Oligopoly Dynamic Pricing: A Repeated Game with Incomplete Information

Yixuan Zhai† and Qing Zhao§

†Electrical and Computer Engineering, UC Davis, Davis, CA, 95616, yxzhai@ucdavis.edu
§School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, 14853, qz16@cornell.edu

***Abstract*—We consider an oligopoly dynamic pricing problem where the demand model is unknown and the sellers have different marginal costs. We formulate the problem as a repeated game with incomplete information. We develop a dynamic pricing strategy that leads to a Pareto-efficient and subgame-perfect equilibrium and offers a bounded regret over an infinite horizon, where regret is defined as the expected cumulative profit loss as compared to the ideal scenario with a known demand model. The resulting equilibrium also reveals a spontaneous collusion among a subset of sellers due to the difference in marginal costs among the sellers.**

***Index Terms*—Dynamic pricing, repeated game, subgame-perfect Nash equilibrium, regret.**

I. Introduction

A. Oligopoly Dynamic Pricing

We consider an oligopoly in which multiple sellers offer a product to a stream of customers arriving sequentially in time. The demand model, characterizing the probability that a customer purchases the product at a given price, is unknown but assumed to take one of M possible forms. At each time, the sellers set their price for the product simultaneously, and the seller(s) offering the lowest price may make a successful sale with a probability determined by the demand model. The objective of each seller is to maximize its long-term profit over an infinite horizon through a dynamic pricing strategy.

Each seller faces the classic tradeoff between learning and earning present in all online learning problems. What complicates the problem at hand is that each seller's selling history is private information. As a result, each seller is learning the demand model from disjoint sets of random observations, and the sellers compete for not only profit but also information for learning the demand model. What further complicates the problem is the heterogeneous marginal costs among the sellers, which may lead to collusion among a subset of sellers.

We formulate the above oligopoly dynamic pricing problem as an infinitely repeated game with incomplete information, where the incomplete information is referring to the uncertainty in the payoff function due to the unknown demand model. We develop a dynamic pricing strategy which leads to a subgame-perfect Nash equilibrium. Furthermore, the developed pricing strategy is Pareto-efficient and offers a bounded regret over an infinite horizon, where regret is defined as the expected cumulative profit loss as compared to the ideal scenario with a known demand model. The resulting equilibrium also reveals a spontaneous collusion among a subset of sellers due to the difference in marginal costs.

[0]This work was supported by the Army Research Office under Grant W911NF-12-1-0271 and by the National Science Foundation under Grant 1549989.

B. Related Work

The problem studied in this paper falls into the category of Bertrand competition in an oligopoly first formulated by Bertrand in 1883 [1]. In Bertrand's original study, a static game was considered, and the demand model was assumed known. Various repeated games for Bertrand competition under known demand models have been formulated and studied in the literature. See [2–6] and references therein.

Dynamic pricing under demand uncertainty has been studied under both monopoly [7–10] and oligopoly [11–13]. In particular, Rotemberg and Saloner [11] considered the pricing problem when the demand fluctuates in different business cycles. Assuming that all sellers have the same marginal cost and demand fluctuations are i.i.d. following a known distribution in each period, Rotemberg and Saloner characterized collusive pricing in an infinitely repeated game. In [12] Bertsimas and Perakis studied parameterized demand learning in oligopoly where the unknown demand function is a known parametric function with unknown parameter values. The objective of [12] is also different from this paper. Rather than formulating a repeated game and explicitly studying equilibria, Bertsimas and Perakis took an optimization approach based on a joint estimate of each firm's demand. In our prior work [13], we considered oligopoly dynamic pricing under demand uncertainty, but assumed that all sellers have the same marginal cost. The symmetry in marginal cost leads to fair competition among all sellers as shown in our prior work. However, the heterogeneous marginal cost results in collusion among a subset of sellers as demonstrated in this paper.

II. Preliminaries

In this section, we briefly review basic concepts of infinitely repeated games with incomplete information. In particular, the payoff functions are determined by an unknown state of nature.

