Dynamic Learning in Strategic Pricing Games

John R. Birge
Matthew Stern
The University of Chicago
Booth School of Business
How to Price and Learn?

• Situation
  – New products (e.g., games, insurance)
  – Unlimited capacity (e.g., digital goods)
  – Multiple (differentiated) suppliers

• Questions
  – What is optimal (equilibrium) pricing behavior and how is it different from the monopoly setting?
  – What is the impact on consumers, social welfare, and intermediaries?
Outline

• Monopoly pricing
• Collusion, competition, and oligopoly results
• Results from simple policies
• Regret due to learning and influence
• Strategic incomplete learning in simple policies
• “Willful ignorance” in prisoner’s dilemma
• Implications
Monopoly Pricing Results

• Monopolist needs to vary price sufficiently to learn demand

• Monopolist can learn a demand function with:
  \[ \text{Regret}(T, \theta, \sigma) = O(\sqrt{T \log T}) \]

where \( \text{Regret}(T, \theta, \sigma) = T \pi^* - E[\sum_{t=1}^{T} \pi_t(\theta, \sigma)] \)

\( \theta \) are parameters, \( \sigma \) is policy, and \( \pi_t \) is revenue with maximum expected value \( \pi^* \).

(Keskin and Zeevi (2013))
Necessity for Price Variation

• Myopically optimizing prices and estimating demand can lead to incomplete learning and poor prices

• Example (Lai and Robbins 1982):

\[ q_0, q_1, q_2, \ldots \]

True demand

\[ 0 = p_0, U = p_1 = p_2 = \ldots \]

Estimated demand
Questions for Competition

• Can sellers in competition learn as efficiently as monopolists?

• Is efficient learning always an optimal strategy?

• How much can optimal equilibrium strategies with competition differ from optimal monopolistic pricing strategies?
Relationship to Pricing Collusion

Results

• Stigler (1964) – can maintain collusive prices with mechanism to detect secret price-cutting
• Maskin/Tirole (1988) – can maintain collusion with kinked demand curves
• Cooper et al. (2013) – may obtain collusion by assuming monopoly model
• Quantity/Prisoner’s Dilemma (Kreps et al. (1982), Green/Porter (1984), Abreu/Pierce/Stacchetti (1986), Sannikov/Skrzypacz (2007), Rahmann (2015))
Setup for Investigation

• Model:
  
  - Two firms face noisy linear demands over $T$ periods
    
    $$d_{it} = \alpha_i + \beta_i p_{it} + \gamma_i p_{-it} + \epsilon_{it}$$
  
  - Estimate the parameters via OLS: $\hat{\theta}_{it} = [a_{it}, b_{it}, c_{it}]$
  
  - Forecast the price of their competitor: $p_{-it}^e$
  
  - Choose prices: $p_{it} \in [l, u], 0 < l < u < \infty$

• Simple candidate policy: best-response plus noise
Myopic (No Noise) Result

*Best Response:*

$$BR_{\theta}(p_{-it}^e) = \frac{a_{it}}{-2b_{it}} + \frac{c_{it}}{-2b_{it}} p_{-it}^e$$

**Competitive Equilibrium**

*Figure: Myopic strategies fail to learn demand*
Classifying Learning: Notions of Regret

Regret

\[ \Delta_i(T) = \sum_{t=1}^{T} \mathbb{E}[p_i^{NE}(\alpha_i + \beta_ip_i^{NE} + \gamma_ip_{-i}^{NE}) - p_{it}(\alpha_i + \beta_ip_{it} + \gamma_ip_{-it})] \]

Learning:

\[ \sum_{t=1}^{T} \mathbb{E} \left[ (p_i^{NE} - p_{it})^2 \right] \]

Influence:

\[ \sum_{t=1}^{T} \mathbb{E} \left[ p_{it} \left( \mathcal{BR}_\theta(p_{-i}^{NE}) - \mathcal{BR}_\theta(p_{-it}) \right) \right] \]
Measure Learning

Let \( x_t := [1 \ p_{it} \ p_{-it}]^T \) is the public price vector.

- **Fisher Information Matrix:**

\[
\mathcal{J}_t := \sum_{s=1}^{t} x_s x_s^T = \sum_{s=1}^{t} \begin{bmatrix} 1 & p_{is} & p_{-is} \\ p_{is} & p_{is}^2 & p_{is} p_{-is} \\ p_{-is} & p_{is} p_{-is} & p_{-is}^2 \end{bmatrix}
\]

- **Firm \( i \)'s OLS estimates are**

\[
\hat{\theta}_{it} = \mathcal{J}_t^{-1} \left( \sum_{s=1}^{t} d_{is} x_s \right)
\]

- **Estimation Error:**

\[
\hat{\theta}_{it} - \theta_i = \mathcal{J}_t^{-1} \sum_{s=1}^{t} \epsilon_{is} x_t
\]
Information and Error Bounds

Lemma (Keskin and Zeevi 2014)

If the demand shocks \( \{\epsilon_{is}\}_{i=1,s=1,\ldots,t} \) have light-tailed distributions, then there exist finite positive constants \( \rho \) and \( k \) such that,

\[
P \left\{ \|\hat{\theta}_{it} - \theta_i\| > \delta, \lambda_{\text{min}}(J_t) \geq m \right\} \leq kt \exp \left( -\rho(\delta \wedge \delta^2)m \right),
\]

for all \( \delta, m > 0 \) and all \( t \geq 3 \).

