J Evol Econ (1999) 9: 27–65

Journal of
Evolutionary
Economics


# An experimental study of adaptive behavior in an oligopolistic market game*

**Rosemarie Nagel[1], Nicolaas J. Vriend[2]**

[1]Universitat Pompeu Fabra, Barcelona, Spain
[2]Department of Economics, Queen Mary and Westfield College, University of London, Mile End Road, London E1 4NS, UK (e-mail: n.vriend@qmw.ac.uk)

**Abstract.** We consider an oligopolistic market game, in which the players are competing firms in the same market of a homogeneous consumption good. The consumer side is represented by a fixed demand function. The firms decide how much to produce of a perishable consumption good, and they decide upon a number of information signals to be sent into the population in order to attract customers. Due to the minimal information provided, the players do not have a well-specified model of their environment. Our main objective is to characterize the adaptive behavior of the players in such a situation.

**Key words:** Market game – Oligopoly – Adaptive behavior – Learning

**JEL-classification:** C72; C91; D83

* We wish to thank Reinhard Selten for encouraging and discussing the experiments, and Klaus Abbink, Joachim Buchta, Cornelia Holthausen, Barbara Mathauschek, Michael Mitzkewitz, and especially Abdolkarim Sadrieh for their indispensable assistance in discussing and organizing the experiments. Steven Klepper helped us improving the paper a great deal by insisting with a number of pertinent questions and suggestions. We are grateful for comments and discussions concerning previous versions of this paper to Antoni Bosch, John Miller, Greg Pollock, Phil Reny, Al Roth, Arthur Schram, Ulrich Witt, and to seminar and conference participants in Pittsburgh, Ames, Long Beach, Barcelona, Augsburg, Trento, Amsterdam, Toulouse, Marseille, London, St. Louis, and Paris. The experiments were made possible by financial support from the Deutsche Forschungsgemeinschaft through Sonderforschungsbereich 303 at the University of Bonn. Stays at the University of Pittsburgh, and the Santa Fe Institute, and financial support through TMR grant ERBFMBICT950277 (NJV) from the Commission of the European Community are also gratefully acknowledged. The usual disclaimer applies.

*Correspondence to*: N.J. Vriend

## 1 Introduction

We consider an oligopolistic market game, in which the players are identical competitors in the same market. In each period, they decide how much to produce of a perishable and homogeneous consumption good, and they decide how many advertising signals to send into the population in order to attract customers. The firms know the parameters of the production and signaling technologies,[1] as well as the fixed price of the good. After every period, each firm observes only its *own market outcomes*. No further information about the environment is given.

Given the minimal information, even a rational player is not in a position to maximize his profits using standard optimization techniques. Following Savage's (1954) terminology, he is in a '*large world*', as opposed to the '*small world*' to which Subjective Expected Utility theory applies. In a large world, the agent's situation is ill-defined in the sense that he does not have a well-specified model of his environment. Hence, instead of *deducing* optimal actions from universal truths, he will need to employ *inductive* reasoning, i.e., proceeding from the actual situation he faces. In Savage's terminology, this is the '*cross that bridge when you meet it*' principle, which is also known as adaptive behavior.[2]

Studying this adaptive behavior in a large world is the main motivation for our simplified experimental setup. To create a large world in a relatively simple oligopoly game, we abstract from the process by which the price is determined, and from the determination of the market demand at that price level. This is perfectly compatible with a standard downward-sloping market demand curve. Notice also that there are many markets in which goods are sold at fixed prices (whether as a result of legislation, of vertically imposed restrictive practices, or of optimizing behavior of the sellers). While a *complete* economic analysis would explain such legislation, restrictive practices, or strategies, by which the prices are fixed, our analysis focuses instead on the learning and adaptive behavior of the firms, and thus applies equally to all the possible ways in which these prices may have been determined. Given the price level, competition can then take place along many dimensions,[3] for example through advertising, and it seems more than plausible that for some of these dimensions the information that an individual firm has about its competitors is far from complete. In our model, as we will show below, the only strategic variable to compete directly with the other firms is the signaling activity. Assuming that firms do not observe the

[1] Notice that we follow the common use of the word 'signaling', and not the more restrictive game-theoretic one related to signaling games.

[2] We would conjecture that many relevant economic problems, when considered at a moderately realistic level, are large-world problems (see, e.g., Arthur, 1992).

[3] Quantity and production capacity are obvious ones. Product differentiation is another one. The quality of a good depends upon many aspects, like, e.g., a warranty, add-ons like frequent flyer miles, or an after sale service. Firms also compete using entry deterring and other restrictive practices, by their choice of technology, location, or the timing of new product lines. Further competitive variables are the firms' R&D decision (including marketing research), and their efforts to build up a reputation.

level of their competitors' signaling activity in our simple model is a first approximation of this fact.[4]

What can be learned from large world experiments that cannot be learned from small-world experiments with fuller information? It is far from certain that you can learn the way in which people behave in large worlds by studying only small worlds. It might very well be that there is no substantial difference between the two as far as the behavior of the players is concerned, but we cannot know this in advance. The only way to check this is to study large-world experiments as well (see Page, 1994, for further arguments along this line). Related to this is the observation that it might be that many apparently small worlds are in fact large worlds as perceived by the players, due to the fact that the agents' rationality is bounded (see, e.g., Simon, 1959), or that their perception is limited (see also Vriend, 1996a). One of the key advantages of laboratory experiments is that one *controls* the players' environment. Hence, it might be true that due to bounded rationality and limited subjective perceptions, some players consider even a full information set-up as a large world, but we can make sure that it is a large world for *all* players by placing some explicit simple restrictions on the information provided to the players.[5]

Our main objective, then, is to characterize the adaptive behavior of the players in such a large world. First of all, we want to characterize the overall market outcomes that result from the interaction of the adaptive players. Second, we want to know whether we can use simple models of adaptive behavior to describe the actual behavior on average. Third, we will examine the distributions of actions and outcomes over the individual players underlying the market averages. As individual behavior is very heterogeneous, we will analyze the reasons for this heterogeneity, despite the market being symmetric.

In order to put the experimental data into perspective we use the following theoretical framework. First, to obtain a game-theoretical benchmark, we derive a stationary symmetric equilibrium, assuming complete information. The second way to put the experimental data into perspective is by using a simple 2-step model of adaptive behavior. Although we will show that this 2-step model is closely related to the game-theoretic analysis, it is very different in the sense that it is based exclusively on the minimal information that the players actually have, while making only minimal assumptions about the agents' reasoning processes. The 2-step model consists of two simple processes; learning direction theory, which has been successfully applied in various experiments (see e.g., Selten and Stoecker, 1986; or Nagel, 1995), and the well-known method of hill climbing, also known as the gradient method. As we will make clear below, these two steps share the following underlying principle. The players' own actions and outcomes in the most recent (two) period(s) give the player information

[4] For a more extensive justification of this type of oligopoly model we refer to Vriend (1996b), and the references cited therein.

[5] Some other 'large-world' experiments can be found in Atkinson and Suppes (1958), Sauermann and Selten (1959), Witt (1986), Malawski (1990), Stewing (1990), and Kampmann and Sterman (1995).

about the direction in which he may find better actions. We will also use this 2-step model to analyze the differences in actions and outcomes between the players.

We expected the players on average to adapt sufficiently to their environment to discover the underlying trading opportunities, and we hoped that the simple 2-step model of adaptive behavior would indeed be able to describe the typical behavior of the players in a satisfactory way. Although we expected to find some spread around the players' average actions and outcomes, we did not expect sharp differences between the players.

How does this study of adaptive behavior fit into more traditional analyses of evolutionary economics? Schumpeterian evolutionary analysis usually focuses on the long-run evolution of economic primitives such as technologies and preferences. In doing so, it tends to abstract from the short-run economic coordination problem by assuming a Walrasian perspective. However, if we are living in a large world, then also in the short-run agents need to be *entrepreneurs*, adaptively discovering and creating trading opportunities. We believe that the outcomes of these short-run evolutionary market processes must in one way or another have consequences for the developments in the longer run, certainly if one observes systematic differences in the players' perceptions of their short-run underlying opportunities such as in our experiment. A *complete* evolutionary economic theory should consider these short-run and long-run developments in a coherent analysis, but that is beyond the scope of the current paper.

The paper is organized as follows. In Section 2, we explain how a large world looks in a small experimental laboratory. In Section 3, we present the theoretical framework within which we will analyze the data. Section 4 contains an analysis of the data, and Section 5 concludes.

## 2 The experiment: model and design

We conducted two series of experiments in the computerized experimental laboratory at the University of Bonn, one with inexperienced, and one with experienced players. Before presenting the experimental design, we will first explain and discuss the oligopoly model used. Table 1 gives the notation used throughout. Superscripts will be used for the time index, and subscripts for the identity of a firm. In addition, Table 1 gives an overview of the parameter values. The last column indicates whether the parameter value was known to the players or not. As we will explain below, in addition to these parameter values, the players did not know the functional form of the demand they faced.

### *a) The oligopoly model*

A fixed number of firms repeatedly interacts in an oligopolistic market. All firms are identical in the sense that they produce the same homogeneous

**Table 1.** Notation and parameter values

| Symbol | Meaning | Value | Known |
|---|---|---|---|
| c | '*marginal*' cost production | 0.25 | yes |
| f | patronage rate satisfied consumers | 0.56 | no |
| g | price minus '*marginal*' cost production | 0.75 | yes |
| k | '*marginal*' cost signaling | 0.08 | yes |
| m | # firms | 6 | no |
| n | # consumers | 712 | no |
| N | total # agents | 718 | no |
| p | price of the commodity | 1.00 | yes |
| Π | profit | – | own |
| q | demand directed towards a firm | – | own |
| Q | aggregate demand | – | no |
| s | # signals sent by a firm | – | own |
| – | maximum value for s | 4999 | yes |
| S | aggregate # signals all firms | – | no |
| $S_{\text{-}}$ | aggregate # signals other firms | – | no |
| V | value | – | no |
| x | sales | – | own |
| z | production | – | own |
| – | maximum value for z | 4999 | yes |
| – | # periods | ± 150 | no |

consumption good, using the same technology (see below). Time is divided into discrete periods. At the beginning of each period, each firm has to decide how many units of the perishable consumption good to produce. The production costs per unit are constant, and identical for all firms. The production decided upon at the beginning of the period is available for sale in that period. The firms also decide upon a number of information or advertising signals to be sent into the population, communicating the fact that they are a firm offering the commodity for sale in that period. Imagine the sending of letters to addresses picked randomly from the telephone directory. The costs per information signal sent to an individual agent are constant, and identical for all firms. The price of the commodity is fixed, invariant for all periods, and identical for all firms and consumers. The choice of the number of units to be produced, and the number of information signals to be sent, is restricted to a given interval.

Consumers in this economy are simulated by a computer program. In each period, when all firms have decided their production and signaling levels, consumers will be '*shopping*', with each consumer wanting to buy exactly one unit of the commodity per period. In fact, the consumer side is represented by the fixed, deterministic demand function given in equation [1].

$$q_i^t = \text{round}\Big(\text{trunc}\big[f \cdot x_i^{t-1}\big] + \frac{s_i^t}{S^t} \cdot \Big[1 - \exp\Big(-\frac{S^t}{N}\Big)\Big] \cdot \Big[n - \sum\nolimits_{j=1}^{m} \text{trunc}\big(f \cdot x_j^{t-1}\big)\Big]\Big) \qquad [1]$$

$$(\quad \text{I} \quad) + (\text{IIa}) \cdot (\quad \text{IIb} \quad) \cdot (\quad \text{IIc} \quad)$$

where

I = demand directed towards firm i by patronizing consumers
IIa = proportion of signals from firm i in aggregate signaling activity
IIb = proportion of individuals reached by one or more signals
IIc = number of 'free', i.e., non-patronizing, consumers
IIa · IIb · IIc = demand directed towards firm i as a result of current signaling

In each period, a fixed fraction of the number of customers satisfied by a given firm during the last period will *patronize* that firm [part I of eq. (1)].[6] The remaining consumers who received at least one signal (part IIc multiplied by IIb) are split between the firms, according to the firms' signaling activity relative to the aggregate signaling in the market (part IIa). Notice that when all firms signal very little, not all consumers will be reached by an information signal, implying that not all consumers will actually be present in the market. Hence, although all signaling has the form of informative advertising, and there is no persuasive signaling (see Stiglitz, 1993), one can distinguish two different effects: a business stealing effect, and an effect on the total demand in the market (see also Petr, 1997). In Vriend (1996b), in a closely related model, we consider explicitly a process of sending, receiving, and choosing individual signals, and show that this leads to a demand function faced by the individual firms that may be approximated by a Poisson distribution. The deterministic function given above equals the expected value of such a Poisson distribution. At the end of the period, all unsold units of the commodity perish. Notice that a firm cannot sell more than it has produced at the beginning of the period. Hence, a firm's profit in period t is given by:

$$\Pi_i^t = P \cdot x_i^t - c \cdot z_i^t - k \cdot s_i^t, \text{ where } x_i^t = \min[z_i^t, q_i^t] \qquad [2]$$

## *b) Information for the individual players*

We now sketch the information available to the individual players, distinguishing technology, market, and experience factors. Appendix A presents the instructions given to the players, and Table 1 above summarizes which parameter values were known and which not.

