Shopbots and Pricebots

Amy R. Greenwald and Jeffrey O. Kephart
amygreen@cs.nyu.edu,kephart@watson.ibm.com

IBM Institute for Advanced Commerce
IBM Thomas J. Watson Research Center
Yorktown Heights, NY 10598

Abstract. *Shopbots* are software agents that automatically gather and collate information from multiple on-line vendors about the price and quality of consumer goods and services. Rapidly increasing in number and sophistication, shopbots are helping more and more buyers minimize expenditure and maximize satisfaction. In response to this trend, it is anticipated that sellers will come to rely on *pricebots*, automated agents that employ price-setting algorithms in an attempt to maximize profits. In this paper, a simple economic model is proposed and analyzed, which is intended to characterize some of the likely impacts of a proliferation of shopbots and pricebots.

In addition to describing theoretical investigations, this paper also aims toward a practical understanding of the tradeoffs between profitability and computational and informational complexity of pricebot algorithms. A comparative study of a series of price-setting strategies is presented, including: game-theoretic (GT), myoptimal (MY), derivative following (DF), and no regret learning (NR). The dynamic behavior that arises among collections of pricebots and shopbot-assisted buyers is simulated, and it is found that game-theoretic equilibria can dynamically arise in our model of shopbots and pricebots.

1 Introduction

Shopbots — automated Web agents that query multiple on-line vendors to gather information about prices and other attributes of consumer goods and services — herald a future in which automated agents are an essential component of electronic commerce [6, 10, 24, 31]. Shopbots outperform and out-inform humans, providing extensive product coverage in a few seconds, far more than a patient and determined human shopper could ever achieve, even after hours of manual search. Rapidly increasing in number and sophistication, shopbots are helping more and more buyers to minimize expenditure and maximize satisfaction.

Since the launch of BargainFinder [26] — a CD shopbot — on June 30, 1995, the range of products represented by shopbots has expanded dramatically. A shopbot available at `shopper.com` claims to compare 1,000,000 prices on 100,000 computer-oriented products. `DealPilot.com` (formerly `acses.com`) is a shopbot that gathers, collates and sorts prices and expected delivery times of books, CDs, and movies offered for sale on-line. One of the most popular shopbots, mysimon.com, compares office supplies, groceries, toys, apparel, and consumer electronics, just to name a few of the items on its product line. As the range of products covered by shopbots expands to include more complex products such as consumer electronics, the level of shopbot sophistication is rising accordingly. On August 16th, 1999, mysimon.com incorporated technology that, for products with multiple features such as digital cameras, uses a series of questions to elicit multi-attribute utilities from buyers, and then sorts products according to the buyer's specified utility. Also on that day, lycos.com licensed similar technology from frictionless.com.

Shopbots are undoubtedly a boon to buyers who use them. Moreover, when shopbots become adopted by a sufficient portion of the buyer population, it seems likely that sellers will be compelled to decrease prices and improve quality, benefiting even those buyers who do not shop with bots. How the widespread utilization of shopbots might affect sellers, on the other hand, is not quite so apparent. Less established sellers may welcome shopbots as an opportunity to attract buyers who might not otherwise have access to information about them, but more established sellers may feel threatened. Some larger players have even been known to deliberately block automated agents from their web sites [8]. This practice seems to be waning, however; today, sellers like Amazon.com and BarnesandNoble.com tolerate queries from agents such as DealPilot.com on the grounds that buyers take brand name and image as well as price into account.

As more and more buyers rely on shopbots to supplement their awareness about products and prices, it is becoming advantageous for sellers to increase flexibility in their pricing strategies, perhaps via *pricebots* — software agents that utilize automatic price-setting algorithms in an attempt to maximize profits. Indeed, an early pricebot is already available at books.com, an on-line bookseller. When a prospective buyer expresses interest in a book at books.com, a pricebot automatically queries Amazon.com, Borders.com, and BarnesandNoble.com to determine the price that is being offered at those sites. books.com then slightly undercuts the lowest of the three quoted prices, typically by 1% of the retail price. Real-time dynamic pricing on millions of titles is impossible to achieve manually, yet can easily be implemented with a modest amount of programming.

As more and more sellers automate their price-setting, pricebots are destined to interact with one another, yielding unexpected price and profit dynamics. This paper reaches toward an understanding of pricebot dynamics via analysis and simulation of a series of candidate price-setting strategies, which differ in their informational and computational demands: game-theoretic pricing (GT), myoptimal pricing (MY), derivative following (DF), and no regret learning (NR). Previously, we studied the dynamics that ensue when shopbot-assisted buyers interact with pricebots utilizing only a subset of these strategies [19, 25, 29]. In this work, we simulate additional, more sophisticated, pricebot strategies, and find that the game-theoretic equilibrium can arise dynamically as the outcome of adaptive learning in our model of shopbots and pricebots.

This paper is organized as follows. The next section presents our model of an economy that consists of shopbots and pricebots. This model is analyzed from a game-theoretic point of view in Sec. 3. In Sec. 4, we discuss the price-setting strategies of interest: game-theoretic, myoptimal pricing, derivative following, and no regret learning. Sec. 5 describes simulations of pricebots that implement these strategies, while Sec. 6 discusses one possible evolution of shopbots and pricebots, and Sec. 7 presents our conclusions.

2 Model

We consider an economy in which there is a single homogeneous good that is offered for sale by S sellers and of interest to B buyers, with $B \gg S$. Each buyer b generates purchase orders at random times, with rate ρ_b, while each seller s reconsiders (and potentially resets) its price p_s at random times, with rate ρ_s. The value of the good to buyer b is v_b; the cost of production for seller s is c_s.

A buyer b's utility for the good is a function of price:

$$u_b(p) = \begin{cases} v_b - p & \text{if } p \leq v_b \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

which states that a buyer obtains positive utility if and only if the seller's price is less than the buyer's valuation of the good; otherwise, the buyer's utility is zero. We do not assume that buyers are utility maximizers; instead we assume that they consider the prices offered by sellers using one of the following strategies:[1]

1. *Any Seller:* buyer selects seller at random, and purchases the good if the price charged by that seller is less than the buyer's valuation.

2. *Bargain Hunter:* buyer checks the offer price of all sellers, determines the seller with the lowest price, and purchases the good if that lowest price is less than the buyer's valuation. (This type of buyer corresponds to those who take advantage of shopbots.)

The buyer population consists of a mixture of buyers employing one of these strategies, with a fraction w_A using the *Any Seller* strategy and a fraction w_B using the *Bargain Hunter* strategy, where $w_A + w_B = 1$. Buyers employing these respective strategies are referred to as type A and type B buyers.

[1] In this framework, it is also possible to consider all buyers as utility maximizers, with the additional cost of searching for the lowest price made explicit in the buyer utility functions. In doing so, the search cost for bargain hunters is taken to be zero, while for those buyers that use the any seller strategy, its value is greater than v_b. The relationship between models of exogenously determined buyer behavior and the endogenous approach which incorporates the cost of information acquisition and explicitly allows for buyer decision-making is further explored in computational settings in Kephart and Greenwald [25]; in the economics literature, see, for example, Burdett and Judd [5] and Wilde and Schwartz [33].

A seller s's expected profit per unit time π_s is a function of the price vector $\boldsymbol{p}$, as follows: $\pi_s(\boldsymbol{p}) = (p_s - c_s)D_s(\boldsymbol{p})$, where $D_s(\boldsymbol{p})$ is the rate of demand for the good produced by seller s. This rate of demand is determined by the overall buyer rate of demand, the likelihood of the buyers selecting seller s as their potential seller, and the likelihood that seller s's price p_s does not exceed the buyer's valuation v_b.[2] If $\rho = \sum_b \rho_b$, and if $h_s(\boldsymbol{p})$ denotes the probability that seller s is selected, while $g(p_s)$ denotes the fraction of buyers whose valuations satisfy $v_b \geq p_s$, then $D_s(\boldsymbol{p}) = \rho B h_s(\boldsymbol{p}) g(p_s)$. Without loss of generality, define the time scale *s.t.* $\rho B = 1$. Now $\pi_s(\boldsymbol{p})$ is interpreted as the expected profit for seller s per unit sold systemwide. Moreover, seller s's profit is such that $\pi_s(\boldsymbol{p}) = (p_s - c_s)h_s(\boldsymbol{p})g(p_s)$. We discuss the functions $h_s(\boldsymbol{p})$ and $g(p)$ presently.

