Divide and Confer:
Aggregating Information without Verification

James Best Daniel Quigley Maryam Saeedi Ali Shourideh

CMU, Oxford

May 28, 2026

KU Leuven, Applied Theory Conference

Information Aggregation

- Fundamental question in economics: how efficiently do various institutions aggregate information dispersed among large populations?

Information Aggregation

- Fundamental question in economics: how efficiently do various institutions aggregate information dispersed among large populations?
- Many papers show how we can almost perfectly extract information from informed parties even when there are large conflicts of interest.
  - Eliciting expert opinions, voting rules, common value auctions, ...
  - *e.g., Cremer & McLean (1988), Feddersen & Pesendorfer (1997), Krishna & Morgan (2002), McLean & Postlewaite (2004), Gerardi et al. (2009).*

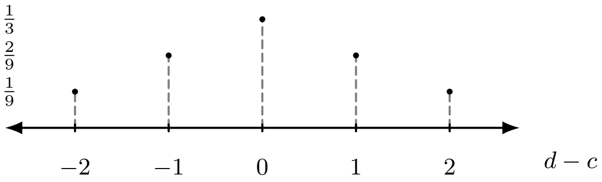
Information Aggregation

- Fundamental question in economics: how efficiently do various institutions aggregate information dispersed among large populations?
- Many papers show how we can almost perfectly extract information from informed parties even when there are large conflicts of interest.
  - Eliciting expert opinions, voting rules, common value auctions, ...
  - *e.g., Cremer & McLean (1988), Feddersen & Pesendorfer (1997), Krishna & Morgan (2002), McLean & Postlewaite (2004), Gerardi et al. (2009).*
- These approaches rely on at least one of the following:
  - Cross-verification: punish (reward) disagreement (agreement).
  - Informational addition: unboundedly increasing the available information by taking the population to infinity.

Informational Division: A Small Example

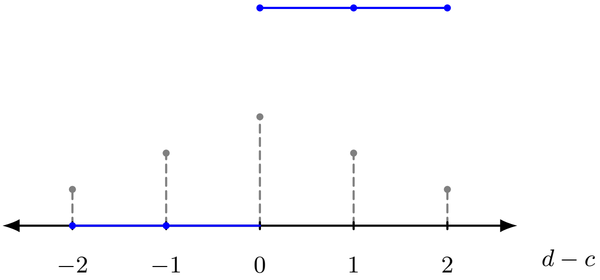
- A **CEO** faces binary decision $a \in \{0, 1\}$: **reject** or **accept** a new project.
- Value of accept depends on cost shock c and demand shock d, each **privately** observed by a department manager
- $c, d \sim U\{-1, 0, 1\}$, independent
- **CEO payoff**: $a(d - c)$
- **Manager payoff**: $a(d - c + 3)$
- **Problem**: CEO may commit to a **direct mechanism** $\sigma : \{-1, 0, 1\}^2 \to \Delta\{0, 1\}$, assigning action as a function of reports $\tilde{c}, \tilde{d}$

