



Marketing Science

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Frontiers: Algorithmic Collusion: Supra-competitive Prices via Independent Algorithms

Karsten T. Hansen, Kanishka Misra, Mallesh M. Pai

To cite this article:

Karsten T. Hansen, Kanishka Misra, Mallesh M. Pai (2021) Frontiers: Algorithmic Collusion: Supra-competitive Prices via Independent Algorithms. Marketing Science 40(1):1-12. <https://doi.org/10.1287/mksc.2020.1276>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2021, INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Frontiers: Algorithmic Collusion: Supra-competitive Prices via Independent Algorithms

Karsten T. Hansen,^a Kanishka Misra,^a Mallesh M. Pai^b

^a Rady School of Management, University of California, San Diego, La Jolla, California 92093; ^b Department of Economics, Rice University, Houston, Texas 77005

Contact: k4hansen@ucsd.edu (KTH); kamisra@ucsd.edu,  <https://orcid.org/0000-0002-0106-1230> (KM); mallesh.pai@rice.edu (MMP)

Received: January 22, 2020

Revised: August 5, 2020

Accepted: August 12, 2020

Published Online in Articles in Advance:
January 8, 2021

<https://doi.org/10.1287/mksc.2020.1276>

Copyright: © 2021 INFORMS

Abstract. Motivated by their increasing prevalence, we study outcomes when competing sellers use machine learning algorithms to run real-time dynamic price experiments. These algorithms are often misspecified, ignoring the effect of factors outside their control, for example, competitors' prices. We show that the long-run prices depend on the informational value (or signal-to-noise ratio) of price experiments: if low, the long-run prices are consistent with the static Nash equilibrium of the corresponding full information setting. However, if high, the long-run prices are supra-competitive—the full information joint monopoly outcome is possible. We show that this occurs via a novel channel: competitors' algorithms' prices end up running correlated experiments. Therefore, sellers' misspecified models overestimate the own price sensitivity, resulting in higher prices. We discuss the implications on competition policy.

History: K. Sudhir served as the editor-in-chief and Harikesh Nair served as associate editor for this article. This paper was accepted through the *Marketing Science* Frontiers review process.

Funding: M. M. Pai gratefully acknowledges financial support from the National Science Foundation [Grant CCF-1763349].

Supplemental Material: The online appendix is available at <https://doi.org/10.1287/mksc.2020.1276>.

Keywords: algorithmic pricing • collusion • behavioral game theory

1. Introduction

A defining feature of many of the most profitable companies in the 21st century is *scale*. For example, the song catalog for the music service Spotify contains more than 50 million songs.¹ The online retailer Amazon.com sells close to 120 million unique products.² The movie streaming service Netflix has more than 130 million subscribing customers.³ This proliferation of products and customers would be impossible to manage if humans did not have help in the form of information technology. Large-scale databases and algorithms for decision making are now an essential ingredient of high-tech business management.

The benefits of this revolution are clear. For example, one consequence of this technology has been the creation of the most valuable companies in human history. At the same time, consumers benefit from recommendation systems engineered for customized product assortments, ease of purchase, peer reviews, online price comparisons, etc. An obvious question to ask is if there are any downsides to algorithmic business management.

In this paper, we analyze a novel, unforeseen consequence of algorithmic management that has received some attention in both the popular press and research: the possibility that companies may end up colluding *through algorithms*. The context we study is

pricing: an area in which online retailers increasingly use machine learning algorithms (The White House 2015). These algorithms set real-time prices for an array of products for which the retailer has incomplete demand information. In this setting, each algorithm is an automated field experiment that learns about potential profit (“exploration”) and sets the product’s profit-maximizing price (“exploitation”). These algorithms are called multiarmed bandits (MABs; each “arm” is a price). Given the complexity and scale of online retail, field experiments in this context are typically analyzed assuming the focal firm is a monopolist (or, alternatively, best-responding to the fixed priced of competitors). Chen et al. (2016) study the best-selling product on Amazon.com and estimate only 2.4% of sellers run pricing algorithms that consider competitive prices. Recent academic empirical pricing studies, including Cheung et al. (2017), Dubé and Misra (2017), and Misra et al. (2019), assume no competitive response to pricing decisions.⁴ Several theoretical investigations into the use of dynamic pricing algorithms also assume the firm is a monopoly, for example, Broder and Rusmevichientong (2012), Handel and Misra (2015), and Ban and Keskin (2021).

This assumption greatly reduces complexity: it simplifies the problem to that of a single choice variable (own price) and estimating a profit curve, which is

only a function of the firm's own price. A consequence of this simplicity is that theoretically optimal algorithms include index algorithms (Gittins 1989, Weber 1992, Auer et al. 2002). These algorithms calculate an index for every arm based on the past performance of the arm (mean, variance, and number of times played). In each round, the algorithm simply selects the arm with the highest current index.

In this paper, we analyze the outcome if competing firms all use this independent multiarmed bandit indexing algorithm as outlined—and act as if they are in a monopolistic market—but, in reality, price into an oligopolistic market. In keeping with the bandit analogy, we call pricing in each period an “experiment.” In particular, we study what long-run prices would result in such a setting. Of course, because a seller's demand (and, therefore, profit) at a particular price depends on competitors' prices, each seller's implicit statistical model is misspecified.

We show that the long-run prices that result depend on the informational value (signal-to-noise ratio or SNR) of the underlying pricing experiments. In markets in which price experiments have low information value, the resulting long-run prices are statistically indistinguishable from Nash equilibrium prices, and the misspecified models achieve nearly the first best profits. However, in markets in which price experiments have high information value, market prices are supra-competitive. We show that more informative pricing experiments result in correlated price experiments across firms. Competitive prices, therefore, become correlated unobservables in each firms' pricing algorithm. By not accounting for the existing competitive pricing, each firm's price sensitivity has an upward bias, resulting in supra-competitive prices.

A growing body of literature has raised concerns that algorithms might induce collusive pricing behavior; see, for example, the white paper from the Organisation for Economic Co-operation and Development (2017) listing policymakers' concerns and the review article of Harrington (2018) on how competition law should adapt. Recent influential papers, such as Calvano et al. (2019, 2020), show that algorithmic collusion can occur in settings in which firms observe each other's prices. We contribute to this literature by identifying a novel channel/mechanism by which such supra-competitive pricing may occur. By contrast, in our setting, each algorithm/firm does not observe competitors' prices, a requirement in the channels identified in the extant literature. We show that collusion can materialize even in this case.

In the operations literature, our result is most closely related to Cooper et al. (2015), who also show non-Nash outcomes as the limit prices if duopoly sellers set prices from a misspecified monopoly model. The main difference between the papers is that our

algorithm considers optimal experimentation, balancing the benefits from learning and earning (reinforcement learning), and Cooper et al. (2015) do not explicitly model experimentation (assumed exogenous) and instead consider an estimation–optimization (certainty equivalence) algorithm by which prices are set as the static optimal prices from a linear ordinary least squares demand model.⁵ In terms of implications for policy, Cooper et al. (2015) suggest that the possibility of collusive prices depends on which parameters of a linear demand model are commonly known by both firms (this cannot be empirically tested). By contrast, our results suggest that it hinges on the underlying signal-to-noise ratio in the demand function, that is, the stochasticity of underlying demand (which *is* empirically testable). We provide a more detailed discussion of the connection to this paper and other prior research (on algorithmic collusion, dynamic pricing, learning in games, behavioral game theory, and algorithmic bias) in the online appendix.

