

Algorithmic Pricing in Securities Markets

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

Workshop on Artificial Intelligence in Finance
IESE Business School, Barcelona
20 March 2025

Roadmap

Introduction

The Market Making Game

Algo-MMs

Algorithmic footprints

Discussion and Conclusion

Artificial Intelligence (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 mortgage/Reuters/Dreamtime

"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 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."**

Financial Times, October 19 2023

Algorithmic Behavioral Finance (Goldstein, Spatt, Ye, 2021)

- ▶ How to model the behavior of algo market makers (algo-MMs)?
 1. AI-powered decision makers behave like any other economic agents with some specific features; e.g., they are faster and more accurate.
 2. AI-powered decision makers use **Reinforcement learning** algorithms (AlphaGo zero, Alpha zero, Robotics, Autonomous Cars etc.): They learn how to make decisions iteratively via trial and errors, without prior knowledge of their environment (“model of the world”).
- ▶ Does the second approach deliver new testable predictions?

Questions

1. How do algo-MMs behave?

- ▶ Can algo-MMs learn to price **adverse selection**? Zillow
- ▶ Can algo-MMs learn to be **competitive**? Major concern in online product markets. Algorithmic Collusion
- ▶ Can algo-MMs learn asset values, i.e., do **price discovery**?

2. What is the footprint of algo-MMs in the data?

What we do

- ▶ We consider a **standard market-making game** but quotes are set by **Q-learning** algorithms (“**algo-MMs**”) with no prior knowledge about the environment.
- ▶ We run **experiments** (a large number of interactions between algo-MMs and their clients) to study how algo-MMs learn to set their prices.
- ▶ We derive **empirical predictions** on the behavior of algo-MMs.

Main Findings - 1

- ▶ **Adverse selection:** Algo-MMs learn to not be adversely selected.
- ▶ **Competition:** Algo-MMs learn to undercut each other only partially, and settle on non-competitive prices.
- ▶ **Price discovery:** Algo-MMs learn to update their quotes based on the order flow (not for today).

Main Findings - 2

- ▶ Algo-MMs are **more competitive when learning (to compete) is easier**.
- ▶ This gives a general principle to spot the footprint of algo-MMs in the data: stronger “signal” / weaker “noise” leads to more competitive prices.
- ▶ Many applications:
 - ▶ Amount of adverse selection.
 - ▶ Asset volatility.
 - ▶ Number of AMMs.
 - ▶ Tick size.
 - ▶ Temporary shocks.

Literature - Economics

- ▶ Burgeoning literature (e.g., Calvano et al. (2020), Asker et al.(2021), Banchio and Skrzypack (2022)) on algo pricing in **product markets**.
 - ▶ Studies how algos choose prices in simulated environments.
 - ▶ 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.
- ▶ Our mechanism relies on a stochastic environment, whereas this literature considers deterministic games.

- ▶ **Dou, Goldstein and Ji (2023):** Consider Kyle (1985) when **informed investors** use Q-learning algorithms.
 - ▶ **The strategic environment is different: market-makers** set prices using Q-learning algorithms.
 - ▶ **Our mechanism is different:** Learning to compete in prices vs. learning to collude.
- ▶ **Pouget (2007):** Walrassian markets with adaptive traders.
- ▶ **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.
 - ▶ Our approach is different: We do not take convergence as a goal in itself and look at what a standard algorithm does.