The probability $h_s(\boldsymbol{p})$ that buyers select seller s as their potential seller depends on the buyer distribution (w_A, w_B) as follows:

$$h_s(\boldsymbol{p}) = w_A f_{s,A}(\boldsymbol{p}) + w_B f_{s,B}(\boldsymbol{p}) \tag{2}$$

where $f_{s,A}(\boldsymbol{p})$ and $f_{s,B}(\boldsymbol{p})$ are the probabilities that seller s is selected by buyers of type A and B, respectively. The probability that a buyer of type A selects a seller s is independent of the ordering of sellers' prices: $f_{s,A}(\boldsymbol{p}) = 1/S$. Buyers of type B, however, select a seller s if and only if s is one of the lowest price sellers. Given that the buyers' strategies depend on the relative ordering of the sellers' prices, it is convenient to define the following functions:

- $\lambda_s(\boldsymbol{p})$ is the number of sellers charging a lower price than s,
- $\tau_s(\boldsymbol{p})$ is the number of sellers charging the same price as s, excluding s itself, and
- $\mu_s(\boldsymbol{p})$ is the number of sellers charging a higher price than s.

Now a buyer of type b selects a seller s iff s is *s.t.* $\lambda_s(\boldsymbol{p}) = 0$, in which case a buyer selects a particular such seller s with probability $1/(\tau_s(\boldsymbol{p}) + 1)$. Therefore,

$$f_{s,B}(\boldsymbol{p}) = \frac{1}{\tau_s(\boldsymbol{p}) + 1}\, \delta_{\lambda_s(\boldsymbol{p}),0} \tag{3}$$

where $\delta_{i,j}$ is the Kronecker delta function, equal to 1 whenever $i = j$, and 0 otherwise.

The function $g(p)$ can be expressed as $g(p) = \int_p^\infty \gamma(x)dx$, where $\gamma(x)$ is the probability density function describing the likelihood that a given buyer has valuation x. For example, suppose that the buyers' valuations are uniformly distributed between 0 and v, with $v > 0$; then the integral yields $g(p) = 1 - p/v$. This case was studied in Greenwald, *et al.* [20]. In this paper, we assume $v_b = v$ for all buyers b, in which case $\gamma(x)$ is the Dirac delta function $\delta(v - x)$, and the integral yields a step function $g(p) = \Theta(v - p)$ as follows:

$$\Theta(v - p) = \begin{cases} 1 & \text{if } p \leq v \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

[2] We assume that buyers' valuations are uncorrelated with their buying strategies.

The preceding results can be assembled to express the profit function π_s for seller s in terms of the distribution of strategies and valuations within the buyer population. Recalling that $v_b = v$ for all buyers b, and assuming $c_s = c$ for all sellers s, yields the following:

$$\pi_s(\boldsymbol{p}) = \begin{cases} (p_s - c)h_s(\boldsymbol{p}) & \text{if } p_s \leq v \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where

$$h_s(\boldsymbol{p}) = w_A \frac{1}{S} + w_B \frac{1}{\tau_s(\boldsymbol{p}) + 1}\, \delta_{\lambda_s(\boldsymbol{p}),0} \tag{6}$$

3 Analysis

In this section, we present a game-theoretic analysis of the prescribed model viewed as a one-shot game.[3] Assuming sellers are profit maximizers, we first show that there is no pure strategy Nash equilibrium, and we then compute the symmetric mixed strategy Nash equilibrium. A Nash equilibrium is a vector of prices $\boldsymbol{p}^* \in \mathbb{R}^S$ at which sellers maximize their individual profits and from which they have no incentive to deviate [28]. Recall that $B \gg S$; in particular, the number of buyers is assumed to be very large, while the number of sellers is a good deal smaller. In accordance with this assumption, it is reasonable to study the strategic decision-making of the sellers alone, since their relatively small number suggests that the behavior of individual sellers indeed influences market dynamics, while the large number of buyers renders the effects of individual buyers' actions negligible.

Traditional economic models consider the case in which all buyers are bargain hunters: *i.e.*, $w_B = 1$. In this case, prices are driven down to marginal cost; in particular, $p_s^* = c$, for all sellers s (see, for example, Tirole [30]). In contrast, consider the case in which all buyers are of type A, meaning that they randomly select a potential seller: *i.e.*, $w_A = 1$. In this situation, tacit collusion arises, in which all sellers charge the monopolistic price, in the absence of explicit coordination; in particular, $p_s^* = v$, for all sellers s. Of particular interest in this study, however, is the dynamics of interaction among buyers of various types: *i.e.*, $0 < w_A, w_B < 1$. Knowing that buyers of type A alone results in all sellers charging the valuation price v, we investigate the impact of buyers of type B, or shopbots, on the marketplace.

[3] The analysis presented in this section applies to the one-shot version of our model, although the simulation results reported in Sec. 5 focus on repeated settings. We consider the Nash equilibrium of the one-shot game, rather than its iterated counterpart, for at least two reasons, including (i) the Nash equilibrium of the stage game played repeatedly is in fact a Nash equilibrium of the repeated game, and (ii) the Folk Theorem of repeated game theory (see, for example, Fudenberg and Tirole [15]) states that virtually all payoffs in a repeated game correspond to a Nash equilibrium, for sufficiently large values of the discount parameter. Thus, we isolate the stage game Nash equilibrium as an equilibrium of particular interest.

Throughout this exposition, we adopt the standard notation $\boldsymbol{p} = (p_s, p_{-s})$, which distinguishes the price offered by seller s from the prices offered by the other sellers. Our analysis begins with the following observation: *at equilibrium, at most one seller s charges $p_s^* < v$.* Suppose that two distinct sellers $s' \neq s$ set their equilibrium prices to be $p_{s'}^* = p_s^* < v$, while all other sellers set their equilibrium prices at the buyers' valuation v. In this case, $\pi_s(p_s^* - \epsilon, p_{-s}^*) = [(1/S)w_A + w_B](p_s^* - \epsilon - c) > [(1/S)w_A + (1/2)w_B](p_s^* - c) = \pi_s(p_s^*, p_{-s}^*)$, for small, positive values of ϵ,[4] which implies that p_s^* is not an equilibrium price for seller s. Now suppose that two distinct sellers $s' \neq s$ set their equilibrium prices to be $p_{s'}^* < p_s^* < v$, while all other sellers set their equilibrium prices at v. In this case, seller s prefers price v to p_s^*, since $\pi_s(v, p_{-s}^*) = [(1/S)w_A](v - c) > [(1/S)w_A](p_s^* - c) = \pi_s(p_s^*, p_{-s}^*)$, which again implies that p_s^* is not an equilibrium price for seller s. In sum, no 2 (or more) sellers charge equal equilibrium prices strictly below v, and no 2 (or more) sellers charge unequal equilibrium prices strictly below v. Therefore, at most one seller charges $p_s^* < v$.

