Mechanism Design

8: Communication with verifiable information

Egor Starkov

Københavns Unversitet

Fall 2023

Introduction

- Throughout the course we dealt with situations where players had some private information that the designer was interested in.
- The players could act based on this private info, but had no way of proving their type (except through the choice of actions).
- Would anything change if the players could disclose hard evidence of their type?
- This lecture is based on (and expands on) the address by Dekel [2016].
- For a broader survey of the literature, see the survey by Dranove and Jin [2010].

Hard evidence

Examples of hard (verifiable) evidence:

- statements about verifiable characteristics of the product:
  - performance,
  - energy efficiency for appliances / fuel efficiency for vehicles,
  - university departments disclose graduates' employability data.
- external ratings and certificates
  - cafes & restaurants have sanitary ratings
  - exchange-traded firms get credit ratings
  - videogames and movies get age ratings, and also critics' reviews

This slide deck:

Disclosure game: basic version [Grossman, 1981]

Let's start with a disclosure game: no design, simply an exploration of how the sender would behave.

1. A firm has a product of privately known quality $\theta \in \Theta$, chooses whether to show a certificate that verifies θ.

Disclosure game: basic version [Grossman, 1981]

Let's start with a disclosure game: no design, simply an exploration of how the sender would behave.

1. A firm has a product of privately known quality $\theta \in \Theta$, chooses whether to show a certificate that verifies θ.
2. A consumer observes evidence (if any) and updates their belief $\phi_c \in \Delta(\Theta)$.

Disclosure game: basic version [Grossman, 1981]

Let's start with a disclosure game: no design, simply an exploration of how the sender would behave.

1. A firm has a product of privately known quality $\theta \in \Theta$, chooses whether to show a certificate that verifies θ.
2. A consumer observes evidence (if any) and updates their belief $\phi_c \in \Delta(\Theta)$.
3. The firm's payoff is $\mathbb{E}[\theta|\phi_c]$.

Disclosure game: basic version [Grossman, 1981]

Let's start with a disclosure game: no design, simply an exploration of how the sender would behave.

1. A firm has a product of privately known quality $\theta \in \Theta$, chooses whether to show a certificate that verifies θ.
2. A consumer observes evidence (if any) and updates their belief $\phi_c \in \Delta(\Theta)$.
3. The firm's payoff is $\mathbb{E}[\theta|\phi_c]$.
   - I.e., the firm wants to induce the highest possible belief (so it can charge higher price, get more consumers, etc – not modelled)

If you want an actual, properly defined game, here it is:

Players: a sender (firm) with private type $\theta \in \Theta \subset \mathbb{R}$; a receiver (consumer) with belief $\phi_0 \in \Delta(\Theta)$.

Actions: sender of type θ chooses a message $m \in \{\varnothing, \theta\}$; the receiver observes m and selects $x \in \mathbb{R}$.

- Some models allow for a richer evidence structure: $m \in M(\Theta)$, where $M(\Theta)$ is some collection of subsets of Θ that include θ. I.e., θ can disclose some but not all information about their type. [Milgrom, 1981]

Payoffs: sender's utility is $u_S(x, \theta) = x$; receiver's utility is $u_R(x, \theta) = -(x - \theta)^2$

- So in eqm, the receiver selects $x = \mathbb{E}[\theta | m]$, and the sender chooses m to maximize $\mathbb{E}[\theta | m]$.

Unraveling

Theorem (Unraveling)

In equilibrium, all firm types θ (except for maybe the lowest one) present evidence, so there is full learning.

- Argument for finite Θ:

Unraveling

Theorem (Unraveling)

In equilibrium, all firm types θ (except for maybe the lowest one) present evidence, so there is full learning.

- Argument for finite Θ:
- Suppose statement is not true: there is a set of types Θ_S that stay silent (do not disclose).

Unraveling

Theorem (Unraveling)

In equilibrium, all firm types θ (except for maybe the lowest one) present evidence, so there is full learning.

- Argument for finite Θ:
- Suppose statement is not true: there is a set of types Θ_S that stay silent (do not disclose).
- Then the firm's payoff when hiding evidence is $\mathbb{E}[\theta|\phi_0, \Theta_S]$, where ϕ_0 is the consumer's prior belief.

