Equilibrium Information Aggregation under Machine Learning

Andrew Ellis (LSE), Michele Piccione (LSE), and Shengxing Zhang (CMU)

June 5, 2026

Model

Given a distribution q over n^i-dimensional feedback $y^i = (v, p, s_1^i, \ldots, s_J^i)$, each agent i:

1. Learns q using the Chow-Liu Tree Algorithm
   - Determines a "best" approximation of q, q_R
2. Observes part of the feedback: s^i & p
3. Updates q_R in accordance with s^i & p
4. Submits demand that maximizes expected utility, i.e.

$$a_i^* \in \arg\max_{a_i} \int u^i(a_i, y^i) dq_R(y^i | s^i, p)$$

q adjusts so that market clears with probability 1 given strategies

Motivation

- Machine Learning: extracts useful information from high-dimensional (big) data
- Makes "better" (but not perfect and not Bayesian) predictions
- Economics: informative, endogenous variables (price)
- Different predictions leads to different relationships
- Equilibrium effects?

Contribution

- Model machine learning via Chow-Liu Trees (1968) (CLT)
- "Proposed in 1968, The Chow–Liu algorithm is widely used in machine learning and statistics as a tool for dimensionality reduction, classification, and as a foundation for algorithm design in more complex dependence structures"
  –Jiao, Han & Weissman (2016, IEEE)
- Unlike e.g. neural network/random forest, CLT has analytic solution

In asset market

- This talk: apply to asset market with dispersed information based on Hellwig (1980)
- Not all information extracted by all algorithms fully reflected in price
- Ex ante identical traders may use different CLTs
  - React asymmetrically to same information
- With low noise, most traders learn only from the (endogenous) price
- With purely public information, traders may update their beliefs in response to price changes
  - Even though prices are a garbling of public information

Related literature

- Information aggregation in markets: Radner (1979), Grossman (1979), Grossman Stiglitz (1980), Hellwig (1980), ...
- Machines playing games: Calavano et al (2020), Klein (2021), Dolgopov (2024), ...
- Stylized AI models: Szentes Ely (2023), Liang (2026), Fudenberg Liang (2026), Bergemann et al (2026)...
- Misspecified learning: Esponda Pouzo (2016), Spiegler (2016), Eliaz et al (2021), Fudenberg et al (2017),...

Chow-Liu Tree

- Consider an n-dimension random vector $Y = (Y_1, \ldots, Y_n)$
- A K-order dependence tree R is
  - (perfect) **directed acyclic graph** with nodes $\{1, \ldots, n\} = N$
  - each variable has **at most K parents**: $|R(j) = \{i : iRj\}| \leq K$.
- Tree R factors q into q_R where

$$q_R(y_1, \ldots, y_n) = q\left(y_{j^*}\right) \prod_{j \neq j^*} q\left(y_j | y_{R(j)}\right).$$

  when j^* is root variable
- E.g. when $R = 1 \to 2 \to 3 \to ...$,

$$q_R(y_1, \ldots, y_n) = q\left(y_1\right) q\left(y_2 | y_1\right) q\left(y_3 | y_2\right) ...$$

- Let $\mathcal{R}$ be the set of 1-order dependence trees
- NB: not causal, can treat as undirected

Chow-Liu Tree

- CLT is a **1**-order dependence tree that **minimizes KL-divergence from q,** i.e.

$$R^* \in \arg\min_{R \in \mathcal{R}} D_{KL}(q||q_R).$$

- When X is discrete

$$D_{KL}(p||q) = \sum_x p(x) \log\left(\frac{p(x)}{q(x)}\right)$$

- When X is continuous

$$D_{KL}(P||Q) = \int p(x) \log\left(\frac{p(x)}{q(x)}\right) dx$$

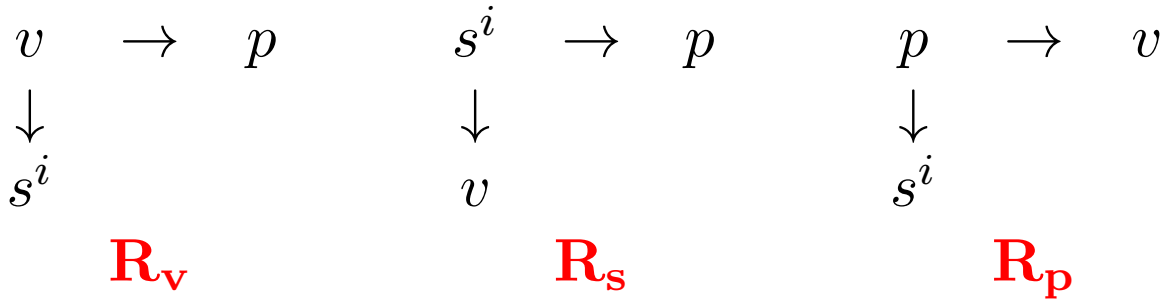
when p is density of P and q is density of Q.

