Algorithmic Pricing in Securities Markets

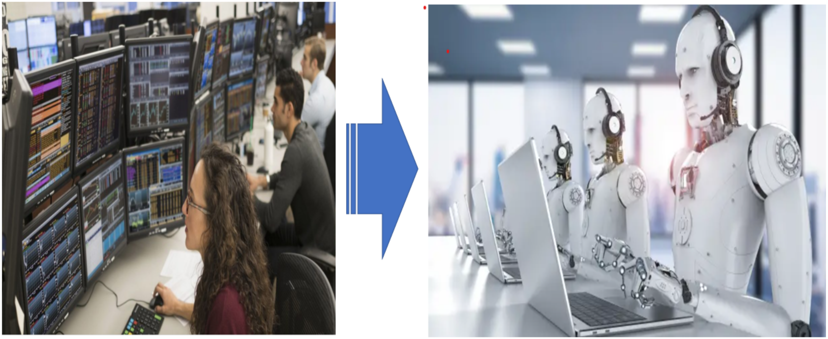
Jean-Edouard Colliard, Thierry Foucault, Stefano Lovo
HEC Paris

Keio University
November 15^{th} 2023

Roadmap

Algorithms in Securities Markets

- Rise of the " Algo market maker".



AI in Securities Markets

Opinion **The FT View**

How to prevent AI from provoking the next financial crisis

New systems have benefits for markets, but risks to stability must be managed

THE EDITORIAL BOARD

Gary Gensler, chair of the US Securities and Exchange Commission, sees regulatory intervention as essential to avoiding a financial crisis caused by AI © FT montage/Reuters/Dreamstime

Amid talk of job cuts due to artificial intelligence, Gary Gensler thinks robots will actually create more work for financial watchdogs. The US Securities and Exchange Commission chair puts the likelihood of an AI-driven financial crisis within a decade as "nearly unavoidable", without regulatory intervention. The immediate risk is more of a new financial crash than a robot takeover.

Gensler's critics argue that the risks posed by AI are not novel, and have existed for decades. But the nature of these systems, created by a handful of hugely powerful tech companies, requires a new approach beyond siloed regulation. Machines may make finance more efficient, but could do just as much to trigger the next crisis.

A new research agenda

- O'Hara (2015, JFE):
  "As a result, trading has changed, and the data that emerge from the trading process are consequently altered [...] these changes call for a new research agenda, one that recognizes how the learning models used in the past are lacking and why traditionally employed empirical methods might no longer be appropriate."
- **How do machines learn?** Reinforcement learning (Gaming (AlphaGo), Robotics, Autonomous cars etc.); an iterative process to learn payoffs of various actions via experimentation and exploitation.
- $\neq$ from Bayesian learning, assumed in standard models of market making with asymmetric information.
- **Does this matter? Does this difference in behavior generate different outcomes? Do we need a new type of "Behavioral Finance/Economics"?**

What we do

- We consider **a standard market-making game (≈ Glosten-Milgrom (1985))** but we assume that quotes are set by **Q-learning** algorithms (**"algo-MMs**) with no prior knowledge about the environment (e.g., intensity of adverse selection).
- We run **experiments** (a large number of interactions between algo-MMs and their clients, holding the environment constant) to study how Algo-MMs learn from experience and set their prices.
- We **benchmark** the observations to the predictions of the Nash equilibrium of the model (standard Bertrand equilibrium with zero expected profits for market makers).

Questions

- **Adverse selection.** Can algo MMs learn to price "adverse selection"? (e.g., iBuyers in the real estate market; Seru et al (2020))

  "Zillow may simply have realized before anyone else that adverse selection is intractable. If so,other iBuyers will eventually fail, too."
  (The Washington Post, November 2021)

- **Competition.** Can algos learn to be competitive (undercut when profitable to do so)? Major concern in online product markets.

- **Price discovery.** Can algos learn asset values ("discover fundamentals")?

Main New Findings

- **Algo-MMs learn not to be adversely selected:** Their average realized spreads are positive.
- **Algo-MMs do <u>not</u> learn to undercut:** They eventually settle on prices less competitive than the least competitive Nash equilibrium of the environment.
- **Algo-MMs** set prices that that are **more** competitive when adverse selection **increases** .
- **Algo-MMs** behave like Bayesian learner (discover asset values) even though they are not coded to be Bayesian.

Literature-Economics

- Burgeoning literature (e.g., Calvano et al. (2020), Asker et al.(2021), Banchio and Skrzypack (2022)) on algo pricing in **product markets**.

