Competitive Demand Learning: a Data-Driven Pricing Algorithm

Yongge Yang[1], Yu-Ching Lee[1⋆], and Po-An Chen[2⋆⋆]

[1] Department of Industrial Engineering and Engineering Management, National Tsing Hua Uiversity, Taiwan.
562514124@qq.com, yclee@ie.nthu.edu.tw
[2] Institute of Information Management, National Chiao Tung University, Taiwan.
poanchen@nctu.edu.tw

Abstract. Dynamic pricing is used to maximize the revenue of a firm over a finite-period planning horizon, given that the firm may not know the underlying demand curve *a priori*. In emerging markets, in particular, firms constantly adjust pricing strategies to collect adequate demand information, which is a process known as price experimentation. To date, few papers have investigated the pricing decision process in a competitive environment with unknown demand curves, conditions that make analysis more complex. Asynchronous price updating can render the demand information gathered by price experimentation less informative or inaccurate, as it is nearly impossible for firms to remain informed about the latest prices set by competitors. Hence, firms may set prices using available, yet out-of-date, price information of competitors.

In this paper, we design an algorithm to facilitate synchronized dynamic pricing, in which competitive firms estimate their demand functions based on observations and adjust their pricing strategies in a prescribed manner. The process is called "learning and earning" elsewhere in the literature. The goal is for the pricing decisions, determined by estimated demand functions, to converge to underlying equilibrium decisions. The main question that we answer is whether such a mechanism of periodically synchronized price updates is optimal for all firms. Furthermore, we ask whether prices converge to a stable state and how much regret firms incur by employing such a data-driven pricing algorithm.

1 Introduction

Dynamic pricing is used to maximize a firm's revenue over a finite-period planning horizon, given that the firm may not know the underlying demand curve *a priori*. In emerging markets, in particular, firms constantly adjust their pricing strategies to collect adequate demand information, which is a process termed *price experimentation* by [5], [13]. Currently, few papers exist that consider competitive environments with unknown demand curves. Such conditions add greater complexity to the analysis of the decision process. Evaluating the influence of decisions made by competitors is challenging. If a firm is sensitive to competitor prices, one might expect a change in the prices set by a firm to provoke an immediate response from competitors. Such asynchronous price updating renders the demand information gathered by price experimentation neither informative nor accurate, as it is almost impossible for each firm to remain informed about the current prices set by competitors. Although firms may still observe the market response of consumers given competitor pricing, the use of out-of-date price information is possible. A mechanism of *synchronized dynamic pricing* thus arose. Such a mechanism ensures that the pricing strategy of each firm is adjusted in a prescribed way to jointly collect demand information and make pricing decisions.

Consider a total of N firms selling an identical product in an oligopolistic market, in which the true underlying demand curve and the presence of demand shocks are unknown. Over a time horizon of T periods, firms make pricing decisions in each period $t = 1, \cdots, T$. Firm i's decision at period t is denoted by p_t^i, and the demand for firm i's product in period t is denoted by $D_t^i(\mathbf{p}_t) = \lambda^i(\mathbf{p}_t) + \varepsilon_t^i$, where $\mathbf{p}_t \equiv [p_t^1, \cdots, p_t^N]$ represents the pricing decisions of all firms. The price set by a firm is the only information that its competitors can observe. Hence, firms are unable to estimate the demand of competitors. Without loss of generality, we assume that the mean value of D_t^i is given by the average rate λ^i, conditional on the price vector of all

⋆ Corresponding, supported by MOST 107-2221-E-007-074-MY3.
⋆⋆ Co-Corresponding, supported in part by MOST 108-2221-E-009-053-MY2.

companies, while ε_t^i measures the demand shock in period t. Note that $\lambda^i(\mathbf{p})$ reflects the fact that firm i acknowledges that the price decisions made by other firms will influence the demand for the product of firm i. By focusing on competition among the N firms, we do not consider capacity limitation, production cost or marginal cost. Each firm is assumed to be selfish and reacts immediately to price changes made by competitors. The goal of each firm is to sequentially set a price to maximize revenue under demand uncertainty and competition.

Finding the equilibrium pricing strategy is a challenge that each firm must tackle. To do so requires that firms constantly adjust prices according to the best response pricing strategy. In particular, when the demand function is unknown, determining a suitable function for demand estimation will directly affect the pricing decision. Such an estimation should consider the behaviour of competitors and the market response. Fortunately, there is evidence that linear functions serve as good approximations of underlying demand functions and allow pricing decisions to converge to an optimum in [3]. The present study designs an algorithm of synchronized dynamic pricing in which firms in competition estimate their demand functions based on observations and adjust their pricing strategies in a prescribed manner to maximize their revenues. The process is termed *learning and earning* elsewhere in the literature. The goal of the process is for the pricing decisions, determined by an estimated demand function, to converge to an underlying equilibrium decision. The main question that we wish to answer is whether such a mechanism of periodically synchronised price updating is optimal for all firms. Specifically, we ask whether the mechanism may allow prices to reach a stable state and how much regret firms incur by employing such a data-driven pricing algorithm.

1.1 Our Results

This paper generalizes the work of [3], who constructed a dynamic pricing algorithm in a monopoly setting in which a single firm chooses a price to maximize expected revenue without knowledge of the true underlying demand curve. Although consideration of a competitive environment is complicated, we are still able to obtain notable results, and, if the market is a monopoly, these findings are consistent with the results of [3].

We propose a data-driven equilibrium pricing (DDEP) algorithm to solve the dynamic pricing decisions of each firm in competition. As the true demand curve is unknown, each firm must estimate a demand curve via a linear approximation scheme in which prices are varied to develop better pricing strategies. Throughout the paper, we use period t or, equivalently, time t. Let p_t^{i*} denote the equilibrium price of firm i at time t, which is obtained by the estimated demand curve of firm i at time t, and let p_t^{-i} denote the prices of other competitors. A *clairvoyant* model implies that a firm has knowledge of the underlying demand curve and the distribution of demand shocks. The goal of learning is to make p_t^i converge to the clairvoyant equilibrium price of firm i, p^{i*}, as t grows large. Note that a learning scheme in which the difference between p_t^i and p^{i*} will eventually converge to zero is called *complete* learning; otherwise, it is termed *incomplete* learning. In a competitive environment, it is not possible for a firm to achieve complete learning alone due to the influence of strategies employed by competitors. Our synchronised pricing algorithm ensures that all firms in the market jointly adjust their prices to achieve complete learning.

In an alternative scenario, in which some firms have knowledge of the demand function and the distribution of demand shocks, such firms may be unwilling to engage in price experimentation. Therefore, we propose a modified DDEP algorithm to account for this. In such cases, the approximate demand function does not estimate the effects of firms with known demand, an idea originating from work by [8]. However, our mechanism requires all that firms set prices simultaneously. As a result of the information asymmetry, the pricing of firms with unknown demand will not be affected by the firms with the knowledge of the demand curves; however, the pricing of the firms with the knowledge of the demand curves will be affected by firms with unknown demand.

The process of learning is often evaluated in terms of *regrets*. In our setting, the definition of regret is quite different from that which is commonly used in a monopoly setting in which a firm is playing against the market demand (without competitors); see [3], [6]. In our study, a firm's regret is described over a time interval given by the difference between the realized revenue and optimal expected revenue, given that other competitors stick to the pricing decisions determined by the pricing algorithm. We also analyze the revenue difference obtained by the algorithm from that obtained by the clairvoyant Nash equilibrium $\mathbf{p}^*$ per algorithm iteration, which are discussed in further detail in Section 2.3.

The main results of our work are as follows:

– The best response function derived by DDEP through the quadratic concave function will generate the sequence $\{\mathbf{p}_t\}$ that converges to $\mathbf{p}^*$ as t grows large. (Theorem 1)
– As time progresses, the accumulated revenue of each firm generated by the pricing policies of DDEP algorithm is asymptotically close to the clairvoyant accumulated revenue. (Theorem 3)
– The revenue difference converges to zero as time progresses and is related to the quantity of firms in competition. The difference implies that realised revenues are sometimes greater than those revenues obtained by the clairvoyant Nash equilibrium $\mathbf{p}^*$. However, we are unable to predict when this will occur. (Theorem 2)
– If some clairvoyant firms with knowledge of the demandprice relationship do not participate in price experimentation and those firms without knowledge of the demand functions use the *partially-flawed linear approximate demand functions* (according to the definition in [8]), then the sequence $\{\mathbf{p}_t\}$, generated by the modified DDEP may still converge to $\mathbf{p}*$. Furthermore, the convergence rate of regret among firms without knowledge of the demand curve is related to the quantity of such firms, while the convergence rate among firms with knowledge of the demand curve is not. (Theorem 5)

1.2 Related Work

To conclude this section, we briefly review the literature that is most related to our work.

– **Learning algorithms for pricing models in a monopoly market.** Dynamic pricing is a critical tool for revenue management. The problem of unknown demand functions has been widely explored in the literature. One common theme of the existing work focuses on the task firms face to arrive at optimal pricing decisions when the underlying demand curve is unknown and how to lower the growth rate of regret. [2] assumed that the demand function belongs to a known parametric family with unknown parameter values and develop a non-parametric pricing policy that achieves asymptotic optimality. The structure relies on price experimentation to infer accurate information about the demand function. Since price experimentation can be costly, some papers focus on a balance between the explorationexploitation trade-off. [5] used maximum-likelihood estimation to build the pricing policy and [4] proposed a policy of gradually reducing the taboo interval to achieve asymptotically optimal pricing. The two policies that they proposed can be regarded as iterated least squares. [13] provided a generally sufficient condition for iterated least squares methods which ensures that regret does not exceed the order of $\sqrt{T}$. A common feature of previous work is that the underlying demand model is assumed to be the same as the parametric demand function and is thus well-specified. [3] proposed a pricing policy constructed on the assumption that the underlying demand curve is non-linear but used a linear approximation to model the market response to the offered price. We refer readers to [6] and [7] which extend the model to coordinate pricing and an inventory control problem.
– **Dynamic pricing in a competitive environment.** Few papers consider dynamic pricing with competition, as this type of problem requires a game-theoretic approach to obtain a solution. Some studies consider non-cooperative competition in revenue management, where the products are differentiated, for example, at the service-level or based on certain attributes of the goods on which consumer choice closely depends. [15] presented a dynamic pricing model in which strategic consumers choose differentiated perishable goods. With firms able to benefit from a lack of complete information among consumers, they provide an analysis of the equilibrium price dynamic under different market settings. [10] also studied how consumer choice depends not only on price, but also on purchase time and product attributes; they showed that the shadow price solved by the deterministic problem can be used to construct an asymptotic equilibrium pricing policy. When different companies sell identical goods, the market response depends mainly on the price. [8] used a "flawed" demand function without incorporating competitor prices and showed that (a) the prices converge to the Nash equilibrium if the slope is known to the seller, (b) the

prices converge to the cooperative prices if the intercept is known to the seller, and (c) there are many potential limit prices that are neither Nash equilibrium nor cooperative prices if the parameters of the demand function are unknown.

- **Learning algorithms for pricing models in a competitive environment.** If the content of learning is incorporated into the pricing policy, companies are able to learn the response of consumers and competitors, continuously adjusting pricing to achieve optimality over time. [1] considered a myopic pricing policy in a competitive oligopolistic environment. In such a case, a given firm must estimate not only its own demand but also the demand and pricing of competitors. A method of mathematical programming with equilibrium constraints is applied to formulate the model to estimate the competitor parameters. [12] proposed further improvements that adjusted the most accurate parameters of demand function periodically based on the equilibrium demands. [11] considered a firm selling a single product with multiple versions and developed a multinomial logit choice demand model, commonly used in the context of non-cooperative competition between firms selling differentiated goods, each seeking to maximize total revenue. Bayesian updating is used to estimate the unknown arrival rate, and maximum likelihood estimation is used to update the core value. They showed that the unknown parameters can be estimated simultaneously while sales are changing. [14] used the observed sales data to estimate the parameters of demand by Kalman filtering, showing that a differential variational inequality can be used to model the non-cooperative competition. [9] conducted two field experiments with randomized prices. The first experiment estimated a consumer choice model in which observations of sales made by competitors are not required, and the second tested a best response pricing strategy.

2 Model and Preliminaries

We consider a periodical equilibrium pricing problem for N firms. In each period $t = 1, \cdots, T$, each firm needs to set prices p_t^i, chosen from a feasible and bounded policy set $\mathcal{P}^i = \left[p^{i,\ell}, p^{i,h}\right]$, $p^{i,\ell} < p^{i,h}$, $\forall i = 1, \cdots, N$. The prices set by firms affect the market response of all firms in the competition. Recall that $\mathbf{p} \equiv (p^i, p^{-i})$ denotes the vector of prices of all firms in the competition. The market response to the price p_t^i for firm i at time t (which is exactly the demand function) is given by $D_t^i(\mathbf{p}_t) = \lambda^i(\mathbf{p}_t) + \varepsilon_t^i$, $\forall i = 1, \cdots, N$, in which $\lambda^i(\mathbf{p}_t)$ is a deterministic twice differentiable function representing the mean demand, conditional on the price $\mathbf{p}_t$, and ε_t^i are zero-mean random variables, assumed to be independent and identically distributed. Hence, the demand curve $\lambda^i(\mathbf{p})$ of firm i not only depends on the price p^i, chosen by itself, but also on the prices of other firms p^{-i}, where $p^{-i} = \left\{p^1, \cdots, p^{i-1}, p^{i+1}, \cdots, p^N\right\}$. Let $\pi^i = (p_1^i, p_2^i, \cdots)$ denote the sequence of pricing policy of firm i and $\mathit{\Pi} = (\pi^1, \cdots, \pi^N)$ denote the admissible pricing policies of all firms. The revenue function r^i of firm i obtained from prices $\mathbf{p}$ is denoted by $r^i(\mathbf{p}) \equiv p^i \mathbb{E}[D^i(\mathbf{p})]$. Each firm seeks to maximize its revenue in a competitive environment. Since the feasible strategy set of firm i depends on the pricing strategies of competitors, the equilibrium prices of all firms are a generalised Nash equilibrium (GNE). We make the following assumptions:

Assumption 1 *For any $p^i \in \mathcal{P}^i$,*
(i) $\dfrac{\partial \lambda^i(\cdot, p^{-i})}{\partial p^i} < 0, \forall i = 1, \cdots, N,$
(ii) $\dfrac{\partial \lambda^i(p^i, p^{-i\backslash j}, \cdot)}{\partial p^j} > 0, \forall j \neq i, i = 1, \cdots, N,$
(iii) $\dfrac{\partial^2 r^i(\mathbf{p})}{\partial^2 \mathbf{p}}, \forall i = 1, \cdots, N$ *is a negative semi-definite matrix.*

Assumption (i) ensures that for every firm i, the underlying demand function $\lambda^i(\cdot, p^{-i})$ is strictly decreasing on p^i given the prices set by other firms, p^{-i}. Assumption (ii) dictates that $\lambda^i(p^i, p^{-i\backslash j}, \cdot)$ is strictly increasing on p^j with p^i and $p^{-i\backslash j}$ given, in which $p^{-i\backslash j}$ represents the vector constituted by all prices except p^i and p^j. Assumption (iii) dictates that the revenue function $r^i(\mathbf{p})$ is a concave function and there exists a unique maximizer for any feasible $\mathbf{p}$.

For every firm i, if the underlying demand curve $\lambda^i(\cdot)$ and the distribution of the random error ε^i are known *a priori*, the optimal pricing decision is obtained by solving the following revenue maximization problem:

$$\begin{aligned} &\text{maximize}_{p^i, D^i}\ r^i(\mathbf{p}) \equiv p^i \mathbb{E}\left[D^i(\mathbf{p})\right] \\ &\text{subject to} \quad \mathbb{E}\left[D^i(\mathbf{p})\right] \geq 0, p^i \in \mathcal{P}^i, \forall i = 1, \cdots, N \end{aligned} \tag{1}$$

A vector $\mathbf{p} \equiv (p^i, p^{-i})$ is feasible for the GNE problem if it satisfies the constraint $\lambda^i(\mathbf{p}) \geq 0$ for all firms $i = 1, \cdots, N$. A feasible vector $\mathbf{p}^*$ is a solution to the problem of GNE if, for all firms $i = 1, \cdots, N$, we have

$$r^i(p^{i*}, p^{-i*}) \geq r^i(p^i, p^{-i*}) \quad \forall p^i : \lambda^i(p^i, p^{-i*}) \geq 0$$

It is well known that if there exists a vector $\mathbf{p}^*$ and multipliers that satisfy the following KKT system for each firm i:

$$\begin{aligned} \left[\lambda^i(\mathbf{p}) + p^i \nabla_{p^i} \lambda^i(\mathbf{p}) + \mathbb{E}(\varepsilon^i)\right] + \mu^{i,1}\left[\nabla_{p^i}\lambda^i(\mathbf{p})\right] - \mu^{i,2} + \mu^{i,3} = 0 \ \forall i \\ \mu^{i,1} \geq 0, \mu^{i,1} \cdot \left[\lambda^i(\mathbf{p}) + \mathbb{E}(\varepsilon^i)\right] = 0, \lambda^i(\mathbf{p}) + \mathbb{E}(\varepsilon^i) \geq 0, \qquad \forall i \\ \mu^{i,2} \geq 0, \mu^{i,2} \cdot (p^{i,h} - p^i) = 0, p^{i,h} - p^i \geq 0 \qquad \forall i \\ \mu^{i,3} \geq 0, \mu^{i,3} \cdot (p^i - p^{i,l}) = 0, p^i - p^{i,l} \geq 0, \qquad \forall i, \end{aligned} \tag{2}$$

then the vector $\mathbf{p}^*$ is a GNE.

Due to the lack of information about the underlying demand curves, firms must collect enough observations to estimate the demand functions. We assume that only the offered prices $\mathbf{p}_t$ are observed by all competitors at time t and that every firm is aware of past sales data. To mitigate the negative effects arising from not knowing the demand function, every firm is motivated to learn of it. Each firm completes the process of learning on its own. The mechanism controls the learning process of all firms and guarantees that prices $\mathbf{p}$ converge to the clairvoyant GNE $\mathbf{p}^*$, which are the solutions to the above system in the case of clairvoyant demand. We describe the learning process as *complete learning* if the gathered information is sufficient and the price of each firm p_t^i converges to p^{i*} in probability as t goes large. Under the mechanism, every firm strictly follows a complete learning process. Otherwise, price p_t^i would not converge to p^{i*}, meaning that the process constitutes *incomplete learning*. Complete learning is critical for the market to stabilise. Then, this mechanism allows firms to share the estimators of the demand curve with each other to generate the equilibrium price, and each firm reprices at the new equilibrium price.