On the other hand, *at equilibrium, at least one seller s charges $p_s^* < v$.* Given that all sellers other than s set their equilibrium prices at v, seller s maximizes its profits by charging $v - \epsilon$, since $\pi_s(v - \epsilon, p_{-s}^*) = [(1/S)w_A + w_B](v - \epsilon - c) > [(1/S)(w_A + w_B)](v - c) = \pi_s(v, p_{-s}^*)$, for small, positive values of ϵ.[5] Thus, v is not an equilibrium price for seller s. It follows from these two observations that at equilibrium, exactly one seller s sets its price below the buyers' valuation v, while all other sellers $s' \neq s$ set their equilibrium prices $p_{s'}^* \geq v$. Note that $\pi_{s'}(v, p_{-s'}^*) = [(1/S)w_A](v - c) > 0 = \pi_{s'}(v', p_{-s'}^*)$, for all $v' > v$, since $w_A > 0$, implying that all other sellers s' maximize their profits by charging price v. The unique form of pure strategy equilibrium which arises in this setting thus requires that a single seller s set its price $p_s^* < v$ while all other sellers $s' \neq s$ set their prices $p_{s'}^* = v$. The price vector (p_s^*, p_{-s}^*), with $p_{-s}^* = (v, \ldots, v)$, however, is not a Nash equilibrium. While v is in fact an optimal response to p_s^*, since the profits of seller $s' \neq s$ are maximized at v given that there exists low-priced seller s, p_s^* is not an optimal response to v. On the contrary, $\pi_s(p_s^*, v, \ldots, v) < \pi_s(v - \epsilon, v, \ldots, v)$, whenever $\epsilon < v - p_s^*$. In particular, the low-priced seller s has incentive to deviate. It follows that there is no pure strategy Nash equilibrium in the proposed shopbot model.[6]

There does, however, exist a symmetric mixed strategy Nash equilibrium. Let $f(p)$ denote the probability density function according to which sellers set their equilibrium prices, and let $F(p)$ be the corresponding cumulative distribution function. Following Varian [32], we note that in the range for which it is defined, $F(p)$ has no mass points, since otherwise a seller could decrease its price by an arbitrarily small amount and experience a discontinuous increase in profits. Moreover, there are no gaps in the said distribution, since otherwise prices would

---

[4] Precisely, $0 < 2\epsilon < \frac{w_B S}{w_A + w_B S}(p_s^* - c)$.

[5] Precisely, $0 < \epsilon < \frac{w_B(S-1)}{w_A + w_B S}(v - c)$.

[6] This argument rests on the fact that price selection is made within a continuous strategy space; the existence of pure strategy Nash equilibria as an outcome of price discretization is discussed in Appendix A.

not be optimal — a seller charging a price at the low end of the gap could increase its price to fill the gap while retaining its market share, thereby increasing its profits. In this probabilistic setting, the event that seller s is the low-priced seller occurs with probability $[1 - F(p)]^{S-1}$. Rewriting Eq. 2, we obtain the demand expected by seller s:[7]

$$h_s(p) = w_A \frac{1}{S} + w_B [1 - F(p)]^{S-1} \tag{7}$$

A Nash equilibrium in mixed strategies requires that (i) sellers maximize individual profits, given the other sellers' strategic profiles, so as there is no incentive to deviate, and (ii) all prices assigned positive probability yield equal profits, otherwise it would not be optimal to randomize. Following condition (ii), we define equilibrium profits $\pi \equiv \pi_s(p) = (p - c)h_s(p)$, for all prices p. The precise value of π can be derived by considering the maximum price that sellers are willing to charge, say p_m. At this boundary, $F(p_m) = 1$, which by Eq. 7 implies that $h_s(p_m) = (1/S)w_A$. Moreover, the function $\pi_s(p)$ attains its maximal value at price $p_m = v$, yielding equilibrium profits $\pi = (1/S)w_A(v - c)$. Now, by setting $(p - c)h_s(p)$ equal to this value and solving for $F(p)$, we obtain:

$$F(p) = 1 - \left[\left(\frac{w_A}{w_B S} \right) \left(\frac{v - p}{p - c} \right) \right]^{\frac{1}{S-1}} \tag{8}$$

which implicitly defines p and $F(p)$ in terms of one another. Since $F(p)$ is a cumulative probability distribution, it is only valid in the domain for which its valuation is between 0 and 1. As noted previously, the upper boundary is $p = v$; the lower boundary is computed by setting $F(p) = 0$ in Eq. 8, which yields:

$$p^* \equiv p = c + \frac{w_A(v - c)}{w_A + w_B S} \tag{9}$$

Thus, Eq. 8 is valid in the range $p^* \leq p \leq v$. A similar derivation of this mixed strategy equilibrium appears in Varian [32]. Greenwald, *et al.* [20] presents various generalizations of this model.

Figs 1 (a) and (b), respectively, exhibit plots of the functions $F(p)$ and $f(p)$ under varying distributions of buyer strategies — in particular, the fraction of shopbot users $w_B \in \{.1, .25, .5, .75, .9\}$ — with $S = 5$, $v = 1$, and $c = 0.5$. When w_B exceeds a critical threshold $w_B^{\text{crit}} = \frac{S-2}{S^2+S-2}$ (equal to 0.1071 for $S = 5$), $f(p)$ is bimodal. In this regime, as either w_B or S increases, the probability density concentrates either just below v, where sellers expect high margins but low volume, or just above p^*, where they expect low margins but high volume, with the latter solution becoming increasingly probable. Since p^* itself decreases under these conditions (see Eq. 9), it follows that both the average price paid by buyers and the average profit earned by sellers decrease. These relationships have a simple interpretation: buyers' use of shopbots catalyzes competition among sellers, and moreover, small fractions of shopbot users induce competition among large numbers of sellers.

[7] In Eq. 7, $h_s(p)$ is expressed as a function of seller s's scalar price p, given that probability distribution $F(p)$ describes the other sellers' expected prices.

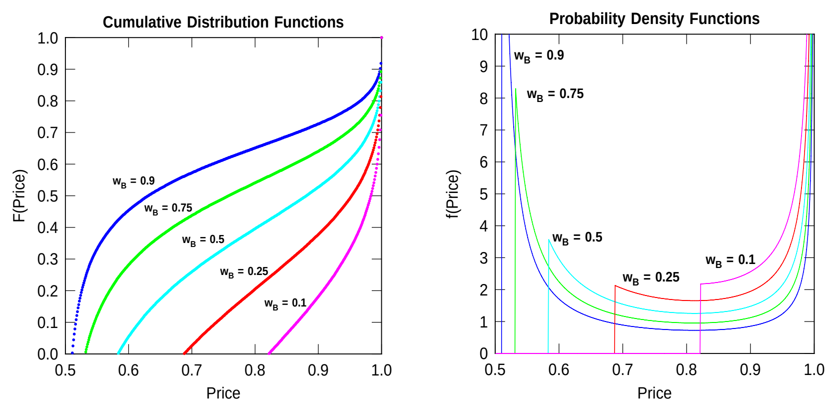

Fig. 1. Nash Equilibria for $S = 5, v = 1, c = .5$, and $w_B \in \{.1, .25, .5, .75, .9\}$

Recall that the profit earned by each seller is $(1/S)w_A$, which is strictly positive so long as $w_A > 0$. It is as though only buyers of type A are contributing to sellers' profits, although the actual distribution of contributions from buyers of type A vs. buyers of type B is not as one-sided as it appears. In reality, buyers of type A are charged less than v on average, and buyers of type B are charged more than c on average, although total profits are equivalent to what they would be if the sellers practiced perfect price discrimination. Buyers of type A exert negative externalities on buyers of type B, by creating surplus profits for sellers.

4 Strategies

When sufficiently widespread adoption of shopbots by buyers forces sellers to become more competitive, it is likely that sellers will respond with the creation of *pricebots* that automatically set prices in attempt to maximize profitability. It seems unrealistic, however, to expect that pricebots will simply compute a Nash equilibrium and fix prices accordingly. The real business world is fraught with uncertainties, undermining the validity of traditional game-theoretic analyses: sellers lack perfect knowledge of buyer demands, and they have an incomplete understanding of their competitors' strategies. In order to be deemed profitable, pricebots will need to learn from and adapt to changing market conditions.

In this section, we introduce a series of pricebot strategies, and which we later simulate in order to compare the resulting price and profit dynamics with the game-theoretic equilibrium. In 1838, Cournot showed that the outcome of learning via a simple best-reply dynamic is a pure strategy Nash equilibrium in a quantity-setting model of duopoly [7]. Recently, empirical studies of more sophisticated learning algorithms have revealed that learning tends to converge to pure strategy Nash equilibria in games for which such equilibria exist [17]. As there does not exist a pure strategy Nash equilibrium in the shopbot model, it is of particular interest to study the outcome of adaptive pricing schemes.

We consider several pricing strategies, each of which makes different demands on the required level of informational and computational power of agents:

GT The *game-theoretic* strategy is designed to reproduce the mixed strategy Nash equilibrium. It therefore generates a price chosen at random according to the probability density function derived in the previous section, assuming its competitors utilize game-theoretic pricing as well, and making full use of information about the buyer population. GT is a constant function that ignores historical observations.