    1. Studies how algos choose prices in simulated environments.
    2. Key result: Standard reinforcement algorithms can "learn" to play "collusive outcomes".

- **Securities markets $\neq$ Product markets:** Adverse selection is central to price formation in securities markets.
- We consider a **market making game with no room for tacit collusion** and yet Algo-MMs set non competitive prices.
- We find that **adverse selection has a bright side:** It makes Algo-MMs more competitive.

Literature - Finance

- Wou, Goldstein and Ji (2023) Consider Kyle (1985) when **informed investors** use Q-learning.
- Cartea et al. (2022a) and Cartea et al. (2022b): Which reinforcement algorithms converge to Nash behavior in market-making environment (without adverse selection) and the role of tick size.
  1. Our approach is different. We do not take convergence as a goal in itself and look at what a standard algorithm does.
  2. Fujiwara-Greve and Nielsen (2021)

Roadmap

The market making game

- **A risky asset** with payoff $\tilde{v} \in \{v_H, v_L\}$ with $\Pr(\tilde{v} = V_H) = 0.5$, $E[\tilde{v}] = \mu$, $Var[\tilde{v}] = \Delta$
- **A client** considers buying one share of the asset. Her valuation for the asset is:

$$\tilde{v}_C = \tilde{w}_C + \tilde{L},$$

  where
  - "fundamentals information ": $\tilde{w}_C \in \{v_H, v_L\}$
  - "liquidity shock": $\tilde{L} \sim \mathcal{N}(0, \sigma^2)$ let $\Phi(x) := \Pr(\tilde{L} < x)$
- 2 **Market Makers** (MMs) X and Y simultaneously quote selling prices a_X and a_Y, not knowing $\tilde{v}$ and $\tilde{v}_C$.
- The client observes $a_{min} = \min_{i \in \{X,Y\}} a_i$ and buys if $v_C \geq a_{min}$.
- If a trade occurs, the aggregate profit of dealers posting the best quote is

$$(a_{min} - \tilde{v})$$

- Dealers who do not post the best quote get 0.

The Client's Demand

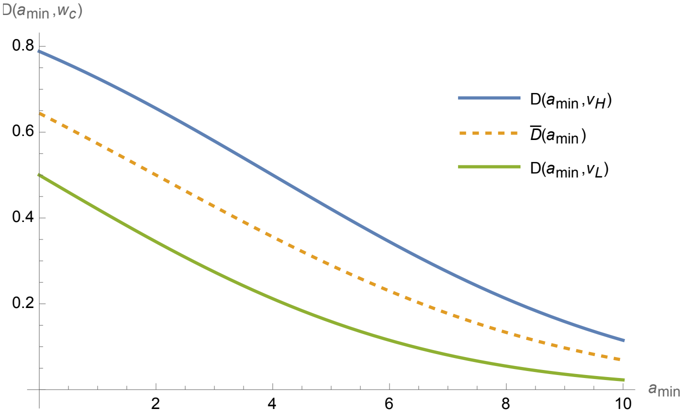
- The client's realized demand is either 1 (buy) or 0 (no trade).
- Conditional on w_C, the likelihood of a trade is:

$$D(a_{min}, w_C) = \Pr(\underbrace{w_C + \tilde{L}}_{\tilde{v}_c} \geq a_{min}) = 1 - \Phi(a_{min} - w_C).$$

- It decreases with a_{min} and it increases with w_C.
- The unconditional likelihood of a trade is:

$$\bar{D}(a_{min}) = \frac{1}{2} D(a_{min}, v_L) + \frac{1}{2} D(a_{min}, v_H).$$

Probability of a Trade



$$v_L = 0, v_H = 4, \sigma = 5$$

Economic Environment

- **Two Cases:**

1. **With adverse selection:** $\tilde{w}_C = \tilde{v}$.
   - $\Rightarrow$ The client is more likely to buy when the asset payoff is high than when it is low:

$$\Delta_D(a_{min}) = D(a_{min}, v_H) - D(a_{min}, v_L) > 0.$$

   - $\Rightarrow$ Dealers are exposed to adverse selection (more likely to sell when the asset payoff is high than low).

2. **Without adverse selection:** $\tilde{w}_C$ and $\tilde{v}$ are i.i.d. The likelihood that a client buys is $\bar{D}(a_{min})$ whether the asset payoff is high or low.