Unraveling

Theorem (Unraveling)

In equilibrium, all firm types θ (except for maybe the lowest one) present evidence, so there is full learning.

- Argument for finite Θ:
- Suppose statement is not true: there is a set of types Θ_S that stay silent (do not disclose).
- Then the firm's payoff when hiding evidence is $\mathbb{E}[\theta|\phi_0, \Theta_S]$, where ϕ_0 is the consumer's prior belief.
- Consider type $z \equiv \max \Theta_S$. Revealing the type (with evidence) yields payoff z, which is higher than the payoff from silence (which is a weighted average of z and lower types).

Unraveling

Theorem (Unraveling)

In equilibrium, all firm types θ (except for maybe the lowest one) present evidence, so there is full learning.

- Argument for finite Θ:
- Suppose statement is not true: there is a set of types Θ_S that stay silent (do not disclose).
- Then the firm's payoff when hiding evidence is $\mathbb{E}[\theta|\phi_0, \Theta_S]$, where ϕ_0 is the consumer's prior belief.
- Consider type $z \equiv \max \Theta_S$. Revealing the type (with evidence) yields payoff z, which is higher than the payoff from silence (which is a weighted average of z and lower types).
- So regardless of which types stay silent, the highest of such types would like to separate. Then the highest of the remaining types would like to do the same, etc – this process is called unraveling.

Unraveling: reasons and robustness

- The opportunity to present evidence leads to all information being revealed.
- Note the buyer-designer would not be able to get this result without evidence:
  - As we discussed, our elicitation methods relied on different agent types θ having different preferences (e.g., single-crossing preferences over multidimensional outcomes, non-monotone preferences over one-dimensional outcome).
  - In this example, only the buyer's preferences depend on the type (or so the story suggests) – note that the firm gets $\mathbb{E}[\theta|\phi_c]$ regardless of true θ – so all types θ have the exact same reporting incentives!

Unraveling: reasons and robustness

- The opportunity to present evidence leads to all information being revealed.
- Note the buyer-designer would not be able to get this result without evidence:
  - As we discussed, our elicitation methods relied on different agent types θ having different preferences (e.g., single-crossing preferences over multidimensional outcomes, non-monotone preferences over one-dimensional outcome).
  - In this example, only the buyer's preferences depend on the type (or so the story suggests) – note that the firm gets $\mathbb{E}[\theta|\phi_c]$ regardless of true θ – so all types θ have the exact same reporting incentives!
- The result works for interval Θ too, though the argument is slightly more subtle.
- We will now look at a couple of variations where unraveling breaks.

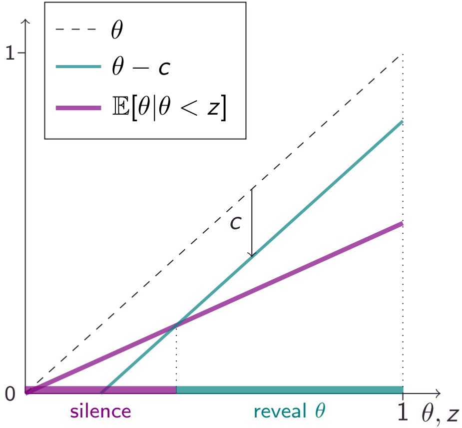
This slide deck:

Costs of disclosure [Verrecchia, 1983]

- Continue the previous example, but suppose now that showing evidence costs $c > 0$ for the firm.
- Then for the low enough types, the profit from showing evidence is not worth the cost.

Costs of disclosure [Verrecchia, 1983]

- Continue the previous example, but suppose now that showing evidence costs $c > 0$ for the firm.
- Then for the low enough types, the profit from showing evidence is not worth the cost.
- **Example:** $\theta \sim U[0,1]$. Suppose types Θ_S silent. Payoff from silence is $\mathbb{E}[\theta \mid \theta \in \Theta_S]$, indep of θ; from disclosure is $\theta - c$, incr in $\theta \Rightarrow$ high θ disclose, low θ silent. Cutoff type z must be indifferent: $z - c = \mathbb{E}[\theta \mid \theta < z] = \frac{z}{2} \Rightarrow z = 2c$. Eqm: $\theta \in [0, 2c)$ silent, $\theta \in [2c, 1]$ disclose.



