consumers are pooled together

Pooling Externalities

The source of the inefficiency

We model a platform as an information intermediary:

- It uses personal data to intermediate consumers and merchants
- In many cases, platform optimally **withholds** some information

When info is withheld, data of different consumers are pooled together

In this case, their payoff may depend on which other consumers choose to join the platform

The source of the inefficiency

We model a platform as an information intermediary:

- It uses personal data to intermediate consumers and merchants
- In many cases, platform optimally **withholds** some information

When info is withheld, data of different consumers are pooled together

In this case, their payoff may depend on which other consumers choose to join the platform

We build a competitive economy around this idea and study if/when it leads to market failure

Literature

We know exogenous correlation in consumers' data can lead to externalities

Choi et al. 19, Bergemann et al. 22, Acemoglu et al. 22

Here, we explore novel externality and its consequences in competitive mkts

- It emerges endogenously from how platform uses data even when data are uncorrelated

Literature

We know exogenous correlation in consumers' data can lead to externalities

Choi et al. 19, Bergemann et al. 22, Acemoglu et al. 22

Here, we explore novel externality and its consequences in competitive mkts

- It emerges endogenously from how platform uses data even when data are uncorrelated

Growing interest in studying these markets, avenue for econ theory

Literature

We know exogenous correlation in consumers' data can lead to externalities

Choi et al. 19, Bergemann et al. 22, Acemoglu et al. 22

Here, we explore novel externality and its consequences in competitive mkts

- It emerges endogenously from how platform uses data even when data are uncorrelated

Growing interest in studying these markets, avenue for econ theory

Our approach: (GLP22, GP23, and this paper)

- Use tools from information-design literature

  Bergemann and Morris 19, Kamenica 19

- To answer questions about data markets

  Acquisti et al 16, Bergemann and Bonatti 19, Bergemann and Ottaviani 21

Plan for Talk

1. Leading example to illustrate main ideas and results
2. General model, general results, open questions & limitations

leading example

(many consumers, one platform, one merchant)

A Simple Data Economy

A **merchant** sells widgets to consumers and charges fee $a \in A$

A Simple Data Economy

A **merchant** sells widgets to consumers and charges fee $a \in A$

A **consumer** has unit demand for the widget and WTP $\omega \in \{1, 2\}$

A Simple Data Economy

A **merchant** sells widgets to consumers and charges fee $a \in A$

A **consumer** has unit demand for the widget and WTP $\omega \in \{1, 2\}$

Type-ω consumers have mass $\bar{q}(\omega)$. Assume $\bar{q}(2) > \bar{q}(1)$

A Simple Data Economy

A **merchant** sells widgets to consumers and charges fee $a \in A$

A **consumer** has unit demand for the widget and WTP $\omega \in \{1, 2\}$

Type-ω consumers have mass $\bar{q}(\omega)$. Assume $\bar{q}(2) = 2, \; \bar{q}(1) = 1$

A Simple Data Economy

A **merchant** sells widgets to consumers and charges fee $a \in A$

A **consumer** has unit demand for the widget and WTP $\omega \in \{1, 2\}$

Type-ω consumers have mass $\bar{q}(\omega)$. Assume $\bar{q}(2) = 2,\ \bar{q}(1) = 1$

Each consumer owns a **data record**, which, if sold to the platform, fully reveals her type ω

A Simple Data Economy

A **merchant** sells widgets to consumers and charges fee $a \in A$

A **consumer** has unit demand for the widget and WTP $\omega \in \{1, 2\}$

Type-ω consumers have mass $\bar{q}(\omega)$. Assume $\bar{q}(2) = 2,\ \bar{q}(1) = 1$

Each consumer owns a **data record**, which, if sold to the platform, fully reveals her type ω

Two periods:

1. Consumers and platform trade data records in a competitive market
2. Platform solves a standard **information-design problem** given acquired data records

First Period

Given prices $p(1)$ and $p(2)$ of data records

- Platform chooses which records to buy $\rightsquigarrow$ **database** $q = (q(1), q(2))$
- Consumer chooses whether to sell her record to platform $\rightsquigarrow$ $\zeta(\omega) \in [0, 1]$

First Period

Given prices $p(1)$ and $p(2)$ of data records

- Platform chooses which records to buy $\rightsquigarrow$ **database** $q = (q(1), q(2))$
- Consumer chooses whether to sell her record to platform $\rightsquigarrow$ $\zeta(\omega) \in [0, 1]$

  If she doesn't sell, she does not trade with merchant. She obtains a final payoff $\varepsilon > 0$

First Period

Given prices $p(1)$ and $p(2)$ of data records

- Platform chooses which records to buy $\rightsquigarrow$ **database** $q = (q(1), q(2))$
- Consumer chooses whether to sell her record to platform $\rightsquigarrow \zeta(\omega) \in [0, 1]$

  If she doesn't sell, she does not trade with merchant. She obtains a final payoff $\varepsilon > 0$

  If she sells, her payoff determined in next period...

First Period

Given prices $p(1)$ and $p(2)$ of data records

- Platform chooses which records to buy $\rightsquigarrow$ **database** $q = (q(1), q(2))$
- Consumer chooses whether to sell her record to platform $\rightsquigarrow$ $\zeta(\omega) \in [0, 1]$

  If she doesn't sell, she does not trade with merchant. She obtains a final payoff $\varepsilon > 0$

  If she sells, her payoff determined in next period...

Second Period

Platform intermediates merchant with consumers by providing former with info about latter, so to influence merchant's fee a

E.g.: full info $\rightsquigarrow$ discriminatory fee; no info $\rightsquigarrow$ uniform fee

Second Period

Platform intermediates merchant with consumers by providing former with info about latter, so to influence merchant's fee a

E.g.: full info $\rightsquigarrow$ discriminatory fee; no info $\rightsquigarrow$ uniform fee

	Consumer:	$g(a,\omega) = \max\{\omega - a, 0\}$
Gross payoffs:	Merchant:	$\pi(a,\omega) = a\mathbb{1}(\omega \geq a)$
	Platform:	$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$

Second Period

Platform intermediates merchant with consumers by providing former with info about latter, so to influence merchant's fee a

E.g.: full info $\rightsquigarrow$ discriminatory fee; no info $\rightsquigarrow$ uniform fee

Gross payoffs:	Consumer:	$g(a,\omega) = \max\{\omega - a, 0\}$
	Merchant:	$\pi(a,\omega) = a\mathbb{1}(\omega \geq a)$
	Platform:	$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$

By standard arguments, platform solves

same as in GLP22

Second Period

Platform intermediates merchant with consumers by providing former with info about latter, so to influence merchant's fee a

E.g.: full info $\rightsquigarrow$ discriminatory fee; no info $\rightsquigarrow$ uniform fee

	Consumer:	$g(a, \omega) = \max\{\omega - a, 0\}$
Gross payoffs:	Merchant:	$\pi(a, \omega) = a\mathbb{1}(\omega \geq a)$
	Platform:	$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$

By standard arguments, platform solves

same as in GLP22

$$\overbrace{U(q)}^{\text{platform's expected payoff}}$$

Second Period

Platform intermediates merchant with consumers by providing former with info about latter, so to influence merchant's fee a

E.g.: full info $\rightsquigarrow$ discriminatory fee; no info $\rightsquigarrow$ uniform fee

Gross payoffs:

Consumer:	$g(a,\omega) = \max\{\omega - a, 0\}$
Merchant:	$\pi(a,\omega) = a\mathbb{1}(\omega \geq a)$
Platform:	$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$

By standard arguments, platform solves

same as in GLP22

$$\overbrace{U(q)}^{\text{platform's expected payoff}} = \max_{x:A\times\Omega\to\mathbb{R}_+} \sum_{\omega,a} u(a,\omega)x(a,\omega)$$

such that:

$$\sum_{\omega} \big(\pi(a,\omega) - \pi(\hat{a},\omega)\big)x(a,\omega) \geq 0 \qquad \forall\, a,\hat{a} \in A$$

$$\sum_{a} x(a,\omega) = q(\omega) \qquad \forall\, \omega \in \Omega$$

Second Period

Platform intermediates merchant with consumers by providing former with info about latter, so to influence merchant's fee a

E.g.: full info $\rightsquigarrow$ discriminatory fee; no info $\rightsquigarrow$ uniform fee

Gross payoffs:

Consumer:	$g(a,\omega) = \max\{\omega - a, 0\}$
Merchant:	$\pi(a,\omega) = a\mathbb{1}(\omega \geq a)$
Platform:	$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$

By standard arguments, platform solves

same as in GLP22

$$\overbrace{U(q)}^{\substack{\text{platform's} \\ \text{expected payoff}}} = \max_{x: A\times\Omega\to\mathbb{R}_+} \sum_{\omega,a} u(a,\omega)x(a,\omega)$$

$$\text{such that:} \quad \sum_{\omega} \big(\pi(a,\omega) - \pi(\hat{a},\omega)\big)x(a,\omega) \geq 0 \qquad \forall\, a, \hat{a} \in A$$

$$\sum_{a} x(a,\omega) = q(\omega) \qquad \forall\, \omega \in \Omega$$

Equilibrium

An **equilibrium** consists of prices $(p^*(1), p^*(2))$, a database $q^* = (q^*(1), q^*(2))$, data supply $(\zeta^*(1), \zeta^*(2))$, and mechanism x^* such that