MY The myopically optimal, or *myoptimal*, [8] pricing strategy (see, for example, [24]) uses information about all the buyer characteristics that factor into the buyer demand function, as well as competitors' prices, but makes no attempt to account for competitors' pricing strategies. Instead, it is based on the assumption of static expectations: even if one seller is contemplating a price change under myoptimal pricing, this seller does not consider that this will elicit a response from its competitors.

The myoptimal seller s uses all available information and the assumption of static expectations to perform an exhaustive search for the price p_s^* that maximizes its expected profit π_s. The computational demands of MY can be reduced greatly if the price quantum ϵ — the smallest amount by which one seller may undercut another — is sufficiently small (see Appendix A). Under such circumstances, the optimal price p_s^* is guaranteed to be either the monopolistic price p_m or ϵ below some competitor's price, limiting the search for p_s^* to S possible values. In our simulations, we choose $\epsilon = 0.002$.

DF The *derivative-following* strategy is less informationally intensive than either the myoptimal or the game-theoretic pricing strategies. In particular, this strategy can be used in the absence of any knowledge or assumptions about one's competitors or the buyer demand function. A derivative follower simply experiments with incremental increases (or decreases) in price, continuing to move its price in the same direction until the observed profitability level falls, at which point the direction of movement is reversed. The price increment δ is chosen randomly from a specified probability distribution; in the simulations described here the distribution was uniform between 0.01 and 0.02.

NR The *no regret* pricing strategies are probabilistic learning algorithms which specify that players *explore* the space of actions by playing all actions with some non-zero probability, and *exploit* successful actions by increasing the probability of employing those actions that generate high profits. In this study, we confine our attention to the no external regret algorithm due to Freund and Schapire [14] and the no internal regret algorithm of Foster and Vohra [12].[9] As the no regret algorithms are inherently non-deterministic, they are candidates for learning mixed strategy equilibria.

[8] In the game-theoretic literature, this strategy is often referred to as Cournot best-reply dynamics [7]; however, price is being set, rather than quantity.

[9] For completeness, the details of these algorithms are presented in App. B.

5 Simulations

We simulated an economy with 1000 buyers and (unless otherwise specified) 5 pricebots. employing the aforementioned pricing strategies. In the simulations depicted below, each buyer's valuation of the good $v = 1$, and each seller's production cost $c = 0.5$. The mixture of buyer types is set at $w_B = 0.75$: *i.e.*, 75% are bargain hunters, or shopbot users. The simulations were asynchronous: at each time step, a buyer or seller was randomly selected to carry out an action (*e.g.*, buying an item or resetting a price). The chance that a given agent was selected for action was determined by its rate; the rate ρ_b at which a given buyer b attempts to purchase the good was set to 0.001, while the rate ρ_s at which a given seller reconsiders its price was 0.00002 (except where otherwise specified). Each simulation was iterated for 100 million time steps.

5.1 GT Pricebots

Simulations verify that, if agents are GT strategists, the cumulative distribution of prices closely resembles the derived $F(p)$ (to within statistical error), and moreover, the time-averaged profit for each seller is $\bar{\pi} = 0.0255 \pm 0.0003$, which is nearly the theoretical value of 0.0250. (Not shown.)

5.2 MY Pricebots

Fig. 2(a) illustrates cyclical price wars that typically occur when 5 pricebots use the myoptimal pricing strategy. Regardless of the initial value of the price vector, a pattern quickly emerges in which prices are positioned near the monopolistic price $v = 1$, followed by a long episode during which pricebots successively undercut one another by ϵ. During this latter phase, no two prices differ by more than $(S-1)\epsilon$, and the prices fall linearly with time. Eventually, when the lowest-priced seller is within ϵ above the value $p^* = 0.53125$, the next seller finds it unprofitable to undercut, and instead resets its price to $v = 1$. The other pricebots follow suit, until all but the lowest-priced seller are charging $v = 1$. At this point, the lowest-priced seller finds that it can maintain its market share but increase its profit dramatically — from $p^* - .5 = 0.03125$ to $0.5 - \epsilon$ — by raising its price to $1 - \epsilon$. No sooner than the lowest-priced seller raises its price does the next seller to reset its price undercut, thereby igniting the next cycle of the price war.

Fig. 2(b) shows the sellers' profits averaged during the intervals between successive resetting of prices. The upper curve represents a linear decrease in the average profit attained by the lowest-priced seller as price decreases, whichever seller that happens to be. The lower curve represents the average profit attained by sellers that are not currently the lowest-priced; near the end of the cycle they suffer from both low market share and low margin. The expected average profit can be computed by averaging the profit given by Eqs. 5 and 6 over one price-war cycle:

$$\pi_s^{\mathrm{MY}} = \frac{1}{S}\left[\frac{1}{2}(v + p^*) - c\right] \tag{10}$$

which yields $\pi_s^{\mathrm{MY}} = 0.053125$ in this instance. The simulation results match this closely: the average profit per time step is 0.0515, which is just over twice the average profit obtained via the game-theoretic pricing strategy.

Since prices fluctuate over time, it is of interest to compute the probability distribution of prices. Fig. 2(a) depicts the cumulative distribution function for myoptimal pricing. This measured cumulative density function has exactly the same endpoints $p^* = 0.53125$ and $v = 1$ as those of the theoretical mixed strategy equilibrium, but the linear shape between those endpoints (which reflects the linear price war) is quite different from what is displayed in Fig. 1(a).

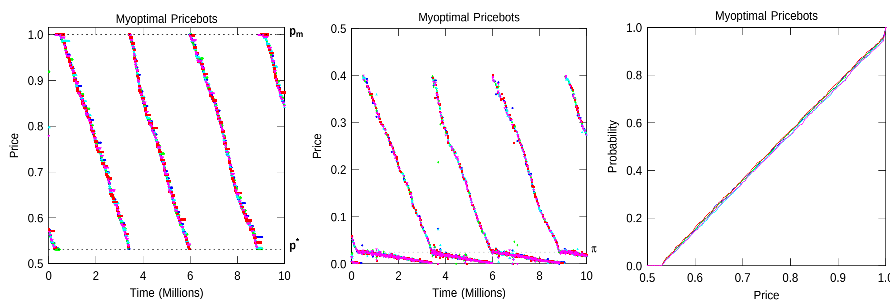

Fig. 2. (a) and (b) Price and profit dynamics, respectively, for 5 MY pricebots. (c) Cumulative distribution of prices observed between times 10 and 100 million.

5.3 DF Pricebots

Fig. 3(a) shows the price dynamics that result when 5 derivative followers are pitted against one another. Recall that derivative following pricebots do not base their pricing decisions on any information that pertains to other agents in the system — neither pricebots' price-setting tendencies nor buyers' preferences. Nonetheless, their behavior tends towards what is in effect a collusive state in which *all* pricebots charge nearly the monopolistic price.[10] This is tacit collusion as defined, for example, in Tirole [30], and so-called because the agents do not communicate at all so there is consequently nothing illegal about their collusive behavior. By exhibiting such behavior, derivative followers accumulate greater wealth than myoptimal or game-theoretic pricebots. According to Fig. 3(b), pricebots that are currently lowest-priced can expect an average profit of 0.30 to 0.35, while the others can expect roughly the game-theoretic profit of 0.025. Averaging over the last 90 million time steps (to eliminate transient effects), we find that the average profit per seller is 0.0841. This value is not far off from the absolute collusive limit of $(1/S)(v - c) = 0.10$.

[10] It has similarly been observed by Huck, *et al.* [23] that derivative followers tend towards collusive behavior in models of Cournot duopoly.

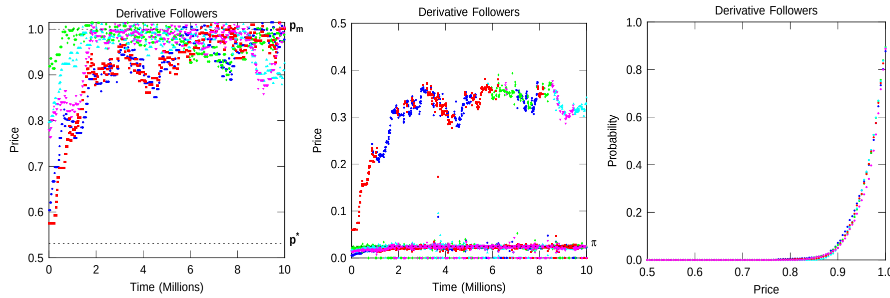

Fig. 3. (a) and (b) Price and profit dynamics, respectively, for 5 DF pricebots. (c) Cumulative distribution of prices observed between times 10 and 100 million.

