AI-Powered Trading, Algorithmic Collusion, and Price Efficiency

Winston W. Dou[◇] Itay Goldstein[◇] Yan Ji[†]

[◇]University of Pennsylvania and NBER

[†]Hong Kong University of Science and Technology

Nordic Fintech Symposium 2023, Sep 26

Algorithmic trading

Algorithmic trading has been prevalent (e.g., HFT)

- Algorithm-based hedge funds are responsible for about 27% of all equity trading of any investor, according to the Tabb Group (WSJ, 2017)
- Algorithm-based hedge funds control more than 30% of all hedge-fund assets, according to HFR Inc. (WSJ, 2017)

AI-powered trading

- Algorithmic trading + reinforcement-learning (RL) algorithms
  - Has the potential to reshape capital markets fundamentally
  - Presents new regulatory challenges

Algorithmic trading

Algorithmic trading has been prevalent (e.g., HFT)

- Algorithm-based hedge funds are responsible for about 27% of all equity trading of any investor, according to the Tabb Group (WSJ, 2017)
- Algorithm-based hedge funds control more than 30% of all hedge-fund assets, according to HFR Inc. (WSJ, 2017)

AI-powered trading

- Algorithmic trading + reinforcement-learning (RL) algorithms
  - Has the potential to reshape capital markets fundamentally
  - Presents new regulatory challenges

Algorithmic trading

Algorithmic trading has been prevalent (e.g., HFT)

- Algorithm-based hedge funds are responsible for about 27% of all equity trading of any investor, according to the Tabb Group (WSJ, 2017)

- Algorithm-based hedge funds control more than 30% of all hedge-fund assets, according to HFR Inc. (WSJ, 2017)

AI-powered trading

- Algorithmic trading + reinforcement-learning (RL) algorithms

  - Has the potential to reshape capital markets fundamentally

  - Presents new regulatory challenges

Defining features of RL algorithms

- Self-learning
  - Not just statistical machine learning, supervised or unsupervised
  - Learning through autonomous trial-and-error experimentation
- Model-free learning
  - No prior knowledge of the environment's parameters or specifications
  - Learning from the outcomes of their own actions
- Behavior-learning
  - Not to learn the environment itself
  - But rather to learn the optimal actions that maximize rewards

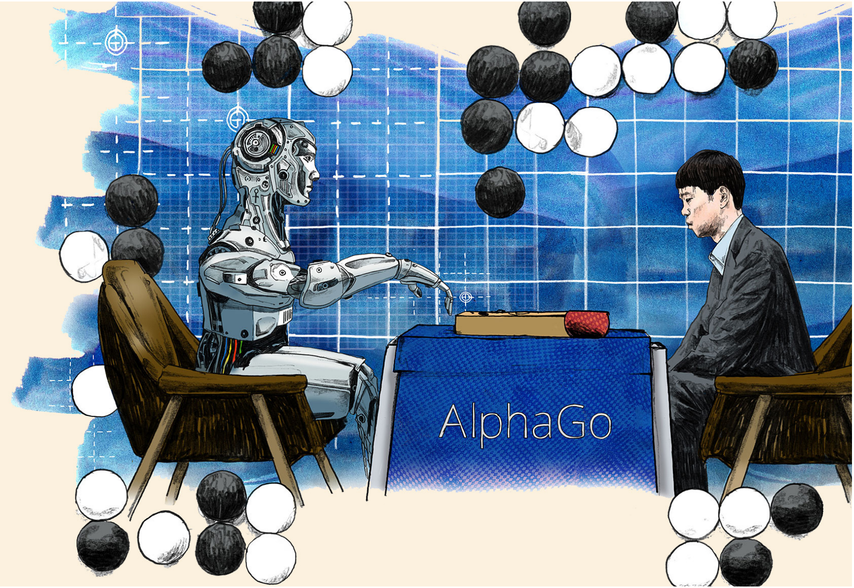
Defining features of RL algorithms

- Self-learning
  - Not just statistical machine learning, supervised or unsupervised
  - Learning through autonomous trial-and-error experimentation
- Model-free learning
  - No prior knowledge of the environment's parameters or specifications
  - Learning from the outcomes of their own actions
- Behavior-learning
  - Not to learn the environment itself
  - But rather to learn the optimal actions that maximize rewards

Defining features of RL algorithms

- Self-learning
  - Not just statistical machine learning, supervised or unsupervised
  - Learning through autonomous trial-and-error experimentation
- Model-free learning
  - No prior knowledge of the environment's parameters or specifications
  - Learning from the outcomes of their own actions
- Behavior-learning
  - Not to learn the environment itself
  - But rather to learn the optimal actions that maximize rewards

RL algorithm is the backbone of “AlphaGo”



AI-powered trading has been on the rise

BarclayHedge poll (2018):

- 56% of hedge fund respondents said they were using artificial intelligence or machine learning in their investment process

Yahoo Finance/Ipsos survey (2023):

- Younger adults are twice as likely to use an AI-powered financial advisor compared to older adults

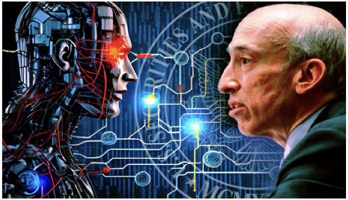
JPMorgan Chase survey (2023):

- More than 50% of respondents, which are 835 institutional and professional traders, said AI technologies would have the most influence on trading the next three years

AI-powered trading as a regulatory priority

SEC Chair, Gary Gensler, has warned of

Financial market instability if big tech-based trading companies monopolize AI development and applications within the financial sector



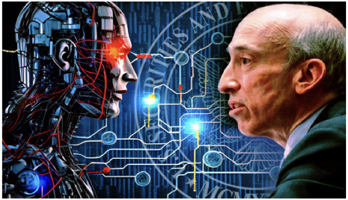
Regulatory challenges:

- Promote competitive and efficient markets amid rapid adoption of AI tech.
- Address the biases in RL algorithms due to factors like artificial stupidity

AI-powered trading as a regulatory priority

SEC Chair, Gary Gensler, has warned of

Financial market instability if big tech-based trading companies monopolize AI development and applications within the financial sector



Regulatory challenges:

- Promote competitive and efficient markets amid rapid adoption of AI tech.
- Address the biases in RL algorithms due to factors like artificial stupidity

"Psychology" of AI traders

AI algorithms do not merely imitate human behavior or intelligence
(e.g., Sargent, 2023)

- Existing theories are built on human behavior
- They are unsuitable to explain the dynamics of AI-powered capital markets

Understanding the implications of AI-powered trading necessitates

- Insights into the AI behavior, akin to the **"psychology" of machines** (Goldstein_Spatt_Ye, 2021)
- But not the **preferences or psychology of human beings**

“Psychology” of AI traders

AI algorithms do not merely imitate human behavior or intelligence
(e.g., Sargent, 2023)

- Existing theories are built on human behavior
- They are unsuitable to explain the dynamics of AI-powered capital markets

Understanding the implications of AI-powered trading necessitates

- Insights into the AI behavior, akin to the **“psychology” of machines** (Goldstein_Spatt_Ye, 2021)
- But not the **preferences or psychology of human beings**

Research questions

How would AI speculators trade under asymmetric information?