Equilibrium

An **equilibrium** consists of prices $(p^*(1), p^*(2))$, a database $q^* = (q^*(1), q^*(2))$, data supply $(\zeta^*(1), \zeta^*(2))$, and mechanism x^* such that

1. Given p^*, database q^* solves platform's problem:
$$q^* \in \arg\max_{q \in \mathbb{R}_+^\Omega} U(q) - \sum_\omega p^*(\omega) q(\omega)$$

Equilibrium

An **equilibrium** consists of prices $(p^*(1), p^*(2))$, a database $q^* = (q^*(1), q^*(2))$, data supply $(\zeta^*(1), \zeta^*(2))$, and mechanism x^* such that

1. Given p^*, database q^* solves platform's problem:
$$q^* \in \arg\max_{q \in \mathbb{R}_+^{\Omega}} U(q) - \sum_{\omega} p^*(\omega) q(\omega)$$
2. Given p^* and q^*, $\zeta^*(\omega)$ solves ω-consumer's problem
$$\zeta^*(\omega) \in \arg\max_{z \in [0,1]} (1-z)\varepsilon + z\Big(\mathbb{E}_{q^*}(g(a,\omega)) + p^*(\omega)\Big)$$

Equilibrium

An **equilibrium** consists of prices $(p^*(1), p^*(2))$, a database $q^* = (q^*(1), q^*(2))$, data supply $(\zeta^*(1), \zeta^*(2))$, and mechanism x^* such that

1. Given p^*, database q^* solves platform's problem:
$$q^* \in \arg\max_{q \in \mathbb{R}_+^\Omega} U(q) - \sum_\omega p^*(\omega) q(\omega)$$
2. Given p^* and q^*, $\zeta^*(\omega)$ solves ω-consumer's problem
$$\zeta^*(\omega) \in \arg\max_{z \in [0,1]} (1-z)\varepsilon + z\Big(\mathbb{E}_{q^*}(g(a,\omega)) + p^*(\omega)\Big)$$
3. Markets clear
$$q^*(\omega) = \zeta^*(\omega)\bar{q}(\omega), \quad \forall \omega$$

Equilibrium

An **equilibrium** consists of prices $(p^*(1), p^*(2))$, a database $q^* = (q^*(1), q^*(2))$, data supply $(\zeta^*(1), \zeta^*(2))$, and mechanism x^* such that

1. Given p^*, database q^* solves platform's problem:
$$q^* \in \arg\max_{q \in \mathbb{R}_+^\Omega} U(q) - \sum_\omega p^*(\omega) q(\omega)$$
2. Given p^* and q^*, $\zeta^*(\omega)$ solves ω-consumer's problem
$$\zeta^*(\omega) \in \arg\max_{z \in [0,1]} (1-z)\varepsilon + z\Big(\mathbb{E}_{q^*}(g(a,\omega)) + p^*(\omega)\Big)$$
3. Markets clear
$$q^*(\omega) = \zeta^*(\omega)\bar{q}(\omega), \quad \forall\omega$$
4. Given q^*, x^* solves platform's information-design problem

Equilibrium

An **equilibrium** consists of prices $(p^*(1), p^*(2))$, a database $q^* = (q^*(1), q^*(2))$, data supply $(\zeta^*(1), \zeta^*(2))$, and mechanism x^* such that

1. Given p^*, database q^* solves platform's problem:
$$q^* \in \arg\max_{q \in \mathbb{R}_+^\Omega} U(q) - \sum_\omega p^*(\omega) q(\omega)$$
2. Given p^* and q^*, $\zeta^*(\omega)$ solves ω-consumer's problem
$$\zeta^*(\omega) \in \arg\max_{z \in [0,1]} (1-z)\varepsilon + z\Big(\mathbb{E}_{q^*}(g(a,\omega)) + p^*(\omega)\Big)$$
3. Markets clear
$$q^*(\omega) = \zeta^*(\omega)\bar{q}(\omega), \quad \forall \omega$$
4. Given q^*, x^* solves platform's information-design problem

discussion

Discussion

Novelty: Endogenous "prior" q in an otherwise standard ID problem

- Consumer data acquired in a **competitive market**

Discussion

Novelty: Endogenous "prior" q in an otherwise standard ID problem

- Consumer data acquired in a **competitive market**

Unrealistic to assume price-taking behavior with single platform

- It allows us to illustrate source of market failure in a simple way

Discussion

Novelty: Endogenous "prior" q in an otherwise standard ID problem

- Consumer data acquired in a **competitive market**

Unrealistic to assume price-taking behavior with single platform

- It allows us to illustrate source of market failure in a simple way

Our general model features:

- Many platforms, many merchants, many types, arbitrary objectives, partially informative records

analysis

Efficiency

In equilibrium, **consumers' welfare** is:

$$W(q^*) \quad = \quad \sum_\omega \bar{q}(\omega)\Big((1-\zeta^*(\omega))\varepsilon + \zeta^*(\omega)\Big(\mathbb{E}_{q^*}(g(a,\omega)) + p^*(\omega)\Big)\Big)$$

Efficiency

In equilibrium, **consumers' welfare** is:

$$
\begin{aligned}
W(q^*) &= \textstyle\sum_\omega \bar{q}(\omega)\Big((1-\zeta^*(\omega))\varepsilon + \zeta^*(\omega)\Big(\mathbb{E}_{q^*}(g(a,\omega)) + p^*(\omega)\Big)\Big) \\
&= \underbrace{\textstyle\sum_\omega (\bar{q}(\omega) - q^*(\omega))\varepsilon + \sum_\omega q^*(\omega)\mathbb{E}_{q^*}(g(a,\omega))}_{\text{consumers' surplus, } CS(q^*)} + \underbrace{\textstyle\sum_\omega q^*(\omega)p^*(\omega)}_{\text{transfers, } T(q^*)}
\end{aligned}
$$

Efficiency

In equilibrium, **consumers' welfare** is:

$$
\begin{aligned}
W(q^*) &= \textstyle\sum_\omega \bar{q}(\omega)\Big((1-\zeta^*(\omega))\varepsilon + \zeta^*(\omega)\Big(\mathbb{E}_{q^*}(g(a,\omega)) + p^*(\omega)\Big)\Big) \\
&= \underbrace{\textstyle\sum_\omega (\bar{q}(\omega) - q^*(\omega))\varepsilon + \sum_\omega q^*(\omega)\mathbb{E}_{q^*}(g(a,\omega))}_{\text{consumers' surplus, } CS(q^*)} + \underbrace{\textstyle\sum_\omega q^*(\omega)p^*(\omega)}_{\text{transfers, } T(q^*)}
\end{aligned}
$$

In equilibrium, the platform must obtain zero payoff

Efficiency

In equilibrium, **consumers' welfare** is:

$$\begin{aligned} W(q^*) &= \sum_\omega \bar{q}(\omega)\Big((1-\zeta^*(\omega))\varepsilon + \zeta^*(\omega)\Big(\mathbb{E}_{q^*}(g(a,\omega)) + p^*(\omega)\Big)\Big) \\ &= \underbrace{\sum_\omega (\bar{q}(\omega) - q^*(\omega))\varepsilon + \sum_\omega q^*(\omega)\mathbb{E}_{q^*}(g(a,\omega))}_{\text{consumers' surplus, } CS(q^*)} + \underbrace{\sum_\omega q^*(\omega)p^*(\omega)}_{\text{transfers, } T(q^*)} \end{aligned}$$

In equilibrium, the platform must obtain zero payoff, i.e.

$$U(q^*) = \underbrace{\sum_\omega q^*(\omega)p^*(\omega)}_{\text{transfers, } T(q^*)}$$

Efficiency

In equilibrium, **consumers' welfare** is:

$$\begin{aligned} W(q^*) &= \sum_\omega \bar{q}(\omega)\Big((1-\zeta^*(\omega))\varepsilon + \zeta^*(\omega)\Big(\mathbb{E}_{q^*}(g(a,\omega)) + p^*(\omega)\Big)\Big) \\ &= \underbrace{\sum_\omega (\bar{q}(\omega) - q^*(\omega))\varepsilon + \sum_\omega q^*(\omega)\mathbb{E}_{q^*}(g(a,\omega))}_{\text{consumers' surplus, } CS(q^*)} + \underbrace{\sum_\omega q^*(\omega)p^*(\omega)}_{\text{transfers, } T(q^*)} \end{aligned}$$

In equilibrium, the platform must obtain zero payoff, i.e.

$$U(q^*) = \underbrace{\sum_\omega q^*(\omega)p^*(\omega)}_{\text{transfers, } T(q^*)}$$

Therefore, $W(q^*) = CS(q^*) + U(q^*)$

Efficiency

In equilibrium, **consumers' welfare** is:

$$\begin{aligned} W(q^*) &= \textstyle\sum_\omega \bar{q}(\omega)\Big((1-\zeta^*(\omega))\varepsilon + \zeta^*(\omega)\Big(\mathbb{E}_{q^*}(g(a,\omega)) + p^*(\omega)\Big)\Big) \\ &= \underbrace{\textstyle\sum_\omega (\bar{q}(\omega) - q^*(\omega))\varepsilon + \sum_\omega q^*(\omega)\mathbb{E}_{q^*}(g(a,\omega))}_{\text{consumers' surplus, } CS(q^*)} + \underbrace{\textstyle\sum_\omega q^*(\omega)p^*(\omega)}_{\text{transfers, } T(q^*)} \end{aligned}$$

