The Design and Price of Information

Dirk Bergemann, Alessandro Bonatti and Alex Smolin('18 AER)

Renjie Zhong

PPE
Renmin University of China

February 2023

Outline

Introduction

Big Picture (BB,2018; Morris,2018)

What Does the Data Broker Sell?

Who Identifies the Prospect?		Only Information	Access to Consumer
	Data Broker	*original lists*	*sponsored search*
	Data Buyer	*data appends*	*retargeting*

Table 1: Classification of Online Information Products

	Single Agent	Many Agent Uninformed Designer	Many Agent Informed Designer
Omniscient	.	Bayesian Solution	BCE
Communicating	Kolotilin et al	Communication Equilibrium	.
Non Communicating	KG informed receiver	Strategic Form Correlated Equilibrium	.

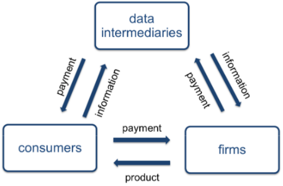


Figure 1: Market for Consumer Data

Research Questions

- Data buyer - a decision maker under uncertainty:
  1. has partial and private information (heterogeneous valuation)
  2. can acquire additional information (Data Appends)
- Data seller offers additional information (Omniscient):
  1. How much information to provide and at what price?
  2. How to provide different information to different data buyers?
- Contractible Elements Information Products

  $=$ experiment (in the statistical sense of Blackwell), which provides statistical information about payoff-relevant state
- Interpretation: selling access to a database as in Acxiom, Bluekai, DoubleClick Ad Exchange

This Paper

Indeed, we can decompose the model in this paper into three basic theoretical toolkits:

- Design optimal selling plan to consumers with heterogeneous valuation: screening
- Design optimal versioning plan:
  1. Blackwell Theorem
  2. bayes correlated equilibrium (BCE)
- The results in this paper is akin to a "combination" of insights from the classical theories above.

Optimal Selling: Screening

- Insights: Tradeoff between Information Rents and Allocation Inefficiency
  1. No distortion at the top
  2. No rent at the bottom, information rent at the top
  3. Profitable Screening only with little θ_H
  4. Virtual value matters
- IC and IR:
  1. Binding IR-L
  2. $q(\cdot)$ is increasing in θ and LDIC-FOC codntion
  3. Ironing and Butching
- But if the seller can (arbitrarily) design some products affecting the valuation?
  1. horizontal differentiation that widens the seller's scope for price discrimination
  2. information is inherently rich and can be modified in many ways

Optimal Versioning: Information Design

- Bayes Correlated Equilibrium (Bergemann and Morris,16,19): a linear programming problem
- Informativeness (Blackwell) means the quality of his decision making under some information structure
  instrumental information is useful as long as it guides the decision of the data buyer
  garbling means less informativeness (Blackwell Theorem)
- Intepretations: Statistical Inference/ Classification

Results

- A menu of experiments is offered:
  1. a coarser menu compared to types
  2. linearity (in probabilities) limits the use of versioning
  3. systematic distortions in information provided:
- Screening facilitated by directional information
  1. either type I/II error
  2. should not observe multiple distortions of the same kind
  3. extract all surplus by versioning

Related Literature

1. Selling Information:
   Admati and Páeiderer (1986, 1990), EsΩo and Szentes (2007), Babaioff (2012)
2. Information Impacts Prices:
   Johnson and Myatt (2006), Bergemann and Pesendorfer (2007)
3. Persuasion:
   Rayo and Segal (2010), Kamenica and Gentzkow (2011)

Model: Setup

Setup

- finite actions $A = \{a_1, ..., a_N\}$
- finite states $\Omega = \{\omega_1, ..., \omega_N\}$
- utility $u(a_i, \omega_j)$

u	a_1	$\cdots$	a_J
ω_1	u_{11}	$\cdots$	u_{1J}
$\vdots$	$\vdots$		$\vdots$
ω_I	u_{I1}	$\cdots$	u_{IJ}

- leading example: $u(a_i, \omega_j) = 1_{i=j} u_{ii}$
- Omniscient and (partial and private) informed buyers: θ
  $\theta \in \Delta\Omega$ is induced by private signal $\lambda : \Omega \to \Delta R$ (with prior belief μ)