- Can they form collusion in the absence of communication?
- If so, what is the mechanism behind the "algo collusion?"

This paper:

- Informed AI speculators can collude and achieve supra-competitive profits
- Two distinct types of collusive behaviors

Mkt. efficiency \ Info. asymmetry	low	high
low	Intelligence	Stupidity
high	Stupidity	Stupidity

  - "Intelligence" ⇒ collusion via "price-trigger punishment"
  - "Stupidity" ⇒ collusion via biased learning + "hub-spoke conspiracy"

Research questions

How would AI speculators trade under asymmetric information?

- Can they form collusion in the absence of communication?
- If so, what is the mechanism behind the “algo collusion?”

This paper:

- Informed AI speculators can collude and achieve supra-competitive profits
- Two distinct types of collusive behaviors

Mkt. efficiency \ Info. asymmetry	low	high
low	Intelligence	Stupidity
high	Stupidity	Stupidity

  - “Intelligence” ⇒ collusion via “price-trigger punishment”
  - “Stupidity” ⇒ collusion via biased learning + “hub-spoke conspiracy”

Research questions

How would AI speculators trade under asymmetric information?

- Can they form collusion in the absence of communication?
- If so, what is the mechanism behind the "algo collusion?"

This paper:

- Informed AI speculators can collude and achieve supra-competitive profits
- Two distinct types of collusive behaviors

		Info. asymmetry: low	Info. asymmetry: high
Mkt. efficiency	low	Intelligence	Stupidity
	high	Stupidity	Stupidity

  - "Intelligence" ⇒ collusion via "price-trigger punishment"
  - "Stupidity" ⇒ collusion via biased learning + "hub-spoke conspiracy"

Research questions

How would AI speculators trade under asymmetric information?

- Can they form collusion in the absence of communication?
- If so, what is the mechanism behind the “algo collusion?”

This paper:

- Informed AI speculators can collude and achieve supra-competitive profits
- Two distinct types of collusive behaviors

		Info. asymmetry: low	Info. asymmetry: high
Mkt. efficiency	low	Intelligence	Stupidity
	high	Stupidity	Stupidity

  - “Intelligence” ⇒ collusion via “price-trigger punishment”
  - “Stupidity” ⇒ collusion via biased learning + “hub-spoke conspiracy”

Research design & approach

Simulation experiments based on a theoretical model

- Proof-of-concept illustration
  - Similar to traditional theoretical studies
- Experimental study on the "psychology" of AI
  - Similar to traditional experimental studies on human psychology

Experimental laboratory is built on Kyle (1985) + Vayanos_Vila (2021)

- Multiple informed speculators + a representative preferred-habitat investor
- However, informed speculators are not rational-expectation agents
- Instead, each informed speculator is a Q-learning algorithm

Research design & approach

Simulation experiments based on a theoretical model

- Proof-of-concept illustration
  - Similar to traditional theoretical studies
- Experimental study on the "psychology" of AI
  - Similar to traditional experimental studies on human psychology

Experimental laboratory is built on Kyle (1985) + Vayanos_Vila (2021)

- Multiple informed speculators + a representative preferred-habitat investor
- However, informed speculators are not rational-expectation agents
- Instead, each informed speculator is a Q-learning algorithm

Outline

1. Theoretical benchmark

2. Simulation experiments

- **Q-learning algorithms**
- **Experimental design**
- **Simulation results**

Four types of agents

(1) Risk-neutral informed speculators, indexed by $i = 1, \cdots, I$

$$\max_{\{x_{i,t}\}_{t\geq 0}} \mathbb{E}\left[\sum_{t=0}^{\infty} \rho^t (v_t - p_t) x_{i,t}\right],$$

where $v_t \sim^{i.i.d.} N(\overline{v}, \sigma_v^2)$ is the fundamental value, and p_t is the market price

(2) A representative preferred-habitat investor with demand curve:

$$z_t = -\xi(p_t - \overline{v}), \quad \text{with } \xi > 0$$

(3) A representative noise trader with order flows: $u_t \sim N(0, \sigma_u^2)$

(4) A market maker who determines the market price p_t.

Four types of agents

(1) Risk-neutral informed speculators, indexed by $i = 1, \cdots, I$

$$\max_{\{x_{i,t}\}_{t\geq 0}} \mathbb{E}\left[\sum_{t=0}^{\infty} \rho^t (v_t - p_t) x_{i,t}\right],$$

where $v_t \sim^{i.i.d.} N(\overline{v}, \sigma_v^2)$ is the fundamental value, and p_t is the market price

(2) A representative preferred-habitat investor with demand curve:

$$z_t = -\xi(p_t - \overline{v}), \quad \text{with } \xi > 0$$

(3) A representative noise trader with order flows: $u_t \sim N(0, \sigma_u^2)$

(4) A market maker who determines the market price p_t.

Four types of agents

(1) Risk-neutral informed speculators, indexed by $i = 1, \cdots, I$

$$\max_{\{x_{i,t}\}_{t\geq 0}} \mathbb{E}\left[\sum_{t=0}^{\infty} \rho^t (v_t - p_t) x_{i,t}\right],$$

where $v_t \sim^{i.i.d.} N(\overline{v}, \sigma_v^2)$ is the fundamental value, and p_t is the market price

(2) A representative preferred-habitat investor with demand curve:

$$z_t = -\xi(p_t - \overline{v}), \quad \text{with } \xi > 0$$

(3) A representative noise trader with order flows: $u_t \sim N(0, \sigma_u^2)$

(4) A market maker who determines the market price p_t.

Four types of agents

(1) Risk-neutral informed speculators, indexed by $i = 1, \cdots, I$

$$\max_{\{x_{i,t}\}_{t\geq 0}} \mathbb{E}\left[\sum_{t=0}^{\infty} \rho^t (v_t - p_t) x_{i,t}\right],$$

where $v_t \sim^{i.i.d.} N(\overline{v}, \sigma_v^2)$ is the fundamental value, and p_t is the market price

(2) A representative preferred-habitat investor with demand curve:

$$z_t = -\xi(p_t - \overline{v}), \quad \text{with } \xi > 0$$

(3) A representative noise trader with order flows: $u_t \sim N(0, \sigma_u^2)$

(4) A market maker who determines the market price p_t.

