Co-evolving 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 9313 7279
bobm@agsm.edu.au

David Midgley,
INSEAD,
France
Tel: +33 1 6071 2654
david.midgley@insead.edu

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 idiosyncratic brand's pricing behaviour using the restriction that marketing strategies depend only on profit-relevant state variables, and use the Genetic Algorithm to search for co-evolved equilibria, where each profit-maximizing brand manager is a stimulus-response automaton, responding to past prices in the asymmetric oligopolistic market. Part of a growing study of repeated interactions and oligopolistic behaviour using the GA.

ACKNOWLEDGMENTS: 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. An earlier version of this paper appeared as "Searching for Markov Perfect equilibria in oligopolistic price wars," INSEAD Working Paper 2000/65/MKT. Two anonymous reviewers gave valuable advice.

Co-evolving Better Strategies in Oligopolistic Price Wars

INTRODUCTION

We use simulated evolution to explore oligopolistic behaviour in a (retail) market with up to four strategic sellers, comparing our simulation results with historical data derived from a retail market for ground, vacuum-sealed coffee beans. We find that our boundedly rational sellers perform well (as measured by their average weekly profits) compared to their historical counterparts, despite their limited memory and constrained marketing actions.

Significant features of our work are: first, our agents are heterogeneous: they respond idiosyncratically to others' actions, they have distinct costs, face distinct demand curves, and so earn distinct profits. For this reason, we cannot ignore the identities of the separate players, which would be convenient, were the players identical. Second, we use the Genetic Algorithm (GA) to model the players' learning. To avoid "social learning" (Vriend 2000), when players drawn from a single population pass information to their "offspring" through the genotype (an extra-market mechanism), we use distinct populations for the four strategic sellers, which precludes extra-market communication and learning. Third, we use stochastic sampling (commonly know as Monte Carlo sampling, see Judd 1998) to generate a distribution of marketing behaviours across the sellers: given the stochastic nature of the GA, and the complexity of the genotypes and phenotypes, we use distinct random seeds to generate 50 distinct outcomes.

Computer scientists have developed machine learning, such as the GA (Holland 1976, 1992; Mitchell 1996; Goldberg 1989) and classifier systems (Holland 1976, 1992) as means of optimising — of finding the argmax of functions not amenable to calculus-based methods of solution. Social scientists have used and developed these tools (Marks 1989, 2002; Arifovic 1993) but less as optimisers, and more as generators of "adaptive plans" or "structures that perform well" in complex systems (Holland 1975, 1992), by modelling "adaptive economic agents" (Holland & Miller 1992) that interact. This chapter demonstrates a use of the GA in this spirit.

OLIGOPOLISTIC THEORY

Rivalry among retail brand managers in a market for vacuum-sealed ground coffee beans can be seen to possess characteristics that clearly reflect the oligopolistic nature of the repeated interaction: the brands are seen as imperfect substitutes by the buyers, the sales of any one brand, if stimulated by heightened marketing actions, will negatively impact on the sales of other brands, and there is no single going market price for coffee. We model Bertrand asymmetric competition among firms, competing with price (and other marketing actions) rather than quantity.

We have access to 78 weeks of supermarket-scanner market data for a city in the U.S. mid-west by supermarket chain. The marketing actions (price, coupons, aisle display, advertising) remain unchanged for seven days, from midnight Saturday, for all brands, a property that lends itself to simulation modelling on a digital computer.

One of us (Cooper) has developed a market model, Casper, which calculates, given all of the nine brands' marketing actions, the volume of sales of each brand, the brands' revenues and profits (Cooper & Nakanishi 1988).[1] The brands differ not only in the demand response of the market (each of their price elasticities of demand is distinct), but also in their costs. The brands are truly hereogeneous, as seen in Tables 1 and 2.

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

	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 2. Asymmetries of the Four Strategic Brands

Casper provides the equivalent of the one-shot payoffs for each of the brands, modelled as playing a repeated game.[2]

Although each brand manager must choose the set of next week's market actions in ignorance of the other brands' action next week, this and preceding weeks' actions are observable by all brands. So the brands can choose to remember the actions of their rivals for one, two, or more weeks. Their depth of memory is a measure of their bounded rationality: an unboundedly rational player would choose to forget nothing, and to use all remembered information including its weekly profits in deciding what marketing actions to undertake next week.

But the brand managers do not have unfettered freedom to choose their marketing actions, since the policies of the supermarket chain constrain them, in two ways. Some actions (including a price well below the "shelf price") result in much higher sales, and higher profits (the lower margins are more than offset by higher volumes of sales). The chain constrains use of these so-called "promotional" actions. First, no brand may use a promotional action set two weeks successively. Second, only one brand may use a promotional action set in any week. The chain acts as the moderator among the brand

1. We can make available the C sources for our programs and the 75 weeks of historical market data on request.
2. With up to four hereogeneous players, each facing a set of up to eight possible actions, the asymmetric $(8 \times 8 \times 8 \times 8)$ payoff matrix is much too large to reproduce here.

managers, who each propose their next week's action set and acquiesce in the supermarket's choice of which brand may promote next week.

Competing against each other, the brand managers are trying to maximise their average weekly profits. The supermarket chain is competing against other chains for sales, although we do not model this rivalry explicitly here. Instead, we model the supermarket as trying to maximise "total category volume" of coffee sales. The reason is that coffee is one of many supermarket categories, but one that might attract more customers to the chain, and so help to sell higher volumes across many categories. We model supermarket moderation in several ways, as discussed in detail below.

The competition among brand managers is asymmetric, because each of the brands is distinct, with distinct price elasticities of demand, distinct unit costs of provision, and distinct responses to the market. Moreover, solution of the Nash equilibrium of the one-week game, let alone solution of Nash equilibria in the repeated game, is not amenable to calculus-based, closed-form techniques.[3] There are nine brand rivals in the chain we focus on, although only four are engaged in what we might call a "rivalrous dance" by altering their marketing actions every week. Figure 1 shows the behaviour of the three major strategic brands, and one minor one.