Uncertain evidence [Dye, 1985, Jung and Kwon, 1988]

- Now return to the case $c = 0$, but the firm only has evidence with probability $\lambda < 1$.
- With probability $1 - \lambda$ the firm has no evidence and is forced to stay silent.
- Then, if types Θ_S stay silent (even with evidence):

$$\mathbb{E}[\theta \mid \text{silence}] = \frac{\lambda \mathbb{P}(\Theta_S)\mathbb{E}[\theta \mid \Theta_S] + (1-\lambda)\mathbb{E}[\theta]}{\lambda \mathbb{P}(\Theta_S) + (1-\lambda)}.$$

  This is higher than in the baseline ($\lambda = 1$), because the consumer understands the firm may not have any evidence. So profit from disclosure is smaller.

- This may again lead to some low types staying silent (pretending to have no evidence).

Uncertain evidence: example

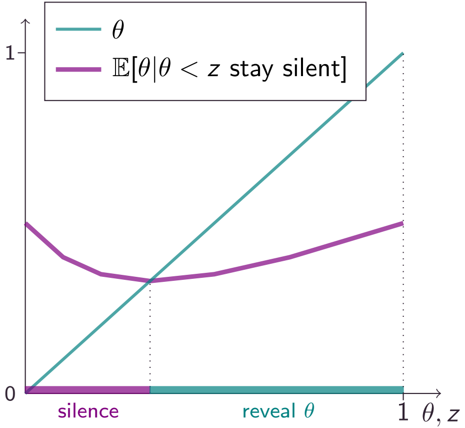
- Suppose again $\theta \sim U[0,1]$ and let $\lambda = 3/4$.
- In eqm, types $\theta \in [z,1]$ disclose their type if they can; types $\theta \in [0,z)$ always silent (same argument as in costly disclosure). Then using LIE and Bayes' rule,

Uncertain evidence: example

- Suppose again $\theta \sim U[0,1]$ and let $\lambda = 3/4$.
- In eqm, types $\theta \in [z,1]$ disclose their type if they can; types $\theta \in [0,z)$ always silent (same argument as in costly disclosure). Then using LIE and Bayes' rule,

$$\mathbb{E}[\theta \mid \text{silence}] = \frac{3z^2+1}{2(3z+1)}$$

- Type θ discloses if $\theta \geq \mathbb{E}[\theta \mid \text{silence}]$ and silent otherwise. Hence $z = \mathbb{E}[\theta \mid \text{silence}] \iff z = 1/3$.
- Eqm: types $\theta \in [1/3, 1]$ disclose if they can; types $\theta \in [0, 1/3)$ always stay silent.



Naive receiver [Milgrom and Roberts, 1986]

- Go back to the base setup, but assume that with probability $\pi \in [0, 1]$ the receiver/consumer is naïve (or, as Eyster and Rabin [2005] call them, cursed).
- A cursed receiver makes no inference from silence: $\mathbb{E}[\theta \mid \text{silence}] = \mathbb{E}[\theta]$; is otherwise same as a rational receiver.

Naive receiver [Milgrom and Roberts, 1986]

- Go back to the base setup, but assume that with probability $\pi \in [0,1]$ the receiver/consumer is naïve (or, as Eyster and Rabin [2005] call them, cursed).
- A cursed receiver makes no inference from silence: $\mathbb{E}[\theta \mid \text{silence}] = \mathbb{E}[\theta]$; is otherwise same as a rational receiver.
- Then if $\theta \sim U[0,1]$ and the firm reveals θ iff $\theta \geq z$, eqm cutoff z must be such that:

$$z = \pi\mathbb{E}[\theta] + (1-\pi)\mathbb{E}[\theta \mid \Theta_S] = \frac{\pi}{2} + (1-\pi)\frac{z}{2} \qquad \Rightarrow \qquad z = \frac{\pi}{1+\pi}.$$

So in equilibrium, the firm reveals θ only if $\theta > \frac{\pi}{1+\pi}$
(i.e., always reveals when $\pi = 0$; reveals only if $\theta > \mathbb{E}[\theta]$ when $\pi = 1$).