Setup

- experiment: $E = \{S, \pi\}$ consists of signals $s \in S$ with $\pi : \omega \to \Delta S$, $\pi_{ik} = Pr[s_k|\omega_i]$
  r and s are independent
- stochastic matrix:

$$\begin{array}{c|ccc} E & s_1 & \cdots & s_K \\ \hline \omega_1 & \pi_{11} & \cdots & \pi_{1K} \\ \vdots & \vdots & & \vdots \\ \omega_I & \pi_{I1} & \cdots & \pi_{IK} \end{array}$$

- a menu of experiments: $\mathcal{M} = \{\varepsilon, t\}$
  ϵ: a collection of experiments
  $t : \varepsilon \to \mathbb{R}_+ t$: associated tariff

Timeline

1. The seller posts a menu $\mathcal{M}$
2. The true state ω and the buyer's type θ are realized
3. The buyer chooses an experiment $E \in \epsilon$ and pays the corresponding price $t(E)$
4. The buyer observes a signal s from experiment E and chooses an action a

Model: Analysis

Value of Information

- without information structure
  1. $a(\theta) \in \arg\max_{a_j \in A}\{\sum_{i=1}^{I} \theta_i u_{ij}\}$
  2. $u(\theta) = \max\{\sum_{i=1}^{I} \theta_i u_{ij}\}$
- with information structure:
  1. $Pr(s_k|\theta) = \sum_{i=1}^{I} \theta_i \pi_{ik}$
  2. $a(s_k \mid \theta) \in \underset{a_j \in A}{\arg\max}\left\{\sum_{i=1}^{I}\left(\frac{\theta_i \pi_{ik}}{\sum_{i'=1}^{I} \theta_{i'} \pi_{i'k}}\right) u_{ij}\right\}$
  3. $u(s_k \mid \theta) \triangleq \max\left\{\sum_{i=1}^{I}\left(\frac{\theta_i \pi_{ik}}{\sum_{i'=1}^{I} \theta_{i'} \pi_{i'k}}\right) u_{ij}\right\}$
- the value of information: $V(E,\theta) \triangleq E[u(s \mid \theta)] - u(\theta) = \sum_{k=1}^{K} \max_j\left\{\sum_{i=1}^{I} \theta_i \pi_{ik} u_{ij}\right\} - u(\theta)$

Example

- three states ω_i $(i = 1, 2, 3)$ and interiem belief $(\theta_1, \theta_2, 1 - \theta_1 - \theta_2)$
- perfect information experiment and imperfect (nosiy) one

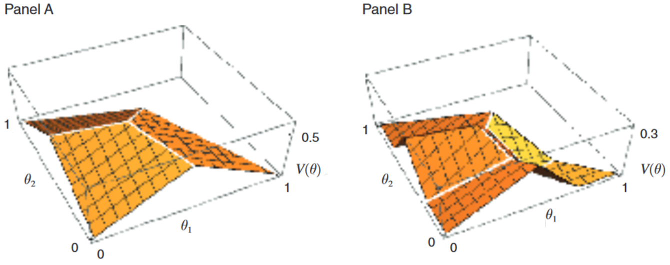


FIGURE 1. VALUE OF FULL AND PARTIAL INFORMATION, $I = J = 3$

- piecewise linearity: the bayesian nature
- convex kinks: the max operater

Seller's Problem

$$\max_{\{E(\theta),t(\theta)\}} \int_{\theta\in\Theta} t(\theta)\, dF(\theta)$$

1. IR: $V(\theta) = V(E(\theta),\theta) - t(\theta) = 0$
2. IC: $V(\theta) \geq V(E(\theta'),)\theta - t(\theta')$, $\forall \theta, \theta' \in \Theta$

Simplication

- a great simplication: responsive or (private) recommendation or "direct revelation" mechanism or maximal cardinality of signals
- responsive: $a(s_k|\theta) = a_k$ for all $s_k \in S(\theta)$
  every signal induces different optimal action

PROPOSITION 1 (Responsive Menus)

The outcome of every menu $\mathtt{M}$ can be attained by a responsive menu.