There are two main purposes of our research. First, we wish to calibrate and validate our model's behaviour to the historical data. To this end, we use the asymmetries implicit in Casper to model the brands' sales, revenues, costs, and profits in any week, given all brands' market actions that week. We allow the model to run for 50 weeks, with up to four "strategic brands" altering their marketing actions from week to week, in response to the state of the market (defined as the set of all players' marketing actions) the previous week. We look for several measures of the simulated competition: weekly profits, weekly Total Category Volume of coffee sales, and the marketing actions employed by the four strategic players.

The marketing actions include price, coupons, aisle display, and flier advertising. Historically, brands' prices varied from \$1.50/lb to \$3/lb, with promotional prices below \$2.25/lb. Coupons reduce the price paid at check-out, and are measured by percentage of stores in the chain that distribute coupons for that brand that week. We net the impact of coupons out of the retail price to simplify the action space. Similarly, aisle display and flier advertising are reported as percentage of stores in the chain that include them for any brand in any week. In practice, as discussed above, the store permits only one brand to promote itself any week, and we see a consistent pattern in coupons, aisle display, and flier advertising: only one promoted brand per week.

We could allow the adaptive brand managers of the model to choose their price from any between 150 and 300 cents per pound, and any percentage of aisle displays and flier advertising, but in practice we believe, first, that this degree of freedom is not necessary to replicate historical performance, and, second, that the practical difficulties of simulating this (such as a huge number of degrees of freedom in the definition of "market state", and the need to execute Casper each simulated week instead of using a much faster compiled look-up table) militate against it.

Instead, we use the historical data to identify, first, four sets, and, second, eight sets of brand-specific actions which are representative of those chosen over the first 50 weeks

3. The Folk Theorem of repeated games (Fudenberg & Maskin 1986) tells us that there is a multiplicity of N.E. of the repeated game; in essence, any individually rational outcome can be a N.E. with a sufficiently low discount rate.

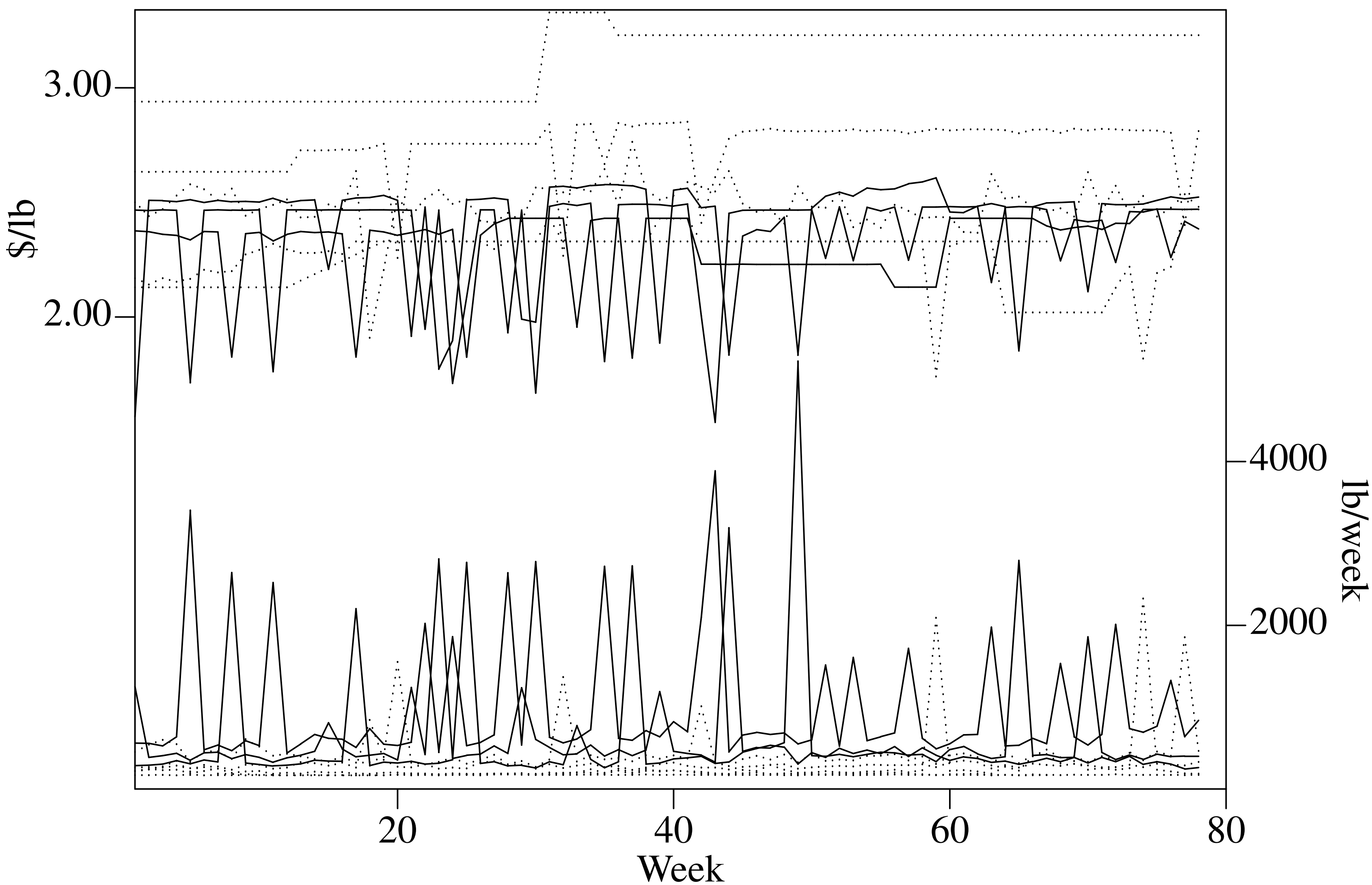


Figure 1: Weekly Prices and Sales

(Solid lines: Folgers, Maxwell House, CFON) of data. Later, we use eight action sets that are identical across the four strategic brands, and find similar results.

The second purpose of our research is to see whether our boundedly rational artificial brand managers can surpass the performance of their historical counterparts, as measured by their weekly profits, handicapped as they are by, first, simple one-week memory, and, two, constrained choice of marketing actions. Necessarily, since we do not have access to actual historical brand managers in order to pit them against our artificial brand managers in a laboratory setting, we must be content with closed-loop experiments, where are artificial brand managers respond to the unfolding history of past rivalries, but where the historical actions cannot respond to our artificial agents' actions. We argue below that both aims are attained.

