LEARNING UNDERSPECIFIED MODELS

IN-KOO CHO AND JONATHAN LIBGOBER

Abstract. This paper considers optimal pricing with a seller who is endowed with a *underspecified* model, in that he does not possess a complete description of how actions translate into payoffs. To save computational cost, a monopolist designs an algorithm delegating the decision to determine a product's price in each period. Not knowing the true demand curve, the algorithm is tasked with ensuring that the optimal price emerges in the long run with sufficiently high probability, uniformly over the set of possible demand curves. The monopolist has a lexicographic preference over the payoff and the complexity cost of the algorithm, seeking an alogrithm with a minimum number of parameters subject to achiving the same level of long run average payoff. We show that for a large class of possible demand curves with strictly decreasing continuous marginal revenue curve, the monopolist selects an algorithms which assume demand is linear even if it is not. The monopolist chooses a misspecified model to save computational cost, while learning the true optimal decision uniformly.

Keywords. Optimal Sales Mechanism, Underspecified model, PAC learnability, Complexity Cost

1. Introduction

Consider a startup company with a new disruptive technology attempting to determine the best price for its product. Although the company may know that the marginal revenue curve is continuous and downward sloping, more details than this piece of information might not be known since the product is new. Given the lack of information, the firm might consider a particular specification of possible mappings when choosing its pricing strategy. But this approach involves some risk that the true demand curve will not be included within the specification of those considered feasible. On the other hand, if the firm can observe how the market reacts to various choices, then it may wish to use a specification which will at least be able to ensure an optimal price emerges eventually. This paper explains how to accomplish this goal, thus suggesting the firm could eschew additional costs which might be involved were it to accomodate additional complexity.

The anecdote features a firm that does not know some critical variable that determines their payoffs but has an opportunity to observe data to learn about demand. Departing from the Bayesian model, we assume that the monopolist lacks any kind of information to form a prior over the set of possible demand curve. We will say an environment is *underspecified* if an agent is *not* (initially) endowed with a complete description of how actions translate into payoffs. In the above story, the firm's decision problem is

Date: April 16, 2022.
Financial support from the National Science Foundation and the National Research Foundation of Korea is gratefully acknowledged.

underspecified because it does not initially fully possess a description of the mapping from prices into levels of profit.

We search for an algorithm that the monopolist can "learn" the optimal strategy for an underspecified demand curve uniformly with a reasonable amount of data, while minimizing the computational cost. Among the algorithms that can learn the optimal price, we look for the one that calculates the minimal number of parameters to derive the optimal price, using readily available aggregate data such as price and quantity.

We show that the simplest algorithm must have at least two parameters to compute. We construct an algorithm with two parameters, that estimates a linear demand curve (which can be paramterized by the slope and the intercept). The algorithm allows the monopolist to learn the optimal price even if the decision problem is underspecified. We also demonstrate that the algorithm is efficient, in that the probability that the forecasting error of the optimal price is larger than a small tolerance bound vanishes at an exponential rate, uniformly over the set of feasible demand curves.

Through the linear demand curves, the seller *can* still learn how to choose the optimal price at an exponential rate, *even though* they may be grossly misspecified because the actual demand curve can be highly non-linear. Yet, within a polynomial time, the monopolistic seller behaves as if he knows the actual demand curve. While a monopolist could possibly use an algorithm calculating more parameters such as the non-parametric estimation (Cole and Roughgarden (2014)), the monopolist chooses a simpler, even if objectively incorrect, model of demand to save computational costs. In this sense, a misspecified model would be a representation of the procedural rationality of the seller (Osborne and Rubinstein (1998)).

Our paper differs from three major approaches to investigate the decision problem with underspecified models. The most prominent approach is that a decision-maker chooses his actions assuming uncertainty about the environment resolves in favor of the worst-case (e.g., Hansen and Sargent (2007), Carroll (2015), Carroll (2017), Du (2018) and Libgober and Mu (2021)). The choice that maximizes the objective under the worst conjecture will typically differ from the optimal solution against the truth; sometimes, significantly so. While data could, in principle, be used to bring the decision-maker closer to optimality, typically, these models are silent on how the decision-maker might do this.[1] In contrast, the monopolist in our model seeks to find a *true* optimal price for an unknown demand curve,using data.

In the second approach, the decision-maker completes the specification by specifying (parametric) model, and estimates the parameters using data. While this approach does allow the decision-maker to learn from data, the specification is fixed exogenously, and might exclude the true mapping from actions into outcomes (e.g., Cho and Kasa (2015), Cho and Kasa (2017), Heidhues, Köszegi, and Strack (2018), Esponda, Pouzo, and Yamamoto (2021), Frick, Iijima, and Ishii (2021) and Fudenberg, Lanzani, and Strack (2021)). Misspecification is generally considered a consequence of behavioral restriction or inability to complete the specification, indicating that if the decision-maker is rational, his model should be correctly specified. The monopolist in our model, by contrast, is not committed to a particular specification. In principle, sets of demand curves richer than linear ones

[1]A remarkable exception is a series of papers by Hansen and Sargent (e.g., Hansen and Sargent (2020)).

could be considered. The monopolist specifies the linear model with two parameters, to learn the optimal decision, while minimzing the computational cost. The linear model is not imposed, but derived as a consequence of optimization by the monpolist.

This paper follows a third approach, inspired by machine learning models (e.g., Huang, Mansour, and Roughgarden (2018) and Cole and Roughgarden (2014)). A typical approach in this literature (Cole and Roughgarden (2014)) assumes that the seller non-parametrically estimates the demand curve. Yhis approach avoids the underspecification issue, and the seller can achieve the true optimal curve as the estimator converges to an actual demand curve. The investigation then focuses on the consistency of the estimator and data complexity.

Following Cole and Roughgarden (2014), we freely borrow the critical concepts developed in the machine learning literature (e.g., Shalev-Shwartz and Ben-David (2014)). We depart from Cole and Roughgarden (2014), however, by considering the complexity of the algorithm and the cost of obtaining data. An essential advantage of the non-parametric estimation technique is to avoid the misspecification of the demand curve. The downside is that the estimator is complex because the estimator requires a possibly unbounded number of parameters to represent a highly non-linear demand curve. If the monopolistic seller incurs the cost of storing the estimator in the memory, he will search for an algorithm that uses a simpler specification of the demand curve.

Cole and Roughgarden (2014) and most papers in computer science assume that the monopolist can extract the reservation value of the buyers for free. Since truthful revelation is a dominant strategy in the revelation game (Myerson (1981)), the monopolistic seller can sample the true valuations of the consumers. The revelation game is a mathematical tool to investigate a Bayesian Nash equilibrium of a trading protocol under asymmetric information. When purchasing a product from a small company's website, it would be rare for a buyer to be asked to report his true valuation. Instead, a buyer pays according to the posted price. The extraction of private information often requires considerable resources to infer private information from the observed action of the informed party. In contrast, the outcome of the trading, such as price and quantity, is often more readily available. Aggregate data such as the quantity of trade is usually cheaper to obtain than microdata, such as the individual valuations of buyers. It would be much more economical to construct an algorithm that uses readily available aggregate data.

The rest of the paper is organized as follows. Section 2 reviews the literature.

In Section 3, we formally describe the problem and define the basic concepts, with a summary of the main result, where we calculate the lower bound of the complexity of algorithms to calculate the optimal price. In Section 4, we construct the algorithm, which is the simplest algorithm among algorithms that learn the optimal pricing rule within a polynomial-time accurately. Instead of a formal analysis, Section 5 informally illustrates how the monopolist can efficiently and uniformly learn the optimal price of a non-linear demand curve through a linear demand curve. The proof is in the appendix. Section 6 reports a numerical exercises to show that the performance of our algorithm is comparable to a more elaborate algorithm of Cole and Roughgarden (2014). Section 7 concludes the paper.

2. Literature Review

The monopoly market provides an ideal framework to investigate the learnability of the optimal price, as Myerson (1981) lays out an elegant benchmark. Cole and Roughgarden (2014), who showed that, by estimating the value distribution and using the optimal reserve on the estimated value distribution, a seller can achieve up to $1-\varepsilon$ times the optimal auction on the known value distribution, with high probability. A sizable literature in computer science following Cole and Roughgarden (2014) relates to how a designer may be able to use draws from a value distribution in order to design an auction that performs well, without assuming knowledge of the true value distribution. We mention only a few papers that appear to be closely related to our paper.[2] Huang, Mansour, and Roughgarden (2018) showed that one can improve their convergence rates by only seeking to find the optimal *auction* without necessarily seeking to find the *revenue* obtained by the auction, and other work has considered extensions to richer settings (e.g., multidimensional ones as in Gonczarowski and Weinberg (2021), richer screening environments as in Goncalves and Furtado (2020) or weaker assumptions on the value distribution as in Roughgarden and Schrijvers (2016)).

After completing the analysis, we learned that the same class of problem was investigated in the operation research literature. Most notably, Besbes and Zeevi (2015) examined the monopoly profit maximization problem, where the monopolist is endowed with a linear demand curve, who estimates the demand curve to choose the price. Besbes and Zeevi (2015) demonstrated that one can construct an algorithm to let the monopolist learn the true optimal price for a class of demand curves satsifying a set of regularity conditions. Our paper differs from Besbes and Zeevi (2015) in two important ways. First, Besbes and Zeevi (2015) shows the "pointwise" convergence to the optimal price, where the number of data needed to approximate the actual optimal price can depend on the demand curve. This paper shows a stronger result. The convergence of the forecast price is uniform over the class of demand functions with strictly decreasing marginal revenue curves. The uniform convergence is essential to PAC guarantee the optimal price, which is the central concept of learnability in the computer science. Second, Besbes and Zeevi (2015) did not consider the complexity of the algorithm, while leaving the question of whether we can achieve the convergence to the true optimal price with a simpler algorithm. In our paper, we measure the complexity roughly in terms of the parameters the algorithm has to remember in each round. Tthe monopolist chooses the simplest algorithm among those that can PAC guarantee the optimal price. We show that our algorithm is the simplest algorithm, which turns out to be estimating a linear demand.[3]

Esponda and Pouzo (2014) provides a useful solution concept, followed by a sizable number of works to investigate the learning with misspecified models (Heidhues, Köszegi, and Strack (2018), Fudenberg, Lanzani, and Strack (2021), Frick, Iijima, and Ishii (2021)

[2]Nisan, Tardos, and Vazirani (2007) reports the early applications of the algorithms to game theretic models including the monopoly problem.

[3]The algorithm of Besbes and Zeevi (2015) has to remember at least 3 parameters: 2 parameters of the demand curve (intercept and slope) and another parameter(s) to indicate the position in one of the two phases of the algorithm. Our algorithm remembers only 2 parameters: intercept and slope of the demand curve.

and Esponda, Pouzo, and Yamamoto (2021)). In the existing studies and Besbes and Zeevi (2015), a misspecified model is imposed upon the decision maker. In our paper, the monopolist can use a non-parametric estimation of the demand curve (thus, can use a correctly specified model), but chooses the linear model of the demand curve, to save the computational cost. The misspecification in our paper is derived from the optimization problem of the monopolist, not assumed by the modeler.

3. Description

3.1. **Demand.** There are N buyers, each of whom is indexed by $i \in \{1, \dots, N\}$ and is endowed with reservation value $v_i \in [\underline{v}, \overline{v}]$. Let $F_i(v_i)$ be the distribution of valuation of buyer i. We assume that v_i and v_j are independent $\forall i \neq j$. Given p, buyer i purchases one unit of the good if $p \leq v_i$. If the seller charges p, the (normalized) aggregate demand is

$$q = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}(v_i \geq p)$$

and

$$\mathbb{E}q = 1 - \frac{1}{N} \sum_{i=1}^{N} F_i(p).$$

Define

$$F(p) = \frac{1}{N} \sum_{i=1}^{N} F_i(p)$$

and

$$\epsilon_2 = q - (1 - F(p)),$$

where $\mathbb{E}\epsilon_2 = 0$.

We interpret $1 - F(p)$ as the expected quantity of sales. The actual amount q of sales as a random variable, whose expected value is $1 - F(p)$ if the price is p. If $p = 0$, $q = 1$ with probability 1, and $F(p) \to 1$ as $p \to \overline{v}$. We can treat F as a distribution function. Let f be the density function of F.

3.2. **What the Monopolist Knows.** At the beginning of the game, the complete specification F is not available to the monopolist. The actual formula of F is not observed by the monopolist. The monopolist does not have a prior over the set of feasible demand curves. Instead, the monopolist knows a certain set of actual properties of F, which is a common knowledge among all players. We write the set of distributions that satisfy the said properties as $\mathcal{F}$.