In equilibrium, the platform must obtain zero payoff, i.e.

$$U(q^*) = \underbrace{\textstyle\sum_\omega q^*(\omega)p^*(\omega)}_{\text{transfers, } T(q^*)}$$

Therefore, $$W(q^*) = CS(q^*) + U(q^*)$$

Definition

An equilibrium data allocation q^* is **efficient** if there is no q such that $W(q) > W(q^*)$

Efficiency

In equilibrium, **consumers' welfare** is:

$$\begin{aligned} W(q^*) &= \textstyle\sum_\omega \bar{q}(\omega)\Big((1-\zeta^*(\omega))\varepsilon + \zeta^*(\omega)\Big(\mathbb{E}_{q^*}(g(a,\omega)) + p^*(\omega)\Big)\Big) \\ &= \underbrace{\textstyle\sum_\omega (\bar{q}(\omega) - q^*(\omega))\varepsilon + \sum_\omega q^*(\omega)\mathbb{E}_{q^*}(g(a,\omega))}_{\text{consumers' surplus, } CS(q^*)} + \underbrace{\textstyle\sum_\omega q^*(\omega)p^*(\omega)}_{\text{transfers, } T(q^*)} \end{aligned}$$

In equilibrium, the platform must obtain zero payoff, i.e.

$$U(q^*) = \underbrace{\textstyle\sum_\omega q^*(\omega)p^*(\omega)}_{\text{transfers, } T(q^*)}$$

Therefore, $$W(q^*) = CS(q^*) + U(q^*)$$

Definition

An equilibrium data allocation q^* is **efficient** if there is no q such that $W(q) > W(q^*)$

Roadmap

What are the competitive equilibria of this economy? Are they efficient?

Roadmap

What are the competitive equilibria of this economy? Are they efficient?

It depends on how platform uses consumers' data, and thus on its incentives

Roadmap

What are the competitive equilibria of this economy? Are they efficient?

It depends on how platform uses consumers' data, and thus on its incentives

To illustrate,

- We will begin from the case $\gamma = 0$, i.e.,

$$u(a, \omega) = \beta \, g(a, \omega) + \gamma \, \pi(a, \omega)$$

  and show that equilibria are **inefficient**

Roadmap

What are the competitive equilibria of this economy? Are they efficient?

It depends on how platform uses consumers' data, and thus on its incentives

To illustrate,

- We will begin from the case $\gamma = 0$, i.e.,

  $$u(a,\omega) = \beta\ g(a,\omega) + \gamma\ \pi(a,\omega)$$

  and show that equilibria are **inefficient**

- We will then discuss the case $\beta = 0$, i.e.,

  $$u(a,\omega) = \beta\ g(a,\omega) + \gamma\ \pi(a,\omega)$$

  and show that equilibria are **efficient**

Roadmap

What are the competitive equilibria of this economy? Are they efficient?

It depends on how platform uses consumers' data, and thus on its incentives

To illustrate,

- We will begin from the case $\gamma = 0$, i.e., $\Leftarrow$

$$u(a, \omega) = \beta \, g(a, \omega) + \gamma \, \pi(a, \omega)$$

  and show that equilibria are **inefficient**

- We will then discuss the case $\beta = 0$, i.e.,

$$u(a, \omega) = \beta \, g(a, \omega) + \gamma \, \pi(a, \omega)$$

  and show that equilibria are **efficient**

Efficient Data Allocation

What is the efficient data allocation in this case?

Record Type
$\omega = 1$
$\omega = 2$

Efficient Data Allocation

What is the efficient data allocation in this case?

Record Type	**Existing Records**
$\omega = 1$	1
$\omega = 2$	2

Efficient Data Allocation

What is the efficient data allocation in this case?

Record Type	Existing Records	Records Retained
$\omega = 1$	1	0
$\omega = 2$	2	1

Efficient Data Allocation

What is the efficient data allocation in this case?

Record Type	Existing Records	Records Retained	Platform's Database
$\omega = 1$	1	0	1
$\omega = 2$	2	1	1

Efficient Data Allocation

What is the efficient data allocation in this case?

Record Type	Existing Records	Records Retained	Platform's Database	Use
$\omega = 1$	1	0	1	$\rightarrow s^L \rightarrow a = 1$
$\omega = 2$	2	1	1	

Efficient Data Allocation

What is the efficient data allocation in this case?

Record Type	Existing Records	Records Retained	Platform's Database	Use	Welfare Benchmark
$\omega = 1$	1	0	1	$\rightarrow s^L \rightarrow a = 1$	$\underbrace{1 + \varepsilon}_{CS(q)} + \underbrace{\beta}_{U(q)}$
$\omega = 2$	2	1	1		

Comments

Given q, platform finds it optimal to withhold info

It does so by pooling all data records in database

Comments

Given q, platform finds it optimal to withhold info

It does so by pooling all data records in database

When data records are pooled, consumer' payoff depend on what other records platform has pooling externality (GLP23)

Comments

Given q, platform finds it optimal to withhold info

It does so by pooling all data records in database

When data records are pooled, consumer' payoff depend on what other records platform has pooling externality (GLP23)

To see this, imagine if some type-1 consumers were to leave database...

Comments

Given q, platform finds it optimal to withhold info

It does so by pooling all data records in database

When data records are pooled, consumer' payoff depend on what other records platform has pooling externality (GLP23)

To see this, imagine if some type-1 consumers were to leave database...

Model "enables" this externality, which will lead to inefficiencies

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Equilibrium: “Too Little Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **“Too little data”**
2. $\beta > \varepsilon$: “Too much data”

Equilibrium: “Too Little Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. **$\beta < \varepsilon$: “Too little data”**
2. $\beta > \varepsilon$: “Too much data”

Record Type	Prices	Records Kept	Platform’s Database
$\omega = 1$	β	1	0
$\omega = 2$	0	2	0

Equilibrium: “Too Little Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **“Too little data”**
2. $\beta > \varepsilon$: “Too much data”

Record Type	**Prices**	**Records Kept**	**Platform’s Database**	**Use** (Trivial)
$\omega = 1$	β	1	0	$\longrightarrow s^L \longrightarrow a = 1$
$\omega = 2$	0	2	0	$\longrightarrow s^H \longrightarrow a = 2$

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	**Prices**	**Records Kept**	**Platform's Database**	**Use** (Trivial)	**Welfare**
$\omega = 1$	β	1	$0 \longrightarrow s^L \longrightarrow a = 1$		3ε
$\omega = 2$	0	2	$0 \longrightarrow s^H \longrightarrow a = 2$		

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	**Prices**	**Records Kept**	**Platform's Database**	**Use** (Trivial)	**Welfare**
$\omega = 1$	β	1	0	$0 \longrightarrow s^L \longrightarrow a = 1$	3ε
$\omega = 2$	0	2	0	$0 \longrightarrow s^H \longrightarrow a = 2$	

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	**Prices**	**Records Kept**	**Platform's Database**	**Use** (Trivial)	**Welfare**
$\omega = 1$	β	1	0	$\longrightarrow s^L \longrightarrow a = 1$	3ε
$\omega = 2$	0	2	0	$\longrightarrow s^H \longrightarrow a = 2$	

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	**Prices**	**Records Kept**	**Platform's Database**	**Use** (Trivial)	**Welfare**
$\omega = 1$	β	1	0	$\longrightarrow s^L \longrightarrow a = 1$	3ε
$\omega = 2$	0	2	0	$\longrightarrow s^H \longrightarrow a = 2$	

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	**Prices**	**Records Kept**	**Platform's Database**	**Use** (Trivial)	**Welfare**
$\omega = 1$	β	1	0	$\longrightarrow s^L \longrightarrow a = 1$	3ε
$\omega = 2$	0	2	0	$\longrightarrow s^H \longrightarrow a = 2$	

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	**Prices**	**Records Kept**	**Platform's Database**	**Use** (Trivial)	**Welfare**
$\omega = 1$	β	1	0	$\longrightarrow s^L \longrightarrow a = 1$	3ε
$\omega = 2$	0	2	0	$\longrightarrow s^H \longrightarrow a = 2$	

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	Prices	Records Kept	Platform's Database	Use (Trivial)	Welfare
$\omega = 1$	β	1	$0 \longrightarrow s^L \longrightarrow a = 1$		3ε
$\omega = 2$	0	2	$0 \longrightarrow s^H \longrightarrow a = 2$		

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. **$\beta < \varepsilon$: "Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	**Prices**	**Records Kept**	**Platform's Database**	**Use** (Trivial)	**Welfare**
$\omega = 1$	β	1	0	$\longrightarrow s^L \longrightarrow a = 1$	3ε
$\omega = 2$	0	2	0	$\longrightarrow s^H \longrightarrow a = 2$	

Comments:

Equilibrium: “Too Little Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **“Too little data”**
2. $\beta > \varepsilon$: “Too much data”