Dimension reduction

- If binary, a probability distribution has $2^n - 1$ dimensions
- If multivariate-normal, there are $n + n + \frac{n(n-1)}{2}$ parameters
- CLT provides a way to reduce dimension or parameters
- q_R lower dimensional than q
- CLT will have $2n$ dimensions or $3n$ parameters

Accuracy/precision tradeoff

With three variables v, s^i, p, three possible dependence trees:



If cares about v, then

- R_v leaves out correlation between p and s^i
  so overreacts to s given p
  but **precise**: relatively certain about v
- R_s (R_p) leaves out correlation between v and p (s)
  so underreacts to p given s (s given p)
  but **accurate**: $E_{R_S}[v|s,p] = E[v|s]$

Bias/variance tradeoff

- Consider K-order tree: $|R(j)| \leq K$.
- Finding best K-order tree is NP-hard for $K \geq 2$ (Chickering, 1996)
- Smaller K: potential bias from omitted relationships
- Larger K: more variance in estimates from outliers
  - More data required: sufficient # of each $K+1$ tuple
- Tradeoff called "regularization": penalty for higher order
- We assume penalty sufficiently large that $K = 1$ is optimal

Chow-Liu Tree

Chow-Liu (1968) shows that CLT can be found using a Greedy Algorithm in finite time.

1. Compute pairwise mutual information for each pair of variables, $I_{ij} = \frac{h(p_{ij})}{h(p_i)h(p_j)}$
   - h is (differential) entropy, subscripts denote marginals
2. Take the two pairs $\{i, j\}$ and $\{k, l\}$ with largest I_{ij} and I_{kl}
3. Add third pair if it doesn't create a cycle
4. Add fourth pair if it doesn't create a cycle
5. ...

Done in at most $n(n+1)/2$ steps with n variables

Illustration

- With 3 normal variables every 1-tree takes form $R = i \to j \to k$ and

$$D(p||p_R) = \frac{1}{2}\ln\left(1-\rho(i,j)^2\right) + \frac{1}{2}\ln\left(1-\rho(j,k)^2\right) + \ln\sigma_i + \ln\sigma_j + \ln\sigma_k + \frac{3}{2}\ln(2\pi e) - h(p)$$

- Minimizing divergence requires only looking at correlations
- Divergence minimized by leaving out the pair with minimal correlation

Asset Market (Hellwig, 1980)

- Risky asset with common value v where $v \sim \mathcal{N}(0, \sigma_v^2)$
- Safe asset with certain unit return & price; unlimited supply
- Supply of risky asset $u \sim \mathcal{N}(0, \sigma_u^2)$
  - Total demand must equal u in equilibrium
  - Negative of "noise trader" demand
- Publicly observed price $p \in \mathbb{R}$

▸ Incomplete information games

Market

- Mass $I \in \{1, 2, 3, \dots\}$ of traders indexed by $i' \in (0, I]$
  Trader $i' \in (i-1, i]$ is type i
- All Type-i traders see same signal $s^i \in S^i = \mathbb{R}^J$
- $s^i_j = v + \epsilon^i_j$ with $\epsilon^i \sim \mathcal{N}(0, \Sigma^i)$ for $i = 1, \dots, I$, independently across types, $v \sim \mathcal{N}(0, \sigma^2_v)$.
- Trader i' buys $x^{i'}$ units of risky asset, and has utility index

$$u^{i'}(x^{i'}, v, p) = -\exp\left(-\frac{1}{r^i} x^{i'} (v - p)\right)$$

  - can take either long or short position in risky asset

Benchmark: Hellwig (1980)

- A **Bayesian equilibrium** consists of $\alpha \in \mathbb{R}^{IJ+1}$ so that $p = \alpha \cdot (\vec{s}, u)$ and:
  1. Type-i trader maximizes expected utility

$$x^i(s^i, p) = \arg\max_x E\left[-\exp\left(-\frac{1}{r^i}x\,(v-p)\right)|s^i, p\right],$$

  2. the asset market clears for every (s, u)

$$\sum_{i=1}^{I} x^i(s^i, p) = u$$

- Expectation uses Bayesian update about v given s^i and p
- Hellwig shows Bayesian equilibrium exists and p becomes as informative as $\vec{s}$ when $\sigma_u \to 0$

Asset pricing

- Type i's feedback / training data: $(s_1^i, \ldots, s_J^i, v, p)$
  - $p = \alpha \cdot (\vec{s}, u)$ results in normal distribution $\nu^i(\alpha)$ over i's feedback
  - input to CLT algorithm
  - "steady state" after many observations with equilibrium distribution
- Let $\mathcal{R}^i$ be possible trees for i

CLT EQUILIBRIUM

- A **CLT equilibrium** consists of an $\alpha \in \mathbb{R}^{IJ+1}$ and $\mu^1, \ldots, \mu^I$ so that:
  1. Type-i trader maximizes utility given tree-R and information s^i and p

$$x_R^i(s^i, p) = \arg\max_x E_R \left[-\exp\left(-\frac{1}{r^i} x\,(v-p) \right) | s^i, p \right],$$

  2. Each trader uses a CLT

$$\mu^i \left(\arg\min_{R \in \mathcal{R}^i} D_{KL}\left(\nu^i(\alpha) || \nu_R^i(\alpha) \right) \right) = 1,$$