*The technology*. The players know that they are identical firms, producing the same homogeneous consumption good, using the same technology. Both the production and signaling technology are common knowledge, and the same applies to the price of the commodity. As to the fact that the choice interval for production and signaling is limited, the players were told that "*this is due only to technological restrictions*, *and has no direct economic meaning*".
*The market*. The players were told that the consumers in this economy would be simulated by a computer program. They did *not* receive the

[6] See also Keser (1992) for duopoly experiments with demand inertia.

specification of the demand function [eq. (1)], and they did not know the number of competing firms,[7] the number of consumers, or the parameter value of the demand inertia. Instead they were given the following general picture of the consumption side. Each consumer wants one unit of the commodity in every period, and so a consumers has to find a firm offering the commodity for sale, and that firm should have at least one unit available at the moment he arrives. The participants were given two considerations concerning consumers' actions. First, a consumer who has received an information signal from a firm *knows* that this firm is offering the commodity for sale in that period, and second, consumers who visited a certain firm, but found only empty shelves, might find that firm's service unreliable. On the other hand, a consumer who succeeded in buying one unit from a firm might remember the good service, and might be more likely to come back. Participants were also told "*experience shows that, in general, the demand faced by an individual firm is below 1000*".[8]

*Experience*. At the end of the period, each firm observes only its own market outcomes, and never the actions and outcomes of the other players. Each firm knows the demand that was directed to it during the period, how much it actually sold, and its profits for that period. Sometimes the market outcomes are such that a firm makes a loss. A firm making cumulative losses is informed about these. Each firm faces a known upper limit for the total losses it may realize, and a firm exceeding this limit is declared bankrupt, with the participant removed from the session. This was known before the experiments started. The players did not know the number of periods to be played, but they knew that the playing time would be about 2½ hours. Given this minimal information, a player is not in a position to maximize his profits on the basis of a well-specified demand function. In other words, he finds himself in a 'large world', and must behave adaptively. During the instructions before the games, some players felt uncomfortable with so much 'mist', and many attempted to get more knowledge about the environment. The usual answer to those questions was 'you just don't know'.

### *c) The experiments*

In the first series, we organized 13 sessions with inexperienced players. In each session, 6 firms were competing in one market, for a total of 78 players. In two of these sessions, the players faced an upper limit of 999 instead of 4999 for their production and signaling decision variables, but these two sessions are excluded from the analysis. The remaining 11 sessions are numbered 1 to 11 throughout this paper. In the second series, we organized 5 sessions with experienced players, with again 6 firms per session, for a total of 30 players. We asked all players to return for a very similar

---

[7] There were at most 12 players at the same time in the lab, but players did not know how many parallel sessions were going on.

[8] This was done to avoid players going bankrupt in one of the initial periods, without impeding them to choose levels above 1000.

oligopoly game with experienced players in order to test whether the players also learn over time to adjust the way in which they adapt to their circumstances.[9]

Most players came from various departments of the University of Bonn. Players sat in front of personal computers, and could not observe the screens of other players. Figure A.1 in appendix A presents an example of a computer screen viewed by a given player. In the sessions with inexperienced players, we played about 150 periods.[10] There was no time limit for the participants' decisions. Each player got a fixed '*show-up*' fee, and was paid according to the total profits realized by his firm. Losses realized were subtracted from the '*show-up*' fee. The total payoff for an individual player was given by: $\alpha + (\alpha/2000) \cdot$ (points realized).[11] Observe that bankrupt players had lost their '*show-up*' fee, and hence got nothing. Each session lasted about 2½ hours, and the average payment over the 66 players was DM 24.83 ($\approx$ \$16.36).

### *d) A closer look at the game*

Besides the minimal information, there are two additional aspects of the structure of this game that are worth noting. First, there is a *positive feedback* mechanism. A fixed proportion f of a firm's satisfied customers will patronize the next period, and so firms having sold more in period t, will get more customers in period t+1. This positive feedback has two effects. First, it makes the game complicated from the individual player's point of view, and second, it may give rise to *lock-in* effects in both directions. For example, say each firm sends 927 signals, and receives 118 customers in a given period t. Of those 118 customers, $0.56 \cdot 118$ will come back in period t+1, $0.56^2 \cdot 118$ in period t+2, etc. In other words, the signals sent in a given period t lead to new customers arriving in the form of a wave, with a steep front that fades out gradually. As a result, in any period t, the demand faced by a firm is composed as follows: 52 customers are there because of a signal received in period t, $0.56 \cdot 52$ because they had reacted to a signal in period t−1, $0.56^2 \cdot 52$ because of a signal in period t−2, etc., up to 1 customer still coming back since period t−8.

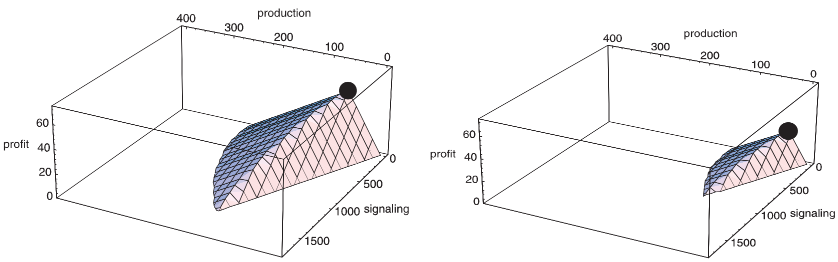
The lock-in effect can be made visible as follows. For a given period, for a given firm, one can calculate for each possible (production, signaling)-pair the immediate profits that pair would realize, taking as given the signaling activity of the other firms and the sales of all firms in the preceding period.

---

[9] An analysis of the data of the sessions with the experienced players can be found in Nagel and Vriend (1999).

[10] The sessions with inexperienced players lasted 151 periods, except for the sessions 7 an 8 (131 periods), 10 (251 periods), and 11 (201 periods). These differences are mainly due to the fact that sessions 1 to 8 were organized with two sessions simultaneously, and that the next period could only start when all twelve players had made a decision.

[11] The value for $\alpha$ was DM 10 in sessions 1 and 2, 15 in sessions 3 to 6, and 20 in sessions 7 to 11, giving an average $\alpha$ of 16.4. The values for $\alpha$ were varied in advance in an effort to keep average payoffs at a level of about DM 15/h.



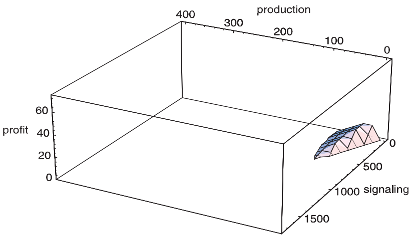




**Fig. 1.** Example immediate profit landscape: shrinking through immediate profit maximization (where • is action chosen)

Hence, we draw a 3-D '*immediate profit landscape*', showing all points that lead to positive immediate profits as an '*island*'. If, on the one hand, firms invest in order to build up a market, this island will grow. On the other hand, a firm's market may collapse if it does not signal enough. For example, it may be tempting for players to seek maximization of their immediate profits, i.e., to search for the peak in their immediate profit landscape. What would happen then? Analyzing every single instance in which an individual player had to make a decision in our experiments, it turns out that very often the global maximum of immediate profits is a corner solution with signaling at zero and production equal to the demand generated. This was the case 84% of the time in the first series of experiments. Hence, if a firm would try to maximize its immediate profits, its market will shrink away under its own eyes in most cases. An example of this effect is shown in Fig. 1, where the *dot* indicates the action chosen by the player considered, and where the other players each send 750 signals.

Figure 2 shows an example within the same environment of the opposite positive feedback effect, where a firm builds up its market by maximizing its sales subject to the condition that its immediate profits are positive.

A second aspect of this game that is worth stressing is the influence that each player's actions have on the outcomes of the other players. While one firm may try to walk up to a peak in its profit landscape, this landscape is deformed continuously by the other players who may be trying to reach their peaks. This coevolutionary process can be seen as a number of players walking simultaneously on one rubber mattress. Figure 3 shows an example, where the aggregate number of signals sent by each of the other players fluctuates from 750 to 1300 to 200. The interaction between the firms through the aggregate signaling activity shows up in the form of noise for an individual firm. If a firm has a larger immediate profit island, it will be

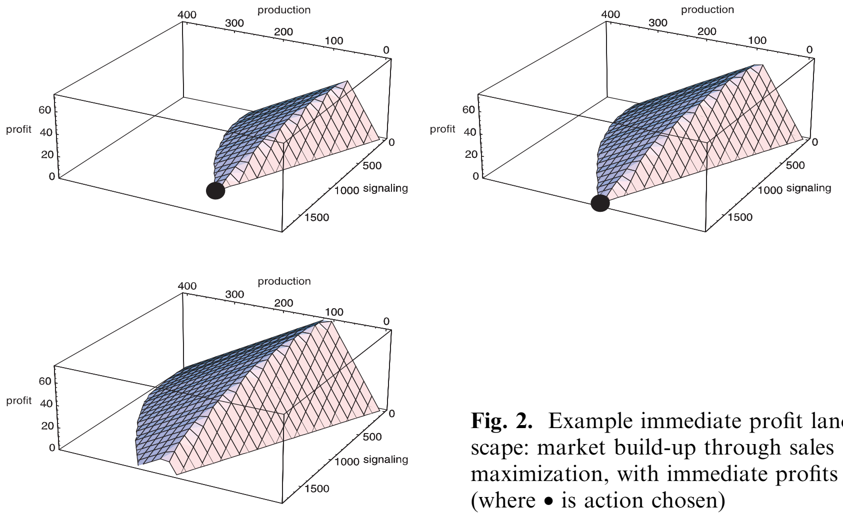



**Fig. 2.** Example immediate profit landscape: market build-up through sales maximization, with immediate profits > 0 (where • is action chosen)

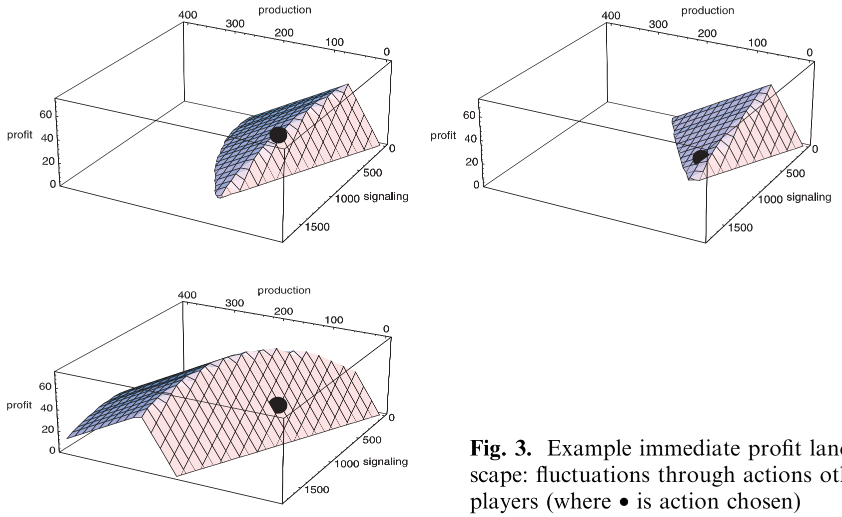



**Fig. 3.** Example immediate profit landscape: fluctuations through actions other players (where • is action chosen)

less vulnerable to this noise in the sense that it will lead less easily to negative profits. (This is because the firm's action can be farther away from the sea, and its island jumps up and down less than smaller islands.) As far as occasional negative profits induce firms to choose inactivity, this implies more positive feedback.

## 3 Theoretical framework

Our analysis of the experimental data will be structured through the following theoretical framework. In Section 3.a we present a game-theoretic analysis of the oligopoly game, and derive an equilibrium strategy. In Section 3.b we outline a simple 2-step model of adaptive behavior based on learning direction theory for the firm's production decisions, and hill climbing for its signaling decisions.

### *a) Game-theoretic analysis*

In order to obtain a theoretical benchmark, we derive the symmetric stationary optimal policy for a given player for any given period, assuming complete information about the demand function.[12] Clearly, this cannot be a normative benchmark, but merely a yardstick. Of course, other theoretical yardsticks are possible, but the stationary symmetric equilibrium for the complete information game is particularly appealing in the sense that it is a simple and well-understood one.

**Proposition 1.** In the symmetric stationary equilibrium, the signaling level for an individual firm i in a given period t is given by:

$$s_i^t = \frac{g}{k} \cdot \left(\frac{m-1}{m^2}\right) \cdot n \qquad [3]$$

and the production is simply equal to the demand thus generated.

*Proof*. We derive the equilibrium signaling level in appendix B (see Table 1 for the notation used). Given this signaling level, the demand for an individual firm is given by equation (1). Since the demand function is deterministic, the optimal production level is simply equal to that demand.

