Breeding Better Strategies in Oligopolistic Price Wars

Robert Marks, *
Australian Graduate School of Management,
UNSW Sydney NSW 2052, Australia
Tel: +61 2 9931 9271
Fax: +61 2 9931 9326
bobm@agsm.edu.au

David Midgley,
INSEAD,
France
Tel: +33 1 6072 4248
david.midgley@insead.fr

Lee Cooper,
Anderson School, UCLA,
Los Angeles, California.
Tel: +1 310 825 4488
Fax: +1 310 206 7422
lcooper@agsm.ucla.edu

* Corresponding author.

ABSTRACT: Using empirical market data from brand rivalry in a retail ground-coffee market, we model each brand's pricing behaviour using the Markov restriction, that marketing strategies depend only on profit-relevant state variables, and use the Genetic Algorithm to search for co-evolved Markov perfect equilibria, where each profit-maximizing brand manager is a stimulus-response automaton, responding to past prices in the oligopolistic market. Part of a growing study of repeated interactions and oligopolistic behaviour using the GA.

ACKNOWLEDGEMENTS: Research support was provided by the Australian Graduate School of Management and the Australian Research Council. Earlier versions of this paper were presented at the Third International Conference on Computing in Economics and Finance, Stanford, June 1997; the Computational Modeling Workshop, U.C.L.A. Anderson School & Division of Social Science, Los Angeles, March 1999; The Genetic and Evolutionary Computation Conference, Orlando, July 1999; and The Industry Economics Conference, Sydney, July 2000.

1. Oligopoly Theory

Oligopolistic behaviour continues to be a challenge to economics theorists. From Friedman (1977) and earlier, oligopolies have been analysed as repeated games, in a literature which has attempted to explain price wars as the results of uncertain demand or other reasons (Green & Porter 1984, Rotemberg & Saloner 1986). A second approach was developed by Maskin & Tirole (1988a, 1988b), that of Markovian strategies. As discussed by Dutta & Sundaram (1998), Markovian games posit the existence of a "state" variable that is designed to capture the environment of the game at each point in time, but that changes through time in response to the actions taken by the players in the game.

The current action of a player could affect his future rewards in two ways: first, the effect the action has on the environment in which future decisions must be made, and, second, its impact on the behaviour of other players in the model. As Dutta & Sundaram (1998) argue, Markovian games thus extend dynamic programming problems (which include only the first effect) and repeated games (which include only the second): Markovian games can include both.

Modelling a duopoly as a differentiated Bertrand game in which players take turns in choosing their price for two periods (short-term commitment), and in which each player's price in any period depends only on the other player's current price, Maskin & Tirole (1988b) derive two kinds of equilibrium: an Edgeworth cycle and a kinked demand curve. It is the former we are concerned with here: it can be characterised as behaviour in which firms successively undercut each other in order to increase their market share (the price-war phase) until the war becomes too costly and one or other of the firms eventually relents by jumping its price up, followed by the other (the relenting phase), until after some time at the high, collusive price the price-war phase begins again.

Such cyclical behaviour is not uncommon (see Slade 1992), and had been examined in theory before Maskin & Tirole, but previously with capacity constraints, which Maskin & Tirole dispensed with in their derivation of Edgeworth cycles as Markovian (or Markov perfect) equilibria (MPE) in a Bertrand duopoly. Similar patterns are observed in the retail coffee data we use as the basis for our empirical work below.[1]

A Markov perfect equilibrium in a general oligopoly can be characterised as follows: starting from any point in the game tree, a player or firm selects the action that maximises its intertemporal profit (whether discounted in an infinite game or limits of means or averaged in a finite game), given the subsequent moves of itself and its rivals. That is, the term "Markov perfect" arises from the simple observation that all Nash equilibria

1. Recent work by Peltzman (2000) does not support the asymmetries of the Edgeworth cycle for supermarket prices in response to cost shocks. But we are not modelling this linkage: rather the pricing rivalry among oligopolists in a differentiated Bertrand oligopoly.

in Markovian situations are also subgame-perfect.

Maskin & Tirole (1988b), by making the Markovian assumption, consider only those equilibria whose strategies depend on the "payoff-relevant" history, in their case, only those actions which enter directly into the one-period profit function. In the stage game, only the actions of the players in that period are payoff relevant, and since the other player's price is known (since they take turns in committing to a price for two periods) each player determines its price as a function of the other player's current price. Moreover, since the stage game is unchanging, time is not a payoff-relevant variable. In their case, each player's state is simply the price chosen by the other player in the previous round. When Maskin & Tirole model a simultaneous-move game, the state variable is null since neither of the prices in the stage-game duopoly is known, and the players use mixed strategies. Clearly, just what is payoff relevant depends on the definition of the payoff in any game, as we explore below.

Maskin & Tirole argue that the Markovian assumption (of including only payoff-relevant variables as state variables) can be justified by two reasons. First, simplicity: firms' strategies depend on as little as possible while still remaining consistent with rationality. Second, the concept, they argue, is closer to the Industrial Organisation conception of reactions by firms not to fulfill some earlier threat, or reacting to earlier history of one's own prices and others', but simply by cutting one's price (during the price-war phase of the Edgeworth cycle) to maintain market share in response to others' cuts.

2. Modelling a Price-Competitive Oligopoly

Maskin & Tirole's (1988b) definition of state in the simultaneous-move differentiated Bertrand oligopoly is too severe. By restricting the payoff-relevant variables to the other player's (already known) price (which, together with one's own chosen price, determines the stage-game payoff), they preclude the modelling of simultaneous-move games as Markovian games. Their discussion of the Edgeworth-cycle equilibrium provides an escape.

In a discussion of game-playing automata, Binmore (1992) remarks that two or more deterministic finite automata (stimulus-response machines, where the stimulus is the state of the market, and the response is player's action) must, by definition, cycle through the same sequence of action combinations, or states. Each player can be assumed to evaluate the income stream it obtains by taking the average of the payoffs it receives during the cycle. Since evaluating in this way is equivalent to using a value function defined by

$$V_i(a) = \lim_{N \to \infty} \sum_{n=1}^{N} \pi_i(a)$$

for player i, it is often referred to as the limits-of-means criterion. Variable V_i is i's payoff value, a is the combination of all players' actions in any period n, and there are N periods.

In the case of cycling, then, the payoff-relevant variables are all players' actions during the cycle. So the state in a Markovian game with cycling should include all previous actions of the players since the period in which the game or market was last in the state attained last round.

There is another reason to include the game's history in the definition of its state. Although the history of play may not alter the stage game directly (although it is possible that, for instance, demand saturation with a durable good may result in past purchases lowering present demand, see Midgley et al. 1997), the past may still affect the outcome of the stage game if one or more players believe that history matters (the un-Fordian view). This is the view of Fudenberg & Tirole (1992) for deterministic games of complete information, where player i has the value function

$$V_i(a^0, a^1, \cdots, a^T) = V_i(h^t, f^t),$$

where a^k is the combination of all players' actions at time k, h^t is the history $(a^0, \ldots, a^{t-1})$, and f^t is the future $(a^t, \cdots, a^T)$ as foreseen at time t.

It is possible that a Markov strategy for player i might be conditioned on less than all player i's information. There are two ways in which the information may be partitioned. First, as discussed by Fudenberg & Tirole (1992), the history may be partitioned into $\{H^t(h^t)\}_{t=0, \cdots, T}$, which, for each date, are mappings from the set of histories into a set of disjoint and exhaustive subsets of the set of possible histories at that date.

Second, as discussed by Marks (1998), the actions at any date may be partitioned in the price space, as, say {"low", "low", "medium"}.

While summarising the history, a partition must not be too coarse. That is, at each date, the players must be able to recover the strategic elements of the following subgame from the element of the partition to which the history belongs. Fudenberg & Tirole's (1992) Definition 13.3 defines a *sufficient* partition, which leads to the definition: the *payoff-relevant history* is the minimal (i.e., coarsest) sufficient partition (their Def. 13.4). For cyclical games, Markov strategies are independent of calendar time, modulo the location within the cycle.

The above provides a foundation for the empirical work described below. In essence, we use historical data to search for two parameters of the model: first, for sufficient partitions of the history (just what is the coarsest partitioning — and so the minimum number of states per player — of past actions sufficient to result in the historical data?), and, second, for Markov perfect equilibria (MPE) by searching for mappings from the state (of history) to each players' actions.

3. Empirical work

Axelrod (1987) modelled players in his discrete repeated Prisoner's Dilemma (RPD) game as stimulus-response automata, where the stimulus was the state of the game, defined as both players' actions over the previous several moves, and the response was the next period's action (or actions). That is, he modelled the game as a state-space game (Fudenberg & Tirole 1992, Slade 1995), in which past play influences current and future actions, not because it has a direct effect on the game environment (the payoff function) but because all (or both) players believe that past play matters. Axelrod's model focused attention on a smaller class of "Markov" or "state-space" strategies, in which past actions influence current play only through their effect on a state variable that summarises the direct effect of the past on the current environment (the payoffs). With state-space games, the state summarises all history that is payoff-relevant, and players' strategies are restricted to depend only on the state and (perhaps) the time.

Specifically, Axelrod's stimulus-response players were modelled as strings, each point of which corresponds to a possible state (one possible history, perhaps truncated by a limited memory) and decodes to the player's action in the next period. The longer the memory, the longer the string, because the greater the possible number of states. Moving to a game with more than the two possible moves of the RPD will lengthen the string, holding the number of players constant. Increasing the number of players will also increase the number of states. Formally, the number of states is given by the formula a^{mp}, where there are a actions, p players, each remembering m periods (Midgley et al. 1997).

Axelrod (1987) used machine learning (specifically, the Genetic Algorithm, or GA, see Holland 1975; and see Marks (2000) for a survey of the use of GAs in the simulation of repeated games) to solve for strategies in the RPD. Although the implicit function of Axelrod's 1987 paper is non-linear and open-form, in one way this pioneering work was limited: the niche against which his players interact is static, so the GA is searching an environment in which the other player's behaviour is determined at time zero: these routines do not learn. This type of problem is characterised as open-loop, and is essentially non-strategic (Slade 1995).