  3. and the asset market clears for every (s, u)

$$\sum_{i=1}^{I} \sum_{R \in \mathcal{R}^i} \mu^i(R) x_R^i(s^i, p) = u$$

CLT equilibrium

Theorem

A CLT equilibrium exists

Demand function for tree R is

$$x_R^i(s,p) = r^i \frac{E_R[v|s^i,p] - p}{Var_R[v|s^i,p]}$$

$$= r^i \left(\sum_{j:\{s_j^i,v\}\in R} \sigma_{i,j}^{-2} s_j^i + \left[\mathbb{I}_R\left(\{p,v\}\right) \sigma_{p|v}^{-2} - Var_R(v|s^i,p)^{-1} \right] p \right)$$

where $\sigma_{i,j}^2 = \Sigma_{jj}^i$ and $\sigma_{p|v}^{-2} = Var(p|v)$

CLT equilibrium

Theorem

A CLT equilibrium exists

Demand function for tree R is

$$x_R^i(s,p) = r^i \frac{E_R[v|s^i,p] - p}{Var_R[v|s^i,p]}$$

$$= r^i \left(\sum_{j:\left\{s_j^i,v\right\}\in R} \sigma_{i,j}^{-2} s_j^i + \left[\mathbb{I}_R\left(\{p,v\}\right) \sigma_{p|v}^{-2} - Var_R(v|s^i,p)^{-1} \right] p \right)$$

where $\sigma_{i,j}^2 = \Sigma_{jj}^i$ and $\sigma_{p|v}^{-2} = Var(p|v)$

Equilibrium: trees $\rightarrow$ price

- fixing distribution of models μ, unique price that clears market
- relative weights on different signals given by

$$\frac{\alpha_{i,j}}{\alpha_{k,l}} = \frac{\mu^i\left(\left\{R : \{v, s^i_j\} \in R\right\}\right)\sigma^{-2}_{i,j} r^i}{\mu^k\left(\{R : \{v, s^k_l\} \in R\}\right)\sigma^{-2}_{k,l} r^k}$$

  - more often used $\implies$ higher weight
  - more precise $\implies$ higher weight
  - more risk loving $\implies$ higher weight
  - Unlike Bayesian, no effect of correlation between different signals and price given μ
- relative weight on u in p inversely proportional to precision of signals weighted by CLT

Equilibrium: price → trees

- together this price determines the correlation between v and p as well as between p and each signal s^i_j
- μ adjusts so that best approximations used
  - knock on effect on pricing and correlations
- price too responsive to signal $\implies$ everyone uses only price
- price too unreactive to every signal $\implies$ market does not clear
- Solve via nested fixed point problem

Information in equilibrium

- If everyone Bayesian, this is reformulation of Hellwig (1980) where $s^i = 1' \Sigma^{i-1} s^i$
  - also Grossman-Stiglitz (1980) without information acquisition
- In Bayesian Equilibrium:
  - p reflects all private signals
  - p is **approximately informationally efficient**:
    As noise σ_u goes to zero,

$$E[v|p] \to E[v|\vec{s}]$$

    beliefs approach what they would be if all private signals were public
    Equivalently, as $\sigma_u \to 0$,

$$\rho(p, v) \to \rho\left(E[v|\vec{s}], v\right)$$

(No) Informational Efficiency

Theorem

Approximate informational efficiency is impossible.
Fixing $I > 1$, $\sigma_v > 0$, and Σ^i for all i, let $\rho^* = \rho(E[v|\vec{s}], v)$. There exists $\delta > 0$ so that for any $\sigma_u > 0$, $\rho(p, v) < \rho^* - \delta$ in any CLT equilibrium.

- Hellwig (1980): $\rho(p, v) \to \rho^*$ as $\sigma_u \to 0$
- Even approximate information efficiency possible only when $\sigma_u = 0$

(No) Informational Efficiency

- Not all information extracted by algorithm reflected in market price
- Non-vanishing private value of machine learning
- Since training is costly, suggests persistent use of it in market
- Relates to Grossman / Grossman-Stiglitz Paradoxes

Proof

- $\rho(p, v) = \rho^*$ if and only if p is sufficient statistic for $(s^1, \ldots, s^I)$
- Suppose that p is (approximately) a sufficient statistic for $(s^1, \ldots, s^I)$
- Data-Processing Inequality: $\rho(p, s^i_j) > \rho(v, s^i_j)$ and $\rho(v, p) > \rho(v, s^i_j)$
- Therefore, unique CLT is

$$\begin{array}{ccccc} v & \leftarrow & p & \rightarrow & s^i_1 \\ & & \downarrow & \searrow & \\ & & s^i_J & & \cdots \end{array}$$

- But then for markets to clear, $p = \alpha_u u$, implying $\rho(p, v) = 0$

CLT equilibrium

We focus on two cases:

1. $I > 1$ and $J = 1$: dispersed private information but only 1 dimensional
   - How does CLT aggregate information?
2. $I = 1$ and $J > 1$: public information but multi-dimensional
   - How does CLT tradeoff between different aspects of signals?

General case has features of both

Dispersed information

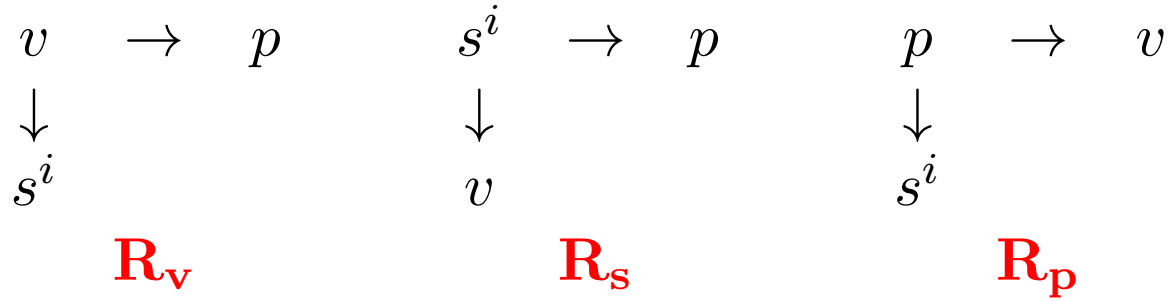
- Suppose that $s^i = v + \epsilon_i$ for $\epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2)$, iid ($j$ suppressed)
- Common risk aversion, $r^i = r$ for all i
- Then, equilibrium price characterized by $\alpha_s, \alpha_u \in \mathbb{R}$ so that

$$\begin{aligned} p &= \alpha_s I^{-1} \sum_i s^i + \alpha_u u \\ &= \alpha_s v + \alpha_s I^{-1} \sum_i \epsilon_i + \alpha_u u \end{aligned}$$