Small Population Example



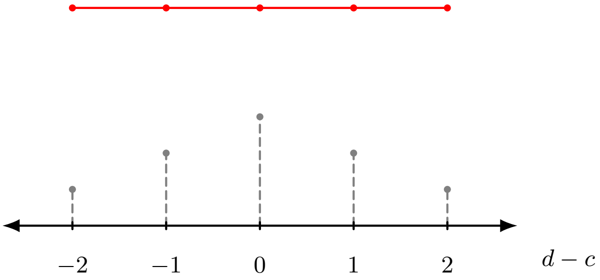
Small Population Example

Receiver Preferred Mechanism $\sigma^R = \mathbf{1}\{d - c \geq 0\}$



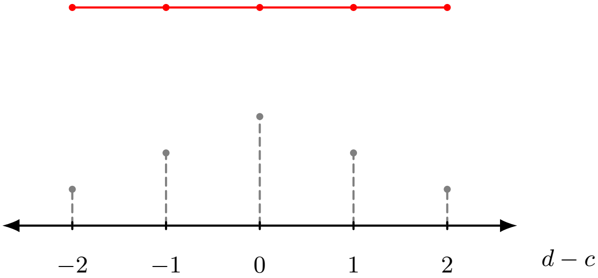
Small Population Example

Sender Preferred Mechanism $\sigma^S = \mathbf{1}\{d - c \geq -3\}$



No Informational Division: 1 Sender

If a single manager observes d and c, cannot do better than σ^S

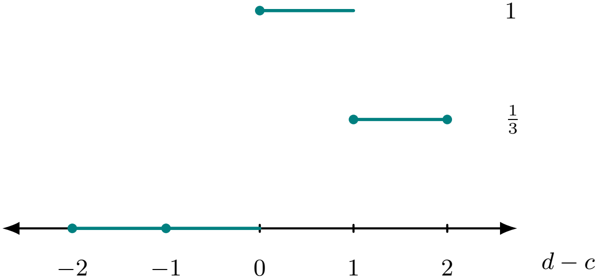


- Sender knows $d - c$, and is the monopoly information source.
- Will always choose report that maximizes σ when $d - c > -3$.

As $d - c > -3$ for sure, σ^S is uninformative in this case—very bad for CEO.

Informational Division: Two Senders

Suppose instead there are two managers: one observes d, the other observes c

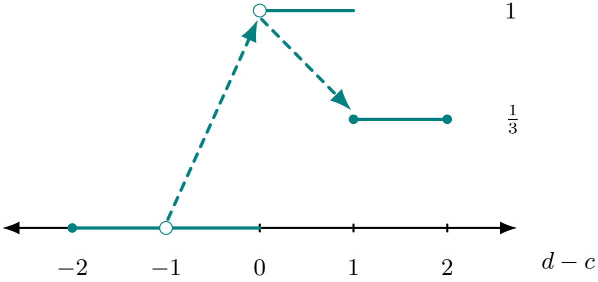


This mechanism:

- is **simple**: outcomes depend only on the payoff-relevant state, $d - c$
- obviously improves CEO payoff, albeit *inefficiently*.
- maintains incentives via *two pivot points*

Informational Division: Two Senders

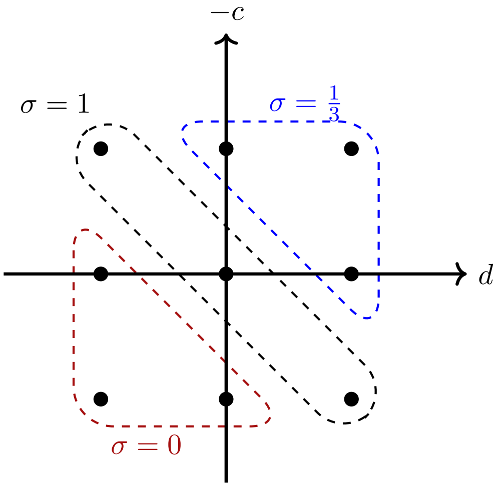
Suppose instead there are two managers: one observes d, the other observes c



This mechanism:

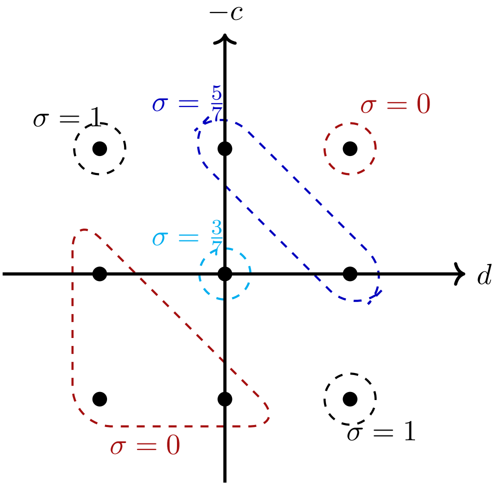
- is **simple**: outcomes depend only on the payoff-relevant state, $d - c$
- obviously improves CEO payoff, albeit *inefficiently*.
- maintains incentives via *two pivot points*

Simple is Restrictive



The Optimal Simple Mechanism.