Dislosure: Interim conclusion

- In the basic game with evidence, unraveling leads to full information revelation.
- Unraveling can be tamed in many ways, including disclosure costs, naïveté, or receivers allowing for a chance of sender having no evidence. (There are, of course, not the only reasons; see Dranove and Jin [2010] for more.)
- Even in those later cases, the idea is simple: reveal good news, hide bad news.
- Think of evidence and incentives to reveal it as an additional tool in your information extraction toolbox.

Good news and bad news

The asymmetry between **good & bad news** is fun to explore:

- it means that silence is bad news (but humans are bad at inferring it, see Jin, Luca, and Martin [2021]),

Good news and bad news

The asymmetry between **good & bad news** is fun to explore:

- it means that silence is bad news (but humans are bad at inferring it, see Jin, Luca, and Martin [2021]),
- that silence leaves more uncertainty than good announcements [Shin, 2003],

Good news and bad news

The asymmetry between **good & bad news** is fun to explore:

- it means that silence is bad news (but humans are bad at inferring it, see Jin, Luca, and Martin [2021]),
- that silence leaves more uncertainty than good announcements [Shin, 2003],
- that bad news are revealed in bunches [Acharya, DeMarzo, and Kremer, 2011]

Good news and bad news

The asymmetry between **good & bad news** is fun to explore:

- it means that silence is bad news (but humans are bad at inferring it, see Jin, Luca, and Martin [2021]),
- that silence leaves more uncertainty than good announcements [Shin, 2003],
- that bad news are revealed in bunches [Acharya, DeMarzo, and Kremer, 2011]
- that the desire to have good news to disclose leads to excessive risk-taking [Ben-Porath, Dekel, and Lipman, 2018]

Good news and bad news

The asymmetry between **good & bad news** is fun to explore:

- it means that silence is bad news (but humans are bad at inferring it, see Jin, Luca, and Martin [2021]),
- that silence leaves more uncertainty than good announcements [Shin, 2003],
- that bad news are revealed in bunches [Acharya, DeMarzo, and Kremer, 2011]
- that the desire to have good news to disclose leads to excessive risk-taking [Ben-Porath, Dekel, and Lipman, 2018]

- There are also a few reasons why sender might want to voluntarily reveal bad news, see an overview in Smirnov and Starkov [2022]

This slide deck:

1. Disclosure with one sender
2. Disclosure with one sender: variations
3. Disclosure with one sender: mechanism design
4. Disclosure with many senders

Evidence and Mechanism design

- We have looked at disclosure games so far, where receiver/designer had to be sequentially rational (maximize $\mathbb{E}u_R$ given $m \Rightarrow$ choose $x = \mathbb{E}[\theta|m]$).
- What if we take a mechanism design perspective? I.e., suppose the receiver can commit to a decision rule $x(m)$. Can this help produce a better outcome?
  - E.g., in the **Dye setting** (uncertain evidence).
  - Note: choosing a strategy $x(m)$ in the no-commitment game can be seen as "a mechanism design problem with a sequential rationality constraint". Extra constraint = worse outcome. Or is it?

Evidence and Mechanism design

- We have looked at disclosure games so far, where receiver/designer had to be sequentially rational (maximize $\mathbb{E}u_R$ given $m \Rightarrow$ choose $x = \mathbb{E}[\theta|m]$).
- What if we take a mechanism design perspective? I.e., suppose the receiver can commit to a decision rule $x(m)$. Can this help produce a better outcome?
  - E.g., in the **Dye setting** (uncertain evidence).
  - Note: choosing a strategy $x(m)$ in the no-commitment game can be seen as "a mechanism design problem with a sequential rationality constraint". Extra constraint = worse outcome. Or is it?

- Hart, Kremer, and Perry [2017] show this is not the case.
- Their result: under some conditions on the sender's evidence and the receiver's preferences (that our setting satisfies):
  - disclosure game has a unique equilibrium,
  - the optimal disclosure mechanism exists and is unique,
  - the two coincide $\Rightarrow$ **no value from commitment**.

Evidence and Mechanism design 2

- Their result (no value from commitment) relies on concavity of the receiver's preferences $+$ assumptions on evidence structure.
- **Counterexample** with preferences not concave in x given θ:
- Let $\theta \in \{1, 2, ..., 9\}$ with $\phi_0(9) = 0.2$ and $\phi_0(\theta) = 0.1$ for $\theta < 9$.
- Suppose types $\theta < 9$ are verifiable: $m \in \{\varnothing, \theta\}$; type $\theta = 9$ has no evidence: $m = \varnothing$.
- Receiver's payoff is $u_R(x, \theta) = \mathbb{I}\{x = \theta\}$; sender's payoff is still $u_S(x, \theta) = x$.