- Hellwig (1980)'s setup

CLT EQUILIBRIUM

- Three possible dependence trees:



- R_v leaves out correlation between p and s^i.
  Overreacts to to private signal given price
- R_s leaves out correlation between v and p.
  Underreacts to price given private signal
- R_p leaves out correlation between v and s^i.
  Underreacts to private signal given price

CLT EQUILIBRIUM

Notice v, s^i, p joint normal with correlations

$$\rho(v, s^i) = \frac{\sigma_v}{\sqrt{\sigma_v^2 + \sigma_\epsilon^2}}$$

$$\rho(v, p) = \frac{\sigma_v}{\sqrt{\sigma_v^2 + I^{-1}\sigma_\epsilon^2 + \frac{\alpha_u^2}{\alpha_s^2}\sigma_u^2}}$$

$$\rho(s^i, p) = \frac{\sigma_v^2 + I^{-1}\sigma_\epsilon^2}{\sqrt{\sigma_v^2 + \sigma_\epsilon^2}\sqrt{\sigma_v^2 + I^{-1}\sigma_\epsilon^2 + \frac{\alpha_u^2}{\alpha_s^2}\sigma_u^2}}$$

- $\rho(v, p)$ and $\rho(s^i, p)$ depend on α only through noise-signal ratio $h \equiv -\frac{\alpha_u}{\alpha_s}$

Market clearing

- Using updating rules for normal distribution and market clearing,

$$h = \frac{\sigma_\epsilon^2}{r^i I \left(1 - \mu(R_p)\right)}$$

- **Correlations depend only on** $\mu(R_p)$
- **Equilibrium effect**: $\rho(p, v) \to \mu(R_p) \to p \to \rho(p, v) \to \ldots$
- ▸ Algebra

Theorem

If $\left(I^2 - 2I\right)\sigma_v^2 > \sigma_\epsilon^2$, *then in equilibrium,*

$$\mu^i(R_v) = \min\left\{1, \sqrt{\frac{\sigma_\varepsilon^2\sigma_u^2\sigma_v^2}{r^2\left(\sigma_\varepsilon^2 + I\sigma_v^2\right)}}\right\} = 1 - \mu^i(R_p)$$

for all i, *and*

$$\rho(p, v) \leq \frac{I\sigma_v^2}{I\sigma_v^2 + \sigma_\epsilon^2}.$$

If $\left(I^2 - 2I\right)\sigma_v^2 < \sigma_\epsilon^2$, *then in equilibrium,*

$$\mu^i(R_s) = \min\left\{1, \sqrt{\frac{\sigma_u^2\sigma_\varepsilon^2}{r^2 I\left(I-1\right)}}\right\} = 1 - \mu^i(R_p)$$