2.1 Performance Measures

For each feasible pricing policy Π generated by the algorithm, we examine it in two ways. The first method is to evaluate whether $\hat{\mathbf{p}}_n$ will converge to $\mathbf{p}^*$; the second is to measure the impacts of $\hat{\mathbf{p}}_t$ on each firm i's revenues. We quantify this impact as *regret* which is defined as

$$\mathcal{R}^i(\pi^i, T) = \mathbb{E}\left[\sum_{t=1}^{T} r^i(p_t^{i*}, p_t^{-i})\right] - \mathbb{E}\left[\sum_{t=1}^{T} r^i(p_t^i, p_t^{-i})\right] \tag{3}$$

We are most concerned with whether the average regret per period, $\mathcal{R}^i(\pi, T)/T$ converges to zero with an acceptable convergence rate. The definition of regret in the monopoly setting is worse-case regret, which is the revenue of a single firm generated by a given pricing policy compared to the optimal revenue generated by the firm with the knowledge of the demand curve.

However, this definition is not suitable for our setting since revenues obtained at clairvoyant GNE $\mathbf{p}^*$ are not optimal. Alternative pricing vectors $\mathbf{p}$ may exist that lead to higher revenues than those at $\mathbf{p}^*$ when firms search for an equilibrium point (see also in [8]). What should be emphasised here is that this situation, in which price vectors are *not* Nash equilibrium points and incentives remain for firms to change their prices, may happen when all firms engage in non-cooperative competition. Compared to the revenues obtained by the pricing policies Π, the revenues obtained by the clairvoyant Nash equilibrium may be higher or lower. We define the *revenue difference* as

$$\mathcal{D}^i(\pi^i, T) = \left|\mathbb{E}[\sum_{t=1}^{T} r^i(p^{i*}, p^{-i*})] - \mathbb{E}[\sum_{t=1}^{T} r^i(p_t^i, p_t^{-i})]\right| = \left|p^{i*}\lambda^i(\mathbf{p}^*)T - \mathbb{E}[\sum_{t=1}^{T} p_t^i D_t^i]\right| \tag{4}$$

to evaluate it. The revenue difference is positive in each period. As it is impossible to determine whether revenues obtained by the clairvoyant Nash equilibrium are less than those obtained by the pricing policies Π, the difference per period between the two is expressed as an absolute value no less than zero.

2.2 Algorithm

We propose a DDEP algorithm to solve the pricing decisions of each firm in competition. DDEP operates in *stages*, which we index by n for $n = 0, 1, 2, \cdots$. Every stage is equally separated as $N + 1$ time intervals, which we index by m for $m = 1, 2, \cdots, N + 1$, and each interval contains I_n periods (i.e., time steps). In the beginning of stage n, firm i publishes a price $\hat{p}_n^i$, the vector of all prices is denoted as $\hat{\mathbf{p}}_n$. As there are $(N + 1)I_n$ periods in one stage, the notation t_n represents period t at the beginning of stage n. In Step 1, firm i's price $p_t^i = \hat{p}_n^i$ will not be changed in NI_n but will be changed in 1 time interval if the index m is equal to the firm index i. Therefore, all firms are required to adjust their prices sequentially and collect sales information. Using the observations in stage n, the demand functions $\lambda^i(\cdot)$ of all firms are then approximated by a simple linear function. At the end of stage n, all firms need to find the current GNE $\hat{\mathbf{p}}_{n+1}$ for the next stage, which is computed by an approximate demand function. Now, we present in detail the DDEP algorithm looping n from 0 until a terminal stage, given as period T.

- Step 0. Preparation: If $n = 0$, input I_0, v, and $\hat{p}_1^i, \forall i = 1, \cdots, N$. If $n > 0$, set $I_n = \lfloor I_0 v^n \rfloor$ and $\delta_n = I_n^{-\frac{1}{4}}$.
- Step 1: Setting prices. Loop m from 1 to $N + 1$. The rule of firm i's price p_t^i at time t is

  if $m \neq i$,
  $p_t^i = \hat{p}_n^i, \qquad \forall t = t_n + 1, \cdots, t_n + iI_n, t_n + (i + 1)I_n + 1, \cdots, t_n + (N + 1)I_n,$
  if $m = i$,
  $p_t^i = \hat{p}_n^i + \delta_n, \forall t = t_n + iI_n + 1, \cdots, t_n + (i + 1)I_n.$

  End the m-loop. Set $t_{n+1} = t_n + (N + 1)I_n$.
- Step 2. Estimating:

$$(\hat{\alpha}_{n+1}^i, \hat{\beta}_{n+1}^{ij}) = \arg\min_{\alpha^i, \beta^{ij}} \left\{ \sum_{t=t_n+1}^{t=t_n+(N+1)I_n} \left[D_t^i - \left(\alpha^i - \beta^{ii} p_t^i + \sum_{j=1, j\neq i}^{N} \beta^{ij} p_t^j \right) \right]^2 \right\}.$$

- Step 3. Computing the equilibrium: We define the following optimization problem for firm i:

$$\max_{p^i} r_{n+1}^i \equiv \max_{p^i} G_{n+1} \left\{ p^i, p^{-i}, \hat{\alpha}_{n+1}^i, \hat{\beta}_{n+1}^{ij} \right\},$$

  where $G_{n+1} \left\{ p^i, p^{-i}, \hat{\alpha}_{n+1}^i, \hat{\beta}_{n+1}^i \right\}$

$$\equiv \left\{ p^i \left(\hat{\alpha}_{n+1}^i - \hat{\beta}_{n+1}^{ii} p^i + \sum_{j=1, j\neq i}^{N} \hat{\beta}_{n+1}^{ij} p^j \right) \middle| \hat{\alpha}_{n+1}^i - \hat{\beta}_{n+1}^{ii} p^i + \sum_{j=1, j\neq i}^{N} \hat{\beta}_{n+1}^{ij} p^j \geq 0, p^i \in \mathcal{P}^i \right\}.$$

Proceeding to solve the system:

$$\begin{aligned}
\left[\hat{\alpha}^i_{n+1} - 2\hat{\beta}^{ii}_{n+1}p^i + \sum_{j,j\neq i}^{N} \hat{\beta}^{ij}_{n+1}p^j\right] + \mu^{i,1}\left[-\hat{\beta}^i_{n+1}\right] - \mu^{i,2} + \mu^{i,3} = 0 \qquad \forall i,\\
\mu^{i,1} \geq 0, \mu^{i,1}\cdot\left(-\hat{\alpha}^i_{n+1} + \hat{\beta}^{ii}_{n+1}p^i - \sum_{j,j\neq i}^{N}\hat{\beta}^{ij}_{n+1}p^j\right) = 0, \hat{\alpha}^i_{n+1} - \hat{\beta}^{ii}_{n+1}p^i + \sum_{j,j\neq i}^{N}\hat{\beta}^{ij}_{n+1}p^j \geq 0 \ \forall i,\\
\mu^{i,2} \geq 0, \mu^{i,2}\cdot\left(p^i - p^{i,h}\right) = 0, p^{i,h} - p^i \geq 0 \qquad \forall i,\\
\mu^{i,3} \geq 0, \mu^{i,3}\cdot\left(p^{i,l} - p^i\right) = 0, p^i - p^{i,l} \geq 0 \qquad \forall i.
\end{aligned} \tag{5}$$

Then, prices for each firm $\hat{p}^i_{n+1}$ are set to the unique solution of this system. Set $n = n + 1$ and **return to Step 0**.

During the DDEP operation cycle in Step 1 to explore the sensitivity of the market to price changes, each firm must adjust the price by adding δ_n. In Step 2, the underlying demand function of all firms is estimated using linear regression. Note that we only use the data gathered in this stage since past data may lead to inaccurate estimation of the demand function. If the underlying demand curve is a linear function, we can use all the cumulative data for the estimation due to the unchanging market sensitivity provided that price changes. However, as the underlying demand curve is not known *a priori*, market sensitivity may keep changing as the prices change if the demand function is misspecified. In Step 3, the mechanism solves the KKT system (5), which includes the revenue optimization problem of all firms, providing a way to compute the equilibrium under the approximated demand. The solutions to system (5) will be used for the next stage. The structure of operation cycles has been frequently utilized elsewhere in research on learning and pricing problems. [3] proposed a pricing policy to learn the demand function and find optimal pricing corresponding to the approximate demand function. DDEP is a generalization of their work, in which there are no competitors ($N = 1$), so the number of periods needed for each stage reduces to be equal to the number in their work. An important piece of later work is that of [6], which considers joint pricing and inventory control with backlog, assuming that the underlying demand model is multiplicative. Their work achieved minimal possible regret growth. In another paper, [7] considered the inventory problem with lost sales in which the underlying demand model is assumed to be additive.

3 Analysis: Convergence, Revenue Difference, and Regret

Although a linear demand function suffices to estimate various underlying demand patterns in a monopolistic setting in [3], the goal of this paper is to investigate whether such a linear model remains suitable in a competitive setting. To this end, we start by analysing the limit point of $\{\hat{\mathbf{p}}_n\}$ given its existence. We divide the analyses into two parts: first, we argue by intuition that when the limit price exists and the noise is absent, no unilateral changes on limit price $\tilde{p}^i$ will maximize the revenue function for an individual player i; that is, limit price is exactly the clairvoyant GNE. Second, we argue that this statement remains true when noise is present.

We formally show in Section 3.1 that the pricing decision of each firm will converge to the clairvoyant GNE $\mathbf{p}^*$ in probability if all firms accept the synchronized control. We find that the average regret defined in (3) converges to zero and the convergence of difference defined in (4) is dependent on the number of competitors. We discuss the revenue difference bound and a regret bound in Section 3.2.

Analysis for the Noiseless Case.

Lemma 1. *Suppose that $\varepsilon^i_t = 0$, $\forall i$ and t, and that the sequence $\{\hat{\mathbf{p}}_n\}$, assuming nonzero demand and that the price is away from the limits, generated by DDEP converges to a limit point $\tilde{\mathbf{p}}$, which satisfies $\tilde{p}^i = -\dfrac{\lambda^i(\tilde{\mathbf{p}})}{\nabla_{p^i}\lambda^i(\tilde{\mathbf{p}})}$. Then, $\tilde{\mathbf{p}}$ is exactly $\mathbf{p}^*$.*

Proof. Without noise, the estimate of the coefficients has a closed-form expression in terms of the rate of demand change with respect to the price increment δ_n:

$$\hat{\beta}^{ii}_{n+1} = -\frac{\lambda^i(\hat{p}^i_n + \delta_n, \hat{p}^{-i}_n;) - \lambda^i(\hat{\mathbf{p}}_n)}{\delta_n}, \qquad \forall i, \text{ and} \tag{6}$$

$$\hat{\beta}^{ij}_{n+1} = \frac{\lambda^i(\hat{p}^j_n + \delta_n, \hat{p}^{-j}_n) - \lambda^i(\hat{\mathbf{p}}_n)}{\delta_n}, \qquad \forall i, j, i \neq j, \tag{7}$$

where $\lambda^i(\hat{p}^j_n + \delta_n, \hat{p}^{-j}_n)$ denotes the the demand faced by firm i at the price vector whose jth element is $\hat{p}^j_n + \delta_n$ and those of others are $\hat{p}^{-j}_n$. It can be inferred from the design of Step 2 in the algorithm that

$$\hat{\alpha}^i_{n+1} = \lambda^i(\hat{\mathbf{p}}_n) + \hat{\beta}^{ii}_{n+1}\hat{p}^i_n - \sum_{j,j\neq i}^{N} \hat{\beta}^{ij}_{n+1}\hat{p}^j_n, \qquad \forall i, \tag{8}$$

and that the closed-form (unconstrained) maximizer of the revenue function is expressed as

$$\hat{p}^i_{n+1} = \frac{\hat{\alpha}^i_{n+1} + \sum_{j\neq i}^{N} \hat{\beta}^{ij}_{n+1}\hat{p}^j_{n+1}}{2\hat{\beta}^{ii}_{n+1}}, \qquad \forall i. \tag{9}$$

Suppose that the sequence $\{\hat{\mathbf{p}}_n\}$ converges to limit vector $\tilde{\mathbf{p}}$. It must be, therefore, that the sequence $\{\hat{\beta}^{ii}_n\}$ converges to $\tilde{\beta}^{ii}$ and $\{\hat{\beta}^{ij}_n\}$ converges to $\tilde{\beta}^{ij}$, where

$$\tilde{\beta}^{ii} = -\nabla_{p^i}\lambda^i(\tilde{\mathbf{p}}) \text{ and } \tilde{\beta}^{ij} = \nabla_{p^j}\lambda^i(\tilde{\mathbf{p}}).$$

Therefore, (8) implies that $\{\hat{\alpha}^i_n\}$ must converge to

$$\tilde{\alpha}^i = \lambda^i(\tilde{\mathbf{p}}) + \nabla_{p^i}\lambda^i(\tilde{\mathbf{p}})\tilde{p}^i - \sum_{j\neq i}^{N} \nabla_{p^j}\lambda^i(\tilde{\mathbf{p}})\tilde{p}^j.$$

Thus, $\tilde{p}^i$ must satisfy the following equation:

$$\tilde{p}^i = \frac{\tilde{\alpha}^i + \sum_{j\neq i}^{N} \tilde{\beta}^{ij}\tilde{p}^j}{2\tilde{\beta}^{ii}} = \frac{\tilde{p}^i}{2} - \frac{\lambda^i(\tilde{\mathbf{p}})}{2\nabla_{p^i}\lambda^i(\tilde{\mathbf{p}})}.$$

The equation above is exactly the first order condition for firm i's revenue maximization problem. In a competitive environment, the best response function can be inferred by the first-order condition for each firms revenue maximisation problem. Hence, $\tilde{\mathbf{p}}$ must be the unique clairvoyant GNE $\mathbf{p}^*$, which can be obtained by solving the concatenated first-order conditions.

Intuitively, the statement in Lemma 1 still holds when the environment is noisy.

Analysis with Noise. In fact, noise ε^i_t always exists over the time horizon T. Hence,

$$\begin{aligned}\hat{\beta}^{ii}_{n+1} &= -\frac{\sum_{t=t_n+iI_n+1}^{t=t_n+(i+1)I_n}\left(\lambda^i(\hat{p}^i_n + \delta_n, \hat{p}^{-i}_n) + \varepsilon^i_t\right) - \sum_{t=t_n+1}^{t=t_n+I_n}\left(\lambda^i(\hat{\mathbf{p}}_n) + \varepsilon^i_t\right)}{I_n\delta_n} \\ &= -\nabla_{p^i}\lambda^i(\hat{\mathbf{p}}_n) + O(\delta_n) + \frac{1}{\delta_n}\frac{1}{I_n}\left(\sum_{t=t_n+iI_n+1}^{t=t_n+(i+1)I_n}\varepsilon^i_t - \sum_{t=t_n+1}^{t=t_n+I_n}\varepsilon^i_t\right), \ \forall i, \text{ and}\end{aligned}$$

$$\hat{\beta}_{n+1}^{ij} = \frac{\sum_{t=t_n+jI_n+1}^{t=t_n+(j+1)I_n} \left(\lambda^i(\hat{p}_n^j + \delta_n, \hat{p}_n^{-j}) + \varepsilon_t^i \right) - \sum_{t=t_n+1}^{t=t_n+I_n} \left(\lambda^i(\hat{\mathbf{p}}_n) + \varepsilon_t^i \right)}{I_n \delta_n}$$

$$= -\nabla_{p^j} \lambda^i(\hat{\mathbf{p}}_n) + O(\delta_n) + \frac{1}{\delta_n}\frac{1}{I_n} \left(\sum_{t=t_n+jI_n+1}^{t=t_n+(j+1)I_n} \varepsilon_t^i - \sum_{t=t_n+1}^{t=t_n+I_n} \varepsilon_t^i \right), \ \forall i, j,$$

where $O(\delta_n)$ denotes the quantity that is, at most, some constant multiplied by δ_n. Derived using Hoeffdings inequality (formally in Appendix B), an exponential bound shows that

$$I_n^{-1} \left(\sum_{t=t_n+iI_n+1}^{t=t_n+(i+1)I_n} \varepsilon_t^i - \sum_{t=t_n+1}^{t=t_n+I_n} \varepsilon_t^i \right) \text{ and } I_n^{-1} \left(\sum_{t=t_n+jI_n+1}^{t=t_n+(j+1)I_n} \varepsilon_t^i - \sum_{t=t_n+1}^{t=t_n+I_n} \varepsilon_t^i \right)$$

may be bounded above by a factor of $(\log(I_n)/I_n)^{1/2}$ with high probability. As n grows, δ_n and thus $\delta_n^{-1}(\log(I_n)/I_n)^{1/2}$ converge to zero (proved formally in Proposition 3), then

$$\hat{\beta}_{n+1}^{ii} \approx -\nabla_{p^i} \lambda^i(\hat{\mathbf{p}}_n) \text{ and } \hat{\beta}_{n+1}^{ij} \approx \nabla_{p^j} \lambda^i(\hat{\mathbf{p}}_n).$$

In particular, when explicitly accounting for noise, the arguments above ensure that the effects of noise vanish as n increases and that the fitted linear model can serve as an estimation of the underlying demand model without being affected by N.

However, we show later that when facing competition, the upper bound of revenue regret, derived in the same way as that of one firm, is scaled by N (i.e., Theorem 3), the upper bound of revenue difference is scaled by N^2 (i.e., Theorem 2), and the deviation between the best responses and the clairvoyant GNE price is upper bounded by a factor of $N^2 I_n^{-\frac{1}{2}}$ (i.e., Inequality (11)).

3.1 Convergence

We clarify that $\hat{\mathbf{p}}_{n+1}$ generated by DDEP is obtained by the KKT system (5) while the clairvoyant GNE $\mathbf{p}^*$ is obtained by (2). We introduce an operator $z^i \left(\breve{\alpha}^i(\mathbf{p}), \breve{\beta}^{ii}(\mathbf{p}), \breve{\beta}^{ij}(\mathbf{p}) \right)$ that represents the best response of firm i when there is *no estimation error*, and we write firm i's revenue as

$$r^i = p^i \left(\breve{\alpha}^i(\mathbf{p}) - \breve{\beta}^{ii}(\mathbf{p}) p^i + \sum_{j=1, j\neq i}^{N} \breve{\beta}^{ij}(\mathbf{p}) p^j \right),$$

where

$$\breve{\alpha}^i(\mathbf{p}) = \lambda^i(\mathbf{p}) - \nabla_{p^i} \lambda^i(\mathbf{p}) p^i + \sum_{j=1, j\neq i}^{N} \nabla_{p^j} \lambda^i(\mathbf{p}) p^j \text{ and } \breve{\beta}^{ij}(\mathbf{p}) = \sum_{j=1}^{N} \nabla_{p^j} \lambda^i(\mathbf{p}) p^j.$$

Let $\mathbf{z}(\mathbf{p}) = \left(z^i \left(\breve{\alpha}^i(\mathbf{p}), \breve{\beta}^{ii}(\mathbf{p}), \breve{\beta}^{ij}(\mathbf{p}) \right) \right)_i^N$ be the collection of best responses of all firms. We state some assumptions associated with our main results.

Assumption 2 *1. For every z^i, we have* $\max_{p^i \in \mathcal{P}^i} \left| \frac{dz^i \left(\breve{\alpha}^i(\mathbf{p}), \breve{\beta}^{ii}(\mathbf{p}), \breve{\beta}^{ij}(\mathbf{p}) \right)}{dp^i} \right| < 1$.