Record Type	Prices	Records Kept	Platform’s Database	Use (Trivial)	Welfare
$\omega = 1$	β	1	0	$\longrightarrow s^L \longrightarrow a = 1$	3ε
$\omega = 2$	0	2	0	$\longrightarrow s^H \longrightarrow a = 2$	

Comments:

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. **$\beta < \varepsilon$: "Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	Prices	Records Kept	Platform's Database	Use (Trivial)	Welfare
$\omega = 1$	β	1	0	$\longrightarrow s^L \longrightarrow a = 1$	3ε
$\omega = 2$	0	2	0	$\longrightarrow s^H \longrightarrow a = 2$	

Comments:

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	Prices	Records Kept	Platform's Database	Use (Trivial)	Welfare
$\omega = 1$	β	1	$0 \longrightarrow s^L \longrightarrow a = 1$		3ε
$\omega = 2$	0	2	$0 \longrightarrow s^H \longrightarrow a = 2$		

Comments:

1. Low-type consumers have no incentive to sell:
   Price $p^*(1) = \beta$ is too low $(\beta < \varepsilon)$

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	Prices	Records Kept	Platform's Database	Use (Trivial)	Welfare
$\omega = 1$	β	1	$0 \longrightarrow s^L \longrightarrow a = 1$		3ε
$\omega = 2$	0	2	$0 \longrightarrow s^H \longrightarrow a = 2$		

Comments:

2. High-type consumers have no incentive to sell:
   Price $p^*(2) = 0$ is too low $(\varepsilon > 0)$

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	**Prices**	**Records Kept**	**Platform's Database**	**Use** (Trivial)	**Welfare**
$\omega = 1$	β	1	$0 \longrightarrow s^L \longrightarrow a = 1$		3ε
$\omega = 2$	0	2	$0 \longrightarrow s^H \longrightarrow a = 2$		

Comments:

3. Platform has no strict incentive to buy at these prices.
   Equilibrium prices = marginal values

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	Prices	Records Kept	Platform's Database	Use (Trivial)	Welfare
$\omega = 1$	β	1	$0 \longrightarrow s^L \longrightarrow a = 1$		3ε
$\omega = 2$	0	2	$0 \longrightarrow s^H \longrightarrow a = 2$		

Comments:

4. Equilibrium welfare is inefficiently low $\left(\varepsilon < \frac{1+\beta}{2}\right)$

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	Prices	Records Kept	Platform's Database	Use (Trivial)	Welfare
$\omega = 1$	β	1	0	$\longrightarrow s^L \longrightarrow a = 1$	3ε
$\omega = 2$	0	2	0	$\longrightarrow s^H \longrightarrow a = 2$	

Comments:

5. High-type consumers would want to subsidize low-type consumers to sell their data, but market is too incomplete

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	Prices	Records Kept	Platform's Database	Use (Trivial)	Welfare
$\omega = 1$	β	1	0	$\longrightarrow s^L \longrightarrow a = 1$	3ε
$\omega = 2$	0	2	0	$\longrightarrow s^H \longrightarrow a = 2$	

Comments:

6. For small ε, consumers welfare would be higher under expropriation!

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	Prices	Records Kept	Platform's Database
$\omega = 1$	0	0	1
$\omega = 2$	0	0	2

Comments:

6. For small ε, consumers welfare would be higher under expropriation!

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

$$
\begin{tikzcd}
\text{Record Type} & \text{Prices} & \text{Records Kept} & \text{Platform's Database} & & \text{Use} \\
\omega = 1 & 0 & 0 & 1 \arrow[r] & s^L \arrow[r] & a = 1 \\
\omega = 2 & 0 & 0 & 2 \arrow[ur, "\frac{1}{2}"] \arrow[r, "\frac{1}{2}"'] & s^H \arrow[r] & a = 2
\end{tikzcd}
$$

Comments:

6. For small ε, consumers welfare would be higher under expropriation!

Equilibrium: "Too Little Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: **"Too little data"**
2. $\beta > \varepsilon$: "Too much data"

Record Type	Prices	Records Kept	Platform's Database	Use	Welfare
$\omega = 1$	0	0	$1 \longrightarrow s^L$	$s^L \longrightarrow a = 1$	1
$\omega = 2$	0	0	$2 \xrightarrow{\frac{1}{2}} s^L$, $2 \xrightarrow{\frac{1}{2}} s^H$	$s^H \longrightarrow a = 2$	

Comments:

6. For small ε, consumers welfare would be higher under expropriation!

Equilibrium: "Too Much Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: "Too little data"
2. $\beta > \varepsilon$: **"Too much data"**

Equilibrium: "Too Much Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: "Too little data"
2. $\beta > \varepsilon$: **"Too much data"**

Record Type	Prices	Records Kept	Platform's Database
$\omega = 1$	β	0	1
$\omega = 2$	0	0	2

Equilibrium: “Too Much Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: “Too little data”
2. $\beta > \varepsilon$: **“Too much data”**

Record Type	Prices	Records Kept	Platform’s Database	Use
$\omega = 1$	β	0	$1 \longrightarrow s^L$	$s^L \longrightarrow a = 1$
$\omega = 2$	0	0	$2 \xrightarrow{\frac{1}{2}} s^H$	$s^H \longrightarrow a = 2$

Equilibrium: “Too Much Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: “Too little data”
2. $\beta > \varepsilon$: **“Too much data”**

Record Type	Prices	Records Kept	Platform’s Database	Use	Welfare
$\omega = 1$	β	0	$1 \longrightarrow s^L$	$s^L \longrightarrow a = 1$	$1 + \beta$
$\omega = 2$	0	0	$2 \xrightarrow{\frac{1}{2}} s^L$, $2 \xrightarrow{\frac{1}{2}} s^H$	$s^H \longrightarrow a = 2$	

Equilibrium: “Too Much Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: “Too little data”
2. $\beta > \varepsilon$: **“Too much data”**

Record Type	Prices	Records Kept	Platform’s Database	Use	Welfare
$\omega = 1$	β	0	$1 \longrightarrow s^L$	$s^L \longrightarrow a = 1$	$1 + \beta$
$\omega = 2$	0	0	$2 \xrightarrow{\frac{1}{2}} s^L$, $2 \xrightarrow{\frac{1}{2}} s^H$	$s^H \longrightarrow a = 2$	

Equilibrium: "Too Much Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: "Too little data"
2. $\beta > \varepsilon$: **"Too much data"**

Record Type	Prices	Records Kept	Platform's Database		Use	Welfare
$\omega = 1$	β	0	1	$\longrightarrow s^L$	$\longrightarrow a = 1$	$1 + \beta$
$\omega = 2$	0	0	2	$\xrightarrow{\frac{1}{2}} s^L$, $\xrightarrow{\frac{1}{2}} s^H$	$s^H \longrightarrow a = 2$	

Equilibrium: “Too Much Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: “Too little data”
2. **$\beta > \varepsilon$: “Too much data”**

Record Type	Prices	Records Kept	Platform’s Database	Use	Welfare
$\omega = 1$	β	0	$1 \longrightarrow s^L$	$s^L \longrightarrow a = 1$	$1 + \beta$
$\omega = 2$	0	0	$2 \xrightarrow{\frac{1}{2}} s^L$, $2 \xrightarrow{\frac{1}{2}} s^H$	$s^H \longrightarrow a = 2$	

Equilibrium: “Too Much Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: “Too little data”
2. $\beta > \varepsilon$**: “Too much data”**

Record Type	Prices	Records Kept	Platform’s Database	Use	Welfare
$\omega = 1$	β	0	$1 \longrightarrow s^L$	$s^L \longrightarrow a = 1$	$1 + \beta$
$\omega = 2$	0	0	$2 \xrightarrow{\frac{1}{2}} s^L$; $2 \xrightarrow[\frac{1}{2}]{} s^H$	$s^H \longrightarrow a = 2$	

Equilibrium: “Too Much Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: “Too little data”
2. **$\beta > \varepsilon$: “Too much data”**

Record Type	Prices	Records Kept	Platform’s Database	Use	Welfare
$\omega = 1$	β	0	$1 \longrightarrow s^L$	$s^L \longrightarrow a = 1$	$1 + \beta$
$\omega = 2$	0	0	$2 \xrightarrow{\frac{1}{2}} s^L$, $2 \xrightarrow{\frac{1}{2}} s^H$	$s^H \longrightarrow a = 2$	

Equilibrium: "Too Much Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: "Too little data"
2. $\beta > \varepsilon$: **"Too much data"**

Record Type	Prices	Records Kept	Platform's Database	Use	Welfare
$\omega = 1$	β	0	1	$\longrightarrow s^L \longrightarrow a = 1$	$1 + \beta$
$\omega = 2$	0	0	2	$\xrightarrow[\frac{1}{2}]{} s^H \longrightarrow a = 2$	

Equilibrium: “Too Much Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: “Too little data”
2. $\beta > \varepsilon$: **“Too much data”**

Record Type	Prices	Records Kept	Platform’s Database	Use	Welfare
$\omega = 1$	β	0	1	$1 \longrightarrow s^L \longrightarrow a = 1$	$1 + \beta$
$\omega = 2$	0	0	2	$2 \xrightarrow{\frac{1}{2}} s^L$; $2 \xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$	