Roadmap

Introduction

The Market Making Game

Algo-MMs

Algorithmic footprints

Discussion and Conclusion

The Market Making Game

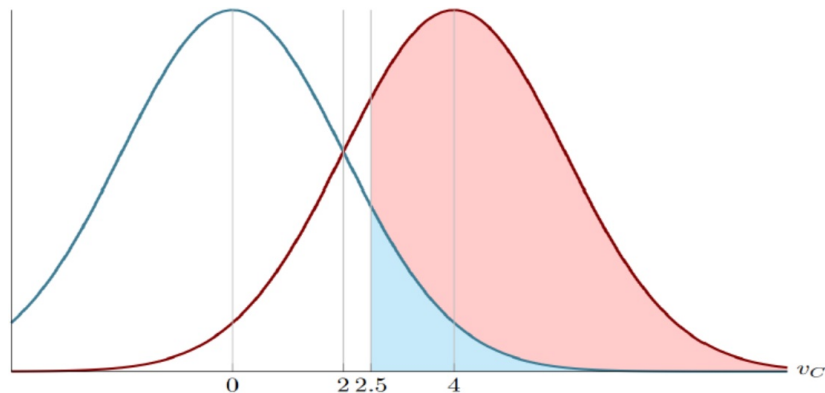
- ▶ **A risky asset** with payoff $\tilde{v} \in \{v_H, v_L\}$, equal probabilities.
- ▶ **A client** considers buying one share of the asset. Her valuation for the asset is:

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

where $\tilde{L} \sim \mathcal{N}(0, \sigma^2)$ (“liquidity shock”).

- ▶ **2 Market Makers** (MMs) X and Y simultaneously post prices a_X and a_Y , not knowing \tilde{v} and \tilde{v}_C .
- ▶ The client buys if $v_C \geq \min(a_X, a_Y)$.
- ▶ If a trade occurs, the MMs posting the best quote earn (in aggregate) $a_{min} - \tilde{v}$, and zero otherwise.

Adverse selection



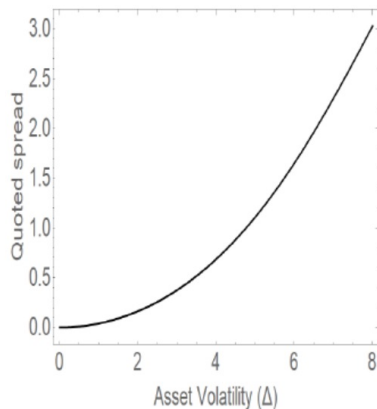
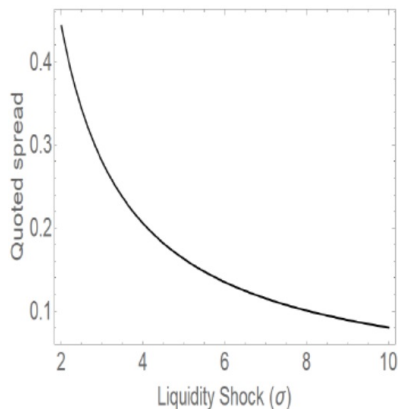
Example: $v_H = 4$, $v_L = 0$, $E(v) = 2$, $a_{\min} = 2.5$

Likelihood of a trade if $v = v_H$ = Blue + Red Areas

Likelihood of a trade if $v = v_L$ = Blue Area

Benchmark: Glosten and Milgrom (1985)

- ▶ Nash Equilibrium: $a_X = a_Y = a^*$ s.t. $\mathbb{E}[Profit] = 0$.
- ▶ Example: $E(v) = 2$, $\Delta_v = 4$, $\sigma = 5$.



Roadmap

Introduction

The Market Making Game

Algo-MMs

Algorithmic footprints

Discussion and Conclusion

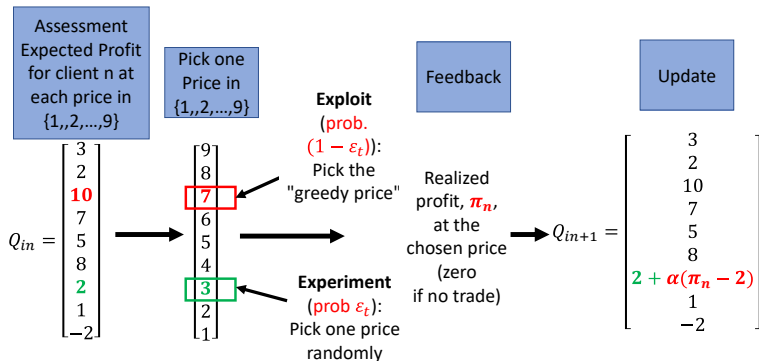
Reinforcement Learning

- ▶ We repeat the game for T independent **episodes**.
- ▶ Goal of market maker n is to maximize:

$$\mathbb{E} \left[\sum_{t=1}^T \pi_{n,t} \right].$$

- ▶ With sufficient information, Nash equilibrium obtains: play a^* in every episode.
- ▶ We assume instead market makers have zero information, they only know the price grid: **multi-armed bandit problem**.
- ▶ We assume market makers use **Q-learning algorithms** in this situation (standard method in bandit problems).

Q-Learning Algorithm: Description

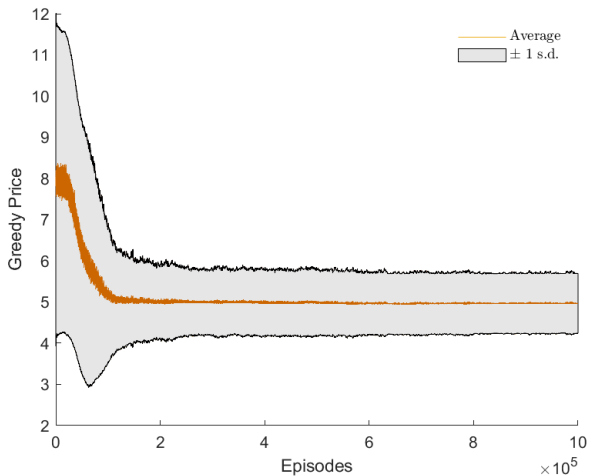


Q-learning Algorithm: Example

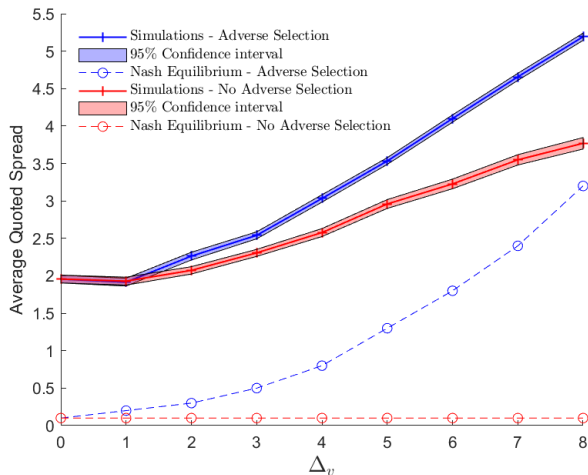
Implementation

- ▶ **Algorithm:** $\alpha = 0.01$, $\epsilon_t = e^{-\beta t}$ with $\beta = 8 \cdot 10^{-5}$, with a random Q_0 .
- ▶ **Baseline Environment:** $\mathbb{E}[\tilde{v}] = 2$, $\Delta_v = 4$, $\sigma = 5$.
- ▶ For each parametrization of the game (σ, Δ_v) , we run $K = 10,000$ “experiments”, each with $T = 10^6$ episodes.
- ▶ **Price Grid:** $\mathcal{P} = \{1.1, 1.2, \dots, 14.9\}$.
- ▶ **We focus on the distribution of prices (mostly mean prices) across experiments** after the last episode.