2. For every firm i, the variance of each firm's ε^i is equal to σ^2.

The first part of the above assumption ensures that $z^i\left(\breve{\alpha}^i(\mathbf{p}), \breve{\beta}^{ii}(\mathbf{p}), \breve{\beta}^{ij}(\mathbf{p})\right)$ constitutes a contraction mapping, and thus there exists a unique fixed point. We observe that p^{i*} is the fixed point of $z^i\left(\breve{\alpha}^i(p^i, p^{-i*}), \breve{\beta}^{ii}(p^i, p^{-i*}), \breve{\beta}^{ij}(p^i, p^{-i*})\right)$, implying that $z^i\left(\breve{\alpha}^i(p^i, p^{-i*}), \breve{\beta}^{ii}(p^i, p^{-i*}), \breve{\beta}^{ij}(p^i, p^{-i*})\right) = p^{i*} = \bar{z}^i(p^i, p^{-i*})$, where $\bar{z}^i(p^i_t, p^{-i}_t)$ denotes the best response of firm i when when its competitors set price p^{-i}_t at time t. We need the second part of the assumption, the homogeneity of variance, to ease our analysis. Hence, we make the following proposition.

Proposition 1

If $p^{i*} = z^i\left(\breve{\alpha}^i(p^i, p^{-i*}), \breve{\beta}^{ii}(p^i, p^{-i*}), \breve{\beta}^{ij}(p^i, p^{-i*})\right)$, then there exists a constant $\gamma \in (0,1)$ such that

$$\|\mathbf{p}^* - \mathbf{z}(\hat{\mathbf{p}}_n)\| \leq \gamma \|\mathbf{p}^* - \hat{\mathbf{p}}_n\|.$$

Proposition 1 is based on a deterministic (mean) demand function, and the convergence result follows directly from the property of a contraction mapping. Now, we focus on a randomized demand function and we aim to establish the convergence result as follows.

Theorem 1

Under Assumptions, the GNE, $\hat{\mathbf{p}}_n$ converges in probability to $\mathbf{p}^*$ as $n \to \infty$.

Theorem 1 states that the prices generated by our algorithm DDEP converge in probability to GNE $\mathbf{p}^*$, which is the solution that satisfies the KKT system (2). The principle technique for establishing the convergence of price in probability is described below. Let the price vector $\hat{\mathbf{p}}_{n+1}$ be the solution of the KKT system (5), which means that $\hat{\mathbf{p}}_{n+1}$ is the Nash equilibrium. Fix $\hat{p}^{-i}_{n+1}$ and let C^i_n denote the difference between $z^i\left(\breve{\alpha}^i(p^i, \hat{p}^{-i}_{n+1}), \breve{\beta}^{ii}(p^i, \hat{p}^{-i}_{n+1}), \breve{\beta}^{ij}(p^i, \hat{p}^{-i}_{n+1})\right)$ and $z^i\left(\hat{\alpha}^i_{n+1}, \hat{\beta}^{ii}_{n+1}, \hat{\beta}^{ij}_{n+1}\right)$, in which the former one represents the best response of firm i to maximize the following revenue function

$$r^i_{n+1} = p^i\left(\breve{\alpha}^i(p^i, \hat{p}^{-i}_{n+1}) - \breve{\beta}^{ii}(p^i, \hat{p}^{-i}_{n+1})p^i + \sum_{j=1, j\neq i}^{N} \breve{\beta}^{ij}(\hat{p}_{n+1})\hat{p}^j_{n+1}\right).$$

While the latter represents the best response of firm i seeking to maximize a second format of revenue function, utilizing fitted rather than the revealed demand is different from the real revenue,

$$r^i_{n+1} = p^i\left(\hat{\alpha}^i_{n+1} - \hat{\beta}^{ii}_{n+1}p^i + \sum_{j=1, j\neq i}^{N} \hat{\beta}^{ij}_{n+1}\hat{p}^j_{n+1}\right).$$

As shown, $\hat{p}^i_{n+1}$ is the fixed point of $z^i\left(\hat{\alpha}^i_{n+1}, \hat{\beta}^{ii}_{n+1}, \hat{\beta}^{ij}_{n+1}\right)$ obtained when other players fix prices as $\hat{p}^{-i}_{n+1}$. We present the following proposition:

Proposition 2

For any given $\hat{p}^i_n \in \mathcal{P}^i$ generated by DDEP, with high probability[a] the following inequality holds

$$\|\mathbf{z}(\hat{\mathbf{p}}_n) - \hat{\mathbf{p}}_{n+1}\| \leq \|C_n\|,$$

where $C_n \equiv \left[C^1_n, \cdots, C^N_n\right]$ is a vector of constants.

[a] The probability lower bound is specified in the proof of Proposition 3.

Proposition 2, which is proved in Appendix A, shows that the difference between these best response functions is bounded with high probability. Note that $\mathbf{z}(\hat{\mathbf{p}}_n)$ denotes the collection of all firms' best responses $z^i\left(\breve{\alpha}^i(\hat{\mathbf{p}}_n), \breve{\beta}^{ii}(\hat{\mathbf{p}}_n), \breve{\beta}^{ij}(\hat{\mathbf{p}}_n)\right)$. The inequality in Proposition 2 may *not* hold because it is reliant on a good event-that the demand shocks ε_t do not cause poor estimation of the parameters. Note that if one firm is unable to make a sufficiently good estimation, other firms will be affected. The effect of price experimentation, in which adding δ_n affects the demand of all participants, should gradually vanish over time. Otherwise, it will cause a poor estimation of the parameters even if demand shocks do not exist.

Broadly speaking, the argument above, together with Proposition 1 provides a constructive way to design effective price experimentation to obtain pricing convergence. The next proposition gives an upper bound for the expectation of the difference between the two best response functions in Proposition 2.

Proposition 3

At stage n, for some suitable constant K_1, the operator $\mathbf{z}(\hat{\mathbf{p}}_n)$ and the DDEP generated $\hat{\mathbf{p}}_{n+1}$ satisfy

$$\mathbb{E}\left[\|\mathbf{z}(\hat{\mathbf{p}}_n) - \hat{\mathbf{p}}_{n+1}\|^2\right] \leq N^2 K_1 I_n^{-\frac{1}{2}}.$$

Proposition 3, proven in Appendix B, also provides an upper bound for the deviation between $\mathbf{z}(\hat{\mathbf{p}}_n)$ and $\hat{\mathbf{p}}_{n+1}$. The upper bound is related to the squared number of competitive firms, N^2, and converges to zero as the number of stages increases.

Proof (Proof of Theorem 1). We obtain an upper bound of $\mathbb{E}[(\hat{\mathbf{p}}_{n+1} - \mathbf{p}^*)^2]$, a result that has been proved previously. Combining the results in Propositions 13, it follows that

$$\begin{aligned}
&\mathbb{E}\left[\|\hat{\mathbf{p}}_{n+1} - \mathbf{p}^*\|^2\right] \\
&\leq \mathbb{E}\left[\left(\|\mathbf{p}^* - \mathbf{z}(\hat{\mathbf{p}}_n)\| + \|\mathbf{z}(\hat{\mathbf{p}}_n) - \hat{\mathbf{p}}_{n+1}\|\right)^2\right] \\
&\leq \mathbb{E}\left[\left(\gamma\|\mathbf{p}^* - \hat{\mathbf{p}}_n\| + \|\mathbf{z}(\hat{\mathbf{p}}_n) - \hat{\mathbf{p}}_{n+1}\|\right)^2\right] \\
&\leq \left(\frac{1+\gamma^2}{2}\right)\mathbb{E}\left[\|\mathbf{p}^* - \hat{\mathbf{p}}_n\|^2\right] + K_2\mathbb{E}\left[\|\mathbf{z}(\hat{\mathbf{p}}_n) - \hat{\mathbf{p}}_{n+1}\|^2\right] \\
&\leq \left(\frac{1+\gamma^2}{2}\right)\mathbb{E}\left[\|\mathbf{p}^* - \hat{\mathbf{p}}_n\|^2\right] + N^2 K_3 I_n^{-\frac{1}{2}}.
\end{aligned} \tag{10}$$

The first inequality uses the triangle inequality. The second inequality is from Proposition 1 with the constant $\gamma \in (0, 1)$. Note that

$$\left(\sqrt{\frac{2\gamma}{1-\gamma^2}\mathbb{E}\left[\|\mathbf{z}(\hat{\mathbf{p}}_n) - \hat{\mathbf{p}}_{n+1}\|^2\right]} - \sqrt{\frac{1-\gamma^2}{2\gamma}\mathbb{E}\left[\|\hat{\mathbf{p}}_n - \mathbf{p}^*\|\right]}\right)^2 \geq 0, \text{and}$$

$$2\mathbb{E}\left[\|\hat{\mathbf{p}}_n - \mathbf{p}^*\|\right]\mathbb{E}\left[\|\mathbf{z}(\hat{\mathbf{p}}_n) - \hat{\mathbf{p}}_{n+1}\|\right] \leq \frac{2\gamma}{1-\gamma^2}\mathbb{E}\left[\|\mathbf{z}(\hat{\mathbf{p}}_n) - \hat{\mathbf{p}}_{n+1}\|^2\right] + \frac{1-\gamma^2}{2\gamma}\mathbb{E}\left[\|\hat{\mathbf{p}}_n - \mathbf{p}^*\|^2\right].$$

Hence, the third inequality is obtained:

$$\mathbb{E}\left[\|\hat{\mathbf{p}}_{n+1} - \mathbf{p}^*\|^2\right] \leq \frac{1+\gamma^2}{2}\mathbb{E}\left[\|\hat{\mathbf{p}}_n - \mathbf{p}^*\|^2\right] + K_2\mathbb{E}\left[\|\mathbf{z}(\hat{\mathbf{p}}_n) - \hat{\mathbf{p}}_{n+1}\|^2\right].$$

Selecting some appropriate constant K_3 and combining it with the results in Proposition 3, we obtain the last inequality. Let $\eta = \frac{1+\gamma^2}{2}$, and we obtain that

$$\begin{aligned}\mathbb{E}\left[\|\hat{\mathbf{p}}_{n+1} - \mathbf{p}^*\|^2\right] &\leq \eta^n \|\mathbf{p}_1 - \mathbf{p}^*\|^2 + N^2 K_3 \sum_{h=0}^{n-1} \eta^h I_{n-h}^{-\frac{1}{2}} \\ &\leq \eta^n \|\mathbf{p}_1 - \mathbf{p}^*\|^2 + N^2 K_3 \sum_{h=0}^{n} \eta^h I_{n-h}^{-\frac{1}{2}} \\ &\leq N^2 K_4 (v^{-\frac{1}{2}})^n \sum_{h=0}^{n} \eta^h v^h,\end{aligned}$$

where v is selected to satisfy $\eta v^{\frac{1}{2}} < 1$ and there exists some appropriate K_5, such that

$$\mathbb{E}\left[\|\hat{\mathbf{p}}_{n+1} - \mathbf{p}^*\|^2\right] \leq N^2 K_5 I_n^{-\frac{1}{2}}. \tag{11}$$

Since $\eta < 1$ and $I_n^{-\frac{1}{2}} \to 0$ as $n \to \infty$, $\mathbb{E}\left[\|\hat{\mathbf{p}}_{n+1} - \mathbf{p}^*\|^2\right] \to 0$.

The convergence implies that if N is considered fixed, as $n \to \infty$, the difference of $\hat{\mathbf{p}}_{n+1}$ and $\mathbf{p}^*$ will tend to zero. Hence, we demonstrate that the generalized equilibrium price will converge to the clairvoyant GNE $\mathbf{p}^*$ in probability.

3.2 Revenue Difference and Regret

Theorem 1 provides the asymptotically equilibrium result for the equilibrium pricing problem: under the designed mechanism, it suffices to show that the linear model guarantees convergence to clairvoyant GNE $\mathbf{p}^*$. This result offers a fundamental basis for the following analysis. We first analyze the revenue difference, as it may yield useful insights into regret analysis. Recall that the revenue difference denotes the difference in the revenue obtained when all firms set the price to $\mathbf{p}^*$ and that obtained when the price is set at $\mathbf{p}_t$.

Theorem 2: Revenue Difference

Under Assumptions, the sequence of the generalized Nash equilibrium $\{\hat{\mathbf{p}}_t : t \geq 1\}$ satisfies

$$\mathbb{E}\left[\sum_{t=1}^{T} \left[\left|r^i(p^{i*}, p^{-i*}) - r^i(p_t^i, p_t^{-i})\right|\right]\right] \leq N^2 K_6 T^{\frac{1}{2}}, \quad \forall i = 1, \cdots, N,$$

for some positive constant K_6, $T \geq 2$, and $N \geq 2$.

Theorem 2, whose proof is in Appendix C, implies that by using this pricing policy, even if the underlying demand function is unknown, the difference between the revenue $r^i(p^{i*}, p^{-i*})$ at the clairvoyant $\mathbf{p}^*$ and the revenue $r^i(p_t^i, p_t^{-i})$ at $\mathbf{p}_t$ obtained by DDEP is at most $N^2 K_6 T^{\frac{1}{2}}$, in which N represents the number of participating firms. In other words, the revenue difference per period converges to zero. When N equals to 1, meaning there are no competitors in the market, the upper bound is $\mathcal{O}(T^{1/2})$.

The common approach in regret analysis in a monopolistic setting, such as that of [3], [7] and [6], involves a Taylor expansion to let the revenue function $r^i(p_t^i, p_t^{-i})$ be approximated at $\mathbf{p}^*$; that is,

$$\begin{aligned}r^i(p_t^i, p_t^{-i}) = r^i(p^{i*}, p^{-i*}) &+ \nabla r^i(p^{i*}, p^{-i*})^T (\mathbf{p}^* - \mathbf{p}) \\ &+ \frac{1}{2}(\mathbf{p}^* - \mathbf{p}_t)^T \nabla^2 r^i(p^{i*}, p^{-i*})(\mathbf{p}^* - \mathbf{p}_t).\end{aligned}$$

Suppose that no competitors exist and the term $\nabla r^i(p^{i*})$ is equal to zero, then regret can be expressed as $\nabla^2 r^i(p^{i*})|p^{i*} - p_t^i|^2$ and the upper bound can be inferred by the result of (11). However, as the term

$\nabla r^i(p^{i*}, p^{-i*})$ is a vector and not equal to zero in our setting, the traditional approach is not suitable for our analysis. It is possible that certain price vectors may yield a higher revenue than Nash equilibrium prices. Despite this, we can prove the same result using the negative semi-definiteness property of the revenue function.

Theorem 2 gives us an upper bound on revenue difference, and we now analyze the revenue regret. It is known that the price p_t^i of firm i made by DDEP is not necessarily the optimal price that truly maximize i's own revenue. Such a true optimal price p_t^{i*} is derived from the true best response function with respect to the true demand. Suppose firm i realizes the true demand function at some time T. In such a case, the firm will regret making the price decision at p_t^i rather than p_t^{i*} during the past time interval $t = 1, \ldots, T$. Hence, we obtain the following theory concerning the regret bound:

Theorem 3: Regret

Under the defined assumptions, the sequence of optimal decisions $\{p_t^{i*} : t \geq 1\}$ satisfies

$$\mathbb{E}\left[\sum_{t=1}^{T}\left[r^i(p_t^{i*}, p_t^{-i}) - r^i(p_t^i, p_t^{-i})\right]\right] \leq K_7 N T^{\frac{1}{2}}, \quad \forall i = 1, \cdots, N,$$

for some positive constant K_7, $T \geq 2$ and $N \geq 2$.

Theorem 3 shows that after time T, the revenue gap (regret) between making pricing at p_t^i and making pricing at p_t^{i*} is at most $NK_7T^{\frac{1}{2}}$, which is also contingent on the number of firms. The revenue differences have a worse bound than those obtained based on the regret of a single firm. There are two major underlying reasons for this. First, the Taylor expansion of the revenue functions is different in the analysis of these two cases. In this case, the Taylor expansion allows the revenue function to be approximated at p_t^{i*} which is the optimal pricing at time t, when the price of other firms p_t^{-i} is fixed. The regret bound can be expressed as $\frac{1}{2}\nabla^2 r^i(p_t^{i*}, p_t^{-i})|p_t^{i*} - p_t^i|^2$ while the revenue difference bound is $\frac{1}{2}\nabla^2 r^i(p_t^{i*}, p_t^{-i*})\|\mathbf{p}^* - \mathbf{p}_t\|^2$. Hence, the difference between convergence rates of p_t to p_t^{i*} and $\mathbf{p}^*$ to $\mathbf{p}_t$ causes the results to differ. Second, the revenue used to compute a revenue difference with the algorithm relies on the revenue obtained at a clairvoyant GNE $\mathbf{p}^*$. The revenue used to compute regret is the revenue obtained at p_t^{i*} which is the true revenue maximizer at time t. The optimal pricing p_t^{i*} is virtual and is therefore not useful in capturing sales information. Thus, firm i cannot predict the subsequent pricing of other firms and, moreover, the convergence of p_t to p_t^{i*} is unrelated to the convergence of p_t^{-i} to p_t^{-i*}. In fact, when p_t^{i*} no longer changes, it means that p_t^{i*} becomes a fixed point of the best response function, z^i, indicating that pricing decisions have reached the clairvoyant GNE $\mathbf{p}^*$.

