Algorithmic Pricing in Securities Markets

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

BSE Summer Forum Workshop on Computational and Experimental Economics
June 6^{th} 2024

Roadmap

Introduction

The Market Making Game

Algo MMs

Research question

Fact: Artificial intelligence (AI) algorithms based on reinforcement learning play an increasingly important role in many sectors of the economy

- **General question**: Do economic theory predictions based on Bayesian Nash equilibria concept hold when it is AI algorithms who interact rather than rational Bayesian agents.
- **This paper 's focus**: Do 'classical' theories predictions about price formation in the stock market hold when it is AI algorithm who set stock prices?

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.

What we do

- We consider **a standard market-making game ($\approx$ Glosten-Milgrom (1985))** but we assume that quotes are set by **Q-learning** algorithms (**"algo-MMs**) with no prior knowledge about the environment i.e. algorithms only know the set of actions.
- We run **experiments** (a large number of interactions between algo-MMs, holding their clients' demand function 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 not 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 similarly to 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.

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

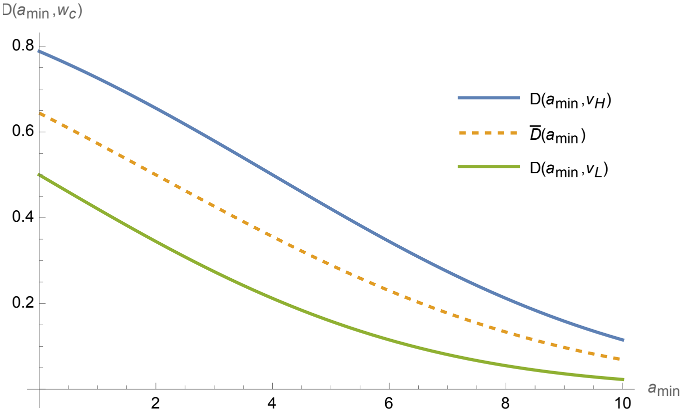
- The client's realized demand is either 1 (buy) or 0 (no trade).
- Conditional on w_C, the likelihood of a trade is:

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

- 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 $t = 200,000$ 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?

- Will AMMs eventually learn the true expected payoff resulting from an action profile?
- Fix prices (a_Y, a_X), and suppose AMMs Y an X keep playing those prices forever, then
  $Q_{Y,t}(a_Y)$ **does not converges to** $\Pi_Y(a_Y, a_X)$ **as** t **goes to infinity.**

Convergence?

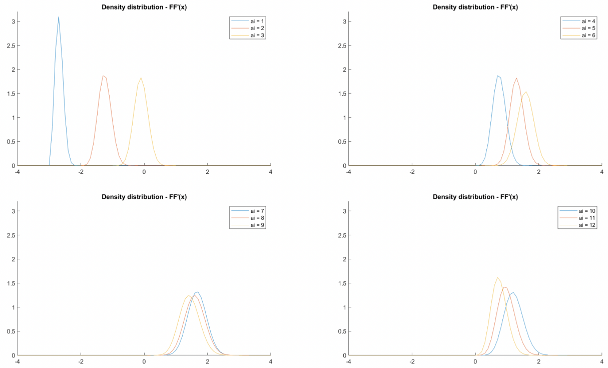
- Will AMMs eventually learn the true expected payoff resulting from an action profile?
- Fix prices (a_Y, a_X), and suppose AMMs Y an X keep playing those prices forever, then $Q_{Y,t}(a_Y)$ **does not converges to** $\Pi_Y(a_Y, a_X)$ **as** t **goes to infinity.**'

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*

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

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

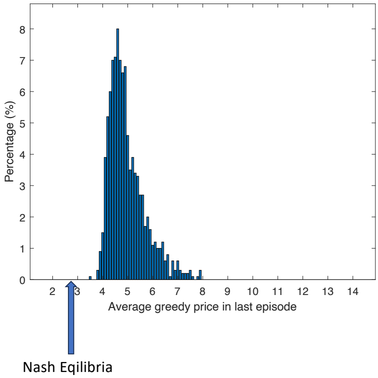
Intrinsic difference between rational agent and AI

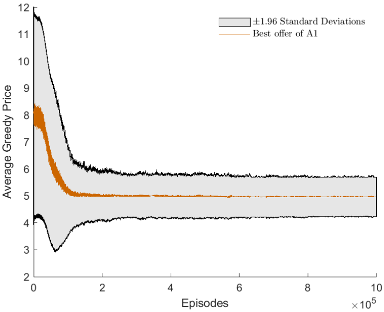
- The rational agent payoff related to an action is a real number
- The AI MM payoff related to an action is a distribution

Implementation

- **Algorithm:** $\alpha = 0.01$, $\beta = 0.0008$, random Q_0.
- **Baseline Environment:** $v_L = 0$, $v_H = 4$, $E[\tilde{v}] = 2$, $\Delta = 4$, $\sigma = 5$ (baseline)
- **Price Grid:** 15 prices centered around 8, tick size $\delta = 0.1$: $\mathcal{A}= \{0, 0.1, 0.2, ..., 8.5, 8.6..., 14\}$.
- $\Rightarrow$ 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).
- **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.

