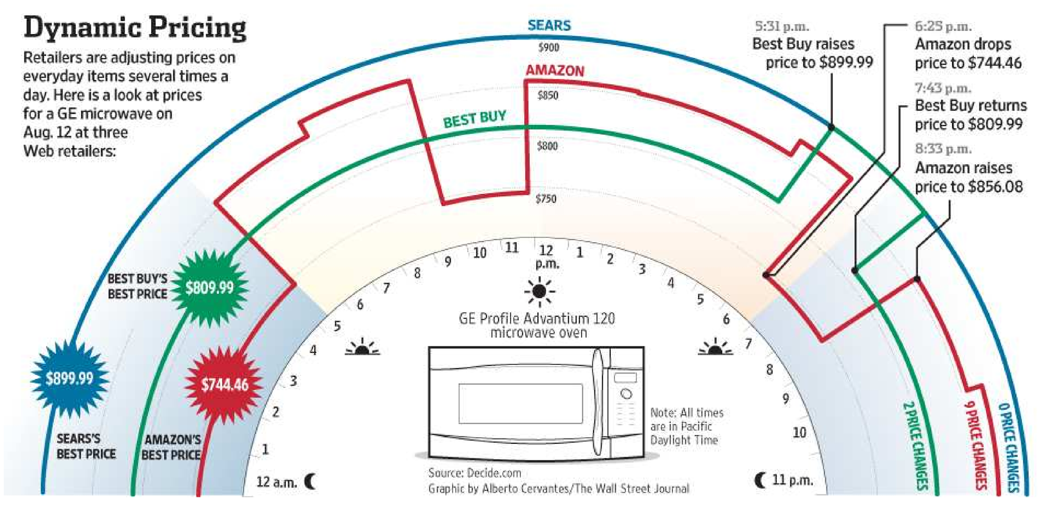
Oligopoly Dynamic Pricing: A Repeated Game with Incomplete Information

Yixuan Zhai — University of California, Davis

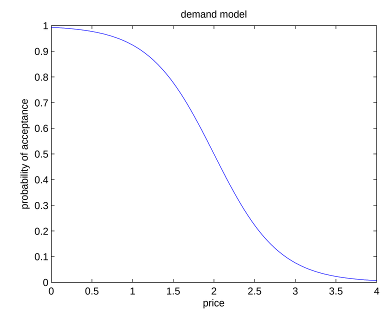
Qing Zhao — Cornell University

Oligopoly Dynamic Pricing



Single Seller Dynamic Pricing

- The *profit-maximization* problem for a seller with an *unlimited* supply of identical goods.
- The seller offers prices sequentially to a stream of potential customers.
- The marginal cost is c.
- For the t-th costumer, the seller chooses a price $p(t) \in [c, p_u]$.
- The seller experiences either success or failure.
- The probability of success at price p at any given time is $\rho(p)$.



Single Seller Dynamic Pricing

Unknown Demand Model with Finite Space Uncertainty

- The unknown underlying demand model is ρ^{ω}.
- $\rho(p)$ is unknown but belongs to a known finite set $\{\rho^{(m)}(p)\}_{m=1}^{M}$.
- $\rho(p)$ is strictly decreasing and satisfies the increasing generalized failure rate (IGFR).
- Intuitively, IGFR means given a seller can sell at price p, the probability of sale at $(1+1\%)p$ is decreasing in p.

Profit

- The expected profit at time t: $r^{(\omega)}(p, c) = (p - c)\rho^{(\omega)}(p)$.
- The profit-maximizing price $p^{(\omega)}(c) = \arg\max_{p \in [c, p_u]} r^{(\omega)}(p, c)$.

Lemma 1. *Given $\rho(p)$ is strictly decreasing and satisfies IGFR. We have a unique $p^{(\omega)}(c)$ and $r^{(\omega)}(p, c)$ is continuous and strictly increasing with p over $[c, p^{(\omega)}(c)]$.*

Oligopoly Dynamic Pricing

- N sellers.
- The marginal cost for seller i is c_i.
- Without loss of generality, we assume $c_1 < c_2 < \ldots < c_N$.
- Unknown demand model $\{\rho^{(m)}(p)\}_{m=1}^{M}$.
- Sellers propose their prices simultaneously.
- The customer accepts the lowest price p with probability $\rho(p)$.
- If multiple sellers offer the same price: one is chosen with equal probability.

An Infinitely Repeated Game

Public and private history

- Sellers' price offers are public.
- Each seller's own sale history is private.

Strategy of Seller i: A mapping σ from public and seller i's private history to a price offer at current time.

Seller i's one-shot payoff $u_i(p_i)$

- If seller i is the single lowest price offerer:
  - $u_i(p_i) = (p_i - c_i)\rho(p_i)$.
  - $u_j(p_j) = 0 \; \forall j \neq i$.
- If K sellers offer the same lowest price, then each has $u_i(p) = \frac{1}{K}(p - c)\rho(p)$.

Seller i's payoff for the infinitely repeated game:

$$U_i^{(\omega)}(\sigma) = \liminf_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} u_i^{(\omega)}(a^t(\sigma)).$$

Equilibria and Efficiency

Equilibria

- Nash Equilibrium (NE)
  - A strategy σ.
  - $U_i(\sigma) \geq U_i(\sigma_i', \sigma_{-i})$ for every seller i, every demand model ρ_m and all strategies σ'.
- Subgame perfect equilibrium
  - NE may not be sequential rational
  - For all history h^t, $\sigma_i|_{h^t}(h^\tau) = \sigma_i(h^t h^\tau)$ is a NE of the repeated game.