Proof (Proof of Theorem 3). Fix a time horizon T, and firm i, and let $k = \inf\left\{h \geq 1 : (N+1)\sum_{i=1}^{h} I_n \geq T\right\}$. The regret after T periods is given by

$$R^i = \mathbb{E}\left[\sum_{t=1}^{T}\left[r^i(p_t^{i*}, p_t^{-i}) - r^i(p_t^i, p_t^{-i})\right]\right].$$

Note that price set by other firms p_t^{-i} is not changable in the definition of regret. Let the revenue function $r^i(p_t^i, p_t^{-i})$ be approximated at p_t^{i*}; hence,

$$\begin{aligned} r^i(p^i, p^{-i}) &= r^i(p_t^{i*}, p_t^{-i}) + \nabla_{p^i} r^i(p_t^{i*}, p_t^{-i})(p_t^{i*} - p_t^i) + \frac{1}{2}\nabla_{p^i}^2 r^i(p^{i*}, p_t^{-i})(p_t^{i*} - p_t^i)^2 \\ &= r^i(p_t^{i*}, p_t^{-i}) + \frac{1}{2}\nabla_{p^i}^2 r^i(p^{i*}, p_t^{-i})(p_t^{i*} - p_t^i)^2. \end{aligned}$$

Since p_t^{i*} is the optimal solution at time t, the term $\nabla_{p^i} r^i(p_t^{i*}, p_t^{-i})$ is equal to zero. It should be emphasized that p_t^{i*} is the fixed point of $z^i\big(\breve{\alpha}(p_t^{i*}, p_t^{-i}), \breve{\beta}(p_t^{i*}, p_t^{-i})\big)$ since $\hat{p}_t^{-i}$ is considered fixed. According to Lemma 1, we have

$$p_t^{i*} = -\frac{\lambda^i(p_t^{i*}, p_t^{-i})}{\nabla_{p^i}\lambda^i(p_t^{i*}, p_t^{-i})} = \frac{\breve{\alpha}(p_t^{i*}, p_t^{-i}) + \sum_{j,j\neq i}^{N} \breve{\beta}^{ij}(p_t^{i*}, \hat{p}_t^{-i})p_t^j}{2\breve{\beta}^{ii}(p_t^{i*}, p_t^{-i})}$$

where $\breve{\alpha}(p_t^{i*}, p_t^{-i}) = \lambda^i(p_t^{i*}, p_t^{-i}) + \nabla_{p^i}\lambda^i(p_t^{i*}, p_t^{-i})p_t^{i*} - \sum_{j,j\neq i}^{N} \nabla_{p^j}\lambda^i(p_t^{i*}, p_t^{-i})p_t^j$, $\breve{\beta}^{ij}(p_t^{i*}, p_t^{-i}) = \nabla_{p^j}\lambda^i(p_t^{i*}, p_t^{-i})$ and $\breve{\beta}^{ii}(p_t^{i*}, p_t^{-i}) = \nabla_{p^i}\lambda^i(p_t^{i*}, p_t^{-i})$. We divide the price experimentation in Step 1 into three parts and analyze the regret separately. First, in the first I_n periods, there is no adjustment of price. For $t = t_n + 1, \cdots, t_n + I_n$, we have that $\mathbf{p}_t = \hat{\mathbf{p}}_n$; thus

$$\begin{aligned}\left|p_t^{i*} - p_t^i\right| &= \left|z^i\big(\breve{\alpha}(p_t^{i*}, \hat{p}_n^{-i}), \breve{\beta}(p_t^{i*}, \hat{p}_n^{-i})\big) - z^i\big(\hat{\alpha}(\hat{\mathbf{p}}_n), \hat{\beta}(\hat{\mathbf{p}}_n)\big)\right| \\ &= \left|p_t^{i*} - z^i\big(\breve{\alpha}(\hat{\mathbf{p}}_n), \breve{\beta}(\hat{\mathbf{p}}_n)\big)\right| + C_n^i = C_n^i, \end{aligned} \tag{12}$$

where C_n^i can be derived from (19). Second, in the periods when the firm i has changed its price, the optimal price is the same as the previous one, as competitors do not change their strategies. For all $t = t_n + iI_n + 1, \cdots, t_n + (i+1)I_n$, we have that $\mathbf{p}_t = (\hat{p}_n^i + \delta_n, \hat{p}_n^{-i})$. Similar arguments show

$$\left|p_t^{i*} - p_t^i\right| = \left|z^i\big(\breve{\alpha}(p_t^{i*}, \hat{p}_n^{-i}), \breve{\beta}(p_t^{i*}, \hat{p}_n^{-i})\big) - z^i\big(\hat{\alpha}(\hat{p}_n^i, \hat{p}_n^{-i}), \hat{\beta}(\hat{p}_n^i, \hat{p}_n^{-i})\big) - \delta_n\right| = C_n^i + \delta_n. \tag{13}$$

Third, there are the periods when the other firms, except firm i, change their prices. Suppose that firm j has also adjusted its price, for all $t = t_n + jI_n + 1, \cdots, t_n + (j+1)I_n$, we have that $\mathbf{p}_t = (\hat{p}_n^j + \delta_n, \hat{p}_n^{-j})$, and it follows that

$$\begin{aligned}\left|p_t^{i*} - p_t^i\right| &= \left|z^i\big(\breve{\alpha}(p_t^{i*}, \hat{p}_n^{-i}), \breve{\beta}(p_t^{i*}, \hat{p}_n^{-i})\big) - z^i\big(\hat{\alpha}(\hat{\mathbf{p}}_n), \hat{\beta}(\hat{\mathbf{p}}_n)\big)\right| \\ &= \left|z^i\big(\breve{\alpha}(p_t^{i*}, \hat{p}_n^{-i\backslash j}, \hat{p}_n^j + \delta_n), \breve{\beta}(p_t^{i*}, \hat{p}_n^{-i\backslash j}, \hat{p}_n^j + \delta_n)\big) - z^i\big(\hat{\alpha}(\hat{\mathbf{p}}_n), \hat{\beta}(\hat{\mathbf{p}}_n)\big)\right| \\ &= \left|z_{p^j}^i\delta_n + z^i\big(\breve{\alpha}(p_t^{i*}, \hat{p}_n^{-i}), \breve{\beta}(p_t^{i*}, \hat{p}_n^{-i})\big) - z^i\big(\hat{\alpha}(\hat{\mathbf{p}}_n), \hat{\beta}(\hat{\mathbf{p}}_n)\big)\right| \\ &= \left|z_{p^j}^i\delta_n + p_t^{i*} - z^i\big(\breve{\alpha}(\hat{\mathbf{p}}_n), \breve{\beta}(\hat{\mathbf{p}}_n)\big)\right| + C_n^i = C_n^i + K_{44}\delta_n, \forall j. \end{aligned} \tag{14}$$

As discussed in the proof of Theorem 1, value of C_n^i is divisible into two situations: good and bad. However, there is no need to consider the possibility of a good event for all firms since the prices p_t^{-i} cannot change. The optimal pricing p_t^{i*} cannot be used to capture sales information,opportunity to predict the optimal pricing of other firms. Combining the results of (12)-(14) gives the following:

$$\mathbb{E}\left[\left\|p_t^{i*} - \hat{p}_t^i\right\|^2\right] = \left|z^i\big(\breve{\alpha}(p_t^{i*}, \hat{p}_t^{-i}), \breve{\beta}(p_t^{i*}, \hat{p}_t^{-i})\big) - z^i\big(\hat{\alpha}(\mathbf{p}_t), \hat{\beta}(\mathbf{p}_t)\big)\right|^2 = NK_{45}I_n^{-\frac{1}{2}}.$$

Fixed a time horizon T and considering i, while letting $k = \inf\left\{h \geq 1 : (N+1)\sum_{i=1}^{h} I_n \geq T\right\}$, and satisfy $\log_v\left(\frac{v-1}{(N+1)I_0 v}T + 1\right) \leq k < \log_v\left(\frac{v-1}{(N+1)I_0 v}T + 1\right) + 1$, we obtain that

$$R^i \leq \mathbb{E}\Bigg[\sum_{t=1}^{T}\Big(\big[r^i(p_n^{i*}, \hat{p}_n^{-i}) - r^i(\hat{p}_n^i, \hat{p}_n^{-i})\big] + \big[r^i(p_n^{i*}, \hat{p}_n^{-i}) - r^i(\hat{p}_n^i, \hat{p}_n^1 + \delta_n, \hat{p}_n^{-i,-i\neq 1})\big]$$

$$+ \cdots + \big[r^i(p_n^{i*}, \hat{p}_n^{-i}) - r^i(\hat{p}_n^i, \hat{p}_n^N + \delta_n, \hat{p}_n^{-i,-i\neq N})\big]\Big)I_n\Bigg]$$

$$\leq K_{46}\sum_{n=1}^{k}\Big((N+1)\mathbb{E}\left\|p_n^{i*}, \hat{p}_n^{-i}\right\|^2 + N\delta_n^2\Big)I_n$$

$$\leq NK_7T^{\frac{1}{2}}.$$

4 Partially-Clairvoyant Model and Results

Given the results above, a question arises as to whether the regret and difference bounds still hold when some firms have knowledge of the underlying demand curves $\lambda^i(\cdot)$ and the distributions of demand shocks ε^i.

[8] considered a linear model in which decision makers ignore the impact of competitors, that is, the *flawed* model $\alpha^i - \beta^{ii}p^i$ in which the prices of the other firms do not appear in an estimated demand curve. An important property in their work is the demand consistency property. They considered a limit point $\tilde{\mathbf{p}}$ which is the clairvoyant GNE and assumed that all firms use these flawed estimated demand curves. The estimated demand and the expected demand at $\tilde{\mathbf{p}}$ are equal, that is, $\alpha^i - \beta^i\tilde{p}^{i*} = \mathbb{E}[D^i(\tilde{\mathbf{p}})]$. However, using this kind of flawed model may lead the firms to converge to some potential limit points that are not Nash equilibria. Thus, the flawed estimated demand curves *cannot* be used directly for our purpose.

Suppose that N' firms do not know the demand curve. We thus define the partially-flawed estimated demand curves under the partially-clairvoyant model for the firms without the knowledge of underlying demand curves as

$$\alpha^i - \beta^{ii}p^i + \sum_{j=1, j\neq i}^{N'} \beta^{ij}p^j,$$

in which the effects of the firms with the knowledge of underlying demand curves do not appear in this model. We justify the choice of such partially-flawed estimated demand curves in the following. If the firms without the knowledge of demand curves use the full estimated demand curves, the number of price combinations in one stage should be $(N+1)$ to determine the values of $N+1$ unknown parameters. However, the clairvoyant firms have no need to estimate the demand curves and are not motivated to attend the price experimentation which may lead to lower revenues for them. The price experimentation can only consist of $(N'+1)$ prices combinations. If the firms without the knowledge of demand curves considered the full estimated demand curves while the clairvoyant firms do not attend the price experimentation, the result of experimentation would be under-determined equations and might result in inaccurate parameters estimation. In addition, in order to ensure the convergence of the parameters, we also require the firms without the knowledge of underlying demand curves to adjust their prices (by adding δ_n) and set equilibrium prices synchronously.

N' firms accept the mechanism to perform price experimentation. The modified Step 1 in DDEP dictates that price experimentation now only involves N' firms while Step 3 remains the same. The index k represents the firms with known demand, $k = N'+1, \cdots, N$. In each iteration, N' firms keep learning the demand function and N firms compute the equilibrium simultaneously. The modified DDEP procedure is given as follows:

– Step 0. Preparation: If $n = 0$, input I_0, v, and $\hat{p}_1^i, \forall i = 1, \cdots, N'$, $\forall k = N'+1, \cdots, N$. If $n > 0$, set $I_n = \lfloor I_0 v^n \rfloor$ and $\delta_n = I_n^{-\frac{1}{4}}$.

– Step 1. Setting prices: Firm k's pricing p_t^k, for $k = N'+1, \cdots, N$, is:

$$p_t^k = \hat{p}_n^k \quad \forall t = t_n + 1, \cdots, t_n + (N'+1)I_n.$$

Loop m from 1 to $N'+1$. The rule of firm i's pricing p_t^i at time t is :

$$\begin{array}{ll} \text{if } m \neq i, & \\ p_t^i = \hat{p}_n^i, & \forall t = t_n + 1, \cdots, t_n + iI_n, t_n + (i+1)I_n + 1, \cdots, t_n + (N'+1)I_n, \\ \text{if } m = i, & \\ p_t^i = \hat{p}_n^i + \delta_n, \forall t = t_n + iI_n + 1, \cdots, t_n + (i+1)I_n. & \end{array}$$

End the m-loop. Set $t_{n+1} = t_n + (N'+1)I_n$.

– Step 2. Estimating:

$$\left(\hat{\alpha}_{n+1}^i, \hat{\beta}_{n+1}^{ij}\right) = \arg\min_{\alpha^i, \beta^{ij}} \left\{ \sum_{t=t_n+1}^{t=t_n+(N'+1)I_n} \left[D_t^i - \left(\alpha^i - \beta^{ii} p_t^i + \sum_{j=1, j\neq i}^{N'} \beta^{ij} p_t^j \right) \right]^2 \right\}.$$

– Step 3. Computing the equilibrium: We define the optimization problem for firm i as follows:

$$\max_{p^i} r_{n+1}^i \equiv \max_{p^i} G_{n+1} \left\{ p^i, p^{-i}, \hat{\alpha}_{n+1}^i, \hat{\beta}_{n+1}^{ij} \right\}$$

where $G_{n+1} \left\{ p^i, p^{-i}, \hat{\alpha}_{n+1}^i, \hat{\beta}_{n+1}^i \right\}$

$$\equiv \left\{ p_n^i \left(\hat{\alpha}_{n+1}^i - \hat{\beta}_{n+1}^{ii} p_n^i + \sum_{j=1, j\neq i}^{N'} \hat{\beta}_{n+1}^{ij} p_n^j \right) \middle| \hat{\alpha}_{n+1}^i - \hat{\beta}_{n+1}^{ii} p_n^i + \sum_{j=1, j\neq i}^{N'} \hat{\beta}_{n+1}^{ij} p_n^j \geq 0, p^i \in \mathcal{P}^i \right\}.$$

Solve the following two systems simultaneously:

$$\begin{aligned} \left[\hat{\alpha}_{n+1}^i - 2\hat{\beta}_{n+1}^{ii} p^i + \sum_{j, j\neq i}^{N'} \hat{\beta}_{n+1}^{ij} p^j \right] + \mu^{i,1} \left[-\hat{\beta}_{n+1}^i \right] - \mu^{i,2} + \mu^{i,3} = 0, \forall i, & \\ \mu^{i,1} \geq 0, \mu^{i,1} \cdot \left(-\hat{\alpha}_{n+1}^i + \hat{\beta}_{n+1}^{ii} p^i - \sum_{j, j\neq i}^{N'} \hat{\beta}_{n+1}^{ij} p^j \right) = 0, \qquad & \forall i, \\ -\hat{\alpha}_{n+1}^i + \hat{\beta}_{n+1}^{ii} p_n^i - \sum_{j, j\neq i}^{N'} \hat{\beta}_{n+1}^{ij} p^j \leq 0, \qquad & \forall i, \\ \mu^{i,2} \geq 0, \mu^{i,2} \cdot \left(p^i - p^{i,h} \right) = 0, p^i - p^{i,h} \leq 0 \qquad & \forall i, \\ \mu^{i,3} \geq 0, \mu^{i,3} \cdot \left(p^{i,l} - p^i \right) = 0, p^{i,l} - p^i \leq 0 \qquad & \forall i, \end{aligned} \tag{15}$$

and

$$\begin{aligned} [\lambda^k(\mathbf{p}) + p^k \nabla_{p^k} \lambda^k(\mathbf{p}) + \mathbb{E}(\varepsilon^k)] + \mu^{k,1} [\nabla_{p^k} \lambda^k(\mathbf{p})] - \mu^{k,2} + \mu^{k,3} = 0 \ \forall k, & \\ \mu^{k,1} \geq 0, \mu^{k,1} \cdot \left(-\lambda^k(\mathbf{p}) - \mathbb{E}(\varepsilon^k) \right) = 0, -\lambda^k(\mathbf{p}) - \mathbb{E}(\varepsilon^k) \leq 0, \qquad & \forall k, \\ \mu^{k,2} \geq 0, \mu^{k,2} \cdot \left(p^k - p^{k,h} \right) = 0, p^k - p^{k,h} \leq 0 \qquad & \forall k, \\ \mu^{k,3} \geq 0, \mu^{k,3} \cdot \left(p^{k,l} - p^k \right) = 0, p^{k,l} - p^k \leq 0 \qquad & \forall k. \end{aligned} \tag{16}$$

Then, the prices for each firm $\hat{p}_{n+1}^i$ and $\hat{p}_{n+1}^k$ are set to the unique solution of the above simultaneous equations (15) and (16). Set $n := n + 1$ and **return to Step 0**.

4.1 Analyses

In the above algorithm, the linear model is different from that described in Section 2. We adjust it according to the number of firms participating in price experimentation. The mechanism requires that firms with knowledge of the demand functions modify prices simultaneously; this is important for those firms lacking knowledge of the demand functions. There are other major changes in Step 3, although the equilibrium price is calculated simultaneously. The prices of the clairvoyant firms would have no influence on the pricing decisions of the firms without knowledge of the demand functions while the latter's pricing decisions affect the former's. Similar to Lemma 1, we show that the linear function

$$\alpha^i - \beta^{ii} p^i + \sum_{j=1, j\neq i}^{N'} \beta^{ij} p^j$$

is sufficient for our learning. An analogy of Lemma 1 in this partially-clairvoyant setting is as follows:

Lemma 2. *Suppose that* $\varepsilon_t^i = 0$, $\forall i = 1, \cdots, N'$ *and* t, *and that the sequence* $\{\hat{\mathbf{p}}_n\}$ *generated by DDEP converges to a limit point* $\tilde{\mathbf{p}}$, *which satisfies* $\tilde{p}^i = -\frac{\lambda^i(\tilde{\mathbf{p}})}{\nabla_{\tilde{p}^i}\lambda^i(\tilde{\mathbf{p}})}$. *Then,* $\tilde{\mathbf{p}}$ *is exactly* $\mathbf{p}^*$.

Proof. Similar to the proof of Lemma 1, we apply simple variation in this scenario. The proof only focuses on the case where $\varepsilon^i = 0$. Parameters $\hat{\beta}_{n+1}^{ii}$ and $\hat{\beta}_{n+1}^{ij}$ can be derived in the same way; however, $\hat{\alpha}_{n+1}^i$ has changed. It therefore follows that

$$\hat{\alpha}_{n+1}^i = \lambda^i(\hat{\mathbf{p}}_n) + \hat{\beta}_{n+1}^{ii} \hat{p}_n^i - \sum_{j, j\neq i}^{N'} \hat{\beta}_{n+1}^{ij} \hat{p}_n^j, \qquad \forall i, j = 1, \cdots, N', \tag{17}$$

$$\hat{p}_{n+1}^i = \frac{\hat{\alpha}_{n+1}^i + \sum_{j\neq i}^{N'} \hat{\beta}_{n+1}^{ij}}{2\hat{\beta}_{n+1}^{ii}}, \qquad \forall i, j = 1, \cdots, N'. \tag{18}$$

As shown, $\hat{p}_{n+1}$ remains the same. The rest of the proof is same as that of Lemma 1.

Let us briefly explore the characteristics of the flawed linear model in [8]. Unlike our previous linear model in Section 2, the impact of prices representing other competitors will not be fully expressed as the impact of firms with known demand functions is omitted. At stage n, firm i, $i = 1, ..., N'$ needs to participate in a price experimentation to learn the demand function, while firm k, $k = N' + 1, \cdot, N$ does not need to. This means that the parameters β^{ik} cannot be estimated correctly. For firm i, totally unaware of firm k's price, the demand response from firm k is captured by the parameter α^i. Firm k has a known demand function and is still very sensitive to price changes by competitors. To better distinguish our models from the flawed model of [8], we define the partially-flawed estimated demand curve as $\alpha^i - \beta^{ii} p^i + \sum_{j, j\neq i}^{N'} \beta^{ij} p^j$. From the best response function perspective, $\hat{p}_{n+1}^i$ is set by

$$\arg\max p^i \left(\hat{\alpha}_{n+1}^i - \hat{\beta}_{n+1}^{ii} p^i + \sum_{j, j\neq i}^{N'} \hat{\beta}_{n+1}^{ij} \hat{p}_{n+1}^j \right)$$

at the end of stage n. Compared with the previous arguments in Section 3.1, the main difference is the parameter $\hat{\alpha}_{n+1}^i$ in both models. Hence, during the estimation cycle, an obvious result can be used to describe the difference between them, namely $\sum_{k=N'+1}^{N} \nabla_{p^k} \lambda^i(\mathbf{p}) p^k$. Therefore, as the previous convergence result naturally applies to this model, we turn our attention to regret analysis. To demonstrate the influence of the firms with knowledge of the underlying demand functions, we show the following results.