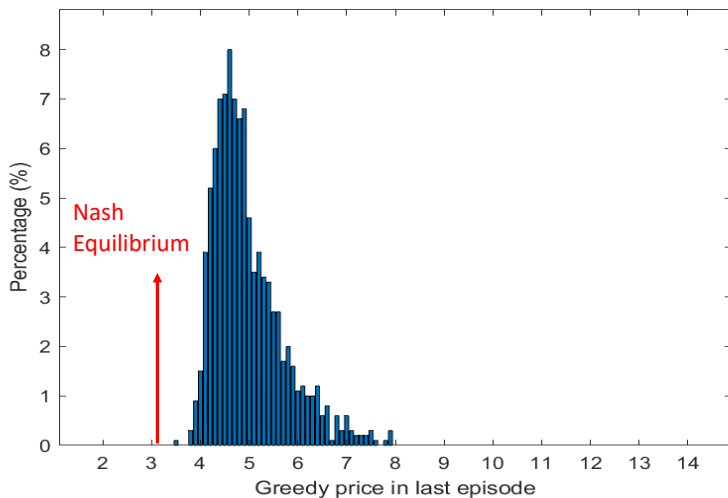
Learning



Observation 1: Algo-MMs account for adverse selection

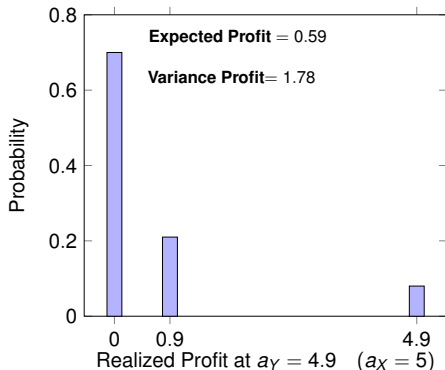


Observation 2: Algo-MMs' prices are not competitive



What's going on?

- ▶ If the algo-MMs post $a = 5$, their expected profit is **0.3**. Undercutting gives:



- ▶ **The Q-value of undercutting is a noisy estimate of expected profit.**
- ▶ With finite experimentation, **more noise** \Rightarrow **less undercutting** in a given time.

Limited experimentation

Limited experimentation plays a role here.

We know limited experimentation is a feature of the optimal solution in bandit problems (Gittins 1979, Easley and Kiefer 1988).