Consider a repeated game played by N players over an infinite horizon. Before the game starts, a state of nature $\omega \in \Omega$ is chosen, which is unknown to the players, and fixed over the infinite horizon. Each player i has a set of actions $\mathcal{A}_i$ and a set

978-1-4799-9988-0/16/$31.00 ©2016 IEEE

of private signals $\mathcal{O}_i$. At each time t, under the action profile $a = (a_i)_{i=1}^N$ of all players, player i receives a private signal $o_i^{(\omega)}(a) \in \mathcal{O}_i$ and a payoff $u_i^{(\omega)}(a)$, both depending on the unknown state ω.

Let $\mathcal{A}^t$ be the t-fold Cartesian product of $(\mathcal{A}_i)_{i=1}^N$, i.e., an element of $\mathcal{A}^t$ is a list of action profiles specifying the actions played by all players from time 0 to t. Similarly, let $\mathcal{O}_i^t$ denote the t-fold Cartesian product of $\mathcal{O}_i$. Let $\mathcal{H}_i^t = \mathcal{A}^t \times \mathcal{O}_i^t$ denote the history space of player i at time t. A strategy σ_i of player i is a sequence of mappings, one for each t, that specifies the action of player i at time t under each history $h_i^t \in \mathcal{H}_i^t$. We define the payoff of a strategy profile $\sigma = (\sigma_i)_{i=1}^N$ using the limit of means criterion

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

where $a^t(\sigma)$ is the action profile induced by σ at time t.

Let $\sigma_{-i} = (\sigma_j)_{j\neq i}$. A strategy profile σ is a Nash equilibrium if for all i and ω, we have

$$U_i^{(\omega)}(\sigma) \geq U_i^{(\omega)}(\sigma_i', \sigma_{-i}), \tag{2}$$

where σ_i' is any strategy for player i. For a repeated game, there exist Nash equilibria that violate the notion of optimality at out-of-equilibrium information sets. Subgame perfection enforces sequential rationality which requires optimal behavior in both in-equilibrium and out-of-equilibrium information sets [14].

Definition 1. A strategy profile σ is a subgame-perfect equilibrium of a repeated game if for any history $h^t = \cup_i h_i^t$, the induced continuation strategy $\sigma|_{h^t}$ is a Nash equilibrium of the continuation game that starts at t following history h^t.

Definition 2. A Nash equilibrium σ is Pareto-efficient if no strategy can improve the payoff of one player without decreasing the payoff of at least one other player.

III. Problem Formulation

An oligopoly is a market where a small number of sellers (referred to as oligopolists) dominate the market. Comparing to a perfectly competitive market, oligopolists are price setters rather than price takers. Each seller's objective is to maximize its own long run profit (revenue minus cost) through a pricing strategy.

Consider an oligopoly with N sellers offering a certain product to a stream of customers. The marginal cost for seller i to produce one unit of the product is c_i. The marginal cost of each seller is common knowledge to all sellers. Without loss of generality, we assume $c_1 < c_2 < \ldots < c_N$. At each given time t, seller i $(i = 1, \ldots, N)$ proposes a price $p_i(t)$ from $\mathcal{A}_i = [c_i, p_u]$ (note that a price below the seller's marginal cost c_i is not an option and p_u is the maximal possible price). The customer purchases from the seller offering the lowest price $\underline{p} = \min_i p_i(t)$ with probability $\rho(\underline{p})$ which is referred to as the demand model. The demand model is unknown and takes one of the M possible forms $\{\rho^{(\omega)}(p)\}_{\omega=1}^M$. If the lowest price is offered by multiple sellers, the customer randomly chooses one of them with equal probability if he decides to make the purchase. The objective of each seller is a dynamic pricing strategy that maximizes its profit over an infinite horizon.

The above problem can be formulated as a repeated game with incomplete information as follows. The unknown demand model $\rho^{(\omega)}(p)$ is determined by the state of nature. The set of private signals (i.e., observations) for seller i is given by $\mathcal{O}_i = \{0, 1\}$ where 1 indicates a successful sale and 0 an unsuccessful one. At each time t, only those n $(n \geq 1)$ sellers offering the lowest price $\underline{p}$ receive a private observation which is a realization of a Bernoulli random variable with mean $\rho^{(\omega)}(\underline{p})/n$. The one-shot payoff for each of these n sellers is given by the expected profit $u_i^{(\omega)}(\underline{p}) = (\underline{p} - c_i)\rho^{(\omega)}(\underline{p})/n$.