Theorem 4

Under the previously stated assumptions, suppose N' firms do not know the underlying demand function, the sequence of the GNE $\{\mathbf{p}_t : t \geq 1\}$ converges to $\mathbf{p}^*$, and satisfies

$$\mathbb{E}\left[\sum_{t=1}^{T}\left[r^i(p_t^{i*}, p_t^{-i*}) - r^i(p_t^i, p_t^{-i})\right]\right] \leq N'^2 K_8 T^{\frac{1}{2}}, \quad \forall i = 1, \cdots, N',$$

for some positive constant K_8, $T \geq 2$ and $N \geq 2, 1 \leq N' \leq N$.

The aggregate difference over T periods between the revenues $r_i(p_t^{i*}, p_t^{-i*})$ and $r_i(p_t^i, p_t^{-i})$ is at most $N'^2 K_8 T^{\frac{1}{2}}$. Comparing this to the results contained in Theorem 2, one apparent difference is that N^2 changes into N'^2, which implies that the revenue difference is dependent on the number of the firms without knowledge of the demand curve. The upper bound of this difference is applicable to any firms, including those that do and do not know the underlying demand function. It stands to reason that if one firm does not know the demand function, it must expend effort to learn the demand function and so the equilibrium will gradually converge to the clairvoyant GNE $\mathbf{p}^*$. If one firm does not reach the GNE $\mathbf{p}^*$, there is no way for all firms to stabilize pricing. Similar to the results in Theorem 3, we derive a further theorem.

Theorem 5

Under the previously stated assumptions, the sequence of optimal decisions $\left\{p_t^{i*} : t \geq 1\right\}$ of the firms who do not know the demand curve satisfies

$$\mathbb{E}\left[\sum_{t=1}^{T}\left[r^i(p_t^{i*}, p_t^{-i}) - r^i(p_t^i, p_t^{-i})\right]\right] \leq N' K_9 T^{\frac{1}{2}}, \quad \forall i = 1, \cdots, N',$$

for some positive constant K_9, $T \geq 2$ and $N \geq 2, 1 \leq N' \leq N$.
The sequence of optimal decisions $\left\{p_t^{i*} : t \geq 1\right\}$ of the firms who know the demand curve satisfies

$$\mathbb{E}\left[\sum_{t=1}^{T}\left[r^i(p_t^{i*}, p_t^{-i}) - r^i(p_t^i, p_t^{-i})\right]\right] \leq K_{10} T^{\frac{1}{2}}, \quad \forall i = N'+1, \cdots, N,$$

for some positive constant K_{10}, $T \geq 2$ and $N \geq 2, 1 \leq N' \leq N$.

It should be emphasized that the two kinds of firms have different regret bounds as shown in Theorem 5. The price decisions of the clairvoyant firms always follows the best responses according to the true demand curves, and there is no estimation error to affect the pricing decisions. However, their pricing decisions cannot affect firms without knowledge of the demand curves. Theorem 5 shows that the upper bound on the regret for firms without knowledge of the demand curves is still associated with the number of those firms while the upper bound on the regret for the clairvoyant firms is not.

5 Numerical Experiments

We examine the performance of the DDEP algorithm and present the numerical results related to regret. We assume that the underlying demand curve is multinomial logit:

$$\frac{\exp^{\alpha^i-\beta^i p^i}}{1+\sum_{i=1}^{N}\exp^{\alpha^i-\beta^i p^i}}, \alpha^i \in [\underline{\alpha}^i, \bar{\alpha}^i], \beta^i \in [\underline{\beta}^i, \bar{\beta}^i],$$

where $[\underline{\alpha}^i, \bar{\alpha}^i] = [3, 4]$ and $[\underline{\beta}^i, \bar{\beta}^i] = [0.4, 0.5]$ for each firm $i = 1, \cdots, N$. The random error of ε_t^i is assumed to follow a normal distribution with mean 0 and variance σ^2. Owing to the properties of the multinomial logit model, each firms demand does not exceed 1. Moreover, as the number of firms increases, setting the standard deviation at a fixed value will lead to unrealistic experimental settings. Hence, we set the standard deviation as the ratio (κ) of the average demand ($\bar{\lambda}$) of all firms, i.e., $\sigma = \kappa \cdot \bar{\lambda}$, and it will vary with the number of firms participating in price experimentation.

For the above specifications, we randomly select 100 combinations of the parameters α^i and β^i under the uniform distributions for each firm i, $i = 1, \cdots, N$. Each combination provides the parameters for the underlying demand curve of all firms. In all experiments, we set an initial batch size $I_0 = N + 1$ and price interval $\mathcal{P}^i = [p^{i,l}, p^{i,h}] = [0, 6]$. The initial price $\hat{p}_1^i$ is also randomly selected from $\mathcal{P}^i$, for each firm $i = 1, \cdots, N$. To present the result of the convergence of regret, for each period we compute the fraction of optimal revenue as follows:

$$\frac{\sum_{t=1}^{T} p_t^i D_t^i}{\sum_{t=1}^{T} p_t^{i*} \lambda^i(p_t^{i*}, p_t^{-i})}.$$

We test the performance of DDEP with different numbers of firms ($N = 2, 3, 4, 5$) and the value of v is set at 2. Table 1 summarizes the results for the convergence of regret.

Table 1. The performance of DDEP in terms of the fraction of optimal revenue

	Period ($\kappa = 5\%$)			Period ($\kappa = 10\%$)			Period ($\kappa = 15\%$)		
N=2	100	1000	10000	100	1000	10000	100	1000	10000
firm 1	0.9442	0.9905	0.9978	0.9005	0.9718	0.9925	0.7975	0.9289	0.9814
firm 2	0.9419	0.9891	0.9974	0.9159	0.9765	0.9926	0.8130	0.9410	0.9815
N=3	100	1000	10000	100	1000	10000	100	1000	10000
firm 1	0.8941	0.9880	0.9929	0.8527	0.9563	0.9885	0.7957	0.9103	0.9740
firm 2	0.8995	0.9832	0.9973	0.8497	0.9671	0.9877	0.8023	0.9007	0.9818
firm 3	0.8964	0.9727	0.9948	0.8600	0.9412	0.9856	0.8018	0.9091	0.9796
N=4	100	1000	10000	100	1000	10000	100	1000	10000
firm 1	0.8489	0.9703	0.9948	0.8279	0.9354	0.9874	0.8004	0.9134	0.9779
firm 2	0.8878	0.9779	0.9957	0.8144	0.9444	0.9900	0.7893	0.9003	0.9727
firm 3	0.8824	0.9733	0.9957	0.8324	0.9431	0.9887	0.7812	0.9072	0.9751
firm 4	0.8562	0.9701	0.9952	0.8129	0.9453	0.9888	0.8049	0.9054	0.9745
N=5	100	1000	10000	100	1000	10000	100	1000	10000
firm 1	0.8284	0.9618	0.9939	0.7978	0.9179	0.9824	0.8032	0.8805	0.9600
firm 2	0.8491	0.9631	0.9938	0.7855	0.9151	0.9816	0.8094	0.8783	0.9551
firm 3	0.8527	0.9616	0.9936	0.8009	0.9211	0.9840	0.7956	0.8587	0.9650
firm 4	0.8179	0.9542	0.9928	0.7985	0.9179	0.9828	0.8186	0.8964	0.9647
firm 5	0.8386	0.9623	0.9936	0.8053	0.9116	0.9835	0.7826	0.8799	0.9624

In Table 1, as the number of firms increases, the fraction of the optimal revenues grows significantly. Additionally, as the demand shocks become larger, the fraction is reduced.

For the partially-clairvoyant model, we let $I_0 = N' + 1$ to test the performance of the modified DDEP with different numbers of firms ($N = 3, 4, 5$).

Table 2. The performance of the modified DDEP in terms of the fraction of optimal revenue

	Period ($\kappa = 5\%$)			Period ($\kappa = 10\%$)			Period ($\kappa = 15\%$)		
N=3, $N' = 2$	100	1000	10000	100	1000	10000	100	1000	10000
firm 1*	0.9655	0.9964	0.9996	0.9676	0.9970	0.9996	0.9593	0.9954	0.9997
firm 2	0.9229	0.9862	0.9971	0.9164	0.9732	0.9938	0.8775	0.9495	0.9867
firm 3	0.9421	0.9885	0.9971	0.9159	0.9762	0.9941	0.9014	0.9596	0.9881
N=4, $N' = 3$	100	1000	10000	100	1000	10000	100	1000	10000
firm 1*	0.9412	0.9943	0.9994	0.9317	0.9932	0.9993	0.9267	0.9910	0.9989
firm 2	0.9044	0.9826	0.9966	0.8848	0.9654	0.9926	0.8170	0.9318	0.9837
firm 3	0.8997	0.9811	0.9964	0.8657	0.9693	0.9928	0.8264	0.9452	0.9839
firm 4	0.8889	0.9801	0.9962	0.8671	0.9614	0.9918	0.8201	0.9180	0.9832
N=4, $N' = 2$	100	1000	10000	100	1000	10000	100	1000	10000
firm 1*	0.9637	0.9964	0.9996	0.9546	0.9955	0.9994	0.9640	0.9967	0.9997
firm 2*	0.9704	0.9970	0.9997	0.9643	0.9961	0.9996	0.9595	0.9957	0.9995
firm 3	0.9263	0.9849	0.9966	0.9042	0.9729	0.9933	0.8547	0.9621	0.9885
firm 4	0.9388	0.9865	0.9969	0.9175	0.9695	0.9924	0.8737	0.9553	0.9868
N=5, $N' = 4$	100	1000	10000	100	1000	10000	100	1000	10000
firm 1*	0.9138	0.9912	0.9992	0.9017	0.9906	0.9990	0.8674	0.9879	0.9987
firm 2	0.8568	0.9705	0.9948	0.8233	0.9485	0.9890	0.8044	0.9302	0.9803
firm 3	0.8539	0.9703	0.9949	0.8542	0.9603	0.9909	0.8224	0.9212	0.9781
firm 4	0.8643	0.9713	0.9950	0.8401	0.9590	0.9918	0.7576	0.9189	0.9779
firm 5	0.8816	0.9731	0.9954	0.8362	0.9590	0.9920	0.7923	0.9217	0.9781
N=5, $N' = 3$	100	1000	10000	100	1000	10000	100	1000	10000
firm 1*	0.9402	0.9940	0.9993	0.9252	0.9921	0.9993	0.9285	0.9929	0.9992
firm 2*	0.9318	0.9931	0.9993	0.9270	0.9923	0.9991	0.9294	0.9932	0.9993
firm 3	0.8999	0.9805	0.9960	0.8642	0.9616	0.9920	0.8401	0.9382	0.9844
firm 4	0.8730	0.9752	0.9955	0.8739	0.9601	0.9922	0.8238	0.9375	0.9864
firm 5	0.8852	0.9775	0.9958	0.8617	0.9589	0.9920	0.8214	0.9330	0.9829
N=5, $N' = 2$	100	1000	10000	100	1000	10000	100	1000	10000
firm 1*	0.9602	0.9959	0.9996	0.9662	0.9965	0.9998	0.9643	0.9971	0.9998
firm 2*	0.9624	0.9962	0.9996	0.9575	0.9958	0.9995	0.9627	0.9962	0.9997
firm 3*	0.9583	0.9955	0.9996	0.9569	0.9954	0.9994	0.9601	0.9963	0.9995
firm 4	0.9102	0.9794	0.9957	0.9075	0.9755	0.9941	0.8895	0.9530	0.9883
firm 5	0.9258	0.9837	0.9959	0.9069	0.9739	0.9930	0.8951	0.9573	0.9874

* indicates the firms with knowledge of the demand function.

Comparing each entry in Table 2 with the corresponding entry in Table 1, when the number of competing firms is the same, using the modified DDEP causes the fraction of optimal revenue for firms without knowledge of demand to increase. This verifies what we previously illustrated in Theorems 3 and 5. Since there exist some firms with knowledge of the demand functions, fewer firms actually participate in price experimentation. As the number of firms with knowledge of the demand functions increases, there are two significant observations:

1. The fraction of optimal revenue for firms with knowledge of the demand is not close to or equal to one.
2. The fraction of optimal revenue for firms without knowledge of the demand will increase.

Step 3 in the modified DDEP is the underlying reason for the first observation, even if the pricing of firms with knowledge of the demand curve are passive to reprice and their pricing does not affect the pricing of the firms without knowledge of the demand function, the firms with the knowledge of the demand curves still price according to each firm's respective best response function. The pricing established in Step 3 will be used in the next stage, while the optimal price will be changed once competitors have adjusted prices. In the first stage, the initial price of each firm is selected randomly in the price interval, such that the optimal price p_1^{i*} in the first stage may be far away from the initial price. Hence, the fraction of optimal revenue for firms with knowledge of the demand function is neither close nor equal to one.

The reason for the second observation is that fewer parameters need to be estimated. Turning our attention to the cases when N' is fixed, one pertinent observation is that the change in N exerts significant effects on the fraction of optimal revenue. The numerical results are consistent with our analysis.

Remark. A phenomenon which appears numerically but is not theoretically guaranteed is that the fraction of optimal revenue for firms with knowledge of the demand curve increases as the number of firms without knowledge of the demand curve decreases. An explanation for this numerical phenomenon is that in our setting, the number of firms participating in price experimentation also affects regret. For different experiments, such as the number of companies participating in the price experimentation, the number of periods in one stage is also different.

For example, there are 18 periods in the first stage when $N' = 2$, while there are 32 periods when $N' = 3$. This means that when we want to compare these results, the number of finishing stages is not the same as in the case of a fixed time interval T. The smaller the number of firms participating in price experimentation, the more stages required before the experimentation finishes.

6 Conclusions

In this paper, we have studied a non-cooperative game among the firms each seeking to maximize revenue. The underlying demand-price information is not known *a priori* for each firm and the demand response is determined by consumers and competitors. Based on past observations, estimated demand models are used to make pricing decisions. The equilibrium pricing algorithm is developed for such a system and is combined with estimation and optimization cycles. The estimation cycle uses a linear regression model to estimate the linear approximation of the underlying demand model at a previously given price. We show that this converges to the equilibrium pricing if the algorithm is designed properly. The role of the algorithm is to guarantee the convergence of pricing decisions. We also conduct an analysis of the convergence rate of regret, which is unlike worst-case regret used elsewhere in the literature. Regret is defined, in our case, as a comparison with the optimal price that the firm could make when full information is available.

We use a linear approximation to estimate the unknown demand function. We analyze the linear demand model in two general scenarios: one being when all competing firms have unknown demand functions, the other being when a subset of demand functions are unknown. In the first scenario, the linear demand model captures the impact of all competitors. In the second scenario (partially clairvoyant model), only the impact of competitors with unknown demand functions is captured. We show that the linear demand model in both scenarios ensures that pricing decisions converge to the true equilibrium pricing as a result of the algorithm operation cycle. The analysis and numerical results show that the regret of firms without knowledge of the demand curve is associated with the number of such firms, while the regret of the firms with knowledge of the demand curve is not.

Since our work focuses on the pricing decisions of firms in a competitive environment, we do not consider limitations in market size. In the future, this aspect of could be considered jointly with pricing and inventory management, adding additional complexity to the problem but making it more realistic.

Acknowledgements. We would like to thank Yue Dai, Chiao-Wei Li, Wen-Ying Huang, and Ling-Wei Wang for useful discussions.

References

1. Bertsimas, D., and Perakis, G. (2006). Dynamic pricing: A learning approach. In *Mathematical and computational models for congestion charging* (pp. 4579). Springer.

2. Besbes, O., and Zeevi, A. (2009). Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Operations Research*, 57, 1407–1420.
3. Besbes, O., and Zeevi, A. (2015). On the (surprising) sufficiency of linear models for dynamic pricing with demand learning. *Management Science*, 61, 723–739.
4. den Boer, A. V., and Zwart, B. (2014). Simultaneously learning and optimizing using controlled variance pricing. *Management science*, 60, 770–783.
5. Broder, J., and Rusmevichientong, P. (2012). Dynamic pricing under a general parametric choice model. *Operations Research*, 60, 965–980.
6. Chen, B., Chao, X., and Ahn, H.-S. (2019). Coordinating pricing and inventory replenishment with nonparametric demand learning. *Operations Research*, 67, 1035–1052.
7. Chen, B., Chao, X., and Shi, C. (2015). Nonparametric algorithms for joint pricing and inventory control with lost-sales and censored demand. *Mathematics of Operations Research*, Major Revision.
8. Cooper, W. L., Homem-de Mello, T., and Kleywegt, A. J. (2015). Learning and pricing withmodels that do not explicitly incorporate competition. *Operations research*, 63, 86–103.
9. Fisher, M., Gallino, S., and Li, J. (2018). Competition-based dynamic pricing in online re-tailing: A methodology validated with field experiments. *Management Science*, 64, 2496–2514.
10. Gallego, G., and Hu, M. (2014). Dynamic pricing of perishable assets under competition. *Management Science*, 60, 1241–1259.
11. Gallego, G., and Talebian, M. (2012). Demand learning and dynamic pricing for multi-version products. *Journal of Revenue and Pricing Management*, 11, 303–318.
12. Kachani, S., Perakis, G., and Simon, C. (2007). Modeling the transient nature of dynamic pricing with demand learning in a competitive environment. In *Network science, nonlinear science and infrastructure systems* (pp. 223–267). Springer.
13. Keskin, N. B., and Zeevi, A. (2014). Dynamic pricing with an unknown demand model: Asymptotically optimal semi-myopic policies. *Operations Research*, 62, 1142–1167.
14. Kwon, C., Friesz, T. L., Mookherjee, R., Yao, T., and Feng, B. (2009). Non-cooperative competition among revenue maximizing service providers with demand learning. *European Journal of Operational Research*, 197, 981–996.
15. Levin, Y., McGill, J., and Nediak, M. (2009). Dynamic pricing in the presence of strategic consumers and oligopolistic competition. *Management science*, 55, 32–46.