Timeline within each period t

- "Beginning":
  - The noise trader submits its order to buy u_t quantity of the asset
  - Informed speculator i observes v_t, but not u_t
  - Informed speculator i submits order $x_{i,t}$
- "End":
  - The market maker sets price p_t

$$\min_{p_t} \mathbb{E}\left[(y_t + z_t)^2 + \theta(p_t - v_t)^2 \middle| y_t \right], \quad \text{with } y_t = \sum_{i=1}^{I} x_{i,t} + u_t$$

    knows the demand curve for z_t, observes y_t, but not single flows or v_t
  - Liquidation value v_t is realized, so are trading profits for all agents

Timeline within each period t

- "Beginning":
  - The noise trader submits its order to buy u_t quantity of the asset
  - Informed speculator i observes v_t, but not u_t
  - Informed speculator i submits order $x_{i,t}$

- "End":
  - The market maker sets price p_t

$$\min_{p_t} \mathbb{E}\left[(y_t + z_t)^2 + \theta(p_t - v_t)^2 \middle| y_t \right], \quad \text{with } y_t = \sum_{i=1}^{I} x_{i,t} + u_t$$

    knows the demand curve for z_t, observes y_t, but not single flows or v_t
  - Liquidation value v_t is realized, so are trading profits for all agents

Market making

The FOC of the market maker's optimal pricing problem is

$$p_t = \overline{v} + \frac{\xi}{\xi^2 + \theta} y_t + \frac{\theta}{\xi^2 + \theta} \left(\mathbb{E}\left[v_t | y_t \right] - \overline{v} \right)$$

Economic interpretation:

- If $\xi \approx 0$ or $\theta \approx \infty$, the market maker focuses on minimizing pricing errors:

$$p_t = \mathbb{E}\left[v_t | y_t \right],$$

where efficient prices prevail, like in Kyle (1985)

- If $\xi \approx \infty$ or $\theta \approx 0$, the market maker focuses on minimizing inventory costs:

$$p_t = \overline{v} + \xi^{-1} y_t,$$

where prices are inefficient, like in Kyle_Xiong (2001)

Market making

The FOC of the market maker's optimal pricing problem is

$$p_t = \overline{v} + \frac{\xi}{\xi^2 + \theta} y_t + \frac{\theta}{\xi^2 + \theta} \left(\mathbb{E}\left[v_t | y_t \right] - \overline{v} \right)$$

Economic interpretation:

- If $\xi \approx 0$ or $\theta \approx \infty$, the market maker focuses on minimizing pricing errors:

$$p_t = \mathbb{E}\left[v_t | y_t \right],$$

  where efficient prices prevail, like in Kyle (1985)

- If $\xi \approx \infty$ or $\theta \approx 0$, the market maker focuses on minimizing inventory costs:

$$p_t = \overline{v} + \xi^{-1} y_t,$$

  where prices are inefficient, like in Kyle_Xiong (2001)

Market making

The FOC of the market maker's optimal pricing problem is

$$p_t = \overline{v} + \frac{\xi}{\xi^2 + \theta} y_t + \frac{\theta}{\xi^2 + \theta} \left(\mathbb{E}\left[v_t | y_t \right] - \overline{v} \right)$$

Economic interpretation:

- If $\xi \approx 0$ or $\theta \approx \infty$, the market maker focuses on minimizing pricing errors:

$$p_t = \mathbb{E}\left[v_t | y_t \right],$$

  where efficient prices prevail, like in Kyle (1985)

- If $\xi \approx \infty$ or $\theta \approx 0$, the market maker focuses on minimizing inventory costs:

$$p_t = \overline{v} + \xi^{-1} y_t,$$

  where prices are inefficient, like in Kyle_Xiong (2001)

Equilibria

Informed speculator *i*'s order is

$$x_{i,t} = \chi(v_t - \overline{v}).$$

Different types of equilibrium:

- Non-collusive: each speculator sets its χ, taking others' χ as given
- Perfect cartel: all I speculators submit orders jointly, like a monopoly
- Price-trigger collusive: speculators agree on some trading strategy, with observations of abnormal prices get punished

If price-trigger collusive equilibrium exists, it must be

$$\chi^M \leq \chi^C < \chi^N$$

Equilibria

Informed speculator *i*'s order is

$$x_{i,t} = \chi(v_t - \overline{v}).$$

Different types of equilibrium:

- Non-collusive: each speculator sets its χ, taking others' χ as given
- Perfect cartel: all I speculators submit orders jointly, like a monopoly
- Price-trigger collusive: speculators agree on some trading strategy, with observations of abnormal prices get punished

If price-trigger collusive equilibrium exists, it must be

$$\chi^M \leq \chi^C < \chi^N$$

Equilibria

Informed speculator *i*'s order is

$$x_{i,t} = \chi(v_t - \overline{v}).$$

Different types of equilibrium:

- Non-collusive: each speculator sets its χ, taking others' χ as given
- Perfect cartel: all I speculators submit orders jointly, like a monopoly
- Price-trigger collusive: speculators agree on some trading strategy, with observations of abnormal prices get punished

If price-trigger collusive equilibrium exists, it must be

$$\chi^M \leq \chi^C < \chi^N$$

Equilibria

Informed speculator *i*'s order is

$$x_{i,t} = \chi(v_t - \overline{v}).$$

Different types of equilibrium:

- Non-collusive: each speculator sets its χ, taking others' χ as given
- Perfect cartel: all I speculators submit orders jointly, like a monopoly
- Price-trigger collusive: speculators agree on some trading strategy, with observations of abnormal prices get punished

If price-trigger collusive equilibrium exists, it must be

$$\chi^M \leq \chi^C < \chi^N$$

Equilibria

Informed speculator *i*'s order is

$$x_{i,t} = \chi(v_t - \overline{v}).$$

Different types of equilibrium:

- Non-collusive: each speculator sets its χ, taking others' χ as given
- Perfect cartel: all I speculators submit orders jointly, like a monopoly
- Price-trigger collusive: speculators agree on some trading strategy, with observations of abnormal prices get punished

If price-trigger collusive equilibrium exists, it must be

$$\chi^M \leq \chi^C < \chi^N$$

Property #1 of collusion

Proposition (Impossibility of Collusion If Efficient Prices Prevail)

If prices are efficient, no collusive Nash equilibrium can be sustained by price-trigger strategies.

Intuition:

- Sustaining collusion through price-trigger strategies requires 2 conditions:
  i. Monitoring requires high price informativeness
  ii. Informational rents require low price impact of informed trading flows

- If efficient prices prevail, these 2 conditions cannot simultaneously hold

  e.g., In Kyle (1985), price informativeness is low and fixed.

In the model, price efficiency is high if θ is large or ξ is small

Property #1 of collusion

Proposition (Impossibility of Collusion If Efficient Prices Prevail)

If prices are efficient, no collusive Nash equilibrium can be sustained by price-trigger strategies.

Intuition:

- Sustaining collusion through price-trigger strategies requires 2 conditions:
    i. Monitoring requires high price informativeness
    ii. Informational rents require low price impact of informed trading flows
- If efficient prices prevail, these 2 conditions cannot simultaneously hold

  e.g., In Kyle (1985), price informativeness is low and fixed.

