Information Span in Credit Market Competition

Zhiguo He, Jing Huang, Cecilia Parlatore

Discussion by Savitar Sundaresan

Imperial College

June 8, 2024

Would you give this guy a loan?

Would you give this guy a loan?

- Suppose that $\theta \in \{0, 1\}$ is whether or not you get repaid.
- Suppose further that $\theta = \theta_H \times \theta_S$ where θ_H and $\theta_S \in \{0, 1\}$.
- θ_H could be: Will Savitar have the money to repay the loan?
- θ_S could be: Is Savitar a good guy?
- 4 possibilities.

Would you give this guy a loan?

Would you give this guy a loan?

- θ_H and θ_S unobservable.
- Can collect hard data to help figure out θ_H:
  - How much money does Savitar have now?
  - What is his salary?
  - etc...
- Can collect soft data to help figure out θ_S:
  - Does Savitar have friends?
  - Do his students/colleagues like him?
  - etc...
- Can get signals from hard/soft data on θ_H and θ_S.

What happens if there's overlap?

- Suppose that $\theta = \theta_H \times \theta_O \times \theta_S$ (θ_O also Bernoulli).
- The hard data/signal you collect tells you about $\theta_H \times \theta_O$.
- The soft data/signal you collect tells you about $\theta_O \times \theta_S$.
- θ_O is a variable that impacts both hard and soft signals.
- θ_O could be: Will Savitar be alive to repay the loan?

What happens if there's overlap?

- Medical records used to be unavailable, but now available.
- What's the difference between:

$$\begin{aligned} \theta &= \theta_H \times \theta_O \times \theta_S \\ \theta &= (\theta_H \times \theta_O) \times (\theta_S \times \theta_O) \end{aligned}$$

- Is the θ_O hard or soft?

SETUP

Compare two settings:

$$\theta = \overbrace{\theta_H}^{\text{hard}} \times \underbrace{\theta_O \times \theta_S}_{\text{soft}}$$

and

$$\theta = \overbrace{\theta_H \times \underbrace{\theta_O}_{\text{soft}}}^{\text{hard}} \times \theta_S$$

Three Agents with Different Info (first setting)

Firm:

$$\hat{\theta}_{firm} = \theta$$

Specialized Lender:

$$\hat{\theta}_{spec} = f(\hat{\theta}_H, \hat{\theta}_{soft})$$

General Lender:

$$\hat{\theta}_{gen} = f(\hat{\theta}_H, \mathbb{E}\left[\theta_O \times \theta_S\right])$$

Three Agents with Different Info (second setting)

Firm:

$$\hat{\theta}_{firm} = \theta \quad \text{(unchanged)}$$

Specialized Lender:

$$\hat{\theta}_{spec} = f(\hat{\theta}_H, \hat{\theta}_{soft}) \quad \Rightarrow \quad f(\hat{\theta}_{hard}, \hat{\theta}_{soft})$$

General Lender:

$$\hat{\theta}_{gen} = f(\hat{\theta}_H, \mathbb{E}\left[\theta_O \times \theta_S\right]) \quad \Rightarrow \quad f(\hat{\theta}_{hard}, \mathbb{E}\left[\theta_S\right])$$

Key Mechanisms

As a θ_O moves from being 'soft' to 'hard':

1. More hard data
   - $\Rightarrow \hat{\theta}_{gen} \rightarrow \hat{\theta}_{firm}$ less asymmetry between lender and borrower
   - $\Rightarrow \hat{\theta}_{gen} \rightarrow \hat{\theta}_{spec}$ less asymmetry between lenders
   - $\Rightarrow$ improved learning from competition.
   - $\Rightarrow$ reduced loan rates, improved profits for general lender.
2. More 'span'
   - $\Rightarrow \rho(\hat{\theta}_{hard}, \hat{\theta}_{soft}) > 0$
   - $\Rightarrow$ Cross-learning from signals by specialized lender.

Setup: Reminder

Compare two settings:

$$\theta = \overbrace{\theta_H}^{\text{hard}} \times \underbrace{\theta_O \times \theta_S}_{\text{soft}}$$

and

$$\theta = \overbrace{\theta_H \times \underbrace{\theta_O}_{\text{soft}}}^{\text{hard}} \times \theta_S$$

For some concreteness, suppose $q_H = q_O = q_S = 0.8$ (span goes from 0 to 0.2).

Philosophical Issue 1: What if most rich guys are jerks?