A Proof of Proposition 2

Proof. We first analyze the convergences of the estimated parameters. Define that

$$W_n^{i,1} = \frac{1}{I_n}\sum_{t=t_n+1}^{t_n+I_n} \varepsilon_t^i, \quad W_n^{i,2} = \frac{1}{I_n}\sum_{t=t_n+I_n+1}^{t_n+2I_n} \varepsilon_t^i, \cdots, W_n^{i,m} = \frac{1}{I_n}\sum_{t=t_n+(m-1)I_n+1}^{t_n+mI_n} \varepsilon_t^i,$$

for $m = 1, 2, ..., N+1$. The recursion for the decisions can be written as

$$\hat{\mathbf{p}}_{n+1} = \mathbf{z}(\hat{\mathbf{p}}_n) + C_n, \quad \text{with} \quad C_n = \hat{\mathbf{z}} - \mathbf{z}(\hat{\mathbf{p}}_n),$$

where $\hat{\mathbf{z}}$ denotes the collection of all firms $z^i(\hat{\alpha}_{n+1}^i, \hat{\beta}_{n+1}^{ij})$, which is exactly $\hat{\mathbf{p}}_{n+1}$. To better describe the relationship between parameter convergence and ε^i, we denote the estimators of the noiseless case, as in equations (6), (7), and (8) by $\breve{\beta}^{ii}_{n+1}$, $\breve{\beta}^{ij}_{n+1}$ and $\breve{\alpha}^{i}_{n+1}$ respectively. Hence, we obtain that

$$\begin{aligned}
\hat{\beta}^{ii}_{n+1} &= -\frac{\lambda^i(\hat{p}^i_n+\delta_n,\hat{p}^{-i}_n)-\lambda^i(\hat{\mathbf{p}}_n)}{\delta_n} - \frac{1}{\delta_n}\left(W^{i,j}_n - W^{i,1}_n\right)\\
&= -\left[\nabla_{p^i}\lambda^i(\mathbf{p}_t)+\frac{1}{2}\nabla^2_{p^i}\lambda^i(\mathbf{p}')\delta_n\right] - \frac{1}{\delta_n}\left(-W^{i,1}_n + W^{i,j}_n\right)\\
&= \breve{\beta}^{ii}(\hat{p}^i_n) + C^i_{n,\beta^{ii}},\\
\hat{\beta}^{ij}_{n+1} &= \frac{\lambda^i(\hat{p}^j_n+\delta_n,\hat{p}^{-j}_n)-\lambda^i(\hat{\mathbf{p}}_n)}{\delta_n} - \frac{1}{\delta_n}\left(W^{i,j}_n - W^{i,1}_n\right)\\
&= \left[\nabla_{p^j}\lambda^i(\mathbf{p}_t)+\frac{1}{2}\nabla^2_{p^j}\lambda^i(\mathbf{p}')\delta_n\right] - \frac{1}{\delta_n}\left(W^{i,j}_n - W^{i,1}_n\right)\\
&= \breve{\beta}^{ij}(\hat{p}^j_n) + C^i_{n,\beta^{ij}},
\end{aligned}$$

where the second equality follows from Taylors theorem applied to λ^i_n with $\mathbf{p'} \equiv (p^i, p^{-i})$, $p^i \in [\hat{p}^i_n, \hat{p}^i_n + \delta_n]$

$$\begin{aligned}
C^i_{n,\beta^{ii}} &= -\frac{1}{2}\nabla^2_{p^i}(\mathbf{p}')\delta_n - \frac{1}{\delta_n}\left(-W^{i,1}_n + W^{i,j}_n\right),\\
C^i_{n,\beta^{ij}} &= \frac{1}{2}\nabla^2_{p^j}(\mathbf{p}')\delta_n - \frac{1}{\delta_n}\left(+W^{i,j}_n - W^{i,1}_n\right).
\end{aligned}$$

Similarly, we can obtain that

$$\begin{aligned}
\hat{\alpha}^i_{n+1} &= \bar{\lambda}^i + \hat{\beta}^i_{n+1}\bar{p}^i_n - \sum_{j,j\neq i}^{N} \hat{\beta}^{ij}_{n+1}\bar{p}^j_n\\
&= \frac{1}{N+1}\sum_{t=t_n+1}^{t_{n+1}} \lambda^i(\mathbf{p}_t) + \frac{1}{N+1}\left(\sum_{m=1}^{m=N+1} W^{i,m}_n\right) - \hat{\beta}^i_{n+1}\left(\hat{p}^i_n + \frac{\delta_n}{2}\right) + \sum_{j,j\neq i}^{N} \hat{\beta}^{ij}_{n+1}\left(\hat{p}^j_n + \frac{\delta_n}{2}\right)\\
&= \frac{1}{N+1}\sum_{t=t_n+1}^{t_{n+1}} \lambda^i(\hat{\mathbf{p}}_n) - \nabla_{p^i}\lambda^i(\hat{\mathbf{p}_n})\hat{p}^i_n + \sum_{j=1,j\neq i}^{N} \nabla_{p^j}\lambda^i(\hat{\mathbf{p}_n})\hat{p}^j_n\\
&+ \frac{1}{2}\sum_{j}^{N}\nabla_{p^j}\lambda^i(\mathbf{p}'')\delta_n - \left[\frac{1}{2}\nabla_{p^i}\lambda^i(\hat{\mathbf{p}}_n)\delta_n + \frac{1}{2}\sum_{j=1,j\neq i}^{N}\nabla_{p^j}\lambda^i(\hat{\mathbf{p}}_n)\delta_n\right]\\
&+ \frac{1}{N+1}\left(\sum_{m=1}^{m=N+1} W^{i,m}_n\right) + C^i_{n,\beta^{ii}}\left(\hat{p}^i_n + \frac{\delta_n}{2}\right) + \sum_{j=1,j\neq i}^{N} C^i_{n,\beta^{ij}}\left(\hat{p}^i_n + \frac{\delta_n}{2}\right)\\
&= \breve{\alpha}^i(\hat{\mathbf{p}}_n) + C^i_{n,\alpha},
\end{aligned}$$

where the third equality follows from Taylors theorem applied to λ^i_n with $\mathbf{p''} \equiv (p^i, p^{-i})$, where $p^i \in [\hat{p}^i_n, \hat{p}^i_n + \delta_n]$, we also obtain

$$C^i_{n,\alpha^i} = \frac{1}{2}\sum_{j}^{N}\nabla_{p^j}\lambda^i(\mathbf{p}'')\delta_n - \left[\frac{1}{2}\nabla_{p^i}\lambda^i(\hat{\mathbf{p}}_n)\delta_n + \frac{1}{2}\sum_{j=1,j\neq i}^{N}\nabla_{p^j_n}\lambda^i(\hat{\mathbf{p}}_n)\delta_n\right]$$

$$+ \frac{1}{N+1}\left(\sum_{m=1}^{m=N+1} W^{i,m}_n\right) + C^i_{n,\beta^{ii}}(\hat{p}^i_n + \frac{\delta_n}{2}) + \sum_{j=1,j\neq i}^{N} C^i_{n,\beta^{ij}}(\hat{p}^i_n + \frac{\delta_n}{2}).$$

Considering event ω of specific occurrences for random variables ε^i_t and thus $W^{i,m}_n$, we define a collection of "good events" in which the accumulated impact of noise on demand is below a given ceiling x, which may be set later, as follows:

$$\mathscr{A}^i_n = \left\{\omega : \left|W^{i,m}_n\right| \leq x,\ m = 1, \cdots, N+1\right\}, \quad \forall i.$$

We define the following constants:

$$K^i_{11} = \min_{p^i\in\mathcal{P}^i}\left|\lambda^i(\mathbf{p})\right|, \quad K^i_{12} = \min_{p^i\in\mathcal{P}^i}\left|\nabla_{p^i}\lambda^i(\mathbf{p})\right|, \quad K^i_{13} = \min_{p^j\in\mathcal{P}^i}\left|\nabla_{p^j}\lambda^i(\mathbf{p})\right| \quad \forall i,j$$

$$K^i_{14} = \max_{p^i\in\mathcal{P}^i}\left|\nabla_{p^i}\lambda^i(\mathbf{p})\right|, \quad K^i_{15} = \max_{p^j\in\mathcal{P}^i}\left|\nabla_{p^j}\lambda^i(\mathbf{p})\right|, \quad \forall i,j, \text{ and}$$

$$K^i_{16} = \max_{p^i\in\mathcal{P}^i}\left|\nabla^2_{p^i}\lambda^i(\mathbf{p})\right|, \quad K^i_{17} = \max_{p^j\in\mathcal{P}^i}\left|\nabla^2_{p^j}\lambda^i(\mathbf{p})\right|, \quad \forall i,j.$$

Let $\mathit{\Gamma}^{i,1}_n = \frac{1}{2}K^i_{16}\delta_n + 2x\delta_n^{-1}$, $\mathit{\Gamma}^{j,1}_n = \frac{1}{2}K^i_{17}\delta_n + 2x\delta_n^{-1}$ and $\mathit{\Gamma}^{i,2}_n = \frac{1}{2}K^i_{14}\delta_n + \sum_{j=1,j\neq i}^{N}\frac{1}{2}K^i_{15}\delta_n + x\delta_n^{-1} + \mathit{\Gamma}^{i,1}_n\left(\hat{p}^{i,h}_n + \frac{\delta_n}{2}\right) + \sum_{j=1,j\neq i}^{N}\mathit{\Gamma}^{j,1}_n\left(\hat{p}^{j,h}_n + \frac{\delta_n}{2}\right)$ be the upper bound of $C^i_{n,\beta^{ii}}, C^i_{n,\beta^{ij}}$ and C^i_{n,α^i}. As n increases, $\mathit{\Gamma}^{i,1}_n$, $\mathit{\Gamma}^{j,1}_n$ and $\mathit{\Gamma}^{i,2}_n$ converge to zero. This dictates that the upper bound of the difference between $\breve{\alpha}^i$ and $\hat{\alpha}^i_n$, $\breve{\beta}^{ii}$ and $\hat{\beta}^{ii}_n$, $\breve{\beta}^{ij}$ and $\hat{\beta}^{ij}_n$ will converge to zero when n grows. Hence, we suppose that the upper bound of this difference will converge to zero at some point. Let $n_0 = \min\left\{n \geq 0 : \mathit{\Gamma}^{i,1}_n < \frac{K^i_{12}}{2}, \mathit{\Gamma}^{j,1}_n < \frac{K^i_{13}}{2} \text{ and } \mathit{\Gamma}^{i,2}_n < \frac{K^i_{11}}{2}\right\}$. For $n \geq n_0$ and $\omega \in \mathscr{A}^i_n$, $\left|C^i_{n,\beta^{ii}}\right| < \frac{K^i_{12}}{2}$, $\left|C^i_{n,\beta^{ij}}\right| < \frac{K^i_{13}}{2}$ and $\left|C^i_{n,\alpha^i}\right| < \frac{K^i_{11}}{2}$. Based on the previous assumption about z^i, we obtain

$$C^i_n = z^i\left(\breve{\alpha}^i(\hat{p}^i_n) + C^i_{n,\alpha^i}, \breve{\beta}^{ii}(\hat{p}^i_n) + C^i_{n,\beta^{ii}}, \breve{\beta}^{ij}(\hat{p}^i_n) + C^i_{n,\beta^{ij}}\right) - z^i\left(\breve{\alpha}^i(\hat{p}^i_n), \breve{\beta}^{ii}(\hat{p}^i_n), \breve{\beta}^{ij}(\hat{p}^i_n)\right)$$

$$= z^i_{\alpha^i}\left(a^{i,1}_n, \cdots, a^{i,N}_n\right)C^i_{n,\alpha^i} + z^i_{\beta^{ii}}\left(a^{i,1}_n, \cdots, a^{i,N}_n\right)C^i_{n,\beta^{ii}} + \sum_{j=1,j\neq i}^{N} z^i_{\beta^{ij}}\left(a^{i,1}_n, \cdots, a^{i,N}_n\right)C^i_{n,\beta^{ij}},$$

for some $\left(a^{i,1}_n, \cdots, a^{i,N}_n\right)$ on the line segment joining $\left(\breve{\alpha}^i(\hat{p}^i_n), \breve{\beta}^{ii}(\hat{p}^i_n), \breve{\beta}^{ij}(\hat{p}^i_n)\right)$ to $\left(\breve{\alpha}^i(\hat{p}^i_n) + C^i_{n,\alpha^i}, \breve{\beta}^{ii}(\hat{p}^i_n) + C^i_{n,\beta^{ii}}, \breve{\beta}^{ij}(\hat{p}^i_n) + C^i_{n,\beta^{ij}}\right)$. For some positive K^i_{18}, with a probability lower bond in terms of x

$$\left|C^i_n\right| \leq NK^i_{18}\left[\delta_n + \frac{x}{\delta_n}\right], \tag{19}$$

Thus, by taking the norm of the recursive decision function *with a probability lower bound in terms of* x, as set out in Proposition 3, we obtain

$$\|\hat{\mathbf{p}}_{n+1} - \mathbf{z}(\hat{\mathbf{p}}_n)\| \leq \|C_n\|, \tag{20}$$

where $C_n = \left[C^1_n\, C^2_n\, \cdots\, C^N_n\right]^T$, and C^i_n can be set at $NK^i_{18}\left[\delta_n + \frac{x}{\delta_n}\right]$ according to (19).

B Proof of Proposition 3

Proof. To start our analysis, we first analyse the probability for a firm obtaining "good" estimators. According to the definition of the difference $C^i_{n,\beta^{ij}}$ and the Hoeffding inequality, to measure the effectiveness of $\hat{\beta}^i_{n+1}$, we have the following bound:

$$\mathbb{P}\left\{\left|\hat{\beta}^{ii}_{n+1} - \breve{\beta}^{ii}_{n+1}(\hat{p}^i_n)\right| \leq K^i_{19}\left(\delta_n + \frac{x}{\delta_n}\right)\right\} \geq 1 - 4e^{-2I_n x^2}, \forall i. \tag{21}$$

A similar argument shows that

$$\mathbb{P}\left\{\left|\hat{\beta}^{ij}_{n+1} - \breve{\beta}^{ij}_{n+1}(\hat{p}^j_n)\right| \leq K^j_{19}\left(\delta_n + \frac{x}{\delta_n}\right)\right\} \geq 1 - 4e^{-2I_n x^2}, \forall j. \tag{22}$$

By the definition of the difference C^i_{n,α^i}, we obtain

$$\begin{aligned}
&\mathbb{P}\left\{\left|\hat{\alpha}^i_{n+1} - \breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n)\right| \leq K_{20} N\left(\delta_n + \frac{x}{\delta_n} + x\right)\right\} \\
\geq &\mathbb{P}\left\{\left|-W^{i,i}_{n+1} + W^{i,1}_{n+1}\right| \leq 2x, \cdots, \left|W^{i,N+1}_{n+1} - W^{j,1}_{n+1}\right| \leq 2x, \left|\sum_{m=1}^{N+1} W^{i,m}_{n+1}\right| \leq 2x\right\} \\
\geq &1 - 8Ne^{-2I_n x^2}, \forall i.
\end{aligned} \tag{23}$$

We also have

$$\mathbb{P}\left\{\left|\hat{\beta}^{ii}_{n+1} - \breve{\beta}^{ii}_{n+1}(\hat{p}^i_n + \delta_n)\right| \leq K^i_{21}\left(\delta_n + \frac{x}{\delta_n}\right)\right\} \geq 1 - 4e^{-2I_n x^2}, \forall i. \tag{24}$$

$$\mathbb{P}\left\{\left|\hat{\beta}^{ij}_{n+1} - \breve{\beta}^{ij}_{n+1}(\hat{p}^j_n + \delta_n)\right| \leq K^j_{21}\left(\delta_n + \frac{x}{\delta_n}\right)\right\} \geq 1 - 4e^{-2I_n x^2}, \forall i, \text{and} \tag{25}$$

$$\mathbb{P}\left\{\left|\hat{\alpha}^i_{n+1} - \breve{\alpha}^i_{n+1}\left(\hat{\mathbf{p}}_n + \mathit{\Delta}_n\right)\right| \leq NK_{22}\left(\delta_n + \frac{x}{\delta_n} + x\right)\right\} \geq 1 - 8Ne^{-2I_n x^2}, \forall i, \tag{26}$$

where $\hat{\mathbf{p}}_n + \mathit{\Delta}_n$ denotes the price when some firm changes their price to $\hat{p}^i_n + \delta_n$ in price experimentation. Taking the intersection of the inequalities above, we obtain

$$\begin{aligned}
\mathbb{P}\Bigg\{ &\left|\hat{\alpha}^i_{n+1} - \breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n)\right| \leq NK_{20}\left(\delta_n + \frac{x}{\delta_n} + x\right), \left|\hat{\beta}^{i1}_{n+1} - \breve{\beta}^{i1}_{n+1}(\hat{p}^1_n)\right| \leq K^1_{19}\left(\delta_n + \frac{x}{\delta_n}\right) \\
&\cdots \left|\hat{\beta}^{iN}_{n+1} - \breve{\beta}^{iN}_{n+1}(\hat{p}^N_n)\right| \leq K^N_{19}\left(\delta_n + \frac{x}{\delta_n}\right), \left|\hat{\alpha}^i_{n+1} - \breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n + \mathit{\Delta}_n)\right| \leq K_{22}\left(\delta_n + \frac{x}{\delta_n} + x\right), \\
&\left|\hat{\beta}^{i1}_{n+1} - \breve{\beta}^{i1}_{n+1}(\hat{p}^1_n + \delta_n)\right| \leq K^1_{21}\left(\delta_n + \frac{x}{\delta_n}\right) \cdots \left|\hat{\beta}^{iN}_{n+1} - \breve{\beta}^{iN}_{n+1}(\hat{p}^N_n + \delta_n)\right| \leq K^N_{21}\left(\delta_n + \frac{x}{\delta_n}\right) \Bigg\} \\
\geq &\ 1 - 24Ne^{-2I_n x^2}, \forall i.
\end{aligned} \tag{27}$$

As the quantities on both sides of probability event in (21) - (23) are nonnegative, (21) - (23) imply that

$$\mathbb{P}\left\{\left|\hat{\beta}^{ij}_{n+1} - \breve{\beta}^{ij}_{n+1}(\hat{p}^i_n)\right|^2 \leq K^i_{23}\left(\delta^2_n + \frac{x^2}{\delta^2_n}\right)\right\} \geq 1 - 4e^{-2I_n x^2}, \forall i, j, \text{and} \tag{28}$$

$$\mathbb{P}\left\{\left|\hat{\alpha}^i_{n+1}-\breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n)\right|^2 \le NK_{24}\left(\delta_n^2+\frac{x^2}{\delta_n^2}+x^2\right)\right\} \ge 1-8Ne^{-2I_n x^2}, \forall i. \tag{29}$$

Similar implications of (24) - (26) are

$$\mathbb{P}\left\{\left|\hat{\beta}^{ij}_{n+1}-\breve{\beta}^{ij}_{n+1}(\hat{p}^j_n+\delta_n)\right|^2 \le K^j_{25}\left(\delta_n^2+\frac{x^2}{\delta_n^2}\right)\right\} \ge 1-4e^{-2I_n x^2}, \forall i,j \text{ and} \tag{30}$$

$$\mathbb{P}\left\{\left|\hat{\alpha}^i_{n+1}-\breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n+\mathit{\Delta}_n)\right|^2 \le NK_{26}\left(\delta_n^2+\frac{x^2}{\delta_n^2}+x^2\right)\right\} \ge 1-8Ne^{-2I_n x^2}, \forall i. \tag{31}$$