We assume that each possible demand model $\rho^{(\omega)}(p)$ is continuously differentiable and strictly decreasing over $[c_1, p_u]$. Define

$$r^{(\omega)}(p, c) = (p - c)\rho^{(\omega)}(p), \tag{3}$$

which is the expected profit for a seller with marginal cost c and offering price p under monopoly. Let

$$p^{(\omega)}(c) = \arg\max_{p\in[c,p_u]} r^{(\omega)}(p, c) \tag{4}$$

denote the profit-maximizing price under demand model $\rho^{(\omega)}(p)$ and marginal cost c in monopoly. We assume that $p^{(\omega)}(c)$ is unique and $r^{(\omega)}(p, c)$ is continuous and strictly increasing with p over $[c, p^{(\omega)}(c)]$. This assumption imposes only a mild condition on the demand model and follows from the IGFR (increasing generalized failure rate) assumption commonly adopted on demand models [15].

IV. Oligopoly Dynamic Pricing under Known Demand Model

We first consider the case when the demand model is known. The results obtained here will be used in subsequent sections. When the demand model $\rho^{(\omega)}(p)$ is known, the dynamic pricing problem is a repeated Bertrand game and a subgame-perfect trigger strategy with pure actions exists. As shown in the theorem below, a subgame-perfect grim trigger strategy with pure actions is given by a collusion of a subset of K sellers with the lowest marginal cost $c_1, \ldots, c_K$. The number K of colluding sellers is determined by the marginal costs $\{c_i\}_{i=1}^N$ and the demand model $\rho^{(\omega)}(p)$ as specified below.

In this colluding strategy (referred to as σ_C), seller 1, the one with the lowest marginal cost, sets the price and forms the optimal collusion to maximize its own profit. Specifically, if seller 1 decides to collude with seller 2 to seller k, the profit-maximizing price $\hat{p}_k^{(\omega)}$ is given by

$$\hat{p}_k^{(\omega)} = \arg\max_{p\in(c_k,c_{k+1}]} \frac{1}{k} r^{(\omega)}(p, c_1). \tag{5}$$

Since $r^{(\omega)}(p, c_1)$ is continuous and strictly increasing with p, we have

$$\hat{p}_k^{(\omega)} = \begin{cases} c_{k+1} & k < N \\ p^{(\omega)}(c_1) & k = N, \end{cases} \tag{6}$$

where without loss of generality, we have assumed that $p^{(\omega)}(c_1) > c_N$ and set $c_{N+1} = p^{(\omega)}(c_1)$ for notation convenience. The optimal collusion K for seller 1 is thus given by

$$K = \arg \max_{k=1,\ldots,N} \frac{1}{k} r^{(\omega)}(\hat{p}_k^{(\omega)}, c_1). \tag{7}$$

For the ease of notation, let $\hat{p}^{(\omega)}$ denote the profit-maximizing colluding price (i.e., the subscript K is omitted). Thus in σ_C, K sellers with marginal cost $c_i < \hat{p}^{(\omega)}$ collusively offers price $\hat{p}^{(\omega)}$ and any deviations will trigger a everlasting punishment. Specifically, seller $j > 1$ will punish deviations by offering its marginal cost c_j. Seller 1 will punish deviations by offering price $p = c_2 - \epsilon$ for an arbitrarily small positive ϵ.

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

V. Oligopoly Dynamic Pricing under Unknown Demand Model

In this section, we develop an oligopoly dynamic pricing strategy under an unknown demand model to approach the performance of σ_C given in Sec. IV. The learning efficiency of a dynamic pricing strategy σ is measured by regret R_σ, which is defined as the accumulated total profit loss as compared to σ_C under a known demand model.

$$R_\sigma = \sum_{i=1}^{N} \sum_{t=1}^{\infty} \left(u_i^{(\omega)}(a^t(\sigma_C) - u_i^{(\omega)}(a^t(\sigma)) \right). \tag{8}$$