Let us list the properties we are interested in, and define the set of distributions accordingly.

IH The support of F is $[\underline{p}, \overline{p}]$ and $f(p) > 0$ $\forall p \in [\underline{p}, \overline{p}]$. $\forall F \in \mathcal{F}$, its density function f is continuous over $[\underline{p}, \overline{p}]$ and its hazard rate

$$\frac{f(p)}{1 - F(p)}$$

is increasing.

Let $\mathcal{F}^0$ be the set of all demand curves satisfying the increasing hazard rate property. In order to avoid the technical problems, we assume that the demand curve is strictly downward sloping without any jump. A convenient way to ensure the property would be to assume the Lipschitz continuity of the density function.

LC $\forall f$ $\exists \eta > 0$ such that $\forall p, p' \in [\underline{p}, \overline{p}]$,

$$\left|f(p) - f(p')\right| < \eta |p - p'|.$$

We consider a compact convex subset of $\mathcal{F}^0$ by imposing the Lipschitz continuity uniformly.

UL Define

$$\mathcal{F}^\eta = \left\{ F \in \mathcal{F}^0 \mid \exists \eta > 0,\ \forall p, p' \in [\underline{p}, \overline{p}],\ \left|f(p) - f(p')\right| < \eta |p - p| \right\}$$

Lemma 3.1. (1) *$\forall \eta > 0$, $\mathcal{F}^\eta$ is (sequentially) compact.*

(2) *$\forall \eta > 0$, there exists a compact set K in the interior of $\mathbb{R}^2_+$ so that $\forall F \in \mathcal{F}^\eta$, $(b^*(F), 1 - F(b^*(F))) \in K$.*

(3) *$\mathcal{F}^p = \cup_{\eta > 0} \mathcal{F}^\eta$*

$\mathcal{F}^\eta$ with a large $\eta > 0$ can be considered as a large subset of $\mathcal{F}^0$. The main result of this paper is based on the assumption that the seller knows $\mathcal{F}^\eta$ and all of parameters that can be deduced from $\mathcal{F}^\eta$.

3.3. **The Seller's Objective.** In the equilibrium analysis, the monopolist knows the distribution F of the valuations of the buyer, which is a common knowledge among players. Instead, let us assume that the monopolist only knows that the true distribution is an element of $\mathcal{F}$

$$F \in \mathcal{F}^\eta$$

and the monopolist knows $\mathcal{F}^\eta$, which is a common knowledge among all players. All buyers are fully rational. The seller has no other information such as the prior probability over $\mathcal{F}$.

To incorporate the learning process by the monopolist, we consider a dynamic version of the static model, where the long run monopolist seller is facing a sequence of short run consumers with IID draws of their valuations.

Time is discrete: $t = 1, 2, 3, \ldots$. In each period, the monopolist post price p_t and N consumers enters the market, each of whom is endowed with reservation value $v_{i,t}$ drawn according to distribution function $F_i(\cdot)$. Upon entering the market in period t, buyer i purchases a single unit of the good paying p_t if $v_{i,t} > p_t$ and leaves the market forever. Since a buyer in period t plays only once, his dominant strategy is to buy if $v_{i,t} > p_t$. Define q_t as the number of buyers whose valuation in period t is more than p_t, which is the aggregate demand in period t.

In each period, the seller observes data D_t which may include (normalized) quantity q_t and price p_t charged in period t. We differentiate the outcome from the data. The outcome specifies what a player can observe. By the data, we mean a subset of the outcome which the decision maker uses as an input for the decision making process (or algorithm). If processing information is costly, the decision maker can choose to ignore some outcomes. By the definition, data is a subset of outcome. Let $\dim D_t$ is the number of components

in D_t, which cannot be larger than the number of components of observed outcomes in period t. The configuration of data is a part of the strategic choice of the monopolist in designing an algorithm.

3.4. **Algorithms.** If $(F_1, \ldots, F_N)$ or the aggregate distribution F is common knowledge and the valuations of consumers are independent, the decision problem is little more than the problem of finding an optimal sales mechanism in a Bayesian game of Myerson (1981). Departing from the Bayesian games, the monopolistic seller only knows $F \in \mathcal{F}^\eta$, and that the monopolist does not have a prior probability distribution over $\mathcal{F}^\eta \subset \mathcal{F}^0$. The smaller$\mathcal{F}^\eta$ becomes, the more information the seller knows about the actual distribution F. One extreme is to assume that the seller knows the actual distribution F, as in the equilibrium analysis. Another extreme is that the seller has no information other than $\mathcal{F}^0$ in which the maxmin type analysis is sensible.

While the monopolist is lacking a precise information about the actual distribution F, the monopolist can generate data D_t in period t to learn $F \in \mathcal{F}^\eta$. Let $\mathcal{D}_t = (D_1, \ldots, D_{t-1})$ be a history at the beginning of period t, and $\mathcal{D}$ be the set of all histories.

Let Θ be the set of parameters that "models" $F \in \mathcal{F}$. $\theta \in \Theta$ is a forecast (or a model) of the underlying aggregate demand curve. Let K be the number of parameters to model $F \in \mathcal{F}$.

Assumption 3.2. *Θ is a compact subset of $\cup_{K \geq 1} \mathbb{R}^K$.*

Θ is closed and bounded. By assuming that Θ is bounded, we assume that the decision maker has some knowledge about the underlying demand curve. That is, if some component of θ is extremely large, then he knows the forecast makes no sense and discards the forecast. On the other hand, it is a technical assumption that Θ is closed. K can be interpreted as the number of parameters for a model. We impose no upper bound for the number of paramters, admitting the non-parametric estimator as a feasible model.

We admit the model of the demand can be misspecified. If $\mathcal{F}$ is a collection of non-linear functions, θ may have infinitely many components. If $\Theta \subset \mathbb{R}^K$ for finite K, then the model could be misspecified. As the number of the parameters increases, the model can better approximate the true demand function. We call $\mathcal{A}(\mathcal{D}_t)$ the forecast conditioned on $\mathcal{D}_t$, or simply, the forecast.

To save the cost of decision making, the seller delegates his decision to an algorithm

$$\mathcal{A} : \mathcal{D} \to \Theta.$$

The number of components in $\mathcal{A}(\mathcal{D}_t)$ or D_t can change over time, if the monopolist's forecast about F evolves from a simple linear function to a non-linear function. In this paper, we only consider the algorithm in which the number of components in $\mathcal{A}(\mathcal{D}_t)$ and D_t do not change over time.

The statistical properties of $\mathcal{A}(\mathcal{D}_t)$ can change over time as well. If the monopolist estimates the parameters of F, the second moment of the estimates may decrease as more data become available. Our formulation of algorithm $\mathcal{A}$ is sufficiently general to cover the construction of new specifications over time as well as the estimation of the parameters. We impose several conditions on $\mathcal{A}$. We sometimes write $\mathcal{A}(\mathcal{D}_t : F)$ for $F \in \mathcal{F}$ to emphasize that the data is generated by the true F.

We impose a series of conditions for a feasible algorithm.

A1. Translating function.
Since $\theta \in \Theta$ can be arbitrary, $\mathcal{A}$ must accompany a "translating function"

$$\varphi(\mathcal{A}(\mathcal{D}_t)) = (\varphi_p(\mathcal{A}(\mathcal{D}_t)), \varphi_q(\mathcal{A}(\mathcal{D}_t))) = (p_t, q_t)$$

which translates θ into the (human readable) forecast about the optimal price p_t and the expected demand q_t at the optimal price (thus, the maximized expected profit). For example, if $\mathcal{A}(\mathcal{D}_t)$ is the empirical distribution of F, then the translation function calculates the estimated optimal price and the expected quantity (thus, estimated profit).

Since the goal of the monopolist is to find the optimal price, it is natural that $\mathcal{A}(\mathcal{D}_t)$ includes information about the forecast about the optimal price conditioned on $\mathcal{D}_t$. On the other hand, it is a substantive restriction to require the algorithm should be able to provide information to forecast the expected quantity (and therefore, the expected profit). To substantiate the forecast about the optimal price, we require that the algorithm should be able to explain how to calculate the optimal price and what would be the consequence of following the forecast from the algorithm.

A2. Lipschitz.
We assume that the translating function φ is Lipschitz continuous. A small change of the forecast should not drastically change the forecast of the optimal price and the expected quantity.

A3. Recursive algorithm.
$\mathcal{A}$ is a recursive algorithm, if $\exists \Psi$ such that

$$\mathcal{A}(\mathcal{D}_t) = \Psi(\mathcal{A}(\mathcal{D}_{t-1}), D_t).$$

The algorithm needs to remember only the most recent forecast from the algorithm and the market data, instead of the entire history. The cost of storing past information is minimized.

A4. PAC guarantee.
For $F \in \mathcal{F}$, define $(b^*(F), q^*(F))$ as

$$b^*(F) = \arg\max_p p(1 - F(p))$$

and $q^*(F) = 1 - F(b^*(F))$. Thanks to the increasing hazard rate property, the optimization problem admits a unique solution $b^*(F)$.

Definition 3.3. *$\mathcal{A}$ PAC guarantees $\mathcal{F}$ if $\forall \mu > 0$, $\forall \lambda \in (0,1)$, $\exists T(\mu, \lambda)$ such that*

$$\mathbb{P}\left(\left|\varphi(\mathcal{A}(D_{T(\mu,\lambda)})) - (b^*(F), q^*(F))\right| \geq \mu\right) \leq \lambda \tag{3.1}$$

where $T(\mu, \lambda) \sim \mathsf{O}\left(-\frac{\log \lambda}{\mu^p}\right)$ for some $p > 0$.

Let $\mathcal{D}_t = (D_1, \ldots, D_{t-1})$ be the history at the beginning of period t. Let $\mathcal{D}$ be the set of all histories. Algorithm $\mathcal{A}$ dictates the monopolist to charge p_t conditioned on $\mathcal{D}_t$. Let (p_t, q_t) be the pair of realized delivery price and realized quantity in period t. The long run average payoff from algorithm $\mathcal{A}$ is

$$\mathcal{U}(\mathcal{A}) = \lim_{T\to\infty} \frac{1}{T} \mathbb{E} \sum_{t=1}^{T} p_t q_t$$

if the limit exists, and 0, otherwise.

3.5. **Existence.** Cole and Roughgarden (2014) proved the existence of an algorithm that PAC guarantees $\mathcal{F}^\eta$. While the original algorithm of Cole and Roughgarden (2014) is defined in non-recursive form, the algorithm can easily adapted to a recursive form, which we present as the benchmark for our analysis.[4]

In each period, the monopolistic seller asks buyer $i \in \{1, \ldots, N\}$ reports his valuation. Since the truthful revelation is a dominant strategy of a buyer in period t, we assume that each buyer in period t reports his valuation $v_{i,t}$ truthfully, because the truthful reporting is a dominant strategy. Let $D_t = (v_{1,t}, \ldots, v_{N,t})$ be the data collected by the monopolist in period t and

$$\hat{F}_{i,t} = \frac{1}{t}\#\{t' \leq t \mid v_{i,t'} \leq v\}$$

be the empirical distribution of buyer i based on t observations, where $\#$ is the number of elements in the set. Given $\mathcal{D}_t = (D_1, \ldots, D_{t-1})$ and D_t, the monopolistic seller updates $\hat{F}_{i,t}$ recursively according to

$$\hat{F}_{i,t}(v) = \hat{F}_{i,t-1}(v) + \frac{1}{t}\left[\mathbb{I}(v_{i,t} \leq v) - \hat{F}_{i,t-1}(v)\right] \qquad \forall v.$$

The monopolist estimates the (normalized) aggregate demand

$$\hat{F}_t(v) = \frac{1}{N}\sum_{i=1}^{N} \hat{F}_{i,t}(v) \qquad \forall v$$

Based on $\hat{F}_t$, the monopolists offers p_t that solves

$$p_t = \arg\max_{p \in [\underline{p}, \bar{p}]} p(1 - \hat{F}_t(p))$$

and forecasts the expected amount of sales as

$$1 - \hat{F}_t(p_t).$$

Let $\mathcal{A}_{CR}$ be the algorithm we just described. Let us state the main result of Cole and Roughgarden (2014) adapted to our framework.

Theorem 3.4. *(Cole and Roughgarden (2014)) $\mathcal{A}_{CT}$ PAC guarantees $\mathcal{F}^\eta$.*

Since v_t is i.i.d., the empirical distribution $\hat{F}_{i,t}(v)$ converges to the true distribution $F_i(v)$ pointwise almost everywhere, and satisfies the large deviation property so that the estimation error vanishes at an exponential rate (Dembo and Zeitouni (1998)). The estimation method is recursive and efficient. By the nature of non-parametric estimation, the monopolist's model of the demand is correctly specified.

