CMSC 35401: The Interplay of Economics and ML (Winter 2024)

The Value and Pricing of Information

Instructor: Haifeng Xu

THE UNIVERSITY OF CHICAGO

1890

Outline

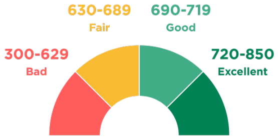
- Bayesian Persuasion and Information Selling
- Sell to a Single Decision Maker
- Sell to Multiple Decision Makers

Motivation: Selling Information

- Car/house inspections

- Financial advices

- Credit report



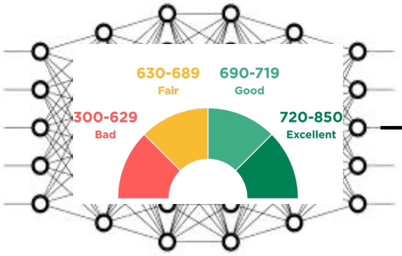
- Consumer data

Motivation: Selling Information

- Car/house inspections

- Financial advices

- Credit report



→ Prob. of default

- Consumer data

→ Prob. of purchase/conversion

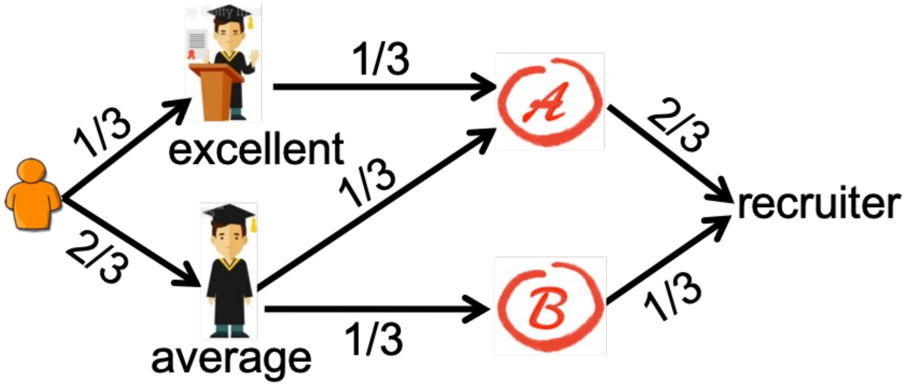
Persuasion vs Information Selling

- In persuasion, we selectively reveal information to induce actions that we like

When selling information, we reveal information for a profit

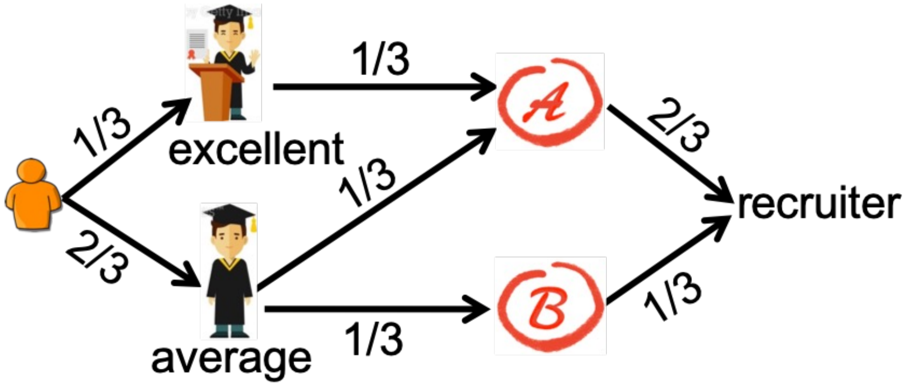
Recap: Model of Bayesian Persuasion

- Two players: persuader (Sender, she), decision maker (Receiver he)
  - Example: advisor = sender, recruiter = receiver
- Receiver looks to take an action $i \in [n] = \{1, 2, \ldots, n\}$
  - Receiver utility $r(i, \theta)$ $\theta \in \Theta$ is a random state of nature
  - Sender utility $s(i, \theta)$
- Both players know $\theta \sim prior\ dist. \mu$, but Sender has an informational advantage – she can observe realization of θ
- Sender reveal partial information via a signaling scheme