First setting:

$$\begin{aligned} Cov(\theta_h, \theta_s) &= Cov(\theta_H, \theta_O \times \theta_S) \\ &= 0 \end{aligned}$$

Second setting:

$$\begin{aligned} Cov(\theta_h, \theta_s) &= Cov(\theta_H \times \theta_O, \theta_O \times \theta_S) \\ &= Var(\theta_O) \\ &= 0.16 > 0 \end{aligned}$$

Hard and soft data must be positively correlated in model. Feature/bug of O-Ring setup.

Philosophical Issue 2: How does hard data processing work?

The process of 'hardening' means new hard data is available (digitized medical records)

- *Must* get aggregated signal about $\theta_H \times \theta_O$.
- Independent factors, but all data bundled. (Modulo section 5.2)
- Specialization does not improve α on new hard data.
- Soft data still collected as before.
- What is α?

Philosophical Issue 2: How does hard data processing work?

What if we allowed for characteristic-specific signals?

- In first setting:
  - h_H for θ_H
  - s_O and s_S for θ_O and θ_S, respectively.
- in second setting
  - h_H and h_O for θ_H and θ_O, respectively.
  - s_O and s_S for θ_O and θ_S, respectively.
- Hardening θ_O means the 'more' results hold. The 'span' results do not hold.

Philosophical Issue 2: How does hard data processing work?

How can we preserve span results? Perhaps reasonable to get aggregated signal on soft data:

- In first setting:
  - h_H for θ_H
  - s for $\theta_O \times \theta_S$.
- in second setting
  - h_H and h_O for θ_H and θ_O, respectively.
  - s for $\theta_O \times \theta_S$.
- Hardening θ_O means the 'more' and 'span' results hold. But now θ_H is pointless.

Philosophical Issue 3: Precise definitions

What is data? What is a signal? What is a fundamental state? What is information?

- θs are the fundamental states.
- Directly input into 'information technology'!
- IT also takes data (not modelled) generated by fundamental and creates signal.
- Hardening described as IT progress? α held fixed.
- 'hard objective data points, i.e. hard signals'.
- How can it be objective and different across lenders?

Non-Philosophical Issue

- Before θ_O is 'hardened', $q_h = 0.8$, $q_s = 0.64$, $\eta = 0$.
- After θ_O is 'hardened', $q_h = 0.64$, $q_s = 0.64$, $\eta = 0.2$
- Probability of fail signal before is $0.2\alpha + (1 - \alpha)0.8$.
- Probability of fail signal after is $0.36\alpha + (1 - \alpha)0.64$.
- Strictly higher! (if $\alpha = 1$, probability of *LL* goes from 4% to 12.96%).
- Countervailing force: conditional on high signal, more cross-learning about soft!

Non-Philosophical Issue

- Hardening of data reduces credit? Possibly counterfactual.
- Also contradicts paper's Figure 5.
- In this example η goes from 0 to 0.2, with q_H, q_O, q_S fixed.
- So $\eta \uparrow$ implies $q_h \downarrow$, $q_s = q_S \times q_O$ unchanged.
- But all figures in paper hold q_h and q_s fixed while varying η. So $q_h^h \uparrow$ as $\eta \uparrow$?
- To keep q_h fixed at 0.8, requires simultaneous increase of q_H from 0.8 to 1.

Miscellaneous - Free Disposal

- O-Ring theory of credit seems difficult to justify. Not necessarily super important in terms of being pivotal for your results (apart from the correlation point earlier).
- Hard to understand purpose (mathematical or intuitional) of having differential signal structures for hard and soft information. Hard data can surely get you a continuous signal, just as a soft report can give you a go/no go recommendation. Is the soft continuity to get smooth changes in results?
- The fact that soft data collection is unaffected by span seems to imply that results will be sensitive to (soft) information acquisition costs. Some discussion of this would be helpful.
- A clear example of what a hardened fundamental θ could be would be extremely useful to fixing ideas, and would help understand what definition you ascribe (data/signal/etc) to what real-world features.

Summary

- Really interesting framework, with some novel mechanisms at play.
- Thinking about information span as opposed to precision or quantity is important.
- Concerns:
  - Side-effects of O-ring theory.
  - How *precisely*, α is meant to be interpreted vis-a-vis data processing.
  - Confusing language around overlapping terms (fundamental, information, data, signal).
  - Confusion around figures with varying η while fixing q_h.