In the model, price efficiency is high if θ is large or ξ is small

Property #1 of collusion

Proposition (Impossibility of Collusion If Efficient Prices Prevail)

If prices are efficient, no collusive Nash equilibrium can be sustained by price-trigger strategies.

Intuition:

- Sustaining collusion through price-trigger strategies requires 2 conditions:
  i. Monitoring requires high price informativeness
  ii. Informational rents require low price impact of informed trading flows
- If efficient prices prevail, these 2 conditions cannot simultaneously hold

  e.g., In Kyle (1985), price informativeness is low and fixed.

In the model, price efficiency is high if θ is large or ξ is small

Property #1 of collusion

Proposition (Impossibility of Collusion If Efficient Prices Prevail)

If prices are efficient, no collusive Nash equilibrium can be sustained by price-trigger strategies.

Intuition:

- Sustaining collusion through price-trigger strategies requires 2 conditions:
  i. Monitoring requires high price informativeness
  ii. Informational rents require low price impact of informed trading flows

- If efficient prices prevail, these 2 conditions cannot simultaneously hold

  e.g., In Kyle (1985), price informativeness is low and fixed.

In the model, price efficiency is high if θ is large or ξ is small

Property #2 of collusion

Proposition (Existence of Collusion with Inefficient Prices)

If prices are inefficient, the collusive Nash equilibrium sustained by price-trigger strategies exists for small σ_u/σ_v and I.

Intuition:

- Small information asymmetry facilitates monitoring

  e.g., Abreu_Milgrom_Pearce (1991) and Sannikov_Skrzypacz (2007)

In the model, price efficiency is low if θ is small or ξ is large

Property #2 of collusion

Proposition (Existence of Collusion with Inefficient Prices)

If prices are inefficient, the collusive Nash equilibrium sustained by price-trigger strategies exists for small σ_u/σ_v and I.

Intuition:

- Small information asymmetry facilitates monitoring

  e.g., Abreu_Milgrom_Pearce (1991) and Sannikov_Skrzypacz (2007)

In the model, price efficiency is low if θ is small or ξ is large

Property #2 of collusion

Proposition (Existence of Collusion with Inefficient Prices)

If prices are inefficient, the collusive Nash equilibrium sustained by price-trigger strategies exists for small σ_u/σ_v and I.

Intuition:

- Small information asymmetry facilitates monitoring

  e.g., Abreu_Milgrom_Pearce (1991) and Sannikov_Skrzypacz (2007)

In the model, price efficiency is low if θ is small or ξ is large

Properties #3 of collusion

Proposition (Supra-competitive nature of collusion)

If a price-trigger collusive equilibrium exists, the average trading profits of informed speculators satisfy:

$$\pi^M \geq \pi^C > \pi^N,$$

Define $\Delta^C \equiv \frac{\pi^C - \pi^N}{\pi^M - \pi^N}$*, inequalities above imply*

$$\Delta^C \in (0, 1].$$

Property #4 and #5 of collusion

Proposition (Price informativeness of collusion)

If a price-trigger collusive equilibrium exists, the price informativeness measures (i.e., logged signal-noise ratios of prices) satisfy:

$$\mathcal{I}^M \leq \mathcal{I}^C < \mathcal{I}^N.$$

Proposition (Determinants of collusion capacity)

If a price-trigger collusive equilibrium exists, the collusion capacity and price informativeness satisfies the following properties:

(i) $\xi \uparrow \Longrightarrow \Delta^C \uparrow \quad \& \quad \mathcal{I}^C \downarrow$

(ii) $\sigma_u/\sigma_v \uparrow \quad \Longrightarrow \quad \Delta^C \downarrow \quad \& \quad \mathcal{I}^C \uparrow$

(iii) $I \uparrow \quad \Longrightarrow \quad \Delta^C \downarrow \quad \& \quad \mathcal{I}^C \uparrow$

(iv) $\rho \uparrow \Longrightarrow \Delta^C \uparrow \quad \& \quad \mathcal{I}^C \downarrow$

Property #4 and #5 of collusion

Proposition (Price informativeness of collusion)

If a price-trigger collusive equilibrium exists, the price informativeness measures (i.e., logged signal-noise ratios of prices) satisfy:

$$\mathcal{I}^M \leq \mathcal{I}^C < \mathcal{I}^N.$$

Proposition (Determinants of collusion capacity)

If a price-trigger collusive equilibrium exists, the collusion capacity and price informativeness satisfies the following properties:

(i) $\xi \uparrow \Longrightarrow \Delta^C \uparrow \quad \& \quad \mathcal{I}^C \downarrow$

(ii) $\sigma_u/\sigma_v \uparrow \quad \Longrightarrow \quad \Delta^C \downarrow \quad \& \quad \mathcal{I}^C \uparrow$

(iii) $I \uparrow \quad \Longrightarrow \quad \Delta^C \downarrow \quad \& \quad \mathcal{I}^C \uparrow$

(iv) $\rho \uparrow \Longrightarrow \Delta^C \uparrow \quad \& \quad \mathcal{I}^C \downarrow$

Outline

1. Theoretical benchmark

2. Simulation experiments

   – Q-learning algorithms

   – Experimental design

   – Simulation results

From Bellman equation to Q-function

Consider the intertemporal optimization problem:

$$\max_{\{x_{i,t}\}_{t\geq 0}} \mathbb{E}\left[\sum_{t=0}^{\infty} \rho^t (v_t - p_t) x_{i,t}\right],$$

The rational-expectations agent uses the Bellman equation:

$$V_i(s) = \max_{x\in\mathcal{X}} \left\{\mathbb{E}\left[(v-p)x|s,x\right] + \rho\mathbb{E}\left[V_i(s')|s,x\right]\right\},$$

The Q-function, $Q_i(s,x)$, captures scenarios even off the equilibrium path:

$$Q_i(s,x) = \mathbb{E}\left[(v-p)x|s,x\right] + \rho\mathbb{E}\left[V_i(s')|s,x\right].$$

A recursive relation of Q-function $Q_i(s,x)$**:**

$$Q_i(s,x) = \mathbb{E}\left[(v-p)x|s,x\right] + \rho\mathbb{E}\left[\max_{x'\in\mathcal{X}} Q_i(s',x')\middle|s,x\right].$$

From Bellman equation to Q-function

Consider the intertemporal optimization problem:

$$\max_{\{x_{i,t}\}_{t\geq 0}} \mathbb{E}\left[\sum_{t=0}^{\infty} \rho^t (v_t - p_t) x_{i,t}\right],$$

The rational-expectations agent uses the Bellman equation:

$$V_i(s) = \max_{x \in \mathcal{X}} \left\{\mathbb{E}\left[(v-p)x | s, x\right] + \rho \mathbb{E}\left[V_i(s') | s, x\right]\right\},$$