The structure of the chapter is as follows. After a discussion of the GA, we describe our historical market data, and then describe the results of a set of computer experiments, as we increase the number of strategic brands from three to four, and the number of possible marketing actions per brand from four to eight. We present the open-loop results of playing our best co-evolved artificial brands against history, and introduce the Holyfield-Tyson effect of pitting more evolved agents against less evolved agents. We discuss the implications of our results for insights into Managerial Learning.

BORROWING FROM NATURE: THE GENETIC ALGORITHM

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.

We have been using versions of the GA since 1988 to explore oligopolistic behaviour.[4] As we describe above, we model the artificial brand managers as stimulus–response automata, in effect, where the stimulus is this week's market state (defined by the marketing actions of all players, and particularly the four strategic brands), and the response is the brand's proposed market actions next week. The eventual market actions per brand are the outcome of a moderating process performed by the supermarket chain, responding to the four proposals of the brand managers.

We use the GA to search simultaneously for better automata for each of the four strategic brands, using their weekly profits as a measure of performance or fitness. Each brand manager is modelled as a binary string. If there are eight possible marketing actions to choose from (correlating aisle display and flier advertising with promotional prices), then we can use three bits on the string to code for next week's marketing action.

---

4. Differentiated Bertrand oligopolistic competition closely resembles an asymmetric *n*-person Prisoner's Dilemma (Fudenberg & Tirole 1992).

How many triples are sufficient for the model? With four strategic players, each with eight possible marketing actions, there are a^{mp} possible states (Midgley et al. 1997), where a = the number of actions (8), m = the number of weeks remembered (1), and p = the number of strategic players (4), a total of 4,096 possible states, each state mapping to a triple of bits on the artificial player's bit-string "chromosome", which requires each string to be 12,288 bits long. Adding an additional 12 bits for the "phantom memory" at the first of the 50 weeks (to endogenise the initial conditions of the brand's belief in the previous week's market state) gives us 12,300 bits per string. 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.

As is well known (see Goldberg, 1989, Mitchell 1996, or the second edition of Holland 1992), the GA borrows from our understanding of evolution to search for solutions to problems not easily solved otherwise. An initial population of solutions is generated; the fitness score of each individual is determined; a subset of individuals is elected to be the "parents" of the next generation; the "crossover" of pairs of parents is simulated; and each bit is flipped from zero to one or vice versa ("mutated") with a small probability (here 1%). The fitness of each member of the new population is determined. And the process repeats until convergence.

The GA has been used by engineers as an optimisation tool. Social scientists have used it in a slightly different way: as a means of simulating co-evolution. In our model, each brand manager learns from its rivals' behaviour, and from its rivals' responses to its own actions. This mutual leaning means that the competitive environment changes, even as each artificial brand manager learns to compete more effectively. As a result, there is no necessary increase in weekly profits, even as the GA winnows the succeeding generations of their worst performing strings.

Co-evolution requires a separate population for each of the strategic players.[5] A single population would allow extra-market communication and learning to occur via the genetic operations of selection and cross-over. Not only would this be illegal under antitrust laws, but such social learning (Vriend 2000) is not what we want to model. Necessarily, four separate populations requires a much more complex GA program, but only a co-evolving GA is appropriate. We extensively rewrote the GA software (GAucsd, based on John Grefenstette's GENESIS package) (Schraudolph & Grefenstette 1992) to allow the simultaneous simulation of up to four populations of agents (modelled as bit strings).

We use a population size of 25, each string being 12,300 bits long, with four populations.[6] This is a non-trivial simulation, but we manage to obtain 2,500 generations, each of 5.5 million weekly interactions, every 50 minutes on a Mac G5 dual-2Ghz Unix workstation.

5. Were our players identical, we would have a symmetric game, and could follow the modelling simplification of Yao & Darwen (1994), as many computer scientists have done. But our players are not identical: their identity matters, as seen in Tables 1–4.

6. The GA parameters include: Crossover Rate = 13.0, Mutation Rate = 0.01; see Schraudolph & Grefenstette 1992.

THE HISTORICAL DATA: THE RETAIL GROUND-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. Table 2 presents further data on the heterogeneity of the strategic players: their own-price elasticities of market share and their Average Variable Costs (AVC). Figure 1 shows the historical prices (top half) and quantity of sales (bottom half) by brand over 75 weeks. The solid lines map the prices and sales of the three strategic brands, Folgers, Maxwell House, and CFON; the dotted lines map the other brands. The data are aggregated on a supermarket chain. As mentioned above, each marketing action comprises four "marketing instruments":

1. prices (the price that week of the brand);
2. flier features (the percentage of stores in the chain featuring the brand's item in their distributed advertising);
3. in-store aisle 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 adjusted the price in any week by the percentage of coupons distributed.

COMPUTER EXPERIMENTS

We model the brand managers as artificial agents. The computational experimenter can 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, are asymmetric).

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

First Results

This chapter builds on work reported in Midgley et al. (1997). There we considered the three most interactive players in the market: Folgers, Maxwell House, and Chock Full O Nuts (CFON). We allowed each agent four action sets, as derived from an analysis of their historical prices and other marketing actions. Table 3 shows the four possible action sets for each of the three agents.

	Folgers			Maxwell House			CFON		
A	P ($/lb)	F (%)	D (%)	P ($/lb)	F (%)	D (%)	P ($/lb)	F (%)	D (%)
p_1	1.87*	95*	69*	1.96*	95*	69*	1.89*	100*	77*
p_2	2.07	83	0	2.33	83	0	2.02	100	65
p_3	2.38	0	0	2.46	0	0	2.29	0	0
p_4	2.59	0	0	2.53	0	0	2.45	0	0

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

TABLE 3. The Four Sets of Actions of the Three Strategic Brands

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 co-evolution 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 asymmetric $4 \times 4 \times 4$ payoff matrices for the three strategic players. 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) (Table 4).