How do derivative followers manage to collude? Like myoptimal pricebots, DFs are capable of engaging in price wars; such dynamics are visible in Fig. 3(a). These price wars, however, are easily quelled, making upward trends more likely than downward trends. Suppose X and Y are the two lowest-priced pricebots engaged in a mini-price war. Assume X's price is initially above Y's, but that X soon undercuts Y. This yields profits for seller X obtained from the entire population of type B buyers while it is lower-priced, and from its share of type A buyers all throughout. Now suppose Y undercuts X, but soon after X again undercuts Y. This yields profits for seller X once again obtained from the entire population of type B buyers during the period in which it is lower-priced, and from its share of type A buyers all throughout. In other words, given equivalent rates of price adjustment for both pricebots, market share remains fixed during mini-price wars of this kind. Thus, the only variable in computing profits is price, leaving pricebots with the incentive to increases prices more often than not. The tendency of a society of DF pricebots to reach and maintain high prices is reflected in the cumulative distribution function, shown in Fig. 3(c).

5.4 NR Pricebots

Finally, we present simulation results for two no regret pricing strategies, namely no external regret (NER) and no internal regret (NIR). As described in [16], there are a number of learning algorithms that satisfy the no external regret optimality criterion (*e.g.*, Foster and Vohra [11] and Freund and Schapire [14]); similarly, the no internal regret optimality criterion is satisfied by algorithms due to both Foster and Vohra [12] and Hart and Mas-Colell [22]. In this section, we discuss simulations of NER pricebots *á la* Freund and Schapire and NIR pricebots *á la* Foster and Vohra. Rather than consider 5 pricebots as above, we limit our attention to merely 2 NR pricebots, since the dynamics of 2 such pricebots converges more readily than does that of 5. As no regret algorithms are inherently non-deterministic as well as myopic, they are candidates for learning mixed strategy equilibria of stage games.

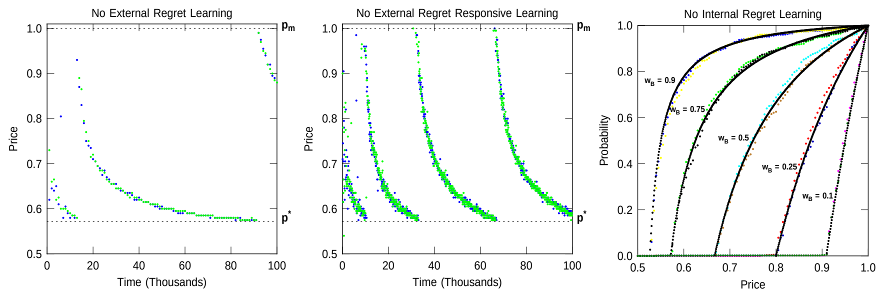

Fig. 4. (a) and (b) Price dynamics for 2 NER pricebots with learning rate $\beta = 0.1$, and $\gamma = 0$ and $\gamma = 0.0001$, respectively. (c) CDFs generated from simulation tape of 2 NIR pricebots with theoretical Nash equilibrium overlayed.

Although the no external regret algorithm of Freund and Schapire has been observed to converge to Nash equilibrium in games of 2 actions (*e.g.*, the Santa Fe bar problem, see Greenwald, *et al.* [18]), NER pricebots cycle exponentially from price p_m to price p^* in the prescribed model, which entertains $n > 2$ possible actions[11] (see Fig. 4(a)). In fact, the outcome of play of NER pricebots in the shopbot game is reminiscent of the outcome of both NER learning and fictitious play in the Shapley game, a game of 3 strategies for which there is no pure strategy Nash equilibrium (see Greenwald, *et al.* [17]). Fig. 4(b) depicts simulations of NER pricebots in which we introduce a responsiveness parameter $0 < \gamma \leq 1$ that exponentially weights the observed history, effectively limiting its length to the finite value $1/\gamma$, thereby allowing NER pricebots to more readily respond to environmental changes. Interestingly, the empirical frequency of play during these finite cycles mimics the symmetric mixed strategy Nash equilibrium of the stage game; this result will be further explored in a future publication.

We now investigate the properties of no internal regret (NIR) learning in our model of shopbots and pricebots. Learning algorithms that satisfy the no internal regret optimality criteria are known to converge to correlated equilibria [2, 12], a superset of the set of Nash equilibria which allows for dependence among players' strategies. Despite several negative theoretical results on the rational learning of Nash equilibria (*e.g.*, Foster and Young [13], Nachbar [27], and Greenwald *et al.* [18]), in practice, no internal regret learning — a form of boundedly rational learning — seems to learn Nash equilibria. In our present simulations we observe convergence to the symmetric mixed strategy Nash equilibrium. Fig. 4(c) depicts the cumulative distribution functions generated from simulation tapes of the no internal regret learning algorithm designed by Foster and Vohra [12]. In these simulations, we consider 2 NIR pricebots, and we let buyer distributions range from $(w_A, w_B) = (0.1, 0.9)$ to $(w_A, w_B) = (0.9, 0.1)$. These plots closely match the theoretical Nash equilibria given 2 sellers, which are overlayed in this figure.

[11] Technically, there is a continuum of prices in our model. For the purpose of simulating no regret learning, this continuum was discretized into 100 equally sized intervals.

6 Evolution of Shopbots and Pricebots

We now revisit the situation in which all five sellers utilize the myoptimal pricing strategy, but we allow one of the sellers to reset its price five times as quickly as all the others. These price dynamics are illustrated in Fig. 5(a). As in Fig. 2(a), we experience price wars; in this case, however, they are accelerated, which is apparent from the increase in the number of cycles that occur during the simulation run. From the individual profit curves, which in Fig. 5(b) depict cumulative profits rather than instantaneous profits, we notice that the fast myoptimal agent obtains substantially more profit than the others because it undercuts far more often than it itself is undercut. Analysis yields that the expected profit for a given myoptimal seller s who resets its prices at rate ρ_s, assuming all other sellers are myoptimal, is as follows:

$$\pi_s^{\mathrm{my}}(p_s) = \left[\frac{1}{S}w_A + \frac{\rho_s}{\sum_{s'}\rho_{s'}}w_B\right]\left[\frac{1}{2}(v+p^*) - c\right] \tag{11}$$

Eq. 11 predicts average profits of $119/960 = 0.1240$ for the fast seller and $34/960 = 0.0354$ for the slower ones, values which compare reasonably well with those obtained by averaging over the last 10 complete cycles of the price war, namely 0.1111 and 0.0351, respectively.

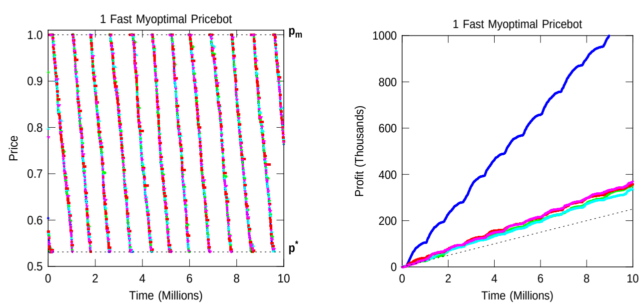

Fig. 5. (a) and (b) Price and profit dynamics, respectively, with 1 Fast MY pricebot.

Evidently, myoptimal pricebots stand to gain by resetting their prices at faster, rather than slower, rates, particularly when a large proportion of the buyer population is shopbot-assisted (see Eq. 11). If MY pricebots were to reprice their goods with ever-increasing frequency, a sort of arms race would develop, leading to arbitrarily fast price-war cycles. This observation is not specific to myoptimal agents. In additional simulations, we have observed sufficiently fast DF pricebots who obtain the upper hand over slower derivative following and myoptimal agents. In the absence of any throttling mechanism, it is advantageous for pricebots to re-price their goods as quickly as possible.