Optimal Mechanism is Complex



The Fully Optimal Mechanism.

Moving to the Large

Dividing information between two managers benefitted the CEO, even though the information about demand and cost was **not cross-verifiable**.

In general, division among finite set of senders with large type space

- may not always offer the receiver a strict improvement.
- optimal mechanism may be very complex!

Main analysis: scope for decision maker to benefit when information is divided finely among arbitrarily many senders.

Model

- N senders, $i \in \{1, 2, \ldots, N\}$, and a single receiver
- Receiver faces binary decision: accept or reject $a \in \{0, 1\}$
- Each sender i observes private signal s_i
- Signals are i.i.d.:
  - $s_i \in S = \{t_1, \ldots, t_K\}$, with distribution $\mathbf{f} = (f_1, \ldots, f_K)$
  - Assume $\mathbb{E}[s_i] = 0$, $\text{Var}(s_i) = 1$ (normalizations)
- Payoff relevant state: $\omega = \frac{\sum_{i=1}^{N} s_i}{\sqrt{N}}$
- Payoffs:
  - Receiver: $u_R(a, \omega) = a(\omega + r)$
  - Senders: $u_S(a, \omega) = a(\omega + b)$, where $b > \max\{r, 0\}$
- Senders are biased in favour of $a = 1$

Informational Division:

- We capture informational division via the formulation of the state

$$\omega = \frac{\sum_{i=1}^{N} s_i}{\sqrt{N}}.$$

- As N grows:
  - aggregate uncertainty (measured by $\mathrm{Var}\,(\omega) = 1$) remains constant;
  - each sender becomes less informed about ω. Indeed, $\frac{s_i}{\sqrt{N}} \to 0$.

- Additionally: as signals s_i are i.i.d., one agents report cannot be used to verify the report of another.

The Mediated Communication Game

A mediator commits to a (direct) communication mechanism.

$$\sigma : S^N \to \Delta\{0,1\}$$

Senders report $\tilde{s}_i \to \sigma$

Space

σ recommends action $\to$ receiver

Space

receiver chooses a.

We only allow for mediation: no transfers.

Problem

$$\max_{\sigma: S^N \to \Delta\{0,1\}} \mathbb{E}\left[\sigma\left(\mathbf{s}\right)\left(\omega(\mathbf{s}) + r\right)\right] \quad \text{subject to}$$

Truth-telling maximizes each sender payoff. (IC)

Obedience maximizes receiver payoff. (OB)

Reframing the Problem

- Direct mechanisms are unwieldy for large-N analysis
- **Solution**: reframe problem in **frequency domain**

Reframing the Problem

- Direct mechanisms are unwieldy for large-N analysis
- **Solution**: reframe problem in **frequency domain**
- Capture empirical distribution via

$$h_k(\mathbf{s}) = \sqrt{N}\left(\frac{|\{i|s_i = t_k\}|}{N} - f_k\right),$$

- Normalized Empirical Frequencies (NEF): $\mathbf{h}^N(\mathbf{s}) = (h_1(\mathbf{s}), \cdots, h_K(\mathbf{s}))$

Symmetry $\implies$ w.l.o.g., may consider mechanisms in frequency domain

$$\sigma(\mathbf{h}^N)$$

Reframing the Problem

1. Domain ($\mathbb{R}^K$) stays constant, even as N gets large.
2. In the limit as $N \to \infty$, NEFs behave nicely:

$$\mathbf{h}^N \xrightarrow{d} N\left(\mathbf{0}, \Sigma\right),$$

with $\Sigma_{kl} = f_k \left(\mathbf{1}\left[k = l\right] - f_l\right)$. Moreover, $\omega = \mathbf{h}^N \cdot \mathbf{t}$ &