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

TABLE 4. The Fixed Prices of the Other Six Brands

With one-week memory, the agents were 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 played a 50-round game with each possible combination of the other two players. The GA used 25 mappings (or strings) per population for each agent. Therefore, testing each generation required 8125 50-round games, or 325 games per string per generation. Each agent had complete information of all previous actions in each 50-round game, but not others' weekly profits (payoffs).

Figure 2 shows three 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 Pricer (EDLP); Maxwell House exhibiting Wide Pulsing (WP). In Pattern 3, CFON is exhibiting Promote to the Max (PttM).

Consult Midgley et al. (1997) for a discussion of the patterns of behaviour of the unconstrained and constrained brands, and the issue of demand saturation over time that the single-week estimates of Casper evoke. After constraining the brands (as discussed above) and accounting for demand saturation, our three-brand, four-action model generates patterns of behaviour similar to Figure 1: brands alternate (roughly) in pricing at p_1, while the other two price at p_2, p_3, or p_4.

Having co-evolved populations of each of the three strategic agents over one

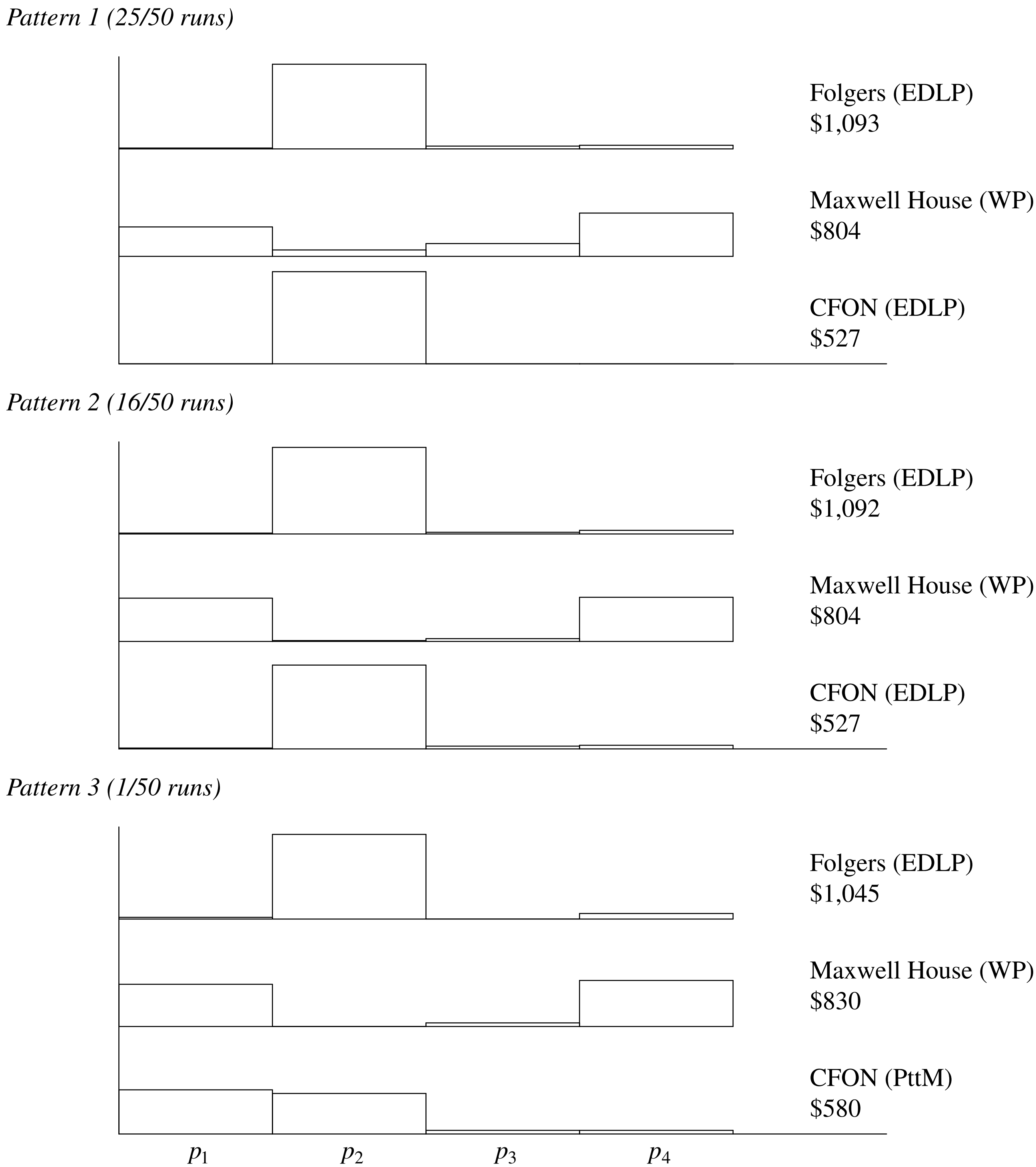


Figure 2: Three Agents, Four Actions.

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.

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.[7] 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 3 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. Surprisingly, the total weekly profits of the first three brands rise when a fourth player is introduced, at least for the 40-odd patterns of Figures 2 and 3. What these simulations demonstrate is that a small player (as measured by market share) isn't necessarily insignificant strategically. In Pattern 1, Maxwell House is exhibiting High Pricer (HP), and in Pattern 3, Shelf Price (ShP).

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-

---

7. 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.