Let us carry the arms race scenario a bit further. In a world in which sellers are inclined to reset prices at ever-increasing rates, human price setters would undoubtedly be too inefficient. Sellers, therefore, would necessarily come to rely on pricebots, perhaps sophisticated variants of one of the strategies proposed in Sec. 4. Quite possibly, pricebot strategies would utilize information about the buyer population, which could be purchased from other agents. Even more likely, pricebot strategies would require knowledge of their competitors' prices. How would up-to-date information be obtained? From shopbots, of course!

With each seller seeking to re-price its products faster than its competitors, shopbots would quickly become overloaded with requests. Imagine a scenario in which a large player like `amazon.com` were to use the following simple price-setting strategy: every 10 minutes, submit 2 million or so queries to one or more shopbots (one for each title carried by `Amazon.com`), then charge 1 cent less than the minimum price returned on each title! Since the job of shopbots is to query individual sellers for prices, it would in turn pass this load back to `Amazon.com`'s competitors. The rate of pricing requests made by sellers could easily dwarf the rate at which similar requests would be made by human buyers, thereby eliminating the potential of shopbots to ameliorate market frictions.

A typical solution to an excess demand for shopbot services would be for shopbots to charge pricebots for price information. Today, shopbots tend to make a living by selling advertising space on their Web pages. This appears to be an adequate business model so long as requests are made by humans. Agents, however, are unwelcome customers because they are are not influenced by advertisements; as a result, agents are either barely tolerated or excluded intentionally. By charging for the information services they provide, shopbots would be economically-motivated agents, creating the proper incentives to deter excess demand, and welcoming business from other agents.

Once shopbots begin to charge for pricing information, it would seem natural for sellers — the actual owners of the desired information — to themselves charge shopbots for price information. In fact, sellers could use another form of pricebot to dynamically price this information. This scenario illustrates how the need for agents to dynamically price their services could quickly percolate through an entire economy of software agents. The alternative is "meltdown" due to overload which could occur as agents become more prevalent on the Internet. Rules of etiquette followed voluntarily today by web crawlers and related programs [9] could be trampled in the rush for competitive advantage.

7 Conclusion

Game-theoretic analysis of a model of a simple commodity market established a quantitative relationship between the degree of shopbot usage among buyers and the degree of price competition among sellers. This motivated a comparative study of various *pricebot* algorithms that sellers might employ in an effort to gain an edge in a market in which the presence of shopbots has increased the degree of competition.

MY pricebots were shown to be capable of inducing price wars, yet even so they earn profits that are well above those of GT strategists. DF pricebots were observed to exhibit tacit collusion, leading to cartel-level profits. Finally, game-theoretic equilibria arose dynamically as the outcome of repeated play among certain NR pricebots. In related work (see [20]), we explore the dynamics of prices and profits among pricebots that use non-myopic learning algorithms, such as Q-learning, and we directly compare the profitability of various pricing strategies by simulating heterogeneous collections of pricebots. In future work, we intend to study the dynamics of markets in which more sophisticated shopbots base their search on product attributes as well as price, and in which pricebot strategies are extended accordingly.

Acknowledgments

The authors are grateful to Fredrik Ygge for his careful reading of an earlier version of this paper, in which he isolated some serious flaws in the original analysis. Any remaining errors are surely the fault of the authors. In addition, the authors thank Adam Kalai for discovering certain pure strategy Nash equilibria in the discretized model, which inspired the analysis presented in Appendix A.

References

1. P. Auer, N. Cesa-Bianchi, Y. Freund, and R. Schapire. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *Proceedings of the 36th Annual Symposium on Foundations of Computer Science*, pages 322–331. ACM Press, November 1995.
2. R. Aumann. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1:67–96, 1974.
3. A. Banos. On pseudo games. *The Annals of Mathematical Statistics*, 39:1932–1945, 1968.
4. D. Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6:1–8, 1956.
5. K. Burdett and K. L. Judd. Equilibrium price dispersion. *Econometrica*, 51(4):955–969, July 1983.
6. Anthony Chavez and Pattie Maes. Kasbah: an agent marketplace for buying and selling goods. In *Proceedings of the First International Conference on the Practical Application of Intelligent Agents and Multi-Agent Technology*, London, U.K., April 1996.
7. A. Cournot. *Recherches sur les Principes Mathematics de la Theorie de la Richesse*. Hachette, 1838.
8. J. Bradford DeLong and A. Michael Froomkin. The next economy? In Deborah Hurley, Brian Kahin, and Hal Varian, editors, *Internet Publishing and Beyond: The Economics of Digital Information and Intellecutal Property*. MIT Press, Cambridge, Massachusetts, 1998.
9. David Eichmann. Ethical web agents. In *Proceedings of the Second World Wide Web Conference '94: Mosaic and the Web*, 1994.

10. Joakim Eriksson, Niclas Finne, and Sverker Janson. Information and interaction in MarketSpace — towards an open agent-based market infrastructure. In *Proceedings of the Second USENIX Workshop on Electronic Commerce*, November 1996.
11. D. Foster and R. Vohra. A randomization rule for selecting forecasts. *Operations Research*, 41(4):704–709, 1993.
12. D. Foster and R. Vohra. Regret in the on-line decision problem. *Games and Economic Behavior*, 21:40–55, 1997.
13. D. Foster and P. Young. When rational learning fails. Mimeo, 1998.
14. Y. Freund and R. Schapire. Game theory, on-line prediction, and boosting. In *Proceedings of the 9th Annual Conference on Computational Learning Theory*, pages 325–332. ACM Press, May 1996.
15. D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, Cambridge, 1991.
16. A. Greenwald. *Learning to Play Network Games*. PhD thesis, Courant Institute of Mathematical Sciences, New York University, New York, May 1999.
17. A. Greenwald, E. Friedman, and S. Shenker. Learning in network contexts: Results from experimental simulations. *Games and Economic Behavior: Special Issue on Economics and Artificial Intelligence*, Forthcoming 1999.
18. A. Greenwald, B. Mishra, and R. Parikh. The Santa Fe bar problem revisited: Theoretical and practical implications. Presented at *Stonybrook Festival on Game Theory: Interactive Dynamics and Learning*, July 1998.
19. A.R. Greenwald and J.O. Kephart. Shopbots and pricebots. In *Proceedings of Sixteenth International Joint Conference on Artificial Intelligence*, volume 1, pages 506–511, August 1999.
20. A.R. Greenwald, J.O. Kephart, and G.J. Tesauro. Strategic pricebot dynamics. In *Proceedings of First ACM Conference on Electronic Commerce*, To Appear, November 1999.
21. J. Hannan. Approximation to Bayes risk in repeated plays. In M. Dresher, A.W. Tucker, and P. Wolfe, editors, *Contributions to the Theory of Games*, volume 3, pages 97–139. Princeton University Press, 1957.
22. S. Hart and A. Mas Colell. A simple adaptive procedure leading to correlated equilibrium. Technical report, Center for Rationality and Interactive Decision Theory, 1997.
23. S. Huck, H-T. Normann, and Jörg Oechssler. Learning in a cournot duopoly – an experiment. Unpublished Manuscript, July 1997.
24. J. O. Kephart, J. E. Hanson, D. W. Levine, B. N. Grosof, J. Sairamesh, R. B. Segal, and S. R. White. Dynamics of an information filtering economy. In *Proceedings of the Second International Workshop on Cooperative Information Agents*, 1998.
25. J.O. Kephart and A.R. Greenwald. Shopbot economics. In *Proceedings of Fifth European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pages 208–220, July 1999.
26. B. Krulwich. The BargainFinder agent: Comparison price shopping on the Internet. In J. Williams, editor, *Agents, Bots and Other Internet Beasties*, pages 257–263. SAMS.NET publishing (MacMillan), 1996. URLs: http://bf.cstar.ac.com/bf, http://www.geocities.com/ResearchTriangle/9430.
27. J. Nachbar. Prediction, optimization, and learning in repeated games. *Econometrica*, 65:275–309, 1997.
28. J. Nash. Non-cooperative games. *Annals of Mathematics*, 54:286–295, 1951.
29. G.J. Tesauro and J.O. Kephart. Pricing in agent economies using multi-agent q-learning. In *Proceedings of Fifth European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pages 71–86, July 1999.