[4]The result of Cole and Roughgarden (2014) is stronger than what we state here, because Cole and Roughgarden (2014) constructed an algorithm that can PAC guaranteee $\mathcal{F}^0$ which includes $\mathcal{F}^\eta$. Since $\cup_{\eta>0}\mathcal{F}^\eta$ is a dense subset of $\mathcal{F}^0$, we implicitly assume that $\eta > 0$ is large so that the difference between $\mathcal{F}^0$ and $\mathcal{F}^\eta$ is "small."

3.6. **Hidden Cost.** We note that $\mathcal{A}_{CR}$ assumes away two sources of computational costs. First, $\mathcal{A}_{CR}$ presumes that the extraction of private information is free. In the revelation game, each player reports his valuation to the mechanism, which assigns the good with a specific probability conditioned on the profile of the reported types from the buyers. However, the revelation game is a mathematical tool to investigate an actual trading protocol. When we buy a good from a monopolist, we are rarely asked to report our reservation value. Instead, we pay the post price, or place a bid instead of reporting our reservation value. To infer the underlying valuation, the monopolistic seller has to invert the strategy of a buyer, which is not a trivial exercise for a boundedly rational seller who does not have a full specification of the demand curve. To extract private information, the monopolist has to implement a protocol to communicate with the buyers in addition to the existing rule, which can be quite expensive. It would be more economical, if the algorithm relies only on the aggregate outcome from the trading instead of the private and micro level data such as the reservation value of an individual buyer. A good example of an aggregate data would be the total quantity traded and the price at which the good is delivered in each period. Price and quantity would be the natural outcomes of any trading protocol, which do not require a special institution to infer the private information from the observed action.

Second, the estimated distribution $\hat{F}_{i,t}$ is generally a non-linear function, which requires many parameters to identify. A recursive algorithm uses the estimator from the previous round, and needs to store the present forecast for the next period. It would be costly to remember a large number of parameters in each period. A boundedly rational seller would prefer an alternative algorithm that can PAC guarantees $\mathcal{F}$ which requires the smallest number of parameters to remember.

3.7. **Complexity Cost.** We modify the preference of the monopolistic seller, incorporating the complexity cost. Let A be the collection of recursive algorithms satisfying **A1**$-$**A4**. The monopolist has a lexicographic preference over the long run average payoff and the input data complexity of an algorithm as in Rubinstein (1986). The monopolist chooses an algrithm in A which requires a minimal input data for the algorithm.

Since $\mathcal{A}$ is a recursive algorithm, the input in period t is $(\mathcal{A}(\mathcal{D}_{t-1}), D_t)$. Define

$$\dim(\mathcal{A}(\mathcal{D}_{t-1}), D_t) = \dim(\mathcal{A}(\mathcal{D}_{t-1})) + \dim(D_t)$$

as the number of variables the algorithm needs in period t. We only consider D_t which is a subset of the outcomes of the game in period t. D_t does not include any private information of the buyer.

Invoking the iterative law of conditional expectations, we re-write the average payoff the monopolist generated by algorithm $\mathcal{A}$ as

$$\mathcal{U}(\mathcal{A}) = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\sum_{t=1}^{T} \mathbb{E}_t\varphi_p(\mathcal{A}(\mathcal{D}_t))\varphi_q(\mathcal{A}(\mathcal{D}_t))$$

if the limit is well defined, and 0 otherwise.

In the equilibrium model as well as in the conventional learning models, the specification of the model is fixed. We can define $\mathbb{E}_t$ with respect to the probability distribution induced by the given specification. In the equilibrium model, the specification is accurate and

therefore, $\mathbb{E}_t$ operator is defined by the actual distribution. In our case, the specification is evolving over time, and thus, the probability distribution to define $\mathbb{E}_t$ changes, as the monopolist observe more data. Since we require $\mathcal{A}$ PAC guarantees F, the monopolist learns the actual distribution in the long run.

In principle, the monopolist searches the optimal price through an algorithm

$$\max_{\mathcal{A}\in\mathsf{A}} \mathcal{U}(\mathcal{A}) \tag{3.2}$$

subject to the constraint that if for any $\mathcal{A}' \in \mathsf{A}$ and sequence of data $\{D'_t\}$ generated by $\mathcal{A}'$, $\dim(\mathcal{A}'(\mathcal{D}'_{t-1}), D'_t) < \dim(\mathcal{A}(\mathcal{D}_{t-1}), D_t)$, then

$$\mathcal{U}(\mathcal{A}) > \mathcal{U}(\mathcal{A}').$$

Given that a recursive algorithm can achieve an optimal solution in the long run, no other algorithm with simpler input data complexity can achieve the same long run payoff as $\mathcal{A}$.

Cole and Roughgarden (2014) proved that the non-parametric estimation algorithm of the distribution PAC guarantees $\mathcal{F}^\eta$. Important questions are whether a simpler algorithm, in particular a parametric estimation algorithm, can PAC guarantee $\mathcal{F}^\eta$, and if one exists, what would be the simplest algorithm. We offer a positive answer to the question by constructing a simplest algorithm that can PAC guarantee $\mathcal{F}^\eta$.[5]

We first show that the any algorithm in A that can PAC guarantee $\mathcal{F}^\eta$ must take at least 2 observations from the market, such as a pair of the delivery price and the aggregate realized demand.

Proposition 3.5. *If $\mathcal{A} \in \mathsf{A}$ solves (3.2), then $\dim D_t \geq 2$ and $\dim \mathcal{A}(\mathcal{D}_{t-1}) \geq 2$.*

Proof. See Appendix A. □

4. Construction

We construct a recursive algorithm with $\dim(\mathcal{A}(\mathcal{D}_{t-1})) = \dim(D_t) = 2$, which PAC guarantees $\mathcal{F}^\eta$ while using the price and the actual demand as inputs.

4.1. Linear Recursive Algorithm.

$D_t = (q_t, p_t)$ is the pair of quantity q_t and the price p_t charged in period t. Recall that q_t is a random variable with $\mathbb{E}q_t = 1 - F(p_t)$. We can write

$$q_t = 1 - F(p_t) + \epsilon_{2,t}.$$

$\mathbb{E}\epsilon_{2,t} = 0$ and $\mathbb{E}\epsilon_{2,t}^2 < \infty$ uniformly, but the actual size of the second moment can depend on p_t and $F \in \mathcal{F}$.

Let

$$\mathcal{D}_t = (D_1, \ldots, D_{t-1})$$

be the history at the beginning of period t. The monopolist assume that the aggregate demand is a linear function:

$$q = \beta_0 + \beta_1 p$$

and estimates (β_0, β_1) according to the least square estimation over $\mathcal{D}_t$. Let $\mathcal{H}$ be the set of all linear demand functions, parameterized by (β_0, β_1).

[5]We admit the possibility that more than one simplest algorithm may exist.

Since $F \in \mathcal{F}^\eta$, the optimal solution $b^*(F)$ must generate a positive profit. Thanks to the uniform bound η, there exists a compact set K in the inteior of $\mathcal{R}^2_+$ such that $b^*(F) \in K$ $\forall F \in \mathcal{F}^\eta$. Thus, (β_0, β_1) must be such that the optimal price and the expected quantity under the linear demand curve parameterized by (β_0, β_1) must be contained in K.

Recall that if (β_0, β_1) can support $bj(F)$ for some $F \in \mathcal{F}^\eta$,

$$1 - F(b^*(F)) = \frac{\beta_0}{2} \text{ and } b^*(F) = -\frac{\beta_0}{2\beta_1}$$

or equivalently,

$$\beta_0 = 2(1 - F(b^*(F)) \text{ and } \beta_1 = -\frac{1 - F(b^*(F))}{b^*(F)}.$$

Since $(1 - F(b^*(F)), b^*(F)) \in K$, there exists a compact set $\mathsf{B} \subset (0, \infty) \times (-\infty, 0)$ such that $(\beta_0, \beta_1) \in \mathsf{B}$ if the linear demand can support a true optimal price. If $(\beta_0, \beta_1) \notin \mathsf{B}$, then the monopolist can conclude that the estimated demand is wrong, based on what the monopolist knows.

Let $\mathcal{H}_\mathsf{B} \subset \mathcal{H}$ be the collection of the linear demand, which induces the market outcome in K. The seller knows that the market demand curve is in $\mathcal{H}_\mathsf{B}$. Since $\mathcal{F}$ typically contains non-linear demand curves, $\mathcal{H}_\mathsf{B}$ is a misspecified model. But, the number of parameters the seller has to keep track of is minimal.

Since the seller does not know (β_0, β_1), the seller estimates the parameters using the least square estimation method recurisvely, while choosing the pricing rule based on the estimated linear demand curve. Let $(\beta_{0,t-1}, \beta_{1,t-1})$ be the least square estimator at the end of period $t-1$. Given the estimated demand curve

$$q = \beta_{0,t-1} + \beta_{1,t-1} p,$$

the monopolist calculates the optimal price

$$-\frac{\beta_{0,t-1}}{2\beta_{1,t-1}}$$

but incurs an implementation error $\epsilon_{1,t}$ so that the actual price in period t is

$$p_t = -\frac{\beta_{0,t-1}}{2\beta_{1,t-1}} + \epsilon_{1,t}$$

where $\epsilon_{1,t}$ is i.i.d. with $\mathbb{E}\epsilon_{1,t} = 0$ and $\mathbb{E}\epsilon^2_{1,t} = \sigma^2_1$. We interpret $\epsilon_{1,t}$ as the implementation error, or the small experimentation by the monopolist seller. We choose $\epsilon_{1,t}$ from a small interval $[-\epsilon, \epsilon]$ according to a fixed distribution, say the uniform distribution. We control the size of $\epsilon > 0$ to achieve the desired level of accuracy of the algorithm.

The monopolist forecasts that the sales quantity will be

$$\beta_{0,t-1} + \beta_{1,t-1}\left[-\frac{\beta_{0,t-1}}{2\beta_{1,t-1}} + \epsilon_{1,t}\right] = \frac{\beta_{0,t-1}}{2} + \beta_{1,t-1}\epsilon_{1,t}$$

but the actual quantity is

$$1 - F\left(-\frac{\beta_{0,t-1}}{2\beta_{1,t-1}} + \epsilon_{1,t}\right) + \epsilon_{2,t}.$$

The forecasting error is

$$\phi(\beta_{t-1},\epsilon_t) = 1 - F\left(-\frac{\beta_{0,t-1}}{2\beta_{1,t-1}}\right) + \epsilon_{1,t} + \epsilon_{2,t} - \frac{\beta_{0,t-1}}{2} - \beta_{1,t-1}\epsilon_{1,t}$$

where $\beta_{t-1} = (\beta_{0,t-1}, \beta_{1,t-1})$ and $\epsilon_t = (\epsilon_{1,t}, \epsilon_{2,t})$.

Since the quantity must be in closed interval $[0,1]$, we first take care of the cases of "corner solution" before moving to "interior solution." If actual quantity

$$q_t = 1 - F\left(-\frac{\beta_{0,t-1}}{2\beta_{1,t-1}}\right) + \epsilon_{1,t}$$

is at the boundary, we directly update $(\beta_{0,t-1}, \beta_{1,t-1})$. If $q_t = 0$, then the algorithm concludes that the forecast price was too high, and adjust accordingly:

$$\beta_{0,t} = \beta_{1,t} - a \quad \text{and} \quad \beta_{1,t} = \beta_{1,t-1}$$

so that

$$b_{t+1} = -\frac{\beta_{0,t}}{2\beta_{1,t}} = -\frac{\beta_{0,t-1}}{2\beta_{1,t-1}} + \frac{a}{2\beta_{1,t-1}} = b_t + \frac{a}{2\beta_{1,t-1}} < b_t,$$

where $a > 0$ is a parameter which will be determined to satisfy the precision and the confidence requirements. Similarly, if $q_t = 1$,

$$\beta_{0,t} = \beta_{1,t} + a \quad \text{and} \quad \beta_{1,t} = \beta_{1,t-1}$$

so that

$$b_{t+1} = b_t - \frac{a}{2\beta_{1,t-1}} > b_t.$$

If $0 < q_t < 1$, then

$$\begin{bmatrix}\beta_{0,t}\\ \beta_{1,t}\end{bmatrix} = \begin{bmatrix}\beta_{0,t-1}\\ \beta_{1,t-1}\end{bmatrix} + a_t R_{t-1}^{-1} \begin{bmatrix}1\\ -\frac{\beta_{0,t-1}}{2\beta_{1,t-1}} + \epsilon_{1,t}\end{bmatrix} \phi(\beta_{t-1},\epsilon_t)$$

and

$$R_t = R_{t-1} + a_t \left(\begin{bmatrix} 1 & -\frac{\beta_{0,t-1}}{2\beta_{1,t-1}} + \epsilon_{1,t} \\ -\frac{\beta_{0,t-1}}{2\beta_{1,t-1}} + \epsilon_{1,t} & \left(-\frac{\beta_{0,t-1}}{2\beta_{1,t-1}} + \epsilon_{1,t}\right)^2 \end{bmatrix} - R_{t-1} \right).$$

The algorithm recursively calculates the least square estimator if $a_t = \frac{1}{t}$. We need to impose a bound to $(\beta_{0,t}, \beta_{1,t})$ to keep the estimator within a compact set.