- Since \( J_t \) is public, firms know (with high probability) how close their estimates are to the true value.

- **Question:** How can we grow the minimum eigenvalue of \( J_t \)?
Adding Noise for Learning

Best Response with Random Dithering:

\[ \rho_{it} = \mathcal{BR}_{\theta}(p_{-it}^e) + \nu_{it} \]

\[ \nu_{it} \] conditionally independent, mean zero noise.
Learning Result

Theorem

Under best-response pricing with random dithering, the worst-case regret due to learning is $O(\sqrt{T \log T})$ for both players under the following assumptions:

- Cournot Adjustment: $p_{-it}^e = p_{-i(t-1)}$
- $\text{Var}(\nu_{it}) = O(\frac{1}{\sqrt{t}})$

• Implication and question:
  Can achieve the monopolist’s learning rate with competitors.

But, is this optimal?
Is Learning Optimal?

• With limited noise, may improve outcomes:

Static Nash equilibrium

Higher revenues
Is Incomplete Learning Optimal for General Strategies?

• Simplified Case with Two Players:

Consider a Prisoner’s Dilemma game with deterministic payoffs.

\[
\begin{array}{cc}
p_0 & p_1 \\
p_0 & (R_k, R_k) & (T_k, S_k) \\
p_1 & (S_k, T_k) & (P_k, P_k)
\end{array}
\]

\[T_k > R_k > P_k > S_k, \quad k = 1, 2\]

• Common prior:

\[
\pi := Pr(\text{Scenario} = 1)
\]

• Uncertainty between cooperation or competition:

\[R_1 = P_2\]
Game Setup

\[ \begin{array}{ccc}
  p_0 & p_1 \\
  p_0 & (R_k, R_k) & (T_k, S_k) \\
  p_1 & (S_k, T_k) & (P_k, P_k) \\
\end{array} \]

\[ T_k > R_k > P_k > S_k, \quad k = 1, 2 \]

- Represents pricing game where \( p_0 \) and \( p_1 \) are the competitive and cooperative equilibrium prices.
- Assume that the firms choose \((p_1, p_1)\) in the first period and observe \( R_1 = P_2 \) payoffs.
- Choice: Continue charging \((p_1, p_1)\) or deviate and learn the game.
Game Reduction

Result:
The $T$ period game can be reduced to a single period game by incorporating the future value into the current beliefs.

**Expected Payoff Parameters:**
- $x_\pi$, the expected payoff to firm $i$ with beliefs $\pi$ of $(p_1, p_1)$
- $y_\pi$, the expected payoff to firm $i$ with beliefs $\pi$ of $(p_1, p_0)$.

\[
x_\pi := (1 - \pi)(P_2 - S_2) + \pi(R_1 - T_1)
\]
\[
y_\pi := (1 - \pi)(T_2 - R_2) + \pi(S_1 - P_1)
\]
Stable Dynamics of the Game

Expected Payoff Parameters:

- $x_\pi$, the expected payoff to firm $i$ with beliefs $\pi$ of $(p_1, p_1)$
- $y_\pi$, the expected payoff to firm $i$ with beliefs $\pi$ of $(p_1, p_0)$
Alternate Stable Dynamics

*Expected Payoff Parameters:*

- $x_\pi$, the expected payoff to firm $i$ with beliefs $\pi$ of $(p_1, p_1)$
- $y_\pi$, the expected payoff to firm $i$ with beliefs $\pi$ of $(p_1, p_0)$.
Result of Simple Game

• Willful ignorance is optimal (cooperative equilibrium) if the firms observe the uninformative revenue outcomes
• Firms choose not to experiment and learn the full market information
• Extends to cases with uncertainty in the outcomes
Implications for Dynamic Pricing

• Firms may prefer not to experiment with prices, particularly for new products where products may be complements or substitutes

• Examples in digital goods, e.g., video games

• Uncertain makes observation and enforcement difficult but can also ensure collaboration
Policies for Pricing with Competition

• Use initial priors for optimal prices
• Base decisions to vary prices and learn on support for alternatives
• Issues for government policy:
  Caution for regulations restricting pricing adjustments (e.g., auto insurance in Maryland)
Conclusions

• Simple best-response policies with noisy adjustments can yield efficient learning as in monopolistic pricing

• Learning policies are not be optimal within the class of best-response and adjust policies

• In general policies, learning may not be optimal for firms

• Policies restricting price adjustments can reduce consumer value
Additional Issues

- How extensive is the set of situations in which firms may choose not to learn the market?
- What should happen in markets with nonstationary behavior?
- What role do intermediaries play in ensuring learning and efficiency?
- How should policy be used to ensure efficiency and maximum social welfare?
Thank you!

Questions?