$$\omega \xrightarrow{d} N\left(0, 1\right)$$

Reframing the Problem

1. Domain ($\mathbb{R}^K$) stays constant, even as N gets large.
2. In the limit as $N \to \infty$, NEFs behave nicely:

$$\mathbf{h}^N \xrightarrow{d} N\left(\mathbf{0}, \Sigma\right),$$

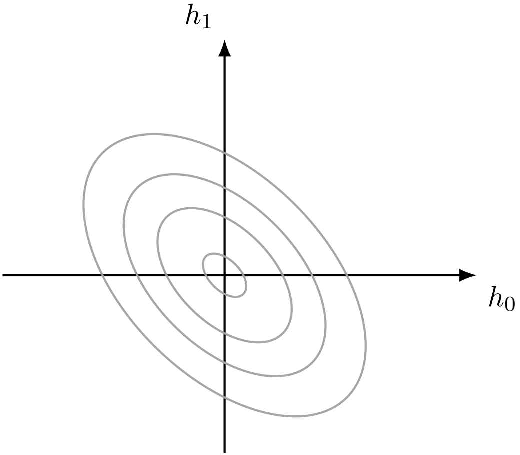
with $\Sigma_{kl} = f_k \left(\mathbf{1}\left[k = l\right] - f_l\right)$. Moreover, $\omega = \mathbf{h}^N \cdot \mathbf{t}$ &

$$\omega \xrightarrow{d} N\left(0, 1\right)$$

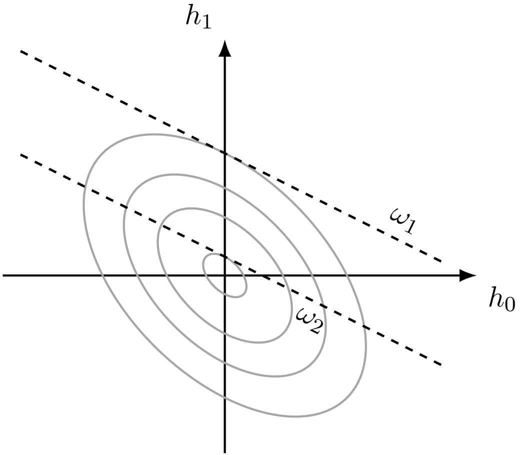
Incentive Compatibility in the Frequency Domain

$$\mathbb{E}\left[\sigma\left(\mathbf{h}^N\right)\left(\mathbf{h}^N \cdot \mathbf{t} + b\right) | s_i = \tilde{s}_i\right] \geq \mathbb{E}\left[\sigma\left(\tilde{\mathbf{h}}^N\right)\left(\mathbf{h}^N \cdot \mathbf{t} + b\right) | s_i, \tilde{s}_i\right], \qquad \text{(F-IC)}$$

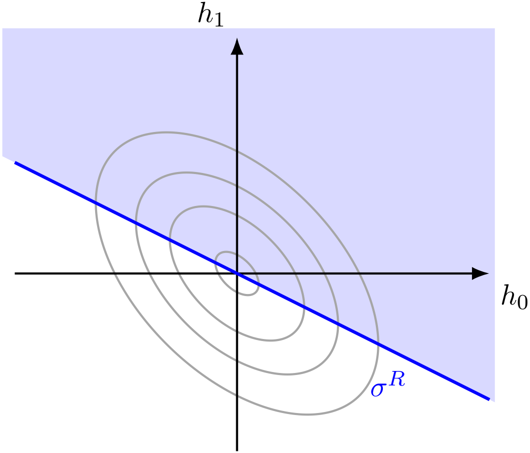
Frequency Domain: 3 types



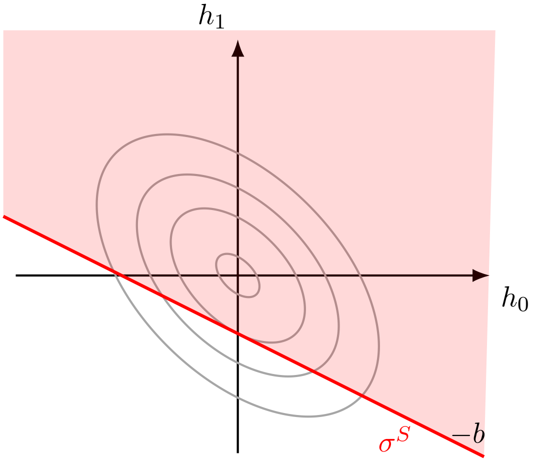
Frequency Domain: 3 types



Frequency Domain: 3 types



Frequency Domain: 3 types



Incentive Compatibility in the Large (ICL)