Let $\mathcal{B}$ be a compact convex set which contains B in the interior of $\mathcal{B}$ so that the Hausdorff distance between $\mathcal{B}$ and B is positive. If $(\beta_{0,t}, \beta_{1,t}) \notin \mathcal{B}$, then the seller can conclude that the estimator is out of the line and needs to adjust the estimator by pushing it back to B.[6]

We modify the baseline updating scheme to construct the formal updating scheme for $(\beta_{0,t}\beta_{1,t})$.

$$\begin{bmatrix}\beta_{0,t}\\ \beta_{1,t}\end{bmatrix} = \begin{bmatrix}\beta_{0,t-1}\\ \beta_{1,t-1}\end{bmatrix} + a_t R_{t-1}^{-1} \begin{bmatrix}1\\ -\frac{\beta_{0,t-1}}{2\beta_{1,t-1}} + \epsilon_{1,t}\end{bmatrix} \phi(\beta_{t-1},\epsilon_t)$$

[6]This mapping is known as the projection facility in the literature of the stochastic approximation (Kushner and Yin (1997)).

if the right hand side is in $\mathcal{B}$. Otherwise, $(\beta_{0,t}, \beta_{1,t}) = (\beta_0, \beta_1) \in \mathsf{B}$ for some fixed (β_0, β_1) in the interior of B.

Treating the estimated demand curve

$$q = \beta_{0,t-1} + \beta_{1,t-1} p$$

as the real demand curve, the seller sets the price

$$p_t = -\frac{\beta_{0,t-1}}{2\beta_{1,t-1}} + \epsilon_{1,t}$$

where $\epsilon_{1,t}$ is an i.i.d. white noise uniformly distributed over $[-\epsilon, \epsilon]$ for a fixed $\epsilon > 0$ so that $\mathbb{E}\epsilon_{1,t} =$ and $\mathbb{E}\epsilon_{1,t}^2 = \sigma_1^2$. Given p_t, the quantity in period t

$$q_t = 1 - F(p_t) + \epsilon_{2,t}$$

is realized. Using (q_t, p_t), the seller updates $(\beta_{0,t-1}, \beta_{1,t-1})$ to $(\beta_{0,t}, \beta_{1,t})$. The translating function

$$\varphi(\beta_0, \beta_1) = \left(-\frac{\beta_{0,t-1}}{2\beta_{1,t-1}}, 1 - F\left(-\frac{\beta_{0,t-1}}{2\beta_{1,t-1}} + \epsilon_{1,t} \right) \right)$$

maps the estimated linear demand curve into the mean forecast price and quantity.

We need to specify one remaining parameter of the algorithm: a_t. Instead of $a_t = 1/t$, we set $a_t = a > 0$. We have to choose a to achieve the desired level of accuracy and confidence. Let $\mathcal{A}_a$ be the recursive algorithm with constant gain $a_t = a > 0$ $\forall t \geq 1$. $\forall a \geq 0$, algorithm $\mathcal{A}_a$ produces β_t following history $\mathcal{D}_t$, where

$$\mathcal{D}_t = ((q_1, p_1), \ldots, (q_{t-1}, p_{t-1}))$$

is the sequence of aggregate market outcomes up to period $t-1$. The constructed algorithm is recursive: $\mathcal{A}_a(\mathcal{D}_t) = \beta_t$ is the output of the algorithm based on D_t and $\mathcal{A}_a(\mathcal{D}_{t-1})$. The input complexity

$$\dim(\mathcal{A}_a(\mathcal{D}_{t-1})) = 2 \text{ and } \dim(D_{t-1}) = 2,$$

which is the minimum among all algorithms in A.

To emphasize the fact that the optimal price b^* is a function of the underlying (aggregate) distribution F, we sometimes write $b^*(F)$ instead of b^*. Let $\beta^*(F)$ be the pair of estimators that induce the optimal price for F:

$$b^*(F) = -\frac{\beta_0^*(F)}{2\beta_1^*(F)} \text{ and } q^*(F) = 1 - F(b^*(F)).$$

Theorem 4.1. *$\forall \mu > 0$, $\exists \tau(\mu) > 0$, $\rho > 0$ and $\overline{a} > 0$ such that $\forall a \in (0, \overline{a})$, if $T = \lceil \frac{\tau(\mu)}{a} \rceil$ $\forall a \in (0, \overline{a})$,*

$$\mathbb{P}\left(\left| \varphi\left(\mathcal{A}_a(\mathcal{D}_{T(a,\mu)}) \right) - (b^*(F), q^*(F)) \right| > 4\mu \right) \leq e^{-\rho T(a,\mu)} \qquad \forall F \in \mathcal{F}^\eta$$

and $\tau(\mu) \sim -\log \mu$ for small $\mu > 0$ and $\overline{a} = \mathsf{O}\left(-\frac{\mu}{\log \mu} \right)$.

Proof. See Appendix B. □

For fixed $\lambda > 0$ less than 1, we can choose $\overline{a} > 0$ sufficiently small so that

$$e^{-\frac{\rho\tau(\mu)}{\overline{a}}} = \lambda$$

and therefore,

$$\frac{\tau(\mu)}{\overline{a}} = -\frac{\log\lambda}{\rho}.$$

Since

$$T(\overline{a},\mu) = \left\lceil \frac{\tau(\mu)}{\overline{a}} \right\rceil = \left\lceil -\frac{\log\lambda}{\rho} \right\rceil,$$

$T(\overline{a},\mu)$ increases at the logarithmic speed with respect to $1/\lambda$. An important observation is that the accuracy of the estimator depends only on $\tau(\mu)$ and the approximation error vanishes linearly as $a \to 0$. Thus, the number of the minimum time step to satisfy the accuracy requirement increases at the rate of $-\frac{\log\mu}{a}$. Theorem 4.1 implies that $\mathcal{A}_a$ PAC guarantees $\mathcal{F}^\eta$.

5. Heuristics

Let us provide a heuristic explantion about how a monopolist can learn to choose an optimal price of a non-linear demand curve through a linear demand curve, *uniformly over the set of non-linear demand curves* satisfying a few basic properties. Suppose F be a distribution function of buyer's valuations with increasing hazard rate property where the support of F is $[\underline{p},\overline{p}]$ with $\overline{p} > \underline{p}$. The crucial implications of the increasing hazard rate property is that

$$\frac{1-F(p)}{f(p)} \tag{5.3}$$

is decreasing. Thus,

$$\frac{1-F(p)}{f(p)} - p \tag{5.4}$$

is strictly decreasing, whose slope is steeper than -1.

Let $\mathcal{F}^0$ be the collection of all possible distribution of valuations satisfying (5.3). $\forall F \in \mathcal{F}^0$, the optimal price $b^*(F)$ associated with F is unique and completely characterized by the first order condition

$$\frac{1-F(b^*(F))}{f(b^*(F))} - b^*(F) = 0. \tag{5.5}$$

Since a buyer with valuation v purchase the product only if $v \geq p$, $1-F(p)$ is the market demand at p, and $1-F(p)-pf(p)$ is the marginal revenue at p, where $f(p) = F'(p)$ is the density function of F. We are interested in $F \in \mathcal{F}^0$ whose marginal revenue curve is a strictly decreasing Lipschitz continuous marginal revenue curve. Note that $\mathcal{F}^0$ is a bounded convex set with respect to the pointwise convergence topology. Let us consider a convex compact subset $\mathcal{F}^\eta$ of $\mathcal{F}^0$ where the marginal revenue curve is uniformly Lipschitz continuous:

$$|f(p)-f(p')| \leq \eta|p-p'| \qquad \forall p,p' \in [\underline{p},\overline{p}],\ \forall F \in \mathcal{F}^\eta.$$

Note that $\mathcal{F}^\eta \subset \mathcal{F}^{\eta'}$ if $\eta \leq \eta'$, and $\cup_{\eta\geq 0}\mathcal{F}^\eta$ is a dense subset of $\mathcal{F}^0$.

Suppose that the monopolist knows $F \in \mathcal{F}^\eta$ for some $\eta > 0$, even though the monopolist does not know exactly what F is. Let $b^*(F)$ be the optimal price if F were known to the

monopolist. Thanks to the increasing hazard rate property, $b^*(F)$ is the unique of (5.5) for $\forall F \in \mathcal{F}^\eta \subset \mathcal{F}^0$. Since the monopolist needs an estimated demand curve to choose a price, she assumes the market demand curve is

$$q = \beta_0 + \beta_1 p$$

and estimates (β_0, β_1). Since $F \in \mathcal{F}^\eta$ is typically non-linear, the linear model entertained by the monopolist is misspecified in a classical econometric sense. On the other hand, the linear demand is the simplest model to formulate a non-trivial downward sloping market demand function.

Let $(\beta_{0,t-1}, \beta_{1,t-1})$ be the estimator at the end of period $t-1$. The monopolist updates the estimator to $(\beta_{0,t}, \beta_{1,t})$ after she observes price p_t and quantity q_t in period t. The optimal price associated with linear demand $q = \beta_{0,t-1} + \beta_{1,t-1} p$ is

$$b_t = -\frac{\beta_{0,t-1}}{2\beta_{1,t-1}}.$$

Let us assume the monopolist uses a simple error correction process to update the estimator, which include the least square esstimation and many other widely used estimation processes. Given $(\beta_{0,t-1}, \beta_{1,t-1})$, the monopolist charges price according to

$$p_t = b_t + \epsilon_{1,t}. \tag{5.6}$$

The monopolist experiments a little by adding a small noise $\epsilon_{1,t}$ which is drawn from a binomial distribution over $\{-\epsilon, \epsilon\}$ for small $\epsilon > 0$. Given b_t, the quantity forecast is

$$\beta_{0,t-1} + \beta_{1,t-1} p_t$$

but the actual demand is

$$q_t = 1 - F(p_t).$$

The forecasting error is therefore

$$\phi_t = 1 - F(p_t) - \left[\beta_{0,t-1} + \beta_{1,t-1} p_t\right]$$

which can be written as

$$\phi_t = 1 - F(b_t + \epsilon_{1,t}) - \left[\beta_{0,t-1} + \beta_{1,t-1} b_t + \beta_{1,t-1}\epsilon_{1,t}\right]$$

The monopolist adjusts $(\beta_{0,t}, \beta_{1,t})$ according to the value of ϕ_t and $(\beta_{0,t-1}, \beta_{1,t-1})$.

The noise added allows the monopolist to locally estimate the slope of the demand curve, while the forecast error allows for local estimation of the intercept. This intuition, however, still falls short of providing a sense of how much confidence the monopolist might have. Our first question is whether

$$\lim_{t\to\infty} b_t = b^*(F).$$

in probability, or equivalently $\forall \mu > 0$, $\exists T(\mu, F)$ such that

$$\mathbb{P}\left(\left|b_t - b^*(F)\right| > \mu\right) < \mu.$$

We often require that the confidence $1-\mu$ of the forecast improve exponentially: $\exists \rho$ such that $\forall t \geq T(\mu, F)$,

$$\mathbb{P}\left(\left|b_t - b^*(F)\right| > \mu\right) < e^{-\rho t}.$$

This type of questions has been extensively investigated in the existing recursive learning literature (e.g., Evans and Honkapohja (2001) and Fudenberg and Levine (1998)).

Our task is more challenging, as we learn $\mathcal{F}^\eta$ instead of an individual $F \in \mathcal{F}^\eta$. To this end, we need to discipline $T(\mu, F)$. The existing notion of convergence allows $T(\mu, F)$ to depend on F. The same algorithm can give an accurate answer with a small number of draws for some F, while it can require extremely large number of draws to meet the accuracy bound for other F. If so, it is difficult to say that the monopolist learns $\mathcal{F}^\eta$.