Combining (28) - (31), we obtain

$$\begin{aligned}&\mathbb{P}\Bigg\{\left|\hat{\alpha}^i_{n+1}-\breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n)\right|^2+\sum_{j=1}^N\left|\hat{\beta}^{ij}_{n+1}-\breve{\beta}^{ii}_{n+1}(\hat{p}^i_n)\right|^2+\left|\hat{\alpha}^i_{n+1}-\breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n+\mathit{\Delta}_n)\right|^2\\&+\sum_{j=1}^N\left|\hat{\beta}^{ij}_{n+1}-\breve{\beta}^{ij}_{n+1}(\hat{p}^j_n+\delta_n)\right|^2 \le K_{27}N\left(\delta_n^2+\frac{x^2}{\delta_n^2}+x^2\right)\Bigg\}\\&\ge 1-24Ne^{-2I_n x^2}, \forall i,\end{aligned} \tag{32}$$

which is equivalently rewritten as

$$\begin{aligned}&\mathbb{P}\Bigg\{\left(\frac{NK_{28}}{\delta_n^2}+NK_{29}\right)^{-1}\Bigg(\left|\hat{\alpha}^i_{n+1}-\breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n)\right|^2+\sum_{j=1}^N\left|\hat{\beta}^{ij}_{n+1}-\breve{\beta}^{ii}_{n+1}(\hat{p}^i_n)\right|^2+\left|\hat{\alpha}^i_{n+1}-\breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n+\mathit{\Delta}_n)\right|^2\\&+\sum_{j=1}^N\left|\hat{\beta}^{ij}_{n+1}-\breve{\beta}^{ij}_{n+1}(\hat{p}^j_n+\delta_n)\right|^2-NK_{27}\delta_n^2\Bigg)\ge x^2\Bigg\}\\&< 24Ne^{-2I_n x^2}, \forall i.\end{aligned} \tag{33}$$

Let $x=\sigma(2\log(I_n))^{\frac{1}{2}}I_n^{-\frac{1}{2}}$, we integrate both sides to obtain

$$\begin{aligned}&\mathbb{E}\Bigg[\left(\frac{NK_{28}}{\delta_n^2}+NK_{29}\right)^{-1}\Bigg(\left|\hat{\alpha}^i_{n+1}-\breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n)\right|^2+\sum_{j=1}^N\left|\hat{\beta}^{ij}_{n+1}-\breve{\beta}^{ii}_{n+1}(\hat{p}^i_n)\right|^2+\left|\hat{\alpha}^i_{n+1}-\breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n+\mathit{\Delta}_n)\right|^2\\&+\sum_{j=1}^N\left|\hat{\beta}^{ij}_{n+1}-\breve{\beta}^{ij}_{n+1}(\hat{p}^j_n+\delta_n)\right|^2-NK_{27}\delta_n^2\Bigg)\Bigg]\\&\le\int_0^{+\infty}24Ne^{-2I_n x^2}dx=\frac{12NK_{30}}{I_n}, \forall i.\end{aligned} \tag{34}$$

Hence, we have that

$$\begin{aligned}
&\mathbb{E}\Bigg[\left|\hat{\alpha}^i_{n+1}-\breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n)\right|^2+\sum_{j=1}^{N}\left|\hat{\beta}^{ij}_{n+1}-\breve{\beta}^{ii}_{n+1}(\hat{p}^i_n)\right|^2+\left|\hat{\alpha}^i_{n+1}-\breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n+\varDelta_n)\right|^2\\
&+\sum_{j=1}^{N}\left|\hat{\beta}^{ij}_{n+1}-\breve{\beta}^{ij}_{n+1}(\hat{p}^j_n+\delta_n)\right|^2\Bigg]\\
&\leq\left(\frac{12NK_{30}}{I_n}+\left(\frac{NK_{28}}{\delta_n^2}+NK_{29}\right)^{-1}NK_{27}\delta_n^2\right)\left(\frac{NK_{28}}{\delta_n^2}+NK_{29}\right)\\
&\leq NK_{31}I_n^{-\frac{1}{2}},\forall i,
\end{aligned}\tag{35}$$

where the first inequality is obtained by taking the constant out of the expectation, and the second inequality is obtained by letting $\delta=I_n^{-\frac{1}{4}}$.

We must define a random event to represent the situation when the estimation is sufficiently good. Let $\mathscr{A}^i_n$ denote the event where all the estimated parameters satisfy $\Big\{|\hat{\alpha}^i_{n+1}-\breve{\alpha}^i|\leq\delta_n+x/\delta_n+x, |\hat{\beta}^{ij}_{n+1}-\breve{\beta}^{ij}|\leq\delta_n+x/\delta_n+x,\ \forall j=1,\cdots,N\Big\}$ for firm i. The probability of this event happening is shown in (27), which can be transformed into

$$\mathbb{P}\left\{\mathscr{A}^i_n\right\}\geq 1-24Ne^{-2I_nx^2}.$$

Let $\mathscr{B}_n=\{\mathscr{A}^i_n \text{ occurs},\ i=1,\cdots,N\}$ represent the random event where all N firms obtain "good" estimation outcomes $\mathscr{A}^i_n$. Let $\mathscr{B}^c_n=\left\{\text{at least one } \mathscr{A}^i_n \text{ does not occur}\right\}$.

$$\mathbb{E}\left[\|\hat{\mathbf{p}}_n-\mathbf{p}^*\|^2\right]\leq\mathbb{E}\left[\|\hat{\mathbf{p}}_n-\mathbf{p}^*\|^2\mid\mathscr{B}_n\right]\mathbb{P}\left\{\mathscr{B}_n\right\}+\mathbb{E}\left[\|\hat{\mathbf{p}}_n-\mathbf{p}^*\|^2\mid\mathscr{B}^c_n\right]\mathbb{P}\left\{\mathscr{B}^c_n\right\}.$$

Denote index set $M_k=\left\{i|\mathscr{A}^i_n \text{ occurs in the } k\text{th repeated permutation}\right\}$ to collect the indices i's such that $\mathscr{A}^i_n$ occurs in the kth repeated permutation. Note that the number of index sets M_k's is equal to 2^N-1. Let event $\mathscr{G}^{M_k}_n=\{\mathscr{A}^i_n \text{ occurs } \forall i\in M_k \text{ and } \mathscr{A}^{jc}_n \text{ occurs } \forall j\notin M_k\}$. Thus, we obtain that

$$\mathbb{P}\left\{\mathscr{B}^c_n\right\}=\mathbb{P}\left\{\bigcup_{k=1}^{2^N-1}\mathscr{G}^{M_k}_n\right\}=\sum_{k=1}^{2^N-1}\mathbb{P}\left\{\mathscr{G}^{M_k}_n\right\}\leq\sum_{1}^{N}\mathbb{P}\left\{\mathscr{A}^{ic}_n\right\}.$$

Hence, we obtain

$$\mathbb{E}\,\|\hat{\mathbf{p}}_n-\mathbf{p}^*\|^2\leq\mathbb{E}\left[\|\hat{\mathbf{p}}_n-\mathbf{p}^*\|^2\mid\mathscr{B}_n\right]\mathbb{P}\{\mathscr{B}_n\}+\left\|\mathbf{p}^h-\mathbf{p}^l\right\|^2\sum_{1}^{N}\mathbb{P}\left\{\mathscr{A}^{ic}_n\right\}.$$

For some suitably large positive constant K_{30}, we can obtain the probability bound of a bad event occurring in the estimation of the demand function, which is

$$\mathbb{P}\left\{\mathscr{B}_n{}^c\right\}\leq 24N^2e^{-2I_nx^2}\leq\frac{24N^2K_{32}}{I_n}.\tag{36}$$

Therefore, combining the results of (35) and (36), it follows that

$$
\begin{aligned}
&\mathbb{E}\left[\|\mathbf{z}-\hat{\mathbf{p}}_{n+1}\|^2\right] \\
&\leq \mathbb{E}\Bigg[K_{33}^2 \sum_{i=1}^{N}\Bigg(\left|\hat{\alpha}_{n+1}^{i}-\breve{\alpha}_{n+1}^{i}(\hat{\mathbf{p}}_n)\right| + \sum_{j=1}^{N}\left|\hat{\beta}_{n+1}^{ij}-\breve{\beta}_{n+1}^{ii}(\hat{p}_n^i)\right| \\
&+\left|\hat{\alpha}_{n+1}^{i}-\breve{\alpha}_{n+1}^{i}(\hat{\mathbf{p}}_n+\varDelta_n)\right| + \sum_{j=1}^{N}\left|\hat{\beta}_{n+1}^{ij}-\breve{\beta}_{n+1}^{ij}(\hat{p}_n^j+\delta_n)\right|\Bigg)^2\Bigg] + \frac{24N^2K_{32}}{I_n}\left\|\mathbf{p}^h-\mathbf{p}^l\right\|^2 \\
&\leq \mathbb{E}\Bigg[K_{34} \sum_{i=1}^{N}\Bigg(\left|\hat{\alpha}_{n+1}^{i}-\breve{\alpha}_{n+1}^{i}(\hat{\mathbf{p}}_n)\right|^2 + \sum_{j=1}^{N}\left|\hat{\beta}_{n+1}^{ij}-\breve{\beta}_{n+1}^{ii}(\hat{p}_n^i)\right|^2 \\
&+\left|\hat{\alpha}_{n+1}^{i}-\breve{\alpha}_{n+1}^{i}(\hat{\mathbf{p}}_n+\varDelta_n)\right|^2 + \sum_{j=1}^{N}\left|\hat{\beta}_{n+1}^{ij}-\breve{\beta}_{n+1}^{ij}(\hat{p}_n^j+\delta_n)\right|^2\Bigg)\Bigg] + \frac{24N^2K_{32}}{I_n}\left\|\mathbf{p}^h-\mathbf{p}^l\right\|^2 \\
&\leq N^2K_1I_n^{-\frac{1}{2}}. \qquad (37)
\end{aligned}
$$

In the first inequality, the first term of the summation stems from rewriting the ℓ-2 norm and the second term is a rough bound of the expectation under the event $\mathscr{B}_n^c$. The second inequality is derived from (35) for some proper constant K_{34}. Selecting some suitable constant K_1, we obtain the final inequality.

C Proof of Theorem 2

Proof. We first note that constants introduced depend on σ. Fix a time horizon T and firm i, and let $k = \inf\left\{h \geq 1 : (N+1)\sum_{i=1}^{h} I_n \geq T\right\}$, in which k satisfies $\log_v\left(\frac{v-1}{(N+1)I_0 v}T+1\right) \leq k < \log_v\left(\frac{v-1}{(N+1)I_0 v}T+1\right)+1$ based on the summation of a geometric sequence. The expected revenue difference after T periods is therefore given by

$$D^i = \mathbb{E}\left[\sum_{t=1}^{T}\left[r_i(p^{i*},p^{-i*}) - r_i(p_t^i,p_t^{-i})\right]\right].$$

According to the assumption, the revenue function r^i is a concave function on $(p^{i,l},p^{i,h})$ and twice differentiable. Applying a Taylor approximation for $r_i(p^i,p^{-i})$ around $\mathbf{p}^*$,

$$
\begin{aligned}
r^i(p^i,p^{-i}) &= r^i(p^{i*},p^{-i*}) + \nabla r^i(p^{i*},p^{-i*})^T(\mathbf{p}-\mathbf{p}^*) \\
&+\frac{1}{2}(\mathbf{p}-\mathbf{p}^*)^T\nabla^2 r^i(p^{i*},p^{-i*})(\mathbf{p}-\mathbf{p}^*).
\end{aligned}
$$

Note that $\nabla r_i(p^{i*},p^{-i*})^T = \left[p^{i*}\nabla_{p^1}\lambda^i(\mathbf{p}^*),\cdots,p^{i*}\nabla_{p^i}\lambda^i(\mathbf{p}^*)+\lambda^i(\mathbf{p}^*),\cdots,p^{i*}\nabla_{p^N}\lambda^i(\mathbf{p}^*)\right]$, where $p^{i*}\nabla_{p^i}\lambda^i(\mathbf{p}^*)+\lambda^i(\mathbf{p}^*) = 0$ at the maximum revenue. Since p^{i*} is a maximizer of $r_i(p^{i*},p^{-i*})$, $\left|\nabla r_i(p^{i*},p^{-i*})\right|$ can be bounded by some positive constant K_{35} and $\frac{1}{2}\left\|\nabla^2 r_i(p^{i*},p^{-i*})\right\|$ can be bounded by some positive constant K_{36}. Notice that we cannot tell whether the term $\nabla r^i(p^{i*},p^{-i*})^T(\mathbf{p}-\mathbf{p}^*)$ is lesser or greater than $\frac{1}{2}(\mathbf{p}-\mathbf{p}^*)^T\nabla^2 r^i(p^{i*},p^{-i*})(\mathbf{p}-\mathbf{p}^*)$. We have

$$
\begin{aligned}
\left|r^i(p^i,p^{-i}) - r^i(p^{i*},p^{-i*})\right| &\leq K_{35}\left|\mathbf{p}-\mathbf{p}^*\right| + K_{36}\left\|\mathbf{p}^*-\mathbf{p}\right\|^2 \\
&\leq K_{37}\left[\left|\mathbf{p}^*-\mathbf{p}\right| + \left\|\mathbf{p}-\mathbf{p}^*\right\|^2\right],
\end{aligned}
$$

where $K_{37} = \max\{K_{35}, K_{36}\}$. The ℓ-1 norm is used here since the term $\nabla r^i(p^{i*}, p^{-i*})^T(\mathbf{p} - \mathbf{p}^*)$ can be a positive or negative scalar.

It can be derived that the revenue difference of firm i for any stage n under the synchronous price control is

$$\begin{aligned}
&\left[r^i(\mathbf{p}^*) - r^i(\hat{\mathbf{p}}_n)\right] + \left[r^i(\mathbf{p}^*) - r^i(\hat{p}_n^i + \delta_n, \hat{p}_n^{-i})\right] + \left[r^i(\mathbf{p}^*) - r^i(\hat{p}_n^i, \hat{p}_n^1 + \delta_n, \hat{p}_n^{-i,-i\neq 1})\right] \\
&\quad + \cdots + \left[r^i(\mathbf{p}^*) - r^i(\hat{p}_n^i, \hat{p}_n^N + \delta_n, \hat{p}_n^{-i,-i\neq N})\right] \\
&= \Big[r^i(\mathbf{p}^*) - r^i(\hat{\mathbf{p}}_n) + r^i(\mathbf{p}^*) - r^i(\hat{p}_n^i + \delta_n, \hat{p}_n^{-i}) + r^i(\mathbf{p}^*) - r^i(\hat{\mathbf{p}}_n) - \big(r^i(\mathbf{p}^*) - r^i(\hat{\mathbf{p}}_n)\big)\Big] \\
&\quad + \Big[r^i(\mathbf{p}^*) - r^i(\hat{p}_n^i, \hat{p}_n^1 + \delta_n, \hat{p}_n^{-i,-i\neq 1}) + r^i(\mathbf{p}^*) - r^i(\hat{\mathbf{p}}_n) - \big(r^i(\mathbf{p}^*) - r^i(\hat{\mathbf{p}}_n)\big)\Big] \\
&\quad + \cdots + \Big[r^i(\mathbf{p}^*) - r^i(\hat{p}_n^i, \hat{p}_n^N + \delta_n, \hat{p}_n^{-i,i\neq N}) + r^i(\mathbf{p}^*) - r^i(\hat{\mathbf{p}}) - \big(r^i(\mathbf{p}^*) - r^i(\hat{\mathbf{p}})\big)\Big] \\
&\quad + \nabla r^i(\mathbf{p}^*)^T(\hat{\mathbf{p}}_n - \mathbf{p}^*) + \frac{1}{2}(\hat{\mathbf{p}}_n - \mathbf{p}^*)^T \nabla^2 r^i(\mathbf{p}^*)(\hat{\mathbf{p}}_n - \mathbf{p}^*).
\end{aligned}$$

For all $t = t_n + iI_n + 1, \cdots, t_n + (i+1)I_n$, we have $\mathbf{p}_t = (\hat{p}_n^i + \delta_n, \hat{p}_n^{-i})$. Hence,

$$\begin{aligned}
&\Big[r^i(\mathbf{p}^*) - r^i(\hat{p}_n^i + \delta_n, \hat{p}_n^{-i})\Big] \\
&= \Big[r^i(\mathbf{p}^*) - r^i(\hat{p}_n^i + \delta_n, \hat{p}_n^{-i}) + r^i(\mathbf{p}^*) - r^i(\hat{\mathbf{p}}_n) - \big(r^i(\mathbf{p}^*) - r^i(\hat{\mathbf{p}}_n)\big)\Big] \\
&= \left[r^i(\mathbf{p}^*) - r^i(\hat{\mathbf{p}}_n)\right] + \Big[r^i(\mathbf{p}^*) - \Big(r^i(\mathbf{p}^*) + \nabla r^i(\mathbf{p}^*)^T\big((\hat{p}_n^i + \delta, \hat{p}_n^{-i}) - \mathbf{p}^*\big) \\
&\quad + \frac{1}{2}\big((\hat{p}_n^i + \delta, \hat{p}_n^{-i}) - \mathbf{p}^*\big)^T \nabla^2 r^i(\mathbf{p}^*)\big((\hat{p}_n^i + \delta, \hat{p}_n^{-i}) - \mathbf{p}^*\big)\Big) - \Big(r^i(\mathbf{p}^*) - (r^i(\mathbf{p}^*) \\
&\quad + \nabla r^i(\mathbf{p}^*)^T(\hat{\mathbf{p}}_n - \mathbf{p}^*) + \frac{1}{2}(\hat{\mathbf{p}}_n - \mathbf{p}^*)^T \nabla^2 r^i(\mathbf{p}^*)(\hat{\mathbf{p}}_n - \mathbf{p}^*)\Big)\Big] \\
&= [r^i(\mathbf{p}^*) - r^i(\mathbf{p})] + \Big[\nabla r^i(\mathbf{p}^*)^T\big(\hat{\mathbf{p}}_n - (\hat{p}_n^i + \delta, \hat{p}_n^{-i})\big) \\
&\quad + \frac{1}{2}\big(\hat{\mathbf{p}}_n - (\hat{p}_n^i + \delta, \hat{p}_n^{-i})\big)^T \nabla^2 r^i(\mathbf{p}^*)\big(\hat{\mathbf{p}}_n - (\hat{p}_n^i + \delta, \hat{p}_n^{-i})\big)\Big] \\
&= [r^i(\mathbf{p}^*) - r^i(\mathbf{p})] - \frac{1}{2}\big((\hat{p}_n^i + \delta, \hat{p}_n^{-i}) - \mathbf{p}^*\big)^T \nabla^2 r^i(\mathbf{p}^*)\big((\hat{p}_n^i + \delta, \hat{p}_n^{-i}) - \mathbf{p}^*\big) \\
&\quad + \frac{1}{2}(\hat{\mathbf{p}}_n - \mathbf{p}^*)^T \nabla^2 r^i(\mathbf{p}^*)(\hat{\mathbf{p}}_n - \mathbf{p}^*) \\
&\leq \left[r^i(\mathbf{p}^*) - r^i(\mathbf{p})\right] + K_{36}\delta^2.
\end{aligned}$$