Equilibrium: “Too Much Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: “Too little data”
2. $\beta > \varepsilon$: **“Too much data”**

Record Type	Prices	Records Kept	Platform’s Database	Use	Welfare
$\omega = 1$	β	0	$1 \longrightarrow s^L$	$s^L \longrightarrow a = 1$	$1 + \beta$
$\omega = 2$	0	0	$2 \xrightarrow{\frac{1}{2}} s^H$	$s^H \longrightarrow a = 2$	

Comments:

Equilibrium: “Too Much Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: “Too little data”
2. $\beta > \varepsilon$: **“Too much data”**

Record Type	Prices	Records Kept	Platform’s Database	Use	Welfare
$\omega = 1$	β	0	$1 \longrightarrow s^L$	$s^L \longrightarrow a = 1$	$1 + \beta$
$\omega = 2$	0	0	$2 \xrightarrow{\frac{1}{2}} s^L$, $2 \xrightarrow{\frac{1}{2}} s^H$	$s^H \longrightarrow a = 2$	

Comments:

Equilibrium: "Too Much Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: "Too little data"
2. $\beta > \varepsilon$: **"Too much data"**

Record Type	Prices	Records Kept	Platform's Database	Use	Welfare
$\omega = 1$	β	0	1	$\longrightarrow s^L \longrightarrow a = 1$	$1 + \beta$
$\omega = 2$	0	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$	

Comments:

Equilibrium: "Too Much Data"

Let's compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: "Too little data"
2. **$\beta > \varepsilon$: "Too much data"**

Record Type	Prices	Records Kept	Platform's Database	Use	Welfare
$\omega = 1$	β	0	1	$\longrightarrow s^L \longrightarrow a = 1$	$1 + \beta$
$\omega = 2$	0	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$	

Comments:

1. Too many high-type consumers sell. Attracted by expected gain $(\frac{1}{2})$, they decrease each other payoffs

Equilibrium: “Too Much Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: “Too little data”
2. $\beta > \varepsilon$: **“Too much data”**

Record Type	Prices	Records Kept	Platform’s Database	Use	Welfare
$\omega = 1$	β	0	1	$\longrightarrow s^L \longrightarrow a = 1$	$1 + \beta$
$\omega = 2$	0	0	2	$\xrightarrow{\frac{1}{2}} s^L$; $\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$	

Comments:

2. Welfare is inefficiently low

Equilibrium: “Too Much Data”

Let’s compute the **competitive equilibrium** of this data economy:

Two cases to consider, both leading to market failure:

1. $\beta < \varepsilon$: “Too little data”
2. $\beta > \varepsilon$: **“Too much data”**

Record Type	Prices	Records Kept	Platform’s Database	Use	Welfare
$\omega = 1$	β	0	$1 \longrightarrow s^L$	$s^L \longrightarrow a = 1$	$1 + \beta$
$\omega = 2$	0	0	$2 \xrightarrow{\frac{1}{2}} s^L$, $2 \xrightarrow{\frac{1}{2}} s^H$	$s^H \longrightarrow a = 2$	

Comments:

3. Negative price on high-type consumers? Again, not an equilibrium...

Summary

Equilibrium is generically inefficiency (i.e., for all ε, β)

Consumer welfare can be even lower than under expropriation

- Perverse consequence of empowering consumers

Summary

Equilibrium is generically inefficiency (i.e., for all ε, β)

Consumer welfare can be even lower than under expropriation

- Perverse consequence of empowering consumers

Two kinds of failures:

- "Too little data:" Low-type consumer fails to internalize positive externality of selling
- "Too much data:" High-type consumer fails to internalize negative externality of selling

Summary

Equilibrium is generically inefficiency (i.e., for all ε, β)

Consumer welfare can be even lower than under expropriation

- Perverse consequence of empowering consumers

Two kinds of failures:

- "Too little data:" Low-type consumer fails to internalize positive externality of selling
- "Too much data:" High-type consumer fails to internalize negative externality of selling

Both failures originates from same source:

- Platform has incentives to withhold information

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$$

Record Type	Prices	Records Kept	Platform's Database
$\omega = 1$	γ	0	1
$\omega = 2$	2γ	0	2

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	**Prices**	**Records Kept**	**Platform's Database**	**Use** (Full Info)
$\omega = 1$	γ	0	1	$\longrightarrow s^L \longrightarrow a = 1$
$\omega = 2$	2γ	0	2	$\longrightarrow s^H \longrightarrow a = 2$

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use (Full Info)	Welfare
$\omega = 1$	γ	0	1	$\longrightarrow s^L \longrightarrow a = 1$	5γ
$\omega = 2$	2γ	0	2	$\longrightarrow s^H \longrightarrow a = 2$	

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use (Full Info)	Welfare
$\omega = 1$	γ	0	1	$\longrightarrow s^L \longrightarrow a = 1$	5γ
$\omega = 2$	2γ	0	2	$\longrightarrow s^H \longrightarrow a = 2$	

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use (Full Info)	Welfare
$\omega = 1$	γ	0	1	$\longrightarrow s^L \longrightarrow a = 1$	5γ
$\omega = 2$	2γ	0	2	$\longrightarrow s^H \longrightarrow a = 2$	

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use (Full Info)	Welfare
$\omega = 1$	γ	0	1	$\longrightarrow s^L \longrightarrow a = 1$	5γ
$\omega = 2$	2γ	0	2	$\longrightarrow s^H \longrightarrow a = 2$	

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use (Full Info)	Welfare
$\omega = 1$	γ	0	1	$\longrightarrow s^L \longrightarrow a = 1$	5γ
$\omega = 2$	2γ	0	2	$\longrightarrow s^H \longrightarrow a = 2$	

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use (Full Info)	Welfare
$\omega = 1$	γ	0	1	$\longrightarrow s^L \longrightarrow a = 1$	5γ
$\omega = 2$	2γ	0	2	$\longrightarrow s^H \longrightarrow a = 2$	

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use (Full Info)	Welfare
$\omega = 1$	γ	0	1	$\longrightarrow s^L \longrightarrow a = 1$	5γ
$\omega = 2$	2γ	0	2	$\longrightarrow s^H \longrightarrow a = 2$	

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use (Full Info)	Welfare
$\omega = 1$	γ	0	$1 \longrightarrow s^L \longrightarrow a = 1$		5γ
$\omega = 2$	2γ	0	$2 \longrightarrow s^H \longrightarrow a = 2$		

Comments:

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use (Full Info)	Welfare
$\omega = 1$	γ	0	1	$\longrightarrow s^L \longrightarrow a = 1$	5γ
$\omega = 2$	2γ	0	2	$\longrightarrow s^H \longrightarrow a = 2$	

Comments:

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use (Full Info)	Welfare
$\omega = 1$	γ	0	$1 \longrightarrow s^L$	$\longrightarrow a = 1$	5γ
$\omega = 2$	2γ	0	$2 \longrightarrow s^H$	$\longrightarrow a = 2$	

Comments:

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use (Full Info)	Welfare
$\omega = 1$	γ	0	$1 \longrightarrow s^L \longrightarrow a = 1$		5γ
$\omega = 2$	2γ	0	$2 \longrightarrow s^H \longrightarrow a = 2$		

Comments:

1. Platform does not withhold information from the merchant $\rightsquigarrow$ price discrimination

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use (Full Info)	Welfare
$\omega = 1$	γ	0	1	$\longrightarrow s^L \longrightarrow a = 1$	5γ
$\omega = 2$	2γ	0	2	$\longrightarrow s^H \longrightarrow a = 2$	

Comments:

2. Crucially, payoff of a consumer is independent of decisions of other consumers. No externalities

Equilibrium

What happens when platform has no incentive to withhold? $\rightsquigarrow$ **Efficiency**

Suppose that $\beta = 0$ and $\gamma > \varepsilon$: as if platform *is* merchant

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use (Full Info)	Welfare
$\omega = 1$	γ	0	1	$\longrightarrow s^L \longrightarrow a = 1$	5γ
$\omega = 2$	2γ	0	2	$\longrightarrow s^H \longrightarrow a = 2$	

Comments:

3. Consumers get fully compensated: They get payoff that platform makes with their data. $\rightsquigarrow$ Equilibrium is efficient

Beyond the Example

In multi-sided markets, platforms balance complex objectives

In some cases, this leads to information withholding $(\beta > \gamma)$

This generates externalities that **can** make equilibrium in data markets inefficient

In a world without intermediaries, data markets **would be** efficient

remedies

An Economy with Lindhal Prices

How can we fix inefficiencies discussed so far?

An Economy with Lindhal Prices

How can we fix inefficiencies discussed so far?

A classic solution: following e.g. Arrow 69, Laffont 78

- Platform has to buy record for a **specific purpose** (i.e. an action recommendation a)
- More complete markets: Prices are $p(\omega, a)$

An Economy with Lindhal Prices

How can we fix inefficiencies discussed so far?

A classic solution: following e.g. Arrow 69, Laffont 78

- Platform has to buy record for a **specific purpose** (i.e. an action recommendation a)
- More complete markets: Prices are $p(\omega, a)$

Note: This presumes data use is contractible

An Economy with Lindhal Prices

How can we fix inefficiencies discussed so far?