We strengthen the convergence criterion accordingly. The monopoly learns to behave optimally over $\mathcal{F}^\eta$ if $\forall \mu > 0$, $\exists \rho > 0$, $\exists T(\mu)$ such that $\forall t \geq T(\mu)$,

$$\mathbb{P}\left(\left|b_t - b^*(F)\right| > \mu\right) < e^{-\rho t}. \tag{5.7}$$

This notion of convergence turns out to be equivalent to PAC (probably almost correct) learnability, which is the fundamental notion of convergence in the machine learning literature (Shalev-Shwartz and Ben-David (2014)). To discipline the number of data to achieve the required accuracy, we often require that $T(\mu)$ increases at a polynomial rate with respect to $1/\mu$. This paper shows that the least square estimation algorithm of the linear demand satisfies (5.7).

We state (5.7) in a more convenient form for numerical analysis, albeit slightly weaker.[7] Given accuracy bound $\mu > 0$ and confidence requirement $\lambda > 0$ for small $\lambda < 1$, we construct a statistical procedure $\mathcal{A}$ and a stopping time (μ, λ) such that the answer generated by $\mathcal{A}$ at the end of $T(\mu, \lambda)$ satisifes

$$\mathbb{P}\left(\left|b_{T(\mu,\lambda)} - b^*(F)\right| > \mu\right) < \lambda \tag{5.8}$$

where $T(\mu, \lambda) \sim \mathsf{O}\left(-\frac{\log \lambda}{\mu}\right)$. We regard (5.7) as the uniform convergence in probability, while (5.8) as the uniform weak convergence.

We are now in a position to provide some more precise details regarding how we proceed with the analysis, as well as some of the technical difficulties we face in obtaining our main results. Roughly speaking, we decompose the dynamics of b_t into two pieces:

$$b_t - b^*(F) = \mathbb{E}b_t - b^*(F) + (b_t - \mathbb{E}b_t)$$

and approximate the probability in (5.7) by averging b_t over time:

$$\frac{1}{T}\sum_{t=1}^{T}\left[b_t - b^*(F)\right] = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[b_t - b^*(F)\right] + \frac{1}{T}\sum_{t=1}^{T}\left[b_t - \mathbb{E}b_t\right].$$

To investigate the asymptotic properties of

$$\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[b_t - b^*(F)\right],$$

we examine the dynamics of the mean of b_t through continous time process. Since the dynamics of b_t is driven by the dynamics of $(\beta_{0,t}, \beta_{1,t})$, we need to examine the mean dynamics of $(\beta_{0,t}, \beta_{1,t})$. Let us fix (β_0, β_1) and calculate the expected value of the estimator

[7] (5.7) is essentially the uniform convergence to the truth in probability, while (5.8) is the uniform convergence to the trueth in distribution (weak convergence). We can modify our algorithm satisfying (5.8) in a straightforward way to satisfy (5.7) and vice versa. See Section 7.1.

in the "next period" if the monopolist chooses the estimator to minimize the forecasting error. Given (β_0, β_1), the next period's quantity q' and price p' are

$$(q', p') = \begin{cases} \big(b+\epsilon, 1-F\left(b+\epsilon\right)\big) & \text{with probability 0.5} \\ \big(b-\epsilon, 1-F\left(b-\epsilon\right)\big) & \text{with probability 0.5.} \end{cases}$$

where

$$b = -\frac{\beta_0}{2\beta_1}. \tag{5.9}$$

The monopolist chooses the new coefficients (β_0', β_1') to fit the observed data best, the new regression line must pass through $\big(b+\epsilon, 1-F\left(b+\epsilon\right)\big)$ and $\big(b-\epsilon, 1-F\left(b-\epsilon\right)\big)$. A simple calculation shows

$$\begin{aligned} \beta_0' &= 1 - F\left(b\right) + bf(b) \\ \beta_1' &= -f(b) \end{aligned}$$

modulo linear approximation error at the order of ϵ^2.

The stochastic approximation theory (Kushner and Yin (1997)) shows that the asymptotic properties of the mean of $(\beta_{0,t}, \beta_{1,t})$ is dictated by the dynamic properties of the associated ordinary differential equation (ODE)

$$\begin{aligned} \dot{\beta}_0 &= \beta_0' - \beta_0 = -f(b) - \beta_0 \\ \dot{\beta}_1 &= \beta_1' - \beta_1 = 1 - F(b) + bf(b) - \beta_1. \end{aligned} \tag{5.10}$$

Since (5.9) holds in every period, we take the time derivative on both sides of the equality to have

$$\dot{b} = -\frac{1}{2\beta_1}\left(\dot{\beta}_1 + 2b\dot{\beta}_0\right).$$

After substituting $\dot{\beta}_1$ and $\dot{\beta}_0$, we have

$$\dot{b} = -\frac{f(b)}{2\beta_1}\left[\frac{1-F(b)}{f(b)} - b\right]. \tag{5.11}$$

Since the demand curve $1 - F(p)$ is strictly decreasing, $\beta_1 < 0$. Thus, $-\frac{f(b)}{2\beta_1} > 0$. The term inside of the bracket has a unique solution $b^*(F)$, which is the profit maximizing price for distribution F. By the increasing hazard rate property, the term in the bracket is strictly decreasing with respect to b, which makes $b^*(F)$ a stable stationary solution of (5.11). Thus, the least square learning algorithm converges to $b^*(F)$ (Kushner and Yin (1997)).

So far, we only proves the convergence "pointwise" with respect to F, allowing the number of data needed to achieve the desired level of accuracy can depend on F. To show the uniform convergence over $\mathcal{F}^\eta$, we need to do additional work. Since $\frac{1-F(b^*(F))}{f(b^*(F))} - b^*(F) = 0$,

$$\frac{1-F(b)}{f(b)} - b = \frac{1-F(b)}{f(b)} - b - \left(\frac{1-F(b^*(F))}{f(b^*(F))} - b^*(F)\right).$$

By (5.4),

$$\frac{1-F(b)}{f(b)} - b - \left(\frac{1-F(b^*(F))}{f(b^*(F))} - b^*(F)\right) \leq -(b - b^*(F)).$$

We can show that $\exists c > 0$ such that[8]

$$\dot{b} = (b - \dot{b^*}(F)) \leq -c(b - b^*(F)) < 0$$

if $b > b^*(F)$. Similarly, if $b < b^*(F)$, then

$$\dot{b} \geq -c(b - b^*(F)) > 0$$

for any $F \in \mathcal{F}^\eta$. Since $\sup_{F \in \mathcal{F}^\eta} b^*(F) \leq \overline{p}$, the initial condition of the ordinary differential equation can be selected from a compact set. Since the ditance $|b - b^*(F))|$ vanishes uniformly at the order of e^{-ct}, it takes $\mathsf{O}(-\log \mu)$ amount of time for b to enter the μ neighborhood of $b^*(F)$. From this observation, we prove that the amount of data to approximate the optimal price is uniform over $\mathcal{F}^\eta$. We can also show that the number of data to achieve μ accuracy increases at the polynomial rate of $1/\mu$.

The remaining step is to examine $b_t - \mathbb{E}b_t$ or more precisely, the distribution of

$$\frac{1}{T}\sum_{t=1}^{T} \left[b_t - \mathbb{E}_{t-1} b_t\right].$$

The average converges to 0, but we need to find $\rho > 0$ such that

$$\mathbb{P}\left(\left|\frac{1}{T}\sum_{t=1}^{T} \left[b_t - \mathbb{E}_{t-1} b_t\right]\right| > \mu\right) \leq e^{-\rho T} \qquad \forall F \in \mathcal{F}^\eta.$$

This part of the exercise is essentially to calculate the tail portion of probability distribution of b_t. For a fixed F, the existence of $\rho > 0$ can be proved by investigating the large deviation properties (Dembo and Zeitouni (1998)) of a recursive algorithm (Dupuis and Kushner (1989)). The exercise is more challenging, because we are searching for $\rho > 0$ *uniformly* over the set of feasible distributions.

If the obseved data were IID as in Cole and Roughgarden (2014), we could invoke the large deviation property of the sample average, such as Chernoff's bound, to prove that the tail probability vanishes at the exponential rate uniformly over $\mathcal{F}^\eta$.

In our case, the algorithm uses (q_t, p_t). Even though the valuation of buyers are drawn IID over time, (q_t, p_t) is not IID (not even martingale), because of $p_t = b_t + \epsilon_{1,t}$ and b_t is responding to the realized sequence of data. The data generating process is endogenous, which makes stochastic process (q_t, p_t) highly non-stationary. As a result, $b_t - \mathbb{E}_{t-1} b_t$ is not IID but a martingale difference. We need to invoke Azuma-Hoeffding-Bennett inequality (Dembo and Zeitouni (1998)) to calculate the uniform exponential rate for all feasible distributions of buyer's valuation, which proves that our algorithm is efficient (Shalev-Shwartz and Ben-David (2014)).

Through the linear demand curves, the seller *can* still learn how to choose the optimal price at an exponential rate, *even though* they may be grossly misspecified in terms of the demand curves that they consider. Thus, within a polynomial time, the monopolistic seller behaves as if he knows the actual demand curve. Even though a non-parametric estimation of the demand curve is feasible as in Cole and Roughgarden (2014), the monopolist chooses

[8]If we assume that $\inf_p f(p) > 0$ uniformly, the proof is easy. Without this assumption, we need some additional work (Lemma B.1).

to use a simple, yet misspecified, model of demand curve to choose his price to save the computational cost.

6. Numerical Experiments

While our algorithm $\mathcal{A}_a$ has the same asymptotic properties as the algorithm of Cole and Roughgarden (2014), we need to rely on the numerical analysis to compare the performance of the algorithm with a finite number of data. It is possible to write the algorithm of Cole and Roughgarden (2014) in a recursive form. Practically, their algorithm is best implemented in an off-line form. In a recursive formulation of the algorithm of Cole and Roughgarden (2014), the algorithm has to store the empirical distribution of the valuation after t period, which almost always requires to remember t data points. On the other hand, our algorithm $\mathcal{A}_a$ only remembers $(\beta_{0,t}, \beta_{1,t})$ after t period.

Since the recursive algorithm relies on a summary of the past information, the implementation of the algorithm is simpler than the off-line algorithm. By the same token, it is possible that the convergence can be slower than the off-line algorithm which relies on the information practically from the entire history.

We generates $\mathcal{F}^\eta$ from a truncated Gaussian distribution at the mean, where the Gaussian distribution has mean 10. By changing the standard deviation of the Gaussian distribution, we can generate different distributions of the valuations, each of which satisfies the increase hazard rate property and other regularity properties. More specifically, we select a standard deviation from 10000 samples, ranging from 9 to 19 with increment 0.001. For each distribution of valuations, we calculate the true optimal price. As the standard deviation becomes larger, so does the true optimal price.

We assume that there are 100 buyers, whose reservatin value is drawn from the same distribution $F \in \mathcal{F}$. The actual amount of delivery for a price is thus random, whose mean is $1 - F(p)$.

For each distribution, we run our algorithm $\mathcal{A}_a$ with $a = 0.00015$ and $\epsilon = 0.75$ (the size of price perturbation) for $T = 150,000$ rounds. At the end of T rounds, we calculate the price forecast and compare it to the true optimal price to calculate the forecasting error.

We calculate the optimal price forecast from Cole and Roughgarden (2014) by drawn K numbers of valuations each period for T rounds, to calculate the empirical distribution and the optimal price from the empirical distribution. We calculate the forecast error.

Per Period	CR		CL	
	Mean	Variance	Mean	Variance
2	−0.0008	0.0173	0.0085	0.0062
4	−0.0033	0.0116	0.0097	0.0062
10	−0.0016	0.0060	0.0086	0.0063

Table 1: The first column reports the number of valuations the algorithm of Cole and Roughgarden [2014] takes per period. By CR and CL, we mean the algorithm of Cole and Roughgarden [2014] and this paper. The true profit maximizing price is roughly ranging from 10 to 20.

Table 1 reports means and variances from the numerical exercies. Note that the first and the second moments of CL remain stable. Because the gain function a of $\mathcal{A}_a$ remains constant, the forecating error converges to a non-degenerate stationary distribution. Because of linear approximation of a non-linear demand curve, the linear approximation error contributes the forecasting error which vanishes as we reduce the size of price perturbation $\epsilon > 0$.

Because the non-parametric estimation method in Cole and Roughgarden (2014) does not incur any linear approximation error, the mean of the forecasting error is closer to 0. Interestingly, the size of the variance is larger for the same number of data as $\mathcal{A}_a$, which takes two data points (price and quantity) in each period. If Cole and Roughgarden (2014) allows the algorithm to receive 10 valuation reports (which is five times as much as data $\mathcal{A}_a$ uses), the variance becomes comparable.