(Simplified) Model of Selling Information

- Two players: seller ~~persuader~~ (Sender, she), decision maker (Receiver he)
  - Example: advisor = sender, recruiter = receiver
- Receiver looks to take an action $i \in [n] = \{1, 2, \dots, n\}$
  - Receiver utility $r(i, \theta)$ $\theta \in \Theta$ is a random state of nature
  - Sender utility ~~$s(i, \theta)$~~ payment from the receiver
- Both players know $\theta \sim prior\ dist.\ \mu$, but Sender has an informational advantage – she can observe realization of θ
- Sender reveal partial information via a signaling scheme



How to Sell Information Optimally?

- For any signaling scheme, seller knows how much it improves buyer's expected utility
  - The value of any signaling scheme is known

    1. Receiver utility under no information: $\max_i \sum_{\theta\in\Theta} r(i,\theta)\cdot\mu(\theta)$
    2. Receiver utility under any π: $\sum_\sigma \Pr(\sigma)\cdot R(\sigma)$

    where $R(\sigma) = \max_{i\in[n]} [\sum_{\theta\in\Theta} r(i,\theta)\cdot\frac{\pi(\sigma,\theta)\cdot\mu(\theta)}{\Pr(\sigma)}]$

- How to maximize revenue?
  - Reveal full information helps the buyer the most. Why?
  - So OPT is to charge him following amount and then reveal θ directly

$$\text{Payment} = \boxed{\sum_{\theta\in\Theta}\mu(\theta)\cdot[\max_i u(i,\theta)] -} \max_i \sum_{\theta\in\Theta}\mu(\theta)\cdot u(i,\theta)$$

Buyer expected utility if learns θ precisely

How to Sell Information Optimally?

- For any signaling scheme, seller knows how much it improves buyer's expected utility
  - The value of any signaling scheme is known

  1. Receiver utility under no information: $\max_i \sum_{\theta \in \Theta} r(i,\theta) \cdot \mu(\theta)$
  2. Receiver utility under any π: $\sum_\sigma \Pr(\sigma) \cdot R(\sigma)$

     where $R(\sigma) = \max_{i \in [n]} \left[\sum_{\theta \in \Theta} r(i,\theta) \cdot \frac{\pi(\sigma,\theta) \cdot \mu(\theta)}{\Pr(\sigma)}\right]$

- How to maximize revenue?
  - Reveal full information helps the buyer the most. Why?
  - So OPT is to charge him following amount and then reveal θ directly

$$\text{Payment} = \sum_{\theta \in \Theta} \mu(\theta) \cdot [\max_i u(i,\theta)] - \max_i \sum_{\theta \in \Theta} \mu(\theta) \cdot u(i,\theta)$$

Q: Are we done?

No – in pricing problems, we typically do not know how much buyer values our "product"

Outline

- Bayesian Persuasion and Information Selling
- Sell to a Single Decision Maker
- Sell to Multiple Decision Makers

(True) Model of Selling Information

- Sender = seller, Receiver = buyer who is a decision maker
- Buyer takes an action $i \in [n] = \{1, \cdots, n\}$
- Buyer has a utility function $u(i, \theta; t)$ where
  - $\theta \sim dist.\ \mu$ is a random state of nature
  - $t \sim dist. f$ captures buyer's (private) utility type

Remarks:

- u, μ, f are public knowledge
- Assume θ, t are independent
- Seller observes θ but does not know buyer's type t
- Buyer knows his own type t but does not know θ

Key Challenge

The class of mechanisms is too broad

- The mechanism will: (1) elicit private info from buyer; (2) reveal info based on realized θ; (3) charge buyer
- May interact with buyer for many rounds
- Buyer may misreport his private type t

Key Challenge

The class of mechanisms is too broad

. . . but, at the end of the day, the buyer of type t is charged some amount x_t in expectation and learns a posterior belief about θ

Theorem (Revelation Principle). Any information selling mechanism is “equivalent” to a direct and truthful revelation mechanism:

1. Ask buyer to report type t
2. Charge buyer x_t and reveal info to buyer via signaling scheme π_t that use n signals (as action recommendations)