Definition (Incentive Compatibility in the Large (ICL))

Mechanism σ is ICL if there is a sequence of **incentive compatible** N-sender mechanisms σ^N whose outcome distributions converge as follows:

$$(a^N, \mathbf{h}^N) \xrightarrow{d} (a, \mathbf{h}).$$

Characterizing ICL Mechanisms

Theorem

A mechanism σ is ICL if and only if

$$\frac{dU}{dh_k} \equiv \mathbb{E}\left[\sigma\left(\mathbf{h}\right)\left(\mathbf{h}\cdot\mathbf{t}+b\right)h_k/f_k\right] = \mathbb{E}\left[\sigma\left(\mathbf{h}\right)\right]t_k, \quad 1 \leq k \leq K \qquad \text{(ENV)}$$

$$\mathbb{E}\left[\sigma\left(\mathbf{h}\right)\frac{h_k}{f_k}\right] \text{ non-decreasing in } k, \qquad \text{(MON)}$$

Characterizing ICL Mechanisms

Theorem

A mechanism σ is ICL if and only if

$$\frac{dU}{dh_k} \equiv \mathbb{E}\left[\sigma\left(\mathbf{h}\right)\left(\mathbf{h}\cdot\mathbf{t}+b\right)h_k/f_k\right] = \mathbb{E}\left[\sigma\left(\mathbf{h}\right)\right]t_k, \quad 1 \leq k \leq K \tag{ENV}$$

$$\mathbb{E}\left[\sigma\left(\mathbf{h}\right)\frac{h_k}{f_k}\right] \text{ non-decreasing in } k, \tag{MON}$$

Proof sketch

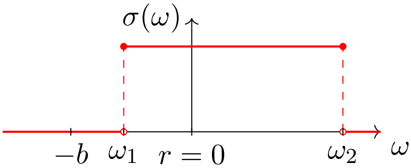
The Optimal Mechanism in the Large

Theorem (Optimal Mechanism)

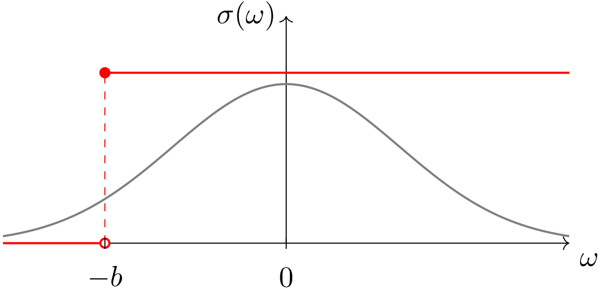
The following simple mechanism is the (essentially) unique optimal ICL mechanism:

$$\sigma^*(\omega) = \mathbf{1}\{\omega \in [\omega_1, \omega_2]\}$$

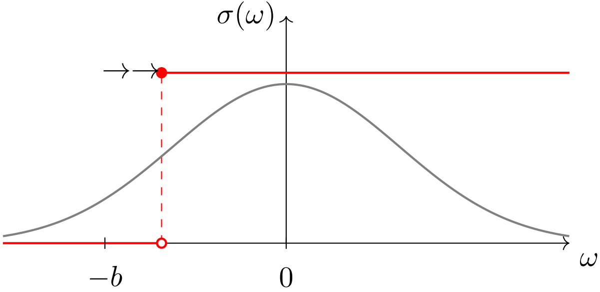
for some cutoffs ω_1, ω_2*, which satisfy* $\omega_1 \in (-b, -r)$, $\omega_2 \in (\omega_1, \infty)$.