The numerical values of this equilibrium implied by the parameters of the model are a production level of 118 and signaling level of 927.

[12] The symmetry feature is justified by the fact that the firms were identical. We do not consider the optimal strategy for the incomplete information game, because the literature on *monopolies* with uncertain, but linear, demand shows that it is often too complicated not only to determine the optimal pricing strategy (in order to maximize the discounted sum of profits) but also to establish convergence as such (see, e.g., Kiefer and Nyarko, 1989). Basically, the reason is that for each action there is a trade-off between the payoff a firm gets in the form of information which may lead to future profits, and the payoff in the form of immediate profits. As Kirman (1993) argues, trying to incorporate this problem into an oligopolistic model, in which there is also strategic interaction, seems unmanageable for the moment (see also, e.g., Green, 1983; or Kirman, 1983). Notice also that in the literature on double oral auctions with private information, it is the complete information outcome that is used as the theoretical benchmark (see Davis and Holt, 1992).

*b) A simple model of adaptive behavior*

Some recent models of adaptive learning and evolutionary dynamics in the economics literature are, for example, Ellison (1993), Kandori et al. (1993), and Young (1993). Marimon (1993) discusses the basic properties of such dynamic models. In the evolutionary dynamic models mentioned, adaptive behavior is basically a one-step error correction mechanism. The agents have a well-specified model of the game, they can reason what, given the actions of the other players, the optimal action would be, completely independent of any payoff actually experienced, and they play a best-response strategy against the frequency distribution of a given (sub-)population of other players (cf. fictitious play). The evolutionary dynamics consist of a coevolutionary adaptive process, players adapting to each others' adaptation to each other ..., plus experimentation in the form of trembling. In our game, the scope of such learning techniques is limited. The agents do not have a well-specified model of their environment, and they do not know what would be the best response. Hence, the very first task for our players, is to *learn* which actions would be good.

An important advance in the theoretical economics literature on learning involves models of Bayesian updating, in which the players optimize their actions based on present beliefs about the state of nature, the types of other players and the actions of other players, while updating these beliefs using Bayes' rule. McKelvey and Palfrey (1992), for example, explain the behavior in centipede games by a learning model in which players have a common belief about the existence of altruists in the population, and a common error rate about beliefs and actions which declines over time. While this kind of model requires a high amount of rationality, there is also a deeper problem: Bayesian updating applies only to small worlds, without surprises. It is mainly a dynamic consistency requirement. The real learning question is where the priors come from. In a small world they can be reasoned backwards using Bayes' rule, but clearly, this cannot apply in a world full of surprises (see also Binmore, 1991). Hence, we cannot apply such a model since a rational player would be unable to construct a plausible probability distribution of priors concerning his environment.

Adaptive behavior and learning have become important topics in experimental economics in the last decade. Learning in experimental economics is usually defined as a systematic change of behavior over time as a function of past information. Very few studies address, in this context, the question of which kind of adaptation is optimal.[13] Some papers have focused on comparing different learning models and finding which of these models describe best average behavior (see, e.g., Camerer and Ho, 1996; Stahl, 1996; Tang, 1996; or Nagel and Tang, 1998). Not surprisingly this turns out to depend on the game being played. This is supported by a recent debate in computer science and AI, about the alleged superiority of various search and learn algorithms. Macready and Wolpert (1995) prove a

[13] An important exception is Selten et al. (1997) who classify in great depth computer strategies submitted for Cournot supergames, and find that the best strategy against actual strategies is a simple measure-for-measure strategy.

so-called No Free Lunch theorem, which basically says that no such Holy Grail can exist, and that the success of an algorithm depends ultimately on the specifics of the search problems at hand.

Since searching for the ultimate learning model does not appear to be a promising strategy, some people began to search for simple models. In reinforcement learning models (see, e.g. Arthur, 1991; Roth and Erev, 1995), which are based on the psychological literature, no knowledge of the game or any beliefs of opponents' behavior is required, but only information on the actual payoffs experienced. Roth and Erev (1995) do not try to come up with the ultimate learning model, but instead take a simple model that has some plausibility, and start asking for what games does this model give a reasonably correct description of people's behavior *on average*. While Roth and Erev (1995) focus more on average behavior, Easley and Ledyard (1993) and Selten and Stoecker (1986) seek to make predictions for individual period-to-period choices based on the plausibility of some very weak qualitative assumptions concerning individual behavior. These models of adaptive behavior do not imply that the players are aware of the optimum, but only that they are continuously engaged in a process of adaptation in the direction of better actions (see also Holland, 1992). The fact that such models are based on some common general principles, and the plausibility of weak assumptions implies not only that they are parsimonious, but also that they are coarse, and that they avoid the problem of idiosyncracy. One would expect more specific learning models to share many of the qualitative conclusions of these simple models. The model of adaptive behavior that we will use fits into this approach.

As shown in the formal game-theoretic analysis, in case of complete information, the only choice variable for a firm is the number of signals to be sent, whereas production should be simply adjusted to the demand generated by these signals. This suggests a 2-step decision problem for the players in our experiment. The *first* step concerns the *number of signals* to be sent, while the *second* step adjusts the *production level* to the level of the demand generated. As we will see below, just as in the game-theoretic analysis, in this 2-step model the two-dimensional decision problem is basically reduced to one dimension, since production just tracks the observed demand. We will first analyze this second step.

## Production: learning direction theory

Given the demand generated by a players' signals sent in the current and previous periods, the production level that would yield the highest profit would be equal to this demand. We conjecture that the players use a simple algorithm to achieve this. This is sometimes known in the experimental literature as '*learning direction theory*' (see, e.g., Selten and Stoecker, 1986; or Nagel, 1995). It is perhaps best illustrated by the following example given in Selten and Buchta (1994): "*(C)onsider the example of a marksman who tries to shoot an arrow at the trunk of a tree. If he misses the trunk to the right, he will shift the position of the bow to the left and if he misses to the left, he will shift the position of the bow to the right. The marksman looks at his experience from his last trial and adjusts his behavior according to a simple*

**Table 2.** Predictions learning direction theory

| If | | Then |
|---|---|---|
| (1) | $\text{production}_t < \text{demand}_t$ | $\text{production}_{t+1} \geq \text{production}_t$ |
| (2) | $\text{production}_t > \text{demand}_t$ | $\text{production}_{t+1} \leq \text{production}_t$ |
| (3) | $\text{production}_t = \text{demand}_t$ | n.a. |

*qualitative picture of the causal relationship between the position of his bow and the path of the arrow*." (p. 9). Given an action, and the corresponding feedback from his environment, it is assumed that the player has enough knowledge of the structure of the game and the payoff function to reason in which direction better actions could have been found (see also Selten, 1997). Notice that the feedback is not necessarily the specific value of the payoff generated. The player is supposed to move directly in his choice parameter space, but it is not necessary for learning direction theory to be applicable that a player knows exactly where the optimum is. Often only the direction is known. Therefore learning direction theory concerns only a qualitative learning mechanism.[14] Notice that although, on the one hand, the theory offers only a general qualitative prediction, it is, on the other hand, very precise in the sense that it predicts a player's action on the basis of his most recent action *alone*.

In our game, this direction learning mechanism can be applied as follows. If a firm faced more demand than it had produced, it *knows* that a higher production level would have led to higher profits. And if a firm faced less demand than it had produced, it *knows* that a lower production level would have led to higher profits. Therefore, in our model, learning direction theory would lead to the *predictions* given in Table 2. Notice that if production and demand were equal, the theory does not predict the direction of the change in production. Remember that, given the 2-step model (setting signals and adjusting production), these predictions are for a *given* demand level. Clearly, as we will analyze below, the demand depends upon the signaling level. Therefore, here we only consider those cases in which the players did not move in the opposite direction with their signaling level in order to induce a demand change.[15]

Under learning direction theory, the reasoning of the players is supposed to be boundedly rational in that it only considers what *would have been* a better action, that is, it considers actions *ex post*. In our formal analysis we explained that the demand was generated by a fixed deterministic demand function. Since this was not known to the players, there was subjective uncertainty. The problem for the players is not so much to maximize their

[14] Notice the similarity with supervised learning algorithms (see Vriend, 1994, for a discussion). With supervised learning it is not assumed that the player himself knows where the better actions are, but it is a supervisor who tells the player where the optimal action would have been. Also most supervised learning algorithms assume a gradual change in the right direction only.

[15] That is, if an increase in production is predicted there should be no decrease in signaling, and the other way round. This condition was satisfied in 63% of the cases.

*ex post* profits, but to maximize their expected *ex ante* profits. If demand is uncertain, and rationing is not all-or-nothing, some overproduction may be profitable, that is, the production that maximizes expected profits may be higher than the expected demand. Given the signaling level, the demand q faced by an individual firm is a stochastic function with p.d.f. f[q]. Hence expected profit E[Π] for a given output level z is: $E[\Pi(z)] = p \cdot \{\Sigma_{q=0}^{z} q \cdot f[q] + z \cdot \Sigma_{q=z}^{\infty} f[q]\} - c \cdot z - k \cdot s$. As can be easily shown: $\Delta E[\Pi(z)]/\Delta z = p \cdot (1 - F[z]) - c$. Hence, expected profit is maximized when $F[z] = 1 - c/p$. That is, if $c/p < 0.5$, as was the case, then we have $F[z] > 0.5$ at the optimal production level. In other words, the *ex ante* optimal production level is higher than the *ex post* average demand. We predict the players to recognize this in our experiment, and hence expect a bias towards 'overproduction' relative to the predictions of learning direction theory.

## Signaling: hill climbing

As noted above, the adaptation of the production level is assumed to take place for a given demand level. Since this demand is generated eventually by the signals sent, it is time to turn to an analysis of the number of signals sent. Learning direction theory cannot predict much with respect to signaling. In the case where demand is higher than production, a firm knows that a lower signaling level would have given higher profits, but it does not know what the optimal signaling level would have been. However, in case production is higher than demand faced, a firm does not even know whether a higher signaling level would have led to higher profits. Perhaps even lower signaling levels would have given higher profits. Also, when the demand faced by a firm equals its production, it does not know in which direction to adjust its signaling. As we showed in Section 2, a player's opportunities could be depicted as a hill. The objective of a player is to find the top of the hill, but he does not know what the hill looks like, and the hill may be changing all the time. A simple way to deal with this problem would be to start walking in one direction, and if one gets a higher payoff, one continues from there; otherwise one goes back to try another direction. Eventually one should reach the top.[16] We conjecture that the players' adaptive behavior in signaling space can be described by such a *hill climbing*, or gradient, algorithm.[17]

In order to explain the essence of hill climbing, and the contrast with learning direction theory, let us continue the example of the marksman trying to hit the trunk of a tree. Now, assume that the marksman is blindfolded. After each trial the only feedback he gets from his environment is the level of enthusiasm with which the crowd of spectators reacts. The

[16] This might be a local top only. Simulated annealing is a more sophisticated variant of hill climbing in that it tries to avoid getting stuck at local optima. To achieve this, the algorithm accepts with some probability downhill moves, whereas uphill moves are always accepted. Since we do not have landscapes with local optima, we do not consider simulated annealing.

[17] See also Bloomfield (1994), Kirman (1993), Roberts (1995), and Merlo and Schotter (1994).

**Table 3.** Predictions hill climbing

| If | | Then |
|---|---|---|
| (1) | $\text{signaling}_t < \text{signaling}_{t-1}$ and $\text{payoff}_t < \text{payoff}_{t-1}$ | $\text{signaling}_{t+1} > \text{signaling}_t$ |
| (2) | $\text{signaling}_t < \text{signaling}_{t-1}$ and $\text{payoff}_t > \text{payoff}_{t-1}$ | $\text{signaling}_{t+1} < \text{signaling}_{t-1}$ |
| (3) | $\text{signaling}_t > \text{signaling}_{t-1}$ and $\text{payoff}_t < \text{payoff}_{t-1}$ | $\text{signaling}_{t+1} < \text{signaling}_t$ |
| (4) | $\text{signaling}_t > \text{signaling}_{t-1}$ and $\text{payoff}_t > \text{payoff}_{t-1}$ | $\text{signaling}_{t+1} > \text{signaling}_{t-1}$ |
| (5) | $\text{signaling}_t = \text{signaling}_{t-1}$ or $\text{payoff}_t = \text{payoff}_{t-1}$ | n.a. |

closer he gets to the optimum, the louder they are expected to shout. Therefore, after two trials he can compare the levels of payoff, and shoot next time in the neighborhood in which the yelling was loudest. In other words, if an action leads to a worse outcome than the previous one, it is rejected as a new starting point. Hill climbers do not use any knowledge of the structure of the game, or of the payoff function. They are myopic local improvers, walking blindly in the direction of the experienced gradient in their payoff landscape. Hence, hill climbers rely completely upon the contours of the payoff landscape, whereas direction learning takes place directly in the space of actions. A deterministic variant of hill climbing would give the *predictions* presented in Table 3.

