Competitive Markets for Personal Data

Simone Galperti   Tianhao Liu   Jacopo Perego

October 2024

Motivation

Consumers supply a crucial input for modern economy: their **personal data**

Yet, they often have **limited control** over how and by whom their data is used:

- This may lead to inefficiencies and inequality

New legislation gives consumers more control over their data (GDPR, CCPA, ...)

- Lays foundations upon which **data markets** could emerge

What properties would these markets have, and how should they be designed to promote desirable outcomes?

This Paper

Model. A stylized **competitive data market** where

- Consumers own their data and can sell it to intermediaries
- Intermediaries use the acquired consumer data to provide information to merchants and interact them with the consumers

This Paper

Model. A stylized **competitive data market** where

- Consumers own their data and can sell it to intermediaries
- Intermediaries use the acquired consumer data to provide information to merchants and interact them with the consumers

Main Results

1. We identify **novel inefficiency** and show it is generated by <u>how intermediaries use data</u>

- If full disclosure $\Rightarrow$ No Externalities $\Rightarrow$ Efficiency
- If some pooling $\Rightarrow$ Externalities $\Rightarrow$ Inefficiencies

This Paper

Model. A stylized **competitive data market** where

- Consumers own their data and can sell it to intermediaries
- Intermediaries use the acquired consumer data to provide information to merchants and interact them with the consumers

Main Results

1. We identify **novel inefficiency** and show it is generated by how intermediaries use data

- If full disclosure $\Rightarrow$ No Externalities $\Rightarrow$ Efficiency
- If some pooling $\Rightarrow$ Externalities $\Rightarrow$ Inefficiencies

2. Propose three solutions to this market failure:

- Data unions; Data taxes; "Lindahl" pricing for the data

Related Work

Exploit progress in info-design to microfound components of data economy:

- How does intermediary use the data? (Bergmann-Morris '19, Kamenica '19)
- What's the value of data? (GLP '23)

Related Work

Exploit progress in info-design to microfound components of data economy:

- How does intermediary use the data? (Bergmann-Morris '19, Kamenica '19)
- What's the value of data? (GLP '23)

Limited knowledge about **data markets**:

- "Correlation" externality Choi et al ('19), BBG ('22), Acemoglu et al. ('22)
- Our inefficiency not due to exogenous correlation, but to platform's role as info intermediary (indeed, no intermediation $\Rightarrow$ no externality)

Related Work

Exploit progress in info-design to microfound components of data economy:

- How does intermediary use the data? (Bergmann-Morris '19, Kamenica '19)
- What's the value of data? (GLP '23)

Limited knowledge about **data markets**:

- "Correlation" externality Choi et al ('19), BBG ('22), Acemoglu et al. ('22)
- Our inefficiency not due to exogenous correlation, but to platform's role as info intermediary (indeed, no intermediation $\Rightarrow$ no externality)

More broadly, we contribute to the growing literature on the economics of platforms, data, & privacy Bergemann and Ottaviani '21, Baley and Veldkamp '24

model

Basic Ingredients

A unit mass of consumers, one (representative) platform, one merchant

Basic Ingredients

A unit mass of consumers, one (representative) platform, one merchant

Each **consumer** has unit demand for merchant's product

Her preference type is denoted by $\omega \in \Omega$ (finite) (e.g., WTP, location, ...)

Basic Ingredients

A unit mass of consumers, one (representative) platform, one merchant

Each **consumer** has unit demand for merchant's product

Her preference type is denoted by $\omega \in \Omega$ (finite) (e.g., WTP, location, ...)

Let $\bar{q} \in \Delta(\Omega)$ be the distribution of ω in the population

Basic Ingredients

A unit mass of consumers, one (representative) platform, one merchant

Each **consumer** has unit demand for merchant's product

Her preference type is denoted by $\omega \in \Omega$ (finite) (e.g., WTP, location, ...)

Let $\bar{q} \in \Delta(\Omega)$ be the distribution of ω in the population

Each consumer owns a **data record** that fully reveals her corresponding ω

Basic Ingredients

A unit mass of consumers, one (representative) platform, one merchant

Each **consumer** has unit demand for merchant's product

Her preference type is denoted by $\omega \in \Omega$ (finite) (e.g., WTP, location, ...)

Let $\bar{q} \in \Delta(\Omega)$ be the distribution of ω in the population

Each consumer owns a **data record** that fully reveals her corresponding ω

Two periods: 1. Data markets are open 2. Product market is open

Period 1: Competitive Data Markets

Period 1: Competitive Data Markets

The consumers and the platform trade data records at prices $p = (p(\omega))_{\omega \in \Omega}$, which they take as given

Period 1: Competitive Data Markets

The consumers and the platform trade data records at prices $p = (p(\omega))_{\omega \in \Omega}$, which they take as given

The demand side:

– Platform demands **database** $q = (q(\omega))_{\omega \in \Omega}$, for which it pays $\sum_{\omega} q(\omega)p(\omega)$

Period 1: Competitive Data Markets

The consumers and the platform trade data records at prices $p = (p(\omega))_{\omega \in \Omega}$, which they take as given

The demand side:

– Platform demands **database** $q = (q(\omega))_{\omega \in \Omega}$, for which it pays $\sum_{\omega} q(\omega)p(\omega)$

The supply side:

– If a type-ω consumer sells her record to the platform, she is paid $p(\omega)$ and is later intermediated with merchant

Notation: $z(\omega) \in [0, 1]$ probability type-ω consumer sells her record

Period 1: Competitive Data Markets

The consumers and the platform trade data records at prices $p = (p(\omega))_{\omega \in \Omega}$, which they take as given

The demand side:

– Platform demands **database** $q = (q(\omega))_{\omega \in \Omega}$, for which it pays $\sum_\omega q(\omega)p(\omega)$

The supply side:

– If a type-ω consumer sells her record to the platform, she is paid $p(\omega)$ and is later intermediated with merchant