We believe the identification of this novel channel is useful not only for theory, but it also raises fresh practical concerns for managers and policymakers. In our setting, firms independently choose algorithms that use misspecified models of the underlying demand system (firms' algorithms assume they are effectively monopolies although the true market is oligopolistic). Each firm uses these algorithms to structure “relevant learning” (Aghion et al. 1991) or learning about profit at the most profit-relevant prices. Our results show that, when multiple firms employ identical algorithms, then learning can be focused away from competitive prices, resulting in supra-competitive prices. Jointly, firms may have an incentive to be “willfully misspecified.” The question then, is whether all firms choosing such misspecified models constitutes a concern for competition policy.

The remainder of this paper is organized as follows: Section 2 describes the general setup and assumptions. Section 3 outlines our main results via a set of simulations. Section 4 contains our theoretical results (in a stylized setting) and lays out the mechanism by which supra-competitive prices result. In Section 5, we reanalyze Amazon.com pricing data from Chen et al. (2016) and argue that the observed correlations in prices are consistent with our simulations. Section 6 concludes and discusses avenues for future research.

2. Competitive Bandits Setup

In this section, we explicitly state the key demand and supply assumptions in our analysis. These assumptions define both the objective functions and the informational assumptions for the agents. We consider two symmetric, single-product firms with a constant marginal cost (set to zero). Firms compete by setting prices in each period.

We assume that each firm's demand is static and stable, a common assumption in empirical and theoretical work and field experiments; see, for example, Besbes and Zeevi (2009), Dubé and Misra (2017), Misra et al. (2019), and Calvano et al. (2019). Static demand rules out a large number of potential consumer dynamics, such as preference learning, uncertainty, strategic consumers, and stockpiling. These assumptions are important in our setting as we focus on the firms' learning about the demand (equivalently, profit) curve by price experimentation. If demand is not static, then the pricing field experiment would be biased by assumption.

We assume that demand is initially unknown to the firms. We assume there is a discrete set of potential prices $P = \{p_1, p_2, \dots, p_K\}$ that each firm can set. The formal results are for $K = 2$, but we show that similar results are obtained for arbitrary values of K in simulations. A period of the algorithm can be considered as a fixed interval of time (e.g., 15 minutes), as is common in the operations research literature (Besbes and Zeevi 2009), or a fixed number of consumers (e.g., 100 consumers), as is common in the marketing literature (Misra et al. 2019). Upon setting a price (running an experiment), the firm observes profit (with sampling error) in that period but does not observe the rival firm's price and quantity sold. An equivalent assumption is that the firm observes competitors' price/quantity but ignores these in its own statistical model.

Notation is as follows: In period t , each firm j sets a price $p_{j,t} \in P$. The firm observes a resulting profit $\pi_{j,t} \in \mathbb{R}_+$. The additional complication is that, in our setting, firms do not observe (or ignore the effect of) competitors' prices. That is to say that each firm's statistical model is $\pi_{j,t} = \pi^*(p_{j,t}) + \varepsilon_{j,t}$, where $\pi^*(\cdot)$ is the true stable expected profit and $\varepsilon_{j,t}$ is a small sample error or noise from the experiment at time t . However, in truth, $\pi_{j,t} = \pi^*(p_{j,t}, p_{-j,t}) + \varepsilon_{j,t}$. This means that firms' models are misspecified: they attribute the effect of competitors' prices ($p_{-j,t}$) on their own profits to be arising from noise from nature.

Viewing each possible price as an arm in an MAB problem, we suppose that each firm runs an MAB algorithm to balance learning for future and current profits. The objective function of the algorithm is to minimize (undiscounted) dynamic *statistical regret*, defined as the difference between average profits achieved with the algorithm and the ex post optimal profits. The literature provides theoretical analysis and mathematical guarantees of these algorithms.⁶

The bandit algorithm we use is the upper confidence bound (UCB) algorithm, originally from Auer (2002). This provides an asymptotically optimal non-parametric MAB solution (Agrawal 1995, Auer et al. 2002): that is to say, (1) it is guaranteed to find the

optimal arm without any further assumptions in any MAB problem, and (2) no other algorithm achieves lower regret. In our context, this means that a monopoly seller using a UCB algorithm for pricing is guaranteed to find the monopoly optimal price without any parametric assumptions on the relationship between prices and profits. The idea of this kind of algorithm is to maintain an index for each arm. The basic algorithm, for each arm p_k , at each period t tracks the empirical average of the profit from that arm for $\bar{\pi}_{k,t}$ and the number of times that arm has been pulled $n_{k,t}$. The index of arm k at period t , $I_{k,t}$, is defined as $\bar{\pi}_{k,t} + \sqrt{\frac{2 \ln t}{n_{k,t}}}$ (the index of an arm that has never been pulled is defined as ∞). In every period, the algorithm experiments with the price with the current highest index, randomizing if there is a tie.

Viewing the profit draws from a given arm as independent and identically distributed (i.i.d.) draws from a distribution of unknown mean, this index tracks the upper bound of a confidence interval of this unknown mean. A simple concentration argument tells us that, in period t , the true expected mean reward of any arm k is lower than the current index $I_{k,t}$ with probability $(1 - \frac{1}{t})$.

A price is, thus, experimented if either the empirical average of past profits from that price is high (exploitation) or $n_{k,t}$ is low relative to t , that is, that price has not been used sufficiently often to be confident about its reward (exploration).

3. Simulation Results

In this section, we describe the details of our simulations to test implications of independent competing firms running the pricing algorithm.

3.1. Simulation Settings

For our main results, we consider the algorithm called UCB-tuned.⁷ For each price (p_k) in period t , we calculate the index, defined as follows, and select the price with the highest index:

$$V_{k,t} = \bar{\pi}_{k,t} - \bar{\pi}_{k,t}^2 + \sqrt{\frac{2 \log t}{n_{k,t}}},$$

$$\text{UCB-tuned}_{k,t} = \bar{\pi}_{k,t} + \sqrt{\frac{\log t}{n_{k,t}} \min\left(\frac{1}{4}, V_{k,t}\right)},$$

where $n_{k,t}$ is the number of periods in which the firm has charged price p_k up to period t and $\bar{\pi}_{k,t}$ is the empirical mean of profits in those periods. In this algorithm, $V_{k,t}$ can be interpreted as the empirical variance plus an exploration bonus that depends on (decreases with) the number of times price p_k has been experimented with. Further, we account for the UCB improvement (Auer and Ortner 2010), and we allow

for arm elimination. We eliminate an arm k if the upper confidence bound of the arm is lower than the mean minus exploration bonus (i.e., the “lower confidence bound”) of another arm.

We consider a parametric data-generating process and allow firms to run independent UCB algorithms. To be explicit, the data-generating process is known to us (the researchers) but not to the firms: the firms run UCB, which is purely nonparametric. The UCB algorithm considers (undiscounted) finite time regret; this requires a prespecification of the number of periods in which the algorithm will be run.⁸ In all our simulations, we consider the algorithms running for two million periods and analyze the outcomes of the last 1,000 periods as the “long run.”⁹

In the main simulation, we assume a linear demand model:

$$d_j^*(p_j, p_{-j}) = \alpha - \beta p_j + \gamma p_{-j}. \quad (1)$$

Under this assumption, the competitive and joint monopoly (i.e., collusive) prices can be analytically computed. They are

$$\begin{aligned} \text{Competitive: } p^D &= \frac{\alpha}{2\beta - \gamma} \\ \text{Collusive: } p^M &= \frac{\alpha}{2(\beta - \gamma)}. \end{aligned}$$

In any experiment (t), a firm observes a noisy estimate of the true profit, that is,

$$\pi_t(p_j, p_{-j}) = p_j d_j^*(p_j, p_{-j}) + \varepsilon_{j,t},$$

but their models are misspecified, so they think they are observing

$$\pi_t(p_j) = p_j d_j^*(p_j) + \varepsilon_{j,t}.$$

In our simulations, we set the values of $\alpha = 0.48$, $\beta = 0.9$, and $\gamma = 0.6$. With these values, the duopoly and monopoly prices are $p^D = 0.4$ and $p^M = 0.8$. The firms decide between 91 discrete prices between \$0.10 and \$1.00, that is, $P = \{0.10, 0.11, \dots, 0.99, 1.00\}$.