In our experiment there is one problem with hill climbing: as we showed in Section 2, the hills may change over time, even considering constant actions of the other players. Therefore, given the *dynamics* of the demand generated by the signals sent and the patronizing customers, a player should look further ahead than his immediate profits only. We showed that players could boost their immediate profits by signaling very little, i.e., by eating into their pool of customers. But future profits are adversely affected by this action. Of all the customers satisfied in a given period, some fraction will come back 'for free' in the next period, i.e., without the need to send them a signal. A firm can also forego some current profits by investing in the buildup of a pool of customers. The higher the current sales level, the better the firm's future profit opportunities, which was visualized by a larger island in Section 2. Hence, when considering the question of how well a firm performed in a given period, one should not only look at its immediate profits, but also at the change in its current sales level. The value of serving additional customers now (besides the immediate profits) is the profit that can be extracted from them in later periods.[18] Since patronizing customers come back 'for free' (without needing a signal), the profit margin for those customers will be the price minus the unit production costs of the commodity. Formally, the lookahead payoff in a given period is: $\Pi + \Delta x \cdot (p - c) \cdot \sum_{t=1}^{\infty} f^t$.

We will consider both the basic hill climbing variant, in which the players go myopically for their immediate profits only, and the variant in which the players climb hills, taking into account their lookahead payoff. If

[18] There is also an indirect effect related to a change in the player's sales level. It will change the number of 'free' consumers for which the player's signals compete with the other players' signals. This indirect effect will be relatively small because it is spread over the six firms (they compete for the same pool of free consumers), and will be ignored here.

there turns out to be myopic immediate profit hill climbers, we would expect to find them among the firms with low production levels, since they underestimate the value of keeping their sales levels up. Notice that since the players do not know the value of the patronage parameter f, nor the exact specification of the demand function, a priori they are not in a position to calculate explicitly the altitude of their lookahead hill. But during the game they can learn about the value of looking ahead. Hence, without specifying here the exact learning mechanism through which they may have learned this value, we will consider the question of how often the players behave '*as if*' they are hill climbing, having learned these lookahead payoff values correctly.

What kind of average time pattern would this 2-step model of adaptive behavior predict? We consider an unrefined numerical model, in which we use learning direction theory for the players' production decision, and hill climbing for their signaling decision. We start with all players choosing the average production and signaling levels observed during the experiments in the first period (see below), and restrict their choices to the same domain, i.e., 0 to 4999. Players follow the learning direction theory hypotheses for production as outlined above (see Table 2). The step size is equal to $|\varepsilon|$, with $\varepsilon \simeq N(0,5)$. If their production is equal to their demand, then they do not change their demand. And as explained above, the players do not change their production level if their signaling decision for that period points in the opposite direction. For hill climbing we use the lookahead variant explained above (see also Table 3). Comparing the payoffs realized in the preceding two periods, a player takes as the new starting value in the next period that signaling level that generated the highest payoff of the two. When the payoffs in the two preceding periods are equal, the new starting value is the average signaling level in those two periods. When the signaling level is unchanged during the two preceding periods, that value will again be the starting value for the next period. In order to generate the player's new signaling level, some noise is added, which is a draw from a truncated N(0, 10) distribution.[19] All players are modeled identically, but independently, which implies that their paths may diverge over time due to the stochastic factors. In Fig. 7 (in Section 4) we present the average behavior of 11 simulated sessions with 6 players, as well as the actual experimental data.

## 4 Data analysis

We will analyze the experimental data following the theoretical framework outlined above. In Section 4.a we will compare the experimental data with the game-theoretic benchmark presented in Section 3.a. In Section 4.b we will examine the data in comparison to the predictions of the simple 2-step model of adaptive behavior presented in Section 3.b, and the modifications thereof that take into account some specifics of the oligopoly game. As we will see below, perhaps the most striking feature of the data, given that the

[19] Here the truncation was determined each time such that the new signaling value stays at the correct side of the discarded signaling value that led to the lower payoff (see Table 3).

oligopoly game as such is symmetric, is the enormous differences between the individual players' actions and outcomes. Section 4.c will explain these differences.

*a) Comparison experimental data to game-theoretic benchmark*

**Observation 1.** The average actions actually chosen by the players are close to the symmetric optimal policy, but the differences between the players are considerable. The average actions chosen by the players get closer to the equilibrium policy as they play more periods, but the differences between the individual players increase, whereas the differences between the sessions decrease.

Figure C.1.a–c in Appendix C show the time series for the average signaling, production and profits of the 66 players for the periods 1 to 131 (with these variables at zero for bankrupt players),[20] and compare this with the symmetric stationary equilibrium. We observe a steep learning curve in the beginning, which leads to profits close to the equilibrium level early on. We see a lot of fluctuations during most of the history, and at the end we observe a movement towards the equilibrium levels. Table 4 presents some '*snapshots*' of this comparison between the symmetric stationary optimal policy and the actual average actions played in the game. The numbers in parentheses are the standard deviations. For each variable we calculate two standard deviations; one based on the averages for each of the 66 individual players, and the second based on the averages per session. Notice that the variance across sessions is small, and much smaller than across subjects, especially in the last 50 periods.

Given the minimal information about their environment available to the players, they are not in a position to specify the demand function. Hence, a player is not able to maximize his firm's profits directly with standard techniques. As their problem situation is ill-defined, they must learn and behave

**Table 4.** Comparison equilibrium, averages, and standard deviations (*subjects–sessions*)

| | Sign. | (s.d.) | Prod. | (s.d.) | Profit | (s.d.) |
|---|---|---|---|---|---|---|
| Equilibrium | 927 | – | 118 | – | 14.3 | – |
| Period 1 | 864 | (1016–480) | 616 | (443–205) | −107.6 | (120.1–73.2) |
| Period 1–50 | 882 | (867–163) | 160 | (121–13) | 5.8 | (23.3–11.3) |
| Period 81–130 | 938 | (954–113) | 133 | (125–6) | 8.1 | (18.9–9.2) |

[20] Throughout the paper, unless otherwise stated, we adhere to the following policy when computing averages. When the objective is to characterize the behavior of the individual players, or the differences between (categories of) individual players, we take the averages over the periods that a player was active, i.e., until the end of the session or until he went bankrupt, whichever came first. When we want to characterize the average actions and outcomes for one or more sessions as such, e.g., to compare it with the theoretical benchmarks computed, we average over all players, taking zero values for the actions and outcomes of bankrupt players.

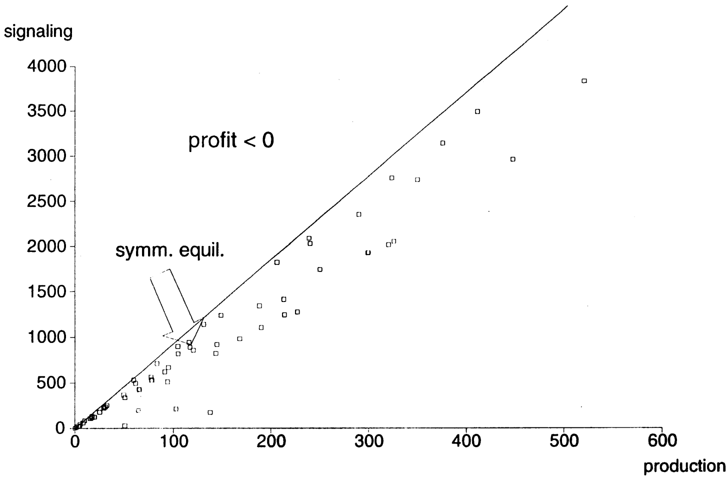



**Fig. 4.** Distribution actions, periods 81–130

adaptively. As we see, the players learn to choose actions that are on average close to a symmetric equilibrium, but there are large differences between these actions. Figure 4 shows the distribution of the individual players' signaling and production levels, averaged over the periods 81–130 (with zero values for bankrupt players). The arrow indicates the symmetric game-theoretic equilibrium given above. The straight line with slope $(p - c)/k$ serves as a benchmark. All combinations of production and signaling above it necessarily lead to negative profits. If every unit produced is actually sold, the net revenue is given by the price minus production costs per unit, multiplied by the production level: $(p - c) \cdot z$. Dividing that number by the cost of a signal (k) gives the number of signals beyond which profits can never be positive. We will return to these differences between the players in Section 4.c.

*b) Comparison experimental data to simple model of adaptive behavior*

We now turn to an analysis of the players' behavior using the 2-step model as a benchmark. We first examine the players' production decision in comparison with the predictions of learning direction theory, and then analyze their signaling decision in comparison with the hill climbing predictions.

**Observation 2.** The players change their production level in a direction that would be wrong according to learning direction theory in only 9% of the cases in which it makes a prediction. But there is an asymmetry in the success of learning direction theory between the cases in which production was too low, and those in which it was too high. This asymmetry seems related to the fact that the players are less boundedly rational than this theory assumes.

Figure 5 summarizes how far learning direction theory predicts correctly, distinguishing the cases of too high and too low production in the preceding period. If production was too low (1250 observations), learning

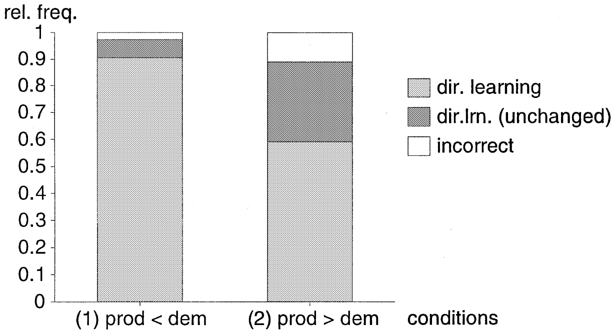



**Fig. 5.** Learning direction theory, with conditions as explained in Table 2

direction theory made a wrong prediction in only 3% of the cases. If production was too high (4571 observations), the relative frequency of wrong predictions was 11%. The weighted average of these two gives the 9% mentioned in observation 2. Production was equal to demand in only 8% of the cases (510 observations).[21]

Figure 5 clearly shows the asymmetry between these cases. As explained above, it seems that the players are more reluctant to decrease their production level when it is too high, because they understand that production should be higher than average demand; the players are less boundedly rational than learning direction theory assumes. Checking the players one by one, we find that 58 out of 66 subjects more often follow the learning direction theory hypothesis in the case in which production is less than demand, than in the case in which production is higher than demand. Also it turns out that *all* players, without any exception, on average overproduce; with the overall average production 1.20 times average demand.

**Observation 3.** Players adjust their signaling level in a way that is *wrong* according to the hypothesis of hill climbing in about a quarter of the cases. This applies equally to myopic (27%) and lookahead (25%) hill climbing. Further, the players seem only slightly inclined to looking ahead.

Figure 6a,b give the percentages of correct and wrong predictions by the hill climbing hypothesis for myopic and lookahead climbing.[22] As we see, Figures 6a, b are very similar. A first explanation is as follows. Analyzing all cases in which a player had changed his signaling level, it turns out that in 71% of the cases the payoff gradient happens to be in the same direction for myopic and lookahead hill climbing. That is, the player's immediate profits

[21] If we neglect the condition that signaling did not move in the opposite direction, considering all players together, the percentages of incorrect predictions would be 5 for the case in which production was less than demand, and 23 for the case in which production was greater than demand.

[22] Notice that if a player had not changed his signaling level during the last two periods, or if his payoff had not changed, there is no gradient, and hill climbing cannot be applied. This is condition (5) in Table 3, and it occurred in 33% of the cases. The absolute frequencies for the cases (1) to (4) in Fig. 6a are 560, 2311, 2997, and 810. In Fig. 6b these frequencies are 1305, 1583, 2121, and 1710.

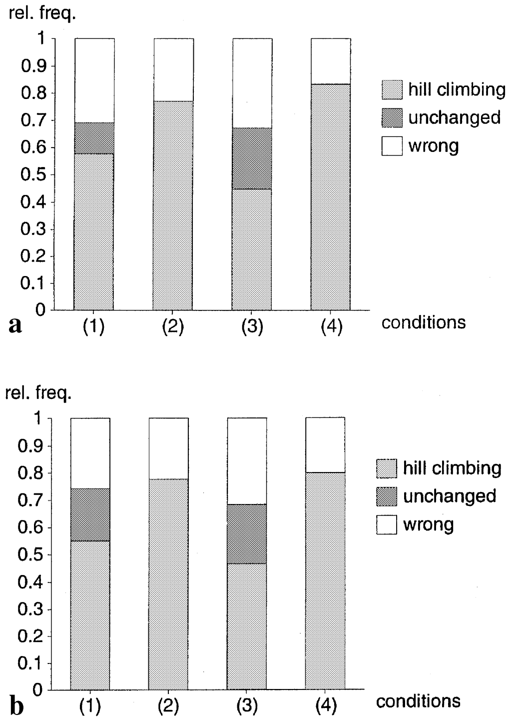



**Fig. 6.** **a** Myopic hill climbing, with conditions as explained in Table 3. **b** Lookahead hill climbing, with conditions as explained in Table 3

as well as his lookahead payoff (taking into account also the future profits related to his current sales level) had increased, or both had decreased. When we consider only the other 29% cases of opposite gradients, the cases in which myopic hill climbing and lookahead hill climbing predict a different change in signaling, we find that on average the players are inclined only slightly towards looking ahead; in 53.2% of those cases they follow the prediction of lookahead hill climbing, and in 48.8 the prediction of myopic hill climbing.

