

Competitive Markets for Personal Data

Simone Galperti Jacopo Perego
UCSD Columbia

February 2023

Personal data fuels multi-billion dollar industries — essential input in digital economy

Status quo: consumers are imperfectly compensated for their data, and have limited control over their use \rightsquigarrow possible market failures (Seim et al. 22)

Establishing competitive markets for personal data could improve efficiency and equity

Growing interest in studying theoretical properties of these markets and their problems e.g, Choi et al. 19, Bergemann et al. 22, Acemoglu et al. 22

We identify a **novel externality** that makes these markets inefficient

Inefficiency can be so severe that consumers would be better off in an economy with expropriation (akin to status quo)

We identify a **novel externality** that makes these markets inefficient

Inefficiency can be so severe that consumers would be better off in an economy with expropriation (akin to status quo)

This inefficiency:

- Present even if markets are otherwise **perfectly competitive**
- Present even if consumers data is **uncorrelated**
- Originates from the specific role that **information intermediaries** play

My Goal Today: Illustrate inefficiency through a simple example

example

(many consumers, one platform, one vendor)

The **vendor** sells a product to consumers and charges fee $a \in \{1, 2\}$

Consumers have unit demand and WTP of $\omega \in \{1, 2\} = \Omega$

Assume mass of consumers: $\bar{q}(\omega = 1) = 1$ and $\bar{q}(\omega = 2) = 2$

The **vendor** sells a product to consumers and charges fee $a \in \{1, 2\}$

Consumers have unit demand and WTP of $\omega \in \{1, 2\} = \Omega$

Assume mass of consumers: $\bar{q}(\omega = 1) = 1$ and $\bar{q}(\omega = 2) = 2$

Each consumer is uniquely identified by a **data record**, which reveals her ω

A collection of records is called a **database**, denoted by $q \in \mathbb{R}_+^\Omega$

Given a database q , the **platform** can intermediate these consumers with the vendor

Specifically, platform acts as **information designer** like in BBM15

- ▶ It sends a signal about a consumer's ω to the vendor to influence the fee he charges, and to maximize its own objective

Given a database q , the **platform** can intermediate these consumers with the vendor

Specifically, platform acts as **information designer** like in BBM15

- It sends a signal about a consumer's ω to the vendor to influence the fee he charges, and to maximize its own objective

Denote:

platform's
expected payoff
from database q

$$\overbrace{U(q)}$$

Given a database q , the **platform** can intermediate these consumers with the vendor

Specifically, platform acts as **information designer** like in BBM15

- It sends a signal about a consumer's ω to the vendor to influence the fee he charges, and to maximize its own objective

Denote:

platform's
expected payoff
from database q

$$\widehat{U(q)}$$

$$= \max_{x: A \times \Omega \rightarrow \mathbb{R}_+} \sum_{\omega, a} u(a, \omega) x(a, \omega)$$

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

$$\sum_a x(a, \omega) = q(\omega) \quad \forall \omega \in \Omega$$

a standard information design problem

Market for data records is assumed to be competitive

Market for data records is assumed to be competitive

1. Taking the **market price** $p(\omega)$ of each ω records as given:

Market for data records is assumed to be competitive

1. Taking the **market price** $p(\omega)$ of each ω records as given:
 - Each **consumer** either:
 - Sells her record to the platform
 - Keeps her record and avoids platform's intermediation

Market for data records is assumed to be competitive

1. Taking the **market price** $p(\omega)$ of each ω records as given:
 - Each **consumer** either:
 - Sells her record to the platform
 - Keeps her record and avoids platform's intermediation
 - **Platform** acquires data records from consumers

Market for data records is assumed to be competitive

1. Taking the **market price** $p(\omega)$ of each ω records as given:
 - Each **consumer** either:
 - Sells her record to the platform
 - Keeps her record and avoids platform's intermediation
 - **Platform** acquires data records from consumers
2. Platform intermediates consumers in its database with the vendor