As we previewed earlier, our results suggest that the long-run distribution of prices depends on the informativeness of the price experiments, that is, the magnitude of the noise. To that end, we vary the distributional assumptions about the error term, $\varepsilon_{j,t}$. In all simulations, we assume that $\varepsilon_{j,t} \sim U[-\frac{1}{\delta}, \frac{1}{\delta}]$, where we vary the value of δ across simulations. If δ is small (large) then the experiment is uninformative (informative), and δ signifies the SNR in an experiment (by construction, the highest signal is 1.0). In our simulation experiments, we vary SNR between $\frac{1}{10}$ (large noise) and 10 (large signal).

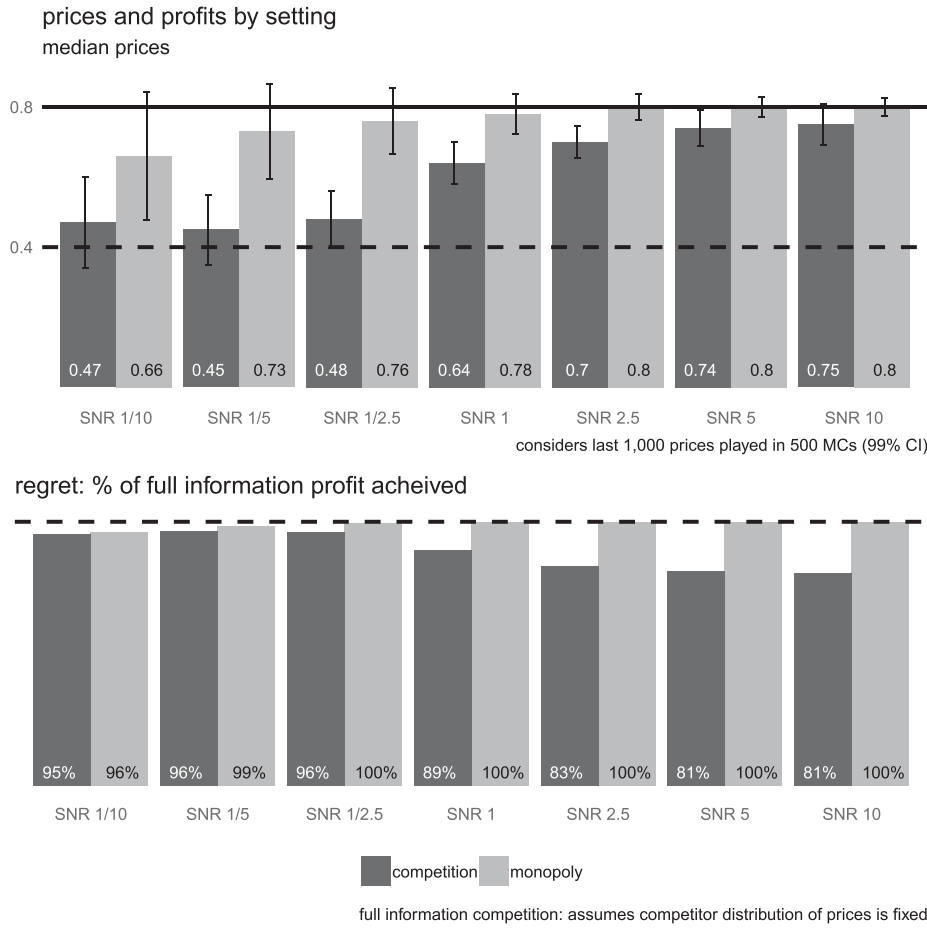
3.2. Main Results

Our main results are shown in Figure 1. The dark gray bars represent a setting with two competing single-product firms, and the light gray bars represent a setting with a monopolist jointly maximizing profits for both products. The top panel considers the long-run prices from our simulations. For all levels of SNR, the long-run price when the firm is a monopoly is not statistically distinguishable from the true monopoly price of 0.8. With a small SNR, the long-run prices in a duopoly are not statistically different from the Nash equilibrium price of 0.4. However, as the SNR increases, we find that the estimated duopoly price increases to supra-competitive levels. When the SNR levels are large, the estimated duopoly prices are statistically different from the Nash equilibrium and, indeed, statistically indistinguishable from the full-information joint monopoly (collusive) outcome. This suggests that, in markets in which pricing experiments are very (less) informative, independent algorithms result in long-run prices that are supra-competitive (competitive). In Section A.2, we show that this result is robust to a different demand system (logit demand) and a different index algorithm (Gittins index). Our main findings continue to hold in both these simulations: in markets with low (high) SNR, independent bandits result in competitive (supra-competitive) prices.

To understand the mechanics that generate supra-competitive duopoly prices, we consider the joint distribution of long-run prices. Figure 2 displays the distribution of long-run duopoly prices by SNR. Our point of emphasis here is the shape of the distribution and how this changes with SNR. For small SNR values, the shape of this distribution implies independent prices. However, as the SNR increases, the distribution indicates correlated prices across the two firms.¹⁰ In Section 4, we explain both the origins of this correlation and how it impacts long-run prices.

This suggests that alternative pricing algorithms that force no (force large) correlation in prices should result in competitive (supra-competitive) prices across all SNR settings. The long-run prices for such alternative algorithms are shown in Figure 3. The top panel of this figure corresponds to the market outcomes when both firms induce random price experimentation with the heuristic-based ε -greedy algorithm. In this algorithm, at period t , with probability ε (we set ε to be 0.01), the firm sets a randomly selected price; otherwise, the firm sets the price with the highest mean profits from past experiments. Consistent with this intuition, we find the long-run prices are indistinguishable from competitive Nash equilibrium prices for all levels of SNR. The bottom panel of this figure corresponds to the market outcomes in which we

Figure 1. Estimated Median Prices and Percentage of Optimal Profits by Simulation Setting



Notes. The dark gray bars represent a setting in which two firms are running simultaneous algorithms, and the light gray bars represent a setting in which a monopolist is jointly pricing the products. For each simulation, we consider the last 1,000 periods out of two million periods. In the top chart, the dashed lines reflect the competitive equilibrium prices; the solid lines reflect monopoly prices. In the bottom chart, the dashed line reflects 100% profit achieved. (Algorithm used: UCB tuned.)

force a correlation of one between the firms' prices. Here, we consider a setting in which one firm uses the UCB algorithm for pricing (as in our main results) and the other firm price-matches in real time. The competitive outcomes are indistinguishable from the monopoly outcomes for all levels of SNR.