The numbers in parentheses on the horizontal axis denote the 'if …' conditions as given in Table 3. The light shaded bars give the frequencies when the hill climbing prediction was strictly correct. The dark shaded bars give the frequencies with which players choose signaling in period $t + 1$ equal to signaling in period t. Notice that for conditions (2) and (4), those cases are already included in the strictly correct predictions. For conditions (1) and (3), according to the hill climbing hypothesis, a player should reverse the direction of the change in his signaling level, whereas it would be strictly wrong to continue moving into the same direction that led to a decrease in payoffs. The inertia indicated in the figures by the dark shaded bars is not exactly predicted by the hill climbing hypothesis, but it is also not strictly wrong. Moreover, there might be good reasons for this inertia. First, players might keep their signaling level constant for a period, in order to adjust their production level according to the rules suggested by the learning direction theory. Second, given the noise caused by the other players, it may be wise not to put all the weight on the last period alone. This suggests that a further refinement of the modeling of the players' behavior could be obtained, by considering algorithms taking into account more periods, such as in reinforcement learning (see, e.g., Roth and Erev, 1995).

**Observation 4.** There is an asymmetry between the cases in which a player's payoff had increased and those in which it had decreased. When things are going well, a player will *not* easily *switch* into the wrong direction with his signaling. When, on the other hand, a player's payoff is decreasing, he is more likely to *continue* into the wrong direction with his signaling.

For convenience, we consider here only lookahead hill climbing. Compare in Figure 6a, b the relative frequencies of *wrong* predictions for cases (1) and (3) with cases (2) and (4). In cases (1) and (3), the player's payoff had gone down, and so continuing to change his signaling level in the same direction would be wrong (29% of the times this happened). In cases (2) and (4), the player's payoff had increased, and so going back to his previous signaling level and then moving into the opposite direction would be wrong (21% on average). We used a sign test to analyze whether individual players were more likely to go into a wrong direction in the cases (1) and (3) than in the cases (2) and (4). For 43 out of 66 subjects this was the case (significant at 1.0% level; 1-sided). We conjecture that the fact that unsuccessful courses of actions are more easily continued than are successful ones reversed, is a more general psychological feature.

We have seen that the 2-step model we proposed does not perfectly describe the behavior of the players. But at the same time, the attraction of the model is its simplicity. A question, then, is whether the time-pattern of the average behavior of the players in the experiments fits the pattern predicted by this simple model. In Fig. 7 we present the average behavior of 11 simulated sessions with 6 players, and the average signaling levels observed in the experiments. As we see, the average signaling level not only converges to the same level, but it also shows a similar initial dip.[23]

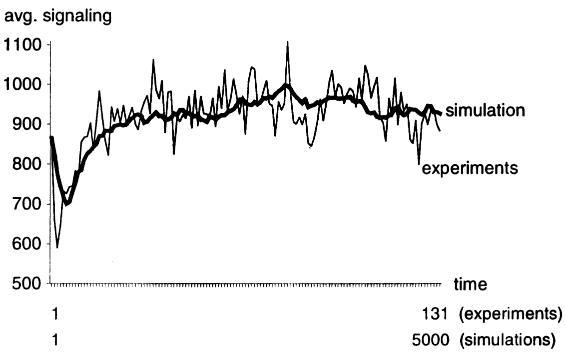



**Fig. 7.** Simulation 2-step model vs. experimental data

[23] It should be stressed that the two curves have a different time scale. Tinkering with the speed of adjustment of the players (e.g., adjusting the speed itself as well), would yield a better fit along this dimension, but that is not our objective. We use an identical and constant (but stochastic) low adjustment speed for all players.

### *c) Explaining the differences between the individual players*

Although the average behavior of the players appears to fit rather well to the symmetric game-theoretic equilibrium, and also to the convergence level and time-pattern of the 2-step model, in Section 4.a we observed that underneath these averages there were strong differences between the players.[24] In this section we will analyze and explain these strong differences.

If we have a look at Fig. 4, showing the distribution of signaling and production levels of the individual players, a first question is how these differences in actions correspond to differences in performance; and a second is how we arrive at this distribution. In other words, in what sense does the behavior of some players differ from that of other players?

**Observation 5.** There are considerable differences in performance among the players. We can distinguish three categories. Category I: the successful players, Category II: the 'nil players', and Category III: the unsuccessful players. The category II players choose relatively low signaling and production levels, and realize profits close to zero. As for the category I players, category III players try higher signaling (and production) levels than category II players, but they are less successful than category I players.

A method to measure the difference in performance among the players is the Gini coefficient (see, e.g., Case and Fair, 1996), which measures the skewness in the wealth distribution of a population, using the Lorenz curve. If the poorest x% of a population has x% of the total wealth of that population for each $0 \leq x \leq 100$, we have an equal distribution, characterized by a Gini coefficient equal to 0. If the richest person in the population has 100% of the total wealth, the Gini coefficient will be 1. The Gini coefficient for the 66 players is 0.41.[25] Given this unequal performance, what does the distribution look like, and what is its relation to the actions chosen? In Figure 8a we order all 66 players in terms of their cumulative profit per period, and in Figure 8b we present for these same players their average signaling.[26] Although these categories can be identified easily visually, they can be derived formally as follows: Having ordered all players on their average profits, calculate average signaling for each player, consider any two possible boundaries yielding three categories, and take those

[24] This is similar to the findings by Keser and Gardner (1998) who observe that aggregate behavior in a common pool experiment is well explained by the subgame perfect equilibrium, although only 5% of the subjects play in accordance to the theory. See also Budd et al. (1993) and Midgley et al. (1996).

[25] In order to allow for a comparison between the different sessions, we consider the same number of periods played for each session, i.e., 131. The wealth for a player is the cumulative profits realized plus the initial 2000 points he could loose before going bankrupt. Hence, bankrupt players have an accumulated wealth of zero. The Gini coefficients per session are available upon request.

[26] These individual averages are taken over the periods in which a player was active, i.e. until he went bankrupt or until the end of the session, whichever came first. Adding production levels would yield little extra information since average production and signaling are almost perfectly correlated.

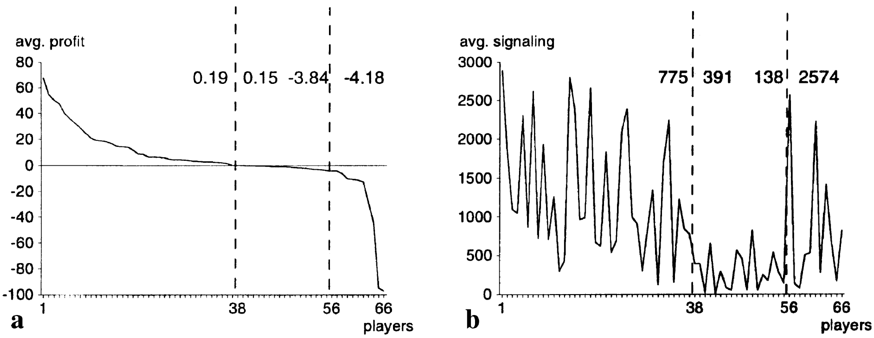



**Fig. 8.** **a** Average profit. **b** Average signaling

**Table 5.** Averages for the three categories

| Players | #Players | Signaling | Production | Profits |
|---|---|---|---|---|
| all | 66 | 951 | 160 | 4.0 |
| cat. I | 37 | 1301 | 194 | 16.6 |
| cat. II | 18 | 290 | 49 | −1.4 |
| cat. III | 11 | 857 | 225 | −29.5 |

boundaries for which the difference between the average signaling in the middle category and the other two categories combined is maximized.[27] We will use these three categories in our subsequent analysis, to see whether we can identify qualitative differences in the adaptive behavior between these three groups of players. The numbers in Fig. 8a, b give the values of profits and signaling respectively for the observations next to the boundaries.

Table 5 illustrates this categorization further by giving the average signaling, production, and profit levels per category as shown in Figure 8a, b. We use the Wilcoxon-Mann-Whitney test (Wilcoxon test from here on) to analyze whether the signaling levels of the individual players in the three categories are drawn from the same population. The alternative hypotheses are that the signaling level is stochastically higher for category I than for category II players (significant at 0.0% level), lower for category II than for category III players (significant at 2.6%), and different for category I and category III players (significant at 5.0%).

The question, then, is from where do these differences between the players' actions and outcomes arise?[28] We will offer three broad explanations. First, we will show how it is related to the dynamics of the oligopoly

[27] We imposed the additional restriction that there should be at least 3 players per category.

[28] The production and signaling technology are characterized by constant marginal costs. Hence, any firm size might seem efficient, and an unequal distribution of firm sizes would not be surprising. Notice, however, that the demand equation (1) implies that the marginal revenue of a signal sent is not constant, and depends upon the firm size.

game, and the players' perception thereof and success in dealing with it. Second, we will show how it is related to the players' initial choices, and the positive feedback inherent in the dynamics of the game. Third, we analyze the differences in the players' ambitions.

**Observation 6.** The observation (see Sect. 4b) that the players are less boundedly rational than learning direction theory assumes applies in particular to category I players.

Figure 9a,b summarize how far learning direction theory predicts correctly, distinguishing the cases of too high and too low production in the preceding period, and distinguishing the three categories of players. Comparing the frequencies of increasing production in those cases in which production was less than demand (Fig. 9a), players in category II increase their production less often than category I players (significant at 3.1% level with 1-sided Wilcoxon test). The difference between category II and category III players is not significant, and the fact that category III players increase their production less often than category I players is significant only at the 7.6% level. Looking instead at the frequencies of decreasing production in those cases in which production was higher than demand (Fig. 9b), players in category II decrease their production more often than category I players (significant at 1.0% level with 1-sided Wilcoxon test), and less often than category III players (significant at 1.1%). The difference between category I and category III players is not significant. Hence, it seems as though category I players understand best the desirability of overproduction, while category II players understand this least well, and as a result more easily become small players.

**Observation 7.** When hill climbing, category I players look ahead most often. Category II players do so least frequently.

Table 6 shows the frequencies with which the players go for immediate profits, and with which they look ahead in those cases in which the hill climbing hypothesis points to opposite directions. We observe that the differences in frequencies between the categories are not large. Category II players look ahead less frequently than category I players (significant at 0.9%; 1-sided Wilcoxon test), and also less frequently than category III

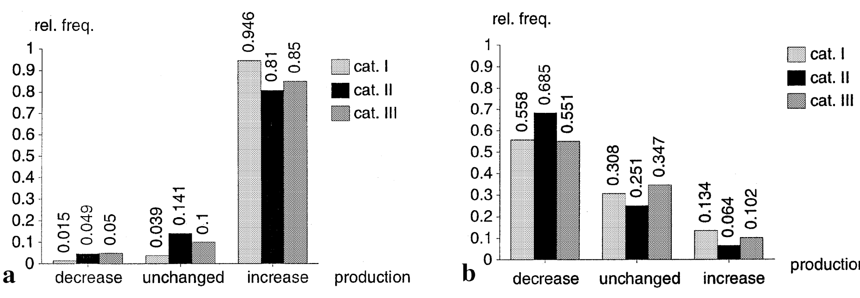



**Fig. 9.** **a** Direction learning after production < demand. **b** Direction learning after production > demand

**Table 6.** Frequencies myopic vs. lookahead hill climbing

| Players | Absolute frequencies | | Rel. frequencies (%) |
|---|---|---|---|
| | Myopic | Lookahead | Lookahead |
| all | 1094 | 1243 | 53.2 |
| cat. I | 595 | 747 | 55.7 |
| cat. II | 383 | 364 | 48.8 |
| cat. III | 117 | 132 | 53.0 |

players (significant at 9%). There is no significant difference between category I and category III players. Hence, category II players are the most myopic, not putting enough resources into building their market, and this partly explains why they are small players.

A second explanation for the differences between the players is related to their choices in the initial periods, and, related to the dynamics of the game, the way in which these initial choices have prolonged effects on the players' behavior.

**Observation 8.** Both production and signaling levels in the first period are concentrated on focal points. Further, the individual players' sales in later periods are positively correlated with their sales in the initial periods. The correlation coefficient between the 66 individual players' average sales levels in the periods 1–10 and the periods 81–130 (taking zero values for bankrupt players) is 0.55 (significant at 0.0% level; 1-sided t-test).