A classic solution: following e.g. Arrow 69, Laffont 78

- Platform has to buy record for a **specific purpose** (i.e. an action recommendation a)
- More complete markets: Prices are $p(\omega, a)$

Note: This presumes data use is contractible

Result. Equilibria in this economy exist and are (first-best) efficient.

Back to Example

To illustrate, return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta \; g(a, \omega) + \gamma \; \pi(a, \omega)$$

Record Type

$\omega = 1$

$\omega = 2$

Back to Example

To illustrate, return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta \; g(a, \omega) + \gamma \; \pi(a, \omega)$$

Record Type
$\omega = 1$
$\omega = 2$

Back to Example

To illustrate, return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta\; g(a, \omega) + \gamma\; \pi(a, \omega)$$

Record Type	Prices
$\omega = 1$	$p^*(1)$
$\omega = 2$	$p^*(2)$

Back to Example

To illustrate, return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta\ g(a, \omega) + \gamma\ \pi(a, \omega)$$

Record Type	**Prices**
$\omega = 1$	$p^*(1, a = 1)$ $p^*(1, a = 2)$
$\omega = 2$	$p^*(2, a = 1)$ $p^*(2, a = 2)$

Back to Example

To illustrate, return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta\ g(a, \omega) + \gamma\ \pi(a, \omega)$$

Record Type	**Prices**
$\omega = 1$	$p^*(1, a = 1)$
	$p^*(1, a = 2)$
$\omega = 2$	$p^*(2, a = 1)$
	$p^*(2, a = 2)$

Back to Example

To illustrate, return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$$

Record Type	**Prices**
$\omega = 1$	$p^*(1,1) = \beta + 1 - \varepsilon$ $p^*(1,2) = 0$
$\omega = 2$	$p^*(2,1) = -(1 - \varepsilon)$ $p^*(2,2) = 0$

Back to Example

To illustrate, return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$$

Record Type	Prices	Records Kept	Platform's Database
$\omega = 1$	$p^*(1,1) = \beta + 1 - \varepsilon$ $p^*(1,2) = 0$	0	1
$\omega = 2$	$p^*(2,1) = -(1 - \varepsilon)$ $p^*(2,2) = 0$	1	1

Back to Example

To illustrate, return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta\; g(a, \omega) + \gamma\; \pi(a, \omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use
$\omega = 1$	$p^*(1,1) = \beta + 1 - \varepsilon$ $p^*(1,2) = 0$	0	1	$s^L \longrightarrow a = 1$
$\omega = 2$	$p^*(2,1) = -(1 - \varepsilon)$ $p^*(2,2) = 0$	1	1	

Back to Example

To illustrate, return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use	Welfare
$\omega = 1$	$p^*(1,1) = \beta + 1 - \varepsilon$ $p^*(1,2) = 0$	0	$1 \longrightarrow s^L$	$\longrightarrow a = 1$	$1 + \beta + \varepsilon$
$\omega = 2$	$p^*(2,1) = -(1-\varepsilon)$ $p^*(2,2) = 0$	1	$1 \nearrow s^L$		

Back to Example

To illustrate, return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$$

Record Type	Prices	Records Kept	Platform's Database	Use	Welfare
$\omega = 1$	$p^*(1,1) = \beta + 1 - \varepsilon$ $p^*(1,2) = 0$	0	$1 \longrightarrow s^L$	$s^L \longrightarrow a = 1$	$1 + \beta + \varepsilon$
$\omega = 2$	$p^*(2,1) = -(1 - \varepsilon)$ $p^*(2,2) = 0$	1	$1 \nearrow s^L$		

Comments

- High-type consumers subsidize the platform to acquire low-type consumers data
- Previously this was not an equilibrium. Why?

How Realistic?

How Realistic?

It captures a qualitative feature of recent privacy-protection policies

- EU's GDPR: "*The* ***specific purposes*** *for which personal data are processed should be explicit and determined at the time of the collection of the personal data*"

How Realistic?

It captures a qualitative feature of recent privacy-protection policies

- EU's GDPR: "*The* ***specific purposes*** *for which personal data are processed should be explicit and determined at the time of the collection of the personal data*"

However, number of markets required is unrealistically large

How Realistic?

It captures a qualitative feature of recent privacy-protection policies

- EU's GDPR: "*The **specific purposes** for which personal data are processed should be explicit and determined at the time of the collection of the personal data*"

However, number of markets required is unrealistically large

Open Questions:

1. Intermediate solutions, partial decentralization?
2. "Non-market" solutions: Data Unions?

to conclusions

general model

Model 1/3

There are I competing **platforms**

Model 1/3

There are I competing **platforms**

A **merchant** is active on platform i and sells product for a_i

Model 1/3

There are I competing **platforms**

A **merchant** is active on platform i and sells product for a_i

Consumers has preference type ω

Model 1/3

There are I competing **platforms**

A **merchant** is active on platform i and sells product for a_i

Consumers has preference type ω

Type-ω consumers are of mass $\bar{q}(\omega) \geq 0$

Model 1/3

There are I competing **platforms**

A **merchant** is active on platform i and sells product for a_i

Consumers has preference type ω

Type-ω consumers are of mass $\bar{q}(\omega) \geq 0$

When a ω-consumer transacts with platform i's vendor, payoffs realize

Model 1/3

There are I competing **platforms**

A **merchant** is active on platform i and sells product for a_i

Consumers has preference type ω

Type-ω consumers are of mass $\bar{q}(\omega) \geq 0$

When a ω-consumer transacts with platform i's vendor, payoffs realize

$$\begin{aligned} \text{consumer:} &\quad g_i(a_i, \omega) \\ \text{merchant:} &\quad \pi_i(a_i, \omega) \\ \text{platform:} &\quad u_i(a_i, \omega) \end{aligned}$$

Model 2/3: Data Records

Each consumer is uniquely identified by a **data record**, which provides:

1. Info about this consumer's WTP ω

Model 2/3: Data Records

Each consumer is uniquely identified by a **data record**, which provides:

1. Info about this consumer's WTP ω (e.g., age, gender, edu, etc.)

Model 2/3: Data Records

Each consumer is uniquely identified by a **data record**, which provides:

1. Info about this consumer's WTP ω (e.g., age, gender, edu, etc.)
2. Exclusive access to this consumer

Model 2/3: Data Records

Each consumer is uniquely identified by a **data record**, which provides:

1. Info about this consumer's WTP ω (e.g., age, gender, edu, etc.)
2. Exclusive access to this consumer (e.g., email, IP address, etc.)

Model 2/3: Data Records

Each consumer is uniquely identified by a **data record**, which provides:

1. Info about this consumer's WTP ω (e.g., age, gender, edu, etc.)
2. Exclusive access to this consumer (e.g., email, IP address, etc.)

Exclusivity is key: Data record is rival good

Model 2/3: Data Records

Each consumer is uniquely identified by a **data record**, which provides:

1. Info about this consumer's WTP ω (e.g., age, gender, edu, etc.)
2. Exclusive access to this consumer (e.g., email, IP address, etc.)

Exclusivity is key: Data record is rival good

Model 2/3: Data Records

Each consumer is uniquely identified by a **data record**, which provides:

1. Info about this consumer's WTP ω (e.g., age, gender, edu, etc.)
2. Exclusive access to this consumer (e.g., email, IP address, etc.)

Exclusivity is key: Data record is rival good

Model 2/3: Data Records

Each consumer is uniquely identified by a **data record**, which provides:

1. Info about this consumer's WTP ω (e.g., age, gender, edu, etc.)
2. Exclusive access to this consumer (e.g., email, IP address, etc.)

Exclusivity is key: Data record is rival good

Model 2/3: Data Records

Each consumer is uniquely identified by a **data record**, which provides:

1. Info about this consumer's WTP ω (e.g., age, gender, edu, etc.)
2. Exclusive access to this consumer (e.g., email, IP address, etc.)

Exclusivity is key: Data record is rival good

Model 2/3: Data Records

Each consumer is uniquely identified by a **data record**, which provides:

1. Info about this consumer's WTP ω (e.g., age, gender, edu, etc.)
2. Exclusive access to this consumer (e.g., email, IP address, etc.)