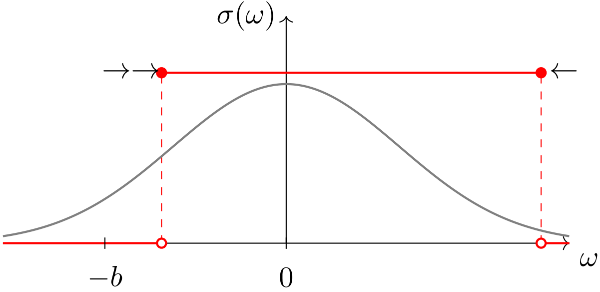
Why are Double Pivot Mechanisms Optimal



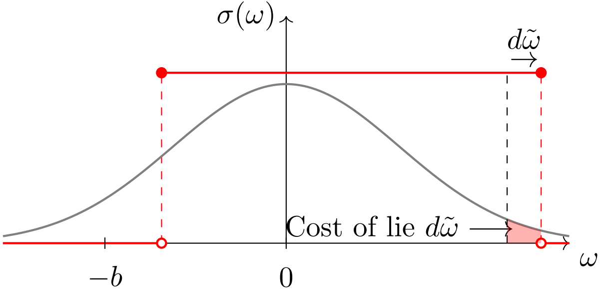
Why are Double Pivot Mechanisms Optimal



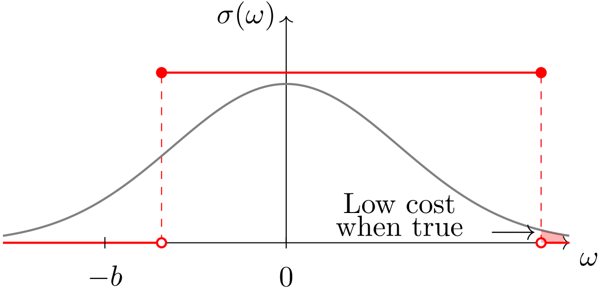
Why are Double Pivot Mechanisms Optimal



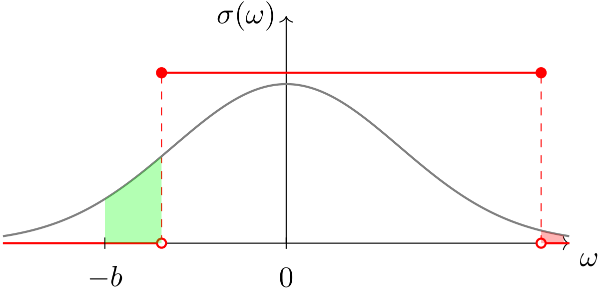
Why are Double Pivot Mechanisms Optimal



Why are Double Pivot Mechanisms Optimal

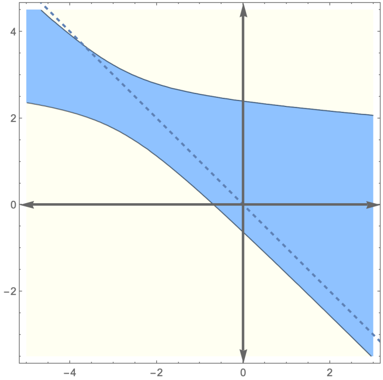


Why are Double Pivot Mechanisms Optimal



Heterogeneous bias

- Finite number of bias classes
- Optimal mechanism simple by class: function of average ω *within each class*, and qualitatively similar.



Related Literature

- **Information Aggregation via Cross-Verification**

  Crémer and McLean (1988); Krishna and Morgan (2001); Battaglini (2002), (2004); Meyer et al. (2019); Gerardi et al. (2009)

- **Aggregation in Voting and Large Economies**

  Feddersen and Pesendorfer (1997)

- **Multi-Sender Models with Binary Types**

  Wolinsky (2002); Kattwinkel and Winter (2024)

- **Convergence Properties of Mechanisms with Rich Data**

  Frick, M., Iijima, R., and Ishii, Y. (2023; 2024)

  - **These papers**: information is **additive** $\implies$ first best attainble as N get large.
  - **Our paper**: information is **divided** $\implies$ first-best unattainable. Can characterize the second-best solution.