Market for data records is assumed to be competitive

1. Taking the **market price** $p(\omega)$ of each ω records as given:
 - Each **consumer** either:
 - Sells her record to the platform
 - Keeps her record and avoids platform's intermediation
 - **Platform** acquires data records from consumers
2. Platform intermediates consumers in its database with the vendor

Market for data records is assumed to be competitive

1. Taking the **market price** $p(\omega)$ of each ω records as given:
 - Each **consumer** either:
 - Sells her record to the platform
 - Keeps her record and avoids platform's intermediation
 - **Platform** acquires data records from consumers
2. Platform intermediates consumers in its database with the vendor

Payoffs:	if consumer:	
	keeps record	sells record
Consumer	ε	
Vendor	0	
Platform	0	

Market for data records is assumed to be competitive

1. Taking the **market price** $p(\omega)$ of each ω records as given:
 - Each **consumer** either:
 - Sells her record to the platform
 - Keeps her record and avoids platform's intermediation
 - **Platform** acquires data records from consumers
2. Platform intermediates consumers in its database with the vendor

Payoffs:	if consumer:		
	keeps record	sells record	
Consumer	ε	$\max\{\omega - a, 0\}$	consumer's surplus
Vendor	0		
Platform	0		

Market for data records is assumed to be competitive

1. Taking the **market price** $p(\omega)$ of each ω records as given:
 - Each **consumer** either:
 - Sells her record to the platform
 - Keeps her record and avoids platform's intermediation
 - **Platform** acquires data records from consumers
2. Platform intermediates consumers in its database with the vendor

Payoffs:	if consumer:		
	keeps record	sells record	
Consumer	ε	$\max\{\omega - a, 0\}$	consumer's surplus
Vendor	0	$a\mathbb{1}(\omega \geq a)$	vendor's profit
Platform	0		

Market for data records is assumed to be competitive

1. Taking the **market price** $p(\omega)$ of each ω records as given:
 - Each **consumer** either:
 - Sells her record to the platform
 - Keeps her record and avoids platform's intermediation
 - **Platform** acquires data records from consumers
2. Platform intermediates consumers in its database with the vendor

Payoffs:	if consumer:		
	keeps record	sells record	
Consumer	ε	$\max\{\omega - a, 0\}$	consumer's surplus
Vendor	0	$a\mathbb{1}(\omega \geq a)$	vendor's profit
Platform	0	$\beta \max\{\omega - a, 0\}$	a fraction of CS

Market for data records is assumed to be competitive

1. Taking the **market price** $p(\omega)$ of each ω records as given:
 - Each **consumer** either:
 - Sells her record to the platform
 - Keeps her record and avoids platform's intermediation
 - **Platform** acquires data records from consumers
2. Platform intermediates consumers in its database with the vendor

Payoffs:	if consumer:		
	keeps record	sells record	
Consumer	ε	$\max\{\omega - a, 0\}$	consumer's surplus
Vendor	0	$a\mathbb{1}(\omega \geq a)$	vendor's profit
Platform	0	$\beta \max\{\omega - a, 0\}$	a fraction of CS

assume: $\beta < \varepsilon < 1/3$

A **competitive equilibrium** is: prices $p^* \in \mathbb{R}_+^\Omega$; database $q^* \in \mathbb{R}_+^\Omega$;
consumers' decisions $\alpha^* \in [0, 1]^\Omega$ such that

A **competitive equilibrium** is: prices $p^* \in \mathbb{R}_+^\Omega$; database $q^* \in \mathbb{R}_+^\Omega$; consumers' decisions $\alpha^* \in [0, 1]^\Omega$ such that

1. Given p^* , platform optimizes.

$$q^* \in \arg \max_q U(q) - \sum_{\omega} p^*(\omega) q(\omega)$$

2. Given p^* and q^* , each consumer optimizes.

$$\alpha^*(\omega) \in \arg \max_{\alpha} (1 - \alpha)\varepsilon + \alpha \left(p^*(\omega) + \mathbb{E}_{q^*}(g(a, \omega)) \right)$$