The Q-function, $Q_i(s,x)$, captures scenarios even off the equilibrium path:

$$Q_i(s,x) = \mathbb{E}\left[(v-p)x | s, x\right] + \rho \mathbb{E}\left[V_i(s') | s, x\right].$$

A recursive relation of Q-function $Q_i(s,x)$**:**

$$Q_i(s,x) = \mathbb{E}\left[(v-p)x | s, x\right] + \rho \mathbb{E}\left[\max_{x' \in \mathcal{X}} Q_i(s', x') \middle| s, x\right].$$

From Bellman equation to Q-function

Consider the intertemporal optimization problem:

$$\max_{\{x_{i,t}\}_{t\geq 0}} \mathbb{E}\left[\sum_{t=0}^{\infty} \rho^t (v_t - p_t) x_{i,t}\right],$$

The rational-expectations agent uses the Bellman equation:

$$V_i(s) = \max_{x \in \mathcal{X}} \left\{\mathbb{E}\left[(v-p)x | s, x\right] + \rho \mathbb{E}\left[V_i(s') | s, x\right]\right\},$$

The Q-function, $Q_i(s, x)$, captures scenarios even off the equilibrium path:

$$Q_i(s, x) = \mathbb{E}\left[(v-p)x | s, x\right] + \rho \mathbb{E}\left[V_i(s') | s, x\right].$$

A recursive relation of Q-function $Q_i(s, x)$**:**

$$Q_i(s, x) = \mathbb{E}\left[(v-p)x | s, x\right] + \rho \mathbb{E}\left[\max_{x' \in \mathcal{X}} Q_i(s', x') \middle| s, x\right].$$

From Bellman equation to Q-function

Consider the intertemporal optimization problem:

$$\max_{\{x_{i,t}\}_{t\geq 0}} \mathbb{E}\left[\sum_{t=0}^{\infty} \rho^t (v_t - p_t) x_{i,t}\right],$$

The rational-expectations agent uses the Bellman equation:

$$V_i(s) = \max_{x\in\mathcal{X}} \left\{\mathbb{E}\left[(v-p)x|s,x\right] + \rho\mathbb{E}\left[V_i(s')|s,x\right]\right\},$$

The Q-function, $Q_i(s,x)$, captures scenarios even off the equilibrium path:

$$Q_i(s,x) = \mathbb{E}\left[(v-p)x|s,x\right] + \rho\mathbb{E}\left[V_i(s')|s,x\right].$$

A recursive relation of Q-function $Q_i(s,x)$**:**

$$Q_i(s,x) = \mathbb{E}\left[(v-p)x|s,x\right] + \rho\mathbb{E}\left[\max_{x'\in\mathcal{X}} Q_i(s',x')\middle| s,x\right].$$

Q-learning algorithms

Challenges:

(1) Unknown conditional distribution $\mathbb{E}[\cdot|s, x]$

(2) Unknown Q values at the off-equilibrium pairs (s, x)

The Q-learning program:

$$\widehat{Q}_{i,t+1}(s_t, x_{i,t}) = (1-\alpha)\underbrace{\widehat{Q}_{i,t}(s_t, x_{i,t})}_{\text{Past knowledge}} + \alpha \underbrace{\left[(v_t - p_t)x_{i,t} + \rho \max_{x \in \mathcal{X}} \widehat{Q}_{i,t}(s_{t+1}, x)\right]}_{\text{Present learning based on a new experiment}},$$

where α governs the “forgetting” rate.

Q-learning algorithms

Challenges:

(1) Unknown conditional distribution $\mathbb{E}[\cdot|s, x]$

(2) Unknown Q values at the off-equilibrium pairs (s, x)

The Q-learning program:

$$\widehat{Q}_{i,t+1}(s_t, x_{i,t}) = (1-\alpha)\underbrace{\widehat{Q}_{i,t}(s_t, x_{i,t})}_{\text{Past knowledge}} + \alpha\underbrace{\left[(v_t - p_t)x_{i,t} + \rho \max_{x \in \mathcal{X}} \widehat{Q}_{i,t}(s_{t+1}, x)\right]}_{\text{Present learning based on a new experiment}},$$

where α governs the “forgetting” rate.

Exploitation-exploration tradeoff

The action $x_{i,t}$ is chosen as follows:

$$x_{i,t} = \begin{cases} \mathrm{argmax}_{x \in \mathcal{X}} \, \widehat{Q}_{i,t}(s_t, x), & \text{with prob. } 1 - \varepsilon_t, \quad \text{(exploitation)} \\ \widetilde{x} \sim \text{uniform distribution on } \mathcal{X}, & \text{with prob. } \varepsilon_t. \quad \text{(exploration)} \end{cases}$$

where $\varepsilon_t = e^{-\beta t}$.

- Exploitation: A greedy approach to exploit what has already been learned
- Exploration: Improve knowledge about each possible action

Exploration generates "off-equilibrium" deviation experimentation

- Crucial for machines to form a collusion

Exploitation-exploration tradeoff

The action $x_{i,t}$ is chosen as follows:

$$x_{i,t} = \begin{cases} \mathrm{argmax}_{x \in \mathcal{X}} \, \widehat{Q}_{i,t}(s_t, x), & \text{with prob. } 1 - \varepsilon_t, \quad \text{(exploitation)} \\ \widetilde{x} \sim \text{uniform distribution on } \mathcal{X}, & \text{with prob. } \varepsilon_t. \quad \text{(exploration)} \end{cases}$$

where $\varepsilon_t = e^{-\beta t}$.

- Exploitation: A greedy approach to exploit what has already been learned
- Exploration: Improve knowledge about each possible action

Exploration generates "off-equilibrium" deviation experimentation

- Crucial for machines to form a collusion

Adaptive market maker

- The market maker does not know the distributions of randomness
- It analyzes historical data on $\{v_{t-\tau}, p_{t-\tau}, z_{t-\tau}, y_{t-\tau}\}_{\tau=1}^{T_m}$ and estimates:

$$z_{t-\tau} = \xi_0 - \xi_1 p_{t-\tau},$$

$$v_{t-\tau} = \gamma_0 + \gamma_1 y_{t-\tau} + \epsilon_{t-\tau}, \text{ with } \tau = 1, \cdots, T_m.$$

- The adaptive pricing rule is

$$p_t(y) = \widehat{\gamma}_{0,t} + \frac{\theta \widehat{\gamma}_{1,t} + \widehat{\xi}_{1,t}}{\theta + \widehat{\xi}_{1,t}^2} y,$$

- Note: Results will not change with a Q-learning market maker

Outline

1. Theoretical benchmark

2. Simulation experiments

– **Q-learning algorithms**

– **Experimental design**

– **Simulation results**

State variables and parameter values

State variables: The minimalist set $s_t = \{p_{t-1}, v_t\}$

Environment parameters:

$$I = 2,\ \overline{v} = 1,\ \sigma_v = 1,\ \sigma_u = 0.1,\ \text{and } \xi = 500$$

Important: Agents do not know any environment parameters

Preference parameters: $\rho = 0.95,$ and $\theta = 0.1$

Simulation parameters: $n_x = 15,\ n_p = 31,\ n_v = 10,$ and $T_m = 10,000$

Hyperparameters: $\alpha = 0.01$ and $\beta = 10^{-5}$

Convergence

Convergence criterion:

- Each speculator's optimal strategy does not change for $1,000,000$ consecutive periods
- All $N = 1,000$ simulation sessions are simulated until convergence

Computation power:

- Implemented in C++
- 9 high-powered-computing servers, with 376 CPU cores

Outline

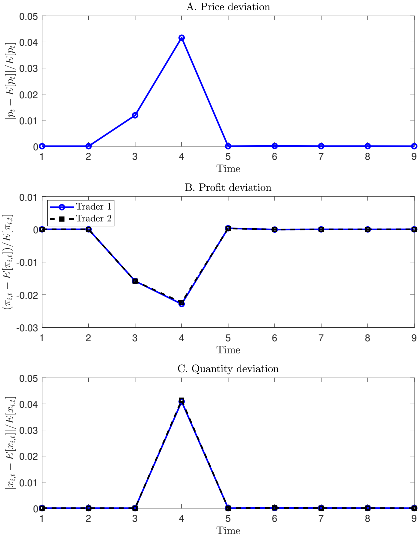
1. Theoretical benchmark

2. Simulation experiments

- **Q-learning algorithms**
- **Experimental design**
- **Simulation results**

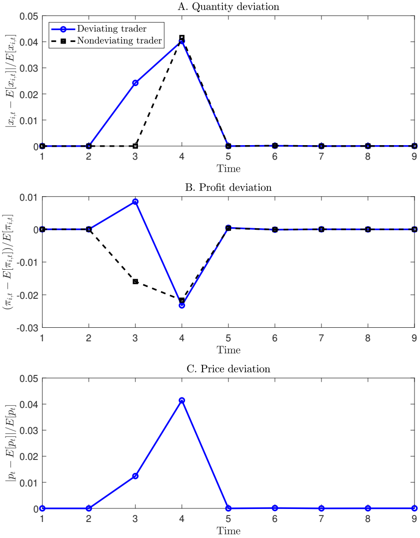
Price-trigger strategy ($\sigma_u/\sigma_v = 0.1$ and $\xi = 500$)

$$\Delta^C = 0.73 \text{ and } \pi^C/\pi^N = 1.09.$$

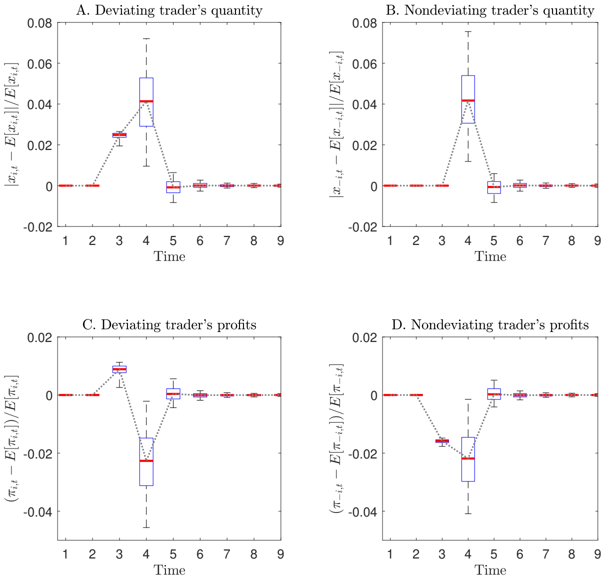


Punishment for deviation ($\sigma_u/\sigma_v = 0.1$ and $\xi = 500$)

$$\Delta^C = 0.73 \text{ and } \pi^C/\pi^N = 1.09.$$

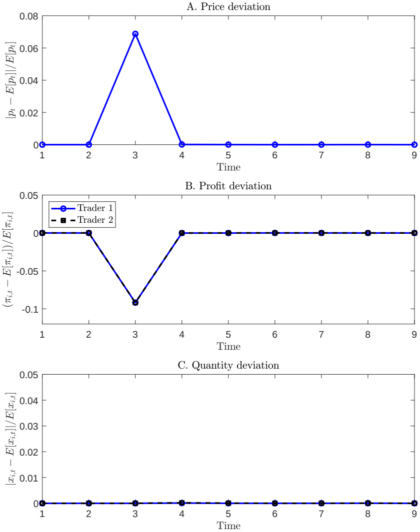


Distributions of IRFs

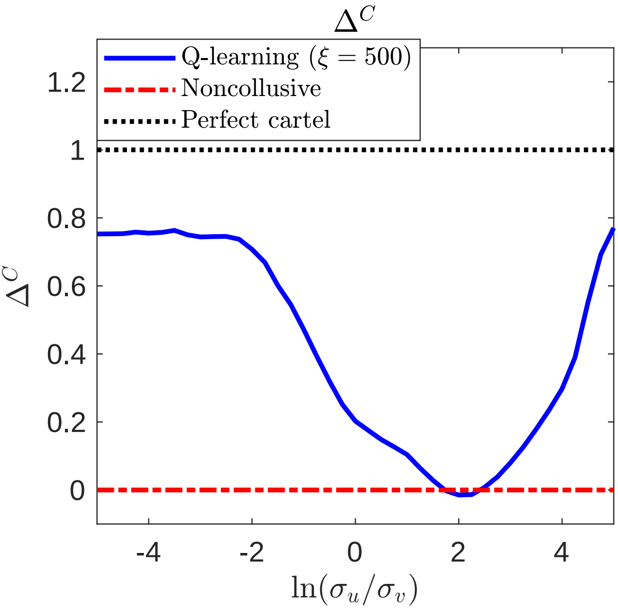


Biased learning ($\sigma_u/\sigma_v = 100$ and $\xi = 500$)

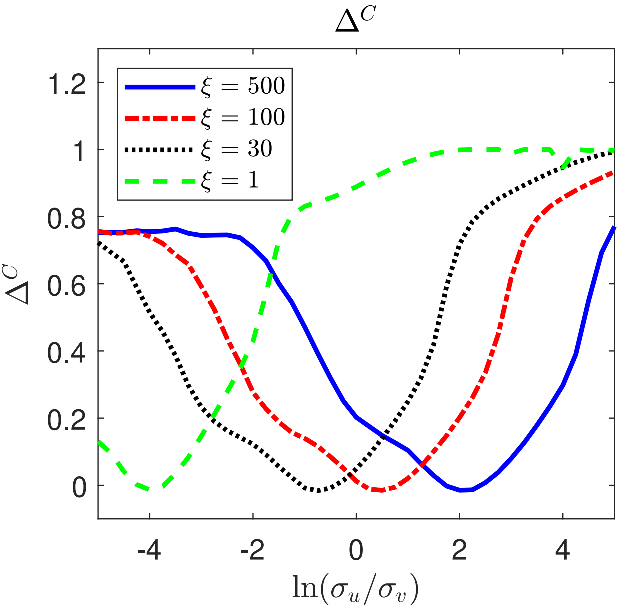
$$\Delta^C = 0.6 \text{ and } \pi^C/\pi^N = 1.075.$$