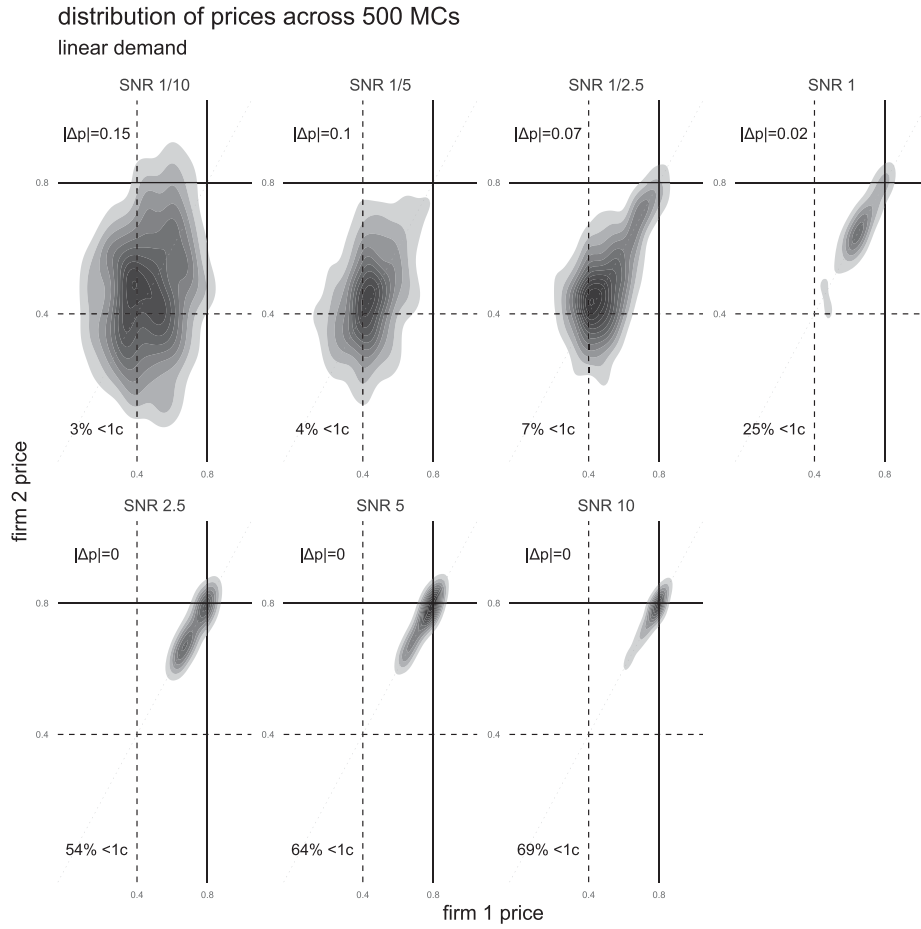
4. Theoretical Foundations

In this section, we explain the theoretical basis for the preceding results. In the online appendix, Section B, we explain the mechanism by a simple intuition for the implication of correlated prices. This is incomplete as the correlation of prices is endogenously generated by the algorithm. We provide correct theoretical foundations in a setting in which each firm chooses between two prices, that is, $K = 2$.

We prove our result in a model in which each firm only chooses between two prices, $p_H > p_L$. A firm's

observed profit when it charges price p_x and its competitor charges p_y , for $x, y \in \{H, L\}$, is $\pi_{xy} + \varepsilon$. Here, ε is a mean-zero shock, independently drawn across periods. The pricing structure, therefore, implies $\pi_{LH} > \pi_{HH}$ and $\pi_{LL} > \pi_{HL}$; that is, charging the low price is a dominant strategy for each firm. However, both firms charging the high price results in higher joint profits, that is, $\pi_{HH} > \pi_{LL}$. Putting this together, we have that the true game for both firms is essentially a prisoner's dilemma.

Both firms do not know the true π ; indeed, their statistical models are misspecified in that they assume the profit when they charge price p_x is $\pi_x + \varepsilon$. They both run UCB algorithms, so from each agent j 's perspective, in any period t , the relevant part of past play can be summarized by four numbers: for each price p_x , the empirical average of profits in past periods in which p_x was charged, $\bar{\pi}_{x,t}$, and the number of

Figure 2. 2-D Density Plot of the Distribution of Prices for the Two Firms Across 500 Monte Carlo (MC) Simulations per Scenario

Notes. Each chart represents a market setting described by SNR; a small SNR means large noise, and a large SNR means small noise. For each simulation, we consider the median price charged in the last 1,000 rounds out of two million rounds. The dashed lines reflect the competitive equilibrium prices; the solid lines reflect monopoly prices. The light gray dotted line presents the 45° line. The number on the top left of each chart shows the median difference between the two firms' prices, and a number on the bottom left represents the percentage of simulations with the difference in price less than one cent. (Algorithm used: UCB tuned.)

times it was charged, $n_{x,t}$. Note that $n_{H,t} + n_{L,t} = (t - 1)$ by definition. The UCB algorithm calculates an index for each price $x \in \{H, L\}$ as

$$I_{x,t} = \bar{\pi}_{x,t} + \sqrt{\frac{2 \ln t}{n_{x,t}}},$$

and always pulls the arm with the highest index. By definition, the index of an arm that has never been pulled is ∞ ; that is, every arm is pulled at least once before any arm is pulled twice. It is clear from the proofs that follow that similar results are obtained for other “index” algorithms in which indices are deterministic functions of $\bar{\pi}$ and n .

We study this system analytically for the extreme case in which the noise in the demand system vanishes; that is, demand in each period is a deterministic function of prices. Our results for each of these cases mirror correlated past-price intuition in Section 3.

Informally, when the noise in the demand vanishes, we show that price paths end up correlated (even though the algorithms are independent), and prices converge to the (p_H, p_H) collusive outcome. Formally,

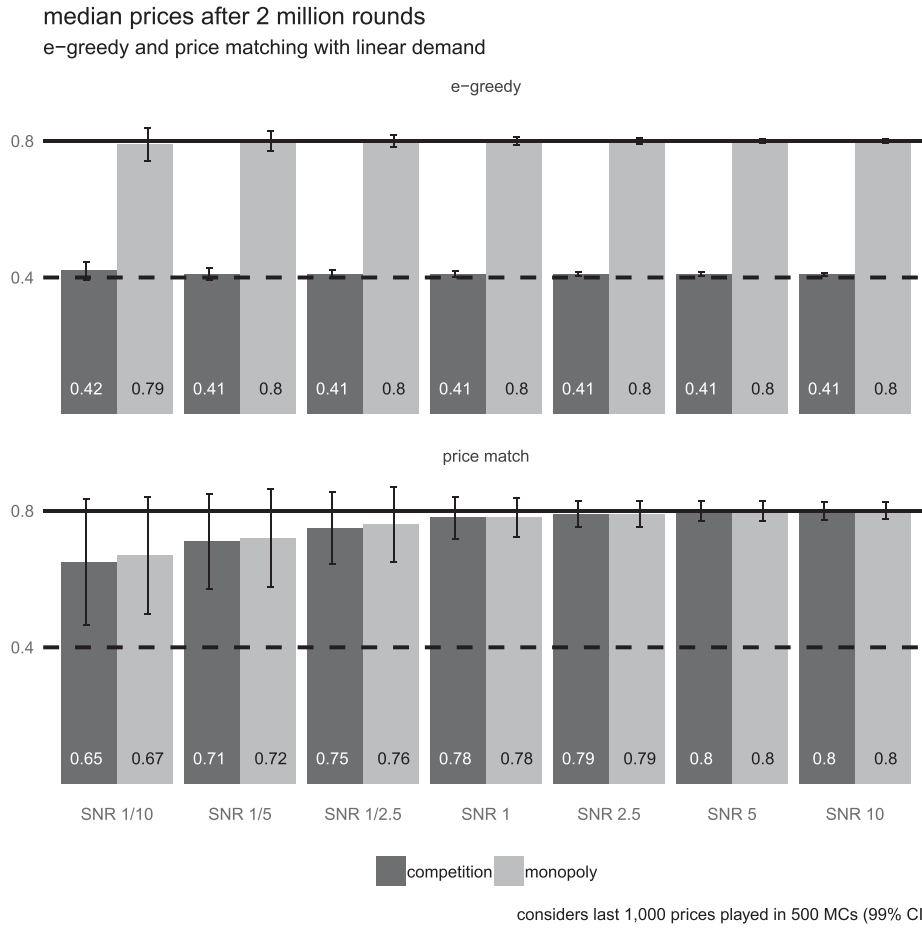
Theorem 1. *Suppose the true demand function is deterministic, and both firms use independent UCB algorithms. Then,*

1. *The prices are always exactly correlated from the third period onward (i.e., in any period $t \geq 3$, both firms charge either p_H or p_L).*
2. *The fraction of times in the first t periods that either seller charges p_L converges to zero as t grows large.*

The proof is provided in Section A.1.