Without loss of generality, we propose an oligopoly dynamic pricing strategy under the assumption that $\hat{p}^{(1)} < \hat{p}^{(2)} < \cdots < \hat{p}^{(M)}$ and $\rho^{(i)}(\hat{p}^{(\omega)}) \neq \rho^{(j)}(\hat{p}^{(\omega)})$ for all $i \neq j$ and all ω. Referred to as Demand Learning under Collusion (DLC), this dynamic pricing strategy partitions the time horizon into fixed-length epochs with epoch length $l \geq 2$ (see Fig. 1). Each epoch starts with a declaration time slot followed by $l-1$ cooperation time slots. In the declaration slot of epoch t, seller 1 first carries out a maximum likelihood estimate $\hat{\omega}(t)$ of the underlying demand model based on its private sale history in the cooperation time slots of the previous $t-1$ epochs ($\hat{\omega}(1)$ can be set to a default value, say 1). Seller 1 then offers the profit-maximizing colluding price $\hat{p}^{(\hat{\omega}(t))}$ based on the estimated demand model in this declaration slot. All other sellers offer the same price ($\hat{p}^{(1)}$ if in the first epoch) they offered in the cooperation slots in the previous epoch. In all the cooperation time slots of epoch t, all sellers with marginal costs lower than $\hat{p}^{\hat{\omega}(t)}$ collusively offer price $\hat{p}^{\hat{\omega}(t)}$. The trigger strategy for punishing any deviation is the same as that in σ_C.

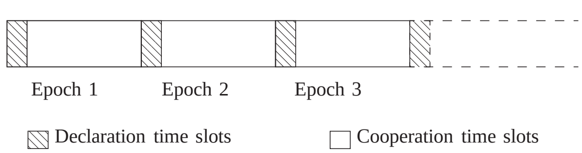

Fig. 1. Demand Learning under Collusion

In DLC, the declaration time slots work like public signals which are used by seller 1 to coordinate with other sellers. The cooperation time slots enable learning by the property that all participating sellers offer the same price. The seller 1's actions are restricted to the M profit-maximizing colluding prices $\{\hat{p}^{(\omega)}\}_{\omega=1}^{M}$. And the actions in the cooperation time are determined by the actions from the declaration time in the same epoch.

In the following theorem, we summarize the properties of the DLC dynamic pricing strategy.

Theorem 2. Properties of DLC:

- 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

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

The proof is omitted due to space limit.

VI. Numerical Example

In this section, we present simulation examples to study the learning efficiency of DLC. The simulation shows the average profit for the dynamic pricing problem with $N = 2$ sellers and $M = 2$ demand models. The demand models are $\rho^{(1)}(p) = 1.3 - 0.6p$ and $\rho^{(2)}(p) = 1.1 - 0.4p$. The maximum price p_u is set to 2. The epoch length in DLC is set to $l = 6$.

Fig. 2 - Fig. 4 clearly demonstrate the convergence of the average profit under DLC toward that under known demand models. While the underlying demand model is the same in Fig. 3 and Fig. 4, seller 1 chooses to share the market with seller 2 by offering price $p^{(1)}(c_1) = 1.33$ in Fig. 3 and to dominate the market by offering price $\hat{p}^{(1)} = 1$. This is due to the change in the marginal cost of seller 2.

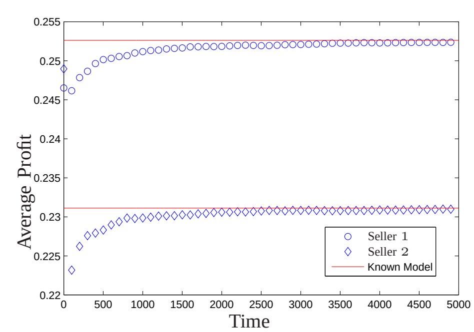

Fig. 2. Average Profit when demand model is $\rho^{(2)}$ ($c_1 = 0.5$, $c_2 = 0.6$)

VII. Conclusion

In this paper, we studied a oligopoly dynamic pricing under an unknown demand model and private observations. We developed a dynamic pricing strategy that was shown to be subgame-perfect and Pareto-efficient. It was further shown that