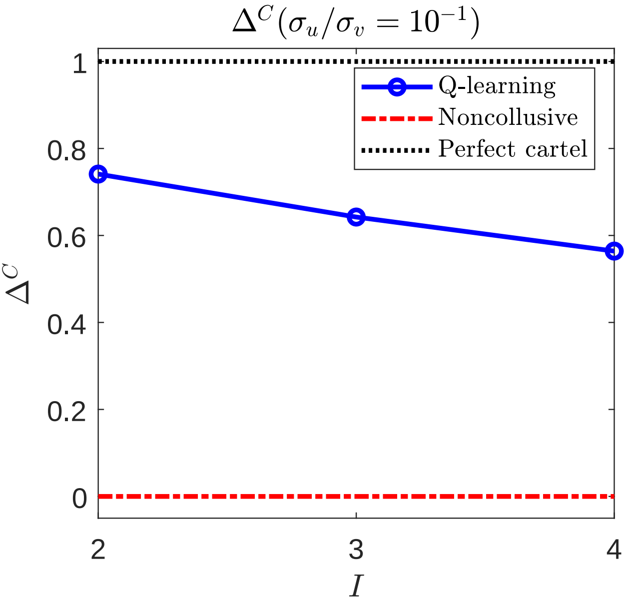
Collusion through artificial intelligence or stupidity



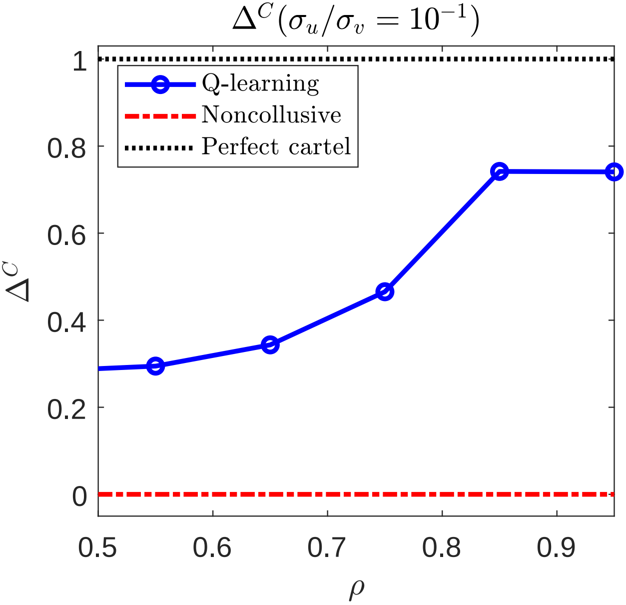
Properties #1, #2, #5(i) and #5(ii)



Properties #5(iii)



Properties #5(iv)



Conclusion

- This paper studies the "psychology" of informed AI speculators
- Algorithmic collusion emerges in the absence of communication
- How does the AI era of financial world look like?
    - Machines are often viewed as superior to humans:
        - Unconscious biases in human decision-making
        - Information-processing limitations of human brains
    - However, AI can hurt market efficiency and price informativeness
        - No matter information asymmetry is low or high
        - Due to "intelligence" or "stupidity"

Conclusion

- This paper studies the “psychology” of informed AI speculators
- Algorithmic collusion emerges in the absence of communication
- How does the AI era of financial world look like?
    - Machines are often viewed as superior to humans:
        - Unconscious biases in human decision-making
        - Information-processing limitations of human brains
    - However, AI can hurt market efficiency and price informativeness
        - No matter information asymmetry is low or high
        - Due to “intelligence” or “stupidity”

Conclusion

- This paper studies the “psychology” of informed AI speculators
- Algorithmic collusion emerges in the absence of communication
- How does the AI era of financial world look like?
  - Machines are often viewed as superior to humans:
    - Unconscious biases in human decision-making
    - Information-processing limitations of human brains
  - However, AI can hurt market efficiency and price informativeness
    - No matter information asymmetry is low or high
    - Due to “intelligence” or “stupidity”

Conclusion

- This paper studies the “psychology” of informed AI speculators
- Algorithmic collusion emerges in the absence of communication
- How does the AI era of financial world look like?
  - Machines are often viewed as superior to humans:
    - Unconscious biases in human decision-making
    - Information-processing limitations of human brains
  - However, AI can hurt market efficiency and price informativeness
    - No matter information asymmetry is low or high
    - Due to “intelligence” or “stupidity”

Literature

AI-powered pricing strategies in product markets

- Calvano_Calzolari_Denicolò_Pastorello (2020), Assad_Clark_Ershov_Xu (2021), Asker_Fershtman_Pakes (2022), Brown_MacKay (2023)
- Findings: Diminish the competitiveness and even lead to collusive behavior
- Difference: No information asymmetry, exogenous and fixed demand curve

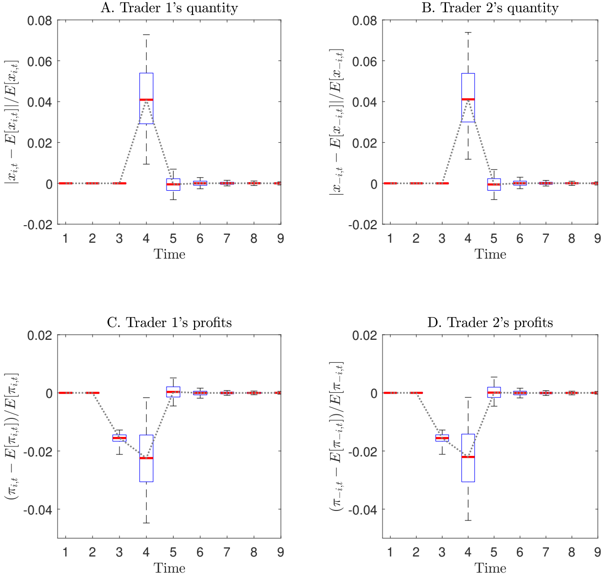
AI-powered trading strategies in financial markets

- Colliard_Foucault_Lovo (2022)
- Findings: Diminish the competitiveness and compromise price efficiency
- Difference: AI market makers, naive non-adaptive informed investors