for all i, *and*

$$\rho(p, v) \leq \frac{\sigma_v}{\sqrt{\sigma_v^2 + \sigma_\epsilon^2}}.$$

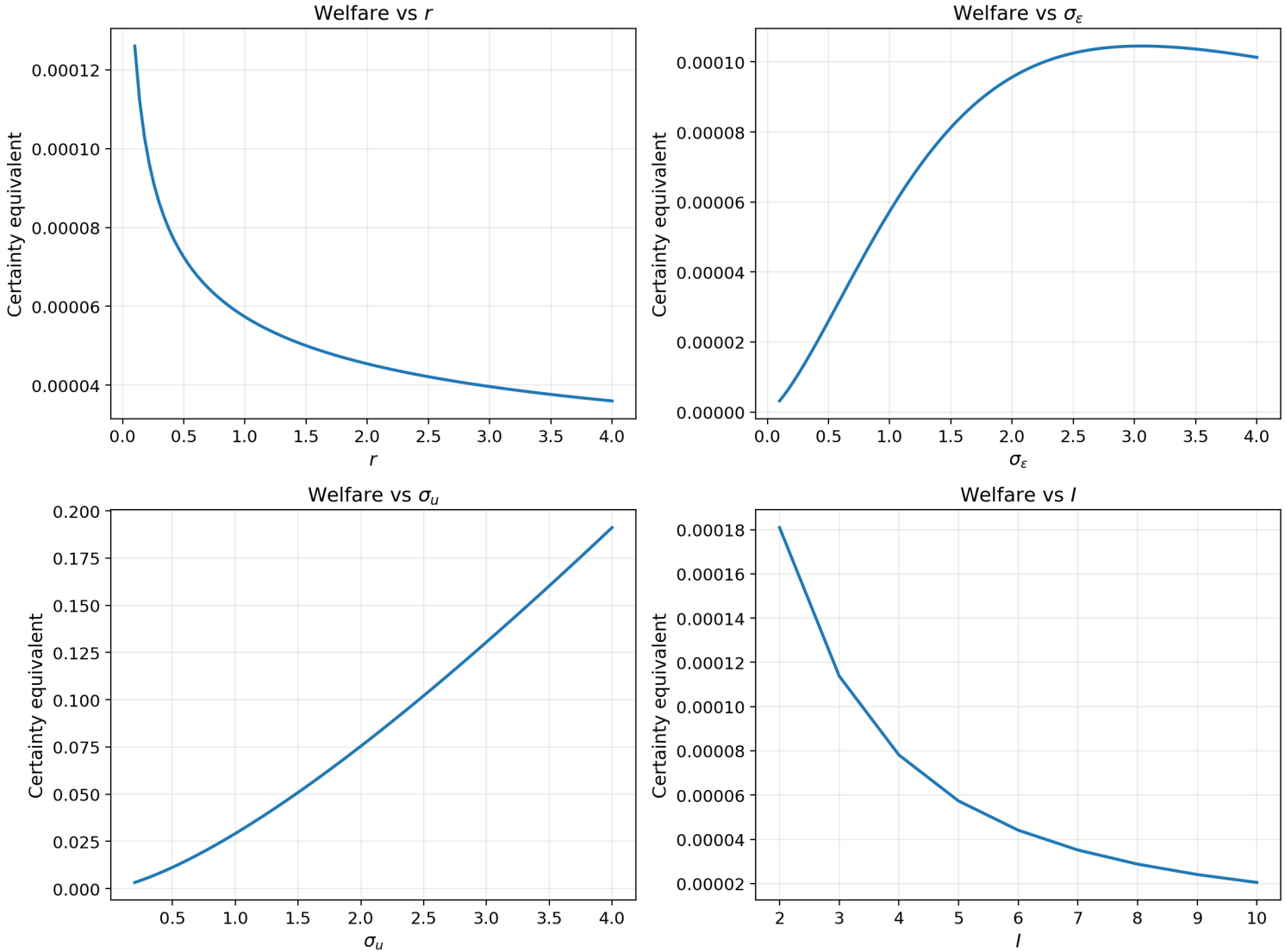
Intuition

- Algebra: $\left(I^2 - 2I\right)\sigma_v^2 \leq \sigma_\epsilon^2 \iff \rho(p,v) \leq \rho(p,s_i) \iff$ R_s better approximation than R_v.
- For concreteness, assume holds without equality, so $\mu(R_v) = 0$
- If $\mu(R_p) = 0$ and $\sigma_u \approx 0$, $\rho(v,p) > \rho(v,s)$, then R_p better than R_s
- If $\mu(R_s) = 0$, then no one uses private information and market cannot clear!
- When σ_u small, $\mu(R_p)$ adjusts so $\rho(v,p) = \rho(v,s)$ and R_p as good as R_s

Intuition

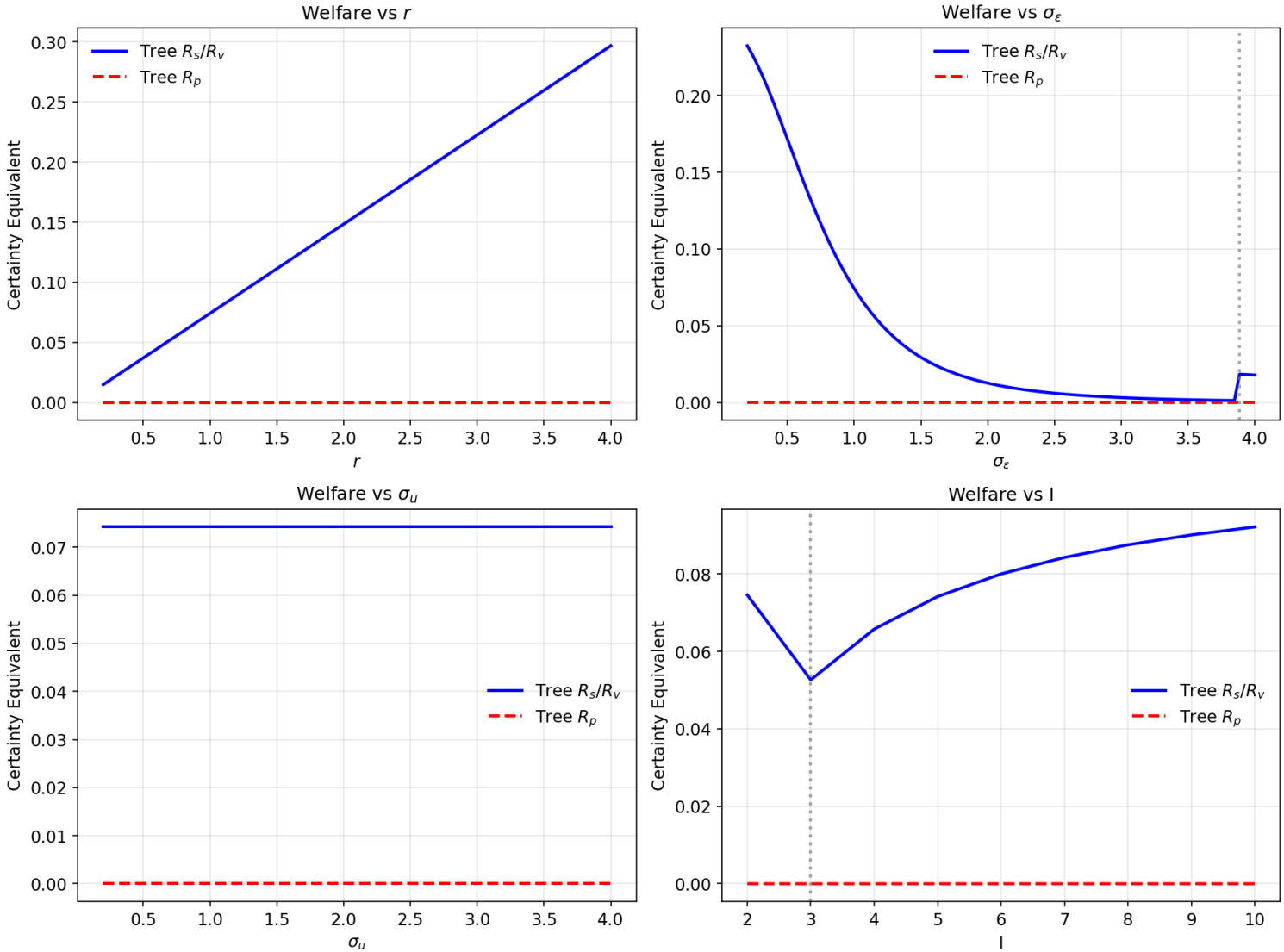
- Algebra: $\left(I^2 - 2I\right)\sigma_v^2 \leq \sigma_\epsilon^2 \iff \rho(p,v) \leq \rho(p,s_i) \iff$ R_s better approximation than R_v.
- For concreteness, assume holds without equality, so $\mu(R_v) = 0$
- If $\mu(R_p) = 0$ and $\sigma_u \approx 0$, $\rho(v,p) > \rho(v,s)$, then R_p better than R_s
- If $\mu(R_s) = 0$, then no one uses private information and market cannot clear!
- When σ_u small, $\mu(R_p)$ adjusts so $\rho(v,p) = \rho(v,s)$ and R_p as good as R_s