Moreover, the mechanism is incentive compatible (IC) – it is the buyer’s best interest to truthfully report t

- Optimal mechanism reduces to computing an IC menu $\{x_t, \pi_t\}_t$
- Proof omitted here

The Optimal Mechanism

The Consulting Mechanism [CXZ, SODA'20]

1. Elicit buyer type t
2. Charge buyer x_t
3. Observe realized state θ and recommend action i to the buyer with probability $\pi_t(\sigma_i, \theta)$

- Will be incentive compatible – reporting true t is optimal
- The recommended action is guaranteed to be the optimal action for buyer t given his information
- $\{x_t, \pi_t\}_t$ is public knowledge, and computed by LP

Theorem. Consulting mechanism with $\{x_t, \pi_t\}_t$ computed by the following program is the optimal mechanism.

Computing the Optimal Mechanism

Optimal $\{x_t, \pi_t\}_t$ can be computed by a convex program

- Variables: $\pi_t(\sigma_i, \theta)$ = prob of sending σ_i conditioned on θ for each t
- Variable x_t is the payment from buyer type t

Expected revenue

$$
\begin{array}{lll}
\max & \boxed{\sum_t f(t) \cdot x_t} & \\
\text{s.t.} & \sum_i \left[\sum_\theta \mu(\theta)\pi_t(\sigma_i,\theta)u(i,\theta;t)\right] - x_t & \\
 & \quad \geq \sum_i \max_j \left[\sum_\theta \mu(\theta)\pi_{t'}(\sigma_i,\theta)u(j,\theta;t)\right] - x_{t'}, & \text{for } t' \neq t \\
 & \sum_i \left[\sum_\theta \mu(\theta)\pi_t(\sigma_i,\theta)u(i,\theta;t)\right] - x_t \geq \max_i \sum_\theta \mu(\theta)u(i,\theta;t), & \text{for } t \\
 & \sum_\theta \mu(\theta)\pi_t(\sigma_i,\theta)u(i,\theta;t) \geq \sum_\theta \mu(\theta)\pi_t(\sigma_i,\theta)u(j,\theta;t), & \text{for } i \neq j, t \\
 & \sum_i \pi_t(\sigma_i,\theta) = 1, & \text{for } \theta, t \\
 & \pi_t(\sigma_i,\theta) \geq 0, & \text{for } t, \sigma_i, \theta
\end{array}
$$

Computing the Optimal Mechanism

Optimal $\{x_t, \pi_t\}_t$ can be computed by a convex program

- Variables: $\pi_t(\sigma_i, \theta)$ = prob of sending σ_i conditioned on θ for each t
- Variable x_t is the payment from buyer type t

Truthfully reporting true t is optimal

$$
\begin{array}{lll}
\max & \sum_t f(t) \cdot x_t & \\
\text{s.t.} & \boxed{\begin{array}{l} \sum_i \left[\sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \right] - x_t \\ \qquad \geq \sum_i \max_j \left[\sum_\theta \mu(\theta) \pi_{t'}(\sigma_i, \theta) u(j, \theta; t) \right] - x_{t'}, \end{array}} & \text{for } t' \neq t \\
& \sum_i \left[\sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \right] - x_t \geq \max_i \sum_\theta \mu(\theta) u(i, \theta; t), & \text{for } t \\
& \sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \geq \sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(j, \theta; t), & \text{for } i \neq j, t \\
& \sum_i \pi_t(\sigma_i, \theta) = 1, & \text{for } \theta, t \\
& \pi_t(\sigma_i, \theta) \geq 0, & \text{for } t, \sigma_i, \theta
\end{array}
$$

Computing the Optimal Mechanism

Optimal $\{x_t, \pi_t\}_t$ can be computed by a convex program

- Variables: $\pi_t(\sigma_i, \theta)$ = prob of sending σ_i conditioned on θ for each t
- Variable x_t is the payment from buyer type t

