

Data Collection by an Informed Seller

Shota Ichihashi Alex Smolin

Queen's University

Toulouse School of Economics

Platform Economics Workshop

August 19, 2022



amazon


Alibaba Group



Cookies

This site uses cookies to offer you a better browsing experience. Find out more on [how we use cookies and how you can change your settings](#).



Log In with Facebook



Log In with Google



Log In with Twitter

Consumers: uncertain about how data affect them

Firms know more

Info asymmetry

Buyer-seller model

- ▶ restrictive but clean result
- ▶ (hopefully) generalizable insight

Outline

1. Example
2. General result
3. Policy implication & literature

Example

A seller (S) and a consumer (C)

S has a product

Value to C is $v \in \{1, 2\}$

Game:

1. Data collection: S requests data \rightarrow C yes/no
2. Transaction: S sets price, C buy/not

Data Collection Stage

Initially, neither S nor C knows v

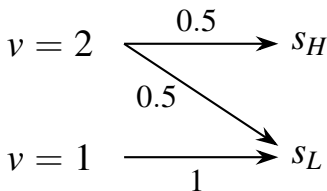
E.g., C hasn't seen the product at registration page

$\mathbb{P}(v = 2)$ is commonly known

Data Collection Stage

S requests “data”

1. No data
2. Full data: S learns $v = 1$ or $v = 2$
3. Partial data; in this example,



(In the paper, we allow arbitrary data)

Game

1. S chooses which data to request
2. C decides whether to provide the data
 - ▶ C knows the requested data
 - ▶ Doesn't know v or the content of data
3. S sets price p
4. C **observes** v and decides whether to buy
5. If C buys, payoffs $v - p$ and p ; if not, both get 0

Perfect Bayesian Equilibrium

If $\mathbb{P}(v = 2) = 2/3$ then $p = 2$ without data

▶ $p = 1 \rightarrow$ revenue 1

▶ $p = 2 \rightarrow$ revenue $4/3$

If $\mathbb{P}(v = 2) = 2/3$ then $p = 2$ without data

- ▶ $p = 1 \rightarrow$ revenue 1
- ▶ $p = 2 \rightarrow$ revenue $4/3$

C gets payoff 0 \rightarrow C prefers to provide **any** data

Equilibrium: S obtains full data

- ▶ efficient
- ▶ consumer surplus = 0

If $\mathbb{P}(v = 2) = 1/3$ then $p = 1$ without data

- ▶ $p = 1 \rightarrow$ revenue 1
- ▶ $p = 2 \rightarrow$ revenue $2/3$

Equilibrium: S obtains **no** data

- ▶ efficient
- ▶ consumer surplus > 0

In either case,

- ▶ Allocation is efficient, but
- ▶ Data collection does not benefit C

What if S knows something C doesn't?

Informed Seller

Now, $\mathbb{P}(v = 2) = 2/3$ or $1/3$

S knows $\mathbb{P}(v = 2)$

C only knows $2/3$ and $1/3$ are equally likely

Why is S better informed?

- ▶ knows market demand better
- ▶ knows consumers better

Informed Seller

1. S observes $\Pr(v = 2) = 2/3$ or $1/3$
2. S chooses data to request
3. C decides whether to accept
4. S chooses a price, then C decides whether to buy

Seller's private belief = Seller's "type"

S with $\Pr(v = 2) = 1/3$ is a bad seller

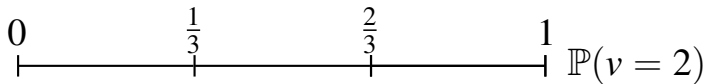
No Data Equilibrium

On-path: Both seller types obtain no information

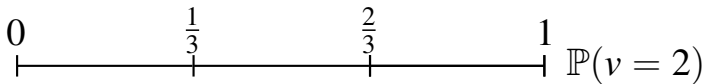
Off-path: C rejects any deviant request,
believing it is from S with $\mathbb{P}(v = 2) = 1/3$

Inefficiency if $\mathbb{P}(v = 2) = 2/3$ as $p = 2$ when $v = 1$

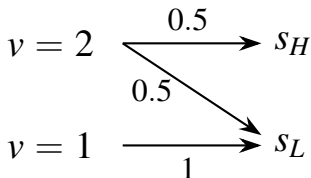
Partial Data Equilibrium



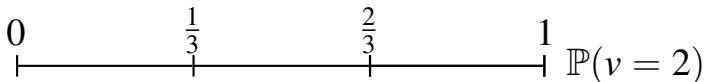
Partial Data Equilibrium



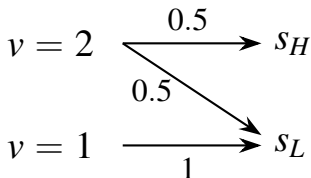
Pooling eqm:
both seller types collect



Partial Data Equilibrium



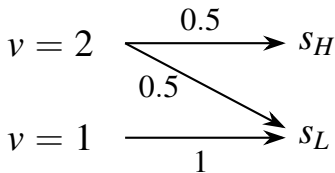
Pooling eqm:
both seller types collect



C is willing to provide data,
because S with $\mathbb{P}(v = 2) = 2/3$ uses it to benefit C

Partial Data Equilibrium

On-path: Both seller types obtain the partial data



Off-path: C rejects any deviant request

Welfare:

- ▶ Inefficient
- ▶ Consumer better than no data collection

General Result

The seller's private information:

- ▶ decreases total surplus
- ▶ decreases profits
- ▶ (weakly) increases consumer surplus

Consumer's skepticism disciplines the seller

Implication: Seller's Strategy

Compare

1. S who knows $\mathbb{P}(v = 2) = 2/3$ or $1/3$
2. S doesn't, i.e., $\mathbb{P}(v = 2) = 1/2$

Uninformed seller collects more data and earns more

Less initial private info \rightarrow more data

Phasing out third-party cookies?

Implication: Privacy Regulation

“Control over data”: C chooses what data to provide

Does it help? Not much:

- ▶ Same data to all seller types
- ▶ Same outcome as one eqm of the original game

Info asymmetry limits the effect of the regulation

Related Literature

Data collection:

Choi Jeon Kim 2019, Fainmesser Galeotti Momot 2021,
Argenziano Bonatti 2021, Bergemann Bonatti Gan 2021,
Acemoglu Makhdoumi Malekian Ozdaglar 2021

- ▶ Consumers have (weakly) more info
- ▶ We study informed firms

Related Literature

Price discrimination & market segmentation

Bergemann Brooks Morris 2015, Roesler Szentes 2017,

Haghpanah Siegel, 2019; Shi and Zhang 2020; Haghpanah and

Siegel 2021; Rhodes and Zhou 2021

- ▶ Typically agnostic about how firms obtain data
- ▶ We endogenize data collection
 - consumer consent + info asymmetry

Final Thought

Which implications are most relevant?

- ▶ Info asymmetry
 - ▶ some firm gets too much data, some gets too little
 - ▶ discourage data sharing
 - ▶ (improve consumer surplus?)
- ▶ “Control” useless without transparency