Summary

- Introduce informational division to large society aggregation problems.
- **Theorem 1**: Develop new tools for studying large-population mechanism design via the notion of incentive compatibility in the large.
- **Theorem 2**: Characterize optimal mechanism for large populations
  - A simple, interval mechanism is optimal. Payoffs are bounded away from the first best, even as $N \to \infty$.
- Demonstrate that information aggregation in large populations is inefficient when we consider division rather than addition.
  - Suggests that it may be valuable to revisit information aggregation in other settings through the lens of informational division.

Thank you!

Proof Sketch: "Only If" Direction

To see why (ENV) must hold for any ICL mechanism σ:

- Fix N, and let U_k^N denote t_k's truth-telling utility.
- IC for types t_k and t_l respectively, imply familiar condition

$$\mathbb{E}\Big[\sigma^N(\mathbf{h}^N)\Big|t_k\Big] \geq \frac{\sqrt{N}\left(U_k^N - U_l^N\right)}{t_k - t_l} \geq \mathbb{E}\Big[\sigma^N(\mathbf{h}^N)\Big|t_l\Big].$$

Proof Sketch: "Only If" Direction

To see why (ENV) must hold for any ICL mechanism σ:

- Fix N, and let U_k^N denote t_k's truth-telling utility.
- Taking $N \to \infty$...

$$\lim_{N\to\infty} \mathbb{E}\left[\sigma^N(\mathbf{h}^N)\middle| t_k\right] \geq \frac{\sqrt{N}\left(U_k^N - U_l^N\right)}{t_k - t_l} \geq \lim_{N\to\infty} \mathbb{E}\left[\sigma^N(\mathbf{h}^N)\middle| t_l\right].$$

Proof Sketch: "Only If" Direction

To see why (ENV) must hold for any ICL mechanism σ:

- Fix N, and let U_k^N denote t_k's truth-telling utility.
- ... sender private information becomes small

$$\mathbb{E}\Big[\sigma(\mathbf{h})\Big] \geq \lim_{N\to\infty} \frac{\sqrt{N}\,\left(U_k^N - U_l^N\right)}{t_k - t_l} \geq \mathbb{E}\Big[\sigma(\mathbf{h})\Big].$$

Proof Sketch: "Only If" Direction

To see why (ENV) must hold for any ICL mechanism σ:

- Fix N, and let U_k^N denote t_k's truth-telling utility.
- ... sender private information becomes small

$$\mathbb{E}\left[\sigma(\mathbf{h})\right] \geq \lim_{N\to\infty} \frac{\sqrt{N}\left(U_k^N - U_l^N\right)}{t_k - t_l} \geq \mathbb{E}\left[\sigma(\mathbf{h})\right].$$

- Marginal utility can be directly calculated from U_k^N

$$U_k^N = \mathbb{E}\left[\left(\mathbf{h}\cdot\mathbf{t} + b\right)\sigma(\mathbf{h})\middle| t_k\right].$$

Proof Sketch: "Only If" Direction

To see why (ENV) must hold for any ICL mechanism σ:

- Fix N, and let U_k^N denote t_k's truth-telling utility.
- ... sender private information becomes small

$$\mathbb{E}\left[\sigma(\mathbf{h})\right] \geq \lim_{N\to\infty} \frac{\sqrt{N}\left(U_k^N - U_l^N\right)}{t_k - t_l} \geq \mathbb{E}\left[\sigma(\mathbf{h})\right].$$

- Applying Bayes' rule

$$U_k^N = \mathbb{E}\left[(\mathbf{h}\cdot\mathbf{t} + b)\,\sigma(\mathbf{h})\left(1 + \frac{h_k}{\sqrt{N} f_k}\right)\right].$$

- From which easy to see

$$\frac{\mathbb{E}\left[(\mathbf{h}\cdot\mathbf{t} + b)\,\sigma(\mathbf{h})\left(\frac{h_k}{f_k} - \frac{h_l}{f_l}\right)\right]}{\cdots} = \mathbb{E}\left[\sigma(\mathbf{h})\right].$$