Notation: $z(\omega) \in [0, 1]$ probability type-ω consumer sells her record

– If type-ω consumer doesn't sell her record, she gets reservation utility $\bar{r}$

Period 2: Product Market

Given acquired database q, platform acts as **information designer**

- It sends signal to merchant about each consumer in its database
- Given signal, the merchant chooses an action $a \in A$ (finite)
- Together, a and ω determine consumer final purchase decision (left implicit)

Period 2: Product Market

Given acquired database q, platform acts as **information designer**

- It sends signal to merchant about each consumer in its database
- Given signal, the merchant chooses an action $a \in A$ (finite)
- Together, a and ω determine consumer final purchase decision (left implicit)

The database q constitutes the "common prior" in the info design problem

Notation for payoffs in period 2:

Consumer's:	$u(a, \omega)$	e.g., trading surplus
Merchant's:	$\pi(a, \omega)$	e.g., profits
Platform's:	$v(a, \omega)$	

Period 2: Product Market

Given acquired database q, platform acts as **information designer**

- It sends signal to merchant about each consumer in its database
- Given signal, the merchant chooses an action $a \in A$ (finite)
- Together, a and ω determine consumer final purchase decision (left implicit)

The database q constitutes the "common prior" in the info design problem

Remark: Info design problem equivalent to a linear program: (BM '16)

$$V(q) = \max_{x:\Omega \to \Delta(A)} \sum_{\omega, a} v(a,\omega) x(a|\omega) q(\omega)$$

$$\text{s.t. } \forall a, a' \colon \sum_{\omega} \Big(\pi(a,\omega) - \pi(a',\omega)\Big) x(a|\omega) q(\omega) \geq 0$$

Period 2: Product Market

Given acquired database q, platform acts as **information designer**

- It sends signal to merchant about each consumer in its database
- Given signal, the merchant chooses an action $a \in A$ (finite)
- Together, a and ω determine consumer final purchase decision (left implicit)

The database q constitutes the "common prior" in the info design problem

Remark: Info design problem equivalent to a linear program: (BM '16)

$$V(q) = \max_{x:\Omega \to \Delta(A)} \sum_{\omega,a} v(a,\omega)x(a|\omega)q(\omega)$$

$$\text{s.t. } \forall a, a': \sum_{\omega} \Big(\pi(a,\omega) - \pi(a',\omega)\Big)x(a|\omega)q(\omega) \geq 0$$

Denote its solutions by $\mathcal{X}(q)$ (standard ID problem, but with endogenous prior)

Equilibrium

Definition

A profile (p^*, z^*, q^*, x^*)

Equilibrium

Definition

A profile (p^*, z^*, q^*, x^*) is an equilibrium of the economy if:

Equilibrium

Definition

A profile (p^*, z^*, q^*, x^*) is an equilibrium of the economy if:

(a). Given p^*, q^* solves the platform's problem in the first period, i.e.,

$$q^* \in \arg\max_{q \in \mathbb{R}_+^\Omega} V(q) - \sum p^*(\omega) q(\omega)$$

Equilibrium

Definition

A profile (p^*, z^*, q^*, x^*) is an equilibrium of the economy if:

(a). Given p^*, q^* solves the platform's problem in the first period, i.e.,

$$q^* \in \arg\max_{q \in \mathbb{R}_+^\Omega} V(q) - \sum p^*(\omega) q(\omega)$$

(b). Given q^*, x^* solves platform's info-design problem, i.e., $x^* \in \mathcal{X}(q)$

Equilibrium

Definition

A profile (p^*, z^*, q^*, x^*) is an equilibrium of the economy if:

(a). Given p^*, q^* solves the platform's problem in the first period, i.e.,

$$q^* \in \arg\max_{q \in \mathbb{R}_+^{\Omega}} V(q) - \sum p^*(\omega) q(\omega)$$

(b). Given q^*, x^* solves platform's info-design problem, i.e., $x^* \in \mathcal{X}(q)$

(c). Given p^* and x^*, ζ^* solves consumers' problem, i.e.,

$$z^*(\omega) \in \arg\max_{\zeta \in [0,1]} \zeta \Big(p^*(\omega) + \underbrace{\sum_a x^*(a|\omega) u(a, \omega)}_{U(\omega, x^*)} \Big) + (1 - \zeta)\bar{r}$$

Equilibrium

Definition

A profile (p^*, z^*, q^*, x^*) is an equilibrium of the economy if:

(a). Given p^*, q^* solves the platform's problem in the first period, i.e.,

$$q^* \in \arg\max_{q \in \mathbb{R}_+^\Omega} V(q) - \sum p^*(\omega) q(\omega)$$

(b). Given q^*, x^* solves platform's info-design problem, i.e., $x^* \in \mathcal{X}(q)$

(c). Given p^* and x^*, ζ^* solves consumers' problem, i.e.,

$$z^*(\omega) \in \arg\max_{\zeta \in [0,1]} \zeta \left(p^*(\omega) + \underbrace{\sum_a x^*(a|\omega) u(a,\omega)}_{U(\omega, x^*)} \right) + (1 - \zeta)\bar{r}$$

(d). Data markets clear, i.e. $\quad q^*(\omega) = \zeta^*(\omega)\bar{q}(\omega) \qquad \forall \omega$

discussion

Discussion

Can accomodate much larger class of **information-intermediation problems**:

- Multiple agents (e.g., competing merchants)
- An arbitrary downstream game (e.g., a second-price auctions, hotelling)
- More than information design (e.g., platform takes a contractible action)

Leading applications: online marketplaces and advertisement auctions

Discussion

Can accomodate much larger class of **information-intermediation problems**:

- Multiple agents (e.g., competing merchants)
- An arbitrary downstream game (e.g., a second-price auctions, hotelling)
- More than information design (e.g., platform takes a contractible action)

Leading applications: online marketplaces and advertisement auctions

Substantive assumptions we made:

- A competitive data market
- Platform is a "gate keeper" alt see BB '23
- A data record combines "access" and information alt see ALV '22

analysis

Mantained Assumption

We focus on economies that are “regular:”

Definition

An economy is **regular** if $\mathcal{X}(q)$ is a singleton for almost all q

This is an assumption on v and π

- Regular economies are generic in the space of economies

Efficiency of the Data Market

The **data market** in which platform and consumers participate is competitive:

Is it efficient? I.e., does it maximize the welfare of its participants?