30. Jean Tirole. *The Theory of Industrial Organization*. The MIT Press, Cambridge, MA, 1988.
31. M. Tsvetovatyy, M. Gini, B. Mobasher, and Z. Wieckowski. MAGMA: an agent-based virtual market for electronic commerce. *Applied Artificial Intelligence*, 1997.
32. H. Varian. A model of sales. *American Economic Review, Papers and Proceedings*, 70(4):651–659, September 1980.
33. L. L. Wilde and A. Schwartz. Comparison shopping as a simultaneous move game. *Economic Journal*, 102:562–569, 1992.

A Pure Strategy Nash Equilibria

This appendix revisits the existence of pure strategy Nash equilibria (PNE) in the prescribed model of shopbots and pricebots whenever $0 < w_A, w_B < 1$. It has previously been established (see Sec. 3) that no PNE exist when prices are selected from a continuous strategy space. Here, we assume that prices are chosen from a strategy space that is discrete rather than continuous, and we derive the set of pure strategy Nash equilibria. This set is symmetric in the case of 2 sellers, but is often asymmetric in the case of $S > 2$ sellers.

Recall from Sec. 2 that the profits for seller s are determined as follows, assuming $v_b = v$ for all buyers b, and $c_s = c$ for all sellers s:

$$\pi_s(\boldsymbol{p}) = \begin{cases} (p_s - c)h_s(\boldsymbol{p}) & \text{if } p_s \leq v \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

where

$$h_s(\boldsymbol{p}) = w_A \frac{1}{S} + w_B \frac{1}{\tau_s(\boldsymbol{p}) + 1} \, \delta_{\lambda_s(\boldsymbol{p}),0} \tag{13}$$

- δ is the Kronecker δ function,
- $\lambda_s(\boldsymbol{p})$ is the number of sellers charging a lower price than s, and
- $\tau_s(\boldsymbol{p})$ is the number of sellers charging the same price as s, excluding s itself.

The equilibrium derivation that follows concerns the case of discrete strategy spaces, characterized by some parameter $\epsilon > 0$, which dictates the sellers' space of strategies as follows: $|P| = \{0, \epsilon, 2\epsilon, \ldots\}$. If we assume $c \bmod \epsilon = 0$, then this strategy space contains prices of the form $p_i = c \pm i\epsilon$, where $i \in \{x \in \mathbb{Z} | x \geq c/\epsilon\}$. For convenience, we further assume $v \bmod \epsilon = 0$.[12]

The derivation of the set of pure strategy Nash equilibria is based on the following observations, which dictate the structure of its elements: at equilibrium,

1. No seller charges price $p_i > v$.
2. No seller charges price $p_i \leq c$.
3. At least two sellers charge prices $c < p_i < v$.
4. Those sellers who charge prices $c < p_i < v$ charge equal prices.

[12] Otherwise, v is everywhere replaced by $v' = c + \epsilon \lfloor \frac{v-c}{\epsilon} \rfloor$ in the discussion that follows.

The first two observations follow from the fact that the profits obtained by charging the monopoly price v are strictly positive, whereas the profits obtained by charging either $p_i > v$ or $p_i \leq c$ are zero. At least two sellers charge $c < p_i < v$ since (i) if all sellers were to charge v, seller s would stand to gain by instead charging $v - \epsilon$ (assuming $\epsilon < v - c$) and (ii) if only one seller were to charge $c < p_i < v$, then p_i must equal $v - \epsilon$, in which case the other sellers would stand to gain by charging $v - 2\epsilon$ (assuming $2\epsilon < v - c$). Finally, if seller s' were to charge $c < p'_i < v$, while seller s were charging $c < p_i < p'_i < v$, then seller s' would prefer price v to price p'_i, implying that p'_i is not an equilibrium price. Therefore, PNE are structured such that $n \geq 2$ sellers charge p_i for $0 < i < (v - c)/\epsilon$, while the remaining $S - n$ sellers charge the monopoly price v (unless $\epsilon \geq v - c$, in which case PNE exist of the form (v, v), or $\epsilon \geq (v - c)/2$, in which case PNE exist of the form $(v - \epsilon, v - \epsilon)$).

Given the prescribed structure, the existence of pure strategy Nash equilibria is ensured whenever the following conditions are satisfied: for all sellers s,

1. No low-priced seller charging $c < p_i < v$ prefers the monopoly price v: *i.e.*, $\pi_s(p_i) \geq \pi_s(v)$, where $\pi_s(p_i)$ is computed assuming p_i is charged by n low-priced sellers. Expanding this condition leads to the following:
$$i\epsilon \geq \frac{w_A(v - c)}{w_A + \frac{1}{n} w_B S} \tag{14}$$
This condition implies that $p_i = c + i\epsilon > p^*$, since $n \geq 2$.[13]

2. No seller charging v prefers to undercut the low-priced sellers charging p_i and charge p_{i-1}: *i.e.*, $\pi_s(v) \geq \pi_s(p_{i-1})$, where $\pi_s(p_{i-1})$ are the profits obtained if seller s is the unique, lowest-priced seller. Expanding this condition yields:
$$\frac{w_A(v - c)}{w_A + w_B S} \geq (i - 1)\epsilon \tag{15}$$
For $i = 1$ this condition reduces to $\pi_s(v) \geq \pi_s(c) = 0$, which is tautological; hence this constraint is only of interest when $i > 1$.

3. No low-priced seller charging p_i prefers to undercut its cohorts by charging p_{i-1}: *i.e.*, $\pi_s(p_i) \geq \pi_s(p_{i-1})$, which incidentally is implied by Conds. 14 and 15, so long as some seller charges v. This yields a constraint on the value of i (or stated otherwise, n) for which PNE exist, namely:
$$\frac{i}{i - 1} \geq \frac{w_A + w_B S}{w_A + \frac{1}{n} w_B S} \tag{16}$$
Like the previous condition, this constraint is only applicable when $i > 1$.

[13] The value of p^* derived in Eq. 9 for the continuous case is applicable in the discrete case, unless $v \bmod \epsilon \neq 0$, in which case v is replaced by v' in Eq. 9 (see Foot. 12).

Together Conds. 14, 15, and 16 are mathematical statements of the conditions for the existence of pure strategy Nash equilibria of the prescribed structure.

We now construct a series of examples, assuming production cost is $c = 0.5$, buyers have constant valuations $v = 1$, and $w_A = 0.25$ and $w_B = 0.75$. Initially, we consider only 2 sellers. By the prescribed structure of PNE, both sellers charge equal prices $c < p_i < v$, for some $0 < i < (v-c)/\epsilon$, assuming $\epsilon < (v-c)/2 = 0.25$. Since no seller charges v, Cond. 15 is not a relevant constraint. Cond. 16 requires that $\frac{i}{i-1} \geq w_A + 2w_B = 1.75$, which is impossible for integer values of $i > 2$. Thus, our interest is confined to values of $i \leq 2$ and $\epsilon < 0.25$ satisfying Cond. 14. In particular, if $i = 1$, then PNE exist whenever $0.25 > \epsilon \geq w_A(v - c) = 0.125$; if $i = 2$, then PNE exist whenever $0.25 > \epsilon \geq w_A(v - c) = 0.0625$. The complete set of PNE for $S = 2$ is listed in Table A. Notice that PNE cease to exist when $|P| > 9$; for $S = 3$, PNE cease to exist when $|P| > 12$; in general, PNE cease to exist whenever $|P| > \lfloor 1 + i^*[1 + (w_B/w_A)(S/n)] \rfloor$ where i^* is the maximum integer value i satisfying Cond. 16, which can be rearranged to give an upper bound on i.

$\|P\|$	$\epsilon = \frac{v-c}{\|P\|-1}$	i	PNE
1	∞	0	(1.0, 1.0)
2	0.5	1	(1.0, 1.0)
3	0.25	1	(0.75, 0.75)
4	$0.1\overline{6}$	1,2	$(0.\overline{6}, 0.\overline{6})$, $(0.8\overline{3}, 0.8\overline{3})$
5	0.125	1,2	(0.625, 0.625), (0.75, 0.75)
6	0.1	2	(0.7, 0.7)
7	$0.08\overline{3}$	2	$(0.\overline{6}, 0.\overline{6})$
8	0.0714	2	(0.643, 0.643)
9	0.0625	2	(0.625, 0.625)
10	$0.0\overline{5}$	-	DNE

Table 1. The set of PNE for $S = 2$. DNE stands for does not exist, implying the non-existence of pure strategy Nash equilibria, although the existence of mixed strategy equilibria is established in Nash [28].