Proof Sketch: "Only If" Direction

To see why (ENV) must hold for any ICL mechanism σ:

- Fix N, and let U_k^N denote t_k's truth-telling utility.
- ... sender private information becomes small

$$\mathbb{E}\left[\sigma(\mathbf{h})\right] \geq \lim_{N\to\infty} \frac{\sqrt{N}\left(U_k^N - U_l^N\right)}{t_k - t_l} \geq \mathbb{E}\left[\sigma(\mathbf{h})\right].$$

- Applying Bayes' rule

$$U_k^N = \mathbb{E}\left[(\mathbf{h}\cdot\mathbf{t}+b)\,\sigma(\mathbf{h})\left(1+\frac{h_k}{\sqrt{N}f_k}\right)\right].$$

- From which easy to see

$$\frac{\mathbb{E}\left[(\mathbf{h}\cdot\mathbf{t}+b)\,\sigma(\mathbf{h})\left(\frac{h_k}{f_k}-\frac{h_l}{f_l}\right)\right]}{\cdots} = \mathbb{E}\left[\sigma(\mathbf{h})\right].$$

Proof Sketch: "If" Direction

Suppose σ satisfies (ENV), & (MON) is slack; its outcomes are $(a, \mathbf{h})$.

- **Step 1**: Find a sequence of N-sender mechanisms σ^N s.t.

$$(a^N, \mathbf{h}^N) \xrightarrow{d} (a, \mathbf{h})$$

Not hard to find, but the σ^N may not be IC...

Proof Sketch: "If" Direction

Suppose σ satisfies (ENV), & (MON) is slack; its outcomes are $(a, \mathbf{h})$.

- **Step 1**: Find a sequence of N-sender mechanisms σ^N s.t.

$$(a^N, \mathbf{h}^N) \xrightarrow{d} (a, \mathbf{h})$$

  Not hard to find, but the σ^N may not be IC...

- **Step 2**: Adjust the σ^N to recover IC
  - Find set H of NEFs s.t. perturbations of σ^N on H can independently 'fix' local IC conditions for each t_k
  - Possible because types' private info about $\mathbf{h}$ is sufficiently different

Proof Sketch: "If" Direction

Suppose σ satisfies (ENV), & (MON) is slack; its outcomes are $(a, \mathbf{h})$.

- **Step 1**: Find a sequence of N-sender mechanisms σ^N s.t.

$$(a^N, \mathbf{h}^N) \xrightarrow{d} (a, \mathbf{h})$$

  Not hard to find, but the σ^N may not be IC...

- **Step 2**: Adjust the σ^N to recover IC
  - Find set H of NEFs s.t. perturbations of σ^N on H can independently 'fix' local IC conditions for each t_k
  - Possible because types' private info about $\mathbf{h}$ is sufficiently different
- **Step 3**: Show necessary adjustments become 'small' as $N \to \infty$
  - Follows from $(a^N, \mathbf{h}^N) \xrightarrow{d} (a, \mathbf{h})$ and σ being ICL.

Proof Sketch: "If" Direction

Suppose σ satisfies (ENV), & (MON) is slack; its outcomes are $(a, \mathbf{h})$.

- **Step 1**: Find a sequence of N-sender mechanisms σ^N s.t.

$$(a^N, \mathbf{h}^N) \xrightarrow{d} (a, \mathbf{h})$$

  Not hard to find, but the σ^N may not be IC...

- **Step 2**: Adjust the σ^N to recover IC
    - Find set H of NEFs s.t. perturbations of σ^N on H can independently 'fix' local IC conditions for each t_k
    - Possible because types' private info about $\mathbf{h}$ is sufficiently different
- **Step 3**: Show necessary adjustments become 'small' as $N \to \infty$
    - Follows from $(a^N, \mathbf{h}^N) \xrightarrow{d} (a, \mathbf{h})$ and σ being ICL.