- Proof Sketch: Merging and Garbling

Illustrative Examples: Binary Situation

- Rewrite the stochastic matrix:

$$\begin{array}{c|ccc} u & a_1 & \cdots & a_J \\ \hline \omega_1 & \pi_{11} & \cdots & \pi_{1J} \\ \vdots & \vdots & & \vdots \\ \omega_I & \pi_{I1} & \cdots & \pi_{IJ} \end{array}$$

- $V(E,\theta) = \sum_{k=1}^{K} \sum_{j=1}^{J} \sum_{i=1}^{I} \theta_i \pi_{ij} u_{ij} - \max\{\sum_{i=1}^{I} \theta_i u_{ij}\}$

An Illustration: Binary States

- binary state, binary action:

$u(a,\theta)$	a_1	a_2
ω_1	u_1	0
ω_2	0	u_2

- $\theta \triangleq Pr(\omega = \omega_1)$, indifferent belief (most uninformed) $\theta^* = \frac{u_2}{u_1+u_2}$

$E(\theta)$	s_1	s_2
ω_1	$\pi_1(\theta)$	$1-\pi_1(\theta)$
ω_2	$1-\pi_2(\theta)$	$\pi_2(\theta)$

- suppose $\pi_1 + \pi_2 \geq 1$ w.l.o.g (monotone likelihood ratio)

Geometry

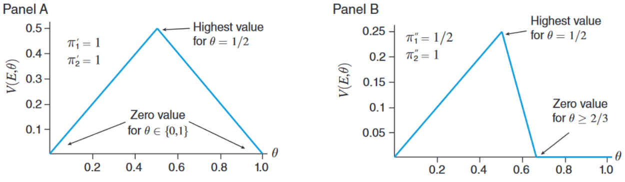


Figure 2. Value of Full and Partial Information ($I = J = 2, u_1 = u_2 = 1$)

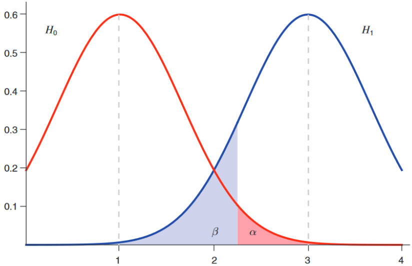
- distance to the uniform belief doesn't necessarily represent the degree of uninformativenss
- different slopes: differential gains of avoiding type 1 errors
- information has horizontal and vertical dimension of differentiation, information is always high-dimensional
- high degree of incompleteness in ranking of information structures

Statistical Inference Interpretation

- a statistical inference interpretation: $H_0 = \{\omega_2\}$

$E(\theta)$	s_1	s_2
ω_1	$1-\beta$	β
ω_2	α	$1-\alpha$

1. FP/Type I Error: $(\omega_2, s_1) : \alpha$
2. FN/Type II Error: $(\omega_1, s_2) : \beta$

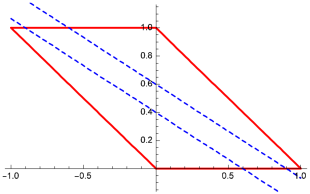


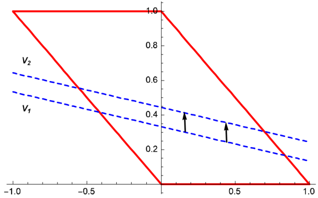
Value of Information Structure

- Value of Information Structure:
  $V(E,\theta) = \theta\pi_1 u_1 + (1-\theta)\pi_2 u_2 - \max\{\theta u_1, (1-\theta)u_2\}$
- Rewrite:
  $V(E,\theta) = \theta(\pi_1 u_1 - \pi_2 u_2) + \pi_2 u_2 - \max\{\theta u_1, (1-\theta)u_2\}$
  1. $\pi_2 u_2$: baseline informativeness (identification)
  2. $\pi_1 u_1 - \pi_2 u_2$: relative informativeness (classification)
- Example: $V(E,\theta) = \theta(\pi_1 - \pi_2) + \pi_2 - \max\{\theta, (1-\theta)\}$

Set of Optimal Experiments

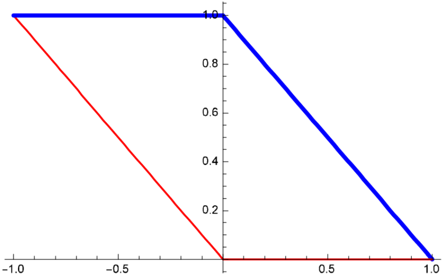
- $V(\pi_1, \pi_2, \theta) = \theta(\pi_1 - \pi_2) + \pi_2 - \max\{\theta, (1-\theta)\}$
- $V(\pi_1 + \delta, \pi_2 + \delta, \theta) - V(\pi_1, \pi_2, \theta) = \delta$
- higher θ have stronger preference for differential $\pi_1 - \pi_2$





Structure of Optimal Menu

- Binary Example gives some intuitive answers to the structure of optimal menu:
  - minimizing type I error: $\pi_2 = 1$
  - minimizing type II error: $\pi_1 = 1$
- only type I/II error



Structure of Optimal Menu

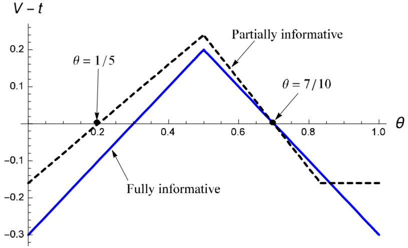
Optimal Menu and Non-Dispersed Information

1. The fully informative experiment $\bar{E}$, $\pi_{ii} = 1$ for all i, is always part of the optimal menu.
2. Every experiment in an optimal menu is non-dispersed, i.e., $\pi_{ij} = 0$ for some $i \neq j$
3. Every experiment in an optimal menu is concentrated, i.e., $\pi_{ii} = 1$ for some i (Matching)

- Proof Sketch: (i) rising the price of highest value
  (ii) cosidering the worst state-utility
  (iii) a direct corollary from (ii)
- finite states and actions, possibly continuum of types ($\surd$)

Binary Types and Binary States: First Example

- binary types: $\theta \in \{\frac{2}{10}, \frac{7}{10}\}$ with equal probability



1. no distortion at the top
2. no rent at the bottom
3. corner solution: no rent at the top

Binary Types and Binary States

- high type mean less informative (without experiments): $|\theta^H - \frac{1}{2}| \leq |\theta^L - \frac{1}{2}|$
- congruent: two types are congruent if they choose the same action given their interim belief, otherwise non-congruent
  $(\theta^H - \frac{1}{2})(\theta^L - \frac{1}{2}) \geq 0$
- recall prior probability of high type: $\gamma = Pr(\theta = \theta_H)$

Informativeness of the Optimal Experiment

Informativeness of the Optimal Experiment

1. congruent priors: the perfectly informative experiment only + both participation only if $\gamma \leq \bar{\gamma} \triangleq \frac{1-\theta^L}{1-\theta^H}$
2. non congruent priors:
   1. the perfectly informative experiment only + both participation only if $\gamma \leq \bar{\gamma}$
   2. high buys the perfect and low buys the partial with $\pi_1 = \frac{2\theta^H - 1}{\theta^H - \theta^L}$ and $\pi_2 = 1$
3. Comparative statics of π_1:
   1. decreasing in the frequency γ of the high type
   2. decreasing in the precision of the low type's prior belief $|\theta^L - \frac{1}{2}|$
   3. increasing in the precision of $|\theta^H - \frac{1}{2}|$ when priors are congruent or the menu is discriminatory

Continuum of Types

- suppose $u_1 = u_2 = 1$
- $q = \pi_1 - \pi_2$, $sgn(q)$ can reveal which $\pi_i = 1$
- $V(q, \theta) = [\theta q - \max\{q, 0\} + \min\{\theta, (1-\theta)\}]^+$
  1. single-crossing suggests q increasing in θ
  2. types $\theta = 0$ and $\theta = 1$ receive zero rents.
  3. consider type $\theta = \frac{1}{2}$, derive additional condition.

Proposition (Necessary Conditions)

For Implementable and Responsive $q(\theta)$:

1. $q(\theta) \in [0, 1]$ is non-decreasing
2. $\int_0^1 q(\theta) d\theta = 0$

Optimal Menu: Cardinality

- $\max_{q(\theta)} \int_0^1 \left[\left(\theta f(\theta) + F(\theta) \right) q(\theta) - \max\{ q(\theta), 0 \} \right] f(\theta) \mathrm{d}\theta$
- Constraints:
  1. $q(\theta) \in [0, 1]$ is non-decreasing
  2. $\int_0^1 q(\theta)\, d\theta = 0$
- Piecewise linear (concave) problem with integral constraint.
  the optimal experiments take values at the kinks
- Absent the integral constraint, corner solutions: all-or-nothing information, flat price

Seller's Problem

Optimal Menu

An optimal menu consists of at most two experiments (coarse menu).

1. The first experiment is fully informative.
2. The second experiment is locally non-dispersed and locally noisefree.

- a continuum of types - yet only a binary choice is provided (coarse menu)
- Optimal mechanism involves at most 2 bunching intervals
- Type $\theta = \frac{1}{2}$ need not get efficient $q = 0$:

Optimal Allocation Rule

- virtual value: $\phi^-(\theta) \triangleq \theta f(\theta) + F(\theta)$ and $\phi^+(\theta) \triangleq (\theta - 1) f(\theta) + F(\theta)$
- ironed virtual value: $\bar{\phi}^-$ and $\bar{\phi}^+$
- λ^*: the multiplier on the integral constraint (shadow cost of providing higher quantity)

Optimal Allocation Rule

The menu $\{q^*(\theta)\}$ is optimal if and only if the following conditions hold:

1. There exists $\lambda^* > 0$ such that, for all θ
$q^*(\theta) \in \arg\max_{q \in [-1,1]} \left[\int_0^q \left(\bar{\phi}(\theta, x) - \lambda^*\right) dx\right]$ for all θ
2. $\{q^*(\theta)\}$ has the pooling property and satisfies integral constraint

Optimal Allocation Rule

COROLLARY 1 Single-Item Menu

1. Almost all types have congruent priors, i.e., $F(\theta^*) \in \{0, 1\}$;
2. The monopoly price for experiment $\bar{E}$ is equal on $[0, \theta^*]$ and $[\theta^*, 1]$;
3. Both virtual values ϕ^- and ϕ^+ are strictly increasing.

- A second experiment is offered only if ironing is required.

Only Fully Informative Experiment: Uniform Distribution

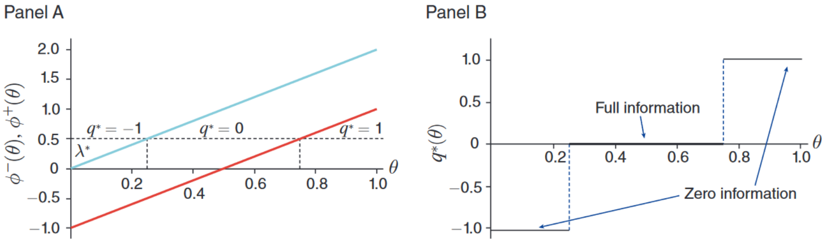


Figure 5. Optimal Allocation with Uniform Distribution

Ironing:Beta Distribution

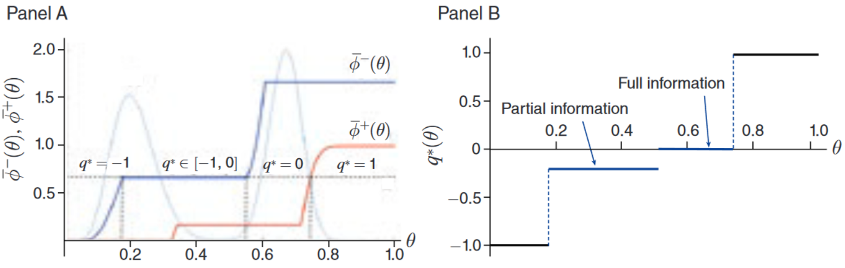


FIGURE 6. PROBABILITY DENSITY FUNCTION AND OPTIMAL ALLOCATION

Conclusion

Conclusion

- selling incremental information to privately informed buyers.
- costless acquisition and transmission, free degrading
  information is inherently rich and can be modified in many ways
- uninterested seller – packaging problem
- bayesian problem for buyers linear in probabilities: limited use of versioning
  horizontal differentiation that widens the seller's scope for price discrimination.
- screening across agents through directional information
- **Insights**: belief and pricing – screening incorporating BCE successfully