Now consider a larger number of sellers; for concreteness, say $S = 100$. We first let $i = 1$, which limits our concern to Cond. 14. It follows from this condition that when the number of sellers is large, PNE exist even for small values of ϵ so long as n is also small. In particular, if $n = 2$ then PNE exist for $\epsilon \geq 0.003311$; specifically, if $\epsilon = 0.00\overline{3}$, then an asymmetric solution arises in which sellers who charge price p_1 earn profits of roughly 0.00126, while sellers who charge price v earn 0.00125. At the other extreme, if $n = 100$, then symmetric PNE exist iff $\epsilon \geq 0.125$. A full range of asymmetric PNE exist when $i = 1$ for the values of ϵ specified by Cond. 14 that arise for values of n ranging from 2 to 100.

Still assuming a large number of sellers, let $i > 1$. Restating Cond. 16 as a bound on n and taking the limit as $S \to \infty$, we find that $n \leq i/(i-1) \leq 2$. But since $n \geq 2$ at equilibrium, it follows that at any PNE exactly 2 sellers charge price p_i. Again rewriting Cond. 16, this time as a bound on i and then taking the limit as $S \to \infty$, we also find it necessary that $i \leq n/(n-1) \leq 2$. Thus, for sufficiently large numbers of sellers, PNE exist in which exactly 2 low-priced sellers charge price p_2, but no PNE exist in which any sellers charge p_i for $i > 2$.

It is nonetheless possible for equilibria to arise in which $i > 2$, however not for the assignments of w_A and w_B assumed throughout our examples. Consider instead $w_A = 0.75$ and $w_B = 0.25$. Now for $S = 2$, an equilibrium arises in which $n = 2$, $i = 5$, and $\epsilon = 0.08\overline{3}$, namely $(0.91\overline{6}, 0.91\overline{6})$. Using Cond. 16, we note that as $w_A \to 1$, i is bounded above only by $(v-c)/\epsilon$; in other words, high equilibrium prices prevail. On the other hand, as $w_B \to 1$, n is bounded above only by S, implying that more and more sellers prefer to charge low prices. Finally, through simulations we have observed pure strategy Nash equilibria to be the outcome of myoptimal pricing and no internal regret learning in the discretized model of shopbots and pricebots.

B No Regret Learning

This appendix describes the no regret learning algorithms which are simulated in Sec. 5.4. There are two no regret criteria of interest, namely no external regret and no internal regret. Computational learning theorists consider the difference between the expected payoffs that are achieved by the strategies prescribed by a given algorithm, as compared to the payoffs that could be achieved by any other fixed sequence of decisions, in the worst-case. If the difference between these two sums is negligible, then the algorithm exhibits *no external regret*. Early no external regret algorithms appeared in Blackwell [4], Hannan [21], and Banos [3]. Game theorists Foster and Vohra [12] study an alternative measure of worst-case performance. If the difference between the cumulative payoffs that are achieved by a sequence of strategies generated by a given algorithm in comparison with the cumulative payoffs that could be achieved by a remapped sequence of strategies is insignificant, then the algorithm is said to exhibit *no internal regret*.[14] No internal regret implies no external regret.

The no regret algorithms are presented from the point of view of an individual player, as if that player were playing a game against nature, where nature is taken to be a conglomeration of all its opponents. From this perspective, let r_i^t denote the payoffs obtained by the player of interest at time t via strategy i. Mixed strategy weights at time t are given by the probability vector (w_i^t), for $1 \leq i \leq S$, where S is the number of strategies.

[14] A sequence is remapped if there is a mapping f of the strategy space into itself $s.t.$ for each occurrence of strategy s_i in the original sequence, the mapped strategy $f(s_i)$ appears in the remapped sequence.

B.1 No External Regret Learning

Freund and Schapire [14] derive an algorithm that achieves no external regret via multiplicative updating. Their algorithm is dependent on the cumulative payoffs achieved by all strategies, including the surmised payoffs of strategies which are not played. In particular, let ρ_i^t denote the cumulative payoffs obtained through time t via strategy i, which is computed as follows: $\rho_i^t = \sum_{x=0}^t r_i^x$. Now the weight assigned to strategy i at time $t+1$, for $\beta > 0$, is given by:

$$w_i^{t+1} = \frac{(1+\beta)^{\rho_i^t}}{\sum_{j=1}^S (1+\beta)^{\rho_j^t}} \tag{17}$$

The multiplicative updating rule given in Equation 17 can be modified to become applicable in naive settings, where complete payoff information is not available, but rather the only payoff information known at time t is that which pertains to the strategy which was in fact employed at time t. Such a variant of this multiplicative updating algorithm appears in Auer, Cesa-Bianchi, Freund, and Schapire [1]. It remains to perform simulations of this naive algorithm in our model of shopbots and pricebots.

B.2 No Internal Regret Learning

We now describe an algorithm due to Foster and Vohra [12] which achieves no internal regret, and a simple implementation due to Hart and Mas-Colell [22]. Learning via the following no internal regret algorithms converges to correlated equilibrium [2, 12].

The regret felt by a player at time t is formulated as the difference between the payoffs obtained by utilizing strategy the player's strategy of choice, say i, and the payoffs that could have been achieved had strategy $j \neq i$ been played instead:

$$\mathrm{R}_{i \to j}^t = \mathbf{1}_i^t (r_j^t - r_i^t) \tag{18}$$

where $\mathbf{1}_i^t$ is the indicator function, which has value 1 if strategy i is employed at time t, and has value 0 otherwise. Now the average regret $\mathrm{R}_{i \to j}^T$ is the summation of regrets from i to j through time T divided by T:

$$\mathrm{R}_{i \to j}^T = \frac{1}{T} \sum_{t=0}^T \mathrm{R}_{i \to j}^t \tag{19}$$

Internal regret is defined as follows:

$$\mathrm{IR}_{i \to j}^T = (\mathrm{R}_{i \to j}^T)^+ \tag{20}$$

where $X^+ = max\{X, 0\}$. Finally, the total internal regret for playing all other strategies but for not having played strategy j throughout the course of a game is given by:

$$\mathrm{IR}_{S \to j}^T = \sum_{i=1}^S \mathrm{IR}_{i \to j}^T \tag{21}$$

Given the above definitions, consider the case of a 2-strategy informed game, with strategies X and Y. The no internal regret learning algorithm updates the components of the weight vector, namely w_X^{t+1} and w_Y^{t+1}, according to the following formulae, which reflect cumulative feelings of regret:

$$w_X^{t+1} = \frac{\mathrm{IR}_{Y \to X}^t}{\mathrm{IR}_{X \to Y}^t + \mathrm{IR}_{Y \to X}^t} \quad \text{and} \quad w_Y^{t+1} = \frac{\mathrm{IR}_{X \to Y}^t}{\mathrm{IR}_{X \to Y}^t + \mathrm{IR}_{Y \to X}^t} \tag{22}$$

If the regret for having played strategy j rather than strategy i is significant, then the algorithm updates weights such that the probability of playing strategy i is increased. In general, if strategy i is played at time t,

$$w_j^{t+1} = \frac{1}{\mu} \mathrm{IR}_{i \to j}^t \quad \text{and} \quad w_i^{t+1} = 1 - \sum_{j \neq i} w_j^{t+1} \tag{23}$$

where μ is a normalizing term that is chosen *s.t.*:

$$\mu > (|S| - 1) \max_{j \in S} \mathrm{IR}_{i \to j}^t \tag{24}$$

This version of the generalized algorithm is due to Hart and Mas-Colell [22].

Like the no external regret algorithm of Freund and Schapire [14], the above no internal regret algorithm depends on complete payoff information at all times t, including information that pertains to strategies that were not employed at time t. The no internal regret learning algorithm has also been studied in naive settings, where complete payoff information is not available (see Foster and Vohra [11] and Greenwald [16]). It remains to simulate the naive variant of the no internal regret learning algorithm in our model of shopbots and pricebots.