Participation is no worse than not

$$
\begin{array}{lll}
\max & \sum_t f(t) \cdot x_t & \\
\text{s.t.} & \sum_i \left[\sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \right] - x_t & \\
 & \qquad \geq \sum_i \max_j \left[\sum_\theta \mu(\theta) \pi_{t'}(\sigma_i, \theta) u(j, \theta; t) \right] - x_{t'}, & \text{for } t' \neq t \\
 & \boxed{\sum_i \left[\sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \right] - x_t \geq \max_i \sum_\theta \mu(\theta) u(i, \theta; t),} & \text{for } t \\
 & \sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \geq \sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(j, \theta; t), & \text{for } i \neq j, t \\
 & \sum_i \pi_t(\sigma_i, \theta) = 1, & \text{for } \theta, t \\
 & \pi_t(\sigma_i, \theta) \geq 0, & \text{for } t, \sigma_i, \theta
\end{array}
$$

Computing the Optimal Mechanism

Optimal $\{x_t, \pi_t\}_t$ can be computed by a convex program

- Variables: $\pi_t(\sigma_i, \theta)$ = prob of sending σ_i conditioned on θ for each t
- Variable x_t is the payment from buyer type t

$$
\begin{array}{lll}
\max & \sum_t f(t) \cdot x_t & \\
\text{s.t.} & \sum_i \left[\sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \right] - x_t & \\
 & \quad \geq \sum_i \max_j \left[\sum_\theta \mu(\theta) \pi_{t'}(\sigma_i, \theta) u(j, \theta; t) \right] - x_{t'}, & \text{for } t' \neq t \\
 & \sum_i \left[\sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \right] - x_t \geq \max_i \sum_\theta \mu(\theta) u(i, \theta; t), & \text{for } t \\
 & \sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \geq \sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(j, \theta; t), & \text{for } i \neq j, t \\
 & \sum_i \pi_t(\sigma_i, \theta) = 1, & \text{for } \theta, t \\
 & \pi_t(\sigma_i, \theta) \geq 0, & \text{for } t, \sigma_i, \theta
\end{array}
$$

Similar to constraints in persuasion

Computing the Optimal Mechanism

Optimal $\{x_t, \pi_t\}_t$ can be computed by a convex program

- Variables: $\pi_t(\sigma_i, \theta)$ = prob of sending σ_i conditioned on θ for each t
- Variable x_t is the payment from buyer type t

- ➢ A convex function of variables
- ➢ Can be converted to an LP

$$
\begin{array}{lll}
\max & \sum_t f(t) \cdot x_t & \\
& & \\
\text{s.t.} & \sum_i \left[\sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \right] - x_t & \\
& \quad \geq \boxed{\sum_i \max_j \left[\sum_\theta \mu(\theta) \pi_{t'}(\sigma_i, \theta) u(j, \theta; t) \right] - x_{t'},} & \text{for } t' \neq t \\
& \sum_i \left[\sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \right] - x_t \geq \max_i \sum_\theta \mu(\theta) u(i, \theta; t), & \text{for } t \\
& \sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(i, \theta; t) \geq \sum_\theta \mu(\theta) \pi_t(\sigma_i, \theta) u(j, \theta; t), & \text{for } i \neq j, t \\
& \sum_i \pi_t(\sigma_i, \theta) = 1, & \text{for } \theta, t \\
& \pi_t(\sigma_i, \theta) \geq 0, & \text{for } t, \sigma_i, \theta
\end{array}
$$

Practical Mechanisms?

What the mechanism is like?

- Generally, the optimal solution to the previous LP has no structure neither any interpretation
- Nevertheless, closed-form optimal solution is possible for more structured problems

Recall Model of Selling Information

- Sender = seller, Receiver = buyer who is a decision maker (DM)
- Buyer takes an action $i \in [n] = \{1, \cdots, n\}$
- Buyer has a utility function $u(i, q; t)$ where
  - $q \sim dist. \mu$ is a random state of nature
  - $t \sim dist. f$ captures buyer's (private) utility type

Remarks:

- u, μ, f are public knowledge
- Assume q, t are independent

Selling Information to a Binary DM

- Sender = seller, Receiver = buyer who is a decision maker (DM)