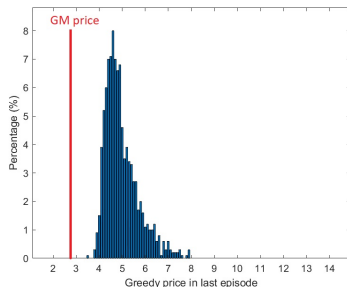


Figure: Baseline

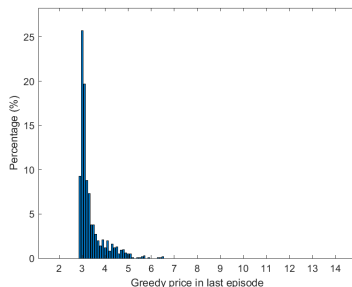


Figure: Experimentation rate $\geq 5\%$

Roadmap

Introduction

The Market Making Game

Algo-MMs

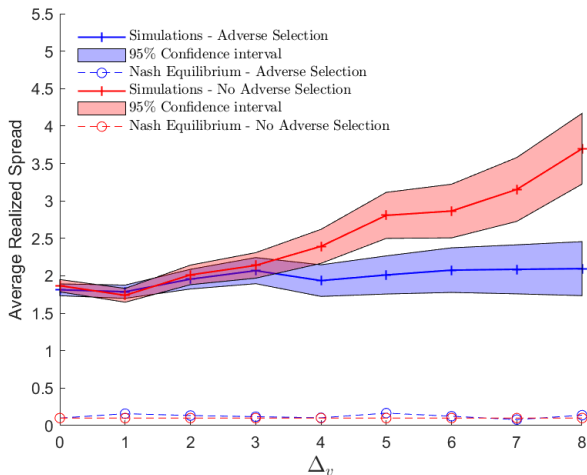
Algorithmic footprints

Discussion and Conclusion

General heuristic

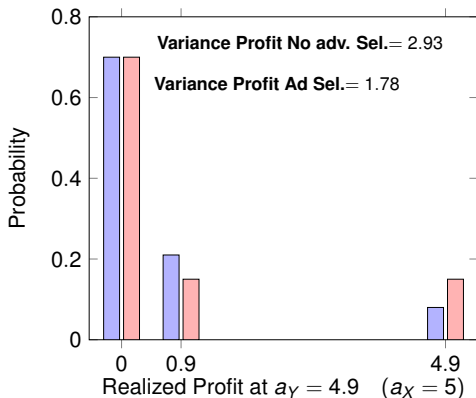
- ▶ Changes in the economic environment that make the payoff from undercutting more volatile (“**noise**”) or lower on average (“**signal**”) lead to less competitive prices.
- ▶ This heuristic can explain a number of (experimental) comparative statics results.
- ▶ While this may sound intuitive, this is not the case in standard Bertrand competition!
⇒ these results are a “footprint” specific to algorithms.

Footprint 1: Adverse selection makes Algo-MMs more competitive



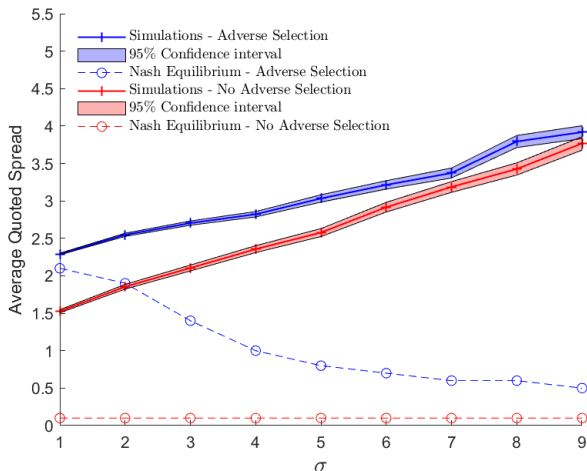
Explanation: Adverse selection reduces noise

- ▶ Same parameters as before but **Blue: Adverse selection** and **Red: No adverse selection**.

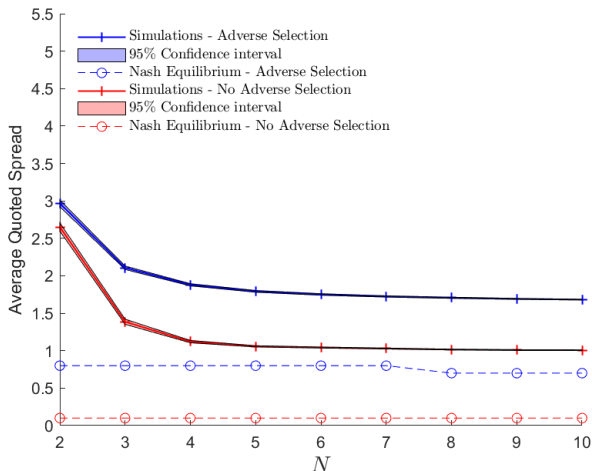


- ▶ **Adverse selection reduces the noise in the feedback received by an Algo-MM if it undercuts**
⇒ **Makes learning to undercut easier.**

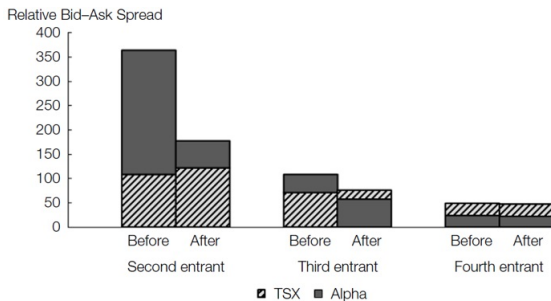
Footprint 2: quoted spreads increase with the dispersion of liquidity shocks