3.1 Our Work

Marks and his coauthors have been involved in a project to use the GA to solve for MPE using historical data on the interactions among ground-coffee brands at the retail level, where these players are modelled as Axelrod finite automata in a differentiated Bertrand oligopoly. This work is a generalisation of Axelrod (1987) and Marks (1992), and uses the ability of the GA to search the highly disjoint space of strategies, as Fudenberg & Levine (1998) have suggested.

The first initial work (Midgley et al. 1997) built on a "market model" that was estimated from historical supermarket-scanner data on the sales, prices, and other marketing instruments of up to nine distinct brands of canned ground coffee, as well as using brand-specific cost data. The market model in effect gives the one-shot payoffs (weekly profits) for each brand, given the simultaneous actions of all nine (asymmetric) brands. There was natural synchrony in brands' actions: the supermarket chains required their prices and other actions to change together, once a week, at midnight on Saturdays.

Modelling the three most rivalrous historical brands as artificial adaptive agents (or Axelrod strings), the authors use the GA to simulate co-evolution of these brands' strings and behaviours, with the actions of the other six brands taken as unchanging, in a repeated interaction to model the oligopoly. After convergence to an apparent MPE in their actions, the best of each brand is played in open-loop competition with the historical data of the other six in order to benchmark its performance. The authors conclude that the historical brand managers could have improved their performance given insights from the response of the stimulus-response artificial managers, although this must be qualified with the observation that closed-loop experiments (Slade 1995), allowing the other managers to respond, would be more conclusive.

Marks et al. (1998) refine the earlier work by increasing the number of strategic players to four, by using four distinct populations in the simulation (against the one-string-fits-all single population of the earlier work), and by increasing the number of possible actions per player from four to eight, in order to allow the artificial agents to learn which actions (prices) are almost always inferior and to be avoided.

Marks (1998) explores some issues raised by the discrete simulations and the curse of dimensionality, and uses entropy as a measure of the loss of information in partitioning the continuous historical data (in prices etc.).

3.2 The Retail Ground-Coffee Market

Using supermarket-scanner data, we are able to examine a market in which several distinct players (in this case, the brands) interact. The historical data describe players who are asymmetric: they have distinct costs; they exhibit distinct responses to the history of interactions with other brands, as summarised in the state of the market, which may also be distinct; and they are perceived as distinct by the supermarket customers, who respond differently in aggregate to each brand, cet. par. (Their responses are prices, as well as their participation in the other marketing actions, including supermarket-aisle displays, distributing coupons to local households, and featuring their specials in the supermarket advertising fliers.)

Each brand's price (and other marketing instruments) is held constant for a period of seven days, and then all brands change their prices (etc.) at the same time, orchestrated by the supermarket chain. We call this market a "synchronised iterated oligopoly".

Using the Casper model (Cooper & Nakanishi, 1988), from any combination of each brand's choice of price, display, coupons, and feature, we can closely predict: the set of the brands' sales per week, the set of the brands' revenues per week, and the set of the brands' profits per week (equivalent to the stage-game payoffs). (The price is is cents per pound, adjusted down by the value of the coupons that week, weighted by the proportion of supermarkets where coupons had been distributed; the display and the feature marketing instruments are percentages of stores.)

3.2.1 The Coffee Market The data refer to a local U.S. retail market for ground-caffeinated coffee. There are nine brands or players. Table 1 gives the average prices ($/lb) and market shares for each of the nine. The data are aggregated on a supermarket chain.

Brand	Price	Market Share
Folgers	$2.33	21%
Maxwell House	$2.22	20%
Chock Full O Nuts	$2.02	11%
MH Master Blend	$2.72	10%
Chase & Sanbourne	$2.34	4%
Hills Bros.	$2.13	4%
Yuban	$3.11	1%
All Other Branded	$1.96	3%
All Other Private Labels	$1.95	27%

TABLE 1. The Nine Brands: Average Price and Market Share

As mentioned above, there are four “marketing instruments”:

1. prices (the price that week of the brand);
2. newspaper features (the percentage of stores in the chain featuring the brand’s item in their distributed advertising);
3. in-store displays (the percentage of stores in the chain featuring the brand’s item as an aisle display); and
4. coupons, which are distributed to households in the district, for redemption of the brand’s product at the supermarket chain. (We found a high correlation between discounted prices and distribution of coupons — use of which lowers the effective price still more — so we adjusted the price in any week by the percentage of coupons distributed.)

4. Computer Experiments

The coffee market is a complex adaptive system: there are several interacting sellers; dynamic, aggregate behaviour (patterns of price competition) emerge; and sellers act to increase profit (or, perhaps positively, market share). For this reason, and given the technical difficulties of solving for optimal (profit-maximising) behaviour in closed-form, we simulate the market by simulating the behaviours of the brands’ managers.

We speak of the use of artificially intelligent adaptive economic agents (AIAEAs) in computer experiments. Such modelled agents are flexible but consistent. The computational experimenter can therefore carefully control the agents’

- information (what they know when);
- learning (how information about their own and others’ behaviour alters their future responses);
- degree of bounded rationality (in particular, their memory of past weeks’ actions and outcomes, perhaps aggregated into coarser partitions);

- sets of possible actions (their deterministic responses to the perceived state of the market); and
- payoffs (which, like their information, learning, memory, partitioning, and actions, can be asymmetric).

Simulation, although it cannot establish necessity, does enable exploration of the *sufficient* conditions for the emergence of particular aggregate market phenomena, given players' micro behaviour.

4.1 First Results

To begin with, we considered the three most interactive players in the market: Folgers, Maxwell House, and Chock Full O Nuts. We decided to allow each agent four actions, as derived from an analysis of their historical prices and other marketing actions. Table 2 shows the four allowed prices for each of the three agents.

	Folgers	Maxwell House	Chock Full O Nuts
LO	1.87+	1.96+	1.89+
lo	2.07+	2.33+	2.02+
hi	2.38	2.46	2.29
HI	2.59	2.53	2.45

TABLE 2. The Four Possible Actions of the Three Strategic Brands

The plus sign (+) indicates that other marketing instruments (features, displays, and coupons) are also being used at the marked prices, which are net of coupons.

Our intention was to pit the three strategic brands against each other, while the other brands were unchanging or non-strategic players, in order to examine the coevolution of the three agents' behaviour. We would need to distinguish convergence of behaviour (phenotype) from structure (genotype).

We used the Casper market model to derive the three $4 \times 4 \times 4$ payoff matrices for the three players Folgers, Maxwell House, and Chock Full O Nuts (CFON). The payoff matrix indicates any brand's weekly profit for each of the 64 combinations of price given in Table 3, given the non-strategic prices of the other six brands ($/lb).

Master Blend	Hills Bros	Yuban	C&S	AOB	APL
2.90	2.49	3.39	2.39	3.68	2.19

TABLE 3. The Fixed Prices of the Other Six Brands

With one-week memory, the agents are modelled as bit strings of length $2 \times 4^3 + 6 = 134$ bits. (The 6 bits of phantom memory endogenise initial conditions: each agent has four possible actions coding to 2 bits, and there are three strategic players.)

Each agent plays a 50-round game with each possible combination of the other two players. The Genetic Algorithm uses 25 mappings (or strings) per population for each agent. Therefore, testing each generation requires 8125 50-round games, or 325 games per string per generation. Each agent has complete information of all previous actions in each 50-round game, but not others' weekly profits (payoffs). See Midgley et al. (1997) for more details.

4.2 Unconstrained agents.

Our prior was that unconstrained co-evolution of strategic agents in repeated interaction would result in convergence to a collusive, high price, at which monopoly profits would be shared among the three members, in effect, of the cartel. This did not happen. Instead, we obtained convergence of all agents to pricing at their *LO* prices.

On reflection, we could see two reasons for this outcome. First, the actions were exogenous to the coevolution and were based on the historically observed actions of actual brand managers, who had been operating in a market with legal and market constraints. Specifically, antitrust statutes would have precluded collusion, even at arm's length. But, more importantly, the existence of outlets offering identical products for sale at competitive prices would have led to a kinked demand curve, with no sales at exorbitant prices, as customers bought from rival supermarket chains.

The historical data show that most sales and profits were made at *LO* prices with promotions, as is evident in Figure 1.

4.3 Institutional constraints

As a result of the first, unconstrained experiment, we added institutional constraints: no agent could follow a *LO* price with another *LO* price in consecutive weeks, and only one agent per week could price *LO* (see Table 2).

This resulted in an interesting pattern of behaviour, in which agents alternated (roughly) in pricing *LO*, with the other two pricing *lo, hi,* or *HI*.

Below, when we increase the number of agents to four, and the number of possible actions to eight, we find that running the simulation to see which illegal behaviour emerge as phenotypes is too slow, and instead we use a filter which screens the structure of the agents' strings (the genotype) to eliminate individual agents which are programmed to necessarily behave illegally before play begins.

The results of filtering are seen in Figure 2A (50 Monte Carlo runs with random initial strings), and Figure 2B (50 Monte Carlo runs with legal initial strings). It is readily seen that in Figure 2A convergence is slow, not occurring until 72 generations, while in Figure 2B (with filtering) convergence is three times as fast, at 24 generations.