Figure 1 reports the distribution of the forecasting errors of CL in blue and CR in orange, when CR takes two valuation reports per period. Note that the distribution of forecasting errors of CR is spread out more than that of CL. Figure 2 reports the distribution of forecasting errors when CR takes ten reported values in each period. The two distribution of the forecasting errors become closer as Table 1 indicates.

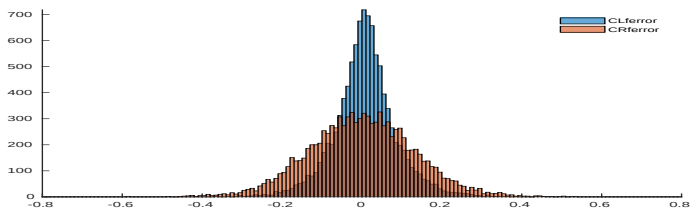

Figure 1: Blue and orange bars represent the density of forecasting errors of our algorithm $\mathcal{A}_a$ (marked as CL) and the algorithm of Cole and Roughgarden [2014] (marked as CR). If the algorithm of Cole and Roughgarden [2014] takes 2 reports per period, the variance is twice as large as that of $\mathcal{A}_a$.

The performance of our algorithm $\mathcal{A}_a$ is comparable to that of Cole and Roughgarden (2014). We can reduce the linear approximation error by reducing the price perturbation, choosing a smaller $\epsilon > 0$. We can reduce the variance of the forecasting error by reducing the gain $a > 0$. On the other hand, Cole and Roughgarden (2014) can reduce the variance of the forecasting error simply by collecting more data about the valuations.

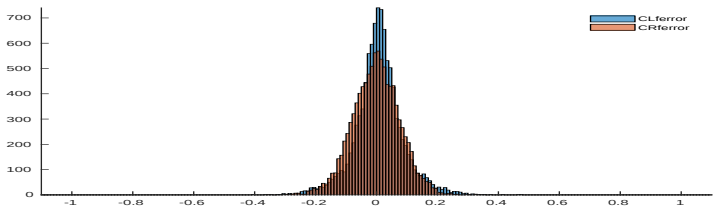

Figure 2: If the algorithm of Cole and Roughgarden [2014] takes 10 reports, the variance is comparable to that of $\mathcal{A}_a$.

The key advantanges of our algorithm are the source of data and the recursive nature of the algorithm. Instead of extracting the private information of the buyer, our algorithm relies on the public outcome of the market in each period. In each period, the algorithm needs to remember just two parameters of the linear demand, while in Cole and Roughgarden (2014), the algorithm has to remember the empirical distribution which essentially requires to remember the total number of reported valuations.

7. Concluding Remarks

7.1. **Convergence in Probability.** A standard notion of learnability in the machine learning literature is PAC learnability.

Definition 7.1. *$\mathcal{A}$ is PAC learnable of $\mathcal{F}$ if $\forall \mu > 0$, $\forall \lambda \in (0,1)$, $\exists T(\mu,\lambda)$ such that*

$$\mathbb{P}\left(\left|\varphi(\mathcal{A}(D_t) - (b^*(F), q^*(F))\right| \geq \mu \quad \forall t \geq T(\mu,\lambda)\right) \leq \lambda \tag{7.12}$$

where $T(\mu,\lambda) \sim \mathsf{O}\left(-\frac{\log \lambda}{\mu^p}\right)$ for some $p > 0$.

We can modify $\mathcal{A}_a$ to satsify PAC learnability. Instead of a positive constant gain function $a > 0$, we can choose $a = \frac{1}{t^\omega}$ where $\omega \in (0,1)$. Let $\mathcal{A}_{\frac{1}{t^\omega}}$ be the modified algorithm often called a decreasing gain algorithm.

Invoking a well known result in the stochastic approximation (Kushner and Yin (1997)), we can show that (7.12) holds "pointwise" with respect to $F \in \mathcal{F}$. Following the same argument as the proof of Theorem 4.1, we obtain the uniform convergence to show that $\mathcal{A}_{\frac{1}{t^\omega}}$ PAC learns $\mathcal{F}^\eta$ where $T(\mu,\lambda) \sim \mathsf{O}\left(-\frac{\log \lambda}{\mu^{3-2\omega}}\right)$.

The key difference between (3.1) and (7.12) is that we use the convergence in distribution for (3.1) while (7.12) uses the convergence in probability. while (3.1) is a stronger notion of convergence than (3.1), we chose $\mathcal{A}_a$ over $\mathcal{A}_{\frac{1}{t^\omega}}$ for two practical reasons. First, we have to specify the stopping time of an algorithm to produce a forecast which the monopolist can use. By the time when the algorthm reaches $T(\mu, \lambda)$ rounds, the two algorithms produce a strong forecast with minimal difference. Second, the numerical analysis of $\mathcal{A}_a$ is much more robust than $\mathcal{A}_{1/t^\omega}$. As the gain function $1/t^\omega$ converges to 0 as t becomes larger, the numerical outcome tends to be very sensitive to a small change of other parameters. In contrast, the gain function of $\mathcal{A}_a$ remains constant, which makes it suitable for long simulations.

7.2. **Final Comments.** Throughout this paper, we regard the linear demand as a misspecified model of the actual demand, which is generally non-linear. A classical notion of the misspecification is that if a model does not include the true state, then the model is misspecified. We demonstrate that the decision maker can play the equilibrium outcome through the linear demand, as if he knows the actual demand curve. For an outsider, the decision maker behaves as if his model is correctly specified.

We claim that whether the linear demand function is a misspecified model depends on the object of interest of the economist. If the exercise of the economist involves the interaction between the model and the realized otcome, a misspecified model in the classical sense may not be misspecified any more.

A well known example would be the short term Phillips curve. If the true data is generated by the expectation augmented Phillips curve, the short term Phillips curve should be misspecified, in the sense that the short term Phillips curve misses the expectation of the agent in the true data generating mechanism of the expectation augmented Phillips curve.[9] This view was challenged by Esponda and Pouzo (2014) among others. If the agent is endowed with the short term Phillips curve and makes a decision based on his own model, the interaction between the agent's model and the true data generating process leads to a stable stationary point. Interestingly, the stable stationary point coincides with the Nash equilibrium of the model in which the agent is endowed with a correctly specified model, thus the short term Phillips curve is viewed by a correctly specified model by Esponda and Pouzo (2014).

Another example would be a direct revelation game, which is a mathematical model critical for the analysis of a broad class of auctions. We construct the direct revelation game, by replacing the bidding function of the agent by the reporting strategy. Thus, the revelation game does not have institutional details of the actual auction. Does it make the revelation game a misspecified model?

If we are interested in the equilibrium of the auction, the revelation game is a correctly specified model, because the revelation game induces the same outcome as the auction (by the construction of the revelation game). Suppose that we are interested in how the change of the action space affects the observed outcome. For example, imagine an economist who understand the realized difference of outcomes between the Dutch auction and the first price auction. Because of framing effect, the outcome tends to differ. The revelation game

[9]See Sargent (1999) for detailed discussion about the short term Phillips curve.

is not a correctly specified model for this type of exercise, because the revelation game does not have a parameter which captures the instiutional details of the actual auctions.

We maintain that the linear demand curve is a misspecified but simple model of the true demand curve, because the actual demand curve is non-linear. We demonstrate that a misspecified model can become a correctly specified model, without fully specifying the demand curve.

Appendix A. Proof of Proposition 3.5

We first prove that $\dim D_t \geq 2$. Since the market outcome in period t is (p_t, q_t), we show that if the algorithm uses only one of the two elements, the algorithm cannot PAC gurantee $\mathcal{F}$. We only show that if $D_t = \{p_t\}$, the algorithm fails to PAC gurantee $\mathcal{F}$. The proof for the case where $D_t = \{q_t\}$ follows the symmetric argument.

Let F and F^α be the distribution functions of the uniform distribution over intervals $[\underline{p}, \underline{p}+1]$ and $[\underline{p}+\alpha, \underline{p}+\alpha+1]$. Note that the optimal price of F^α is

$$b^*(F^\alpha) = \frac{1+\underline{p}+\alpha}{2}.$$

Suppose that $D_t = \{p_t\}$. Then,

$$\mathcal{A}(\mathcal{D}_t : F^\alpha) = \mathcal{A}(\mathcal{D}_t : F) \qquad \forall \alpha.$$

Both $\mathcal{A}(\mathcal{D}_t : F^\alpha)$ and $\mathcal{A}(\mathcal{D}_t : F)$ have the same input and therefore, generate the identical forecast for any F^α.

$$\varphi_p(\mathcal{A}(\mathcal{D}_t : F^\alpha)) = \varphi_p(\mathcal{A}(\mathcal{D}_t : F)) \qquad \forall \alpha, \forall t \geq 1. \tag{A.13}$$

where φ is the translating function of the forecast of $\mathcal{A}$ to the optimal price. Since $\mathcal{A}$ PAC guarantees $\mathcal{F}$,

$$\lim_{t\to\infty} \varphi_p(\mathcal{A}(\mathcal{D}_t : F^\alpha)) = b^*(F^\alpha) = \frac{1+\underline{p}+\alpha}{2} \qquad \forall \alpha$$

with probability 1. If so, (A.13) cannot hold.

Next, we show that $\dim(\mathcal{A}(\mathcal{D}_{t-1})) \geq 2$. Suppose that $\dim(\mathcal{A}(\mathcal{D}_{t-1})) = 1$. Let

$$\theta_t = \mathcal{A}(\mathcal{D}_t) \in \mathbb{R}$$

by the hypothesis of the proof. Since Θ is compact, $\{\theta_t\}$ has a convergent subsequence. After re-numbering the subsequence, let

$$\theta_t \to \theta^* \in \Theta.$$

Since $\mathcal{A}$ PAC guarantees $\mathcal{F}$, $\forall F \in \mathcal{F}$,

$$\theta_t = \mathcal{A}(\mathcal{D}_t : F) \to \theta^*$$

implies that

$$\varphi(\theta_t) \to \varphi(\theta^*) = (b^*(F), q^*(F)).$$

We can therefore consider the following set:

$$\{(a,b) \mid \exists \theta \in \Theta \text{ such that } a = \varphi_p(\theta), b = \varphi_q(\theta)\} \tag{A.14}$$

$\theta \in \mathbb{R}$ is single dimensional by the hypothesis of the proof. Lipschitz continuity of φ implies that the Hausdorff dimension of the set Θ cannot increase under the parameterization (see, for instance, Proposition 3.1.5 of Ambrosio and Tilli (2004)). Hence this set is one dimensional.

Since $\mathcal{A}$ PAC guarantees $\mathcal{F}$, (A.14) is exactly

$$\{(a,b) \mid \exists F \in \mathcal{F} \text{ such that } a = b^*(F), b = 1 - F(b^*(F))\} \tag{A.15}$$

To obtain the contradiction, it suffices to observe that (A.15) is in fact not one dimensional. This can be seen using linear demand curves (or the uniform distribution function F) which is a subset of $\mathcal{F}$. That is, if $q = b_0 + \beta_1 p$, then $q^* = \frac{\beta_0}{2}$ and $p^* = -\frac{\beta_0}{2\beta_1}$. Since for every fixed β_0, (a) the set of (q^*, p^*) generated by some β_1 is a one dimensional set, (b) these sets are also disjoint for distinct β_0, and (c) the set of valid β_0 is itself one dimensional, we therefore have that the set in (A.15) is two dimensional for this class of $\mathcal{F}$. This contradiction establishes that there cannot be a single-dimensional parameter which is used in the algorithm.

Appendix B. Proof of Theorem 4.1

The projection facility is only used to ensure the tightness of the set of the sample paths, and does not alter the asymptotic properties such as the stability and the large deviation properties of the algorithm (Dupuis and Kushner (1989)). Thanks to the projectin facility, we can assume tht $(\beta_{0,t}, \beta_{1,t})$ is contained in a compact convex set. For the remainder of the paper, we suppress the projection facility to simplify the exposition, when we examine the asymptotic properties of the algorithm.

B.1. **Preliminaries.** Fix $\tau > 0$ and consider an interval $[0, \tau)$ of real time. Fix small $a > 0$, and divide the interval into subintervals of size a, with a possible exception of the last subinterval. Define

$$T(a) = \left\lceil \frac{\tau}{a} \right\rceil$$

be the number of the subintervals (treating the last subintervals as the full size subinterval) in $[0, \tau)$, where a is the gain coefficient of the updating term in the recursive formula. Recall that $\epsilon_{2,t}$ is distributed uniformly over $[-\epsilon, \epsilon]$ where $\epsilon > 0$. We will choose τ, a, ϵ to meet the accuracy and the confidence requirement of the algorithm, which in turn determines $T(a)$.