Dealers' Expected Profits

- $\bar{\Pi}(a_i, a_{-i})$: Dealer i's expected profit if dealer i posts a_i and the other dealer posts a_{-i} for $i \in \{X, Y\}$.
- Let
$$I(a_i, a_{-i}) = \begin{cases} 1 \text{ if } a_i < a_{-i} \\ \frac{1}{2} \text{ if } a_i = a_{-i} \\ 0 \text{ if } a_i > a_{-i} \end{cases}$$
- **Adverse Selection Case:**
$$\bar{\Pi}(a_i, a_{-i}) = I(a_i, a_{-i})(\frac{1}{2}D(a_i, v_H)(a_i - v_H) + \frac{1}{2}D(a_i, v_L)(a_i - v_L)),$$
$$\Leftrightarrow$$
$$\bar{\Pi}(a_i, a_{-i}) = I(a_i, a_{-})\bar{D}(a_i)\left[(a_i - \mathrm{E}(\tilde{v})) - \underbrace{\frac{Cov(D(a_i, v), v)}{\bar{D}(a_i)}}_{AdverseSelection \quad Cost}\right],$$
where $Cov(D(a_i, v), v) = \frac{\Delta \times \Delta_D}{2} > 0$.
- **No Adverse Selection Case:** $Cov(D(a_{min}, v), v)) = 0$ because the likelihood of a buy does not vary with v.

Benchmark

- **Competitive price**: a^* such that $\overline{\Pi}(a^*, a^*) = 0$
- **Nash equilibrium with at least 2 MMs:**
  1. **Without adverse selection:** All MM play $a^* = \mathrm{E}(\tilde{v})$. Independent of the risk of the asset Δ and the standard deviation of L (σ).
  2. **With adverse selection:** $a^* = \mathrm{E}(\tilde{v} \mid \mathrm{Buy}) > \mathrm{E}(\tilde{v})$ (No regret quotes, as in Glosten and Milgrom (1985)).
- Empiricists often use **two measures of illiquidity**.
  1. **Dealers' average quoted spread:** $a^* - \mathrm{E}(\tilde{v})$.
  2. **Dealers' average realized spread:** $\mathrm{E}(a^* - \tilde{v} \mid Buy)$.
  3. Dealers' total expected profit $=$ Likelihood of a trade $\times$ Average realized spread.

Experimental Hypotheses

1. **H.1.** Dealers' **average quoted spread**

$$a^* - E(\tilde{v})$$

   is strictly **larger with adverse selection.**

2. **H.2.** Dealers' **average quoted spread** declines with σ and increases with Δ if adverse selection. Without adverse selection, these parameters have no effect on the quoted spread.

3. **H.3.** Dealers' **average realized spreads** is **zero** with and without adverse selection.

- Are these (very standard) hypotheses satisfied when prices are set by Algo MMs?

Effects of exposure to adverse selection

$\tilde{v} \in \{0, 4\}$					
σ	0.5	1	3	5	7
Quoted Spread	2.00	2.00	1.24	0.68	0.47
Prob. Trade	25%	25%	37%	45%	47%
Cov$(\tilde{D}, \tilde{v})$	0.5	0.49	0.45	0.3	0.22
Realized Spread (expected payoff)	0	0	0	0	0
$\sigma = 5$					
Δv	0	2	4	6	8
Quoted Spread	0	0.16	0.68	1.65	3.02
Prob. Trade	50%	48%	45%	39%	32%
Cov$(\tilde{D}, \tilde{v})$	0	0.07	0.3	0.64	0.99
Realized spread (expected payoff)	0	0	0	0	0

Roadmap

Why Reinforcement Learning

- The theory assumes that the market makers (MMs) know $\bar{\Pi}(a_i, a_{-i})$.
- This requires a lot of knowledge on the environment: (i) The distribution of client's valuation (v_C), (ii) the distribution of the asset payoff, (iii) the number of competitors etc.

- **An alternative approach: Algo MMs:** They have no prior knowledge about the environment and learn to play the market making game via "trial and errors" ("reinforcement learning").
- We focus on the simplest type of reinforcement learning algorithm: **Q-learning**.

Clarifying our question

- **We ask**: If this standard market microstructure game is played by standard AI learning algorithms, do the resulting prices differ from those in the Nash equilibrium?

- **We DO NOT ask:** How to design the AI so that machines play the Nash equilibrium?