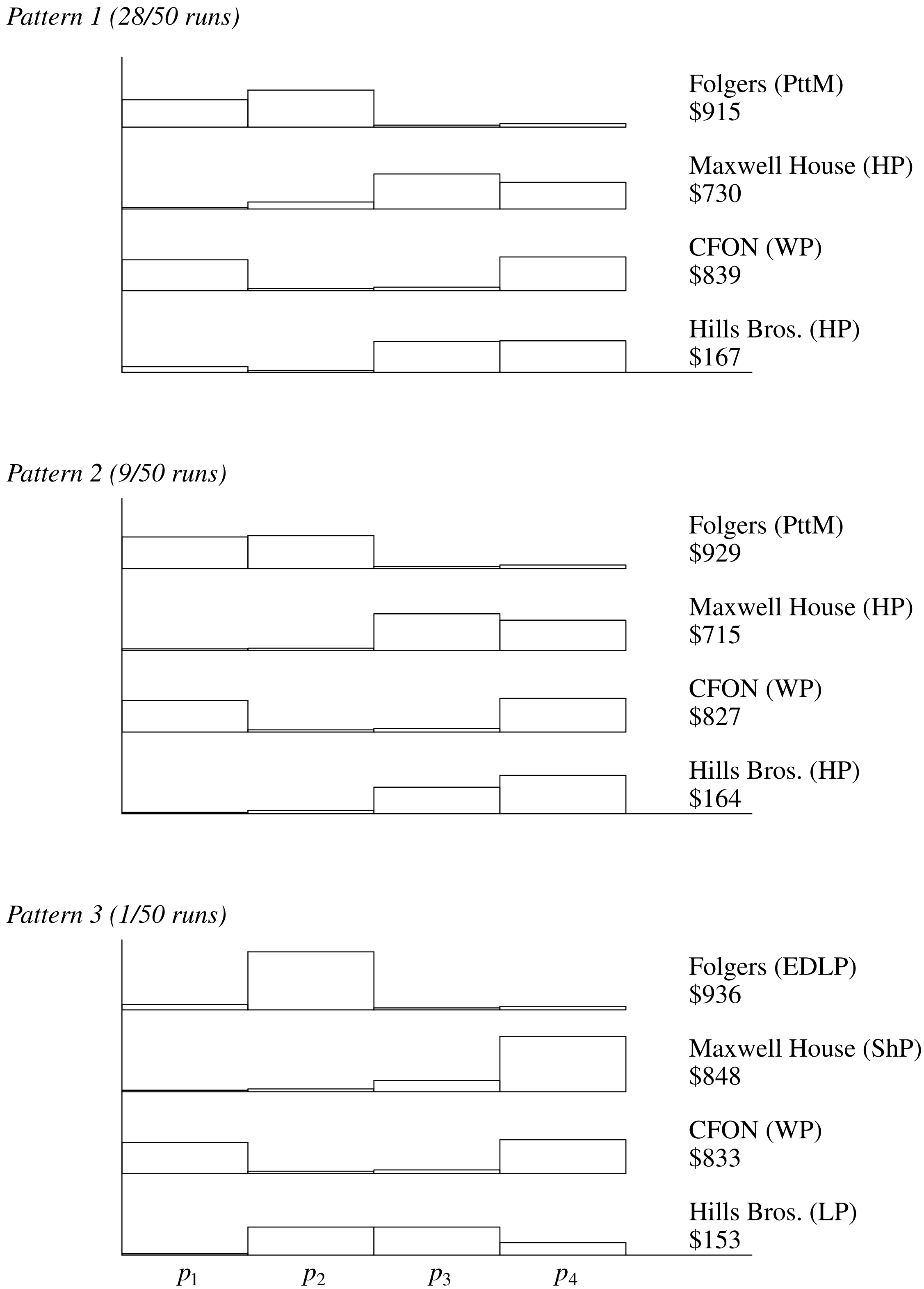


Figure 3:, Four Agents, Four Historical Actions — Hundredth Generation

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.[8]

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 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			Hills Bros.		
A	P ($/lb)	F (%)	D (%)	P ($/lb)	F (%)	D (%)	P ($/lb)	F (%)	D (%)	P ($/lb)	F (%)	D (%)
p_1	1.62*	67*	67*	1.60*	97*	97*	1.64	0	0	1.86*	100*	74*
p_2	1.83*	97*	96*	1.87*	94*	91*	1.89*	97*	97*	1.91	0	73
p_3	1.96	0	0	2.06*	88*	76*	1.89*	98*	29*	1.95*	100*	87*
p_4	2.03*	79*	77*	2.33	79	0	2.01	0	0	2.09*	100*	0*
p_5	2.04*	85*	0*	2.38	54	0	2.02*	97*	62*	2.19	0	0
p_6	2.22	96	33	2.52	0	0	2.31	0	49	2.42	0	0
p_7	2.57	0	0	2.53	0	53	2.33	0	0	2.49	0	100
p_8	2.78	0	0	2.59	0	13	2.49	0	0	2.56	0	14

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

TABLE 5. Four Brands: Sets of Eight Possible Marketing Actions.

Figure 4 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 black diamonds ♦ in the figures correspond to the asterisks in Table 5: actions subject to store moderation.)

After four generations, starting from a uniform distribution of actions (because the bit strings are chosen randomly to begin with, apart from filtering against the actions of promoting two weeks in succession), 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 eight possible actions used frequently: agents have *co-learnt* the two or three actions that are most profitable, given others' behaviour. The actions are brand-specific.

Specifically, with three strategic agents: CFON is pulsing between Shelf Price (high) and Promotional 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. CFON is pulsing with two action: wide or narrow.

From a 50-run Monte Carlo simulation of four agents and eight possible actions, we observe in Figure 5 for 44 runs that the four agents exhibit different behaviour: Folgers and CFON show Wide Pulsing, from high to low, promotional prices (indicated by the

8. 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.