For a fixed $F \in \mathcal{F}^\eta$, we are interested in

$$\beta_{T(a)} - \beta^*(F).$$

Recall that for $t \geq 1$, we can write the recursive formula as

$$\beta_t = \beta_{t-1} + a\varphi(\beta_{t-1}, p_t, \epsilon_t)$$

since the updating term is determined by the old estimate β_{t-1}, the price in period t and the realized quantity, where the last two variables are subject to two shocks $(\epsilon_{1,t}, \epsilon_{2,t})$. Let

$$\varphi(\beta_{t-1}, p_t, \epsilon_t) = \mathbb{E}_{t-1}\varphi(\beta_{t-1}, p_t, \epsilon_t) + \xi_t$$

where ξ_t is the martingale difference. Since $\beta_t \in \mathsf{B}$ which is comapct, ξ_t is uniformly bounded: $\exists \xi > 0$ such that

$$|\xi_t| \leq \xi.$$

Define

$$\overline{b}_{t-1}(\beta_{t-1}) = \mathbb{E}_{t-1}\varphi(\beta_{t-1}, p_t, \epsilon_t).$$

As shown in (B.19), the functional form of b_{t-1} is not affected by $t-1$ and is a Lipschitz continuous function of β_{t-1}. To simplify notation, we write $\overline{b}(\beta_{t-1})$ in place of $\overline{b}_{t-1}(\beta_{t-1})$, dropping the time subscript from $\overline{b}_{t-1}$. We can write the recursive formula as

$$\beta_t = \beta_{t-1} + a\left[\overline{b}(\beta_{t-1}) + \xi_t\right].$$

Given $\beta_0, \beta_1, \ldots, \beta_{T(a)}$, define a continous time process obtained by the linear interpolation: $\forall s \in [a(t-1), at)$,

$$\beta^a(s) = \frac{(s - a(t-1))\beta_t + (at - s)\beta_{t-1}}{a}.$$

Define

$$\beta(s) = \lim_{a \to 0} \beta^a(s)$$

pointwise. The existence of the limit point is guaranteed by the fact that β_t is contained in a compact set and $\overline{b}$ is a Lipschitz continuous function.

Define $\beta^*(F) = (\beta_0^*(F), \beta_1^*(F))$ as the intercept and the slope of a linear demand curve that generates the optimal price $b^*(F)$ and the expected quantity $1 - F(b^*(F))$, that solves

$$1 - F(b^*(F)) = \frac{\beta_0^*(F)}{2} \quad \text{and} \quad f(b^*(F)) = -\beta_1^*(F).$$

We can write

$$\begin{aligned}\beta_{T(a)} - \beta^*(F) &= \beta_0 + a\sum_{t=1}^{T(a)} \overline{b}(\beta_{t-1}) + a\sum_{t=1}^{T(a)} \xi_t - \beta^*(F) \\ &= \beta_0 + \int_0^\tau \overline{b}(\beta(s))ds - \beta^*(F) \qquad \text{(B.16)} \\ &\quad + a\sum_{t=1}^{T(a)} \overline{b}(\beta_{t-1}) - \int_0^\tau \overline{b}(\beta(s))ds \qquad \text{(B.17)} \\ &\quad + a\sum_{t=1}^{T(a)} \xi_t \qquad \text{(B.18)}\end{aligned}$$

We examine (B.16), (B.17) and (B.18) one by one.

B.2. **Convergence and Stability.** We can write

$$\beta(\tau) = \beta(0) + \int_0^\tau \overline{b}(\beta(s))ds$$

where $\beta(0) = \beta_0$, which is often written as

$$\dot{\beta} = \overline{b}(\beta).$$

To simplify notation, we write

$$b_t = -\frac{\beta_{0,t-1}}{2\beta_{1,t-1}} \qquad \forall t \geq 1$$

and $b(\tau)$ as the continuous process constructed from b_t via linear interpolation $\forall \tau \geq 0$. If the meaning is clear from the context, we drop τ to write b instead of $b(\tau)$. The same convention applies to all other variables such as $\beta_{0,t}$ and $\beta_{1,t}$.

We examination the properties of the ordinary differential equation (ODE):

$$\dot{\beta} = R^{-1}\mathbb{E}\begin{bmatrix} 1 \\ -b_t + \epsilon_{1,t} \end{bmatrix} \phi(\beta_{t-1}, \epsilon_{1,t})$$

where

$$R = \begin{bmatrix} 1 & b_t \\ b_t & b_t^2 + \sigma_1^2 \end{bmatrix}$$

and

$$\phi(\beta_{t-1}, \epsilon_t) = 1 - F\left(b_t + \epsilon_{1,t}\right) + \epsilon_{2,t} - \beta_{0,t-1} - \beta_{1,t-1}b_t - \beta_{1,t-1}\epsilon_{1,t}.$$

Since $\epsilon_{1,t}$ has a small support, and F is differentiable, it is more convenient to write

$$F\left(b_t + \epsilon_{1,t}\right) = F\left(b_t\right) + f\left(b_t\right)\epsilon_{1,t} + \mathsf{O}(\epsilon^2).$$

We are interested in the column vector

$$\mathbb{E}\begin{bmatrix} 1 \\ b_t + \epsilon_{1,t} \end{bmatrix} \phi(\beta_{t-1}, \epsilon_t).$$

The first component is

$$1 - F\left(b_t\right) - \beta_{0,t-1} - \beta_{1,t-1}b_t + \mathsf{O}(\epsilon^2).$$

The second component is

$$\left[-f\left(\frac{\beta_{0,t-1}}{2\beta_{1,t-1}}\right) - \beta_{1,t-1}\right]\sigma_1^2 + \mathsf{O}(\epsilon^3).$$

We can write

$$\dot{\beta} = R^{-1}\begin{bmatrix} 1 - F\left(b\right) - \beta_0 - \beta_1 b + \mathsf{O}(\epsilon^2) \\ b\left[1 - F\left(b\right) - \beta_0 - \beta_1 b\right] - \left[f\left(b\right) + \beta_1\right]\sigma_1^2 + \mathsf{O}(\epsilon^3) \end{bmatrix}. \qquad \text{(B.19)}$$

At the stationary point $b^*(F)$ where the right hand side of ODE vanishes,

$$\begin{aligned} 1 - F\left(b^*(F)\right) &= \frac{\beta_0}{2} - \mathsf{O}(\epsilon^2) \\ f\left(b^*(F)\right) &= -\beta_1 + \frac{\mathsf{O}(\epsilon^3)}{\mathsf{O}(\epsilon^2)}. \end{aligned}$$

Thus,

$$\frac{1 - F\left(b^*(F)\right)}{f\left(b^*(F)\right)} = b^*(F) + \mathsf{O}(\epsilon). \tag{B.20}$$

To simplify notation, let us ignore O term and treat $b^*(F)$ as the solution of

$$\frac{1 - F\left(b^*(F)\right)}{f\left(b^*(F)\right)} = b^*(F).$$

The uniqueness of the solution is implied by the increasing hazard rate property of F. Since $\forall F \in \mathcal{F}$, f is Lipschitz continous

$$|f(x) - f(x')| \leq \eta |x - x'|,$$

$|\mathsf{O}(x)| \leq \eta |x|$. By reducing the size of the support of $\epsilon_{1,t}$, we can achieve the desired level of accuracy uniformly.

Let us proceed the calculation after suppressing O terms. Note that

$$R^{-1} = \frac{1}{\sigma_1^2} \begin{bmatrix} b^2 + \sigma_1^2 & -b \\ -b & 1 \end{bmatrix}$$

We write ODE of $\beta = (\beta_0, \beta_1)$ without linear approximation error O as

$$\begin{bmatrix} \dot{\beta}_0 \\ \dot{\beta}_1 \end{bmatrix} = \begin{bmatrix} (1 - F(b) - \beta_0 - \beta_1 b) + b(f(b) + \beta_1) \\ -(f(b) + \beta_1) \end{bmatrix}.$$

By definition of b

$$\beta_0 + 2\beta_1 b = 0$$

at every moment of time. Thus,

$$\dot{\beta}_0 + 2\dot{\beta}_1 b + 2\beta_1 \dot{b} = 0.$$

After substituting $\dot{\beta}_0$ and $\dot{\beta}_1$, we have

$$\dot{b} = -\frac{f(b)}{2\beta_1} \left[\frac{1 - F(b)}{f(b)} - b \right] \equiv \mathsf{R}(b) \tag{B.21}$$

modulo linear approximation errors $\mathsf{O}(\epsilon)$.

Lemma B.1. *$\exists c > 0$ such that*

$$\mathsf{R}(b) \geq -c(b - b^*(F)) \qquad b \leq b^*(F)$$

and

$$\mathsf{R}(b) \leq -c(b - b^*(F)) \qquad b \geq b^*(F).$$

Proof. We constructed the projection facility so that $\beta_1 < 0$ and $\beta_0 > 0$ and moreover,

$$\sup_{F \in \mathcal{F}^\eta} \beta_1 < 0.$$

If $\inf_{F \in \mathcal{F}^\eta} f(b) > 0$, then the proof is trivial. Since we only assume that $F \in \mathcal{F}^\eta$, we need more work.

Consider an iso-(expected) profit curve in the space of (q, p)

$$\Pi = pq.$$

Its slope is

$$-\left.\frac{dq}{dp}\right|_{\Pi} = \frac{1 - F(p)}{p}.$$

If $b = b^*(F)$, then the slope of the iso-profit curve must be equal to the slope of the demand curve $f(p)$ at $p = b^*(F)$:

$$\frac{1 - F(b^*(F))}{b^*(F)} = f(b^*(F)).$$

Since $b^*(F) \in [\underline{v}, \overline{v}]$, the slope of an iso-profit curve at the optimal price must be uniformly bounded: $\exists \underline{M}, \overline{M} > 0$ such that

$$\underline{M} \leq f(b^*(F)) = \frac{1 - F(b^*(F))}{p^*} \leq \overline{M}.$$

Since $F \in \mathcal{F}^\eta$ is uniformly Lipschitz continous, for a sufficiently small $\epsilon > 0$,

$$f(b^*(F) - \epsilon) \geq \underline{M} - \eta\epsilon > \frac{\underline{M}}{2}$$

and similarly,

$$f(b^*(F) + \epsilon) \leq -\underline{M} + \eta\epsilon < -\frac{\underline{M}}{2}.$$

We prove that the right hand side of ODE

$$\dot{b} = -\frac{f(b)}{\beta_1}\left(\frac{1 - F(b)}{f(b)} - b\right)$$

is strictly bounded away from 0 over $b < b^*(F) - \epsilon$ and $b > b^*(F) + \epsilon$.

If $f(b) = 0$, the increasing hazard rate property implies that $f(b') = 0 \ \forall b' < b$, and in particular $f(\underline{v}) = 0$. Since $f(b^*(F)) > 0$, $b < b^*(F)$. By the construction of the algorithm along the boundary, $\dot{b} = -\frac{1}{2\beta_1} > 0$ uniformly, because $\sup_{F \in \mathcal{F}^\eta} \beta_1 < 0$.