Q-Learning Algorithm - Description

- We restrict to a finite set of possible prices: the price grid $\mathcal{A}$.
- Holding parameters (Δ and σ) fixed, the market making game is repeated over T different episodes (i.i.d. realization of the asset payoff $\tilde{v}$and clients type $\tilde{v}_c$ across episode).
- **Q-Matrix:**$\mathcal{A} \to \mathbf{R}$, $Q_{it}(a)$ is AMM i's assessment, at the beginning of episode t, of the payoff resulting from playing price a.
- **Action:**: In episode t, AMM i chooses her price as follows:
  - **With probability** $\epsilon_t = e^{-\beta t}$**: Explore**: Pick randomly a price $a_t \in \mathcal{A}$
  - **With probability** $(1 - \epsilon_t)$**: Exploit:** Play $a_{it} = arg \max_{a \in \mathcal{A}} Q_{it}(a)$.
- **Feedback:** After choosing a_{it}, AMM i obtains her realized profit π_{it} (and has no further information).
- **Learning** AMM i updates the cell of the Q-matrix for a_{it} (and a_{it} only):

$$Q_{it+1}(a_t) = \alpha \pi_t + (1 - \alpha) Q_{it}(a_t).$$

- **Initialization:** $Q_{i0}(a)$ is chosen randomly.

One episode t simulation

In every episode t:

1. All AMMs simultaneously set their prices:
   Each AMM i choose its own price a_{it} based on its own Q-matrix at time t,
   - with Prob ϵ_t, it "explores" : random price
   - with Prob $1 - \epsilon_t$, it "exploits": $a_{it} \in \arg\max_a Q_{it}(a)$
2. Nature choses $\tilde{v}(t)$ and $\tilde{v}_c(t)$ (following the distribution function defined by parameters v_L, v_H, σ).
3. Given $a(t)$, $\tilde{v}(t)$ and $\tilde{v}_c(t)$, there is a trade or not and each AMM observes its own realized payoff resulting from its own price $a_i(t)$ (with no observation of the other AMMs prices and payoff)
4. Each AMM i updates its Q-matrix

$$Q_{it+1}(a_t) = \alpha\pi_t + (1 - \alpha)Q_{it}(a_t).$$

Looping this until $t = 200,0000$ gives us 1 experiment. We run 10,000 experiments.

Q-Learning Algorithm - Example

- **Parameters environment:** $\mathrm{E}(v) = 2$, $\Delta = 4$, $\sigma = 5$. Adverse selection ($\tilde{w}_C = \tilde{v}$) $\Rightarrow$ The competitive price is 2.68 in theory. $\mathcal{A} = \{3, 3.1\}$.
- **Parameters algorithm:** $\alpha = 0.5, \beta = 0.1$.
- Dealer X's price is fixed at $a_X = 3.1$ to simplify (will not be the case in actual experiments).
- **Expected profits for dealer Y in theory:**

$$\Pi(a_Y, 3.1) = \begin{cases} 0.12 \text{ for } a_Y = 3 \\ 0.075 \text{ for } a_Y = 3.1 \end{cases}$$

$\Rightarrow$ undercutting dealer X is optimal.

- Will AMM Y eventually learn to undercut if it uses a Q-learning algorithm?

Q-Learning Algorithm - Example

- **Initialization of the Q-matrix**

$$Q_{Y0} = \begin{pmatrix} Q_{Y0}(3) \\ Q_{Y0}(3.1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0.01 \end{pmatrix}$$

- $t = 1$**:** $\tilde{v} = v_L = 0$, $\epsilon_1 = 0.90$
  Explore
  $a = 3$
  Trade occurs $(v_L + \tilde{L} \geq 3)$.

$$\begin{aligned} Q_{Y1}(3) &= \alpha \times [3 - v_L] + (1 - \alpha) \times Q_{Y0}(3) = 1.5. \\ Q_{Y1}(3.1) &= Q_{Y0}(3.1) = 0.01 \end{aligned}$$

Q-Learning Algorithm - Example

- Reminder: Parameters algorithm: $\alpha = 0.5, \beta = 0.1$.

$$Q_{Y1} = \begin{pmatrix} 1.5 \\ 0.01 \end{pmatrix}$$

- $t = 2$: $\tilde{v} = v_L = 0$, $\epsilon_2 = 0.82$
  Explore
  $a = 3.1$
  Trade occurs ($v_L + \tilde{L} \geq 3.1$).

$$\begin{aligned} Q_{Y2}(3) &= 1.5 \\ Q_{Y2}(3.1) &= \alpha \times [3.1 - v_L] + (1 - \alpha) \times Q_{Y1}(3.1) = 1.55 \end{aligned}$$