Evidence and Mechanism design 2

- Their result (no value from commitment) relies on concavity of the receiver's preferences + assumptions on evidence structure.
- **Counterexample** with preferences not concave in x given θ:
- Let $\theta \in \{1, 2, ..., 9\}$ with $\phi_0(9) = 0.2$ and $\phi_0(\theta) = 0.1$ for $\theta < 9$.
- Suppose types $\theta < 9$ are verifiable: $m \in \{\varnothing, \theta\}$; type $\theta = 9$ has no evidence: $m = \varnothing$.
- Receiver's payoff is $u_R(x, \theta) = \mathbb{I}\{x = \theta\}$; sender's payoff is still $u_S(x, \theta) = x$.
- Then commitment outcome is better than no commitment:
  - In eqm **w/o commitment**, after silence the receiver chooses $x(\varnothing) = 9$, so sender never reveals θ. $\Rightarrow \mathbb{E}u_R = 0.2$.
  - **With commitment**, receiver can commit to $x(\varnothing) = 1 \Rightarrow$ sender reveals θ if possible $\Rightarrow \mathbb{E}u_R = 0.8$.

This slide deck:

Many senders, same evidence [Milgrom and Roberts, 1986]

- Go back to the game with **naïve receiver**, but suppose that in addition to the firm, there is now also a competitor.
- Both observe firm quality θ and can disclose it in a verifiable way.
- The competitor's payoff is $-\mathbb{E}[\theta|\phi_c]$, the opposite of the firm's.

Many senders, same evidence [Milgrom and Roberts, 1986]

- Go back to the game with **naïve receiver**, but suppose that in addition to the firm, there is now also a competitor.
- Both observe firm quality θ and can disclose it in a verifiable way.
- The competitor's payoff is $-\mathbb{E}[\theta|\phi_c]$, the opposite of the firm's.

- Then the firm wants to reveal θ iff $\theta > \pi\mathbb{E}[\theta] + (1-\pi)\mathbb{E}[\theta \mid \text{silence}]$.
- Similarly, the competitor wants to reveal θ iff $\theta < \pi\mathbb{E}[\theta] + (1-\pi)\mathbb{E}[\theta \mid \text{silence}]$.
- Regardless of θ, one of them wants to reveal $\Rightarrow$ someone always reveals θ $\Rightarrow$ full info in equilibrium.

Many senders, same evidence

- Conclusion: if players have a conflict of interest and access to the same info, more information is revealed.
- This complements the results without evidence: the receiver can exploit the conflict of interest between players to extract more info (we saw this in the Battaglini [2002] model).

Item allocation with evidence [Ben-Porath, Dekel, and Lipman, 2019]

- Evidence+competition help elicit info even with no common info.
- E.g., consider an **item allocation problem:**
  - N bidders with private, verifiable types θ_i; designer chooses allocation $x \in \Delta(N)$;
  - bidder i's utility: $u_i(x, \theta) = x_i \theta_i$;
  - designer's utility: $u_0(x, \theta) = \sum_{i=1}^{N} x_i \theta_i$ (designer wants to allocate to highest θ_i).

Item allocation with evidence

- **Without evidence,** we'd have to require payments as a proof of high θ_i.
- **With evidence,** can simply ask a player to show proof of their high θ_i
  $\Rightarrow$ efficient allocation is implementable even without transfers!
- The same is true even if $u_i(x, \theta) = \mathbb{I}\{x_i = 1\}$
  - i.e., if players don't care about their θ_i, only the principal does
  - no single-crossing in this case $\Rightarrow$ transfers wouldn't implement the principal's desired allocation
  - Example: investor only wants to fund projects that are genuinely good, but all entrepreneurs think their projects are genuinely good

Item allocation with evidence

- **Without evidence,** we'd have to require payments as a proof of high θ_i.
- **With evidence,** can simply ask a player to show proof of their high θ_i
  $\Rightarrow$ efficient allocation is implementable even without transfers!
- The same is true even if $u_i(x, \theta) = \mathbb{I}\{x_i = 1\}$
  - i.e., if players don't care about their θ_i, only the principal does
  - no single-crossing in this case $\Rightarrow$ transfers wouldn't implement the principal's desired allocation
  - Example: investor only wants to fund projects that are genuinely good, but all entrepreneurs think their projects are genuinely good
- Bottom line:
  - Evidence helps info elicitation
  - Evidence makes commitment power unnecessary

Evidence: Conclusion

- Evidence helps info elicitation, and may even lead to unraveling of all of players' private info
- Some factors (like direct cost of disclosure or a milder penalty for silence) may hinder disclosure; competition stimulates it.
- The intrinsic incentives to disclose evidence may be strong enough to render commitment useless for the principal/receiver
  - Good news for receiver – can achieve commitment outcome even without any commitment power!

References I

V. V. Acharya, P. DeMarzo, and I. Kremer. Endogenous Information Flows and the Clustering of Announcements. *American Economic Review*, 101(7):2955–2979, December 2011. ISSN 0002-8282. doi: 10.1257/aer.101.7.2955.

M. Battaglini. Multiple Referrals and Multidimensional Cheap Talk. *Econometrica*, 70(4): 1379–1401, July 2002. ISSN 1468-0262. doi: 10.1111/1468-0262.00336.

E. Ben-Porath, E. Dekel, and B. L. Lipman. Disclosure and Choice. *The Review of Economic Studies*, 85(3):1471–1501, July 2018. ISSN 0034-6527, 1467-937X. doi: 10.1093/restud/rdx064.

E. Ben-Porath, E. Dekel, and B. L. Lipman. Mechanisms With Evidence: Commitment and Robustness. *Econometrica*, 87(2):529–566, 2019. ISSN 0012-9682. doi: 10.3982/ECTA14991.

References II

E. Dekel. On Evidence in Games and Mechanism Design, 2016. Econometric Society Presidential Address.

D. Dranove and G. Z. Jin. Quality Disclosure and Certification: Theory and Practice. *Journal of Economic Literature*, 48(4):935–963, 2010. ISSN 0022-0515. doi: 10.2307/29779704.

R. A. Dye. Disclosure of nonproprietary information. *Journal of accounting research*, pages 123–145, 1985.

E. Eyster and M. Rabin. Cursed Equilibrium. *Econometrica*, 73(5):1623–1672, September 2005. ISSN 0012-9682, 1468-0262. doi: 10.1111/j.1468-0262.2005.00631.x.

S. J. Grossman. The informational role of warranties and private disclosure about product quality. *The Journal of Law and Economics*, 24(3):461–483, 1981.

References III

S. Hart, I. Kremer, and M. Perry. Evidence Games: Truth and Commitment. *American Economic Review*, 107(3):690–713, March 2017. ISSN 0002-8282. doi: 10.1257/aer.20150913.

G. Z. Jin, M. Luca, and D. Martin. Is No News (Perceived As) Bad News? An Experimental Investigation of Information Disclosure. *American Economic Journal: Microeconomics*, 13 (2):141–173, May 2021. ISSN 1945-7669. doi: 10.1257/mic.20180217.

W.-O. Jung and Y. K. Kwon. Disclosure When the Market Is Unsure of Information Endowment of Managers. *Journal of Accounting Research*, 26(1):146, 1988. ISSN 00218456. doi: 10.2307/2491118.

P. Milgrom and J. Roberts. Relying on the Information of Interested Parties. *The RAND Journal of Economics*, 17(1):18–32, 1986. ISSN 0741-6261. doi: 10.2307/2555625.

References IV

P. R. Milgrom. Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, pages 380–391, 1981.

H. S. Shin. Disclosures and asset returns. *Econometrica*, 71(1):105–133, 2003.

A. Smirnov and E. Starkov. Bad News Turned Good: Reversal under Censorship. *American Economic Journal: Microeconomics*, 14(2):506–560, May 2022. ISSN 1945-7669. doi: 10.1257/mic.20190379.

R. E. Verrecchia. Discretionary disclosure. *Journal of Accounting and Economics*, 5:179–194, January 1983. ISSN 0165-4101. doi: 10.1016/0165-4101(83)90011-3.