- Buyer takes an action $i \in \{0,1\}$: an active action 1 and a passive action 0
  - Active action: approve loan, buyer a car, invest stock X, etc.
- Buyer has a utility function $u(i, q; t)$ where $\begin{cases} u(0, q; t) \equiv 0 \\ u(1, q; t) = v(q, t) \end{cases}$
  - $q \sim dist. \mu$ is a random state of nature
  - $t \sim dist. f$ captures buyer's (private) utility type
- Further assume $v(q, t)$ is linear and non-decreasing in t

Remarks:

- u, μ, f are public knowledge
- Assume q, t are independent

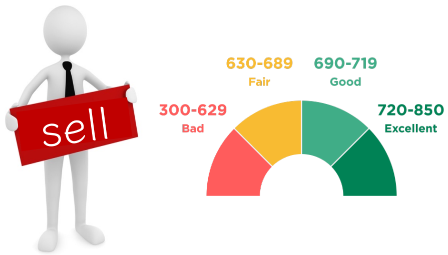
Selling Information to a Binary DM

- Sender = seller, Receiver = buyer who is a decision maker (DM)
- Buyer takes an action $i \in \{0,1\}$: an active action 1 and a passive action 0
  - Active action: approve loan, buyer a car, invest stock X, etc.
- Buyer has a utility function $u(i, q; t)$ where $\begin{cases} u(0, q; t) \equiv 0 \\ u(1, q; t) = v(q, t) \end{cases}$
  - $q \sim dist.\, \mu$ is a random state of nature
  - $t \sim dist.\, f$ captures buyer's (private) utility type
- Further assume $v(q, t)$ is linear and non-decreasing in t

That is: $v(q, t) = v_1(q)[t + \rho(q)]$ for some $v_1(q) \geq 0$

What is the optimal mechanism for this more structured problem?

An Example



- Buyer is a loan company; action is to approve a loan or not
  - If not approving (action 0), payoff is 0
  - If approving (action 1), payoff is

$$v(q,t) = (1 - \underbrace{q}_{\substack{q \in [0,1] \\ \text{default probability}}})\times \underbrace{t}_{\text{Revenue}} - 2 \longrightarrow \text{operation cost}$$

Threshold experiments turn out to suffice

Recall $v(q,t) = v_1(q)[t + \rho(q)]$
(q is the state unknown to buyer)

Def. π_t is a threshold experiment if π_t simply reveals $\rho(q) \geq \theta(t)$ or not for some buyer-type-dependent threshold $\theta(t)$

- Threshold is on $\rho(q)$

The Magical "Virtual Value Functions"

- Virtual value function turns out to naturally arise at optimal mechanism [Myerson'81]

Def. Lower virtual value function: $\underline{\phi}(t) = t - \frac{1-F(t)}{f(t)}$

The Magical "Virtual Value Functions"

- Virtual value function turns out to naturally arise at optimal mechanism [Myerson'81]

Def. Lower virtual value function: $\underline{\phi}(t) = t - \frac{1-F(t)}{f(t)}$

Upper virtual value function: $\bar{\phi}(t) = t + \frac{F(t)}{f(t)}$

Mixed virtual value function: $\phi_c(t) = c\underline{\phi}(t) + (1-c)\bar{\phi}(t)$

Note: "upper" or "lower" is due to

$$\underline{\phi}(t) \leq t \leq \bar{\phi}(t)$$

The Magical "Virtual Value Functions"

- Virtual value function turns out to naturally arise at optimal mechanism [Myerson'81]

Def. Lower virtual value function: $\underline{\phi}(t) = t - \frac{1-F(t)}{f(t)}$

Upper virtual value function: $\bar{\phi}(t) = t + \frac{F(t)}{f(t)}$

Mixed virtual value function: $\phi_c(t) = c\underline{\phi}(t) + (1-c)\bar{\phi}(t)$

- Will assume the virtual value function $\phi(t)$ is monotone (weakly) increasing in t (known as the regularity assumption)
    - Not crucial – if not monotone, there is a standard procedure to adjust it to make it monotone