Q-Learning Algorithm - Example

- Reminder parameters algorithm: $\alpha = 0.5, \beta = 0.1$.

$$Q_{Y2} = \begin{pmatrix} 1.5 \\ 1.55 \end{pmatrix}$$

- $t = 3$: $\tilde{v} = v_H = 4$, $\epsilon_3 = 0.74$
  Explore
  $a = 3$
  Trade occurs.

$$\begin{aligned} Q_{Y3}(3) &= \alpha \times [3 - 4] + (1 - \alpha) \times Q_{Y2}(3) = 1 \\ Q_{Y3}(3.1) &= 1.55 \end{aligned}$$

Q-Learning Algorithm - Example

- Parameters: $\alpha = 0.5, \beta = 0.1$.

$$Q_{Y3} = \begin{pmatrix} 1 \\ 1.5 \end{pmatrix}$$

- $t = 4$**:** $\tilde{v} = v_H = 4$, $\epsilon_4 = 0.67$
  Exploit
  $a = arg \max Q_{Y3} = 3.1$
  Trade does not occur,

$$\begin{aligned} Q_{Y4}(3) &= 1 \\ Q_{Y4}(3.1) &= \alpha \times 0 + (1 - \alpha) \times Q_{Y3}(3.1) = 0.75 \end{aligned}$$

- Etc T times

Convergence?

- Does $Q_{Y,t}(a_Y)$ converges to $\Pi(a_Y, 3.1)$ as t goes to infinity?

Convergence?

- Does $Q_{Y,t}(a_Y)$ converges to $\Pi(a_Y, 3.1)$ as t goes to infinity?

Lemma

Let $F_{a,t}(x) := \Pr(Q_{Y,t}(a_Y) < x)$, then there is F_a with $F_a'(x) > 0$ for all $x \in (a - v_H, a - v_L)$ such that

$$lim_{t\to\infty}||F_{a,t} - F_a|| = 0$$

Convergence?

- Does $Q_{Y,t}(a_Y)$ converges to $\Pi(a_Y, 3.1)$ as t goes to infinity?

Lemma

Let $F_{a,t}(x) := \Pr(Q_{Y,t}(a_Y) < x)$, then there is F_a with $F_a'(x) > 0$ for all $x \in (a - v_H, a - v_L)$ such that

$$lim_{t \to \infty} ||F_{a,t} - F_a|| = 0$$

- Will AMM Y eventually learn that the true expected payoff of $a = 3$ is higher...? No...Not necessarily.

Proof of Lemma

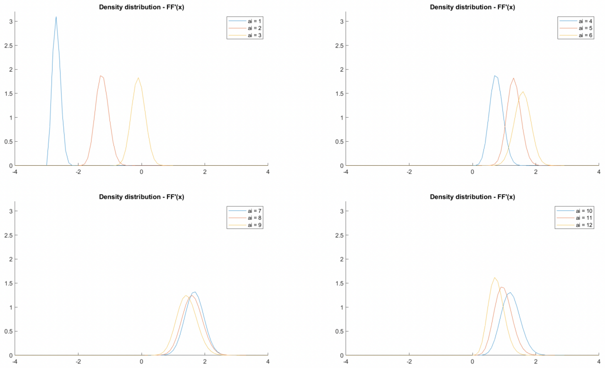
$$
\begin{aligned}
Q_{t+1}(a) &= \alpha\pi_t + (1-\alpha)Q_t(a) \\
F_{a,t}(x) &= \Pr(Q_t(a) < x) \\
&\Downarrow \\
F_{a,t+1}(x) = \Pr(Q_{t+1}(a) < x) &= \Pr\left(Q_t(a) < \frac{x-\alpha\pi_t}{1-\alpha}\right) = F_{a,t}\left(\frac{x-\alpha\pi_t}{1-\alpha}\right)
\end{aligned}
$$

Mapping M from $F_{a,t}$ into $F_{a,t+1}$:

$$
\begin{aligned}
F_{a,t+1}(x) &= \underbrace{\frac{1}{2}D(a,v_L)}_{Pr(\pi_t=a-v_L)}\ F_{a,t}\left(\frac{x-\alpha(a-v_L)}{1-\alpha}\right) + \underbrace{\frac{1}{2}D(a,v_H)}_{Pr(\pi_t=a-v_H)}\ F_{a,t}\left(\frac{x-\alpha(a-v_H)}{1-\alpha}\right) \\
&+ \underbrace{(1-\frac{1}{2}(D(a,v_L)-D(a,v_H))}_{\Pr(\pi_t=0)} F_{t,m,l}\left(\frac{x}{1-\alpha}\right) \qquad (1)
\end{aligned}
$$

is a contraction $\Rightarrow$ it has a unique fixe point $F(a)$.