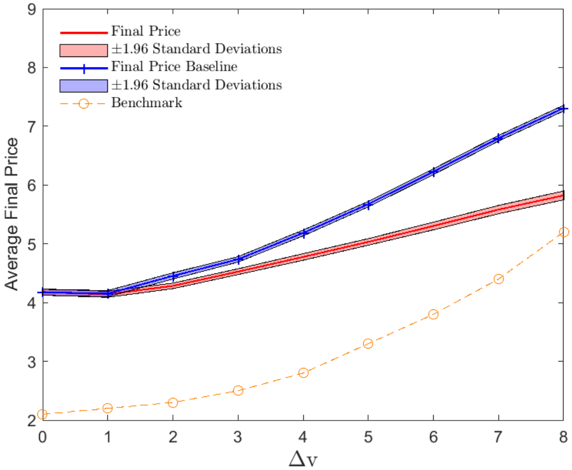
Observation 1: Prices are Not competitive





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

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



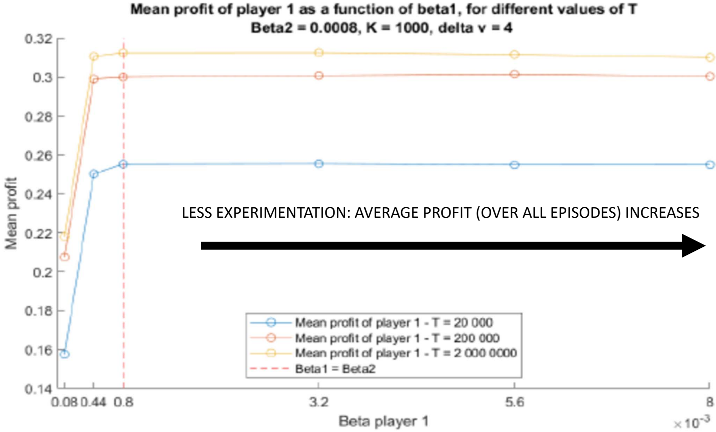
with adverse selection without adverse selection

Why AMM do not learn to undercut each other?

Role of finite Experimentation

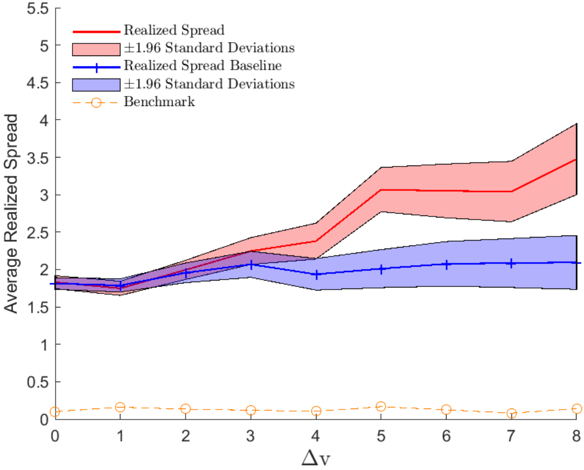
- Because probability of exploration decade exponentially at rate β algos stop exploring before learning to undercut.
- **Algos can learn to undercut if they explore forever**
- **BUT exploration is costly:** It implies that actions that in fact yield low expected payoffs must be chosen for a long time $\Rightarrow$ Experimenting forever is not optimal.

Experimentation is costly



Observation 2: AMMs' profits are not competitive

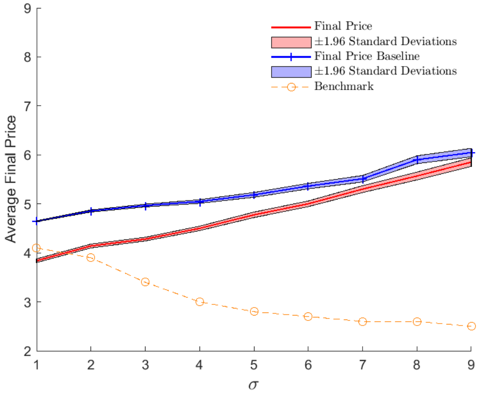
Comparative statics w.r.t. $\Delta = v_H - v_L$



with adverse selection without adverse selection

Observation 3: Adverse selection resulting from low σ makes prices more competitive

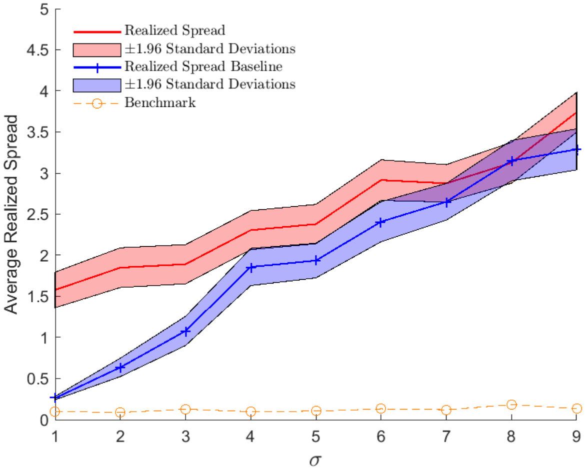
Comparative statics w.r.t. σ



with adverse selection without adverse selection

Observation 3: Adverse selection due to low σ 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.** Nash equilibrium lead to 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.

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.
- **Implication:** Price makers are better-off when algorithm set quotes rather than when playing the static Nash equilibrum, but clients are not

Thank You!