3. And markets clear.

$$q^*(\omega) = \alpha^*(\omega) \bar{q}(\omega), \quad \forall \omega$$

A **competitive equilibrium** is: prices $p^* \in \mathbb{R}_+^\Omega$; database $q^* \in \mathbb{R}_+^\Omega$; consumers' decisions $\alpha^* \in [0, 1]^\Omega$ such that

1. Given p^* , platform optimizes.

$$q^* \in \arg \max_q U(q) - \sum_{\omega} p^*(\omega) q(\omega)$$

2. Given p^* and q^* , each consumer optimizes.

$$\alpha^*(\omega) \in \arg \max_{\alpha} (1 - \alpha)\varepsilon + \alpha \left(p^*(\omega) + \mathbb{E}_{q^*}(g(a, \omega)) \right)$$

3. And markets clear.

$$q^*(\omega) = \alpha^*(\omega) \bar{q}(\omega), \quad \forall \omega$$

In the paper: many platforms, many vendors, arbitrary objectives, partially informative records, multiple types

analysis

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

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

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?



Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Record Type	Records Kept	Platform's Database	Use	Payoffs		
$\omega = 1$	0	1	$\xrightarrow{\frac{1}{2}} s^L \longrightarrow a = 1$	C	P	V
				1	β	4
$\omega = 2$	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$			

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Record Type	Records Kept	Platform's Database	Use	Payoffs		
$\omega = 1$	0	1	$\xrightarrow{\frac{1}{2}} s^L \longrightarrow a = 1$	C	P	V
				1	β	4
$\omega = 2$	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$			

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Record Type	Records Kept	Platform's Database	Use	Payoffs		
$\omega = 1$	0	1	$\xrightarrow{\frac{1}{2}} s^L \longrightarrow a = 1$	C	P	V
				1	β	4
$\omega = 2$	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$			

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Record Type	Records Kept	Platform's Database	Use	Payoffs		
$\omega = 1$	0	1	$\xrightarrow{\frac{1}{2}} s^L \longrightarrow a = 1$	C	P	V
				1	β	4
$\omega = 2$	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$			

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Record Type	Records Kept	Platform's Database	Use	Payoffs		
$\omega = 1$	0	1	$\xrightarrow{\frac{1}{2}} s^L \longrightarrow a = 1$	C 1	P β	V 4
$\omega = 2$	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$			

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Record Type	Records Kept	Platform's Database	Use	Payoffs		
$\omega = 1$	0	1	$\xrightarrow{\frac{1}{2}} s^L \longrightarrow a = 1$	C	P	V
				1	β	4
$\omega = 2$	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$			

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Record Type	Records Kept	Platform's Database	Use	Payoffs		
$\omega = 1$	0	1	$\xrightarrow{\frac{1}{2}} s^L \longrightarrow a = 1$	C 1	P β	V 4
$\omega = 2$	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$			

Comments:

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Record Type	Records Kept	Platform's Database	Use	Payoffs		
$\omega = 1$	0	1	$\xrightarrow{\frac{1}{2}} s^L \longrightarrow a = 1$	C 1	P β	V 4
$\omega = 2$	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$			

Comments:

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Record Type	Records Kept	Platform's Database	Use	Payoffs		
$\omega = 1$	0	1	$\xrightarrow{\frac{1}{2}} s^L \longrightarrow a = 1$	C 1	P β	V 4
$\omega = 2$	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$			

Comments:

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Record Type	Records Kept	Platform's Database	Use	Payoffs		
$\omega = 1$	0	1	$\xrightarrow{\frac{1}{2}} s^L \longrightarrow a = 1$	C 1	P β	V 4
$\omega = 2$	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$			

Comments:

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Record Type	Records Kept	Platform's Database	Use	Payoffs		
$\omega = 1$	0	1	$\xrightarrow{\frac{1}{2}} s^L \longrightarrow a = 1$	C 1	P β	V 4
$\omega = 2$	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$			

Comments:

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Record Type	Records Kept	Platform's Database	Use	Payoffs		
$\omega = 1$	0	1	$\xrightarrow{\frac{1}{2}} s^L \longrightarrow a = 1$	C	P	V
				1	β	4
$\omega = 2$	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$			