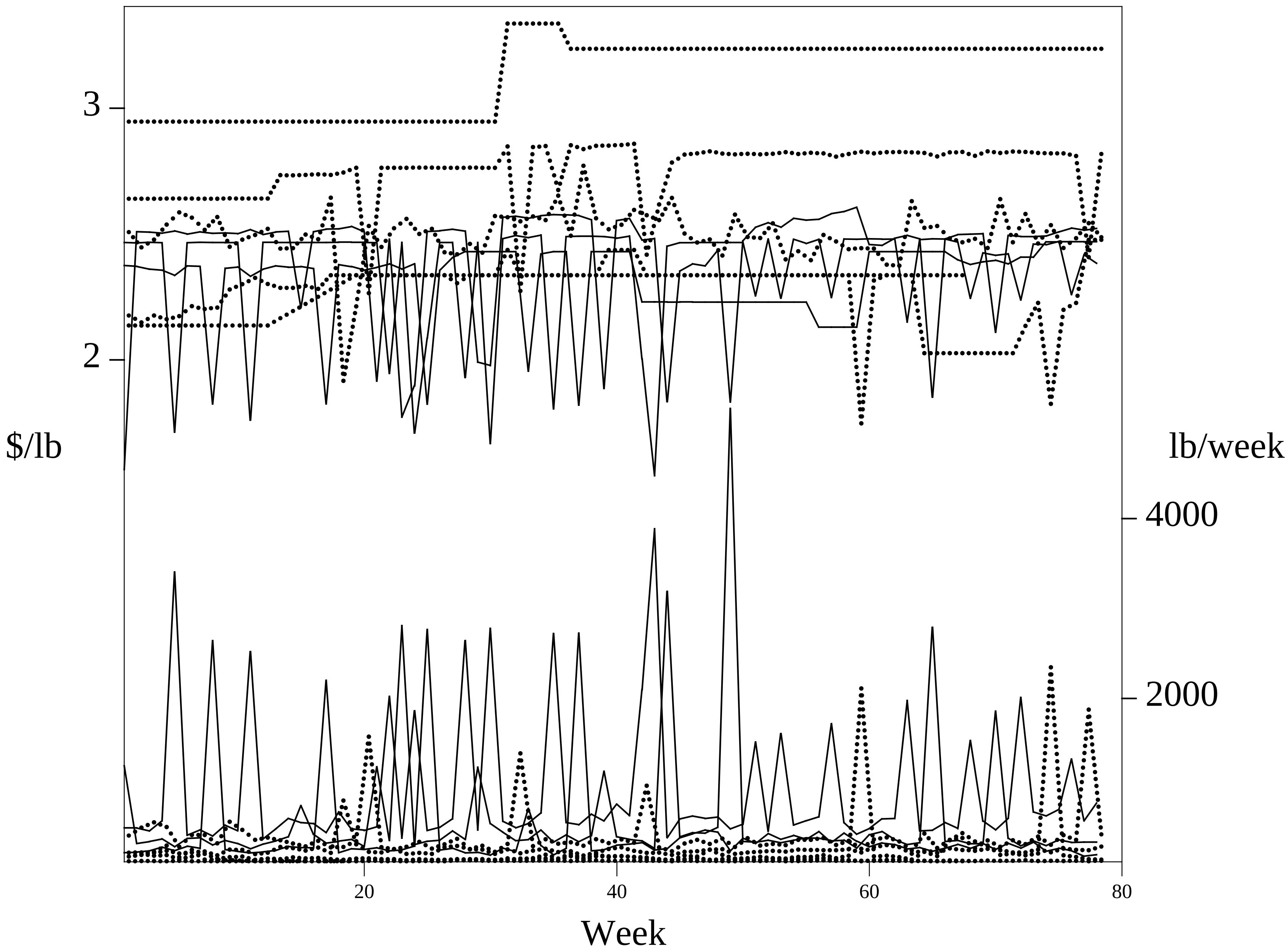


Figure 1: Weekly Prices and Sales (Folgers, Maxwell House, CFON solid)

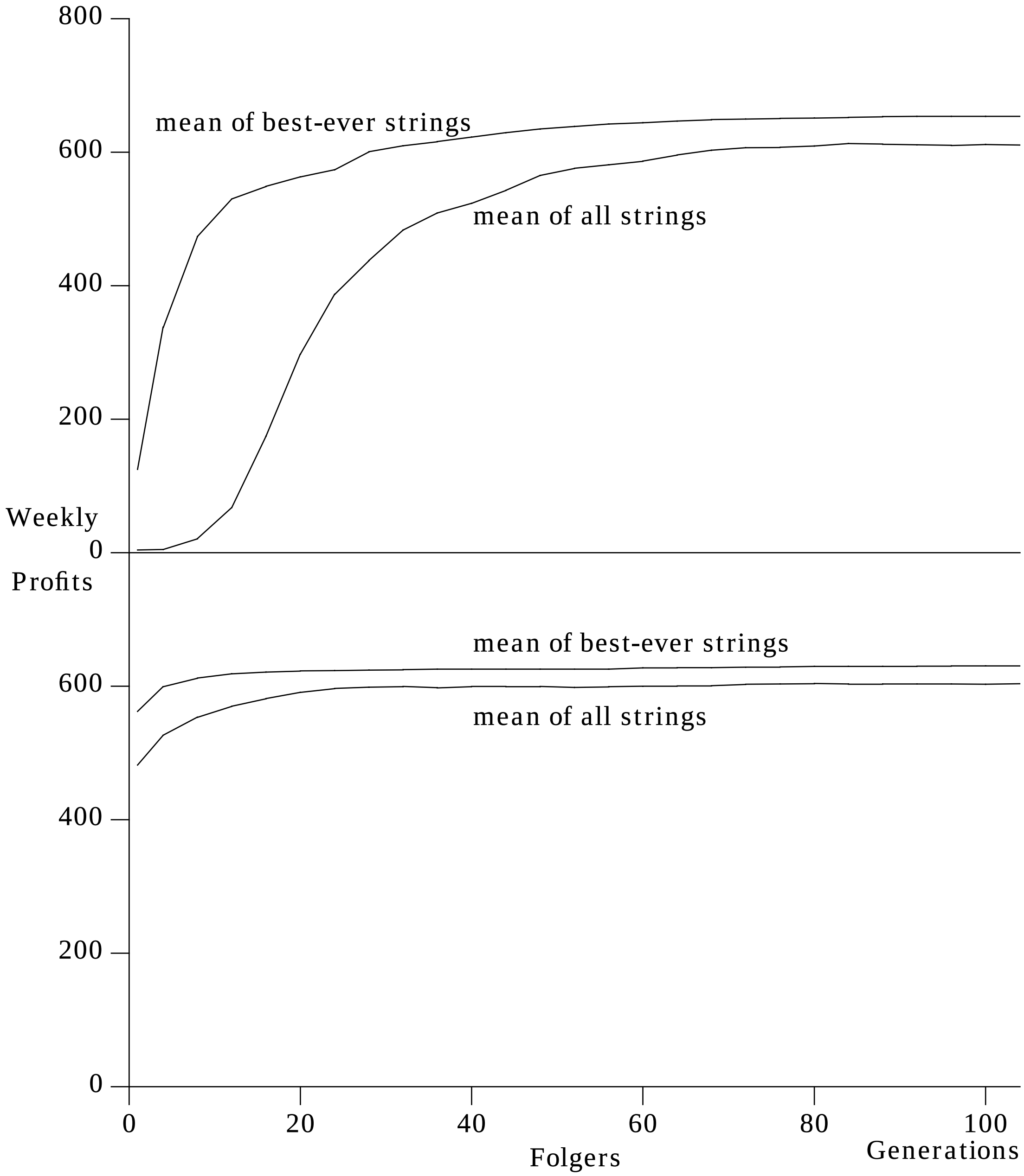


Figure 2: Learning —
with Random Initial Population (above)
with Filtering: Legal Initial Population (below)

4.4 Demand Saturation.

But too frequent pricing of *LO* and *lo* results in "saturation" of demand. That is, the Casper market model is a one-shot model, while the vacuum-sealed coffee is durable (up to seven weeks' shelf life). Casper does not take account of the fact that purchases in any one week may affect purchases up to seven week later. We found that, over several weeks, the demand was unrealistically high.

To counter this, we added demand saturation: after the seventh week the demand (the sum of the demands for each of the nine brands, as predicted by the Casper market model) was prorated by the degree of oversaturation of the past seven weeks' demand.

This altered the sales and hence the payoffs of the agents. In some cases, we got convergence of behaviour to all *lo* pricing. In other cases, we got patterns of behaviour similar to those seen historically in Figure 1.

4.5 Tests against history

Having coevolved populations of each of the three strategic agents over one hundred generations, we decided that one way to demonstrate the extent to which the agents had learnt to act effectively was to use the most profitable agent by brand from the hundredth generation and play it against the history of play of the other strategic brands. In order to do this, we had to partition the historical actions into four intervals for each of the three strategic brands. We measured performance by the average profits over the seventy-five week history.

For Folgers and CFON the agents improved on their historical performance, but Maxwell House sometimes did worse, even on average. But this was an "open-loop" simulation: the historical managers had responded to the historical actions of *all* others, but here could not respond to the agents' actions. Nonetheless, our very simple agents generated reasonable performance in a noisy environment.

4.6 Multiple-population simulations

Heretofore we have simulated rivalry between the three strategic brands (Folgers, Maxwell House, and Chock Full O Nuts) using a GA with a single population of 25 strings. In addition, we have used a single population of strings, in a one-string-fits-all-three-brands simulation, where the dollar payoff depends on which is the designated brand at any stage.

But, as we stated in Tables 1 and 2 above, the brands are heterogenous. This is underlined in Table 4, which includes a fourth brand, Hills Bros.

	Own-Price Elasticity of Market Share	*AVC* ($/lb)
Folgers	–4.4	$1.39
Maxwell House	–3.9	$1.32
CFON	–4.7	$1.19
Hills Bros.	–0.5	$1.18

TABLE 4. Asymmetries of the Four Brands

We extensively rewrote the GA software (based on John Grefenstette's GENESIS package) to allow the simultaneous simulation of up to four populations of agents (modelled as bit strings). The effect of this was to reduce the "noise" in which the strings compete: with a single population of strings, the performance of any string, in an identical environment, depends on which brand it is designated as, since the payoff matrices are distinct: an identical combination of three competing genotypes (structures) may result in different payoffs. With distinct populations, the only noise derives from the GA itself, which allows the simulation to "learn" more effectively.

The results bear this out. Figure 3 shows the results of the simulations for Folgers' weekly profits, using both a common, single-population model and a three-population model, over a 50-run Monte Carlo, with the same three strategic brands as above, with the four actions of Table 2, and one week's memory. Comparing the two, it is readily apparent that the distinct-population simulations converge faster (at 20 generations or 20% sooner), with higher weekly profits (roughly $100 or 16% per week higher), and with less variance across the population's profits.

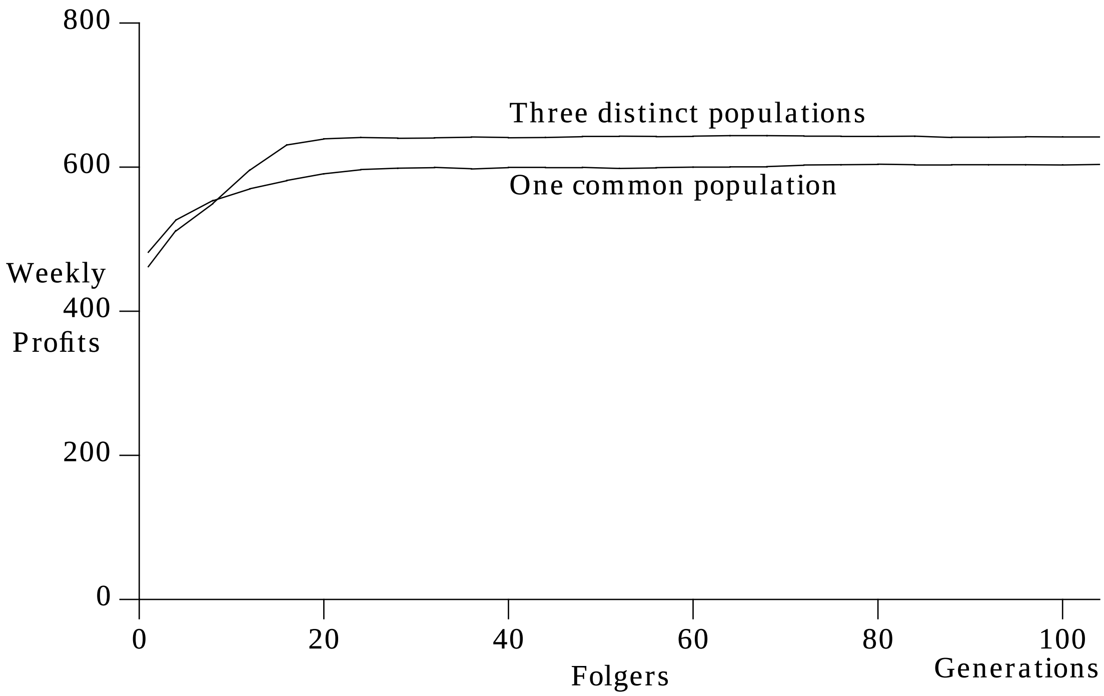


Figure 3: Evolution of Folgers' Profits,
One Common Population and Three Distinct Populations

All of these are consistent with a lower noise level in the GA search. The results of moving from a common-population model to a distinct-populations model are similar for the other two strategic brands.

When we look at the patterns of play, the differences are not as striking. Figure 4 shows the patterns and average weekly profits seen in a 50-run Monte Carlo for the three agents with a common population; Figure 5 shows the patterns and average weekly profits with three distinct populations. For most of the runs, the agents' behaviour is very similar (Folgers and CFON pricing at an Every Day Low Price, or EDLP; Maxwell House exhibiting Wide Pulsing, or WP). For one of the runs, however, the simulations differ, with Folgers moving from pricing to the minimum (PttM) with a common population to EDLP with the distinct-population model; CFON's behaviour also changed, as seen.

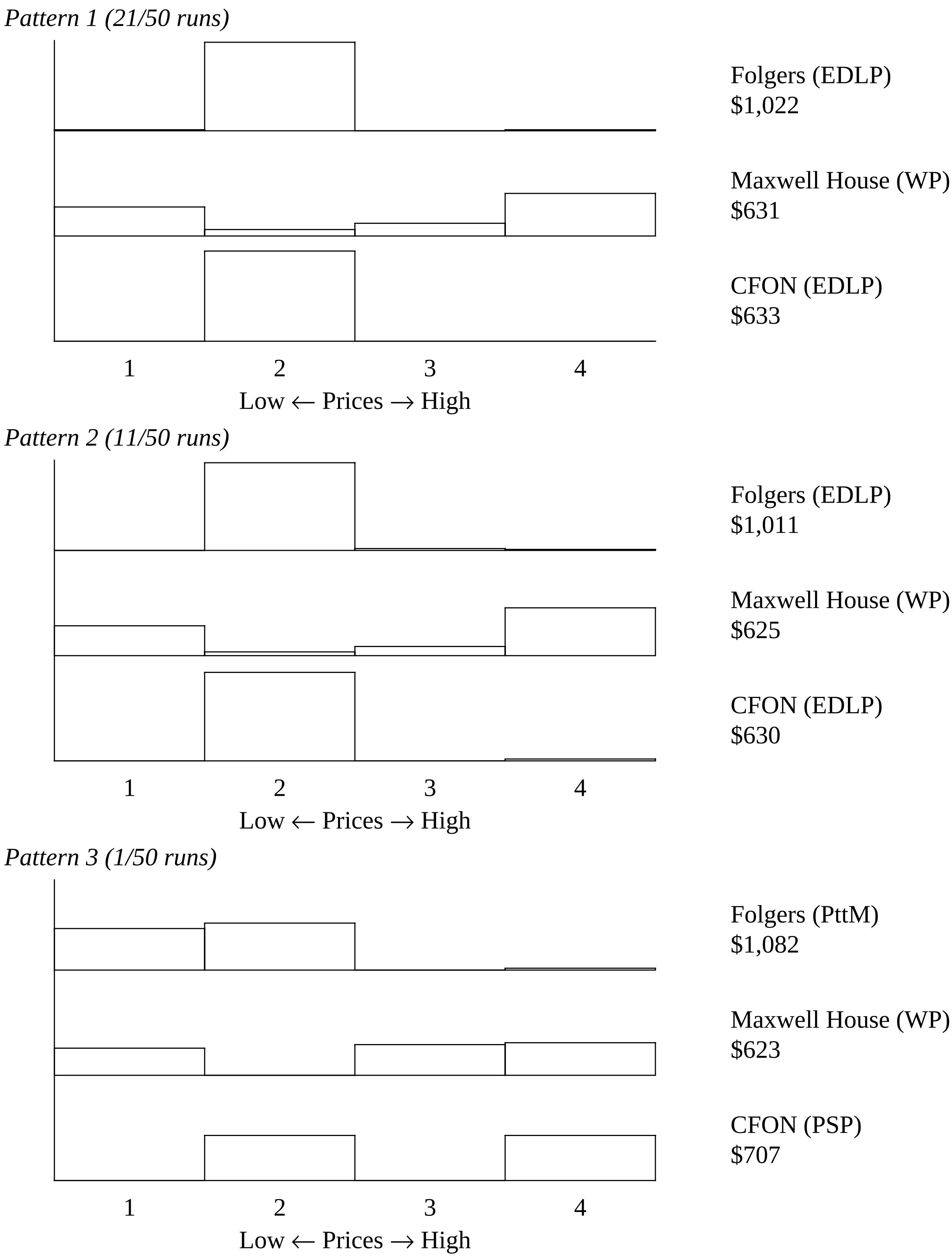


Figure 4: Three Agents, Four Actions, Common Population

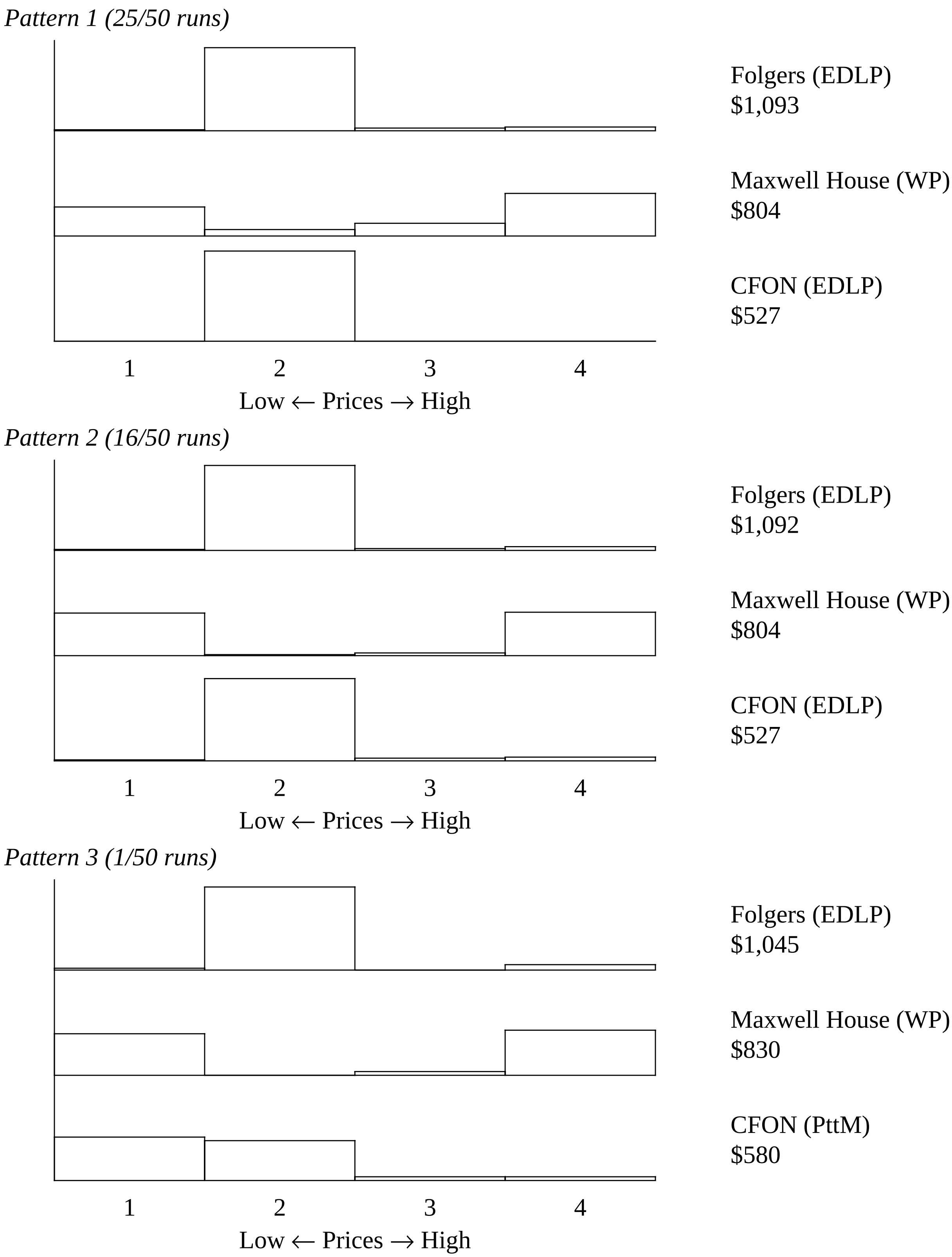


Figure 5: Three Agents, Four Actions, Distinct Populations

4.7 Four strategic players

Previously, we modelled the oligopoly with three strategic players, each with four possible actions, remembering one week back. As discussed above, the agents were modelled as bit strings of length 134 bits. To improve the realism of the simulation, we increase the number of strategic brands to four, by including Hills Bros. This increases the bit-string length from 134 bits to 520 bits.[2] We chose Hills Bros., despite its small market share, as the fourth strategic agent, because the fourth largest brand (Master Blend) is not independent of Maxwell House, and so their strategic actions could be orchestrated by the owner.

The results of introducing the fourth strategic brand are striking. Even though Hills Bros. has a small market share (4%), its introduction is quite significant. The market changes in significant, complex, and asymmetric ways. There are changes in the other brands' behavior as well as in other brands' average weekly profits. Figure 6 shows three patterns and weekly profits which comprise 38 of 50 Monte Carlo runs. The new strategic agent apparently takes up some of the fixed number of opportunities for major promotions, and has differing competitive impacts on the other brands. What these simulations demonstrate is that a small player (as measured by market share) isn't necessary insignificant strategically.

4.8 Eight actions per player

Heretofore the strategic agents (whether three or four) have been constrained by the four possible actions, chosen from the historically observed actions of the actual brand managers. In effect, the agents were given a choice of pricing high or low, with minor variation around the two positions, and they were constrained by the corporate memory and prior learning of the actual brand managers, who had, we assume, learned not to price too high (and sell very little) or too low (and earn little and perhaps spark a price war).

We wanted to increase the choices of the agents. The simplest way was to double the number of possible actions per agent from four to eight. The effect of this on the bit-string length will depend on the number of strategic agents: for three agents, with one-week memory, allowing eight possible actions instead of four increases the length from 134 bits to 1,545; for four agents, the length increases from 520 bits to 12,300 bits.[3]

By increasing the number of actions to eight, we hoped to give our agents the opportunity to demonstrate that the four actions used earlier were robust, and that our assumption of a mature oligopoly were correct, at least

2. Four actions requires 2 bits per action; 4 actions, 4 players, and 1-week memory implies $4^4 = 256$ possible states; phantom memory is $4 \times 2 = 8$ bits. So $2 \times 256+8 = 520$ bits per string.
3. Eight actions requires 3 bits per action; 8 actions, 3 players, and 1-week memory implies $8^3 = 512$ possible states; phantom memory is $3 \times 3 = 9$ bits. So $3 \times 512+9 = 1,545$ bits per string. Eight actions per player and 4 players (while retaining 1-week memory) requires $3 \times 8^4 + 4 \times 3 = 12,300$ bits per string,

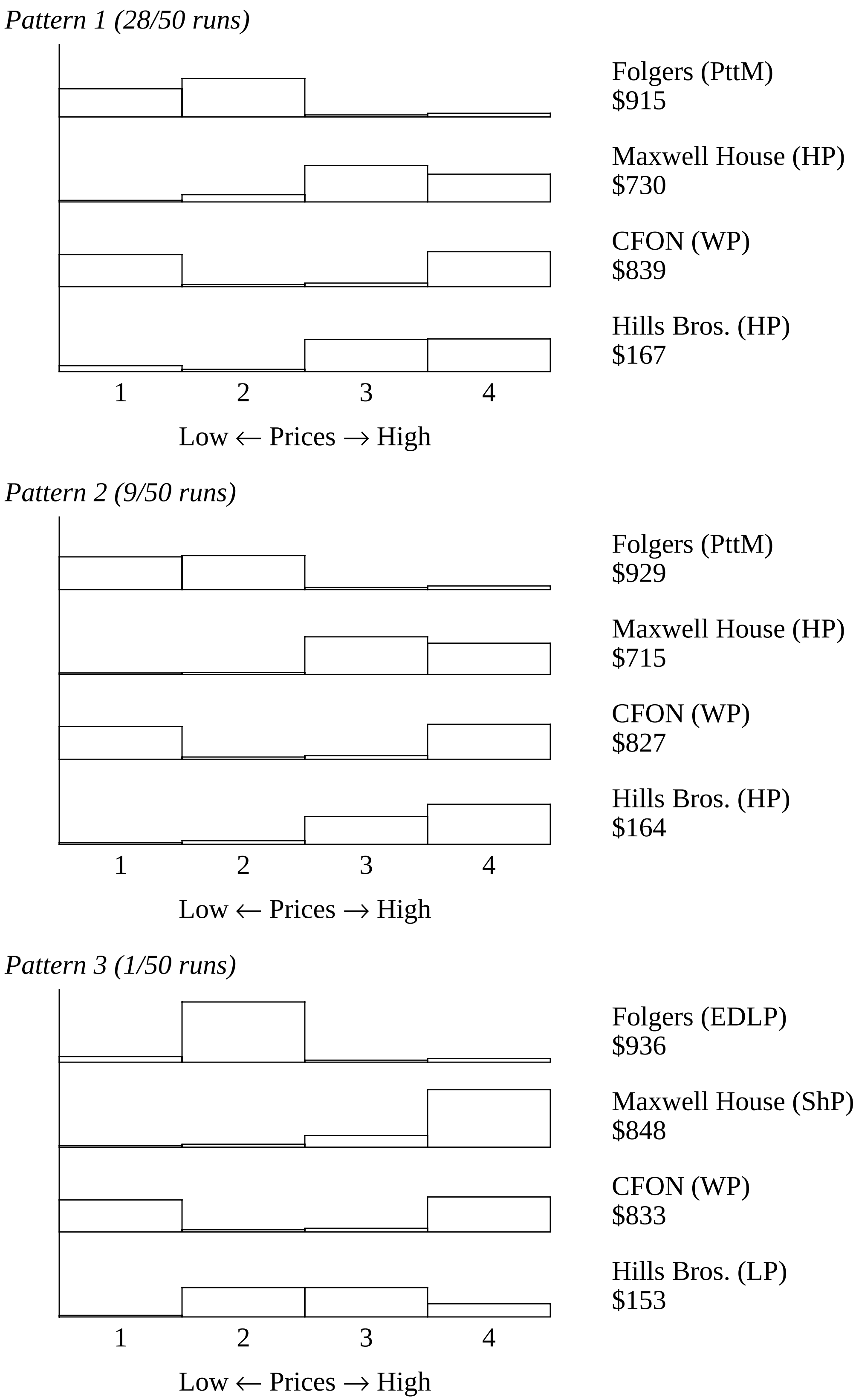


Figure 6:, Four Agents, Four Historical Actions — Hundredth Generation

in terms of the combinations of prices and other marketing actions encountered.

Moving to eight possible actions, especially including some beyond the observed range of actions of the historical brand managers, introduces the possibility of the agents learning anew what was embodied in the historical range: not to price too high or too low.

	Folgers			Maxwell House			CFON		
A	P ($/lb)	F (%)	D (%)	P ($/lb)	F (%)	D (%)	P ($/lb)	F (%)	D (%)
1	1.87*	95*	69*	1.96*	95*	69*	1.89*	100*	77*
2	2.07	83	0	2.33	83	0	2.02	100	65
3	2.38	0	0	2.46	0	0	2.29	0	0
4	2.59	0	0	2.53	0	0	2.45	0	0
1	1.62*	67*	67*	1.60*	97*	97*	1.64	0	0
2	1.83*	97*	96*	1.87*	94*	91*	1.89*	97*	97*
3	1.96	0	0	2.06*	88*	76*	1.89*	98*	29*
4	2.03*	79*	77*	2.33	79	0	2.01	0	0
5	2.04*	85*	0*	2.38	54	0	2.02*	97*	62*
6	2.22	96	33	2.52	0	0	2.31	0	49
7	2.57	0	0	2.53	0	53	2.33	0	0
8	2.78	0	0	2.59	0	13	2.49	0	0

* Asterisked actions are subject to store policy. A is Action, P is Price, F is advertising Feature, D is aisle Display.

TABLE 5. Sets of Four and Eight Possible Actions.

Figure 7 shows the weekly profits and patterns of behaviour, as reflected by the frequency of actions across the three strategic agents. The data refer to 50-run Monte Carlo simulations. (The diamonds in the figure correspond to the asterisks in Table 5: actions subject to store policy.)

After four generations, starting from a uniform distribution of actions (because the bit strings are chosen randomly to begin with, apart from filtering against illegal actions), we see that the frequencies of actions are still almost uniform. After 100 generations, however, the agents have focussed on only two or three main patterns of interaction, with many fewer than the eight possible actions used frequently: agents have *co-learnt* the two or three actions that are most profitable, given others' behaviour. The patterns are brand-specific.

Specifically, with three strategic agents: CFON is pulsing between shelf price (high) and feature price (low). Folgers exhibits three pulsing patterns: P2 — pulsing three actions, P1 — more diverse pulsing, with four actions, and P3 — pulsing with two actions. Maxwell House exhibits a less dynamic choice of Every Day Low Price, and avoids the store constraints

From a 50-run Monte Carlo simulation of four agents and eight possible actions, we observe in Figure 8 for 30 runs that all four agents exhibit what might be termed High Pulsing, and that for 9 runs they exhibit High Entropy.

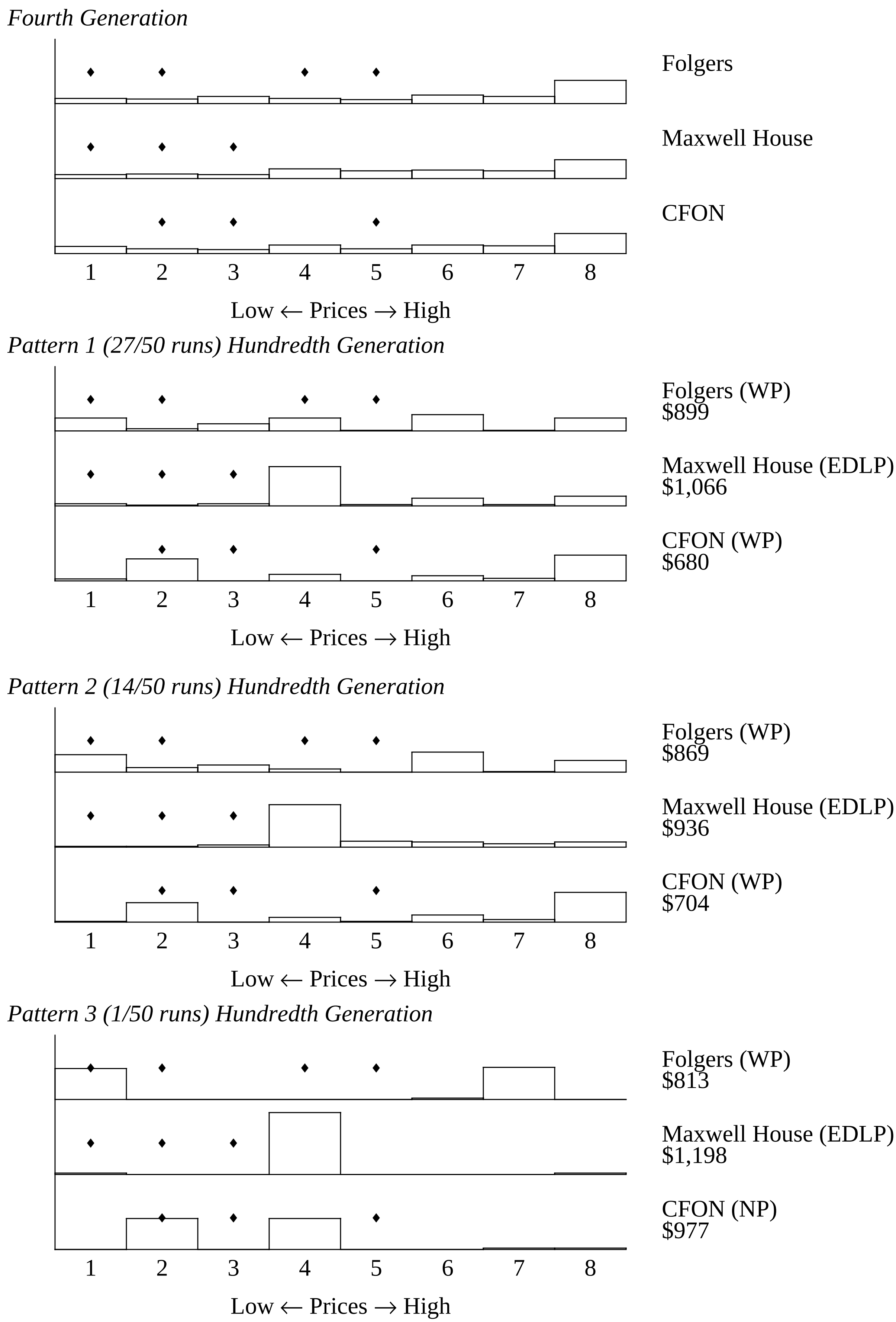


Figure 7: Three Agents, Eight Historical Actions

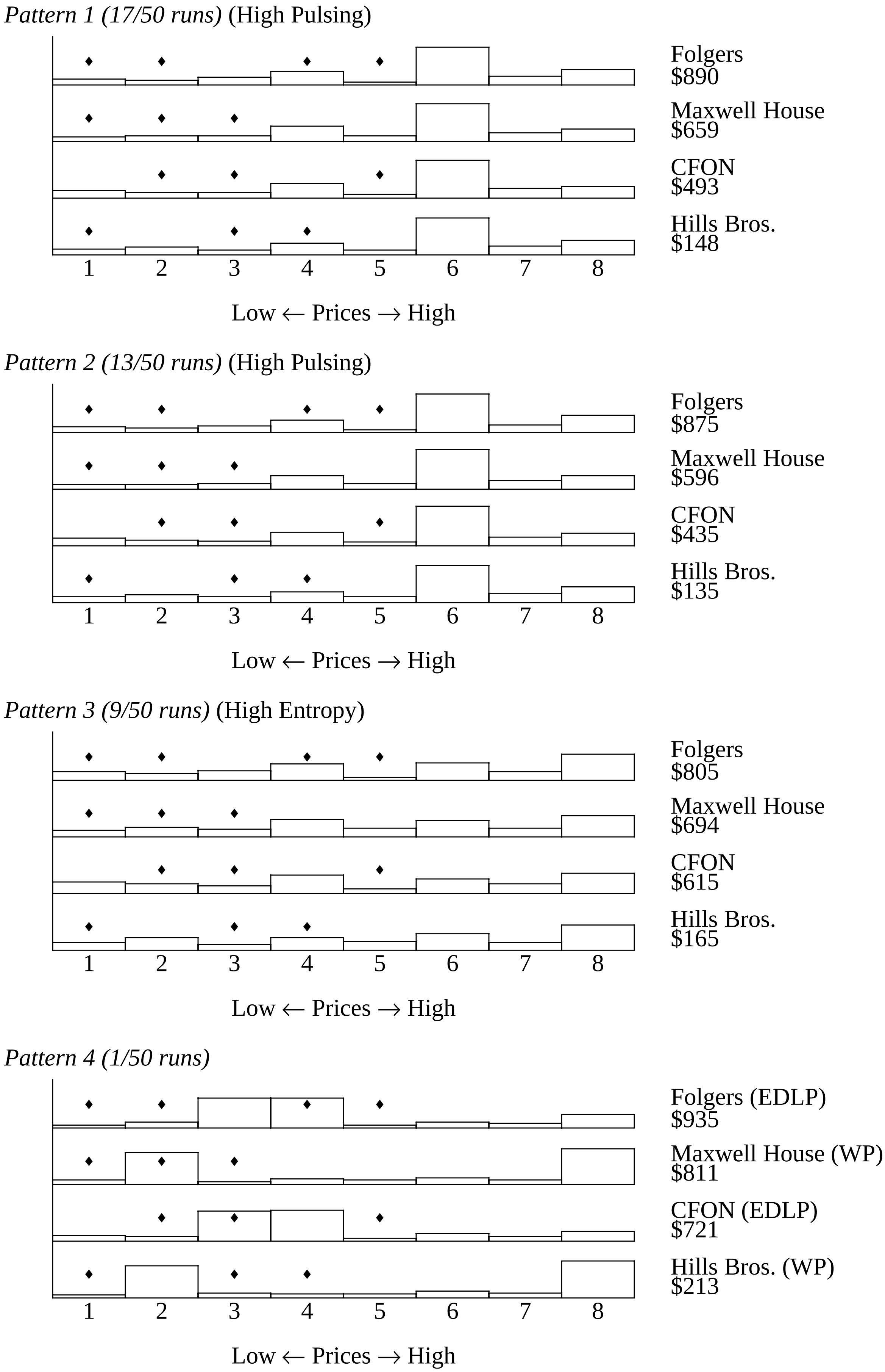


Figure 8: Four Agents, Eight Historical Actions — Hundredth Generation

Overall, we can say that with eight possible actions, a greater degree of heterogeneity emerges, and that a greater number of actions results in a greater number of possible paths of coevolution. Moreover, adding a fourth strategic agent increases the degree of competition in the market, which is reflected in lower average profits, as well as different behaviour. But the choices from eight possible actions appear to allow more effective shifts in agents' behaviour than do four possible actions. For Folgers and Maxwell House, average weekly profits rose as the number of possible actions rose from four to eight, while for CFON, the profits fell.

When we repeated the open-loop plays between the best of the co-evolved three agents with eight possible action and the historical brand managers, we found that the best agents clearly outperformed their historical counterparts: for Folgers by 156%, for MH by 32%, and for CFON by 42%.

The Frankenstein Effect: Agents that showed only a few behaviours in the coevolutionary "lab" were able to evince a wider repertoire when faced with a more variable environment (the history of actual managers' behavior). We dub this the Frankenstein effect because the artificially bred agents were more interesting in the wild than in the lab.

4.9 The Holyfield-Tyson effect

The artificial agents "learn" through application of the evolutionary techniques of the GA. This is clear when the agents are solutions to a static problem, as has been the most usual application of GA techniques in, say, engineering. It is also the case that the first application of GAs in economics (Axelrod 1987) was static, even if stochastic: Axelrod used GAs against a non-evolving but mixed-strategy niche of algorithms derived from the early Prisoner's Dilemma tournaments (Axelrod 1984). But Marks (1992) and others following have bred artificial agents against each other, a process that Marks called "bootstrapping" and biologists term "coevolution".[4]

Against a static environment, progress of the artificial agents is readily revealed by their improving fitness scores, but against a dynamic environment comprised of like artificial agents scores may not rise from generation to generation.

Apart from the growth in average weekly profits see in Figures 2 and 3, there are at least two further ways to demonstrate that the artificial natural selection has improved the agents' performances. In our earlier work we

4. "Nothing is absolutely predictable about the direction of coevolution. How an interaction coevolves depends not only on the current genetic makeup of the species involved but also on new mutations that arise, the population characteristics of each species, and the community context in which the interaction takes places. Under some ecological conditions, an antagonistic interaction between two species can coevolve to enhance the antagonism; the species "build up" methods of defense and attack, much like an evolutionary arms race. Under other ecological conditions, however, the antagonism may be lessened, resulting in reduced antagonism" or coöperation or mutual dependence. *Encyclopædia Britannica*, CD97, Propædia: The Biosphere and Concepts of Ecology, The coevolutionary "arms race" versus reduced antagonism.

attempted to show the greater competence of our artificial agents by pitting them against the historical histories of play of their opponents, but some criticism has been made that this overstates the skills of the artificial agents and understates the skills of the historical agents, who have no opportunity to respond to the actions of the artificial agent: their plays are given, or open-loop.

Here we attempt to show how the artificial agents have learnt by taking agents after 2,500 trials (100 generations) and playing them against not the frozen moves of their historical opponents, but the agents after only 200 trials (8 generations): a process we have termed pitting a "sophisticated" agent against "naïve" agents. How to show that the coevolved agents are learning to respond better (are truly fitter)? Previously: we considered the mean weekly profits, as seen in Figures 2 and 3. Now: in turn we replace the best naïve (at 8 generations) Folgers (respectively, Maxwell House, and CFON) string with the best sophisticate (after 100 generations) Folgers (respectively, Maxwell House and CFON) string.

The procedure followed was:

1. After 8 generations, identify the best string from each of the 3 populations.
2. Play these 3 against each other for a 50-week repeated game; note average weekly profits.
3. Allow the 3 populations to continue co-evolving via the GA.
4. After 100 generations, identify the best strings from the 3 populations, play them against each other as before; note average weekly profits. Table 6 shows these results.

Experiment	Folgers	Maxwell House	CFON	Hills Bros.	Total
1 pop., 4 actions	1,018	637	635	n/a	2,290
3 pop., 4 actions	1,053	793	534	n/a	2,380
3 pop., 8 actions	889	985	694	n/a	2,568
4 pop., 4 actions	915	729	835	164	2,479
4 pop., 8 actions	858	660	509	148	2,175

Average weekly profits computed from 50 Monte Carlo simulations and all combinations of agents. Historical-action sets.

TABLE 6. Performance of Hundredth-Generation Agents Competing with Each Other

5. Replace the best Folgers string after 8 generations by the best Folgers string after 100 generations (i.e. replace the best primitive string by the best sophisticate).

6. Play all combinations of 3 strategic brands, and consider string-by-string the change in average weekly profits with the sophisticated player and without the sophisticated player in one brand.
7. Repeat steps 5 and 6 for the remaining 2 brands.
8. Repeat steps 1-7 50 times. Table 7 shows the performances.

Experiment	Folgers	Maxwell House	CFON	Hills Bros.
Historical actions	188*a*	198*a*	69*a*	?
1 pop., 4 actions	431*b*, 480*c*	201, 236	79,105	
3 pop., 4 actions	410, 468	271, 329	107, 113	
3 pop., 8 actions	523, 806	295, 514	104, 124	
4 pop., 4 actions	430, 469	191, 286	103, 111	?
4 pop., 8 actions (60th gen.)	481, 944	262, 559	98, 110	12, 13

a. Average weekly profits computed from historical actions.
b. Average weekly profits computed from playing the best agents from 50 Monte Carlo simulations against historical actions.
c. Single best performance observed.

Note: The profits derived from historical actions will not be the same as single-period Casper results because of the demand-saturation constraint.

TABLE 7. Performance of Best Agents Competing with the Managers' Histories

Table 8 shows the three combinations of results:

	ΔF	ΔMH	ΔCFON
Folgers	–15.0	41.4	42.0
MH	2.0	–20.0	37.8
CFON	13.9	–29.0	82.3

TABLE 8. Mean Changes in Average Profits with the Best Sophisticates

We would have expected positive diagonals (i.e., that sophisticates do better), and negative off-diagonals (i.e., that others' profits fall). Instead, we see that the CFON sophisticate is the only one to improve on the replaced naïve's performance. In the cases of Folgers and Maxwell House, the sophisticates did worse than did the naïves.

The results of Table 8 are unexpected. One possibility is genetic drift, a phenomenon where lack of selective pressure on many alleles (sites) on the bit strings (because of convergence of behaviour, generation after generation, which means that only a small subset of possible states occur, and hence only a small subset of alleles (sites) are triggered) means that those bits may, through chance and recombination, flip, which is only obvious when, in the hurly-burly of rivalry against the naïves, these states are encountered again, after many generations, and the perhaps effete sophisticates do not always cut the mustard.

We have dubbed this the Holyfield-Tyson effect after the notorious championship bout between the two heavyweights, in which Tyson bit off part of Holyfield's ear.[5]

Genetic drift is inversely proportional to the number of individuals in the population. We increased the population size per brand from 25 strings to 250. This led to very slow convergence, even with the short strings in the three-agent, four-action simulations: not only was there a thousand-fold increase in the number of three-way interactions per generation, but there was apparently lengthy spiralling towards convergence of the GA — only a single run was performed, not a Monte Carlo.

After 80 generations, Table 9 shows the best sophisticates against the naïves:

	ΔF	ΔMH	ΔCFON
Folgers	–28.0	–12.2	–50.2
MH	–86.4	–518.3	166.1
CFON	30.6	11.8	–191.3

TABLE 9. Mean Changes in Average Profits with the Best Sophisticates

All three diagonals are negative, which doesn't suggest that genetic drift is the issue. The GA was still converging at 80 generations but the results after 160 generations were no better, and the GA had still not converged.

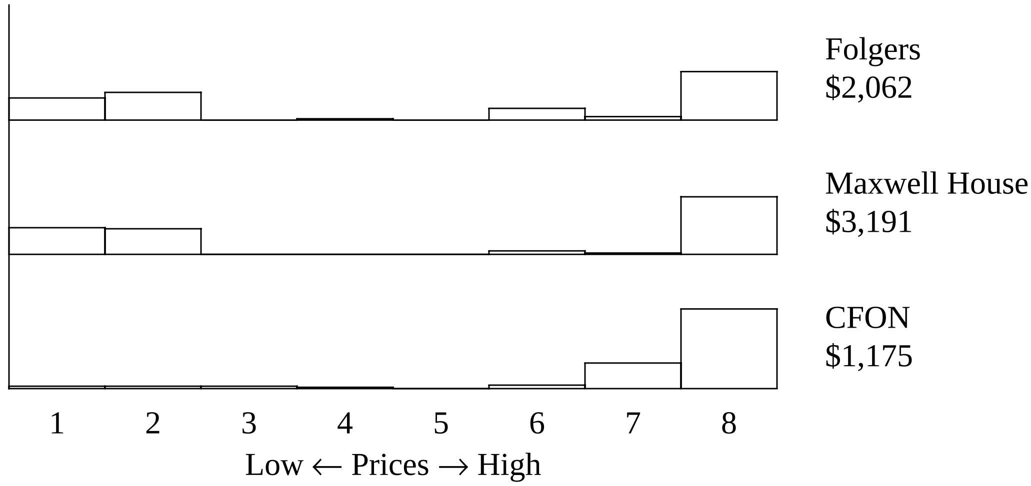
4.10 Managerial Learning

The eight-action sets per player of Section 4.8 were derived from historical actions and so embodied prior learning What if we give the artificial agents a different repertoire of actions — one developed without reference to the historical actions of managers? We used a random experimental design, where the price per pound is stepped in ten-cent increments between $1.60 and $2.80 and feature and display can take on the values of either 0 or 100%.

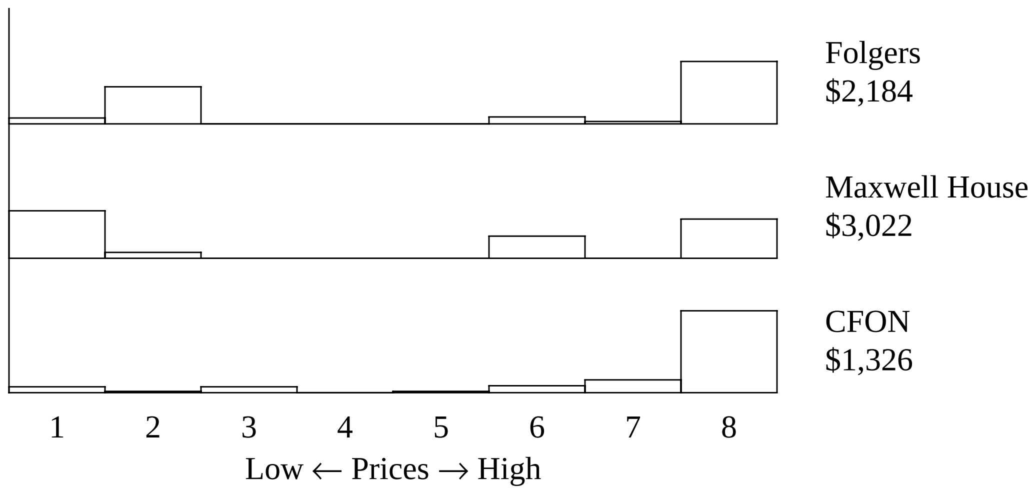
Figure 9 shows three patterns that accounted for 39 of 50 Monte Carlo runs. Note that average weekly profits are much higher than with historical, learned action sets. Note too that in general the agents shun low-price promotions and maintain high prices throughout most interactions. The levels of competition are much lower than with historical-action sets — with these randomly chosen action sets the agents are engaging in the sort of collusion that we'd expected to see in the first simulations in Section 4.2. But we speculate that these results show that inter-chain competition is what our

5. We should like to thank Bernhard Borges for this name.

Pattern 1 (26/50 runs)



Pattern 2 (9/50 runs)



Pattern 3 (4/50 runs)

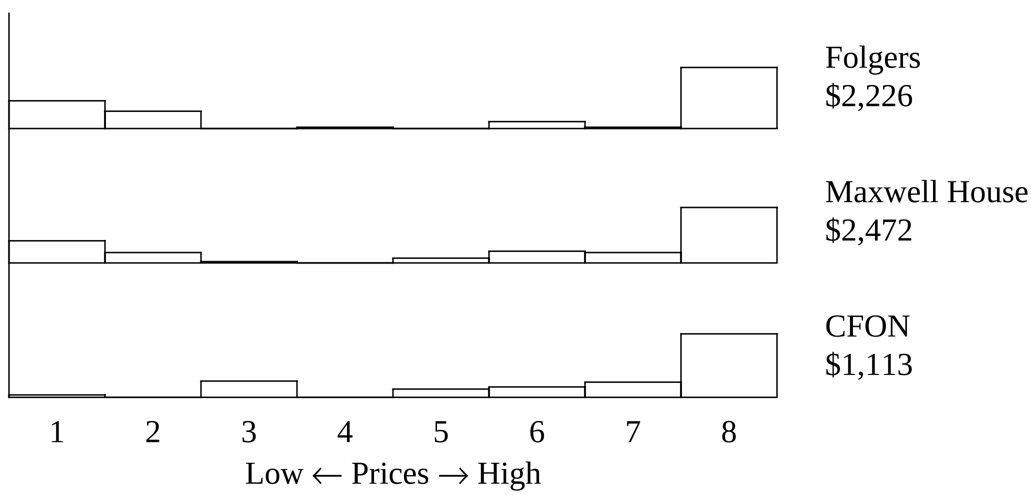


Figure 9: Three Agents, Eight-Random-Action Sets — Hundredth Generation

model (and Casper) lacks — the demand curve for coffee from our supermarket chain must be kinked when potential customers go elsewhere to avoid paying the high prices our artificial agents would like to charge in implicit collusion.

Results of three-player, eight-possible-action simulations reveal two major patterns: much higher average weekly profits, and almost no low, feature pricing, with profits earned at very high pricing. This result is seen in Figure 10, which shows the patterns for the four strategic players under the three regimes: historical frequencies of the brand managers, coevolved agents competing against each other, and the best coevolved agents competing against history. Notice that for Maxwell House and Hills Bros. the coevolved agents' frequencies of actions are very similar to the historical brand managers' frequencies of actions; and for Folgers and CFON the two patterns are similar, with a slightly higher shelf price for the historical managers.

5. Conclusion

We can summarise our experiments on rivalry in a mature differentiated Bertrand oligopoly in two ways: the average weekly profits of the agents, and the patterns of actions.

Table 6 summarises the average weekly profits of the four strategic brands under the different combinations of common and distinct populations and four- or eight-action sets (all derived from the historically observed actions of the brand managers).

Figure 10 summarises the frequencies of chosen actions (eight-action sets, derived from the historically observed actions) under the three conditions of, first, historical actions, second, coevolved agents competing, and, third, agents competing against history (playing the 50 best agents per brand against the historical actions of their three competitors).

Our experiments have revealed some restrictions on the historical brand managers which were not immediately apparent, but more significantly, we have shown that the patterns of interaction among the brand managers were not as profitable as they might have been, even if all strategic players in the oligopoly had been using strategies as finely tuned as our agents had learnt to use, in the simulations using the Genetic Algorithm. We hypothesise that the techniques used here could shed light on the Markov perfect equilibria in similar oligopolies, and on how the actors in those markets might have been able to improve their profits in the past and perhaps in the future.

When John Holland (1975) invented the GA, his original term for it was an "adaptive plan" which looked for "improvement" in complex systems, or "structures which perform well." Despite that, most research effort, particularly outside economics, has been on its use as a function optimiser. But, starting with Axelrod (1987), the GA has increasingly been used as an adaptive search procedure, and latterly as a model of human learning in repeated situations. In the 1992 second edition of his 1975 monograph, Holland expressed the wish that the GA be seen more as a means of improvement and less on its use as an optimiser. The work we report on here

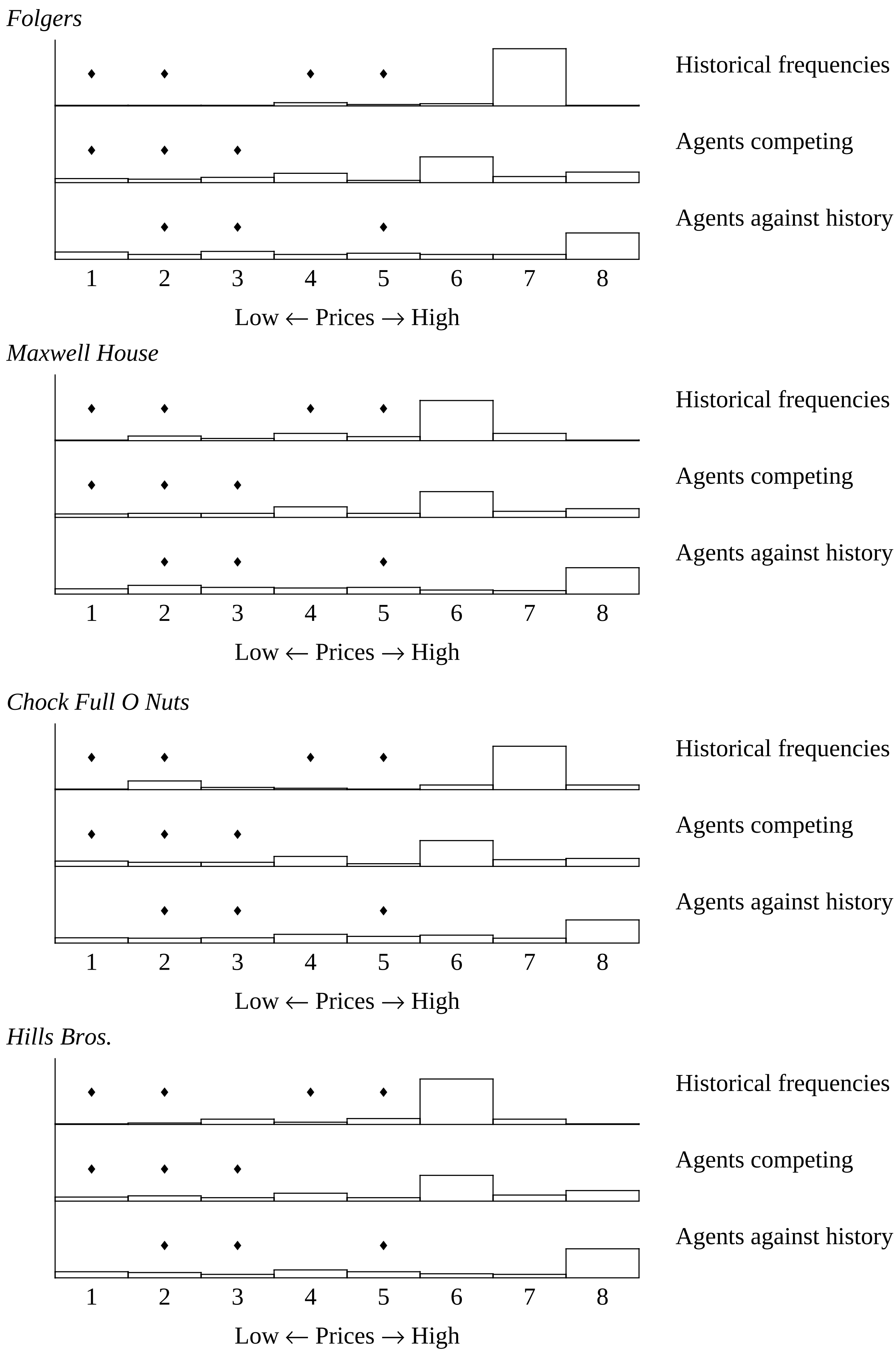


Figure 10: Summary of Patterns

is an example of the usefulness of the GA in a continuing research program about the behaviour of sellers competing in an oligopoly, where the sellers are modelled as automata responding to the past actions of all sellers.

6. References

[1] Axelrod R. (1984) *The Evolution of Cooperation*, New York: Basic Books.

[2] Axelrod R. (1987) The evolution of strategies in the iterated Prisoner's Dilemma, in *Genetic Algorithms and Simulated Annealing*, L. Davis (ed.), London: Pittman.

[3] Binmore K. (1992), *Fun and Games: A Text on Game Theory*, Lexington, Mass.: Heath.

[4] Cooper L.G. & Nakanishi M. (1988) *Market Share Analysis: Evaluating Competitive Market Effectiveness*, Boston: Kluwer.

[5] Dutta P.K. & Sundaram R.K. (1998) The equilibrium existence problem in general Markovian games, Chapter 5 in M. Majumbar (ed.), *Organizations with Incomplete Information: Essays in Economic Analysis, A Tribute to Roy Radner*, Cambridge: C.U.P.

[6] Eaton J. & Engers M. (1990), Intertemporal price competition, *Econometrica*, **58**, 637–660.

[7] Friedman J.W. (1977), *Oligopoly and the Theory of Games*, NY: North-Holland.

[8] Fudenberg D. & Levine D.K. (1998), *The Theory of Learning in Games*, Cambridge: MIT Press.

[9] Fudenberg D. & Tirole J. (1992), *Game Theory*, Cambridge: MIT Press.

[10] Green E. & Porter R. (1984), Noncooperative collusion under imperfect price information, *Econometrica*, **54**: 533–554.

[11] Holland J.H. (1975), *Adaptation in Natural and Artificial Systems*, Ann Arbor: Univ. of Michigan Press. (A second edition was published in 1992: Cambridge: MIT Press.)

[12] Marks R.E. (1992a) Repeated games and finite automata, in *Recent Developments in Game Theory*, ed. by J. Creedy, J. Borland, and J. Eichberger, Aldershot: Edward Elgar.

[13] Marks R.E. (1992b), Breeding hybrid strategies: optimal behaviour for oligopolists, *Journal of Evolutionary Economics*, **2**: 17–38.

[14] Marks R.E. (1998), Evolved perception and behaviour in oligopolies, *Journal of Economic Dynamics and Control*, **22**(8–9): 1209–1233.

[15] Marks R.E. (2000), Playing games with Genetic Algorithms, *Evolutionary Computation in Economics and Finance*, ed. by Shu-Heng Chen, New York: Springer-Verlag, forthcoming.

[16] Marks R.E., Midgley D.F., & Cooper L.G. (1998), Refining the breeding of hybrid strategies, AGSM Working Paper 98–017, December. <http://www.agsm.edu.au/~bobm/papers/wp98017.html>

[17] Maskin E. & Tirole J. (1988a), A theory of dynamic oligopoly I: overview and quantity competition with large fixed costs. *Econometrica*, **56**: 549–570.

[18] Maskin E. & Tirole J. (1988b), A theory of dynamic oligopoly II: price competition, kinked demand curves, and Edgeworth cycles, *Econometrica*, **56**: 571–600.

[19] Midgley D.F, Marks R.E., & Cooper L.G. (1997), Breeding competitive strategies, *Management Science*, **43**(3): 257–275.

[20] Peltzman S. (2000), Prices rise faster than they fall, *Journal of Political Economy*, **108**(3): 466–502.

[21] Rotemberg J. & Saloner G. (1986), A supergame-theoretic model of business cycles and price wars during booms, *American Economic Review*, **76**: 390–407.

[22] Slade M.E. (1992), Vancouver's gasoline price wars: an empirical exercise in uncovering supergame strategies, *Review of Economic Studies*, **28**(2): 368–402.

[23] Slade M.E. (1995), Empirical games: the oligopoly case, *Canadian Journal of Economics*, **28**(2): 368–402.