In the first period, the players have very little information to guide their decisions. Nevertheless, these choices are far from uniformly randomly distributed over the relevant choice domain. First, we look at production. The choice domain ranges from 0 to 4999, but the players was told that the demand faced by an individual firm would in general be below 1000. Only 6 players (9%) chose production levels greater than 1000. 61 out of 66 players (92%) chose a multiple of 50, and 53 (80%) picked production levels that are multiples of 100. The favorite multiple of 100 is 500, chosen by 13 players (20%), followed by 800 (7 players, or 11%), and 1000 (6 players, or 9%). Thus, as observed in many other experiments, the midpoint is a focal point (see, e.g., Ochs, 1994 on coordination games). Next, we look at signaling 61 players (92%) chose multiples of 50 or 100, and 55 (83%) chose multiples of 100, the most frequently chosen being again 500 (8 players, or 12%).

The correlations between the players' initial and later experiences are further illustrated by Table C.1 in Appendix C, where we give for each player his initial period actions and outcomes, and his averages over his whole playing history. The question one has to address is, once we observe such a correlation, where does it stem from? In Section 2 we identified various positive feedback mechanisms. Let us see how they can be related to these positive correlations between initial and later sales. First, we showed the temptation to maximize immediate profits by choosing signaling equal to zero, with production greater than zero. In that way, a firm's costs would

**Table 7.** Shrinking customer pool by not signaling, with production > 0

| Players | # Obs. | Shrinking | Rel. frequency |
|---|---|---|---|
| all | 10174 | 391 | 3.8 |
| cat. I | 5877 | 107 | 1.8 |
| cat. II | 3078 | 156 | 5.1 |
| cat. III | 1219 | 128 | 10.5 |

be greatly reduced because there are no signaling costs, with the patronizing customers showing up 'for free', but a consequence would be the shrinking of its pool of customers, with negative effects on later sales and profitability. How often did the players follow this strategy? And are there differences between the categories?

**Observation 9.** Shrinking the customer pool by not signaling is done regularly by players in all three categories. But there are differences between the categories. Category III players are much more inclined to eat drastically into their customer pool than are category II players, who are in turn much more inclined to do so than are category I players.

Table 7 illustrates this. Notice that category III players do this in more than 10% of their decision periods, that this is almost 6 times as often as category I players, and more than twice as often as category II players. We use the Wilcoxon test to analyze whether these levels of the individual players in the three categories are drawn from the same population. The alternative hypotheses are that shrinking occurs less often for category I than for category II players (significant at 0.9% level), less often for category II than for category III players (significant at 4.0%), and less often for category I than for category III players (significant at 0.0%). Recall that category III players signal on average much more than category II players, that is, they counter the shrinking of their customer pool by extra signaling in the periods following it. This aggressive '*on-off*' signaling behavior might be one of the explanations for their low profits.[29]

A second positive feedback effect presented in Section 2 was related to the fact that small firms would more easily get negative profits. Players on small islands get wet feet easily. Clearly, positive and negative profits are only relative. However, when profits are negative, a player has always the option to play (0, 0) for (signaling, production). Since that leads to a sales level of zero, and no patronizing customers, it implies a strong negative lock-in effect.

**Observation 10.** Excluding bankruptcy cases, *switching* to inactivity is predominantly done by players after observing a loss in the preceding period. There are differences between the categories. Category II players are more

[29] It is not that players deliberately making themselves bankrupt increase these frequencies for category III players. In fact, leaving the bankrupt players out would give an even higher average frequency for shrinking for category III players (11.0%).

skeptical about their opportunities than the other two categories. They switch most easily to inactivity. Once voluntarily inactive, the probability to stay inactive the next period is much higher than the probability of returning to business (84% against 16%).

Table 8 illustrates the voluntary switching to inactivity. Considering the individual players, only 1 player out of 66 switches to inactivity less often after a loss than otherwise. Using the Wilcoxon test to analyze whether the switching-to-inactivity frequencies of the individual players in the three categories are drawn from the same population, we find that category II players switch to inactivity more often than category I players (significant at 0.0% level), and category II players switch to inactivity also more often than category III players (significant at 2.4%), whereas there is no significant difference between category I and category III players. Recall that category III players realized negative profits much more frequently than category II players, so they try hard to improve upon their payoffs by acting rather than staying out.

Up to this point we have discovered two main explanations for the differences between the players. A first factor explaining these differences is their perception of the dynamics of the game, and this is extensively documented in the analysis above. A second factor is that the players' choices and outcomes in the initial periods turned out to be an important explanatory factor for success, or lack thereof, in later periods. The players' actions and outcomes during the initial periods might be just a matter of good or bad luck, but it might also be related to their pre-game experience in real life – what they have learned outside the laboratory – or it might be related to other psychological factors. There is a third factor that might explain some of the differences between the three categories of players. This being the ambitions of the players. To consider this, in a previous paper (Nagel and Vriend, 1998) we carried out an aspiration level analysis. The basic idea of such an analysis is that agents, due to their bounded rationality, are not able to optimize, and therefore will settle for satisficing behavior. Which outcomes are satisficing for a certain agent depends upon his aspiration level, where those levels are a moving target based, for example, on the agent's direct experience, or on the outcomes of other agents. This is a qualitative theory of adaptive behavior, presuming that when an agent's targets are met, he will be satisfied, and hence not change his behavior, whereas if his targets are not met, he will try to improve upon his situation by changing his

**Table 8.** Relative frequency switching to voluntary inactivity

| Players | # Observations | | Rel. frequencies (%) inactivity | |
|---|---|---|---|---|
| | Profit $< 0$ | Profit $\geq 0$ | After profit $< 0$ | After profit $\geq 0$ |
| all | 3434 | 6308 | 1.7 | 0.1 |
| cat. I | 1470 | 4334 | 0.5 | 0.0 |
| cat. II | 1349 | 1427 | 3.4 | 0.4 |
| cat. III | 615 | 547 | 1.1 | 0.4 |

actions.[30] The central question we studied there was whether there are differences between the three categories of players (the successful ones, the unsuccessful, and the 'nil' players). We found, among other things, that there are systematic differences between the players in the three categories as far as their reaction to satisfactory or unsatisfactory outcomes is concerned. In particular, category I players appear to be more ambitious than category II or III players, in the sense that they increase their production and signaling levels even when their aspiration level had been reached, whereas the latter two categories tend to keep production and signaling unchanged when their aspiration levels had been reached.

## 5 Conclusions

There are three main conclusions we can draw from this experimental study of adaptive behavior in an oligopolistic market game. The first is related to the average behavior of the players. Notwithstanding the minimal information the players were provided with, on average they learned to choose actions that were close to the symmetric stationary equilibrium for the complete information variant of the game. The second conclusion concerns the proposed a 2-step model, based on the game-theoretic analysis, in which the players use their signaling level as the basic strategic variable, whereas they adjust their production level towards the demand thus generated. It seems fair to conclude that learning direction theory, combined with the qualification concerning the *ex ante* optimality of overproduction, gives an accurate description of the players' behavior as far as their changes of production levels is concerned. The hill climbing hypothesis with respect to the players' signaling level was slightly less accurate, and made wrong predictions in about a quarter of the cases. In particular we detected an asymmetry in the players' behavior. When payoffs were increasing, players tend to continue their course of action. But when payoffs were decreasing and the players should have reversed the direction their signaling was moving into, they often continued walking downhill. We also showed that inertia in the players' behavior was important. This suggests that a further refinement of the modeling of the players' behavior could be obtained, by considering algorithms taking into account more periods than the most recent alone, e.g. reinforcement learning (see Roth and Erev, 1995).[31] Using the hill climbing hypothesis, we analyzed how far the players were inclined to go myopically for immediate profits: all players were only slightly more inclined to look ahead, and this was true above all for the successful players. A numerical exercise showed that the simple 2-step model seems to offer a reasonable explanation for the average market outcomes, both for the level and time-pattern of convergence.

[30] See Börgers and Sarin (1996) and Hart and Mas-Colell (1996) for two recent examples of aspiration level analyses.

[31] One of the first problems, then, is how to reduce the choice set of the players (see, e.g., Holland et al., 1986). Much more progress needs to be made here.

A third conclusion is that the players' behavior in this symmetric game is highly heterogeneous, much more so than expected. There are strong differences between the players, both with respect to their average actions and to their average payoffs. We showed that three categories of players could be distinguished: the successful ones, the 'nil players', and the unsuccessful players. The actions and outcomes in the initial period turned out to be important for the players' later performance. This could be due to good or to bad luck, but it also could be related to personality issues, both how daring they are in the first periods, and how they react to success or lack of it. We analyzed how this was related to some of the positive feedback mechanisms present in the market, and how the different categories of players dealt with these more or less successfully. In general, with help of the 2-step model, we showed that the players had different rates of success in adapting to their environment. An aspiration level analysis pointed to differences in the players' ambitions as an additional factor explaining their differences in performance.

## Appendix A. Instructions, and computer screen

Table A.1 contains the English version of the instructions given to the players.

**Table A.1.** Instructions to the players

Actors:
* Each of you will be a firm in a market economy.
* The consumers in this economy are simulated by a computer program.

Each day:
* In the morning, firms decide:
  - Identical firms decide upon a number of units of a perishable consumption good (each firm the same good).
  - The production of each unit costs 0.25 point.
  - The production decided upon at the beginning of the day is available for sale on that day.
  - Experience shows that, in general, the demand faced by an individual firm is below 1000.
  - The firms also decide upon a number of information signals to be sent into the population, communicating the fact that they are a firm offering the commodity for sale on that day. Imagine the sending of letters to addresses picked randomly from the telephone book.
  - Sending one information signal to an individual agent always costs 0.08 point.
  - The price of the commodity is 1 point. The price of the commodity is given, it does not change over time, it is equal for all firms and consumers, and known to all agents.
  - It is not possible to enter values greater than 4999 for the number of units to be produced and the number of information signals to be sent. This is due only to technological restrictions, and has no direct economic meaning.

**Table A1** (*Contd.*)

* During the day, consumers are '*shopping*':
  – When all firms have decided their actions, consumers will be '*shopping*'. Each day, each consumer wishes to buy exactly one unit of the commodity. Hence, consumers have to find a firm offering the commodity for sale, and such a firm should have at least one unit available at the moment they arrive.
  – We give you two considerations concerning the consumers' actions:
  *a* A consumer that has received an information signal from you knows that you are a firm offering the commodity for sale on that day.
  *b* Consumers who visited you, but arrived too late and found only empty shelves might find your service unreliable. On the other hand, a consumer who succeeded in buying one unit from you might remember the good service.
* At the end of the day, each consumer and each firm observes his own market outcomes:
  – Consumers turn home satisfied or not, i.e. with or without a unit of the commodity.
  – All unsold units of the commodity perish.
  – Each firm will know the demand that was directed to it during the day, how much it has actually sold (notice that it cannot sell more than it has produced at the beginning of the day), and its profits of that day.
  – It cannot be excluded that sometimes the market outcomes are such that a firm makes a loss. Each firm faces an upper limit of 2000 points for the total losses it may realize. A firm exceeding this limit will be declared bankrupt, implying that it will be forced to inactivity from then on.
  – A firm might have received some information signals sent to random addresses by other firms. These information signals will be listed (senders and numbers of signals), using fictitious names for the sending firms.

Time:
* There is no time limit for your daily decisions. From day 20 on, you will hear a warning sound when you are using more than one minute decision-time.
* The playing-time will be about 2½ hours.

Payment:
* Each player will be paid according to the total profits realized by its firm.
* Each player gets a '*show-up*' fee of DM 20.-.
* In addition, the payoff will be DM 10.- for each 1000 profit points realized.
* Note that losses realized will be subtracted from the DM 20.-.
* Bankrupt players have lost an amount of DM 20.-, and hence get nothing.

Anonymity:
* A player will never know the actions and outcomes of other players.

Keyboard:
* To confirm your choice: Enter [<]
* To delete: Backspace [<--]
* Please, before confirming your choices, always make sure that you did not make a typing-error.

Firm "X": RESULTS day 7

| ACTIONS | | OUTCOMES | | |
|---|---|---|---|---|
| production | signaling | demand | sales | profits |
| 123 | 450 | 114 | 114 | 47.25 |

The NEXT day is:

day 8

production =
signaling =

Firm "X", please enter your choices

price = 1.00; costs/unit produced = 0.25; costs/signal sent = 0.08

**Fig. A.1.** Computer screen firm 'X'

Figure A.1 shows the computer screen as viewed by a player acting as firm 'X' in a given period. At the beginning of day 1 the top part of the screen contained the following message: "Experience (from previous experiments) shows that, in general, the demand faced by an individual firm is below 1000". When a player had negative cumulative profits, he got a warning in the center of the screen saying: "WARNING! Your total losses are 192.25 (total losses greater than 2000.00 imply BANKRUPTCY!)".

## Appendix B. Game-theoretic analysis

*Proof of Proposition 1.* First, we study the finite horizon case, and then obtain the stationary signaling policy as a limit. The profit function and demand function are given by: $\Pi_i^t = p \cdot x_i^t - c \cdot z_i^t - k \cdot s_i^t$, where: $x_i^t = \min[z_i^t, q_i^t]$, and: $q_i^t = \text{round}(\text{trunc}[f \cdot x_i^{t-1} + \frac{s_i^t}{S^t} \cdot [1 - \exp(-\frac{S^t}{N})] \cdot [n - \Sigma_{j=1}^m \text{trunc}(f \cdot x_j^{t-1})]])$. Since the demand function is deterministic, $z_i^t = q_i^t = x_i^t$. Hence, the only control variable is signaling. Assuming the game is played for $T$ periods, the value $V$ of an action in any period $T-t'-1$ equals the sum of the immediate profits $\Pi$ in period $T-t'-1$ and the value $V$ in period $T-t'$: $V_i^{T-t'-1} = \Pi_i^{T-t'-1} + V_i^{T-t'}$, which has to be maximized. The first-order condition is: $\partial V_i^{T-t'-1}/\partial s_i^{T-t'-1} = \partial \Pi_i^{T-t'-1}/\partial s_i^{T-t'-1} + \partial V_i^{T-t'}/\partial s_i^{T-t'-1} = 0$. We consider these two terms on the right hand side separately.[32]

Determination of the first term: The immediate profit in a given period is: $\Pi_i^t = g \cdot q_i^t - k \cdot s_i^t \Rightarrow \partial \Pi_i^t/\partial s_i^t = g \cdot \partial q_i^t/\partial s_i^t - k$. Neglecting the term

[32] Cf. Fudenberg and Tirole (1991) on equilibria in dynamic games.

$[1-\exp(-\frac{S^t}{N})]$, and the roundings and truncations, demand is given by: $q_i^t = f \cdot q_i^{t-1} + s_i^t/S^t \cdot (n - f \cdot Q^{t-1}) \Rightarrow \partial q_i^t/\partial s_i^t = S_{-i}^t/(S^t)^2 \cdot (n - f \cdot Q^{t-1})$, where $S_{-i}^t$ is the aggregate signaling of the other players. Since all consumers visit a firm: $Q^{t-1} = n$. Hence, we get: $\partial \Pi_i^t/\partial s_i^t = g \cdot S_{-i}^t/(S^t)^2 \cdot n \cdot (1-f) - k$. We substitute $T - t' - 1$ for $t$.

Now we turn to the second term. We have to determine $V_i^{T-t'}$. We solve this first for the last period $T$, and then solve the game using backward induction. In the last period, period $T$, we have $\Pi_i^T = V_i^T$, and hence the first-order condition is: $\partial V_i^T/\partial s_i^T = \partial \Pi_i^T/\partial s_i^T = 0$. From above, we know $\partial \Pi_i^T/\partial s_i^T = g \cdot S_{-i}^T/(S^T)^2 \cdot (n - f \cdot Q^{T-1})$. Hence, we get: $\frac{k}{g} = S_{-i}^T/(S^T)^2 \cdot (n - f \cdot Q^{T-1})$. Since $S^T = m \cdot s_i^T$, and $Q^{T-1} = n$, we obtain: $s_i^T = g/k \cdot (m-1)/m^2 \cdot n \cdot (1-f)$, which is the optimal signaling level in the last period. Thus, the value $V_i^T$ in the last period is

$$\begin{aligned} V_i^T &= \Pi_i^T = g \cdot q_i^T - k \cdot s_i^T \\ &= g \cdot [f \cdot q_i^{T-1} + \frac{1}{m} \cdot (n - f \cdot Q^{T-1})] - k \cdot [\frac{g}{k} \cdot (\frac{m-1}{m^2}) \cdot n \cdot (1-f)] \Rightarrow \end{aligned}$$

$V_i^T = g \cdot [f \cdot q_i^{T-1} + \frac{n}{m^2} \cdot (1-f)]$. In other words, $V_i^T = g \cdot [A_0 \cdot q_i^{T-1} + B_0]$, or in general: $V_i^{T-t'} = g \cdot [A_{t'} \cdot q_i^{T-t'-1} + B_{t'}]$, where: $A_0 = f$, $B_0 = \frac{n}{m^2} \cdot (1-f)$ and: $A_{t'+1} = f + f \cdot A_{t'}$. Hence, $\partial V_i^{T-t'}/\partial s_i^{T-t'-1} = g \cdot A_{t'} \cdot \partial q_i^{T-t'-1}/\partial s_i^{T-t'-1} = g \cdot A_{t'} \cdot S_{-i}^T/(S^T)^2 \cdot n \cdot (1-f)$. Combining the two terms we get:

$$\begin{aligned} \frac{\partial V_i^{T-t'-1}}{\partial s_i^{T-t'-1}} &= g \cdot \frac{S_i^{T-t'-1}}{(S_i^{T-t'-1})^2} \cdot n \cdot (1-f) - k] \\ &\quad + [g \cdot A_{t'} \frac{S_{-i}^{T-t'-1}}{(S^{T-t'-1})^2} \cdot n \cdot (1-f)] = 0 \Rightarrow \\ g \cdot (1 + A_{t'}) \cdot \left(\frac{m-1}{m}\right) &\cdot \frac{1}{S^{T-t'-1}} \cdot (1-f) \cdot n = k \Rightarrow \end{aligned}$$

$s_i^{T-t'} = g/k \cdot (1 + A_{t'+1}) \cdot (m-1)/m^2 \cdot (1-f) \cdot n$. Now consider the difference equation $A_{t'+1} = f + f \cdot A_{t'}$, which can be solved as: $A_{t'+1} = [f - f/(1-f)] \cdot f^{t'+1} + f/(1-f)$, with $\lim_{t' \to \infty} A_{t'+1} = f/(1-f)$. Hence, for large enough $t'$ the optimal action in a given period $T - t'$ in the steady state is: $s_i^{T-t'} = g/k \cdot [1 + f/(1-f)] \cdot [(m-1)/m^2] \cdot (1-f) \cdot n = g/k \cdot [(m-1)/m^2] \cdot n$. QED.[33]

[33] See also Stokey and Lucas (1989).

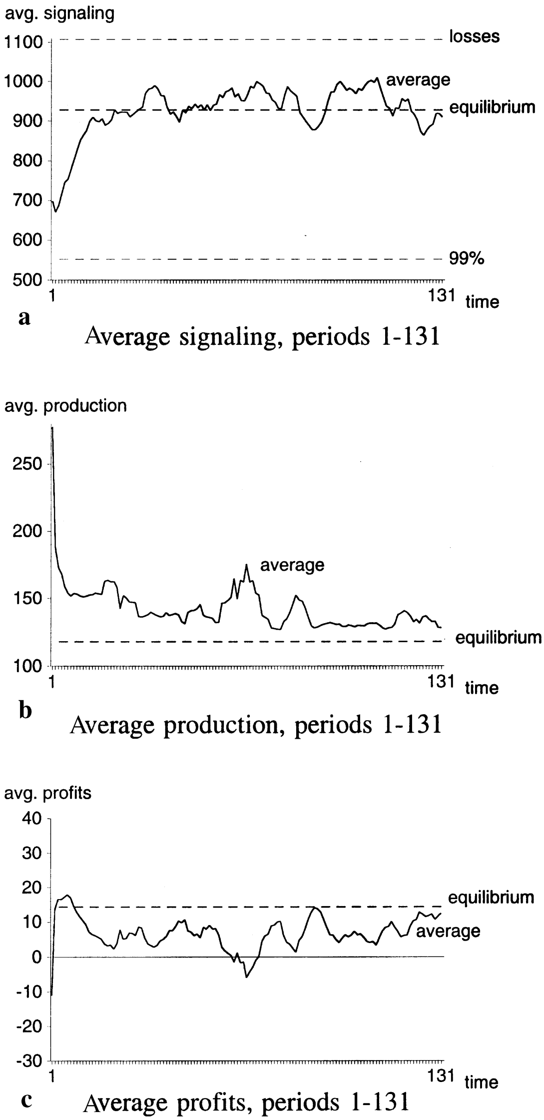



**Fig. C.1.** **a** Average signaling, periods 1–131. **b** Average production, periods 1–131. **c** Average profits, periods 1–131

## Appendix C. Some additional data

Figure C.1.a, c present the time series for signaling, production, and profits for periods 1 to 131 averaged over the 66 players (with all variables at zero for bankrupt players). In all these graphs, we took a five period moving average for presentational reasons, and we added the equilibrium levels as a first benchmark. In the graph for signaling (C.1.a) we added two other benchmarks. The first one is called 'losses', and corresponds to the line drawn in Figure 4, as explained in Section 4. It is the signaling level beyond which positive profits are impossible, given the equilibrium production

**Table C.1.** Summary data individual players

| Session | Player | Period 1 | | | Avg. all periods | | | Profits | Number periods | Also session | Player |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Prod. | Sign. | Sales | Prod. | Sign. | Sales | | | | |
| 1 | 1 | 0 | 100 | 0 | 154 | 868 | 144 | 35.9 | 151 | 21 | 2 |
| 1 | 2 | 400 | 200 | 91 | 215 | 1096 | 192 | 50.4 | 151 | 21 | 1 |
| 1 | 3 | 500 | 500 | 228 | 197 | 1045 | 180 | 47.4 | 151 | 21 | 4 |
| 1 | 4 | 250 | 400 | 183 | 131 | 393 | 64 | 0 | 151 | 21 | 5 |
| 1 | 5 | 100 | 100 | 46 | 129 | 715 | 118 | 28.9 | 151 | – | – |
| 1 | 6 | 20 | 5 | 2 | 467 | 161 | 35 | −94.9 | 22 | – | – |
| 2 | 1 | 1199 | 2199 | 522 | 362 | 2309 | 315 | 39.8 | 151 | – | – |
| 2 | 2 | 300 | 100 | 24 | 38 | 78 | 10 | −6.1 | 151 | – | – |
| 2 | 3 | 200 | 0 | 0 | 207 | 1348 | 163 | 3.4 | 151 | – | – |
| 2 | 4 | 600 | 400 | 95 | 83 | 540 | 70 | 6.6 | 151 | – | – |
| 2 | 5 | 600 | 100 | 24 | 163 | 986 | 134 | 14 | 151 | 21 | 3 |
| 2 | 6 | 350 | 150 | 36 | 12 | 52 | 6 | −1 | 151 | 21 | 6 |
| 3 | 1 | 1000 | 1500 | 113 | 138 | 827 | 99 | −1.8 | 151 | – | – |
| 3 | 2 | 600 | 800 | 91 | 124 | 846 | 100 | 1.5 | 151 | 23 | 3 |
| 3 | 3 | 800 | 4800 | 216 | 9 | 54 | 5 | −2 | 151 | – | – |
| 3 | 4 | 1200 | 2000 | 129 | 335 | 2574 | 286 | −4.2 | 151 | – | – |
| 3 | 5 | 1000 | 800 | 91 | 36 | 170 | 20 | −2.7 | 151 | 23 | 2 |
| 3 | 6 | 200 | 200 | 72 | 213 | 1699 | 192 | 2.6 | 151 | – | – |
| 4 | 1 | 500 | 1500 | 153 | 429 | 2893 | 406 | 67.8 | 151 | 23 | 1 |
| 4 | 2 | 1000 | 1000 | 124 | 439 | 820 | 78 | −97 | 28 | – | – |
| 4 | 3 | 500 | 500 | 95 | 26 | 147 | 20 | 2.2 | 151 | – | – |
| 4 | 4 | 2000 | 2000 | 182 | 344 | 710 | 99 | −44.2 | 68 | – | – |
| 4 | 5 | 900 | 400 | 89 | 50 | 298 | 40 | 4 | 151 | – | – |
| 4 | 6 | 100 | 50 | 69 | 202 | 1268 | 171 | 19.1 | 151 | – | – |
| 5 | 1 | 500 | 2000 | 285 | 356 | 2805 | 330 | 16.8 | 151 | – | – |
| 5 | 2 | 700 | 600 | 85 | 29 | 138 | 14 | −3.8 | 151 | – | – |
| 5 | 3 | 500 | 600 | 85 | 86 | 505 | 52 | −10.3 | 151 | – | – |
| 5 | 4 | 500 | 700 | 100 | 44 | 280 | 30 | −3.3 | 151 | 24 | 6 |
| 5 | 5 | 500 | 500 | 71 | 229 | 1841 | 211 | 6.6 | 151 | – | – |
| 5 | 6 | 400 | 600 | 85 | 97 | 527 | 56 | −10.8 | 151 | 24 | 4 |
| 6 | 1 | 1200 | 800 | 185 | 107 | 618 | 85 | 8.8 | 151 | 24 | 3 |
| 6 | 2 | 800 | 400 | 92 | 112 | 691 | 90 | 6.4 | 151 | – | – |
| 6 | 3 | 550 | 500 | 116 | 391 | 2627 | 340 | 32.3 | 151 | 24 | 2 |
| 6 | 4 | 400 | 300 | 69 | 6 | 20 | 3 | −0.1 | 151 | – | – |
| 6 | 5 | 500 | 800 | 185 | 168 | 1228 | 142 | 2.2 | 151 | 24 | 5 |
| 6 | 6 | 234 | 234 | 54 | 113 | 272 | 37 | −12.6 | 151 | 24 | 1 |
| 7 | 1 | 250 | 150 | 35 | 58 | 457 | 50 | −1.1 | 131 | – | – |
| 7 | 2 | 600 | 900 | 207 | 138 | 991 | 118 | 4.4 | 131 | 25 | 3 |
| 7 | 3 | 400 | 400 | 92 | 274 | 2109 | 243 | 6.2 | 131 | 25 | 1 |
| 7 | 4 | 50 | 100 | 23 | 38 | 289 | 32 | −0.5 | 131 | – | – |
| 7 | 5 | 2000 | 500 | 115 | 43 | 134 | 17 | −4.3 | 131 | – | – |
| 7 | 6 | 800 | 1000 | 230 | 259 | 2230 | 232 | −11.2 | 131 | 25 | 6 |
| 8 | 1 | 875 | 900 | 117 | 94 | 430 | 76 | 18.1 | 131 | – | – |
| 8 | 2 | 1000 | 100 | 13 | 173 | 962 | 135 | 14.5 | 131 | – | – |
| 8 | 3 | 500 | 500 | 65 | 77 | 299 | 62 | 18.8 | 131 | – | – |
| 8 | 4 | 400 | 500 | 65 | 350 | 1836 | 289 | 55 | 131 | – | – |
| 8 | 5 | 900 | 990 | 128 | 28 | 121 | 20 | 3.1 | 131 | – | – |
| 8 | 6 | 2255 | 2500 | 324 | 154 | 705 | 116 | 21 | 131 | – | – |
| 9 | 1 | 800 | 2000 | 334 | 95 | 540 | 64 | −2.9 | 151 | – | – |

**Table C.1** (*Contd.*)

| Session | Player | Period 1 | | | Avg. all periods | | | Profits | Number periods | Also session | Player |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Prod. | Sign. | Sales | Prod. | Sign. | Sales | | | | |
| 9 | 2 | 300 | 100 | 17 | 39 | 246 | 27 | −2.3 | 151 | – | – |
| 9 | 3 | 200 | 600 | 100 | 269 | 2246 | 250 | 2.6 | 151 | 23 | 5 |
| 9 | 4 | 500 | 1000 | 167 | 295 | 2394 | 270 | 4.6 | 151 | 23 | 6 |
| 9 | 5 | 50 | 50 | 8 | 9 | 54 | 5 | −1.2 | 151 | – | – |
| 9 | 6 | 750 | 500 | 84 | 103 | 775 | 88 | 0.2 | 151 | 23 | 4 |
| 10 | 1 | 800 | 400 | 55 | 16 | 84 | 10 | −0.9 | 251 | – | – |
| 10 | 2 | 100 | 300 | 41 | 99 | 668 | 87 | 9.1 | 251 | 22 | 6 |
| 10 | 3 | 800 | 3000 | 414 | 265 | 1934 | 246 | 24.6 | 251 | – | – |
| 10 | 4 | 500 | 750 | 104 | 79 | 565 | 64 | −1 | 251 | – | – |
| 10 | 5 | 350 | 400 | 55 | 354 | 2377 | 293 | 14.7 | 251 | 22 | 5 |
| 10 | 6 | 300 | 300 | 41 | 3 | 11 | 1 | −0.4 | 251 | – | – |
| 11 | 1 | 800 | 1000 | 63 | 57 | 391 | 46 | 0.1 | 201 | 25 | 4 |
| 11 | 2 | 1000 | 3500 | 222 | 363 | 2667 | 316 | 11.9 | 201 | 25 | 5 |
| 11 | 3 | 500 | 450 | 28 | 255 | 1419 | 149 | −28.4 | 84 | – | – |
| 11 | 4 | 500 | 4999 | 316 | 86 | 656 | 74 | −0.2 | 201 | – | – |
| 11 | 5 | 1000 | 1000 | 63 | 117 | 912 | 107 | 4.3 | 201 | 25 | 2 |
| 11 | 6 | 300 | 300 | 19 | 101 | 842 | 96 | 3.6 | 201 | – | – |

**Table C.2.** Session averages

| Session | Period 1 | | | Avg. all periods | | | Profits | Number periods |
|---|---|---|---|---|---|---|---|---|
| | Prod. | Sign. | Sales | Prod. | Sign. | Sales | | |
| 1 | 212 | 218 | 92 | 149 | 690 | 117 | 25 | 151 |
| 2 | 542 | 492 | 117 | 144 | 886 | 116 | 9 | 151 |
| 3 | 800 | 1683 | 119 | 143 | 1028 | 117 | −1 | 151 |
| 4 | 833 | 908 | 119 | 157 | 846 | 116 | 9 | 151 |
| 5 | 517 | 833 | 119 | 140 | 1016 | 116 | −1 | 151 |
| 6 | 614 | 506 | 117 | 150 | 909 | 116 | 6 | 151 |
| 7 | 683 | 508 | 117 | 135 | 1035 | 115 | −1 | 131 |
| 8 | 988 | 915 | 119 | 146 | 726 | 116 | 22 | 131 |
| 9 | 433 | 708 | 118 | 135 | 1043 | 117 | 0 | 151 |
| 10 | 475 | 858 | 118 | 136 | 940 | 117 | 8 | 251 |
| 11 | 683 | 1875 | 119 | 139 | 1010 | 117 | 1 | 201 |

level. The second additional benchmark is called '99%', and corresponds to the average signaling level needed to make sure that 99% of the consumer population receives at least one signal.

Table C.1 presents the individual averages and first period actions for the players. For each individual player, the averages are taken over the periods in which the player actually played. Table C.2 gives the session averages, where the averages are taken over the periods in which the session lasted (with production and signaling levels at zero for bankrupt players).

## References

Arthur WB (1991) Learning and adaptive economic behavior. Designing economic agents that act like human agents: a behavioral approach to bounded rationality. American Economic Review 81: 353–359

Arthur WB (1992) On learning and adaptation in the economy. Working Paper 92-07-038, Santa Fe Institute

Atkinson RC, Suppes P (1958) An analysis of two-person game situations in terms of statistical learning theory. Journal of Experimental Psychology 55: 369–378

Binmore KG (1991) DeBayesing game theory. Mimeo, Lecture for the International Conference on Game Theory, Florence

Bloomfield R (1994) Learning a mixed strategy equilibrium in the laboratory. Journal of Economic Behavior and Organization 25: 411–436

Börgers T, Sarin R (1996) Naive reinforcement learning with endogenous aspirations. ELSE Working Paper, University College London

Budd C, Harris C, Vickers J (1993) A model of the evolution of duopoly: does the asymmetry between firms tend to increase or decrease? Review of Economic Studies 60: 543–573

Camerer C, Ho TH (1996) Experience-weighted attraction learning in games: A unifying approach. Mimeo

Case KE, Fair RC (1996) Principles of microeconomics, 4th edn. Prentice Hall, Upper Saddle River, NJ

Davis DD, Holt CA (1992) Experimental economics. Princeton University Press, Princeton, NJ

Easley D, Ledyard JO (1993) Theories of price formation and exchange in double oral auctions. In: Friedman D, Rust J (eds) The double auction market. Institutions, theories, and evidence. Addison-Wesley, Reading MA, pp 63–97

Ellison G (1993) Learning, local interaction, and coordination. Econometrica 61: 1047–1071

Fudenberg D, Tirole J (1991) Game theory. MIT Press, Cambridge, MA

Green JR (1983) Comment on "A. Kirman, On mistaken beliefs and resultant equilibria". In: Frydman R, Phelps ES (eds) Individual forecasting and aggregate outcomes: Rational expectations examined. Cambridge University Press, Cambridge

Hart S, Mas-Colell A (1996) A simple adaptive procedure leading to correlated equilibrium. Economics Working Paper 200, Universitat Pompeu Fabra

Holland JH (1992) Adaptation in natural and artificial systems. An introductory analysis with applications to biology, control, and artificial intelligence, 2nd edn. MIT Press, Cambridge, MA

Holland JH, Holyoak KJ, Nisbett RE, Thagard PR (1986) Induction: Processes of inference, learning, and discovery. MIT Press, Cambridge, MA

Kampmann C, Sterman JD (1995) Feedback complexity, bounded rationality, and market dynamics. Mimeo

Kandori M, Mailath GJ, Rob R (1993) Learning, mutation, and long run equilibria in games. Econometrica 61: 29–56

Keser C (1992) Experimental duopoly markets with demand inertia: Game-playing Experiments and the strategy method. Lecture Notes in Economics and Mathematical Systems 391. Springer, Berlin Heidelberg New York

Keser C, Gardner R (1998) Strategic behavior of experienced subjects in a common pool resource game. International Journal of Game Theory (forthcoming)

Kiefer NM, Nyarko Y (1989) Optimal control of an unknown linear process with learning. International Economic Review 30: 571–586

Kirman AP (1983) On mistaken beliefs and resultant equilibria. In: Frydman R, Phelps ES (eds) Individual forecasting and aggregate outcomes: Rational expectations examined. Cambridge University Press, Cambridge

Kirman AP (1993) Learning in oligopoly: Theory, simulation, and experimental evidence. In: Kirman AP, Salmon M (eds) Learning and rationality in economics. Blackwell, Oxford

Macready WG, Wolpert DH (1995) No free-lunch theorems for search. Working Paper 95-02-010, Santa Fe Institute

McKelvey R, Palfrey T (1992) An experimental study of the centipede game. Econometrica 60: 803–836

Malawski M (1990) Some learning processes in population games. ICS PAS Reports 678, Institute of Computer Science Polish Academy of Sciences, Warsaw

Marimon R (1993) Adaptive learning, evolutionary dynamics and equilibrium selection in games. European Economic Review 37: 603–611

Merlo A, Schotter A (1994) An experimental study of learning in one and two-person games. Economic Research Reports 94-17, CV Starr Center for Applied Economics, New York University

Midgley DF, Marks RE, Cooper LG (1996) Breeding competitive strategies. Management Science

Nagel R (1995) Unraveling in guessing games. An experimental study. American Economic Review 85: 1313–1326

Nagel R, Tang FF (1998) An experimental study on the centipede game in normal form–An investigation on learning. Journal of Mathematical Psychology (forthcoming)

Nagel R, Vriend NJ (1998) An experimental study of adaptive behavior in an oligopolistic market game. Working Paper No. 388, Queen Mary and Westfield College, University of London

Nagel R, Vriend NJ (1999) Do players really learn in an oligopolistic market game with minimal information? Industrial and Corporate Change (forthcoming)

Ochs J (1995) Coordination problem. In: Kagel J, Roth AE (eds) The handbook of experimental economics. Princeton University Press, Princeton, NJ, pp 195–252

Page SE (1994) Two measures of difficulty. Working Paper 94-12-063, Santa Fe Institute

Petr M (1997) A dynamic model of advertising competition: an empirical analysis of feedback strategies. Mimeo

Roberts M (1995) Active learning: Some experimental results. Mimeo

Roth AE, Erev I (1995) Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. Games and Economic Behavior 8: 164–212

Sauermann H, Selten R (1959) Ein Oligopolexperiment. Zeitschrift für die gesamte Staatswissenschaft 115: 427–471

Savage LJ (1954) The foundations of statistics. Wiley, New York

Selten R (1997) Features of experimentally observed bounded rationality. Mimeo, Presidential Address of the 1997 Meetings of the European Economic Association, Toulouse

Selten R, Buchta J (1994) Experimental sealed bid first price auctions with directly observed bid functions. Discussion Paper No. B-270, University of Bonn

Selten R, Mitzkewitz M, Uhlich G (1997) Duopoly strategies programmed by experienced players. Econometrica 65: 517–555

Selten R, Stoecker R (1986) End behavior in sequences of finite prisoner's dilemma supergames. A learning theory approach. Journal of Economic Behavior and Organization 7: 47–70

Simon HA (1959) Theories of decision-making in economics and behavioral science. American Economic Review 49: 253–283

Stahl DO (1996) Boundedly rational rule learning in a guessing game. Games and Economic Behavior 16: 303–330

Stewing R (1990) Entwicklung, Programmierung und Durchführung eines Oligopolexperiments mit minimaler Information. Masters thesis, University of Bonn

Stiglitz JE (1993) Principles of microeconomics. Norton, New York

Stokey NL, Lucas RE, jr (1989) Recursive methods in economic dynamics. Harvard University Press, Cambridge, MA

Tang FF (1996) Anticipatory learning in two-person games: An experimental study. Part II. Learning. Discussion Paper No. B-363. University of Bonn
Vriend NJ (1994) Artificial intelligence and economic theory. In: Hillebrand E, Stender J (eds) Many-agent simulation and artificial life. IOS, Amsterdam, pp 31–47
Vriend NJ (1996a) Rational behavior and economic theory. Journal of Economic Behavior and Organization 29 (2): 263–285
Vriend NJ (1996b) A model of market-making (Economics Working Paper 184) Universitat Pompeu Fabra, Barcelona
Witt U (1986) How can complex economic behavior be investigated? The example of the ignorant monopolist revisited. Behavioral Science 31: 173–188
Young HP (1993) The evolution of conventions. Econometrica 61: 57–84