Footprint 3: spreads decrease with the number of Algo-MMs

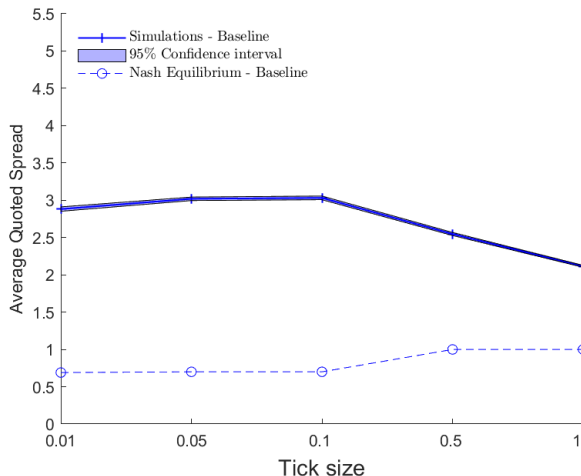


Brogaard and Garriott (2019, JFQA):



Footprint 4: spreads increase with the tick size, but...

Two effects here: signal vs. experimentation.



Footprint 5: larger reaction to loss-triggering shocks

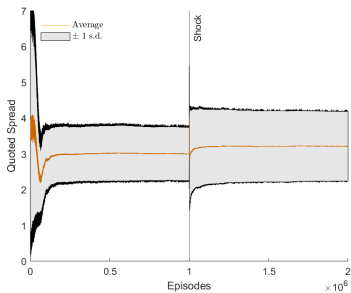


Figure: Temporary $\nearrow \Delta_v$ shock

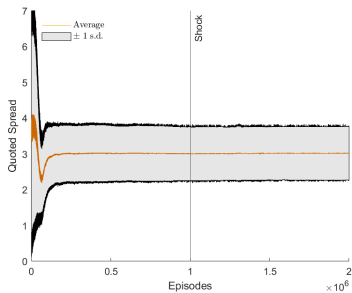


Figure: Temporary $\searrow \Delta_v$ shock

Roadmap

Introduction

The Market Making Game

Algo-MMs

Algorithmic footprints

Discussion and Conclusion

Discussion: Better algorithms?

Should market makers use “more sophisticated” algos that would learn more quickly?

Here the answer is **no**: if anything, more profitable to use slower algos as a way to commit to high prices.

Thought experiment: what if an algo-MM experimented alternative parameters for K' experiments of 1M episodes?

θ'	$K' = 100$	$K' = 200$	$K' = 300$	$K' = 400$	$K' = 500$	$K' = 600$	$K' = 700$	$K' = 800$	$K' = 900$	$K' = 1000$
(α_l, β_l)	1	1	1	1	1	1	1	1	1	1
(α_l, β_m)	1	1	1	1	1	1	1	1	1	1
(α_l, β_h)	1	1	1	1	1	1	1	1	1	1
(α_m, β_l)	0.51	0.63	0.83	0.87	0.93	0.9	0.94	0.95	0.91	0.89
(α_m, β_h)	0.5	0.37	0.28	0.16	0.06	0.06	0.19	0.17	0.3	0.29
(α_h, β_l)	1	1	1	1	1	1	1	1	1	1
(α_h, β_m)	1	1	1	1	1	1	1	1	1	1
(α_h, β_h)	1	1	1	1	1	1	1	1	1	1

Table OA.2: Deviations from (α_m, β_m) : This table gives the p-values for a test of the null hypotheses that (α_m, β_m) and θ' give the same expected payoff to dealer 1, against the alternative hypothesis that θ' is more profitable, for all possible θ' and different values of K' .

Conclusion

- ▶ Q-learning is a reasonable model of algo-MMs:
 - ▶ They learn to not be adversely selected.
 - ▶ They learn to compete to some extent.
 - ▶ They engage in price discovery (not seen today).
- ▶ Reinforcement learning should leave a footprint in the data: environments with lower “signal” / higher “noise” lead to less competitive prices.
- ▶ We derive a number of predictions specific to algo-MMs.
 - ▶ Adverse selection.
 - ▶ Asset volatility.
 - ▶ Competition.
 - ▶ Tick size.
 - ▶ Shocks to adverse selection.

Thank You!

Q-learning Algorithm: Example

- ▶ **Parameters environment:** $E(v) = 2$, $\Delta_v = 4$, $\sigma = 5$.
Adverse selection ($\tilde{w}_C = \tilde{v}$) \Rightarrow The competitive price is 2.68 in theory.
- ▶ **Parameters algorithm:** $\alpha = 0.5$, $\epsilon_t = e^{-\beta t}$ with $\beta = 0.1$. $\mathcal{P} = \{3, 3.1\}$.
- ▶ Dealer X's price is fixed at $a_X = 3.1$ to simplify (will not be the case in actual experiments).
- ▶ **True expected profits for dealer Y:** 0.12 with $a_Y = 3$ and $0.15/2 = 0.075$ with $a_Y = 3.1 \Rightarrow$ undercutting dealer X is optimal.
- ▶ Will AM Y eventually learn to undercut if it uses a Q-learning algorithm?

Q-Learning Algorithm - Example

► Initialization of the Q-matrix

$$Q_{Y_0} = \begin{pmatrix} Q_{Y_0}(3) \\ Q_{Y_0}(3.1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0.01 \end{pmatrix}$$

► $t = 1$: $\tilde{v} = v_L = 0$, $\epsilon_1 = 0.90$

Explore

$a = 3$

Trade occurs ($v_L + \tilde{L} \geq 3$).

$Q_{Y_1}(3) = \alpha \times [3 - v_L] + (1 - \alpha) \times Q_{Y_0}(3) = 1.5.$

$Q_{Y_1}(3.1) = Q_{Y_0}(3.1) = 0.01$

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 ($v_L + \tilde{L} \geq 3.1$).

$$Q_{Y2}(3.1) = \alpha \times [3.1 - v_L] + (1 - \alpha) \times Q_{Y1}(3.1) = 1.55.$$

Q-Learning Algorithm - Example

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

$$Q_{Y_2} = \begin{pmatrix} 1.5 \\ 1.55 \end{pmatrix}$$

- ▶ $t = 3$: $\tilde{v} = v_H = 4$, $\epsilon_3 = 0.74$

Explore

$a = 3$

Trade occurs.

$$Q_{Y_3}(3) = \alpha \times [3 - 4] + (1 - \alpha) \times Q_{Y_2}(3) = 0.25$$

Q-Learning Algorithm - Example

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

$$Q_{Y3} = \begin{pmatrix} 0.25 \\ 1.55 \end{pmatrix}$$

- ▶ $t = 4$: $\tilde{v} = v_H = 4, \epsilon_4 = 0.67$

Exploit

Greedy-price: $a = 3.1$

Trade does not occur,

$$Q_{Y4}(3.1) = \alpha \times 0 + (1 - \alpha) \times Q_{Y3}(3.1) = 0.775.$$

- ▶ Etc T times
- ▶ Will AM Y eventually learn that the true expected payoff of $a = 3$ is higher...? No...Not necessarily: T is finite and experimentation is less likely over time.

Science

Protecting consumers from collusive prices due to AI

Emilio Calvano, Giacomo Calzolari, Vincenzo Denicolò, Joseph E. Harrington Jr. and Sergio Pastorello

POLICY FORUM

TECHNOLOGY AND LAW

Protecting consumers from collusive prices due to AI

Price-setting algorithms can lead to noncompetitive prices, but the law is ill equipped to stop it

- ▶ Hendershott et al. (2010, JoF): Algo trading is associated with an *increase* in “realized spreads” (rents):
“This is surprising because we initially expected that if AT improved liquidity, the mechanism would be competition between liquidity providers.”

Can machines learn to price adverse selection?

THE WALL STREET JOURNAL

English Edition • Print Edition | Video | Audio | Latest Headlines | More

Latest World **Business** U.S. Politics Economy Tech Finance Opinion Arts & Culture Lifestyle Real Estate Personal Finance Health Style

BUSINESS

Zillow Quits Home-Flipping Business, Cites Inability to Forecast Prices

Termination of 'iBuying' comes after company said it was halting new home purchases for rest of 2021

Opinion | Zillow sent its algorithm to take on the housing market. The housing market won.

What if they can't? Zillow may simply have realized before anyone else that **adverse selection** is intractable. If so, other iBuyers will eventually fail, too.

Price Discovery

Learning from Order Flow

- ▶ **Price discovery:** Dealers discover the asset fundamental, v , via repeated interactions with their clients.
- ▶ **Implication:** If dealers receive requests from two clients in sequence, in the Glosten and Milgrom equilibrium:
 1. They raise their offer for the second client after observing a buy from the first (because a buy is more likely if $v = v_H$).
 2. They lower their offer for the second client if the first does not trade (because the first client is more likely not to trade if $v = v_L$).

Example

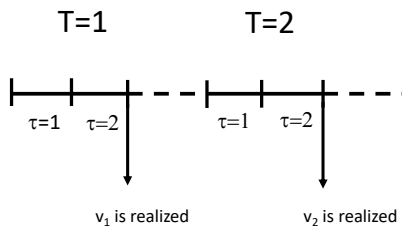
- **Parameters:** Adverse selection, $\Delta_v = 4$, $L_t \sim \mathcal{N}(0, 5)$. a_2^T (a_2^{NT}): 2nd client price after a trade (No Trade) with the first client

	σ				
	0.5	1	3	5	7
$(a_2^T - a_1^*)$	0	0	0.58	0.58	0.45
$(a_2^{NT} - a_1^*)$	0	0	-0.8	-0.6	-0.45
	Δ_v				
	0	2	4	6	8
$(a_2^T - a_1^*)$	0	0.34	0.58	0.95	0.84
$(a_2^{NT} - a_1^*)$	0	-0.36	-0.6	-1.2	-1.36

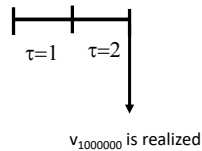
Price Discovery

- ▶ To study whether AMs update their quotes as in the Glosten and Milgrom equilibrium, we extend the baseline model to 2 clients \Rightarrow **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).

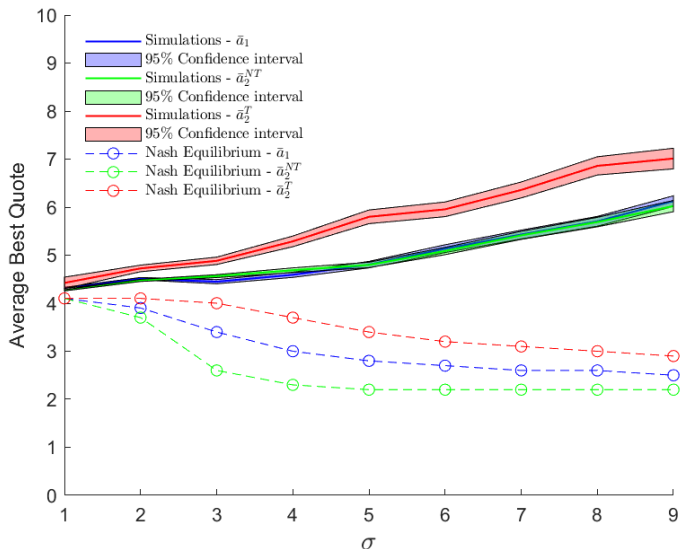
Extension



$T=1000,000$



Price Discovery



Price Discovery: Main Findings

- ▶ 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.
- ▶ Algo-MMs obtain larger rents in the second period.
 1. Consistent with the idea that less adverse selection leads algo-MMs to select less competitive prices.
 2. But lack of experimentation is an even more acute problem for the second client. Holding T fixed, fewer observations of the various states for the second client than the first \Rightarrow less scope for learning.