Limit F



Density distribution - FF'(x)
ai = 1
ai = 2
ai = 3
Density distribution - FF'(x)
ai = 4
ai = 5
ai = 6
Density distribution - FF'(x)
ai = 7
ai = 8
ai = 9
Density distribution - FF'(x)
ai = 10
ai = 11
ai = 12

Q-Learning Algorithm -Remarks

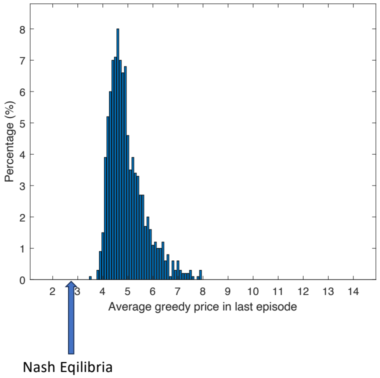
- After each round, the algorithm receives feedback ("learns") about the performance of different actions.
- "Learning" is controlled by β and α.
- Our Q-learning algorithm is not the best way to learn in our environment (e.g., a constant rate of experimentation leads to estimates of true expected payoffs of various actions that are more accurate).
- **BUT** Programmer has no knowledge of the environment$\Rightarrow$ **no possibility to optimize α and β; these are parameters.**
- The goal is not to find the algorithm that converges to the Nash equilibrium (traders do not design their algorithms to eventually play Nash) but to study the outcomes obtained if traders use standard Q-learning algorithms.

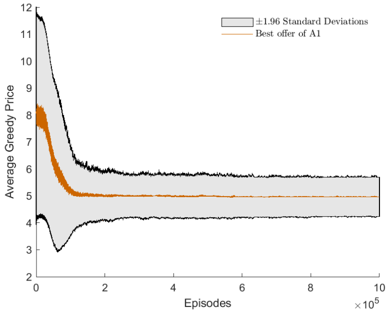
Roadmap

Implementation

- **Algorithm:** $\alpha = 0.01$, $\beta = 0.0008$, random Q_0.
- **Baseline Environment:** $E(v) = 2$, $\Delta = 4$, $\sigma = 5$ (baseline) $\Rightarrow$ such game has two pure Nash equilibrium: $a = 2.7$ and $a = 2.8$.
- For each environment, we run $K = 10,000$ "experiments", each with $T = 200,000$ episodes (clients).
- **Price Grid:** 15 prices centered around 8, tick size $\delta = 0.1$: $\mathcal{A} = \{0, 0.1, 0.2, ..., 8.5, 8.6..., 14\}$.
- **We focus on the distribution of prices (mostly mean prices) across experiments** in the last episode (that is, after learning has taken place).
- **Note:** Price discreteness implies that Nash prices yield small profits and there can be up to two pure Nash equilibria. We account for this in computing benchmarks.

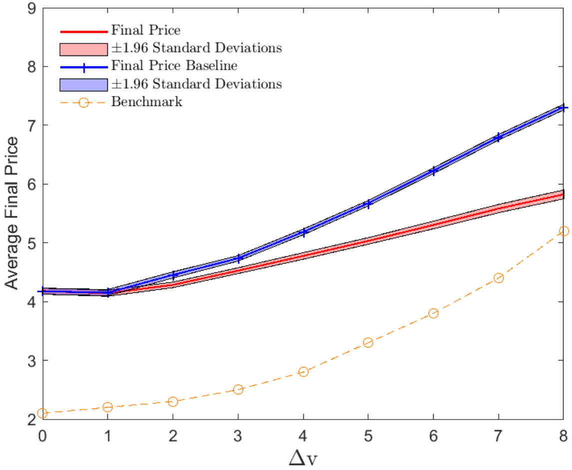
Observation 1: Prices are Not competitive





Observation 1: AMMs' prices are not competitive

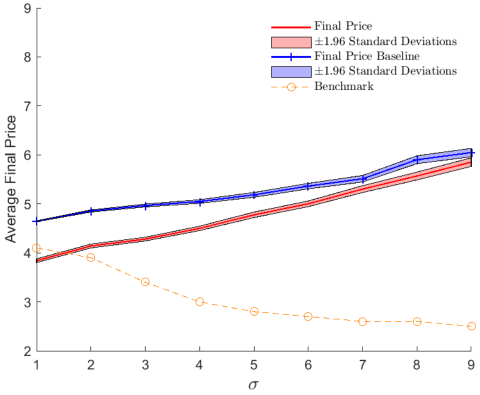
Comparative statics w.r.t. $\Delta v = v_H - v_L$



with adverse selection without adverse selection

Observation 1: AMMs' prices are not competitive and account for adverse selection