Our result is shown for the specific case of two possible prices, deterministic demand, and when both sellers use the UCB algorithm. These features make the system feasible to study analytically. However, the mechanics of this proof and the intuition

Figure 3. Estimated Median Prices by Algorithm and Level of Competition



Notes. The error bars represent the 99% confidence intervals for the estimated median across MC simulations. The dark gray bars represent a setting in which two firms are running simultaneous algorithms, and the light gray bars represent a setting in which a monopolist is jointly pricing the products. The top panel corresponds to firms using the ϵ -greedy algorithm ($\epsilon = 0.01$), and the bottom panel corresponds to one firm using UCB and the other price matching. For each simulation, we consider the median price charged in the last 1,000 period out of two million periods. The dashed lines reflect the competitive equilibrium prices; the solid lines reflect monopoly prices. (Algorithms used: ϵ greedy or price matching with UCB tuned.)

provided illustrate how similar phenomena should be obtained more generally (and, indeed, why such results are obtained in the simulations we conducted previously). Independent algorithms can end up having correlated price paths. Because the algorithms are misspecified, this results in an omitted variable bias (the omitted variable being the competitor's price) and an overestimate of own price sensitivity. The absence of sufficiently large demand shocks to force independent experimentation then results in this being self-reinforcing. Both sellers' algorithms then settle on high prices even though this is neither the equilibrium of the underlying game nor, indeed, even the best response given the competitor's strategy. When demand shocks are small and the sellers' algorithms are close to deterministic, such mislearning may occur, and the competing firms may settle on charging collusive prices. This could occur forever (as in Theorem 1),

or with small amounts of noise, it may occur for arbitrarily long periods or with very high probability.¹¹

So when does such coordination fail? Demand has to be stochastic enough to prevent both firms from settling in to the correlated price paths displayed. It should be clear that, with highly stochastic demand, the early stages of experimentation under UCB are close to independent/stochastic for many periods. Because charging p_L is a dominant strategy in the underlying pricing game, both agents learn this, their associated index of p_L is higher, and their historical average payoff when playing p_L converges to π_{LL} by standard concentration inequalities. Subsequently, the stochastic differences in their estimates of the payoff under p_H ensure that both players are unlikely to switch to playing p_H at the same time. As a result, the historical average when playing p_H fall to $\pi_{HL} < \pi_{LL}$, and therefore, the path of play features both players

playing the low price almost always in the limit. How much noise/stochasticity is needed in demand to ensure this happens with high probability is a subject left for future research.

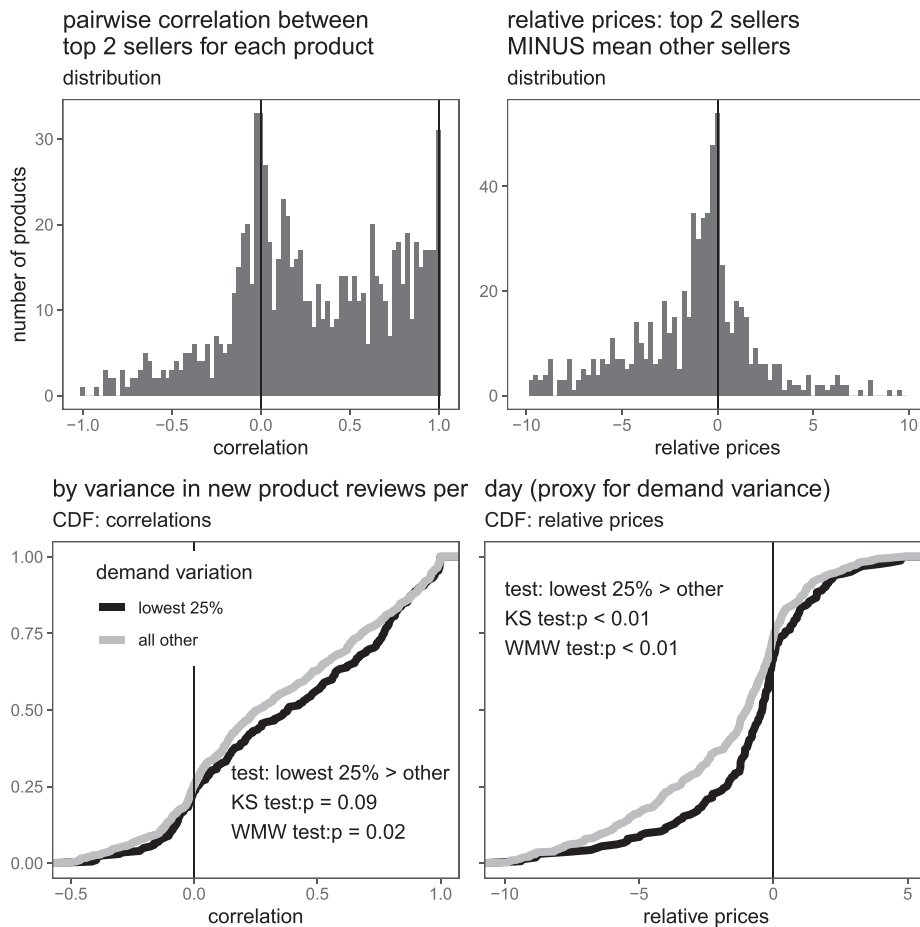
5. Empirical Relevance

To provide the empirical relevance of our results, we reanalyze pricing data from Chen et al. (2016). These data include a price tracker for best-selling products on Amazon.com in two web crawls: from September 15 to December 8, 2014, and from August 11 to September 21, 2015. In each web crawl, individual sellers' prices and ratings were collected every 25 minutes. For each product, we define the top two sellers as the sellers with the highest listings (defined by page and rank within the page) with Prime shipping (to remove variation in shipping prices and times). We consider 830 products with variations in prices for both sellers.

We consider two measures of interest: (1) pairwise correlations in prices for the top two sellers over time and (2) the relative prices of the top two sellers versus all other sellers. For each product, we define the relative price as the mean price for the top two sellers minus the mean price for all other sellers. The results are shown in top row of Figure 4. The top left chart plots the distribution of pairwise correlations and shows that the distribution is bimodal with modes near zero and one. This is consistent with our simulation settings, in which correlations are either near zero (small SNR) or near one (large SNR).¹² The top right chart plots the distribution of relative prices. The median relative price in the data is $-\$1.04$ (the top two sellers price is $\$1.04$ below the average price of all other sellers); however, there is large cross-sectional variation across products.