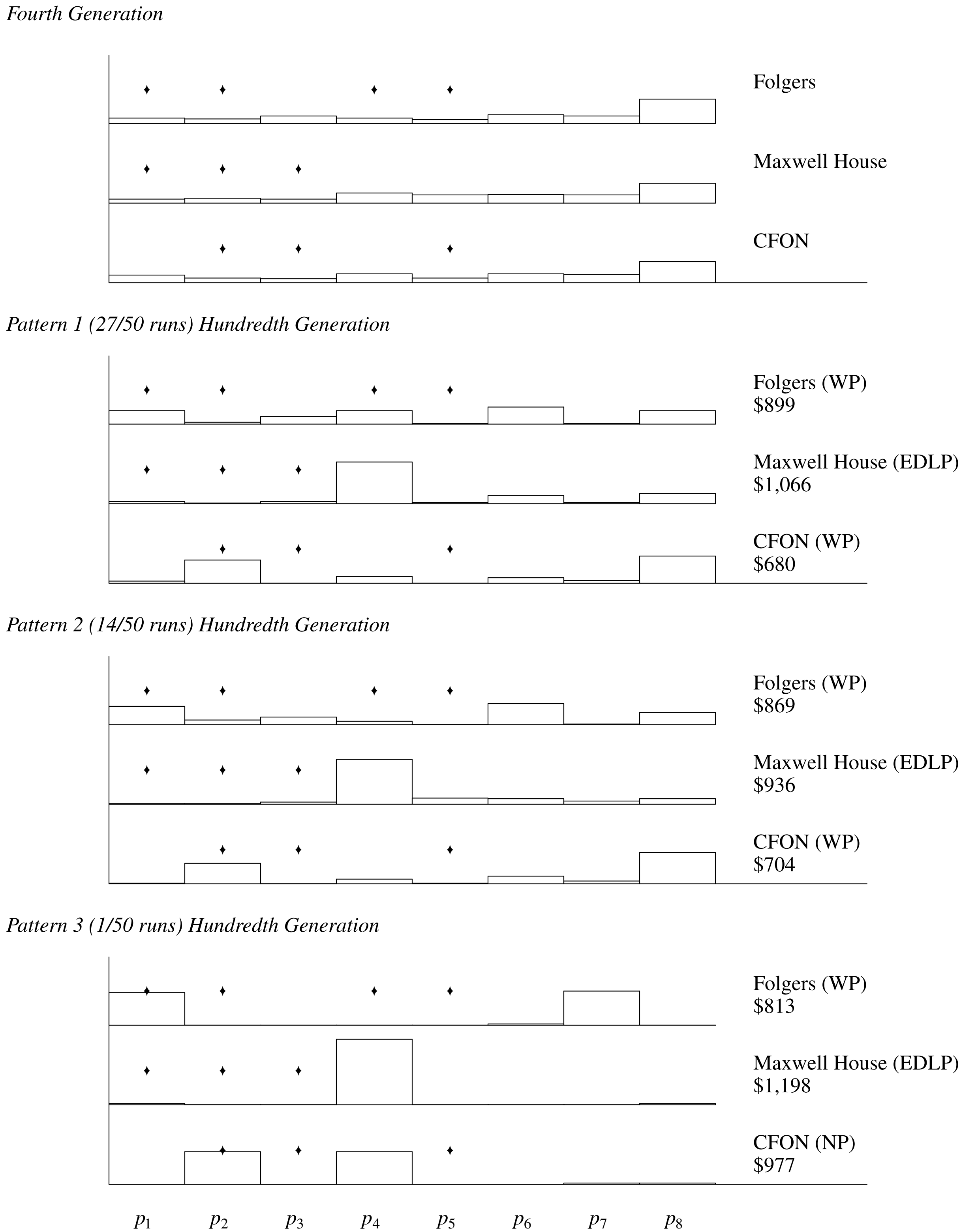


Figure 4: Three Agents, Eight Historical Actions

black diamonds), but Folgers, with 42% of its actions promotion (of a possible maximum of 50%) is almost Promoting to the Maximum, whereas CFON is promoting only 22% of the time; Maxwell House shows High Pulsing, seldom (15%) promoting at low prices; and Hills Bros. shows Shelf Price (p_6) or higher, promoting only 8% of the time.

Pattern 1 (44/50 runs)

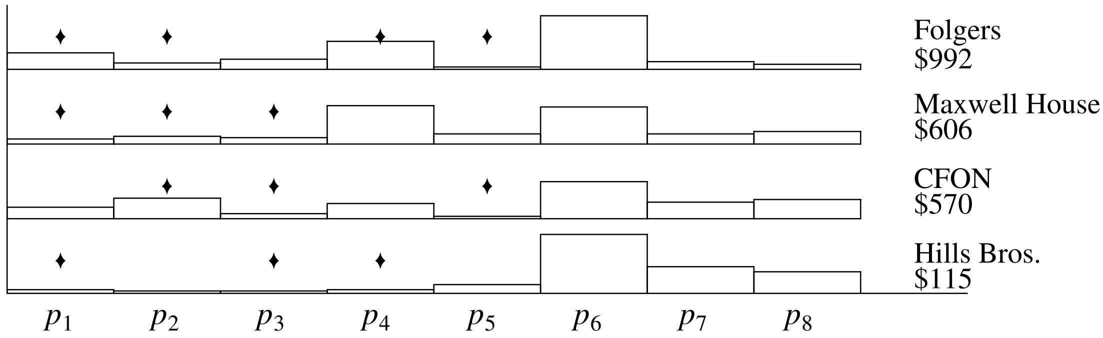


Figure 5: Four Agents, Eight Historical Actions — 2500th Generation

Overall, we can say that, with the eight possible actions of Table 5, a greater degree of homogeneity emerges, with 44 of 50 Monte Carlo runs being identical. Moreover, adding a fourth strategic agent increases the degree of competition in the market, which is here reflected in lower average profits for the first three brands, as well as different behaviour.

Moderation in the runs of Figure 5 is achieved randomly (by a "zero-intelligence" chain moderator), but we explored changing this in two ways: first, by altering the possible actions of Table 5 by eliminating the lowest prices, and, second, by estimating from the historical data just how moderation was achieved and the chain's preferences across brands revealed. We do not report these experiments in detail here, but brands' profits fell, as did the volume of coffee sold.

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 co-evolutionary "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.

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 But Marks (1992) and others following have bred artificial agents against each other, a process that Marks called "bootstrapping" and biologists term "co-evolution".

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. Two questions: Do highly co-evolved players become effete? Will a naïve outperform a sophisticate?

Apart from the growth in average weekly profits, there are at least two further ways to demonstrate that the artificial natural selection has improved the agents' performances. In our earlier work we 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 co-evolved agents are learning to respond better (are truly fitter)? Previously: we considered the mean weekly profits. 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 or 4 populations.
2. Play these 3 or 4 against each other for a 50-week repeated game; note average weekly profits.
3. Allow the 3 or 4 populations to continue co-evolving via the GA.
4. After 100 generations, identify the best strings from the 3 or 4 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
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	992	606	570	115	2,284

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 or 4 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 or 3 strategic 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*	?
3 pop., 4 actions	410*b*, 468*c*	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.[9]

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. The GA was still converging at 80 generations and the results after 160 generations were no better: the GA had still not converged. We cannot confirm genetic drift as an explanation.