Hellwig (1980)



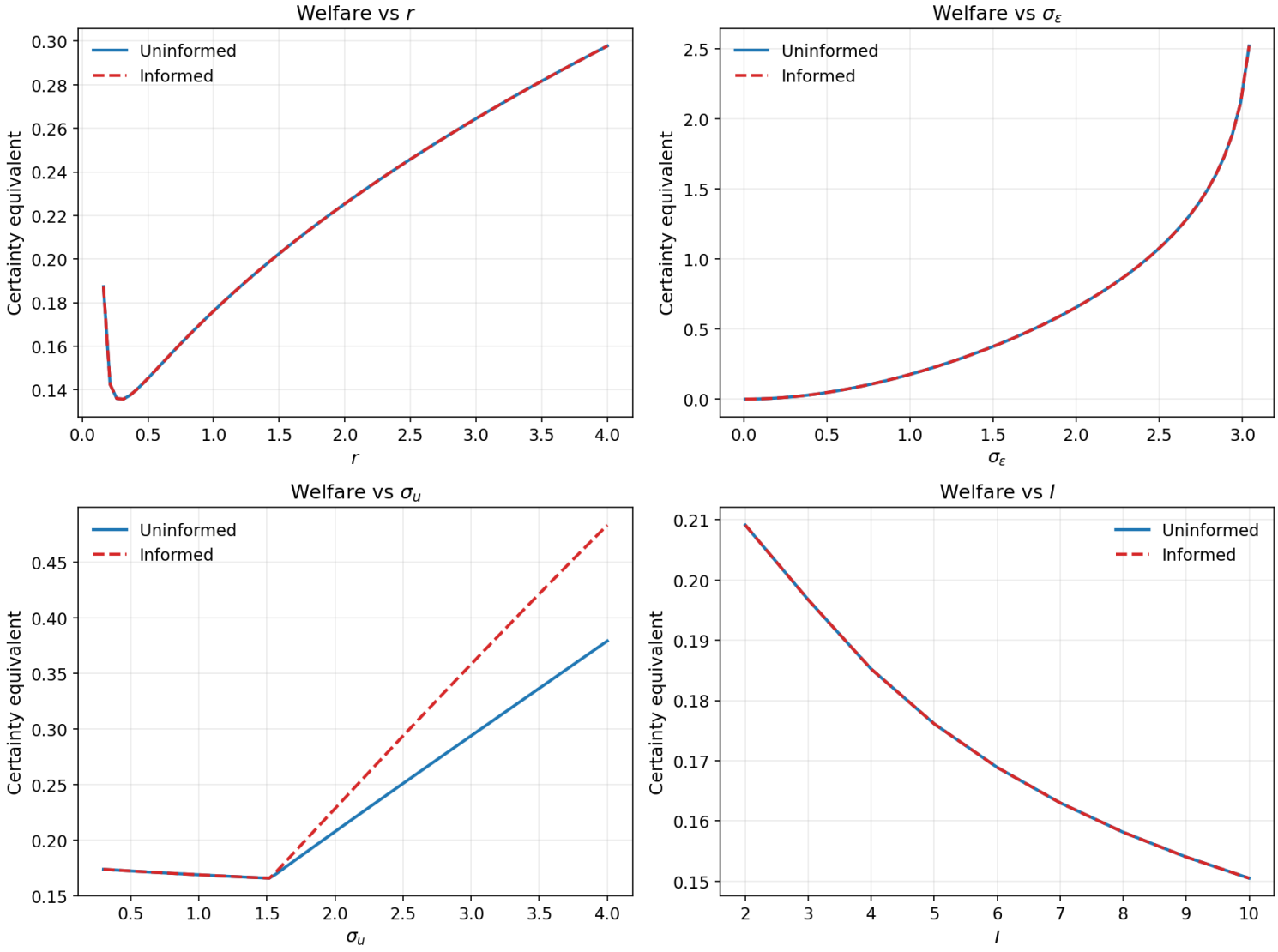
Baseline: $I = 5,\ r = 1,\ \sigma_v = 1,\ \sigma_\epsilon = 1,\ \sigma_u = 0.5$