Taking Taylors expansion for $r^i(\hat{p}_n^i + \delta_n, \hat{p}_n^{-i})$ and $r^i(\hat{\mathbf{p}}_n)$ around $\mathbf{p}^*$, we obtain the first inequality. Since $\nabla r^i(\mathbf{p}^*)^T\big(\hat{\mathbf{p}}_n - (\hat{p}_n^i + \delta, \hat{p}_n^{-i})\big) = 0$, the last equality is obtained. The terms $\frac{1}{2}\big((\hat{p}_n^i + \delta, \hat{p}_n^{-i}) - \mathbf{p}^*\big)^T \nabla^2 r^i(\mathbf{p}^*)\big((\hat{p}_n^i + \delta, \hat{p}_n^{-i}) - \mathbf{p}^*\big)$ and $\frac{1}{2}(\hat{\mathbf{p}}_n - \mathbf{p}^*)^T \nabla^2 r^i(\mathbf{p}^*)(\hat{\mathbf{p}}_n - \mathbf{p}^*)$ are bounded by $-K_{36}\big\|(\hat{p}_n^i + \delta, \hat{p}_n^{-i}) - \mathbf{p}^*\big\|^2$ and $-K_{36}\big\|\hat{\mathbf{p}}_n - \mathbf{p}^*\big\|^2$, respectively. Thus, we obtain the last inequality. We also have $\nabla r^i(\mathbf{p}^*)^T\big(\hat{\mathbf{p}}_n - (\hat{p}_n^j + \delta, \hat{p}_n^{-j})\big) < 0$. Similarly for all $t = t_n + jI_n + 1, \cdots, t_n + (j+1)I_n$, we have $\mathbf{p}_t =$

$(\hat{p}_n^j+\delta_n, \hat{p}_n^{-j})$ and an identical inequality.

$$\begin{aligned}&\Big[r^i(\mathbf{p}^*)-r^i(\hat{p}_n^i,\hat{p}_n^j+\delta,\hat{p}_n^{-i,-i\neq j})+r^i(\mathbf{p}^*)-r^i(\mathbf{p})-\big(r^i(\mathbf{p}^*)-r^i(\mathbf{p})\big)\Big]\\&\leq \big[r^i(\mathbf{p}^*)-r^i(\mathbf{p})\big]+K_{36}\delta^2.\end{aligned}$$

Combining the results above, the revenue difference of firm i becomes

$$\begin{aligned}D^i\leq\ &\mathbb{E}\Bigg[\sum_{n=1}^{k}\Bigg(\big[r^i(\mathbf{p}^*)-r^i(\mathbf{p})\big]+\big[r^i(\mathbf{p}^*)-r^i(p_t^i+\delta_n,p_t^{-i})\big]+\Big[r^i(\mathbf{p}^*)-r^i(p_t^i,p_t^1+\delta_n,p_t^{-i,-i\neq 1})\Big]\\&+\cdots+\Big[r^i(\mathbf{p}^*)-r^i(p_t^i,p_t^N+\delta_n,p_t^{-i,-i\neq N})\Big]\Bigg)I_n\Bigg]\\&\leq\mathbb{E}\left[\sum_{n=1}^{k}\big((N+1)(r^i(\mathbf{p}^*)-r^i(\mathbf{p}))+N\delta^2\big)I_n\right]\\&\leq K_{37}\sum_{n=1}^{k}\Big((N+1)\mathbb{E}\left\|\hat{\mathbf{p}}_n-\mathbf{p}^*\right\|^2+N\delta_n^2+(N+1)\mathbb{E}\left|\hat{\mathbf{p}}_n-\mathbf{p}^*\right|\Big)I_n,\end{aligned}\tag{38}$$

where $k=\inf\left\{h\geq 1:(N+1)\sum_{i=1}^{h}I_n\geq T\right\}$. The first inequality represents the revenue difference generated during the DDEP operations. The second inequality arises as the revenue difference in each time interval I_n has the same upper bound. The third inequality is derived using the Taylor approximation for $r^i(\mathbf{p})$ at $\mathbf{p}^*$. According to the results of Theorem 1, the upper bound of $\mathbb{E}\left[\|\hat{\mathbf{p}}_n-\mathbf{p}^*\|^2\right]$ is $N^2K_5I_n^{-\frac{1}{2}}$. Thus, we obtain the above result. Now, we turn to finding the upper bound of $\mathbb{E}\left[|\hat{\mathbf{p}}_n-\mathbf{p}^*|\right]$. Similar to the arguments in Proposition 3, we change the inequalities (27) into

$$\begin{aligned}&\mathbb{P}\Bigg\{\left|\hat{\alpha}_{n+1}^i-\breve{\alpha}_{n+1}^i(\hat{\mathbf{p}}_n)\right|+\sum_{j=1}^{N}\left|\hat{\beta}_{n+1}^{ij}-\breve{\beta}_{n+1}^{ii}(\hat{p}_n^i)\right|+\left|\hat{\alpha}_{n+1}^i-\breve{\alpha}_{n+1}^i(\hat{\mathbf{p}}_n+\Delta_n)\right|\\&\quad+\sum_{j=1}^{N}\left|\hat{\beta}_{n+1}^{ij}-\breve{\beta}_{n+1}^{ij}(\hat{p}_n^j+\delta_n)\right|\geq K_{38}N\left(\delta_n+\frac{x}{\delta_n}+x\right)\Bigg\}\\&\quad<24N^2e^{-2I_nx^2},\forall i.\end{aligned}\tag{39}$$

Note that $\left(\delta_n+\frac{x}{\delta_n}+x\right)\leq K_{39}x/\delta_n$ and that $\mathbb{E}[X]=\int_0^{+\infty}\mathbb{P}\{X\geq t\}dt$ for every non-negative random variable X. (The last equation is referred to as "tail-integral formula".) Integrating both sides leads to

$$\begin{aligned}&\mathbb{E}\Bigg[\left|\hat{\alpha}_{n+1}^i-\breve{\alpha}_{n+1}^i(\hat{\mathbf{p}}_n)\right|+\sum_{j=1}^{N}\left|\hat{\beta}_{n+1}^{ij}-\breve{\beta}_{n+1}^{ii}(\hat{p}_n^i)\right|+\left|\hat{\alpha}_{n+1}^i-\breve{\alpha}_{n+1}^i(\hat{\mathbf{p}}_n+\Delta_n)\right|\\&\quad+\sum_{j=1}^{N}\left|\hat{\beta}_{n+1}^{ij}-\breve{\beta}_{n+1}^{ij}(\hat{p}_n^j+\delta_n)\right|\Bigg]\\&\leq\delta_n^{-1}\int_0^{+\infty}Ne^{-2I_nx^2}dx\leq NK_{40}I_n^{-\frac{3}{4}},\forall i\end{aligned}\tag{40}$$

for some constant K_{40}. Similar to (37), we have

$$
\begin{aligned}
&\mathbb{E}\left[|\mathbf{z}-\hat{\mathbf{p}}_{n+1}|\right] \\
&\leq \mathbb{E}\Bigg[K_{28}\sum_{i=1}^{N}\left|\hat{\alpha}^i_{n+1}-\breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n)\right| + \sum_{j=1}^{N}\left|\hat{\beta}^{ij}_{n+1}-\breve{\beta}^{ii}_{n+1}(\hat{p}^i_n)\right| \\
&+\left|\hat{\alpha}^i_{n+1}-\breve{\alpha}^i_{n+1}(\hat{\mathbf{p}}_n+\Delta_n)\right| + \sum_{j=1}^{N}\left|\hat{\beta}^{ij}_{n+1}-\breve{\beta}^{ij}_{n+1}(\hat{p}^j_n+\delta_n)\right|\Bigg] + \frac{24NK_{32}}{I_n}\left\|\mathbf{p}^h-\mathbf{p}^l\right\| \\
&\leq NK_{41}I_n^{-\frac{3}{4}}
\end{aligned} \tag{41}
$$

for some constant K_{41}. Using (38), (40) and (41), we can derive as follows:

$$
\begin{aligned}
D^i &\leq K_{42}\sum_{n=1}^{k}N^2(N+1)\left[I_n^{-\frac{1}{2}}+\delta_n^2+I_n^{-\frac{3}{4}}\right]I_n \\
&\leq K_{43}\sum_{n=1}^{k}N^2((N+1)I_n)^{\frac{1}{2}} = N^2K_{43}\frac{((N+1)(I_0v))^{\frac{1}{2}}}{1-v^{\frac{1}{2}}}(1-v^{\frac{n-1}{2}}-1) \\
&\leq N^2K_{43}\frac{(N+1)^{\frac{1}{2}}(I_0v)^{\frac{1}{2}}}{1-v^{\frac{1}{2}}}\left(v^{\log_v(\frac{v-1}{(N+1)I_0v}T+1)+1-1}\right)^{\frac{1}{2}} \leq N^2K_6T^{\frac{1}{2}},
\end{aligned}
$$

where $K_{42}=\max\{K_{37},K_{41}\}$. The first inequality is obtained by employing the results of (41) in (38). Selecting some suitable constant K_{43}, we can make the order of I_n consistent. The equality comes from the summation of a geometric sequence according to the constant k defined previously. This theorem implies that using this pricing policy, even if the underlying demand function is unknown, the difference due to the gap between the equilibrium point obtained at each stage $\hat{\mathbf{p}}_{n+1}$ and the clairvoyant GNE $\mathbf{p}^*$ is at most $N^2K_6T^{\frac{1}{2}}$.

D Proof of Theorem 4

Proof. This result is derived from the proofs of Theorem 2, Lemma 2 and Theorem 1 shown previously. Having established the price convergence when all firms are unaware of the demand functions, we now verify whether there are similar results when some firms know the demand functions.

Suppose $N-N'$ firms already know the demand function, then only N' firms need to estimate their demand function. According to the proof of Lemma 2, it follows that

$$
\begin{aligned}
\hat{\beta}^{ij}_{n+1} &= \frac{\lambda^i(\mathbf{p}_n;\hat{p}^j_n+\delta_n)-\lambda^i(\mathbf{p}_n)}{\delta_n} - \frac{1}{\delta_n}\left(W^{i,j}_n-W^{i,1}_n\right) \\
&= \left[\nabla_{p^j}\lambda^i(\mathbf{p}_t)+\frac{1}{2}\nabla^2_{p^j}\lambda^i(\mathbf{p}')\delta_n\right] - \frac{1}{\delta_n}\left(W^{i,j}_n-W^{i,1}_n\right) \\
&= \breve{\beta}^{ij}_{n+1}+C^{ij}_{n,\beta^{ij}},
\end{aligned}
$$

Hence, the inequality in Proposition 2 changes to

$$
\|\mathbf{p}^*-\mathbf{z}(\hat{\mathbf{p}}_n)\| \leq \|C_n\|,
$$

where $C_n \equiv \left[C^1_n,\cdots,C^{N'}_n\right]$. Recall that N' represents the number of firms who do not know the demand function. Then, the probability in Proposition 3 (36) also changes to

$$
\mathbb{P}\left\{\mathscr{B}_n\right\} \geq 1-\frac{24N'^2K_{32}}{I_n}.
$$

Combine these two inequalities above, the convergence results of equilibrium can be obtained; that is,

$$\mathbb{E}\left[\|\hat{\mathbf{p}}_{n+1}-\mathbf{p}^*\|^2\right] \leq N'^2 K_{47} I_n^{-\frac{1}{2}}. \tag{42}$$

Based on the proof of Theorem 2, similar results can be obtained:

$$R^i \leq N'^2 K_8 T^{\frac{1}{2}}.$$

E Proof of Theorem 5

Proof. We first analyze the regret among firms without knowledge of the demand curve. This proof is similar to that of Theorem 3. The major difference relates to the best response. The best response operator in the partially clairvoyant model is $z^i\big(\hat{\alpha}(\hat{\mathbf{p}}_t), \hat{\beta}(\hat{\mathbf{p}}_t)\big)$ where the inputs $\hat{\alpha}$ and $\hat{\beta}$ are different from the inputs of the best response operator in the full model. Fixing firm i, the parameter $\hat{\alpha}^i_{n+1}$ in the full model is

$$\hat{\alpha}^i_{n+1} = \lambda^i(\hat{\mathbf{p}}_n) + \hat{\beta}^{ii}_{n+1}\hat{p}^i_n - \sum_{j,j\neq i}^{N} \hat{\beta}^{ij}_{n+1}\hat{p}^j_n,$$

while $\hat{\alpha}^i_{n+1}$ in the partially-clairvoyant model is

$$\hat{\alpha}^i_{n+1} = \lambda^i(\hat{\mathbf{p}}_n) + \hat{\beta}^{ii}_{n+1}\hat{p}^i_n - \sum_{j,j\neq i}^{N'} \hat{\beta}^{ij}_{n+1}\hat{p}^j_n.$$

The difference of α between the partially-clairvoyant model and the full model is $\sum_{k=1}^{N-N\prime} \beta^{ik}_{n+1}p^k_n$. For all $t = t_n+1, \cdots, t_n+I_n$, we have that $\mathbf{p}_t = (\hat{p}^i_n, \hat{p}^{-i}_n)$. Thus,

$$\begin{aligned}
&\left|p^{i*}_t - p^i_t\right| \\
&= \left|z^i\big(\breve{\alpha}(p^{i*}_t, \hat{p}^{-i}_n), \breve{\beta}(p^{i*}_t, \hat{p}^{-i}_n)\big) - z^i\big(\hat{\alpha}(\hat{\mathbf{p}}_n), \hat{\beta}(\hat{\mathbf{p}}_n)\big)\right| \\
&= \left|z^i\big(\breve{\alpha}(p^{i*}_t, \hat{p}^{-i}_n), \breve{\beta}(p^{i*}_t, \hat{p}^{-i}_n)\big) - z^i\big(\breve{\alpha}(\hat{\mathbf{p}}_n), \breve{\beta}(\hat{\mathbf{p}}_n)\big) + \frac{\sum_{k=1}^{N-N\prime} \breve{\beta}^{ik}_{n+1}(p^k_n - p^k_{n+1})}{\breve{\beta}^{ii}_{n+1}}\right| + \|C_n\| \\
&= \left|\frac{\sum_{k=1}^{N-N\prime} \breve{\beta}^{ik}_{n+1}(p^k_n - p^{k*} - p^k_{n+1} + p^{k*})}{\breve{\beta}^{ii}_{n+1}}\right| + \|C_n\| \\
&\leq \left|\frac{\sum_{j=1}^{N} \breve{\beta}^{ij}_{n+1}(p^j_n - p^{j*} - p^j_{n+1} + p^{j*})}{\breve{\beta}^{ii}_{n+1}}\right| + \|C_n\| \\
&\leq K_{48}\left|\hat{\mathbf{p}}_{n+1} - \mathbf{p}^*\right| + \|C^i_n\|.
\end{aligned} \tag{43}$$

As p^{i*}_t and p^i_t are the fixed points of $z^i\big(\breve{\alpha}(p^{i*}_t, \hat{p}^{-i}_n), \breve{\beta}(p^{i*}_t, \hat{p}^{-i}_n)\big)$ and $z^i\big(\hat{\alpha}(\hat{\mathbf{p}}_n), \hat{\beta}(\hat{\mathbf{p}}_n)\big)$, respectively, we can obtain the first equality.. The second equality transforms the operator $z^i\big(\hat{\alpha}(\hat{\mathbf{p}}_n), \hat{\beta}(\hat{\mathbf{p}}_n)\big)$ in the flawed demand model into the operator in the full model. The last inequality uses the results of price convergence in the ℓ-1 norm and selects some appropriate constant K_{48} to cover the term $(p_n^j - p^* - p_{n+1}^j + p^{j*})$. Using a result implied in the proof of Theorem 2, we obtain that $\mathbb{E}\left[|\hat{\mathbf{p}}_{n+1} - \mathbf{p}^*|\right] \leq N' K_{49} I_n^{-\frac{3}{4}}$. Combining the inequality above and (43), one obtains

$$\mathbb{E}\left[\left\|p_t^{i*} - p_t^i\right\|^2\right] \leq N' K_{50} I_n^{-\frac{1}{2}}.$$

Similar to the proof of Theorem 3, the regret bound can be obtained as follows:

$$R^i \leq N' K_9 T^{\frac{1}{2}}.$$

We then analyze the regret of the firms with knowledge of the demand curve. The firms with the knowledge of the demand curves do not need to adjust prices. We divide the price experimentation in Step 1 of the modified DDEP into two parts and analyze the regret separately. First, in the first I_n periods, there is no firm adjusting price. According to the proof of Theorem 3, for all $t = t_n + 1, \cdots, t_n + I_n$, $\mathbf{p}_t = (\hat{p}_n^i, \hat{p}_n^{-i})$. Thus,

$$\left|p_t^{i*} - p_t^i\right| = \left|z^i(\breve{\alpha}(p_t^{i*}, \hat{p}_n^{-i}), \breve{\beta}(p_t^{i*}, \hat{p}_n^{-i})) - z^i(\breve{\alpha}(\hat{\mathbf{p}}_n), \breve{\beta}(\hat{\mathbf{p}}_n))\right| = \left|p_t^{i*} - z^i(\breve{\alpha}(\hat{\mathbf{p}}_n), \breve{\beta}(\hat{\mathbf{p}}_n))\right| = 0. \quad (44)$$

Second, in periods in which firms without knowledge of the demand curves have changed their price, suppose that firm j has adjusted its price, for all $t = t_n + jI_n + 1, \cdots, t_n + (j+1)I_n$, we arrive at $\mathbf{p}_t = (\hat{p}_n^j + \delta_n, \hat{p}_n^{-j})$

$$\begin{aligned}
&\left|p_t^{i*} - p_t^i\right| \\
&= \left|z^i\big(\breve{\alpha}(p_t^{i*}, \hat{p}_n^{-i}), \breve{\beta}(p_t^{i*}, \hat{p}_n^{-i})\big) - z^i\big(\hat{\alpha}(\hat{\mathbf{p}}_n), \hat{\beta}(\hat{\mathbf{p}}_n)\big)\right| \\
&= \left|z^i\big(\breve{\alpha}(p_t^{i*}, \hat{p}_n^{-i\backslash j}, \hat{p}_n^j + \delta_n), \breve{\beta}(p_t^{i*}, \hat{p}_n^{-i\backslash j}, \hat{p}_n^j + \delta_n)\big) - z^i\big(\breve{\alpha}(\hat{\mathbf{p}}_n), \breve{\beta}(\hat{\mathbf{p}}_n)\big)\right| \\
&= \left|z^i_{\hat{p}_n^j}\delta_n + z^i\big(\breve{\alpha}(p_t^{i*}, \hat{p}_n^{-i}), \breve{\beta}(p_t^{i*}, \hat{p}_n^{-i})\big) - z^i\big(\breve{\alpha}(\hat{\mathbf{p}}_n), \breve{\beta}(\hat{\mathbf{p}}_n)\big)\right| \\
&= \left|z^i_{\hat{p}_n^j}\delta_n + p_t^{i*} - z^i\big(\breve{\alpha}(\hat{\mathbf{p}}_n), \breve{\beta}(\hat{\mathbf{p}}_n)\big)\right| = K_{51}\delta_n, \forall j. \quad (45)
\end{aligned}$$

In the same way,

$$\mathbb{E}\left[\left\|p_t^{i*} - p_t^i\right\|^2\right] \leq K_{52} I_n^{-\frac{1}{2}}.$$

Hence, the following regret bound is obtained.

$$R^i \leq K_{10} T^{\frac{1}{2}}.$$