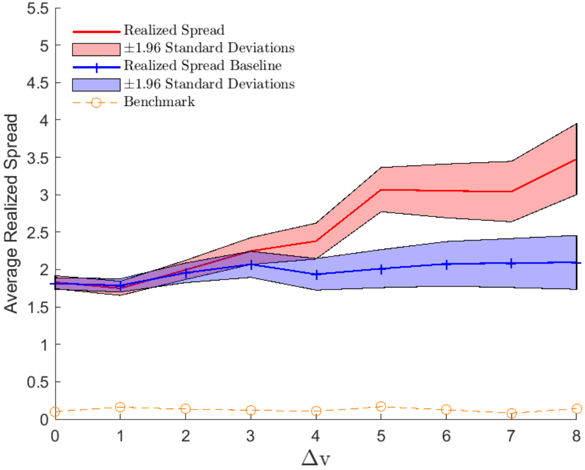
Comparative statics w.r.t. σ



with adverse selection without adverse selection

Observation 2: AMMs' profits are not competitive

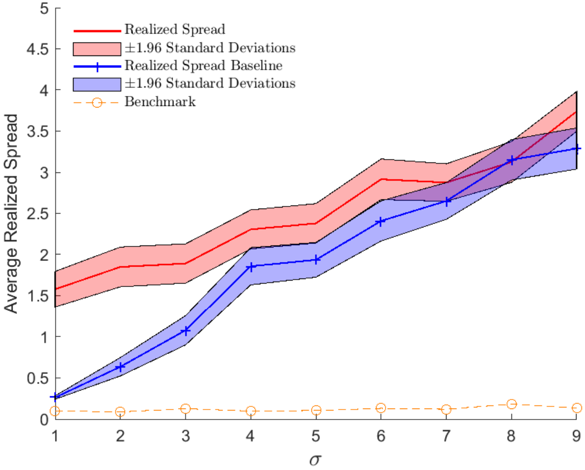
Comparative statics w.r.t. $\Delta v = v_H - v_L$



with adverse selection without adverse selection

Observation 3: Adverse selection reduces rents

AMM profits comparative statics w.r.t. σ



with adverse selection without adverse selection

Puzzles

- **H.1 is satisfied: AMMs charge larger spreads when there is adverse selection.**

- **However, other hypotheses are not satisfied:**

  1. **Realized spreads are not zero.** Algo MMs do not learn to undercut non competitive quotes.

  2. **Realized spreads are larger when there is no adverse selection.** In principle, they should be zero with or without adverse selection.

  3. **Quoted spreads and realized spreads are larger when the dispersion of traders' private valuation is larger**. The theory predicts the opposite.

- **Why does economic theory fail to explain Algo MM's behavior?**

An Explanation: Noisy Learning

- Suppose dealer X plays a price a_X **above** the least competitive Nash equilibrium. If dealer Y undercuts this price by a one tick, she increases her **expected** profit.

- **But her actual profit ($\Pi(a_Y, a_X)$) is uncertain.** In particular, there are always cases (e.g., no trade) such that this profit is less than the expected profit with undercutting.
- This uncertainty is a source of estimation errors: The Q value of undercutting is a noisy estimate of the true expected value $\Rightarrow$ It can be low even if the expected true value of undercutting is large $\Rightarrow$ There are paths on which undercutting does not happen.
- Adverse selection reduces estimation errors while an increase in the dispersion of clients' private valuation increases them (next slide).

An Explanation: Noisy Learning

- **The variance (Var) of profit ($\Pi(a_Y, a_X)$) for dealer Y at $a_Y < a_X$ is:**

  1. **Without adverse selection:**

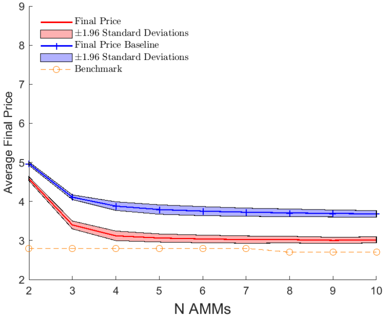
  $$\text{Var}_{Wo} = (a_Y - \mu)^2 \overline{D}(a_Y)(1 - \overline{D}(a_Y)) + \frac{\Delta^2}{4} \bar{D}(a_Y).$$

  2. **With adverse selection:**

  $$\text{Var}_{Wi} = \text{Var}_{Wo} - (\Delta \times \Delta_D)(\frac{\Delta \times \Delta_D}{4} + (a - \mu)(1 - \bar{D}(a_Y))$$

- Hence, $\text{Var}_{Wo} > Var_{Wi}$ for $a_Y < a_X$.
- Moreover both Var_{Wi} and Var_{Wo} increases with σ, holding prices constant.
- **New Insight:** Changes in the environment that makes a dealer's profit at a given price less volatile makes the outcome more competitive.

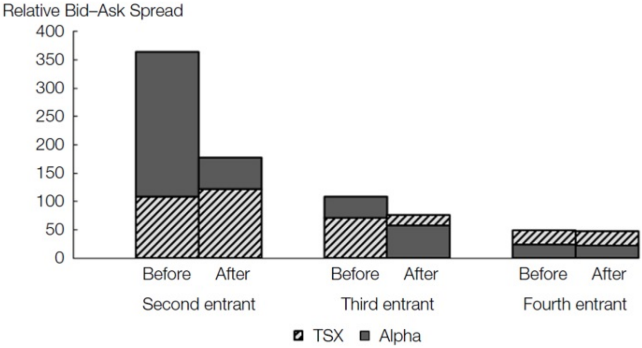
Implication: An increase in the number of AMMs make prices more competitive



- **Not obvious:** In theory, it takes only two dealers to get the Nash equilibrium.
- **Mechanism:** More AMMs reduces the variance of dealers' profit at a non competitive price and increases the average gain from undercutting $\Rightarrow$ Probability of undercutting becomes higher.

Evidence

Brogaard and Garriott (2019, JFQA):

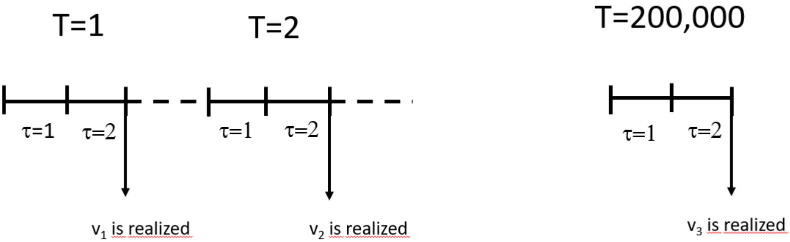


Roadmap

Price Discovery

- **Price discovery:** Dealers progressively discover v via repeated interactions with their clients.
- To study whether price discovery happens, we extend the baseline model to 2 clients: **each episode has two periods** with the same $\tilde{v}$.
- More complex dynamic problem: Dealers face a dynamic optimization problem because the price in period 1 affects the informational content of the trade in period 1 and therefore the choice of the price in period 2 (Leach and Madhavan (1995)).
- Q-learning was precisely developed for this kind of environment (estimate of value function in dynamic optimization problems).

Extension


T=1
T=2
T=200,000
τ=1
τ=2
τ=1
τ=2
τ=1
τ=2
v1 is realized
v2 is realized
v3 is realized

Main Findings

1. Algo-MMs learn to revise their prices according to the order flow: (i) They increase their offer after a buy order in period 1 and (ii) They decrease their offer after no trade in period 1.

2. Algo-MMs obtain larger rents in the second period. Consistent with the idea that less adverse selection leads algo-MMs to select less competitive prices.

Roadmap

Conclusion

- **Can algo learn to price "adverse selection"? Yes**
- **Competition.** Can algos learn to be competitive (undercut when profitable to do so)? **No. BUT:**
  1. Uncertainty on the payoff from undercutting matters. Greater uncertainty $\Rightarrow$ Less competitive outcome
  2. Adverse selection reduces this uncertainty $\Rightarrow$ Generates more competitive outcome.
  3. **Price discovery.** Can algos learn asset values ("discover fundamentals")? **Yes**
- **New important insight:** Adverse selection makes algo-pricers more competitive.

Thank You!

Roadmap

Role of Experimentation

- **Algos can earn to undercut if they explore forever**
- **BUT exploration is costly:** It implies that actions that in fact yield low expected payoffs must be chosen for a long time $\Rightarrow$ Experimenting forever is not optimal.
- See next slide.

Experimentation is costly