MANAGERIAL LEARNING

The eight-action sets per player of above 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 6 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 above. But we speculate that these results show that inter-chain competition is what our 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.

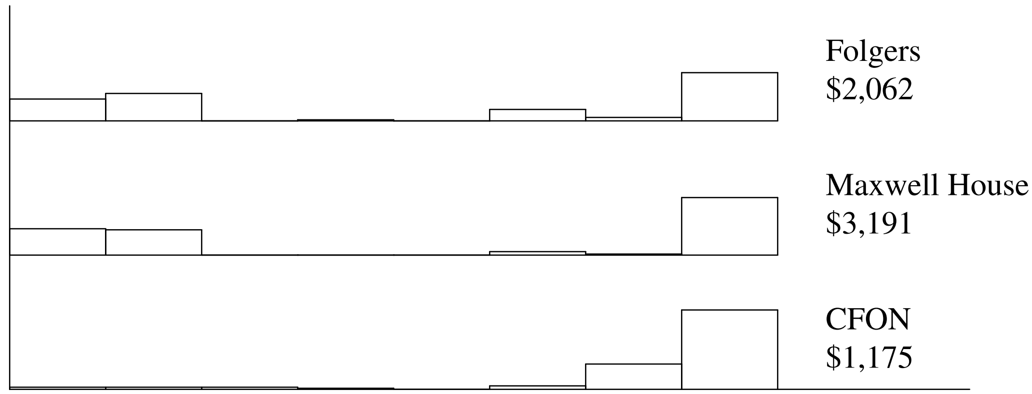
Results of three-player, eight-possible-action simulations reveal two major patterns: much higher average weekly profits, and almost no low, feature pricing, with profits earned at very high pricing. This result is seen in Figure 7, which shows the patterns for the four strategic players under the three regimes: historical frequencies of the brand managers, co-evolved agents competing against each other, and the best co-evolved agents competing against history. Notice that for Maxwell House and Hills Bros. the co-evolved 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.

CONCLUSION

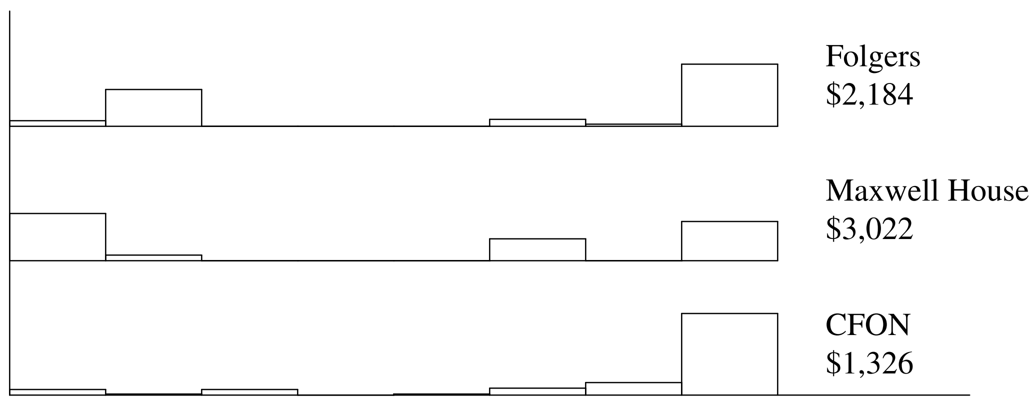
We can summarise our experiments on rivalry in a mature differentiated Bertrand oligopoly in two ways: the average weekly profits of the agents, and the patterns of actions. Table 6 summarises the average weekly profits of the four strategic brands under the different combinations of strategic brands and four- or eight-action sets (all derived from the historically observed actions of the brand managers). Figure 7 summarises the frequencies of chosen actions (eight-action sets, derived from the historically observed

9. 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)