CLT Equilibrium



Baseline: $I = 5$, $r = 1$, $\sigma_v = 1$, $\sigma_\epsilon = 1$, $\sigma_u = 0.5$

Grossman-Stiglitz (1980)



Baseline: $I = 5,\ r = 1,\ \sigma_v = 1,\ \sigma_\epsilon = 1,\ \sigma_u = 0.5$

Big data: Model

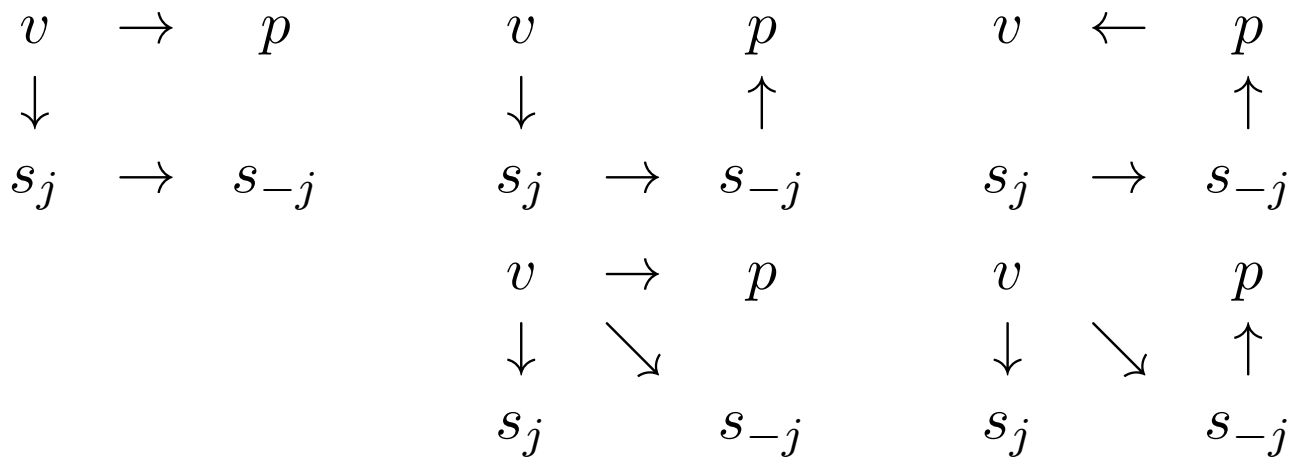
- Suppose instead that $I = 1$ (public information), $J = 2$ (two-dimensional information), $s^i = s = (s_1, s_2)' = (v + \epsilon_1, v + \epsilon_2)'$, and

$$\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \sigma_1^2 & \bar{\rho}\sigma_1\sigma_2 \\ \bar{\rho}\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \Sigma\right)$$

- WLOG, $\sigma_1 \leq \sigma_2$

Big data: trees

- More trees than before:



- These are "representative" since only edges with v matter
- Price noisy function of s_1, s_2 so rational agent does not update from it

Big data: equilibrium

Theorem

In the equilibrium of this model, some trader forms expectation using both s_1 and s_2, i.e., $\mu\left(\left\{R : \left\{\left\{v, s_1\right\}, \left\{v, s_2\right\}\right\} \subset R\right\}\right) > 0$, only if $\bar{\rho} \leq \rho^*\left(\sigma_v, \sigma_\varepsilon\right)$.

- Specialization in information
- Useful information disregarded

Big data

If $\bar{\rho}$ is large enough so that $\rho(s_1, s_2) > \rho(s_2, v)$,
Then **in equilibrium:**

$$\mu\left(\left\{\begin{array}{ccc} v & \to & s_2 \\ \downarrow & & \\ s_1 & \to & p \end{array}, \begin{array}{ccc} v & \to & s_1 \\ \downarrow & \searrow & \\ s_2 & & p \end{array}, \begin{array}{ccc} v & \to & s_2 \\ \downarrow & & \downarrow \\ s_1 & & p \end{array}\right\}\right) = 0.$$

Proof: replacing the edge (v, s_2) with the edge (s_1, s_2) decreases divergence

In equilibrium, only s_1 is used. ▸ Conclusion

If $\bar{\rho}$ is small enough that $(s_1, s_2) < \rho(v, s_2)$, and $\sigma_1^2 \approx \sigma_2^2$
Then **in equilibrium** both signals are used:

$$\mu\left(\{R : \{v, s_1\} \in R\}\right), \mu\left(\{R : \{v, s_2\} \in R\}\right) > 0$$

Proof: If $\mu\left(\{R : \{v, s_2\} \in R\}\right) = 0$, then $\alpha_j = 0$. This implies

$$\rho(s_2, p) = \frac{\sigma(s_2)}{\sigma(p)}\rho(s_1, s_2) < \rho(s_1, s_2) < \rho(v, s_2)$$

Replacing edge $\{p, s_j\}$ with edge $\{v, s_j\}$ decreases divergence.