To investigate the cross-sectional variation across products, we consider a proxy for the variance in

Figure 4. Analysis of Amazon Pricing Data in Chen et al. (2016)



Notes. The left column considers pairwise correlations between the top two sellers (based on highest listings) for each product. The right column considers prices defined as the prices for the top two sellers minus the mean price of all other sellers. The top row plots histograms to show the distribution of the raw data. The bottom chart considers the CDF of each measure by the number of new product reviews per hour (as a proxy for demand variation). We consider the products with the lowest 25% of demand variation and all other products. The two tests performed are the Kolmogorov–Smirnov test and the Wilcoxon–Mann–Whitney test to see if the distributions are shifted to the right for lower demand variation (the CDF of all others lie above that of the lowest 25% demand variation).

demand shocks (i.e., the magnitude of δ in our model). For this, we consider the variance in number of new product reviews per day for each product. Intuitively, if agents leave reviews at roughly the same rate, then variance in the rate of new reviews equates to variance in demand. The charts in the bottom row of Figure 4 plot the cumulative distribution function (CDF) for the products with the lowest demand variation (lowest 25%) versus all other products. The bottom left chart shows the CDF of pairwise correlations, and the bottom right chart shows the CDF for relative prices. We observe that the CDF for products with the lowest demand variation is shifted to the right for both charts (statistically significant for relative prices). This suggests that the top two sellers for products with the lowest demand variation have (a) a higher correlation in prices over time and (b) higher prices relative to other sellers.

Both these trends in the Amazon.com data are consistent with our results in which we find that markets with the higher SNR (lower noise) have the higher correlations in prices and higher levels of prices. Of course, because we do not observe (a) the algorithms used by sellers on Amazon, (b) demand for each product, and (c) supply conditions (marginal costs) and (d) the Chen et al. (2016) data set has a nonrandom selection of products, we do not assert that these correlations and prices are uniquely a result of the mechanism described in our paper.

6. Conclusion

With the growth of e-commerce, we have seen increased usage of algorithms automating pricing decisions. The pricing algorithm runs automated field experiments to learn about the demand curve and each product's profit-maximizing price. Given the complexity and scale of online markets, field experiments typically assume firms are monopolists or oligopolists best responding to fixed competitors' prices. Although this assumption greatly reduces complexity, it also results in a misspecified pricing algorithm. In this paper, we study outcomes in an oligopoly setting in which all competing sellers independently use such misspecified algorithms for pricing.

We show that the long-run prices that result depend critically on the informational value (signal-to-noise ratio) of pricing experiments. If low, the long-run prices are competitive and the misspecified algorithms achieve nearly first-best profits. However, if high, the long-run prices are supra-competitive. We show that this occurs via a novel channel: competitors' algorithms' prices end up being correlated through the experiments. Therefore, sellers' misspecified models overestimate their own price sensitivity, resulting in higher prices. We believe the identification of this novel channel raises important new concerns for competition policy.

In terms of future research, although we reveal a new channel by which collusive-seeming pricing is possible, there is much work needed to understand how robust this finding is. For instance, our theoretical results investigate a stylized model in which symmetric firms simultaneously price using the UCB algorithm and there is no noise. What if firms set prices asynchronously (as in Brown and MacKay 2020) or demand is not stationary (as in Keskin and Zeevi 2017) or with price discrimination (e.g., Dubé and Misra 2017)? How much noise is necessary to “break” the collusion? What are the properties of algorithms other than UCB that can sustain such outcomes? Our simulations provide insight and intuition, but a full theoretical analysis would be desirable. Relatedly, although we provide indicative evidence, it would be interesting to see if empirical work could robustly show the existence of such pricing patterns in the real world.

Acknowledgments

The authors thank Joe Harrington and Jan Svitak as well as seminar participants at the University of California (UC), Berkeley; UC Los Angeles; UC Riverside; and UC San Diego for helpful comments and discussion.

Appendix

A.1. Proof of Theorem 1

We simulate the evolution of the UCB algorithm. By the definition of UCB, each firm, in the first two periods, experiments exactly once with each of the two prices p_H and p_L . There are now exactly two possible class histories at the end of the first two rounds:

1. Matched prices, that is, (p_H, p_H) in one round and (p_L, p_L) in the other.
2. Mismatched prices, that is, (p_L, p_H) in one round and (p_H, p_L) in the other.

We consider each of these in turn.

A.1.1. Case 1: Matched Prices Note that, at the end of the first two periods, the “state”¹³ of each firm's algorithm can be summarized as $(\pi_{HH}, 1, \pi_{LL}, 1)$. To see part (1) of Theorem 1, note that the algorithms and demand system are deterministic, and both firms' algorithms share a common state at the end of period 2. Therefore, the firms take the same action in period 3 and then, inductively, in every subsequent period.

Now $\pi_{HH} > \pi_{LL}$, and both firms charge the high price as long as

$$\pi_{HH} + \sqrt{\frac{2 \ln t}{t-2}} > \pi_{LL} + \sqrt{2 \ln t}$$

$$\text{Or, } \pi_{HH} - \pi_{LL} > \sqrt{2 \ln t} \left(1 - \frac{1}{\sqrt{t-2}} \right).$$

Note that, for some t large enough, this inequality is violated (as the right-hand side is strictly increasing in t and the left-hand side is constant); however, because demand is deterministic, it is violated at the same time for both firms.

At such t , both firms switch to charging p_L . In period $t + 1$, the state of both firms' algorithms is, therefore, $(\pi_{HH}, (t - 2), \pi_{LL}, 2)$. However, because (1) p_H was chosen in $t - 1$ and (2) the exploration bonus decreases with the number of attempts,¹⁴ we must have

$$\pi_{HH} - \pi_{LL} > \sqrt{2 \ln(t+1)} \left(\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{t-2}} \right).$$

In other words, both firms immediately switch back to charging p_H . Further, both firms charge p_L for the n^{th} time in period t_n subject to

$$\pi_{HH} - \pi_{LL} < \sqrt{\frac{2 \ln(t_n)}{n-1}} - \sqrt{\frac{2 \ln(t_n)}{t_n - n}}.$$

Therefore, n cannot be larger than $O(\ln t_n)$ as it grows large. As a result $\frac{n}{t_n}$ goes to zero as n grows large.

A.1.2. Case 2: Mismatched Prices As in case 1, at the end of the first two periods, the two firms share a common state, $(\pi_{HL}, 1, \pi_{LH}, 1)$. As a result, both firms take the same action in

period 3 and then, inductively, in every subsequent period. This shows (1) of Theorem 1.

Consider any subsequent period in which both firms have charged the high price n_H times and the low price n_L times. The corresponding indices in period $t = n_H + n_L + 1$ are

$$I_{H,t} = \left(\frac{1}{n_H} \pi_{HL} + \frac{n_H - 1}{n_H} \pi_{HH} \right) + \sqrt{\frac{2 \ln t}{n_H}}$$