Comments:

1. Since platform's and vendor's objectives are not aligned, it is optimal to withhold some information (key feature of platform mkts: examples)

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Record Type	Records Kept	Platform's Database	Use	Payoffs		
$\omega = 1$	0	1	$\xrightarrow{\frac{1}{2}} s^L \longrightarrow a = 1$	C	P	V
			$\nearrow \frac{1}{2}$	1	β	4
$\omega = 2$	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$			

Comments:

- When data records are pooled, one consumer's payoff depends on what other records platform has

Imagine platform **expropriated** consumers of their records (\approx status quo)

How would the platform use database \bar{q} ?

Record Type	Records Kept	Platform's Database	Use	Payoffs		
$\omega = 1$	0	1	$\xrightarrow{\frac{1}{2}} s^L \longrightarrow a = 1$	C	P	V
			$\nearrow \frac{1}{2}$	1	β	4
$\omega = 2$	0	2	$\xrightarrow{\frac{1}{2}} s^H \longrightarrow a = 2$			

Comments:

- This “pooling externality” (GLP '22) is the root of inefficiency in our data economy

Let's compute the (unique) **competitive equilibrium** of the data economy:

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Let's compute the (unique) **competitive equilibrium** of the data economy:

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	$0 \longrightarrow s^H \longrightarrow a = 2$	

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Comments:

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Comments:

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Comments:

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Comments:

1. Consumers have no incentive to sell: prices too low

$$(\beta < \varepsilon)$$

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Comments:

- Platform has no (strict) incentive to buy: prices too high $(\beta < \varepsilon)$
The value of a pair of records is β

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Comments:

- Equilibrium is inefficient ("too little data"): Consumer welfare is lower than under expropriation (remember $\varepsilon < \frac{1}{3}$)

Let's compute the (unique) **competitive equilibrium** of the data economy:

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

Comments:

- High-type consumers would like to subsidize low-type consumers to sell their data, but this market is too incomplete

Beyond this example?

Beyond this example?

- ▶ Our data economy is generally inefficient, i.e., equilibrium data allocation does not maximize consumers' welfare

Beyond this example?

- ▶ Our data economy is generally inefficient, i.e., equilibrium data allocation does not maximize consumers' welfare
- ▶ The inefficiency can be so severe that consumers' welfare would be higher in an economy with expropriation
 - Unintended consequence of giving consumers control

Beyond this example?

- ▶ Our data economy is generally inefficient, i.e., equilibrium data allocation does not maximize consumers' welfare
- ▶ The inefficiency can be so severe that consumers' welfare would be higher in an economy with expropriation
 - Unintended consequence of giving consumers control
- ▶ “Pooling externalities” are the root of inefficiency:
Economy is efficient when platforms don't have incentives to withhold information (back to example)

Beyond this example?

- ▶ Our data economy is generally inefficient, i.e., equilibrium data allocation does not maximize consumers' welfare
- ▶ The inefficiency can be so severe that consumers' welfare would be higher in an economy with expropriation
 - Unintended consequence of giving consumers control
- ▶ “Pooling externalities” are the root of inefficiency:
Economy is efficient when platforms don't have incentives to withhold information (back to example)
- ▶ This problem distinct and complementary to “social data”

Beyond this example?

- ▶ Our data economy is generally inefficient, i.e., equilibrium data allocation does not maximize consumers' welfare
- ▶ The inefficiency can be so severe that consumers' welfare would be higher in an economy with expropriation
 - Unintended consequence of giving consumers control
- ▶ “Pooling externalities” are the root of inefficiency:
Economy is efficient when platforms don't have incentives to withhold information (back to example)
- ▶ This problem distinct and complementary to “social data”

solutions?

A classic solution:

Arrow (1969), Laffont (1978)

- ▶ Open markets where parties can trade the way data is **used**.
This presumes data use can be contractible

A classic solution:

Arrow (1969), Laffont (1978)

- ▶ Open markets where parties can trade the way data is **used**.
This presumes data use can be contractible

Back to example:

**Record
Type**

$$\omega = 1$$

$$\omega = 2$$

A classic solution:

Arrow (1969), Laffont (1978)

- Open markets where parties can trade the way data is **used**.
This presumes data use can be contractible

Back to example:

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

A classic solution:

Arrow (1969), Laffont (1978)

- Open markets where parties can trade the way data is **used**.
This presumes data use can be contractible

Back to example:

**Record
Type**

Prices

$\omega = 1$

$$p^*(1, a = 1)$$

$$p^*(1, a = 2)$$

$\omega = 2$

$$p^*(2, a = 1)$$

$$p^*(2, a = 2)$$

A classic solution:

Arrow (1969), Laffont (1978)

- Open markets where parties can trade the way data is **used**.
This presumes data use can be contractible

Back to example:

**Record
Type**

Prices

$\omega = 1$

$$p^*(1, a = 1)$$

$$p^*(1, a = 2)$$

$\omega = 2$

$$p^*(2, a = 1)$$

$$p^*(2, a = 2)$$

A classic solution:

Arrow (1969), Laffont (1978)

- Open markets where parties can trade the way data is **used**.
This presumes data use can be contractible

Back to example:

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

A classic solution:

Arrow (1969), Laffont (1978)

- Open markets where parties can trade the way data is **used**.
This presumes data use can be contractible

Back to example:

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

A classic solution:

Arrow (1969), Laffont (1978)

- Open markets where parties can trade the way data is **used**.
This presumes data use can be contractible

Back to example:

Record Type	Prices	Records Kept	Records Used	Payoffs		
$\omega = 1$	$p^*(1, 1) = \beta + 1 - \varepsilon$	0	$x(1, 1) = 1$	C $1 + \varepsilon + \beta$	P 0	V 2
	$p^*(1, 2) = 0$		$x(1, 2) = 0$			
$\omega = 2$	$p^*(2, 1) = -(1 - \varepsilon)$	1	$x(2, 1) = 1$			
	$p^*(2, 2) = 0$		$x(2, 2) = 0$			

A classic solution:

Arrow (1969), Laffont (1978)

- Open markets where parties can trade the way data is **used**.
This presumes data use can be contractible

Back to example:

Record Type	Prices	Records Kept	Records Used	Payoffs		
$\omega = 1$	$p^*(1, 1) = \beta + 1 - \varepsilon$	0	$x(1, 1) = 1$	C $1 + \varepsilon + \beta$	P 0	V 2
	$p^*(1, 2) = 0$		$x(1, 2) = 0$			
$\omega = 2$	$p^*(2, 1) = -(1 - \varepsilon)$	1	$x(2, 1) = 1$			
	$p^*(2, 2) = 0$		$x(2, 2) = 0$			

A classic solution:

Arrow (1969), Laffont (1978)

- Open markets where parties can trade the way data is **used**.

This presumes data use can be contractible

Back to example:

Record Type	Prices	Records Kept	Records Used	Payoffs		
$\omega = 1$	$p^*(1, 1) = \beta + 1 - \varepsilon$	0	$x(1, 1) = 1$	C $1 + \varepsilon + \beta$	P 0	V 2
	$p^*(1, 2) = 0$		$x(1, 2) = 0$			
$\omega = 2$	$p^*(2, 1) = -(1 - \varepsilon)$	1	$x(2, 1) = 1$			
	$p^*(2, 2) = 0$		$x(2, 2) = 0$			

Result. In this economy, every equilibrium data allocation maximizes consumer surplus (and vice versa)

How realistic?

Open Questions:

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

Seim et al. 2022

conclusions

Summary

We study competitive markets for personal data:

1. We identify a novel inefficiency that a data market can suffer from, even when it is otherwise perfectly competitive

It shows how giving consumers control over their data can backfire

2. The inefficiency originates from how data is endogenously used by the intermediaries
“pooling externality”
3. We discuss possible solutions and their limits

thank you