Efficiency of the Data Market

The **data market** in which platform and consumers participate is competitive:

Is it efficient? I.e., does it maximize the welfare of its participants?

Definition

An allocation (q°, x°) is **constrained efficient** if it solves

$$\begin{aligned} W^\circ = \quad & \max_{q,x} \quad V(q) + \sum_\omega q(\omega) U(\omega, x) + \sum_\omega \Big(\bar{q}(\omega) - q(\omega)\Big)\bar{r} \\ & \text{s.t.} \quad q \leq \bar{q} \ \text{ and } \ x \text{ solves platform' problem } \mathcal{P}_q \end{aligned}$$

Efficiency of the Data Market

The **data market** in which platform and consumers participate is competitive:

Is it efficient? I.e., does it maximize the welfare of its participants?

Definition

An allocation (q°, x°) is **constrained efficient** if it solves

$$W^\circ = \max_{q,x} \quad V(q) + \sum_\omega q(\omega) U(\omega, x) + \sum_\omega \Big(\bar{q}(\omega) - q(\omega)\Big)\bar{r}$$

$$\text{s.t.} \quad q \leq \bar{q} \ \text{ and } \ x \text{ solves platform' problem } \mathcal{P}_q$$

To illustrate failure in data market, this less-demanding benchmark is desirable

(We also study "social" welfare and "unconstrained" efficiency discussion)

Efficiency of the Data Market

To characterize the efficiency of equilibria I will compare:

- How records are allocated by the market
- How records are allocated by the social planner

Equilibrium

Fix an equilibrium (p^*, z^*, q^*, x^*)

The **private gain** for a type-ω consumer when she sells her record is

$$G^*(\omega) \triangleq p^*(\omega) + U(\omega, x^*)$$

Equilibrium

Fix an equilibrium (p^*, z^*, q^*, x^*)

The **private gain** for a type-ω consumer when she sells her record is

$$G^*(\omega) \triangleq p^*(\omega) + U(\omega, x^*)$$

In equilibrium, it must hold that $p^*(\omega) = \dfrac{\partial V(q^*)}{\partial q^*(\omega)}$ platform makes no profit

Equilibrium

Fix an equilibrium (p^*, z^*, q^*, x^*)

The **private gain** for a type-ω consumer when she sells her record is

$$G^*(\omega) \triangleq p^*(\omega) + U(\omega, x^*)$$

In equilibrium, it must hold that $p^*(\omega) = \dfrac{\partial V(q^*)}{\partial q^*(\omega)}$ platform makes no profit

Equilibrium

Fix an equilibrium (p^*, z^*, q^*, x^*)

The **private gain** for a type-ω consumer when she sells her record is

$$G^*(\omega) \triangleq \frac{\partial V(q^*)}{\partial q^*(\omega)} + U(\omega, x^*)$$

In equilibrium, it must hold that $p^*(\omega) = \dfrac{\partial V(q^*)}{\partial q^*(\omega)}$ platform makes no profit

Equilibrium

Fix an equilibrium (p^*, z^*, q^*, x^*)

The **private gain** for a type-ω consumer when she sells her record is

$$G^*(\omega) \triangleq \frac{\partial V(q^*)}{\partial q^*(\omega)} + U(\omega, x^*)$$

In equilibrium, it must hold that $p^*(\omega) = \dfrac{\partial V(q^*)}{\partial q^*(\omega)}$ platform makes no profit

Additionally, in eqm, it must hold that consumers optimize

- If $q^*(\omega) > 0$, then $G^*(\omega) \geq \bar{r}$
- If $q^*(\omega) < \bar{q}(\omega)$, then $G^*(\omega) \leq \bar{r}$

Equilibrium

Fix an equilibrium (p^*, z^*, q^*, x^*)

The **private gain** for a type-ω consumer when she sells her record is

$$G^*(\omega) \triangleq \frac{\partial V(q^*)}{\partial q^*(\omega)} + U(\omega, x^*)$$

In equilibrium, it must hold that $p^*(\omega) = \dfrac{\partial V(q^*)}{\partial q^*(\omega)}$ platform makes no profit

Additionally, in eqm, it must hold that consumers optimize

- If $q^*(\omega) > 0$, then $G^*(\omega) \geq \bar{r}$
- If $q^*(\omega) < \bar{q}(\omega)$, then $G^*(\omega) \leq \bar{r}$

Social Planner

How does a marginal increase in $q(\omega)$ affect social welfare?

Social Planner

How does a marginal increase in $q(\omega)$ affect social welfare?

A **social loss** of $\bar{r}$

Social Planner

How does a marginal increase in $q(\omega)$ affect social welfare?

A **social loss** of $\bar{r}$

What's the corresponding **social gain**?

Denote by $\Psi_q \subset \mathbb{R}_+^\Omega$ the supergradients of (a.s. a singleton)

$$V(q) + \sum_\omega q(\omega) U(\omega, x_q)$$

Social Planner

How does a marginal increase in $q(\omega)$ affect social welfare?

A **social loss** of $\bar{r}$

What's the corresponding **social gain**?

Denote by $\Psi_q \subset \mathbb{R}_+^\Omega$ the supergradients of (a.s. a singleton)

$$V(q) + \sum_\omega q(\omega) U(\omega, x_q)$$

For $\psi_q \in \Psi_q$, $\psi_q(\omega)$ captures the social gain of a marginal increase in $q(\omega)$

Constrained-Efficient Allocations

Proposition

Fix an allocation (q, x). If (q, x) is constrained efficient, x solves $\mathcal{P}_q$ and there is a $\psi \in \Psi_q$ s.t.