Exclusivity is key: Data record is rival good

A collection of data records is called a **database**: denoted $q_i \in \mathbb{R}_+^{\Omega}$

Model 3/3: Platforms

Suppose platform i has acquired database q_i

Model 3/3: Platforms

Suppose platform i has acquired database q_i

i.e., i has exclusive access to consumers whose records belong to q_i

Model 3/3: Platforms

Suppose platform i has acquired database q_i

i.e., i has exclusive access to consumers whose records belong to q_i

Platform i is an **information designer**: It sends a signal about each consumers's ω to its vendor to influence his price a

Model 3/3: Platforms

Suppose platform i has acquired database q_i

i.e., i has exclusive access to consumers whose records belong to q_i

Platform i is an **information designer**: It sends a signal about each consumers's ω to its vendor to influence his price a

$$
\begin{aligned}
\max_{x_i : A \times \Omega \to \mathbb{R}_+} \quad & \sum_{\omega, a} u_i(a,\omega) x_i(a,\omega) & \\
\text{such that:} \quad & \sum_{\omega} \big(\pi_i(a,\omega) - \pi_i(\hat{a},\omega)\big) x_i(a,\omega) \geq 0 & \forall\, a, \hat{a} \in A \\
& \sum_{a} x_i(a,\omega) = q_i(\omega) & \forall\, \omega \in \Omega
\end{aligned}
$$

Model 3/3: Platforms

Suppose platform i has acquired database q_i

i.e., i has exclusive access to consumers whose records belong to q_i

Platform i is an **information designer**: It sends a signal about each consumers's ω to its vendor to influence his price a

$$\max_{x_i : A \times \Omega \to \mathbb{R}_+} \quad \sum_{\omega, a} u_i(a,\omega) x_i(a,\omega)$$

$$\text{such that:} \quad \sum_{\omega} \big(\pi_i(a,\omega) - \pi_i(\hat{a},\omega)\big) x_i(a,\omega) \geq 0 \qquad \forall\, a, \hat{a} \in A$$

$$\sum_{a} x_i(a,\omega) = q_i(\omega) \qquad \forall\, \omega \in \Omega$$

Model 3/3: Platforms

Suppose platform i has acquired database q_i

i.e., i has exclusive access to consumers whose records belong to q_i

Platform i is an **information designer**: It sends a signal about each consumers's ω to its vendor to influence his price a

$$\overbrace{U_i(q_i)}^{\text{platform's exptd payoff}} = \max_{x_i : A \times \Omega \to \mathbb{R}_+} \sum_{\omega, a} u_i(a,\omega) x_i(a,\omega)$$

$$\text{such that:} \quad \sum_{\omega} \big(\pi_i(a,\omega) - \pi_i(\hat{a},\omega)\big) x_i(a,\omega) \geq 0 \qquad \forall\, a, \hat{a} \in A$$

$$\sum_{a} x_i(a,\omega) = q_i(\omega) \qquad \forall\, \omega \in \Omega$$

Model 3/3: Platforms

Suppose platform i has acquired database q_i

i.e., i has exclusive access to consumers whose records belong to q_i

Platform i is an **information designer**: It sends a signal about each consumers's ω to its vendor to influence his price a

$$\overbrace{U_i(q_i)}^{\text{platform's exptd payoff}} = \max_{x_i : A\times\Omega\to\mathbb{R}_+} \sum_{\omega,a} u_i(a,\omega)x_i(a,\omega)$$

$$\text{such that:} \quad \sum_{\omega} \big(\pi_i(a,\omega) - \pi_i(\hat{a},\omega)\big)x_i(a,\omega) \geq 0 \qquad \forall\, a,\hat{a}\in A$$

$$\sum_{a} x_i(a,\omega) = q_i(\omega) \qquad \forall\, \omega\in\Omega$$

Denote by $x^*_{q_i}$ a solution (note: it depends on the entire database)

Suppose platform i has acquired database q_i

i.e., i has exclusive access to consumers whose records belong to q_i

Platform i is an **information designer**: It sends a signal about each consumers's ω to its vendor to influence his price a

$$\overbrace{U_i(q_i)}^{\text{platform's exptd payoff}} = \max_{x_i : A \times \Omega \to \mathbb{R}_+} \sum_{\omega, a} u_i(a, \omega) x_i(a, \omega)$$

such that:

$$\sum_{\omega} \big(\pi_i(a, \omega) - \pi_i(\hat{a}, \omega)\big) x_i(a, \omega) \geq 0 \qquad \forall\, a, \hat{a} \in A$$

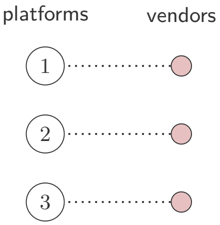
$$\sum_{a} x_i(a, \omega) = q_i(\omega) \qquad \forall\, \omega \in \Omega$$

Denote by $x^*_{q_i}$ a solution (note: it depends on the entire database)

Summing Up

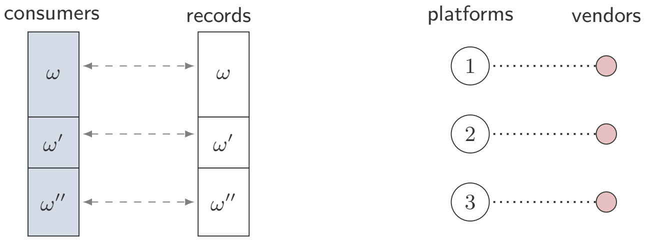
platforms

Summing Up

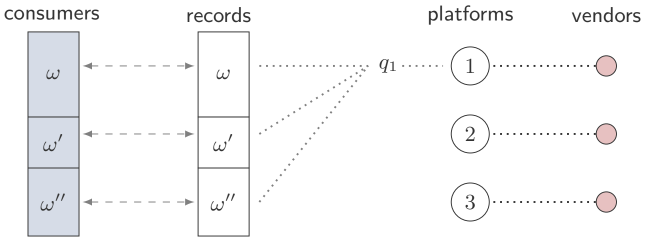


Summing Up

Summing Up


consumers
records
ω
ω'
ω''
ω
ω'
ω''
platforms
vendors
1
2
3

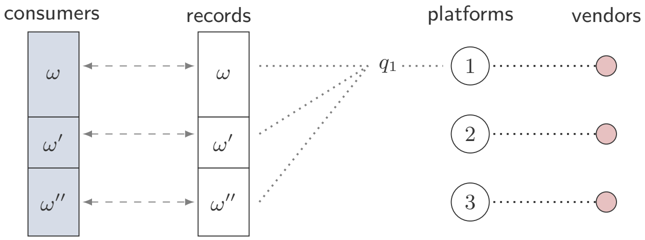
Summing Up



Timing:

1. Platform i acquires q_i in a **competitive market** for data records

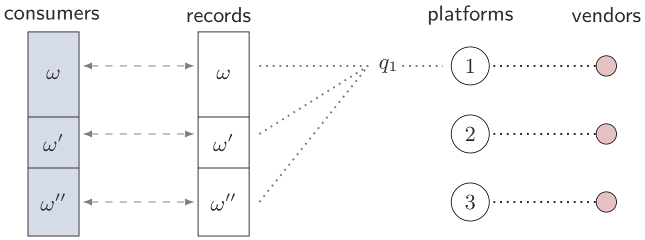
Summing Up



Timing:

1. Platform i acquires q_i in a **competitive market** for data records
2. Given database q_i, platform i chooses $x^*_{q_i}$

Summing Up



Timing:

1. Platform i acquires q_i in a **competitive market** for data records
2. Given database q_i, platform i chooses $x^*_{q_i}$
3. Payoffs realize

results

Three Economies

We study the welfare properties of three different economies:

$\mathcal{E}_1$ An economy with expropriation

Platforms own consumers data and can trade

$\mathcal{E}_2$ An economy with data ownership

Consumers own their data and can trade

$\mathcal{E}_3$ An economy with data ownership and Lindhal prices

Data are priced conditional on how it is used

$\mathcal{E}_1$ – An Economy with Expropriation

In this economy:

- Consumers "expropriated" of their records: no control, imperfect compns
- Platforms trade records among each other, taking prices as given

$\mathcal{E}_1$ – An Economy with Expropriation

In this economy:

- Consumers "expropriated" of their records: no control, imperfect compns
- Platforms trade records among each other, taking prices as given

Definition. Equilibrium in $\mathcal{E}_1$

Prices $p^* \in \mathbb{R}^\Omega$ and a feasible data allocation $q^* \in \mathbb{R}_+^{\Omega \times I}$ are an equilibrium of $\mathcal{E}_1$ if:

1. Platforms maximize given prices $\quad q_i^* \in \arg\max_{q_i} U_i(q_i) - \sum_\omega p^*(\omega) q_i(\omega)$
2. All markets clear $\quad$ for all ω, $p^*(\omega)\left(\bar{q}(\omega) - \sum_i q_i^*(\omega)\right) = 0$

$\mathcal{E}_1$ – An Economy with Expropriation

Platform i's payoff depends only on q_i, not on q_j (exclusivity)

$\mathcal{E}_1$ – An Economy with Expropriation

Platform i's payoff depends only on q_i, not on q_j (exclusivity)

Proposition. Equilibrium Characterization in $\mathcal{E}_1$

Equilibria of $\mathcal{E}_1$ exist and maximize the sum of platforms' payoffs

Every platform-optimal allocation can be supported as an equilibrium of $\mathcal{E}_1$

Equilibria are solutions of a grand-information design (LP) problem

$\mathcal{E}_1$ – An Economy with Expropriation

Platform i's payoff depends only on q_i, not on q_j (exclusivity)

Proposition. Equilibrium Characterization in $\mathcal{E}_1$

Equilibria of $\mathcal{E}_1$ exist and maximize the sum of platforms' payoffs

Every platform-optimal allocation can be supported as an equilibrium of $\mathcal{E}_1$

Equilibria are solutions of a grand-information design (LP) problem

$\mathcal{E}_1$ – An Economy with Expropriation

Platform i's payoff depends only on q_i, not on q_j (exclusivity)

Proposition. Equilibrium Characterization in $\mathcal{E}_1$

Equilibria of $\mathcal{E}_1$ exist and maximize the sum of platforms' payoffs

Every platform-optimal allocation can be supported as an equilibrium of $\mathcal{E}_1$