Big data

If $\bar{\rho}$ **and** σ_u^2 are small enough and $\sigma_1^2 \approx \sigma_2^2$,
Then **in equilibrium** price is used:

$$\mu\left(\{R : \{v, p\} \in R\}\right) > 0$$

Proof: If not, then as $\sigma_u^2 \to 0$

$$\rho(v, p) \to \rho(E[v|s_1, s_2], p) > \rho(p, s_1) > \rho(s_1, v)$$

Replacing edge $\{p, s_j\}$ with edge $\{v, p\}$ decreases divergence.

Conclusions?

- ML good for predictions, not so good at causality
- Using it extensively may have important equilibrium effects on informativeness of prices even if privately beneficial
- Tragedy of the commons with dispersed info: if everyone uses private info, then price informative. But then nobody uses private info.
- Other settings? Other models of ML?

Thanks!

Algebra

Demand using updating rules for normal distribution:

$$\begin{aligned} x^i_{R_v}(\vec{s},p) &= r^i[\sigma_\epsilon^{-2}s^i + \sigma_{\tilde{p}}^{-2}\alpha_s^{-1}p - \tau_v p] \\ x^i_{R_s}(\vec{s},p) &= r^i[\sigma_\epsilon^{-2}s^i + \qquad\qquad - \tau_s p] \\ x^i_{R_p}(\vec{s},p) &= r^i[\qquad\quad + \sigma_{\tilde{p}}^{-2}\alpha_s^{-1}p - \tau_p p \end{aligned}$$

where

$$\begin{aligned} \sigma^2_{\tilde{p}} &= Var(v|p) = I^{-1}\sigma^2_\epsilon + h^2\sigma^2_u \\ \tau_s &= \sigma_\epsilon^{-2} + \sigma_v^{-2} \\ \tau_p &= \sigma_v^{-2} + \sigma_{\tilde{p}}^{-2} \\ \tau_v &= \sigma_v^{-2} + \sigma_{\tilde{p}}^{-2} + \sigma_\epsilon^{-2} \end{aligned}$$

- Suppose that $\mu^i(R) = \mu(R)$ for all i and let

$$\begin{aligned} \mu &= (\mu(R_v), \mu(R_p), \mu(R_s)) \\ \tau &= (\tau^v, \tau^p, \tau^s) \end{aligned}$$

$$\lambda_s\tau = (\lambda_s^v\tau^v, \lambda_s^p\tau^p, \lambda_s^s\tau^s) = \left(\sigma_\epsilon^{-2}, 0, \sigma_\epsilon^{-2}\right)$$

Incomplete information game

There are I players.

1. Fundamental uncertainty Ω resolves
2. Player i observes a n^i dimensional signal $s^i \in S^i = \prod_{j=1}^{n^i} S^i_j$ where each $|S^i_j| \in [2, \infty)$
   - Distributed according to $p \in \Delta\left(\Omega \times \prod_i S^i\right)$
3. Player i takes an action $a^i \in A^i$ where each $|A^i| \geq 2$
   - Player i maximizes $u^i : A^i \times Y^i \to \mathbb{R}$ given s^i
4. Player i observes feedback $y^i \in Y^i \subset \mathbb{R}^{m^i}$ with $m^i \geq n^i$; $Y = Y^1 \times \cdots \times Y^I$
   - First n^i dimensions of y^i correspond to s^i

CLT equilibrium

- Distribution over the feedback to i given by $f^i : A \times \Omega \times S \to \Delta Y^i$
  - Assume that $f^i(a, \omega, s)$ does not depend on i's action a_i
  - Assume that f_i has full support (avoid zero probability updating)
- Behavioral strategy: $\sigma^i : S^i \to \Delta A^i$
- Feedback to i given strategy profile σ distributed according to $\nu^{i,\sigma}$ where

$$\nu^{i,\sigma}(y) = \int_{A\times\Omega\times S} \mu(\omega, s) \prod_{i=1}^{I} \sigma^i(s^i)(a^i) f^i(a, \omega, s)(y) d(a, \omega, s)$$

  indicates the probability of seeing feedback $y \in Y^i$
- Let $\mathcal{R}^i$ be the 1-order dependence trees with over $\{1, \ldots, n^i\}$

Definition

A CLT equilibrium is a strategy profile σ_* and I-distributions μ^i over $\mathcal{R}^i$ so that

$$\mu^i(R) > 0 \implies R \in \arg\max_{R \in \mathcal{R}_i} D_{KL}(\nu^{i,\sigma_*} || \nu_R^{i,\sigma_*})$$

and

$$\sigma_*^i(a, s^i) = \sum_{R \in \mathcal{R}_i} \mu^i(R) \sigma_R^i(a, s^i)$$

for some σ_R^i such that

$$\sigma_R^i(a^i, s^i) > 0 \implies a^i \in \arg\max_{a \in A^i} \sum_{y^i \in Y^i} u^i\left(a, y^i\right) \nu_R^{i,\sigma_*}\left(y^i | s^i\right).$$

- σ_R^i is a best response for beliefs ν_R^{i,σ_*}
- σ_*^i combines CLT (μ^i) and best response given tree (σ_R^i)

CLT equilibrium

Theorem

A CLT equilibrium exists for any incomplete information game when all above sets finite

- Follows from fixed point arguments
- σ_* generates dataset that traders learn from
- Optimal model and strategies based on that data

◂ Back