- If $q(\omega) > 0$, then $\psi(\omega) \geq \bar{r}$
- If $q(\omega) < \bar{q}(\omega)$, then $\psi(\omega) \leq \bar{r}$

Additionally, if $\sum_\omega q(\omega) U(\omega, x_q)$ is concave in q, such condition is also sufficient

Constrained-Efficient Allocations

Proposition

Fix an allocation (q, x). If (q, x) is constrained efficient, x solves $\mathcal{P}_q$ and there is a $\psi \in \Psi_q$ s.t.

- If $q(\omega) > 0$, then $\psi(\omega) \geq \bar{r}$
- If $q(\omega) < \bar{q}(\omega)$, then $\psi(\omega) \leq \bar{r}$

Additionally, if $\sum_\omega q(\omega) U(\omega, x_q)$ is concave in q, such condition is also sufficient

Intuition: Necessity obvious. If planner's problem is concave, "FOC" is also sufficient

Constrained-Efficient Allocations

Proposition

Fix an allocation (q, x). If (q, x) is constrained efficient, x solves $\mathcal{P}_q$ and there is a $\psi \in \Psi_q$ s.t.

- If $q(\omega) > 0$, then $\psi(\omega) \geq \bar{r}$
- If $q(\omega) < \bar{q}(\omega)$, then $\psi(\omega) \leq \bar{r}$

Additionally, if $\sum_\omega q(\omega) U(\omega, x_q)$ is concave in q, such condition is also sufficient

Intuition: Necessity obvious. If planner's problem is concave, "FOC" is also sufficient

Thus, eqm efficiency depends on alignement between G^* and ψ_{q^*}

Data Externality

The equilibrium **social benefit** of selling additional ω-record is:

$$\psi_{q^*}(\omega) \cong \frac{\partial V(q^*)}{\partial q^*(\omega)} + U(\omega, x^*) + \sum_{\omega'} q(\omega') \frac{\partial U(\omega', x^*)}{\partial q^*(\omega)}$$

Data Externality

The equilibrium **social benefit** of selling additional ω-record is:

$$\psi_{q^*}(\omega) \cong \frac{\partial V(q^*)}{\partial q^*(\omega)} + U(\omega, x^*) + \sum_{\omega'} q(\omega') \frac{\partial U(\omega', x^*)}{\partial q^*(\omega)}$$

The equilibrium **private gain** of selling an ω-record is

$$G^*(\omega) = p^*(\omega) + U(\omega, x^*)$$

Data Externality

The equilibrium **social benefit** of selling additional ω-record is:

$$\psi_{q^*}(\omega) \cong \frac{\partial V(q^*)}{\partial q^*(\omega)} + U(\omega, x^*) + \sum_{\omega'} q(\omega') \frac{\partial U(\omega', x^*)}{\partial q^*(\omega)}$$

The equilibrium **private gain** of selling an ω-record is

$$G^*(\omega) = \frac{\partial V(q^*)}{\partial q^*(\omega)} + U(\omega, x^*)$$

Data Externality

The equilibrium **social benefit** of selling additional ω-record is:

$$\psi_{q^*}(\omega) \cong \frac{\partial V(q^*)}{\partial q^*(\omega)} + U(\omega, x^*) + \underbrace{\sum_{\omega'} q(\omega') \frac{\partial U(\omega', x^*)}{\partial q^*(\omega)}}_{\substack{\text{a non-pecuniary} \\ \textbf{data externality}}}$$

The equilibrium **private gain** of selling an ω-record is

$$G^*(\omega) = \frac{\partial V(q^*)}{\partial q^*(\omega)} + U(\omega, x^*)$$

Data Externality

The equilibrium **social benefit** of selling additional ω-record is:

$$\psi_{q^*}(\omega) \cong \frac{\partial V(q^*)}{\partial q^*(\omega)} + U(\omega, x^*) + \underbrace{\sum_{\omega'} q(\omega') \frac{\partial U(\omega', x^*)}{\partial q^*(\omega)}}_{\substack{\text{a non-pecuniary} \\ \textbf{data externality}}}$$

The equilibrium **private gain** of selling an ω-record is

$$G^*(\omega) = \frac{\partial V(q^*)}{\partial q^*(\omega)} + U(\omega, x^*)$$

When selling her data, a consumer does not internalize she affects q, which changes x, which changes the payoff of other consumers $\Rightarrow$ **externality**

Full-Disclosure Mechanisms

Externality enabled by the fact that how platform uses data (x_q^*) may depend on what data it has obtained (q^*)

Full-Disclosure Mechanisms

Externality enabled by the fact that how platform uses data (x_q^*) may depend on what data it has obtained (q^*)

An important class of recommendation mechanisms:

Definition

x_q is a **full-disclosure mechanism** if it recommends only actions the merchant would choose under complete information, i.e., for all ω,

$$q(\omega)x_q(a|\omega) > 0 \qquad \text{only if} \qquad a \in \arg\max_{a \in A} \pi(a, \omega)$$

Full-Disclosure Mechanisms

Externality enabled by the fact that how platform uses data (x_q^*) may depend on what data it has obtained (q^*)

An important class of recommendation mechanisms:

Definition

x_q is a **full-disclosure mechanism** if it recommends only actions the merchant would choose under complete information, i.e., for all ω,

$$q(\omega)x_q(a|\omega) > 0 \qquad \text{only if} \qquad a \in \arg\max_{a \in A} \pi(a, \omega)$$

Efficiency Characterization

Proposition

If a database $q \in \mathbb{R}^{\Omega}_{++}$ exists at which full disclosure is optimal, all equilibria of the economy are **constrained efficient**. Morever, each equilibrium maximizes consumer welfare.