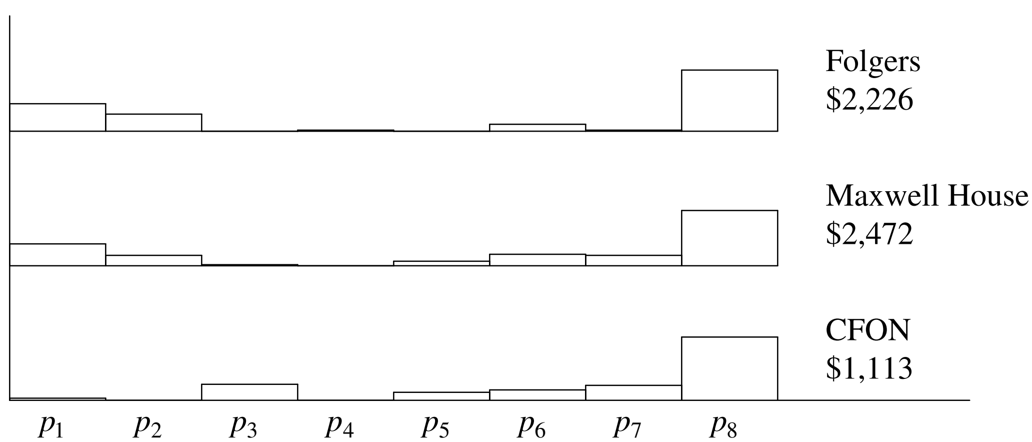


Figure 6: Three Agents, Eight-Random-Action Sets — Hundredth Generation

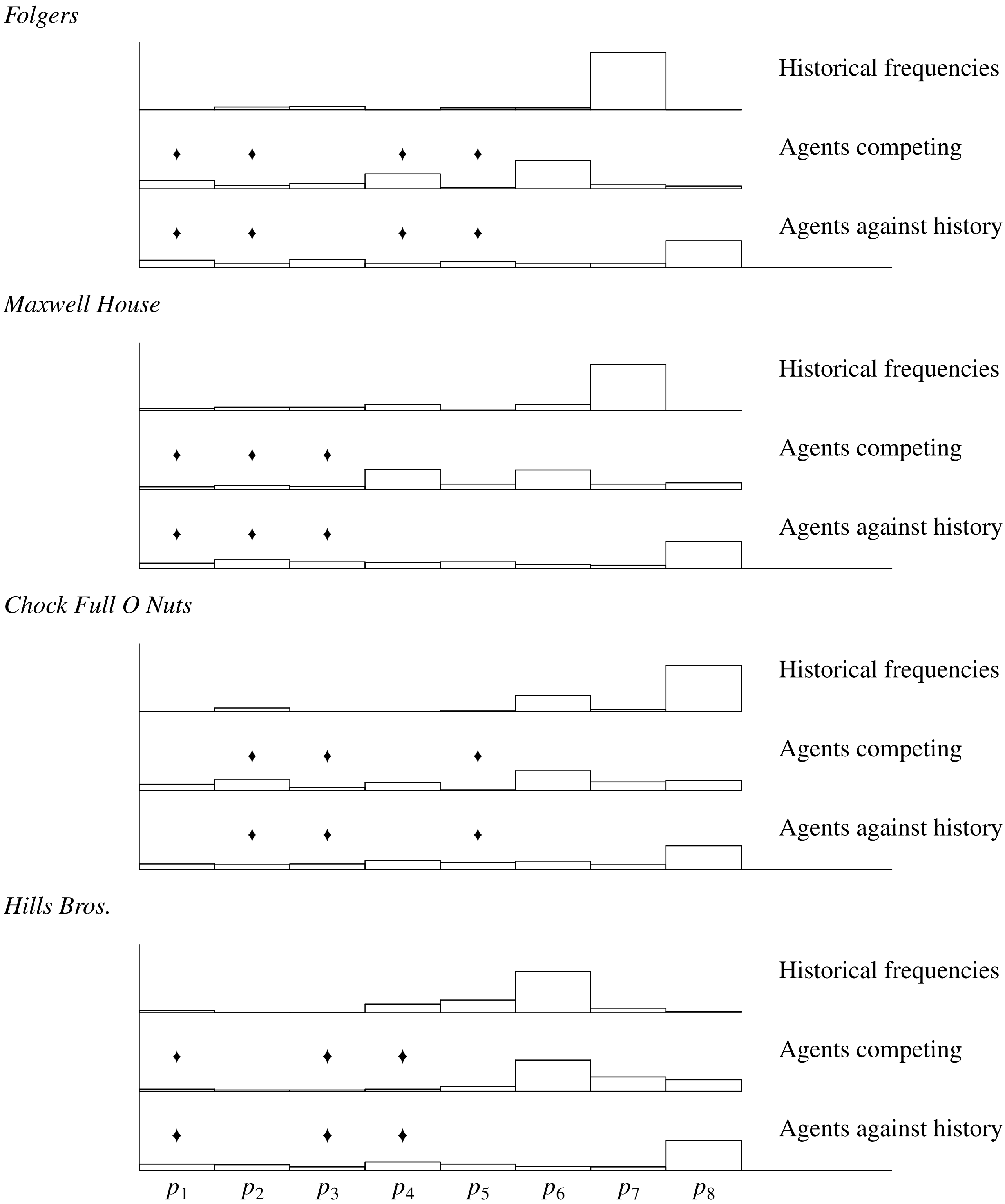


Figure 7: Comparison of Patterns

actions) under the three conditions of, first, historical actions (from Figure 1), second, co-evolved agents competing (from Figure 5), and, third, agents competing against history (playing the 50 best agents per brand against the historical actions of their three competitors). The competitive behaviour of one of our artificial brand managers (Hills Bros.) is similar to the historical frequencies, but the other three artificial brands reveal more strategic behaviour than the historical brands engaged in. For at least one brand, a simple set of possible actions and one-week memory are sufficient to simulate historical behaviour, suggesting a lack of sophistication on the part of historical brand managers. Later work will explore this issue of "zero-intelligence" behaviour (or simple heuristics) further.

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 learnt using the GA. We hypothesise that the techniques used here could shed light on the behaviours in similar asymmetric 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 (Duffy 2006). 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 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.

REFERENCES

[1] Arifovic J. (1994), Genetic Algorithm learning and the cobweb model, *Journal of Economic Dynamics and Control,* **18**(1): 3−18.

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

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

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

[5] Duffy J. (2006), Agent-based models and human subject experiments, in *Handbook of Computational Economics, Vol. 2, Agent-Based Modeling,* ed. by L. Tesfatsion & K.L. Judd, Amsterdam: Elsevier, forthcoming.

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

[7] Fudenberg D. & Maskin E.S. (1986), The Folk Theorem in repeated games with discounting or with incomplete information, *Econometrica,* **54**: 533–556.

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

[9] Goldberg D.E. (1989), *Genetic Algorithms in Search, Optimization & Machine Learning,* Reading: Addison Wesley.

[10] 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.)

[11] Holland J.H. & Miller J.H. (1991), Artificial adaptive agents in economic theory, *American Economic Review: Papers and Proceedings,* **81**(2): 365–370.

[12] Judd K.L. (1998), *Mathematical Methods in Economics,* Cambr.: MIT Press.

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

[14] Marks R.E. (2002), Playing games with Genetic Algorithms, in *Evolutionary Computation in Economics and Finance,* ed. by Shu-Heng Chen, Studies in Fuzziness and Soft Computing, Vol. 100, ed. by J. Kacprzyk, Heidelberg, New York: Physica-Verlag, pp. 31–44.

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

[16] Mitchell M. (1996), *An Introduction to Genetic Algorithms,* Cambridge: MIT Press.

[17] Schraudolph N.N. & Grefenstette J.J. (1992), *A User's Guide to GAucsd 1.4,* Technical Report CS92–249, CSE Department, UC San Diego. www.agsm.edu.au/~bobm/teaching/SimSS/GAucsd.pdf

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

[19] Vriend N.J. (2000), An illustration of the essential difference between individual and social learning, and its consequences for computational analyses, *Journal of Economic Dynamics and Control,* **24**: 1–19.

[20] Yao X. & Darwen P.J. (1994), An experimental study of *N*-person iterated Prisoner's Dilemma games, *Informatica*, **18**: 435–450.