Efficiency

- Pareto Efficient Nash Equilibrium
  - Inefficiency of σ: exist a σ' that for all i, $U_i(\sigma) \leq U_i(\sigma')$.
- Learning efficiency: Regret
  - Regret is defined as the accumulated profit loss in the unknown demand case to the profit gained under the known demand case.

Dynamic Pricing under Known Demand Model

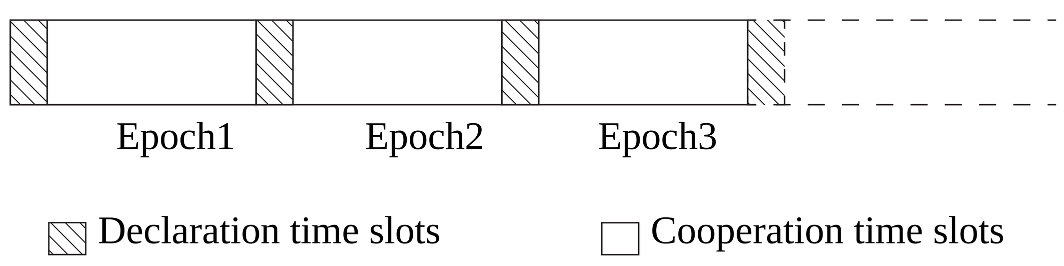
The colluding strategy σ_C under known demand model:

- Let $\rho^{(\omega)}$ be the underlying demand mode , $c_{N+1} = p^{(\omega)}(c_1)$.
- Seller 1 forms the optimal collusion of K sellers ($K = \arg\max_k \frac{1}{k} r^{(\omega)}(c_{k+1}, c_1)$) to maximize its own profit (from Lemma 1).
- Sellers with $c_i < c_{K+1}$ offer the collusive price c_{K+1}.
- Sellers with $c_i \geq c_{K+1}$ don't participate.
- Any deviations will trigger a punishment that seller 1 offers $c_2 - \epsilon$ and seller $i \neq 1$ offers c_i forever.

Theorem 1. *The colluding strategy σ_C is a subgame-perfect and Pareto-efficient Nash equilibrium.*

Demand Learning under Collusion

- The time horizon is partitioned into ***fixed length epochs***, length $l \in \mathbb{N}^+, l \geq 2$.
- Epoch: starts with a ***declaration*** time slot and then ***cooperation*** time slots.



Demand Learning under Collusion

Demand Learning under Collusion

In the declaration slot of epoch t:

- A default estimate $\hat{\omega}(1) = 1$.
- Seller 1 carries out a maximum likelihood estimate $\hat{\omega}(t)$ of the underlying demand model using its private history.
- Seller 1 then offers the profit-maximizing colluding price $\hat{p}^{(\hat{\omega}(t))}$.
- All other sellers offer the same price they offered in the cooperation slots in the previous epoch.

In all the cooperation time slots of epoch t

- Seller i with $c_i < \hat{p}^{\hat{\omega}(t)}$ offers $\hat{p}^{\hat{\omega}(t)}$.
- Not participate if $c_i \geq \hat{p}^{\hat{\omega}(t)}$.

Trigger strategy for punishing any deviations

- Any deviations in declaration slots from seller $2, \ldots, N$ and any deviations in cooperation time slots will trigger a everlasting punishment.
- Punishment is that seller 1 offers $c_2 - \epsilon$ and seller $i \neq 1$ offers c_i forever.

Properties of DLC

Theorem 2.

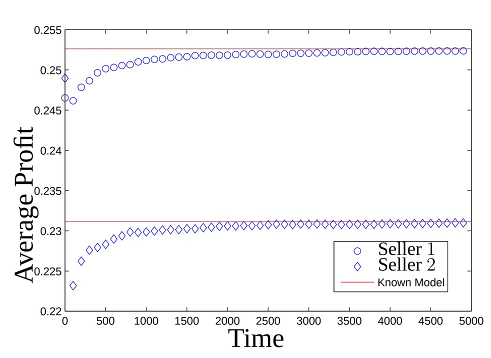
- *DLC is a subgame-perfect Nash equilibrium.*
- *DLC is a Pareto-efficient Nash equilibrium.*
- *DLC achieves a bounded regret, i.e., under any demand model $\rho^{(\omega)} \in \{\rho^{(\omega)}\}_{\omega=1}^{M}$, there exists a positive constant C such that*

$$R_{DLC} \leq C.$$

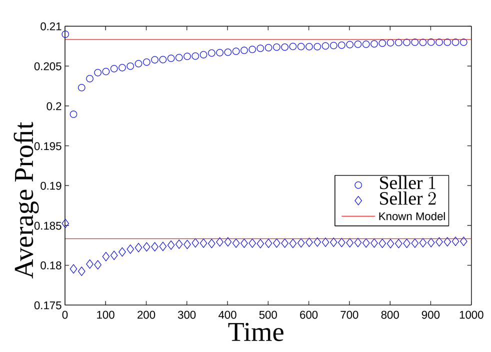
Bounded Regret

- In general, online learning problem has sublinear regret. e.g. $\log(T)$,$\sqrt{T}$.
- The action profiles converges to the optimal action profile played as under the known demand model.

Simulation



Average Profit when demand model is $\rho^{(2)}$ ($c_1 = 0.5, c_2 = 0.6$)



Average Profit when demand model is $\rho^{(1)}$ ($c_1 = 0.5, c_2 = 0.6$)

Conclusion

Oligopoly dynamic pricing

- Infinitely repeated game with private observations.
- Incomplete information: payoff based on the unknown demand model.
- The optimal collusion with a subset of sellers.

Demand Learning under Collusion (DCL):

- Subgame perfect Nash equilibrium.
- Pareto Efficient.
- Efficient online learning with bounded regret.