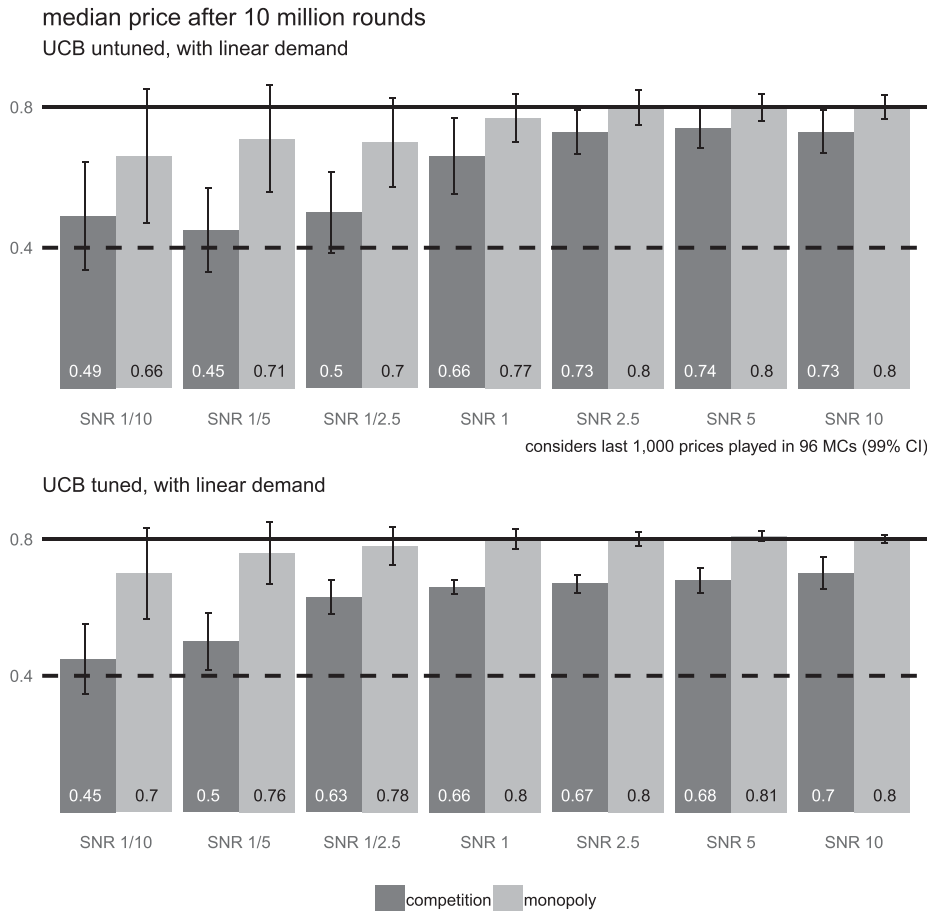
$$I_{L,t} = \left(\frac{1}{n_L} \pi_{LH} + \frac{n_L - 1}{n_L} \pi_{LL} \right) + \sqrt{\frac{2 \ln t}{n_L}}.$$

Observe that the first term of the first equality is increasing to π_{HH} as n_H grows large, and the first term of the second is decreasing to π_{LL} . Part (2) of Theorem 1 now follows from an analogous argument to that presented. \square

A.2. Robustness

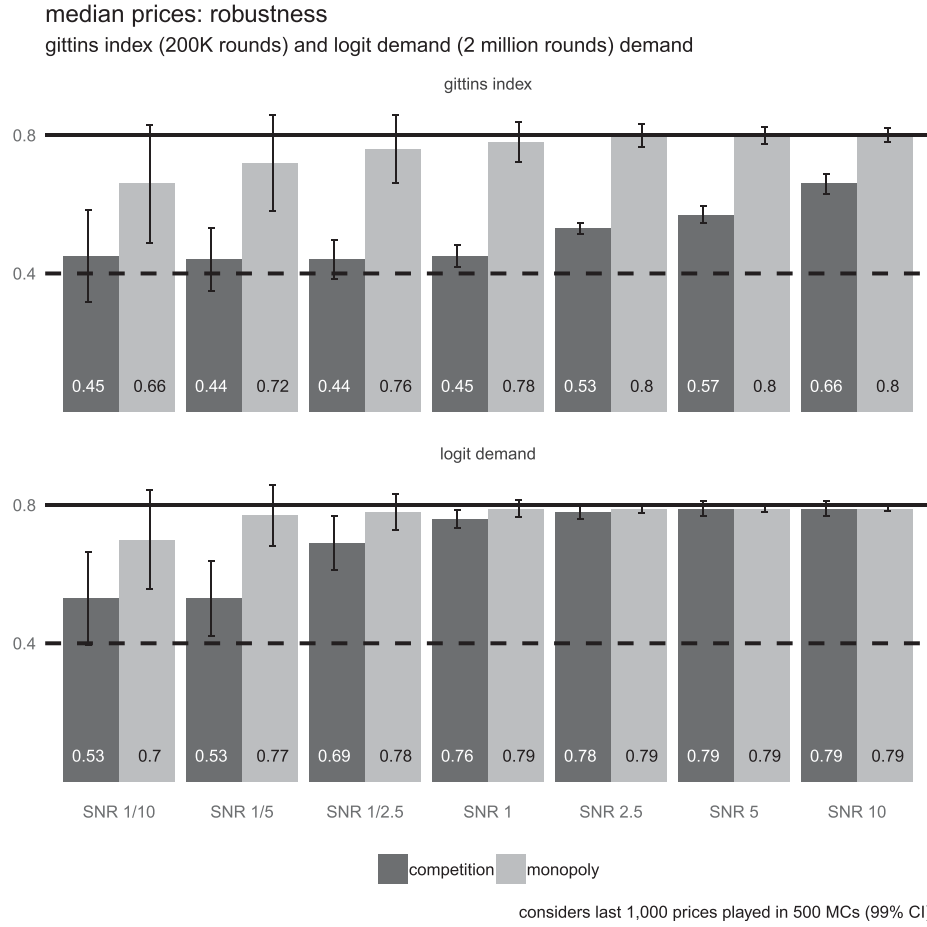
Figure A.1 shows robustness to the specific algorithm in the main paper. (1) We show that our results are robust to using the UCB untuned (Auer et al. 2002) and the number of

Figure A.1. Estimated Median Prices by Algorithm and Level of Competition



Notes. The error bars represent 99% confidence intervals for the estimated median across MC simulations. The dark gray bars represent a setting in which two firms are running simultaneous algorithms, and the light gray bars represent a setting in which a monopolist is jointly pricing the products. For each simulation, we consider the median price charged in the last 1,000 rounds out of 10 million rounds. The dashed lines reflect the competitive equilibrium prices; the solid lines reflect monopoly prices. (Algorithm used: UCB tuned.)

Figure A.2. Estimated Median Prices by Algorithm and Level of Competition



Notes. Gittins (1979): linear demand model using the closed-form Gittins index as suggested in Brezzi and Lai (2002); we assume a beta prior and approximate the profit as an aggregate binomial distribution. Note that this algorithm converges faster than UCB, and hence, we consider 200,000 rounds. Logit demand model: $u_j = 4.1 - 4.74p_j + \epsilon_j$ for $j \in \{1, 2\}$ with an outside option $u_0 = \epsilon_0$ with ϵ i.i.d. type 1 extreme value. Under this demand system, the Nash equilibrium price $P^D = 0.4$ and the monopoly price $P^M = 0.8$. The error bars represent 99% confidence intervals for the estimated median across MC simulations. The dark gray bars represent a setting in which two firms are running simultaneous algorithms, and the light gray bars represent a setting in which a monopolist is jointly pricing the products. For each simulation, we consider the median price charged in the last 1,000 rounds in each MC. The dashed lines reflect the competitive equilibrium prices; the solid lines reflect monopoly prices. (Algorithm used: Gittins index and UCB tuned.)

rounds to run the algorithm. (2) The UCB is based on a finite time analysis of the MAB problem and requires an ex ante specification of number of rounds. In this figure, we run the algorithm for 10 million rounds to show the robustness of our results. The results are statistically indistinguishable from the main results in Figure 1.

The UCB index for the online appendix is given by

$$\text{UCB-untuned}_{kt} = \bar{\pi}_{kt} + \sqrt{\frac{2 \log t}{n_{kt}}}.$$

In the top panel of Figure A.1, we show that our results are robust to considering the UCB-untuned algorithm. Note that, consistent with the prior literature (including in Auer et al. 2002), the untuned version of the algorithm takes longer to converge and requires a larger number of rounds

(10 million). In the bottom panel of Figure A.1, we show that our results are robust to considering 10 million rounds with the UCB-untuned algorithm (as opposed to two million in the main paper).