Efficiency Characterization

Proposition

If a database $q \in \mathbb{R}^{\Omega}_{++}$ exists at which full disclosure is optimal, all equilibria of the economy are **constrained efficient**. Morever, each equilibrium maximizes consumer welfare.

Corollary

Fix an equilibirum (p^*, q^*, z^*, x^*) with $q^* \in \mathbb{R}^{\Omega}_{++}$. If x^* involves full disclosure, the equilibrium is constrained efficient.

Efficiency Characterization

Proposition

If a database $q \in \mathbb{R}^{\Omega}_{++}$ exists at which full disclosure is optimal, all equilibria of the economy are **constrained efficient**. Morever, each equilibrium maximizes consumer welfare.

Corollary

Fix an equilibirum (p^*, q^*, z^*, x^*) with $q^* \in \mathbb{R}^{\Omega}_{++}$. If x^* involves full disclosure, the equilibrium is constrained efficient.

Proof Sketch

- If full disclosure is optimal at interior q, full disclosure is optimal at all q
- By regularity, there is an interior q' where full disclosure is uniquely optimal. Then, full disclosure is uniquely optimal at all q's
- Thus, $\sum_{\omega'} q(\omega') \frac{\partial U(\omega', x^*)}{\partial q^*(\omega)} = 0 \quad \Rightarrow \quad$ no data externalities
- Additionally, $\sum_{\omega} q(\omega) U(\omega, x_q)$ is linear in q. Thus, planner's "FOC" is sufficient for efficiency.
- Thus, any eqm is efficient

Bigger Picture

Main insight. The way platform uses data determines whether eqm is efficient

- Full disclosure is a sufficient condition for efficiency
- Thus, in an inefficient equilibrium platform must do some pooling

Bigger Picture

Main insight. The way platform uses data determines whether eqm is efficient

- Full disclosure is a sufficient condition for efficiency
- Thus, in an inefficient equilibrium platform must do some pooling

Thus, substantive question is: When do info intermediaries have incentives to fully disclose info with their agents? And what happens when they don't?

an application

Mediated Price Discrimination

So far, minimal assumptions on the intermediation problem

More structure needed to further characterize equilibrium (in)efficiency

We specialize setting to canonical application: Price discrimination *à la* BBM '15

Mediated Price Discrimination

So far, minimal assumptions on the intermediation problem

More structure needed to further characterize equilibrium (in)efficiency

We specialize setting to canonical application: Price discrimination *à la* BBM '15

- Let $\omega \in \mathbb{R}_{++}$ be the consumer's WTP for merchant's product
- Let a denote the merchant's price set for the product
- Players payoffs are

$$\begin{aligned} \text{Consumer's:} \quad & u(a,\omega) = \max\{\omega - a, 0\} \\ \text{Merchant's:} \quad & \pi(a,\omega) = a \, \mathbb{1}(\omega \geq a) \\ \text{Platform's:} \quad & v(a,\omega) = \gamma_u \, u(a,\omega) + \gamma_\pi \, \pi(a,\omega) \end{aligned}$$

Equilibrium Characterization

Proposition

- If $\gamma_u < \gamma_\pi$, all equilibria are constrained efficient. Consumers' welfare is maximized.
- If $\gamma_u > \gamma_\pi$, equilibria can be inefficient. In particular, an open set of $\bar{r}$'s exist such that all equilibria in the corresponding economies are inefficient

Equilibrium Characterization

Proposition

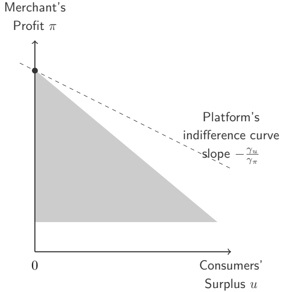
- If $\gamma_u < \gamma_\pi$, all equilibria are constrained efficient. Consumers' welfare is maximized.
- If $\gamma_u > \gamma_\pi$, equilibria can be inefficient. In particular, an open set of $\bar{r}$'s exist such that all equilibria in the corresponding economies are inefficient

Intuition for Efficiency

If $\gamma_u < \gamma_\pi$

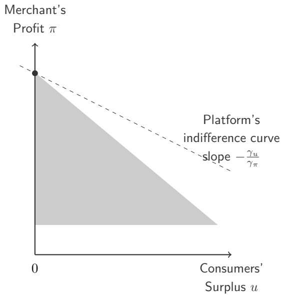
Intuition for Efficiency

If $\gamma_u < \gamma_\pi$



Intuition for Efficiency

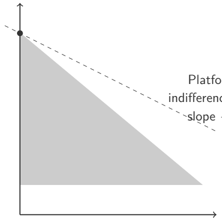
If $\gamma_u < \gamma_\pi$



- At all q, **full disclosure** is optimal
- Merchant extracts surplus from all consumers

Intuition for Efficiency

If $\gamma_u < \gamma_\pi$



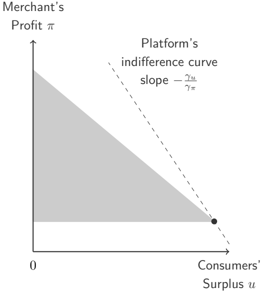
- At all q, **full disclosure** is optimal
- Merchant extracts surplus from all consumers
- Therefore, $x^*(a, \omega)$ does not depend on q
- Therefore, no externality! All equilibria are constrained efficient

Intuition for Inefficiency

If $\gamma_u > \gamma_\pi$

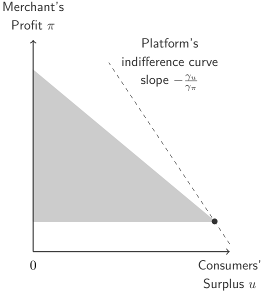
Intuition for Inefficiency

If $\gamma_u > \gamma_\pi$



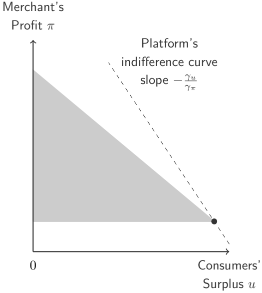
Intuition for Inefficiency

If $\gamma_u > \gamma_\pi$



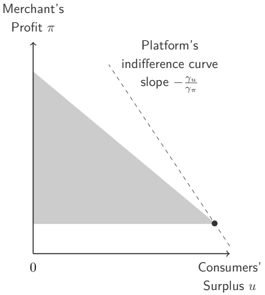
Intuition for Inefficiency

If $\gamma_u > \gamma_\pi$



Intuition for Inefficiency

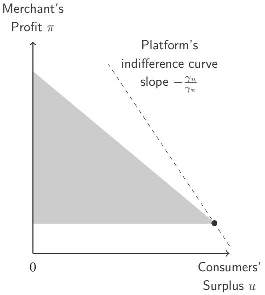
If $\gamma_u > \gamma_\pi$



- Platform **withholds information** from merchant

Intuition for Inefficiency

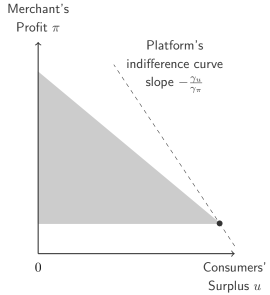
If $\gamma_u > \gamma_\pi$



- Platform **withholds information** from merchant
- Pooling types required to keep merchant indifferent at some as

Intuition for Inefficiency

If $\gamma_u > \gamma_\pi$



- Platform **withholds information** from merchant
- Pooling types required to keep merchant indifferent at some as
- Thus, x_q depends on q
- Thus, $\sum_{\omega'} q(\omega') \frac{\partial U(\omega', x^*)}{\partial q^*(\omega)}$ can be non-zero: data externality

Bigger Picture

Application shows when there is **no conflict of interest** btw platform and merchant $\Rightarrow$ full disclosure is optimal $\Rightarrow$ data market is efficient

Special case: if platform is merchant $\Rightarrow$ no intermediation

The source of the inefficiency is thus the role platforms play as intermediaries

- Platforms typically balance conflicting interests, which they rarely resolve with full disclosure otw, no info-design literature! :)
- Instead, they often garble the data they have collected

This paper shows how this practice can lead to a failure of the first-welfare theorem in a competitive data market

In the Works

Within application, we are working towards tighter conditions for **inefficiency**

Conjecture

Let $\gamma_u > \gamma_\pi$. An equilibrium is efficient if and only if the platform recommends $a = \min \Omega$ with probability 1.

Therefore, any "nontrivial" use of the database would lead to inefficiencies

example

A Simple Example to Illustrate

Suppose:

- $\gamma_u > \gamma_\pi = 0$, i.e. platform only cares about consumers' surplus
- Only two types of consumers: $\Omega = \{1, 2\}$ with $\bar{q}(1) < \bar{q}(2)$
- Suppose $\bar{r} < \frac{1+\gamma_u}{2}$ so that some trade is efficient

A Simple Example to Illustrate

Suppose:

- $\gamma_u > \gamma_\pi = 0$, i.e. platform only cares about consumers' surplus
- Only two types of consumers: $\Omega = \{1, 2\}$ with $\bar{q}(1) < \bar{q}(2)$
- Suppose $\bar{r} < \frac{1+\gamma_u}{2}$ so that some trade is efficient

There is a **unique** constrained-efficient allocation (q°, x°):

- All low-type consumers sell: $q^\circ(1) = \bar{q}(1)$
- Only some high-type consumers sell: $q^\circ(2) = \bar{q}(1) < \bar{q}(2)$
- Platform provides no info to merchant, who charges lowest fee to all consumers in database: $x^\circ(a = 1|\omega) = 1, \ \forall\omega$

A Simple Example to Illustrate

Suppose:

- $\gamma_u > \gamma_\pi = 0$, i.e. platform only cares about consumers' surplus
- Only two types of consumers: $\Omega = \{1, 2\}$ with $\bar{q}(1) < \bar{q}(2)$
- Suppose $\bar{r} < \frac{1+\gamma_u}{2}$ so that some trade is efficient

There is a **unique** constrained-efficient allocation (q°, x°):

- All low-type consumers sell: $q^\circ(1) = \bar{q}(1)$
- Only some high-type consumers sell: $q^\circ(2) = \bar{q}(1) < \bar{q}(2)$
- Platform provides no info to merchant, who charges lowest fee to all consumers in database: $x^\circ(a = 1|\omega) = 1, \ \forall\omega$

A Simple Example to Illustrate

Claim: If $\gamma_u < \bar{r}$, all equilibria are inefficient $\rightsquigarrow$ no trade (Corollary 1)

A Simple Example to Illustrate

Claim: If $\gamma_u < \bar{r}$, all equilibria are inefficient $\rightsquigarrow$ no trade (Corollary 1)

It can be shown that $p^*(\omega) \leq \gamma_u$. This implies that:

– Low-type consumers do not want to sell their records, $q^*(1) = 0$

Why? $U^*(1) = p^*(1) \leq \gamma_u < \bar{r}$

Do not internalize positive externality that selling their record generate for high-type consumers

A Simple Example to Illustrate

Claim: If $\gamma_u < \bar{r}$, all equilibria are inefficient $\rightsquigarrow$ no trade (Corollary 1)

It can be shown that $p^*(\omega) \leq \gamma_u$. This implies that:

- Low-type consumers do not want to sell their records, $q^*(1) = 0$

  Why? $U^*(1) = p^*(1) \leq \gamma_u < \bar{r}$

  Do not internalize positive externality that selling their record generate for high-type consumers

- Hence, high-type consumer do not want to sell either

  Why? $U^*(2) = p^*(2) \leq \gamma_u < \bar{r}$

A Simple Example to Illustrate

Claim: If $\gamma_u < \bar{r}$, all equilibria are inefficient $\rightsquigarrow$ no trade (Corollary 1)

It can be shown that $p^*(\omega) \leq \gamma_u$. This implies that:

- Low-type consumers do not want to sell their records, $q^*(1) = 0$

  Why? $U^*(1) = p^*(1) \leq \gamma_u < \bar{r}$

  Do not internalize positive externality that selling their record generate for high-type consumers

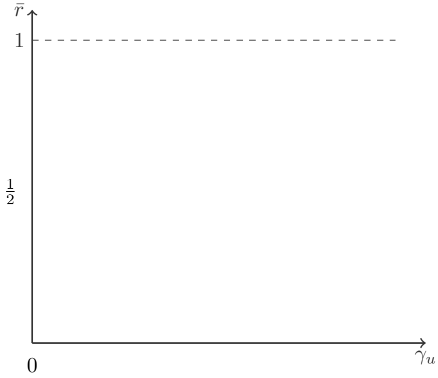
- Hence, high-type consumer do not want to sell either

  Why? $U^*(2) = p^*(2) \leq \gamma_u < \bar{r}$

- Market unravels $\rightsquigarrow$ No trade $\rightsquigarrow$ Inefficiency

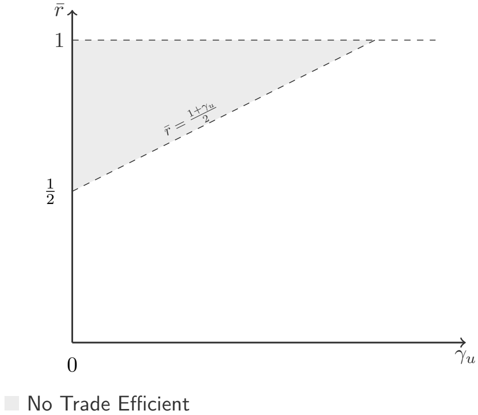
Example

Complete equilibrium characterization for this example:



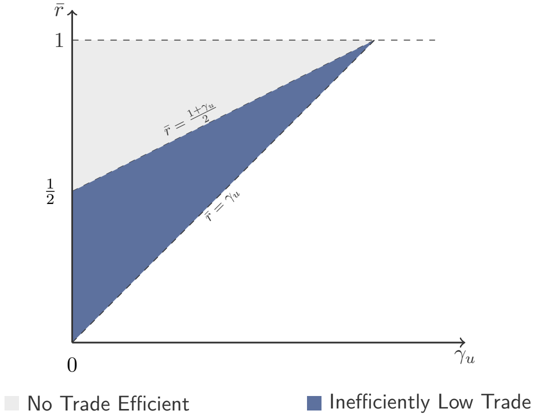
Example

Complete equilibrium characterization for this example:



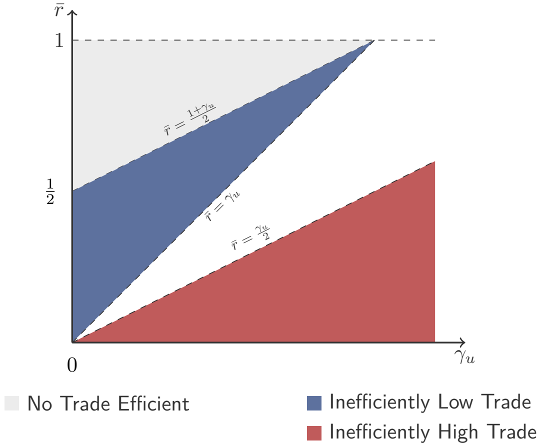
Example

Complete equilibrium characterization for this example:



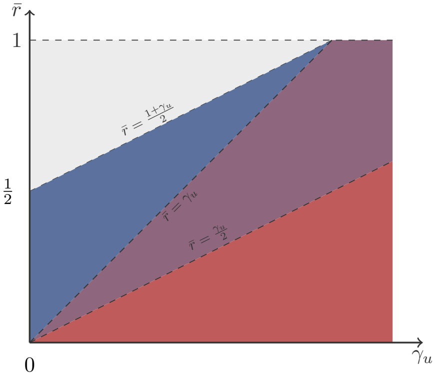
Example

Complete equilibrium characterization for this example:



Example

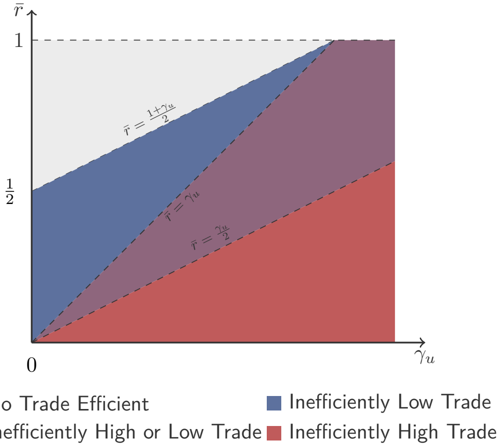
Complete equilibrium characterization for this example:



No Trade Efficient | Inefficiently Low Trade
Inefficiently High or Low Trade | Inefficiently High Trade

Example

Complete equilibrium characterization for this example:



remedies

Remedies

How to fix this market failure?

We explore three alternative market designs:

1. Introducing a **data union**
2. Implementing **data taxes**
3. Making data markets more **complete**

data union

Data Unions

Recent policy proposals for the data economy (Posner-Weyl 18; Bergemann et al 23)

A data union would represent consumers by managing data on their behalf

We offer some theoretical support to these policy proposals

Data Union

How does a data union work?

- Consumers can participate in the union
- If they do, they relinquish their data to the union
- Union sells some of this data to the platform

  Consumers retain reservation utility unless record is sold to platform

- With the proceeds of sale, union compensates all participating consumers (to incentivize their participation)
- Union maximizes welfare of participating consumers

Data Union

Formally, the data union problem is:

$$
\begin{aligned}
\max_{(p,q,x)} \quad & \sum_{\omega} p(\omega)\bar{q}(\omega) + \sum_{a,\omega} u(a,\omega)x(a|\omega)q(\omega) + \sum_{\omega}(\bar{q}(\omega) - q(\omega))\bar{r} \\
\text{such that} \quad & q \leq \bar{q}, \\
\text{and} \quad & \textstyle\sum_{\omega} p(\omega)\bar{q}(\omega) = V(q), \\
\text{and} \quad & x \text{ solves } \mathcal{P}_q, \\
\text{and} \quad & p(\omega) + \tfrac{q(\omega)}{\bar{q}(\omega)} \textstyle\sum_a u(a,\omega)x(a|\omega) + \left(1 - \tfrac{q(\omega)}{\bar{q}(\omega)}\right)\bar{r} \geq \bar{r}.
\end{aligned}
$$

Data Union

Proposition

Equilibria of the data-union economy are constrained efficient and maximize consumers' welfare (and vice versa), regardless of the platform's objective

Some intuition:

Data union coordinates consumers by deciding which records to sell and how to compensate them

By doing so, data union acts as a substitute for the competitive market and avoids market failure

data taxes

Data Taxes

Enrich competitive economy by introducing a simple **data tax**:

- When selling her record, consumer pays tax $\tau(\omega) \in \mathbb{R}$ to the govt

When properly designed, data taxes force consumers to internalize effects that selling their records create on economy

Data Taxes

Enrich competitive economy by introducing a simple **data tax**:

- When selling her record, consumer pays tax $\tau(\omega) \in \mathbb{R}$ to the govt

When properly designed, data taxes force consumers to internalize effects that selling their records create on economy

Proposition

Let (q°, x°) be a constrained-efficient allocation. There exists a profile of taxes τ^*, of prices p^*, and of consumer choices ζ^*, such that $(p^*, \zeta^*, q^\circ, x^\circ)$ is an equilibrium of the economy with taxation τ^* and the government does not run a deficit.

Sketch of the Proof

Let allocation (q°, x°) be constrained efficient

Let p^* be a supergradient of $V(q^\circ)$

Define $\boxed{\tau^*(\omega) \triangleq p^*(\omega) + \sum_a x^\circ(a|\omega)u(a,\omega) - \bar{r}}$

Notice that $U^*(\omega) - \tau^*(\omega) \equiv \bar{r}$

Therefore, all consumers indifferent $\rightsquigarrow$ choose ζ^* to implement q° $\square$

more-complete markets

Lindahl Economy

We let price of data depend not only on its type (i.e., ω) but also on its "intended use" (i.e., a)

> Platform and the consumer trade on **how** record will be used—i.e., which fee a platform will recommend to the merchant

Lindahl Economy

We let price of data depend not only on its type (i.e., ω) but also on its "intended use" (i.e., a)

> Platform and the consumer trade on **how** record will be used—i.e., which fee a platform will recommend to the merchant

This adapts to our setting the standard approach for modeling economies with externalities (Arrow 1969, Laffont 1976)

Lindahl Economy

We let price of data depend not only on its type (i.e., ω) but also on its "intended use" (i.e., a)

> Platform and the consumer trade on **how** record will be used—i.e., which fee a platform will recommend to the merchant

This adapts to our setting the standard approach for modeling economies with externalities (Arrow 1969, Laffont 1976)

This is reminiscent of GDPR: "*The **specific purposes** for which personal data are used should be determined at the time of the collection*"

Lindahl Economy

A market for each (a, ω), where ω-records can be traded for use a at price $p(a, \omega)$

Our equilibrium definition extends naturally to this richer economy

In particular, timing is the same

Lindahl Economy

A market for each (a, ω), where ω-records can be traded for use a at price $p(a, \omega)$

Our equilibrium definition extends naturally to this richer economy

In particular, timing is the same

Proposition

Equilibria of this economy are (unconstrained) efficient and maximize consumers' welfare, regardless of platform's incentives

conclusion

Summary

1. A stylized framework to study competitive markets for personal data

    Rooted in GE tradition but leveraging recent progress in info-design

2. Identify novel inefficiency leading this otherwise perfectly competitive market to fail

    Show how inefficiency critically depends on platform's role as an information intermediary

3. Propose three alternative market designs that fix inefficiency: data unions, data taxes, richer data prices

Competitive Markets for Personal Data

Simone Galperti UCSD
Tianhao Liu Columbia
Jacopo Perego Columbia

Thank You!

Efficiency Benchmark

To illustrate market failure, a less demanding efficiency benchmark is desirable:

Efficiency Benchmark

To illustrate market failure, a less demanding efficiency benchmark is desirable:

1. We require x° to be optimal given q° **for the platform**

   If not, detect inefficiency driven by platform lack of commitment in period 1

   (main results extend to "unconstrained" efficiency)

Efficiency Benchmark

To illustrate market failure, a less demanding efficiency benchmark is desirable:

1. We require x° to be optimal given q° **for the platform**

   If not, detect inefficiency driven by platform lack of commitment in period 1

   (main results extend to "unconstrained" efficiency)

2. We exclude merchant's payoff from $W(q, x)$

   If not, detect inefficiency driven by platform not fully internalizing merchant's payoff (main results extend to "aggregate" welfare)

Efficiency Benchmark

To illustrate market failure, a less demanding efficiency benchmark is desirable:

1. We require x° to be optimal given q° **for the platform**

   If not, detect inefficiency driven by platform lack of commitment in period 1

   (main results extend to "unconstrained" efficiency)

2. We exclude merchant's payoff from $W(q, x)$

   If not, detect inefficiency driven by platform not fully internalizing merchant's payoff (main results extend to "aggregate" welfare)

Bonus: In eqm, platform makes not profits. Thus, $W(q^*, x^*)$ equals consumer welfare. Thus, any constrained-efficient eqm maximizes consumer welfare

back