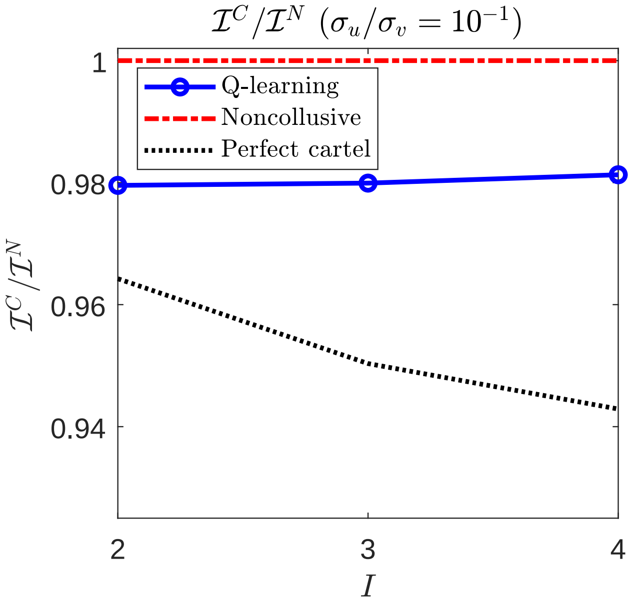
We focus on Q-learning algorithm. Why?

- A foundational framework for numerous RL algorithms
- Popularity among scientists and wall street practitioners
- Simplicity and transparency, with clear economic interpretations

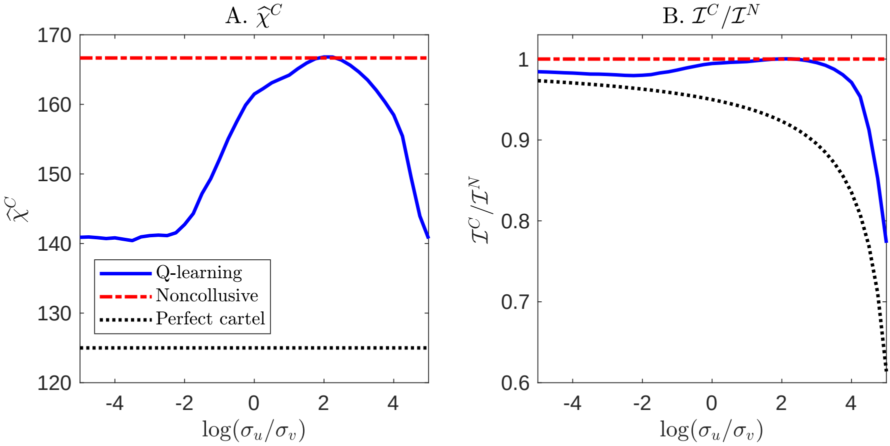
Price-trigger strategy ($\sigma_u/\sigma_v = 0.1$ and $\xi = 500$)



Properties #5(iii)

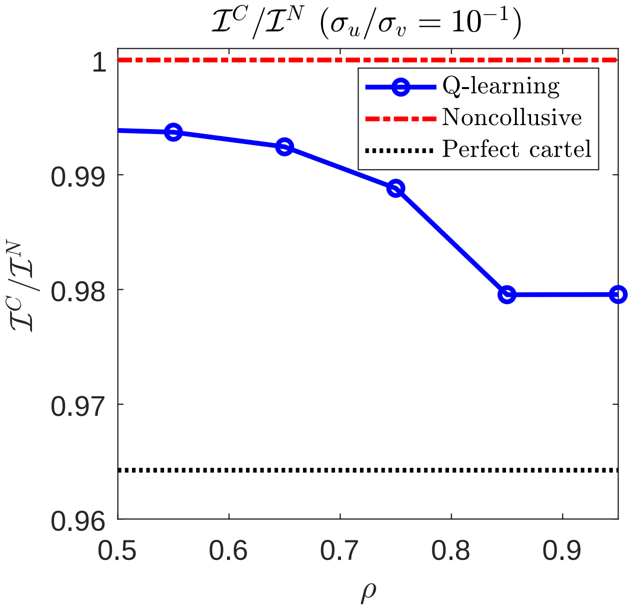


Property #4

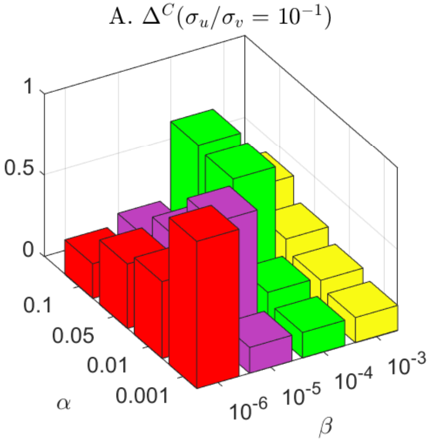


A. $\widehat{\chi}^C$
B. $\mathcal{I}^C/\mathcal{I}^N$
$\widehat{\chi}^C$
$\mathcal{I}^C/\mathcal{I}^N$
170
160
150
140
130
120
1
0.9
0.8
0.7
0.6
-4
-2
0
2
4
$\log(\sigma_u/\sigma_v)$
Q-learning
Noncollusive
Perfect cartel

Properties #5(iv)

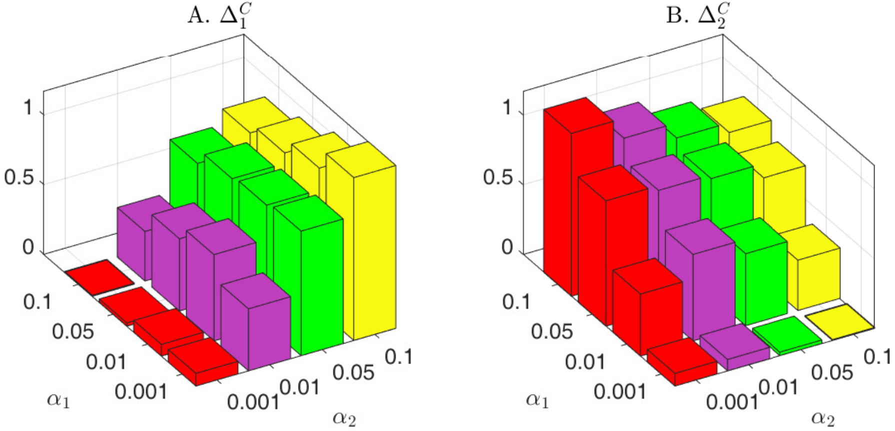


Hyperparameters α and β



$\alpha \downarrow =$ more advanced AI algo that requires higher computational power

How to sustain collusion through biased learning?



- More advanced AI algo wins, while less advanced AI algo loses
- "Hub-spoke conspiracy:" Speculators adopt the same AI algo from the same technology supplier (e.g., Johnson_Sokol, 2023)

Policy implications

To improve market efficiency and price informativeness.

- Provide market makers with more incentives and capacities for pricing error minimization
- Avoid concentration of information technologies (i.e., make sure I is sufficiently large)
- Avoid concentration among the suppliers of the AI technologies