Equilibria are solutions of a grand-information design (LP) problem

$\mathcal{E}_2$ – An Economy with Data Ownership

$\mathcal{E}_2$ is the economy we analyzed earlier

$\mathcal{E}_2$ – An Economy with Data Ownership

$\mathcal{E}_2$ is the economy we analyzed earlier

Equilibrium in $\mathcal{E}_2$:

Prices $p^* \in \mathbb{R}^{\Omega}$, data allocation $q^* \in \mathbb{R}_+^{\Omega \times (I+1)}$, consumers' decisions $\alpha^* \in (\Delta(I))^{\Omega}$ are an equilibrium if:

1. Given p^*, database q_i^* solves platform i's problem
$$q_i^* \in \arg\max_{q_i} U_i(q_i) - \sum_{\omega} p^*(\omega) q_i(\omega)$$
2. Given p^* and q^*, $\alpha^*(\omega)$ solves ω-consumer's problem
$$\alpha^*(\omega) \in \arg \max_{\alpha(\omega) \in \Delta(I)} (1 - \alpha(0|\omega)) r(\omega) + \sum_i \alpha(i|\omega) \left(p^*(\omega) + \mathbb{E}_{q_i^*}(g_i(a_i, \omega)) \right)$$
3. Markets clear
$$q_i^*(\omega) = \alpha^*(i|\omega) \bar{q}(\omega), \quad \forall \omega, i$$

$\mathcal{E}_2$ – An Economy with Data Ownership

What We Know:

- Equilibrium *can* be inefficient $\rightsquigarrow$ our leading example
- Sufficient conditions for efficiency:

Proposition. No-Intermediation Case

When $u_i = \pi_i$ for all i, equilibria in $\mathcal{E}_2$ exist and are efficient

$\mathcal{E}_2$ – An Economy with Data Ownership

What We Know:

- Equilibrium *can* be inefficient $\rightsquigarrow$ our leading example
- Sufficient conditions for efficiency:

Proposition. No-Intermediation Case

When $u_i = \pi_i$ for all i, equilibria in $\mathcal{E}_2$ exist and are efficient

What We Don't Know (yet):

- Sufficient conditions for inefficiency beyond examples?
- Sufficient conditions for existence in the intermediation case?

$\mathcal{E}_3$ – An Economy with Lindhal prices

How can we fix inefficiencies discussed so far?

$\mathcal{E}_3$ – An Economy with Lindhal prices

How can we fix inefficiencies discussed so far?

We enrich our economy by opening "more complete" markets

following e.g. Arrow 69, Laffont 78

- Consumers can sell record for a **specific purpose** (i.e. an action a_i)
- A richer price system: prices $p_i(\omega, a_i)$ depend on record type, on platform identity, and on intended use a_i

$\mathcal{E}_3$ – An Economy with Lindhal prices

How can we fix inefficiencies discussed so far?

We enrich our economy by opening "more complete" markets

following e.g. Arrow 69, Laffont 78

- Consumers can sell record for a **specific purpose** (i.e. an action a_i)
- A richer price system: prices $p_i(\omega, a_i)$ depend on record type, on platform identity, and on intended use a_i

This presumes data use is contractible

$\mathcal{E}_3$ – An Economy with Lindhal prices

How can we fix inefficiencies discussed so far?

We enrich our economy by opening "more complete" markets

following e.g. Arrow 69, Laffont 78

- Consumers can sell record for a **specific purpose** (i.e. an action a_i)
- A richer price system: prices $p_i(\omega, a_i)$ depend on record type, on platform identity, and on intended use a_i

This presumes data use is contractible

Proposition. Equilibrium Characterization in $\mathcal{E}_3$

Equilibria in $\mathcal{E}_3$ exist and are (first-best) efficient.

Every (first-best) efficient data allocation can be supported in an eqm

Equilibria are solutions of a grand-information design (LP) problem

Back to Example

Illustrate functioning of economy $\mathcal{E}_3$ with our earlier example

Return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$$

Record Type
$\omega = 1$
$\omega = 2$

Back to Example

Illustrate functioning of economy $\mathcal{E}_3$ with our earlier example

Return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$$

Record Type

$\omega = 1$

$\omega = 2$

Back to Example

Illustrate functioning of economy $\mathcal{E}_3$ with our earlier example

Return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$$

Record Type	Prices
$\omega = 1$	$p^*(1)$
$\omega = 2$	$p^*(2)$

Back to Example

Illustrate functioning of economy $\mathcal{E}_3$ with our earlier example

Return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	Prices
$\omega = 1$	$p^*(1, a=1)$ $p^*(1, a=2)$
$\omega = 2$	$p^*(2, a=1)$ $p^*(2, a=2)$

Back to Example

Illustrate functioning of economy $\mathcal{E}_3$ with our earlier example

Return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$$

Record Type	**Prices**
$\omega = 1$	$p^*(1, a = 1)$ $p^*(1, a = 2)$
$\omega = 2$	$p^*(2, a = 1)$ $p^*(2, a = 2)$

Back to Example

Illustrate functioning of economy $\mathcal{E}_3$ with our earlier example

Return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	**Prices**
$\omega = 1$	$p^*(1,1) = \beta + 1 - \varepsilon$
	$p^*(1,2) = 0$
$\omega = 2$	$p^*(2,1) = -(1-\varepsilon)$
	$p^*(2,2) = 0$

Back to Example

Illustrate functioning of economy $\mathcal{E}_3$ with our earlier example

Return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$$

Record Type	Prices	Records Kept	Records Used
$\omega = 1$	$p^*(1,1) = \beta + 1 - \varepsilon$ $p^*(1,2) = 0$	0	$1 \rightsquigarrow a = 1$
$\omega = 2$	$p^*(2,1) = -(1 - \varepsilon)$ $p^*(2,2) = 0$	1	$1 \rightsquigarrow a = 1$

Back to Example

Illustrate functioning of economy $\mathcal{E}_3$ with our earlier example

Return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	**Prices**	**Records Kept**	**Records Used**	**Consumers Welfare**
$\omega = 1$	$p^*(1,1) = \beta + 1 - \varepsilon$ $p^*(1,2) = 0$	0	$1 \rightsquigarrow a = 1$	$1 + \beta + \varepsilon$
$\omega = 2$	$p^*(2,1) = -(1 - \varepsilon)$ $p^*(2,2) = 0$	1	$1 \rightsquigarrow a = 1$	

Back to Example

Illustrate functioning of economy $\mathcal{E}_3$ with our earlier example

Return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a,\omega) = \beta\, g(a,\omega) + \gamma\, \pi(a,\omega)$$

Record Type	Prices	Records Kept	Records Used	Consumers Welfare
$\omega = 1$	$p^*(1,1) = \beta + 1 - \varepsilon$ $p^*(1,2) = 0$	0	$1 \rightsquigarrow a = 1$	$1 + \beta + \varepsilon$
$\omega = 2$	$p^*(2,1) = -(1-\varepsilon)$ $p^*(2,2) = 0$	1	$1 \rightsquigarrow a = 1$	

Back to Example

Illustrate functioning of economy $\mathcal{E}_3$ with our earlier example

Return to case of market unravelling ($\beta < \epsilon$, $\gamma = 0$):

$$u(a, \omega) = \beta\, g(a, \omega) + \gamma\, \pi(a, \omega)$$

Record Type	Prices	Records Kept	Records Used	Consumers Welfare
$\omega = 1$	$p^*(1,1) = \beta + 1 - \varepsilon$ $p^*(1,2) = 0$	0	$1 \rightsquigarrow a = 1$	$1 + \beta + \varepsilon$
$\omega = 2$	$p^*(2,1) = -(1 - \varepsilon)$ $p^*(2,2) = 0$	1	$1 \rightsquigarrow a = 1$	

Comments

- High-type consumers subsidize the platform to acquire low-type consumers data
- Previously this was not an equilibrium. Why?

How Realistic is $\mathcal{E}_3$?

How Realistic is $\mathcal{E}_3$?

It captures a qualitative feature of recent privacy-protection policies

- EU's GDPR: "*The* ***specific purposes*** *for which personal data are processed should be explicit and determined at the time of the collection of the personal data*"

How Realistic is $\mathcal{E}_3$?

It captures a qualitative feature of recent privacy-protection policies

- EU's GDPR: "*The **specific purposes** for which personal data are processed should be explicit and determined at the time of the collection of the personal data*"

However, number of markets required is unrealistically large

How Realistic is $\mathcal{E}_3$?

It captures a qualitative feature of recent privacy-protection policies

- EU's GDPR: "*The* ***specific purposes*** *for which personal data are processed should be explicit and determined at the time of the collection of the personal data*"

However, number of markets required is unrealistically large

Open Questions:

1. Intermediate solutions, partial decentralization?
2. "Non-market" solutions: Data Unions?

conclusions

Summary

1. We introduce framework to study competitive markets for personal data and their equilibria
   - Rather general setting: many platforms, many merchants, arbitrary objectives, partially informative records, multiple types
2. We identify a novel externality that can make these markets inefficient
   - The way platforms withhold information creates externalities that can lead to market failures
3. We discuss possible remedies and their limits

thank you