The Optimal Mechanism

Theorem (Informal, see rigorous statement in [LSX, EC'21]**).**

The mechanism with threshold experiments $\theta^*(t) = -\phi_c^+(t)$ and following payment function represents an optimal mechanism:

$$p^*(t) = \int_{q \in Q} \pi^*(q,t)\mu(q)\, v(q,t)\mathrm{d}q - \int_{t_1}^{t} \int_{q \in Q} \pi^*(q,x)\mu(q)\, v_1(q)\mathrm{d}q \ \mathrm{d}x$$

where constant c is chosen such that

$$\int_{t_1}^{t_2} \int_{q:\rho(q) \geq \phi_c^+(x)} \mu(q) v_1(q)\mathrm{d}q \ \mathrm{d}x = \bar{v}(t_2)$$

Remarks

- Threshold mechanisms are common in real life
  - House/car inspections, stock recommendations: information seller only need to reveal it "passed" or "deserves a buy" or not
- Optimal mechanism has personalized thresholds and payments, tailored to accommodate different level of risk each buyer type can take
  - Different from optimal pricing of physical goods

Remarks

What if seller is restricted to sell the same information to every buyer? How will revenue change?

- Revenue can be arbitrarily worse
- $1/e$-approximation of optimal revenue if the *value of full information* as a function of t is “heavy tail”

Outline

- Bayesian Persuasion and Information Selling

- Sell to a Single Decision Maker

- Sell to Multiple Decision Makers

Challenges

- For single decision maker, more information always helps
  - Recall in persuasion, receiver always benefits from signaling scheme
- A fundamental challenge for selling to multiple buyers is that information does not necessarily help them

Example: More Information Hurts Buyers

- Insurance industry: *insurance company* and *customer*
  - Both are potential information buyers
- Two types of customers: Healthy and Unhealthy
  - Publicly know, Pr(Healthy) = 0.9
- Seller is an information holder, who knows whether any customer is healthy or not

Insurance company

customer

	Sell	Not Sell
Buy	(-10, 10)	(-0, 0)
Not Buy	(0 , 0)	(0 , 0)

Healthy customer

Insurance company

	Sell	Not Sell
Buy	(-10, -50)	(-110, 0)
Not Buy	(-111 , 0)	(-111 , 0)

Unhealthy customer

Example: More Information Hurts Buyers

Insurance company

customer

	Sell	Not Sell
Buy	(-10, 10)	(-0, 0)
Not Buy	(0 , 0)	(0 , 0)

Healthy customer, prob = 0.9

Insurance company

	Sell	Not Sell
Buy	(-10, -50)	(-110, 0)
Not Buy	(-111 , 0)	(-111 , 0)

Unhealthy customer

Q: What happens without seller's information ?

- Customer and insurance company will look at expectation
  - Dominant strategy equilibrium is (Buy, Sell)

	Sell	Not Sell
Buy	(-10, 4)	(-11 , 0)
Not Buy	(-11.1, 0)	(-11.1, 0)

Example: More Information Hurts Buyers

Insurance company

customer

	Sell	Not Sell
Buy	(-10, 10)	(-0, 0)
Not Buy	(0 , 0)	(0 , 0)

Healthy customer, prob = 0.9

Insurance company

	Sell	Not Sell
Buy	(-10, -50)	(-110, 0)
Not Buy	(-111 , 0)	(-111 , 0)

Unhealthy customer

Q: What if seller tells (even only) customer her health status ?

- If Healthy, customer will not buy → utility (0,0) for both
- If Unhealthy, customer will buy → Will not sell, utility (-110,0)
- Customer’s reaction reveals his healthy status
- In expectation (-11, 0), and no insurance was sold ever

Recall previous utilities (-10,4)

Example: More Information Hurts Buyers

Insurance company

customer

	Sell	Not Sell
Buy	(-10, 10)	(-0, 0)
Not Buy	(0 , 0)	(0 , 0)

Healthy customer, prob = 0.9

Insurance company

	Sell	Not Sell
Buy	(-10, -50)	(-110, 0)
Not Buy	(-111 , 0)	(-111 , 0)

Unhealthy customer

Q: What if seller tells (even only) customer her health status ?

Lessons Learned

- Existence of insurance is due to ignorance to our health condition
- Such ignorance benefits both us and insurance companies

Thank You

Haifeng Xu

University of Chicago

haifengxu@uchicago.edu