Suppose $f(b) > 0$ and $b \leq b^*(F) - \epsilon$. We know that $\mathsf{R}(b)$ defined in (B.21) is strictly decreasing, because of the increasing hazard rate property. We also know that $\mathsf{R}(b^*(F) - \epsilon) \geq \frac{M}{2}$. Thus,

$$\mathsf{R}(b) \geq \mathsf{R}(b^*(F) - \epsilon) \geq \frac{\underline{M}}{2} > 0.$$

Similarly, if $b \geq b^*(F) + \epsilon$,

$$\mathsf{R}(b) \leq \mathsf{R}(b^*(F) + \epsilon) \leq -\frac{\underline{M}}{2} < 0.$$

The increasing hazard rate property implies that if $b > b^*(F)$,

$$\frac{1 - F(b)}{f(b)} - b = \frac{1 - F(b)}{f(b)} - b - \left(\frac{1 - F(b^*(F))}{f(b^*(F))} - b^*(F)\right) < -(b - b^*(F))$$

and if $b < b^*(F)$,

$$\frac{1 - F(b)}{f(b)} - b > -(b - b^*(F)).$$

Thus,

$$|\mathsf{R}(b) - \mathsf{R}(b^*(F))| \geq \frac{\underline{M}}{2}|b - b^*(F)|$$

over $b \in [b^*(F) - \epsilon, b^*(F) + \epsilon]$. Since $b \in [\underline{v}, \overline{v}]$, $\exists c > 0$ such that

$$\mathsf{R}(b) - \mathsf{R}(b^*(F)) = R(b) \geq -c(b - b^*(F)) \qquad b \leq b^*(F)$$

and

$$\mathsf{R}(b) - \mathsf{R}(b^*(F)) = R(b) \leq -c(b - b^*(F)) \qquad b \geq b^*(F).$$

□

For any initial value $b(0) \in [\underline{v}, \overline{v}]$,

$$\left|b(\tau) - b^*(F)\right| \leq e^{-c\tau}\left|b(0) - b^*(F)\right| \leq e^{-c\tau}(\overline{v} - \underline{v}).$$

Let $\tau(\mu)$ be the first time when

$$\left|b(\tau) - b^*(F)\right| \leq \mu.$$

$(1 - F(b^*(F)), b^*(F)) \in K \ \forall F \in \mathcal{F}$ and K is a compact subset in the interior of $\mathbb{R}^2_+$. Thus,

$$\overline{\tau}(\mu) = \sup_{(\beta_0(0), \beta_1(0)) \in \mathsf{B}} \tau(\mu) < \infty.$$

and

$$\overline{\tau}(\mu) \sim -\log \mu$$

as $\mu \to 0$. Let us choose $\tau = \overline{\tau}(\mu)$.

B.3. **Riemann Residual.** Let us consider (B.17)

$$\mathbf{R}(a, F) = a \sum_{t=1}^{T(a)} \overline{b}(\beta_{t-1}) - \int_0^{\tau} \overline{b}(\beta(s))ds$$

which is the Riemann residual. Since f is uniformly Lipschitz over $\mathcal{F}$, $\overline{b}(\beta)$ is uniformly Lipschitz: $\exists \eta' > 0$ such that

$$\left|\overline{b}(\beta) - \overline{b}(\beta')\right| \leq \eta' |\beta - \beta'| \qquad \forall F \in \mathcal{F}.$$

For each subinterval of size a, the difference between the discrete value and the integration is at most $\frac{\eta' a^2}{2}$. Thus,

$$\mathbf{R}(a, F) \leq \frac{\eta' a^2}{2} \frac{\overline{\tau}(\mu)}{a} = \frac{\eta' a \overline{\tau}(\mu)}{2}.$$

Note that the right hand side is independent of F. Thus, $\forall \mu > 0$, define

$$\overline{a} = \frac{2\mu}{\overline{\tau}(\mu)\eta'} \tag{B.22}$$

so that $\forall a \leq \overline{a}$, $\forall F \in \mathcal{F}$,

$$\mathbf{R}(a, F) \leq \mu. \tag{B.23}$$

Thus,

$$\overline{a} = \mathsf{O}\left(-\frac{\mu}{\log \mu}\right)$$

which implies

$$\frac{\tau(\mu)}{\overline{a}} = \mathsf{O}\left(\frac{(\log \mu)^2}{\mu}\right).$$

B.4. **Lower Bound of Confidence.** Next, we examine (B.18). Consider $\frac{\xi_t}{\xi}$, which is a martingale difference with

$$\left|\frac{\xi_t}{\xi}\right| \leq 1 \quad \text{and} \quad \mathbb{E}\left(\frac{\xi_t}{\xi}\right)^2 \leq 1.$$

Since $\frac{\xi_t}{\xi}$ satisfies the large deviation property, $\forall \mu' > 0$, $\exists \rho(\mu', F) > 0$ (called the rate function) such that

$$\mathbb{P}\left(\left|\frac{1}{T}\sum_{t=1}^{T} \frac{\xi_t}{\xi}\right| > \mu'\right) \leq e^{-\rho(\mu', F)T}.$$

We need a uniform rate function $\rho(\mu', F) > 0$ over $\mathcal{F}$. By Azuma-Hoeffding-Bennett inequality (Corollary 2.4.7 in Dembo and Zeitouni (1998)), we have $\forall \mu' \in (0, 1/2)$,

$$\mathbb{P}\left(\left|\frac{1}{T}\sum_{t=1}^{T} \frac{\xi_t}{\xi}\right| > \mu'\right) \leq e^{-2T\mathsf{H}\left(\mu' + \frac{1}{2} \,|\, \frac{1}{2}\right)}$$

where

$$\mathsf{H}(p \mid p_0) = p \log \frac{p}{p_0} + (1-p) \log \frac{1-p}{1-p_0}$$

for $p, p_0 \in (0, 1)$. Let

$$T(a, \mu) = \left\lceil \frac{\overline{\tau}(\mu)}{a} \right\rceil.$$

Then,

$$\mathbb{P}\left(\left|a\sum_{t=1}^{T(a,\mu)}\xi_t\right| < \overline{\tau}(\mu)\xi\mu'\right) \leq e^{-2T\mathsf{H}\left(\mu'+\frac{1}{2}\,|\,\frac{1}{2}\right)}.$$

Let

$$\mu' = \frac{\mu}{\overline{\tau}(\mu)\xi}.$$

After substitution, we have

$$\mathbb{P}\left(\left|\frac{1}{T(a,\mu)}\sum_{t=1}^{T(a,\mu)}\xi_t\right| > \mu\right) \leq e^{-2T(a,\mu)\mathsf{H}\left(\mu'+\frac{1}{2}\,|\,\frac{1}{2}\right)}.$$

Since $\overline{\tau}(\mu) = \mathsf{O}(-\log\mu)$,

$$\frac{\mu}{\overline{\tau}(\mu)\xi} = \mathsf{O}(\mu)$$

and therefore, $T(a,\mu)$ increases at the polynomial speed as $\mu \to 0$. Let

$$\rho = 2\mathsf{H}\left(\mu'+\frac{1}{2}\,|\,\frac{1}{2}\right) > 0.$$

B.5. **Combine the Pieces.** Fix $\mu > 0$. Recall that we assume that $\epsilon_{1,t}$ is distributed over $[-\epsilon,\epsilon]$. We first choose $\epsilon > 0$ so that the stationary solution (B.20) of ODE is within μ neighborhood of $\beta^*(F)$ $\forall F \in \mathcal{F}$.

Since $\forall f \in \mathcal{F}$

$$|f(p) - f(p')| \leq \eta|p - p'|.$$

The Taylor residual is bounded by $\eta\epsilon$ for small $\epsilon > 0$:

$$|\mathsf{O}(\epsilon^3)| \leq |\mathsf{O}(\epsilon^2)| \leq \eta\epsilon^2 \qquad \forall t \geq 1.$$

Note that the last term is independent of $F \in \mathcal{F}$. Choose $\epsilon > 0$ sufficiently small so that

$$\left|\beta^s - \beta^*(F)\right| < \mu \qquad \forall F \in \mathcal{F}$$

where β^s solves (B.20). We chose $\overline{\tau}(\mu)$ so that

$$|\beta(\overline{\tau}(\mu)) - \beta^s| < \mu.$$

Given $\mu > 0$, we chose $a_1 > 0$ in (B.22) so that the Riemann residual $\mathsf{R}(a,F)$ satisfies (B.23).

By the construction, $\forall a \in (0,a_1)$,

$$\mathbb{P}\left(\left|\frac{1}{T(a,\mu)}\sum_{t=1}^{T(a,\mu)}\xi_t\right| > \mu\right) < e^{-T(a,\mu)\rho}.$$

Combining these results, we have $\forall a \in (0,a_1)$,

$$\mathbb{P}\left(\left|\beta_{T(a,\mu)} - \beta^*(F)\right| \geq 4\mu\right) \leq e^{-T(a,\mu)\rho}$$

where $T(a,\mu)$ increases linearly with respect to $1/a$ and at the polynomial speed as $\mu \to 0$.

References

Ambrosio, L., and P. Tilli (2004): *Topics on Analysis in Metric Spaces*, vol. 25 of *Oxford Lecture Series in Mathematics and Its Applications*. Oxford University Press.

Besbes, O., and A. Zeevi (2015): "On the (Surprising) Sufficiency of Linear Models for Dynamic Pricing with Demand Learning," *Management Science*, 61(4), 723–739.

Carroll, G. (2015): "Robustness and Linear COntracts," *American Economic Review*, 105(2), 536–563.

——— (2017): "Robustness and Separation in Multidimensional Screening," *Econometrica*, 85(2), 453–488.

Cho, I.-K., and K. Kasa (2015): "Learning and Model Validation," *Review of Economic Studies*, 82, 45–82.

——— (2017): "Gresham's Law of Model Averaging," *American Economic Review*, 107(11), 1–29.

Cole, R., and T. Roughgarden (2014): "The Sample Complexity of Revenue Maximization," in *Proceedings of the forty-sixth annual acm symposium on theory of computing*, STOC '14, pp. 243–252, New York, New York. ACM.

Dembo, A., and O. Zeitouni (1998): *Large Deviations Techniques and Applications*. Springer-Verlag, New York, 2nd edn.

Du, S. (2018): "Robust Mechanisms Under Common Valuation," *Econometrica*, 86(5), 1569–1588.

Dupuis, P., and H. J. Kushner (1989): "Stochastic Approximation and Large Deviations: Upper Bounds and w.p.1 Convergence," *SIAM Journal of Control and Optimization*, 27, 1108–1135.

Esponda, I., and D. Pouzo (2014): "An Equilibrium Framework for Players with Misspecified Models," University of Washington and University of California, Berkeley.

Esponda, I., D. Pouzo, and Y. Yamamoto (2021): "Asymptotic Behavior of Bayesian Learners with Misspecified Models," *Journal of Economic Theory*.

Evans, G. W., and S. Honkapohja (2001): *Learning and Expectations in Macroeconomics*. Princeton University Press.

Frick, M., R. Iijima, and Y. Ishii (2021): "Belief Convergence under Misspecified Learning: A Martingale Approach," Discussion paper, Yale University and Pennsylvania State University.

Fudenberg, D., G. Lanzani, and P. Strack (2021): "Limit Points of Endogenous Misspecified Learning," *Econometrica*.

Fudenberg, D., and D. Levine (1998): *Learning in Games*. M.I.T. Press.

Goncalves, D., and B. Furtado (2020): "Statistical Mechanism Design: Robust Pricing and Reliable Projections," Discussion paper, Columbia University.

Gonczarowski, Y. A., and S. M. Weinberg (2021): "The Sample Complexity of Up-to-ϵ Multi-dimensional Revenue Maximization The Sample Complexity of Up-to-ϵ Multi-dimensional Revenue Maximization," *Journal of the ACM*, 68(3).

Hansen, L. P., and T. J. Sargent (2007): *Robustness*. Princeton University Press.

Hansen, L. P., and T. J. Sargent (2020): "Structured ambiguity and model misspecification," *Journal of Economic Theory*, p. 105165.

Heidhues, P., B. Köszegi, and P. Strack (2018): "Unrealistic Expectations and Misguided Learning," *Econometrica*, 86(4), 1159–1214.

Huang, Z., Y. Mansour, and T. Roughgarden (2018): "Making the Most of Your Samples," *SIAM Journal on Computing*, 47(3), 651–674.

Kushner, H. J., and G. G. Yin (1997): *Stochastic Approximation Algorithms and Applications*. Springer-Verlag.

Libgober, J., and X. Mu (2021): "Informational Robustness in Intertemporal Pricing," *Review of Economic Studies*, 88(3), 1224–1252.

Myerson, R. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

Nisan, N., E. Tardos, and V. V. Vazirani (eds.) (2007): *Algorithmic Game Theory*. Cambridge University Press.

Osborne, M. J., and A. Rubinstein (1998): "Games with Procedurally Rational Players," *American Economic Review*, 88(4), 834–847.

Roughgarden, T., and O. Schrijvers (2016): "Ironing in the dark," in *Proceedings of the 17th ACM Conference on Economics and Computation (EC'16)*, pp. 1–18.

Rubinstein, A. (1986): "Finite Automata Play Repeated Prisoners Dilemma," *Journal of Economic Theory*, 39(1), 83–96.

Sargent, T. J. (1999): *The Conquest of American Inflation*. Princeton University Press.

Shalev-Shwartz, S., and S. Ben-David (2014): *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press.

Department of Economics, Emory University, Atlanta, GA 30322 USA
Email address: icho30@emory.edu
URL: https://sites.google.com/site/inkoocho

Department of Economics, University of Southern California, Los Angeles, CA 90089 USA
Email address: libgober@usc.edu
URL: http://www.jonlib.com/