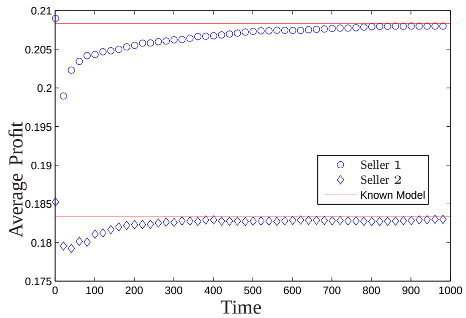

Fig. 3. Average Profit when demand model is $\rho^{(1)}$ ($c_1 = 0.5$, $c_2 = 0.6$)

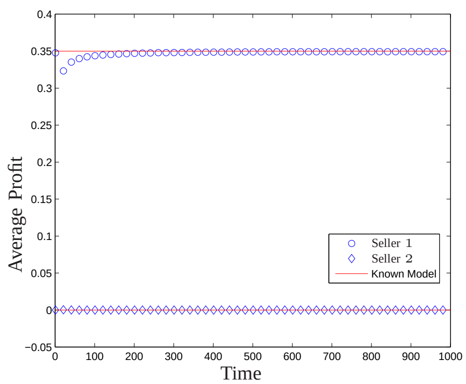

Fig. 4. Average Profit when demand model is $\rho^{(1)}$ ($c_1 = 0.5$, $c_2 = 1$)

the proposed strategy offers bounded regret over an infinite horizon. It is thus order optimal in terms of the efficiency of learning the underlying demand model.

REFERENCES

[1] J. Bertrand, "Theorie mathematique de la richesse sociale," *Journal des Savants*, pp. 499–508, 1883.

[2] M. Osborne, *An introduction to game theory*. Oxford Univ. Press, 2004.

[3] D. Niyato and E. Hossain, "Competitive pricing for spectrum sharing in cognitive radio networks: Dynamic game, inefficiency of nash equilibrium, and collusion," *IEEE Journal on Selected Areas in Communications*, vol. 26, pp. 192–202, Jan 2008.

[4] Z. Ji, W. Yu, and K. Liu, "A game theoretical framework for dynamic pricing-based routing in self-organized manets," *IEEE Journal on Selected Areas in Communications*, vol. 26, pp. 1204–1217, September 2008.

[5] S. L. Puller, "Pricing and firm conduct in california's deregulated electricity market," *The Review of Economics and Statistics*, vol. 89, pp. 75–87, February 2007.

[6] S. Borenstein and A. Shepard, "Dynamic pricing in retail gasoline markets," *RAND Journal of Economics*, vol. 27, pp. 429–451, 1996.

[7] M. Rothschild, "A two-armed bandit theory of market pricing," *Journal of Economic Theory*, vol. 9, pp. 185–202, October 1974.

[8] O. Besbes and A. Zeevi, "Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms," *Operations Research*, vol. 57, no. 6, pp. pp. 1407–1420, 2009.

[9] J. M. Harrison, N. B. Keskin, and A. Zeevi, "Bayesian dynamic pricing policies: Learning and earning under a binary prior distribution," *Management Science*, pp. 1–17, October 2011.

[10] Y. Zhai, P. Tehrani, L. Li, J. Zhao, and Q. Zhao, "Dynamic pricing under binary demand uncertainty: A multi-armed bandit with correlated arms," in *Proc. of the 45th IEEE Asilomar Conference on Signals, Systems, and Computers*, November 2011.

[11] J. Rotemberg and G. Saloner, "A supergame-theoretic model of business cycles and price wars during booms," *AmerPean Eeonomic Review*, vol. 76, pp. 390–407, 1986.

[12] D. Bertsimas and G. Perakis, "Dynamic pricing: A learning approach," in *Mathematical and Computational Models for Congestion Charging* (S. Lawphongpanich, D. Hearn, and M. Smith, eds.), Applied Optimization, pp. 45–79, Springer US, 2006.

[13] Y. Zhai and Q. Zhao, "Competitive dynamic pricing under demand uncertainty," in *Proc. of the 45th IEEE Asilomar Conference on Signals, Systems, and Computers*, pp. 1098–1101, Nov 2014.

[14] G. J. Mailath and L. Samuelson, *Repeated Games and Reputations: Long-Run Relationships*. Oxford Univ. Press, 2006.

[15] M. A. Lariviere, "A note on probability distributions with increasing generalized failure rates," *Operations Research*, vol. 54, pp. 602–604, May-June 2006.