Figure A.2 shows the robustness of our results in two ways. First, we consider a different demand system (logit demand system). Second, we use a different MAB index algorithm (Gittins index). Our main findings (see Figure 1) are replicated in both these simulations: markets with low (high) SNR, independent bandits result in competitive (supra-competitive) prices.

Endnotes

¹Please see <https://newsroom.spotify.com/company-info/>.

²Please see <https://www.scrapehero.com/number-of-products-on-amazon-april-2019/>.

³Please see <https://www.cnn.com/2019/01/17/media/netflix-earnings-q4/index.html>.

⁴See section 4 of den Boer (2015) for a more complete overview of dynamic monopoly pricing with unknown demand. We discuss this literature in more detail in the online appendix.

⁵An implication is that, even if the seller is a monopolist, the limit price need not converge to the correct optimal price if the assumed linear demand is incorrect. The UCB algorithm used in our paper considers (undiscounted) dynamic profits without making parametric demand assumptions.

⁶For an overview, see Sutton and Barto (1998). For applications in similar settings, see Hauser et al. (2009) for a marketing application and Misra et al. (2019) for pricing.

⁷In Section A.2, we show that our results are robust to considering the untuned version (Auer et al. 2002) called UCB1. Note that the tuned version is shown to have better empirical performance; for example, see Auer et al. (2002) (general) or Misra et al. (2019) (application to demand learning). We also show that our results are robust to other index algorithms, namely the Gittins index (Gittins 1989, Brezzi and Lai 2002).

⁸In a setting with infinite time, there is no trade-off between learning and earning as, once found, the best arm is played for an infinite time.

⁹To check robustness, in Section A.2, we present results when the number of periods is increased to 10 million. Our results are statistically indistinguishable from those presented in the main paper.

¹⁰Numerically, we can see this as the decrease in the median difference between the two firms' prices (top left of the chart); this decreases from 0.15 to 0.00. Alternatively, we consider the percentage of simulations in which resultant prices are within one cent of each other (bottom left of chart). Again, we see this number increase from 3% in the low-SNR setting to 69% in the high-SNR setting. This suggests that the signal strength of experiments is critical to coordinating prices across different algorithms.

¹¹Consistent with this intuition, in the online appendix, we show a simulation in which we allow demand shocks to increase or decrease by round and show our results are robust, further suggesting the importance of initial noise to make the algorithms uncorrelated.

¹²We investigate high correlations resulting from sellers identified as "algorithmic competitive sellers" in Chen et al. (2016). We find that seller pairs with no identified sellers have a higher median correlation (0.54) than seller pairs with one (0.22) or two (0.14) identified sellers. This suggests that the higher correlations are not a result of algorithmic competitive sellers.

¹³The state variables are (1) average profits for price H , (2) the number of times price H has been charged, (3) average profits for price L , and (4) the number of times price L has been charged.

¹⁴Formally, (1) $\pi_{HH} - \pi_{LL} > \sqrt{2 \ln(t-1)}(1 - \frac{1}{\sqrt{t-3}})$, and (2) $\frac{\partial(\frac{1}{\sqrt{t}} - \frac{K}{\sqrt{t+1}})}{\partial x} < 0$, $\forall K < 1$.

References

- Aghion P, Bolton P, Harris C, Jullien B (1991) Optimal learning by experimentation. *Rev. Econom. Stud.* 58(4):621–654.
- Agrawal R (1995) Sample mean based index policies with $O(\log n)$ regret for the multi-armed bandit problem. *Adv. Appl. Probab.* 27(4):1054–1078.
- Auer P (2002) Using confidence bounds for exploitation-exploration trade-offs. *J. Machine Learn. Res.* 3:397–422.
- Auer P, Ortner R (2010) UCB revisited: Improved regret bounds for the stochastic multi-armed bandit problem. *Periodica Mathematica Hungarica* 61:55–65.
- Auer P, Cesa-Bianchi N, Fischer P (2002) Finite-time analysis of the multiarmed bandit problem. *Machine Learn.* 47:235–256.
- Ban GY, Keskin NB (2021) Personalized dynamic pricing with machine learning: High dimensional features and heterogeneous elasticity. *Management Sci.* Forthcoming.
- Besbes O, Zeevi A (2009) Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Oper. Res.* 57(6):1407–1420.
- Brezzi M, Lai TL (2002) Optimal learning and experimentation in bandit problems. *J. Econom. Dynamics Control* 27(1):87–108.
- Broder J, Rusmevichientong P (2012) Dynamic pricing under a general parametric choice model. *Oper. Res.* 60(4):965–980.
- Brown Z, MacKay A (2020) Competition in pricing algorithms. Preprint, submitted December 19, <https://dx.doi.org/10.2139/ssrn.3485024>.
- Calvano E, Calzolari G, Denicolò V, Pastorello S (2019) Algorithmic pricing what implications for competition policy? *Rev. Indust. Organ.* 55:155–171.
- Calvano E, Calzolari G, Denicolò V, Pastorello S (2020) Artificial intelligence, algorithmic pricing and collusion. *Amer. Econom. Rev.* 110(10):3267–3297.
- Chen L, Mislove A, Wilson C (2016) An empirical analysis of algorithmic pricing on Amazon Marketplace. *Proc. 25th Internat. World Wide Web Conf. (International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE)*, 1339–1349.
- Cheung WC, Simchi-Levi D, Wang H (2017) Dynamic pricing and demand learning with limited price experimentation. *Oper. Res.* 65(6):1722–1731.
- Cooper WL, Homem-de Mello T, Kleywegt AJ (2015) Learning and pricing with models that do not explicitly incorporate competition. *Oper. Res.* 63(1):86–103.
- den Boer AV (2015) Dynamic pricing and learning: Historical origins, current research, and new directions. *Surveys Oper. Res. Management Sci.* 20(1):1–18.
- Dubé J-P, Misra S (2017) Scalable price targeting. Working Paper 23775, National Bureau of Economic Research, <https://ssrn.com/abstract=3035110>.
- Gittins JC (1979) Bandit processes and dynamic allocation indices. *J. Roy. Statist. Soc. B* 41(2):148–177.
- Gittins JC (1989) *Multi-Armed Bandit Allocation Indices*, 1st ed. (John Wiley and Sons, Chichester, UK).
- Handel B, Misra K (2015) Robust new product pricing. *Marketing Sci.* 34(6):864–881.
- Harrington JE (2018) Developing competition law for collusion by autonomous artificial agents. *J. Competition Law Econom.* 14(3):331–363.
- Hauser JR, Urban GL, Liberali G, Braun M (2009) Website morphing. *Marketing Sci.* 28(2):202–223.
- Keskin NB, Zeevi A (2017) Chasing demand: Learning and earning in a changing environment. *Math. Oper. Res.* 42(2):277–307.
- Misra K, Schwartz EM, Abernethy J (2019) Dynamic online pricing with incomplete information using multiarmed bandit experiments. *Marketing Sci.* 38(2):226–252.
- OECD (2017) Algorithms and Collusion: Competition Policy in the Digital Age. www.oecd.org/competition/algorithms-collusion-competition-policy-in-the-digital-age.htm.
- Sutton RS, Barto AG (1998) *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA).
- Weber R (1992) On the Gittins index for multiarmed bandits. *Ann. Appl. Probab.* 2(4):1024–1033.
- The White House (2015) Big data and differential pricing. Report, Council of Economic Advisors, Washington, D.C., https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf.