An Experimental Study of Adaptive Behavior
in an Oligopolistic Market Game [*]

Rosemarie Nagel & Nicolaas J. Vriend, *Universitat Pompeu Fabra*, Barcelona, Spain

July 1997

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.

J.E.L. classification codes: C7, C9, D8

Keywords: Market Game, Oligopoly, Adaptive Behavior, Learning

1. Introduction

We consider an oligopolistic market game. The players are competitors in the same market. At the beginning of each period, firms decide how much to produce of a perishable consumption good, and they decide upon a number of information or advertising signals to be sent into the population in order to attract customers. The firms know their production and signaling technology,[1] and the price of the commodity, which is fixed. The consumers in this economy are simulated by a computer program. At the end of the day, each firm observes *only* its own market outcomes. Firms do not have any further knowledge about their environment.

Given this minimal information, the players are not in a position to maximize their profits using standard optimization techniques. Following Savage's [1954] terminology, they are in a *'large world'*.[2] 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 inductive reasoning; proceeding from the actual situation he faces (see, e.g., Arthur [1992]). In Savage's terminology, this is the *'cross that bridge when you meet it'* principle, which is also known as *'adaptive behavior'*. We would conjecture that most relevant economic problems, when considered at a moderately realistic level, are large world problems. The point is not so much that the agents' rationality is bounded, but more that their perception is limited (see also Vriend [1996a]). Our main objective is to characterize the adaptive behavior of the players. In particular, we will analyze the adaptive behavior according to two simple processes; learning direction theory, which has been successfully applied in various experiments (see e.g., Selten & Stoecker [1986], or Nagel [1995]), and the well-known method of hill climbing, also known as gradient method. Some other 'large world' experiments can be found in Atkinson & Suppes [1958], Witt [1986], Malawski [1990], and Kampmann & Sterman [1995].

In section 2, we explain how a large world looks like in a small experimental laboratory. In section 3, we present the game-theoretic analysis, and make a first global comparison with the experimental data. In section 4, we take a closer look at the game as such. Section 5 contains an analysis of the data for unexperienced players. In section 6, we analyze the differences between experienced and unexperienced players, and section 7 concludes.

2. The Experiment: Model and Design

We conducted two series of experiments in the computerized experimental laboratory at the University of Bonn. Before presenting the experimental design, we first explain and discuss the oligopoly model used.

The Oligopoly Model

---

[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] As opposed to the *'small world'* to which Subjective Expected Utility theory applies.

   A fixed number of firms repeatedly interacts in an oligopolistic market.[3] All firms are identical in the sense that they produce the same homogeneous 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 following fixed, deterministic demand function.   [1]

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

$$\text{I} \qquad + \text{IIa} \cdot \qquad \text{IIb} \qquad \cdot \qquad \text{IIc}$$

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 that were satisfied by a given firm during the last period will *patronize* that firm (part I of equation [1]). 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. 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

---

[3] See appendix B for notation and all parameter values used.

cannot sell more than it has produced at the beginning of the period. Hence, a firm's profit function is given by:

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

The following characteristic of the specific oligopoly model employed merits some discussion.[4] We have a fixed price, and the consumer side is simulated with a fixed deterministic demand function. That is, we abstract from the process by which the price was determined, and from the determination of the market demand at that price level.[5] These two abstractions are made in order to focus on the behavior of the firms in a relatively stable environment. As we will see in section 4, since the firms influence each other's environment, the task of the firm is already rather complicated.[6] Once we have understood and structured the behavior in such an environment, we can further relax these assumptions. Note, however, 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). Clearly, a *complete* economic analysis would explain such legislation, restrictive practices, or strategies, by which the prices are fixed, as well. But, as said above, that is not the aim of this paper. Instead of explaining the price, our analysis focuses 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 take place along many dimensions. Competing for customers through advertising seems a fundamental one. The essential characteristic of most forms of advertising in a market economy seems that the buyers' attention is drawn to the fact that someone is selling something somewhere sometime.

Information for the Individual Players

We now sketch the information that was available to the individual players, distinguishing technology, market, and experience factors. Appendix A presents the instructions given to the players, and the table in appendix B summarizes which parameter values were known and which not.

1) The Technology. The players knew that they were identical firms, producing the same homogeneous consumption good, using the same technology. Both the production and signaling technology were common knowledge, and the same applies to the price of the commodity. About the fact that the choice interval for production and signaling was limited, the players was told that *"this is due only to technological restrictions, and has no direct economic meaning"*.

---

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

[5] Notice that this is perfectly compatible with a standard downward-sloping market demand curve.

[6] Here is an illustration. During conversations after the sessions, many players were convinced that aggregate demand was fluctuating continuously, with only one player saying he was convinced there was a constant demand function. The experienced players claimed this point slightly less forcefully.

2) The Market. The players was told consumers in this economy are simulated by a computer program. They did *not* receive the specification of the demand function (equation [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. Hence, in each period, 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. 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 was also told *"experience shows that, in general, the demand faced by an individual firm is below 1000".*

3) 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 of 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. A firm exceeding this limit is declared bankrupt, with the participant removed from the session. This was told 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.

How reasonable is it to assume that firms do not observe the signaling decisions of the other firms? Notice that the only variable to compete directly with the other firms in this model is the signaling activity. In reality, competition takes place along many dimensions. 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. Even advertising as such comes in many forms, and might be, e.g., informative or persuasive. The bottom line of this partial list is, that 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. Assuming that firms do not observe the level of their competitors' signaling activity in our simple model is a first approximation of this fact.

---

[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.

Given this minimal information provided to a player, he 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 most questions attempted to get more knowledge about the environment. The usual answer to those questions was 'you just don't know'.

The Experiments

In the first series, we organized 13 sessions (throughout this paper numbered 1 to 11, plus 91 and 92). In each session, 6 firms were competing in one market, for a total of 78 players. In the second series, we organized 5 sessions (numbered 21 to 25), with again 6 firms per session, for a total of 30 players. Most players came from various departments of the University of Bonn. The players in the second series were a sample of experienced players from the first series. In the second series, all sessions were independent from each other with respect to the experience of the players in the first series.[8] In sessions with experienced players, there are two options for choosing the parameter setup; either using the same setup as in the preceding session, or a different parameter setup. The problem with using the same parameter setup for the experienced players is that a player might have found a good strategy by chance during the first experiment, without having learned anything general as to how to play.[9] As we are interested not in specific actions for the game considered, but in adaptive behavior in this type of environment, we chose the second option.[10] Thus, players in the second series were experienced in the sense that they might have learned something about the general *structure* of the game.

Players sat in front of personal computers, and could not observe the screens of other players. Appendix A presents an example of a computer screen viewed by a given player. We played about 150 periods with unexperienced players, and 300 with experienced players.[11] There was no time limit for the participants' decisions. Each player got a fixed *'show-up'* fee. In addition, each player 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).[12] 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 108 players was DM 36.42 ($\approx$ \$ 24.00).

---

[8] All experienced players in session 21 came from the sessions 1 or 2, in session 22 from sessions 91, 92 or 10, in 23 from 3, 4 or 9, in 24 from 5 or 6, and in 25 from 7 or 11.

[9] This problem may be mitigated by giving each experienced player all the statistics of the first series before the second series starts.

[10] See the table in appendix B for the two parameter setups used.

[11] All sessions with unexperienced players lasted 151 periods, except for the sessions 7 an 8 (131 periods), 10 (251 periods), and 11 (201 periods). All sessions with experienced players lasted 301 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.

[12] The value for α was DM 10 in sessions 1 and 2, 15 in sessions 3 to 6, 20 in sessions 7 to 11, 15 in sessions 91 and 92, 25 in session 21, 7 in session 22, and 13 in sessions 23 to 25, giving an average α of 15.6. The values for α were varied in an effort to keep average payoffs at a level of about DM 15/hour.

3. Game-theoretic Analysis and Actual Average Actions and Outcomes

In order to obtain a theoretical benchmark, in appendix C we derive the symmetric stationary optimal policy for a given player for any given period, assuming complete information about the demand function and an infinite number of periods, and considering only symmetric optima.[13] The optimal signaling level for an individual firm is given in equation [3].

$$s^{optimal} \;=\; \frac{g}{k} \;\cdot\; (\frac{m-1}{m^2}) \;\cdot\; n \qquad\qquad [3]$$

Given the signaling level, the demand for an individual firm is given by equation [1]. The optimal production level, then, is simply equal to that demand. Before we take a closer look at the game as such in section 4, we make a first global comparison between this theoretical benchmark and the actual experimental data.

Observation 1: The average actions actually chosen by both the unexperienced and the experienced 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.

Figures D.1.a to D.1.c in appendix D show the time series for the average signaling, production and profits of the 66 unexperienced players for the periods 1 to 131, and compare this with the symmetric stationary equilibrium. Figures D.2.a to D.2.c do the same for the 30 experienced players for the periods 1 to 301. With the unexperienced players, we observe a steep learning curve in the beginning. This leads to profits close to the equilibrium level early on. Experienced players happen to start with production levels closer to the equilibrium level. In both cases we see a lot of fluctuations during most of the history. And at the end we observe a movement towards the equilibrium levels. Table 1 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

---

[13] Notice that the firms are identical, and that the players did not know how many periods the experiment would last. We do not consider the optimal strategy for the incomplete information game. 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 & 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 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 for the moment unmanageable (see also, e.g., Green [1983], or Kirman [1983]). Also 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 & Holt [1992]). Clearly, this is not a normative benchmark, but merely a yardstick.

players, and the second based on the averages per session. We will come back to the differences between the individual players in section 5. Notice that the variance across sessions is rather small, and much smaller than across subjects, especially in the last 50 periods.

|	sign.	*(s.d.)*	prod.	*(s.d.)*	profit	*(s.d.)*
unexperienced:						
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)*
experienced:						
equilibrium	687		232		25.1	
period 1	959	*(1234/319)*	306	*(344/61)*	-109.0	*(176.4/84.1)*
period 1-50	727	*(713/110)*	254	*(251/13)*	5.8	*(38.7/23.0)*
period 81-130	715	*(790/78)*	247	*(279/8)*	11.5	*(32.7/15.9)*
period 251-300	698	*(925/78)*	241	*(311/6)*	17.6	*(32.9/14.1)*

Table 1 comparison equilibria, averages and standard deviations

4. A Closer Look at the Game

Besides the minimal information, there are two additional aspects of the structure of this game that are worth noticing. First, there is a *positive feedback* mechanism. A fixed proportion f of a firm's satisfied customers will patronize the next period. Hence, firms having sold more in period t, will get more customers in period t+1, and the other way round. 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. If we take the symmetric equilibrium strategy given in section 3 for the unexperienced players, each firm sends 927 signals, and receives 118 customers in a given period t. Of those 118 customers, if satisfied, $0.56 \cdot 118$ 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. 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, 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 is the case in 84% of the times in the first series of experiments, and 91% of the times with experienced players. 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 figure 1, where the *dot* indicates the action chosen by the player considered, and where the other players each send 750 signals.
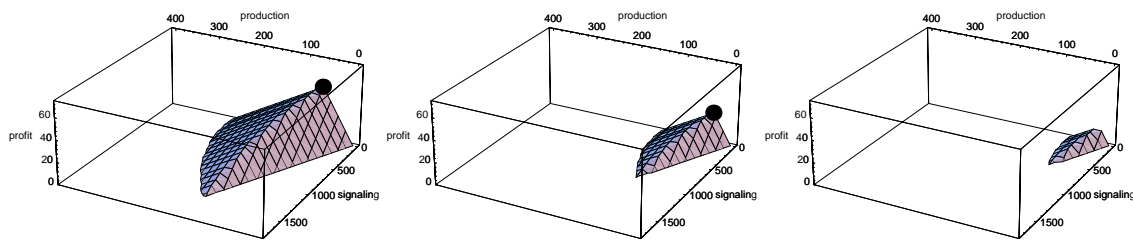
Figure 1 shrinking through immediate profit maximization

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.
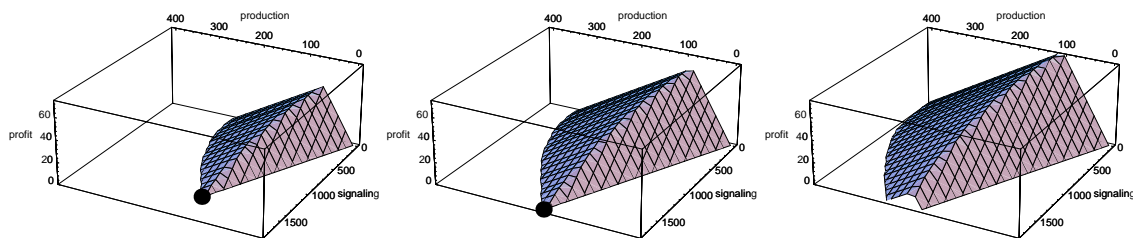
Figure 2 market build-up through sales maximization, with immediate profits>0

A second aspect of this game that is worth to be stressed is the influence that each player's actions have on the outcomes of other players. While one firm might try to walk up to a peak in its profit landscape, this landscape is deformed continuously by the other players who might be trying to reach their peak. 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 less vulnerable to this noise in the sense that it will less easily lead 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 voluntary inactivity, this implies more positive feedback.
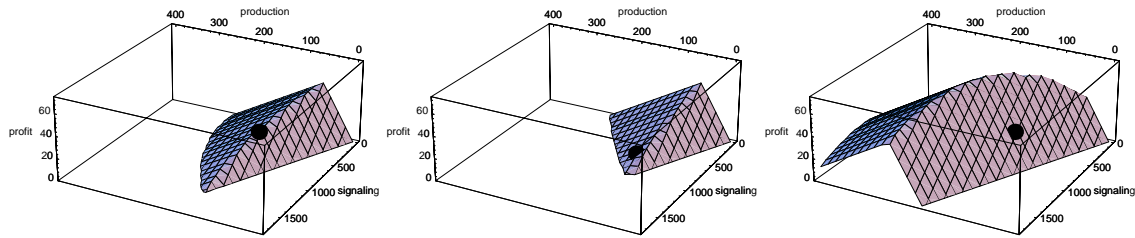
Figure 3 fluctuations through actions other players

5. Data Analysis: Unexperienced Players

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 adaptively. As observed in section 3, the players learn to choose actions that are on average close to a symmetric equilibrium, but there are large differences between the actions of the players. Figure 4 shows the distribution of the individual unexperienced players' signaling and production levels, averaged over the periods 81-130. The arrow indicates the symmetric 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.
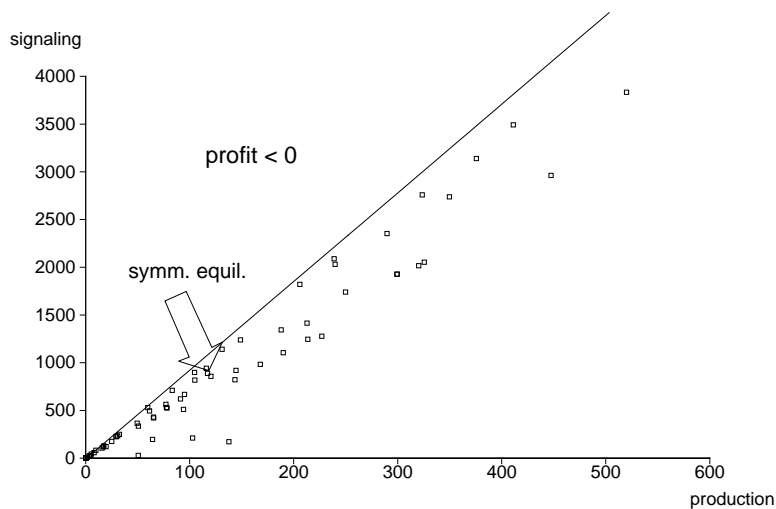
Figure 4 distribution actions unexperienced players, periods 81-130

This leads to two questions. First, how do these differences in actions correspond to differences in performance? Second, how, then, do we arrive at this distribution? In other words, in what sense does the behavior of some players differ from that of other players?

Observation 2: There are considerable differences in performance among the players. We can distinguish three categories of players. 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. Just like the category I players, the category III players try higher signaling (and production) levels than the category II players, but they are less successful than the category I players.

A method to measure the difference in performance among the players is calculating a Gini coefficient (see, e.g., Case & Fair [1996]). A Gini coefficient measures the skewness in the wealth distribution of a population, using the Lorenz curves. 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 unexperienced players is 0.41.[14] Given this unequal performance, what does the distribution look like, and what is its relation to the actions chosen? In figure 5.a we order all 66 unexperienced players on their cumulative profit realized per period, and in figure 5.b we present for these same players their average signaling.[15] 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 boundaries for which the difference between the average signaling in the middle category and the other two categories combined is maximized.[16] We will use these three categories in our subsequent analysis, to see whether we can identify differences in the behavior between these three groups of players. The numbers in figure 5.a and 5.b give the values of profits and signaling respectively for the observations next to the boundaries.

---

[14] 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.

[15] The averages are taken over the periods that a player was active, i.e. until he went bankrupt or the end of the session was reached. Adding production levels would yield little extra information since average production and signaling are almost perfectly correlated.

[16] We imposed the additional restriction that there should be at least 3 players per category.
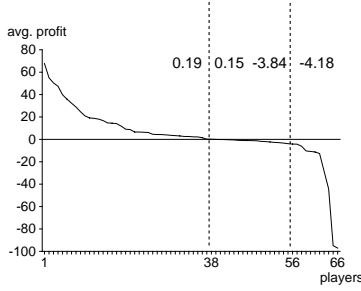
Figure 5.a average profit
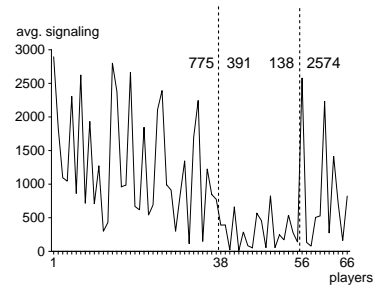
Figure 5.b average signaling

Table 2 illustrates this categorization further by giving the average signaling, production, and profit levels per category as shown in figures 5.a and 5.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%).

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

Table 2 averages for the three categories

The question, then, is where these differences between the players' actions and outcomes come from.[17] First, we take a look at the players' decisions in the very first period, and we analyze whether there is a significant correlation between the individual players' experience during the initial periods and their experience during later periods.

Observation 3: Both production and signaling levels in the first period are concentrated on focal points. And 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 is 0.55 (significant at 0.0% level; 1-sided t-test).

---

[17] 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.

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%) choose production levels greater than 1000. 61 out of 66 players (92%) choose a multiple of 50, and 53 (80%) pick 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%) choose multiples of 50 or 100, and 55 (83%) 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 D.1 in appendix D, where we give for each unexperienced player his initial period actions and outcomes, and his averages over his whole playing history. The question one has to address is, once observed such a correlation, where does it stem from? In section 4 we identified various positive feedback mechanisms in our experiment. 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 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 4: 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 category II players, who are in turn much more inclined to do this than category I players.

Table 3 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.[18]

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

Table 3 shrinking customer pool by not signaling, with production>0

A second positive feedback effect presented in the previous section 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 5: Voluntarily, i.e., 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 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 into business, with category I players least likely to return into business.

Table 4 illustrates the voluntarily switching to inactivity.[19] 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 switch to inactivity more often than category I players (significant at 0.0% level), category II 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.

---

[18] 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%).

[19] Recall that category III players realized lower average profits than category II players.

players	# observations		rel. frequencies (%) inactivity	
	profit<0	profit≥0	after profit<0	after profit≥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

Table 4 relative frequency switching to voluntary inactivity

One can also look at it from the other side. Once switched to inactivity, how often does a player decide to come back into business? Table 5 illustrates this. In 84% of the cases voluntarily inactive players remain inactive in the next period. Comparing the frequencies of returning into business by the individual players in the 3 categories, players in category I return less often into business than category II players (significant at 0.9% level with 2-sided Wilcoxon test). The differences between category I and category III players, and between category II and category III players are not significant. Perhaps the fact that the probability to return is the lowest for category I means that these players almost never choose inactivity, but when they do it, they are slightly more serious about it.

players	absolute frequencies		rel. frequencies (%)
	back into business	stay inactive	back into business
all	59	307	16.1
cat. I	5	31	13.9
cat. II	46	238	16.2
cat. III	8	38	17.4

Table 5 frequencies returning into business after choice of inactivity

After analyzing the players' actions in these particular cases, we now turn to an analysis of the players' behavior in a more general sense. 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 the optimal action would be, given the actions of the other players, completely independent from any payoff actually experienced, and they play a best-response strategy against the frequency distribution of a given (sub-)population of other players. The evolutionary dynamics consist of a co-evolutionary 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 *'supervised'* learning techniques is very limited. The agents do not have a well-specified model of their environment, and they do not know which action would be the best response. Hence, the very first task for our players, is to *learn* what good actions would be. 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 concerning the *number of signals* to be sent, and in the *second* step adjusting the *production level* to the level of the demand generated. 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 & Stoecker [1986], or Nagel [1995]). It is perhaps best illustrated by the following example given in Selten & 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 qualitative picture of the causal relationship between the position of his bow and the path of the arrow."* (Selten & Buchta [1994], p. 9). As we see, learning direction theory is based on the idea that the structure of the payoff function is known by the player, and can be exploited. 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 to reason in which direction better actions could have been found. 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. 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.[20] 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 following sense. It predicts a player's action on the basis of his action and outcome in the last period *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 6. 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

---

[20] 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.

analyze below, the demand depends upon the signaling level. Therefore, here we only consider those cases in which the players did not move into the opposite direction with their signaling level.[21]

if	then
production$_t$ < demand$_t$	production$_{t+1}$ ≥ production$_t$
production$_t$ > demand$_t$	production$_{t+1}$ ≤ production$_t$
production$_t$ = demand$_t$	n.a.

Table 6 predictions learning direction theory

Observation 6: The players change their production level into a direction that would be wrong according to learning direction theory in only 9% of the cases where it makes a prediction. But there is an asymmetry in the success of learning direction theory between the cases where production was too low, and where it was too high. This asymmetry seems related to the fact that the players are less boundedly rational than this theory assumes, and this applies in particular to the category I players.

Figures 6.a and 6.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. If production was too low (1250 observations), learning 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 6. Production was equal to demand in only 8% of the cases (510 observations).
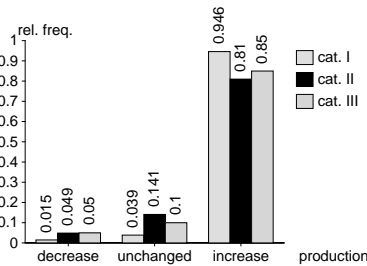
Figure 6.a direction learning
after production < demand
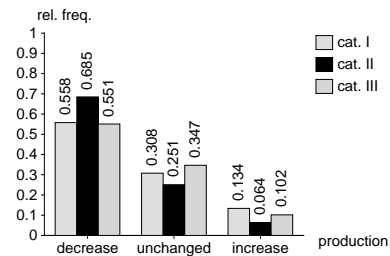
Figure 6.b direction learning
after production > demand

---

[21] 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. If we neglect this condition on signaling, considering the three categories together, the percentages for decrease, unchanged, and increase would be 5, 11, and 84 for the case that production was less than demand, and 47, 31, and 23 for the case that production was greater than demand.

Remains the question how to explain the asymmetry between the cases where production was too low, and where it was too high in the preceding period. Learning direction theory makes the prediction on the basis of the last period. 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 is *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 what their profit could have been (ex post), but to maximize their expected profits (ex ante). 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: $E[\Pi(z)] = p \cdot \left\{ \sum_{q=0}^{z} q \cdot f[q] + z \cdot \sum_{q=z}^{\infty} f[q] \right\} - c \cdot z - k \cdot s$. As can be easily shown: $\Delta E[\Pi]/\Delta z = p \cdot (1 - F[z]) - c$. Hence, expected profit is maximized when $F[z] = 1 - c/p$. That is, $F[z] > 0.5$ if $c/p < 0.5$, which was the case with both unexperienced and experienced players. In other words, the ex ante optimal production level is higher than the ex post average demand. And the lower the marginal costs of production, the more overproduction is profitable. This explains the asymmetry in the learning direction results given above. Players appear more reluctant to decrease their production level when it was too high, because they understand that production should be higher than average demand. That is, 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 that production was less than demand, than in the case in which production was higher than demand. Also it turns out that *all* players on average overproduce; with the overall average production 1.20 times average demand.[22]

But in figures 6.a and 6.b we see also some differences between the categories, as far as the propensity to follow the learning direction theory hypothesis. Comparing the frequencies of increasing production after production was less than demand (figure 6.a), 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 also 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 after production was higher than demand (figure 6.b), players in category II decrease their production more often than category I players (significant at 1.0% level with 1-sided Wilcoxon test), and also less often than category III players (significant at 1.1%). The difference between category I and category III players

---

[22] In all the sessions, there is exactly one player who is an exception to this. The average production for this experienced player was 0.998 times his average demand. He was an extreme direction learner. Out of 221 decisions, he *never* went into the wrong direction with his production level.

is not significant. Hence, it seems category I players understand best the desirability of overproduction, while category II players understand this least well, and as a result become more easily small players.

Signaling: Hill Climbing

As said 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. Notice that learning direction theory cannot predict much with respect to signaling. In the case that demand was 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 was 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. And also when the demand faced by a firm equals its production, it does not know in which direction to adjust its signaling. In our closer look at the game in section 4, we showed that a player's opportunities could be depicted as a hill. The objective of a player is to find the top of the hill. A problem for a player is that he does not know how his 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 into some direction, if one gets a higher payoff one continues from there, otherwise one goes back to try another direction. Eventually one should reach a top.[23] We conjecture that the players' adaptive behavior in signaling space can be described by such a *hill climbing*, or gradient, algorithm.

In order to explain the essence of hill climbing, and the difference 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 blind-folded. After each trial the only feedback he gets from his environment is the level of enthusiasm with which the crowd of spectators reacts. The closer he gets to the optimum, the louder they are expected to shout. Therefore, after each two trials he can compare the levels of payoff, and decides to try next time in the neighborhood of the one where 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 7.

---

[23] This might be a local top only. Simulated annealing is a one step 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.

if		then
(1)	signaling$_t$ < signaling$_{t-1}$ and payoff$_t$ < payoff$_{t-1}$	signaling$_{t+1}$ > signaling$_t$
(2)	signaling$_t$ < signaling$_{t-1}$ and payoff$_t$ > payoff$_{t-1}$	signaling$_{t+1}$ < signaling$_{t-1}$
(3)	signaling$_t$ > signaling$_{t-1}$ and payoff$_t$ < payoff$_{t-1}$	signaling$_{t+1}$ < signaling$_t$
(4)	signaling$_t$ > signaling$_{t-1}$ and payoff$_t$ > payoff$_{t-1}$	signaling$_{t+1}$ > signaling$_{t-1}$
(5)	signaling$_t$ = signaling$_{t-1}$ or payoff$_t$ = payoff$_{t-1}$	n.a.

Table 7 predictions hill climbing

Before analyzing how far the players were hill climbers, we need to define the payoffs that form the hill. The islands shown in section 4 were for immediate profits. But as argued already there, given the *dynamics* of the demand generated by the signals sent and the patronizing customers, a player should look further ahead. Above we showed that players could boost their immediate profits by signaling very little, i.e., by eating into their pool of customers. Future profits are adversely affected by this. The cause is that 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 4. Hence, when considering the question 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 profits that can be extracted from them in later periods.[24] 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: $\Pi + \Delta x \cdot (p\text{-}c) \cdot \sum_{t=1}^{\infty} f^t$. In our analysis of hill climbing, we will distinguish both the case in which players take into account this lookahead payoff, and the case in which players go myopically for their immediate profits only. 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 the altitude of their lookahead hill explicitly. 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 consider the question how often the players do behave *'as if'* they are hill climbing having learned these lookahead payoff values correctly.

---

[24] 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.

Observation 7: 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.
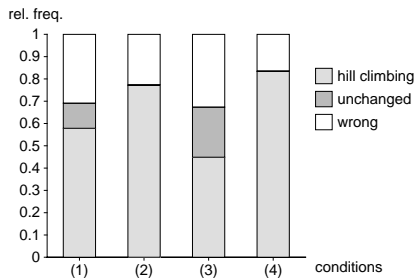
Figure 7.a myopic hill climbing
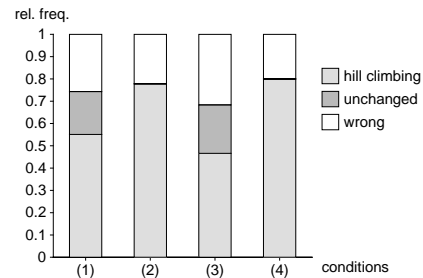
Figure 7.b lookahead hill climbing

Figures 7.a and 7.b give the percentages of correct and wrong predictions by the hill climbing hypothesis for myopic and lookahead climbing.[25] The numbers in parentheses on the horizontal axis denote the 'if ...' conditions as given in table 7. 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. In the 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 according to that hypothesis. 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 discussed above. Second, given the noise caused by the other players, it might be wise not to put all 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, like reinforcement learning (see, e.g., Roth & Erev [1995]).

Observation 8: There is an asymmetry between the cases when a player's payoff had increased and when 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.

---

[25] 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 7, and occurred in 33% of the cases. The absolute frequencies for the cases (1) to (4) in figure 7.a are 560, 2311, 2997, and 810. In figure 7.b these frequencies are 1305, 1583, 2121, and 1710.

For convenience, we consider here only lookahead hill climbing. Compare in figures 7.a and 7.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. Hence, 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. Hence, 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 that successful ones are reversed, is a more general psychological feature.

As we see, figure 7.a and 7.b are very similar. A first explanation is the following. 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, in those cases the player's immediate profits 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. Hence, we now consider only the other 29% cases of opposite gradients. The question is, which players tend to go for the immediate profits, and which players look more often ahead?

Observation 9: Category I players look ahead most often. Category II players are the least frequently looking ahead.

Table 8 shows the frequencies. As we see, on average players are inclined just a little bit towards looking ahead. And this is a second explanation why the myopic and lookahead hill climbing pictures in figures 7.a and 7.b look so similar. We also 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 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.

| 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

Table 8 frequencies myopic vs. lookahead hill climbing

We have seen that the 2-step model we proposed does not perfectly describe the behavior of the players, and there are important differences between the players. But at the same time, the attraction of the model is that it is simple. A question, then, is whether this simple model would generate a pattern of behavior of the players that would converge to the same average level as observed in the experiments. To answer this question, 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 level observed during the experiments in the first period, and restrict their choices to the same domain. Players follow the learning direction theory hypothesis for production as outlined above, where the size of their step is a draw from a N(0,5) distribution, and they do not change their production level if their signaling decision for that period points into the opposite direction. The step size in their hill climbing for signaling is also stochastic, using a N(0,10) distribution for the noise around the new starting value, but always such that they stay at the correct side of the discarded signaling value that led to the lower payoff. All players are modeled identically, but independently, which implies that their paths may diverge over time due to the stochastic factors. In figure 8 we present the average behavior of 11 simulated sessions with 6 players, and the average signaling levels observed in the experiments. It should be stressed that the two curves have a different time scale. The main thing to notice is the convergence to a similar signaling level. Moreover, we also observe that the graph shows a similar initial dip.
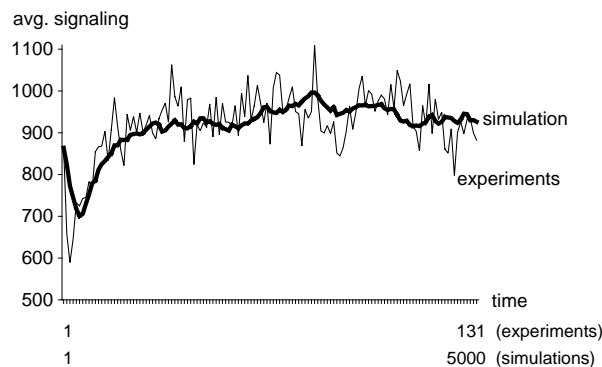
Figure 8 simulation 2-step model vs. experimental data

6. Data Analysis: What Have the Experienced Players Learned?

In this section we turn to an analysis of the experienced players. The basic question we want to address is: what have they learned? Before considering the specifics of the sessions with experienced

players, a relevant question is which players came back.[26] In other words, how far had the players discovered whether the game was worthwhile?

Observation 10: The group of players that came back has a selection bias. The better their average profit as unexperienced player, the more likely they were to come back. None of the four bankrupt players came back. Survival as an unexperienced player did not guarantee survival as an experienced player (five went bankrupt). But the players who got experienced in the strongly reduced strategy space (999 instead of 4999 as maximum values for signaling and production) were at a disadvantage.

Table 9 show the percentages of players that came back, and those that did not, per category.[27] We use a χ^2-test to compare the probability for category I players to come back with this probability for category II and III players combined. Category I players appear much more likely to come back, but due to the small number of observations this is significant only at the 9.3% level. This seems an example of reinforcement learning. They tried the game as an unexperienced player, and the higher the payoff realized, the more likely it is to play again.[28]

players	absolute frequencies		rel. freq. (%)
	yes	no	yes
all	30	48	38
cat. I	22	26	46
cat. II	5	13	28
cat. III	3	9	25

Table 9 which players came back

Just like with the unexperienced players, we can distinguish three categories of players: the successful ones, the 'nil players', and the unsuccessful ones. With the unexperienced players we had one treatment (sessions 91 and 92) in which there was a strongly reduced strategy space: 1000 instead of 5000. Four of the players of those two sessions came back. As unexperienced players these four would have been all four category I players, but as experienced players they ended up all four in category III, and two of them went bankrupt. This suggests that the other players have learned something in the larger strategy space that turned out useful as experienced players.

Recall that the only difference in the setup with the experienced players is that we changed the parameter values (see table B.1 in appendix B). Everything else, also in the instructions, remained the same. We explained this to the players. Hence, we had a game with the same structure, but with different numerical values. The reason was that we are interested in the question whether the players

---

[26] We tried to get all players back. First, they got a letter inviting them to participate in a similar experiment, and a couple of days later we contacted them all by phone.

[27] Including the sessions 91 and 92.

[28] Garvin & Kagel [1994] also find that aggressive, unsuccessful players are less likely to play again.

had learned something of the general structure of the game, and not whether players that had happened to find good actions as unexperienced players would simply continue to choose those specific numbers. A first question, then, is whether the players understood right from the start that their opportunities were different in a quantitative sense.

Observation 11: The experienced players' choices for production and signaling levels are less concentrated on focal points than with the unexperienced players. But that is not because the players simply continue with those values they had learned during the game as unexperienced players.

The production level chosen in the very first period is again rarely greater than 1000; only 2 players (7%) against 1 unexperienced (3%). Multiples of 50 are again popular (22 players, or 73%). But multiples of 100 are less popular with experienced players (17 players, or 57%) than with unexperienced (23 players, or 77%), which is a difference significant at 10% (χ^2-test). The most chosen production levels are 100 (4 times), 200 (3), and 300 (3). Hence, the midpoint of 500 has no focal point character anymore for production levels. This is a significant difference with the unexperienced players at 1.0% (χ^2-test). With respect to signaling, multiples 50 and 100 are still chosen often (25 times, or 83%, and 18 times, or 60%). The most chosen signaling levels are 300, 500, 1000 (3 times each, or 10%). Concerning the ratio between signaling and production, with unexperienced we had 16 players out of 30 (53%) who choose signaling greater than production, while 24 experienced players (80%) choose signaling greater than production. This is significant at 3.0% (χ^2-test). Thus, they have learned that signaling is relatively important.

The correlation coefficient between the numbers learned as unexperienced players (taking for each individual player his average signaling in the last 50 periods), and the values chosen in the initial period as experienced players is only 0.14 and not significant (t-test).[29] Hence, already in the first period, the experienced players understand that the different parameters imply different opportunities. Table D.2 in appendix D gives the initial period actions and outcomes for the experienced players, as well as their later averages. The next question then is, how well do the players on average discover these changed overall market opportunities during the experiment? And is their market experience during the initial periods again an important factor explaining their performance in later periods?

Observation 12: Also for the experienced players there is a significant positive correlation between sales in the initial periods and sales in later periods. The correlation coefficient between the 30 individual players' average sales levels in the periods 1-10 and the periods 81-130 is 0.57 (significant at 0.0% level; 1-sided t-test), and even between sales levels in the periods 1-10 and the periods 251-300 there is a positive correlation (coefficient of 0.37; significant at 2.2%).

---

[29] Including the sessions 91 and 92.

Next, we turn to our 2-step model of learning direction theory for production, and hill climbing for signaling. Do the experienced players behave differently? And in particular, we want to know whether the players have learned to look ahead more often than before.[30]

Observation 13: The experienced players adaptive behavior as described by our 2-step model is similar to that of the unexperienced players.

With respect to the learning direction hypothesis concerning the adjustment of the production level, there is little difference between unexperienced and experienced players. Unexperienced adjusted their production level into the wrong direction in 11.4% of the cases, and experienced in 9.5%. Using the Wilcoxon test to analyze whether these levels of the individual players are drawn from the same distribution for unexperienced and experienced players, we find that the null hypothesis 'no difference' cannot be rejected. The frequencies with which the experienced players tend to climb hills are similar to those for the unexperienced players. They go into a strictly wrong direction with signaling in 28.5% of the cases, compared with 24.2% for the unexperienced for lookahead hill climbing. Using the Wilcoxon test as above, this difference is not significant. And there is also no significant difference between the unexperienced and experienced players as far as their inclination to look ahead is concerned (Wilcoxon test); 54.3% for unexperienced against 54.2% for experienced in those cases in which the payoff gradient for immediate profits and the one for immediate plus future profits pointed into a different direction.

In section 5, we showed that although average actions were close to the equilibrium ones, there were strong differences between the individual players. Since the experienced players came from different sessions, the following two questions arise. First, did the experienced players have a significantly different background? And if so, second question, does this imply differences for their behavior as experienced players?

Observation 14: There were differences in the players' experience as far as their environment was concerned. And players that got their experience in a more cooperative environment tend to play more cooperatively as experienced players.

Table D.3 in appendix D gives the average signaling per period per player for each of the 11 sessions with unexperienced players. Clearly, in some sessions there was more aggregate signaling than in other sessions. If players signal very little, not all potential customers might be reached, and some might not

---

[30] To avoid selection bias, when comparing with the unexperienced players, we consider here only those 30 unexperienced players that came back. This also implies we now include the four players of sessions 91 and 92 that came back, and that were excluded in the analyses in the previous section. More details concerning this comparison are available upon request.

show up. But very soon almost all customers are reached.[31] If firms were cooperative, they could keep signaling low, but each individual firm would have an incentive to increase its signaling. Hence, we can see the average signaling level in a session as a measure of cooperation among the firms in that session. What is the effect for the individual firms of this level of average signaling? Whatever their sales level, an increase in the aggregate signaling implies that the signaling costs per unit of sales go up. There are two causes for this. First, a firm's sales go down. Second, when a firm reacts, increasing its signaling, its signaling costs go up. Take as an example the two sessions 1 and 9. In session 1 the average signaling per firm was 690, whereas it was 1043 in session 9. The result of this aggressive competition in session 9 was that the signaling costs per unit sold were much higher. This applies approximately equally to all players in a session, whatever their size. On average signaling costs per unit of sales were 0.47 in session 1 (with a minimum of 0.37, and a maximum of 0.49), and 0.71 in session 9 (ranging from 0.68 to 0.82). Since there are also costs of unsold stocks, in session 9 the aggressive signaling environment makes that the profit margin was almost completely eroded, and it was difficult to make positive profits.

The question, then, is whether this different background implies differences for their behavior as experienced players? Table 10 shows the average signaling per firm for the sessions with experienced players, the numbers of the sessions where they had got their experience, and a weighted average of the average signaling activity in those sessions.[32] The correlation coefficient between the measure of cooperation as unexperienced and as experienced is 0.94 (significant at 2.9%; 1-sided t-test).

session experienced	avg. sign. experienced	from sessions unexperienced	weighted avg. sign. unexperienced
21	592	4x#1, 2x#2	755
23	770	2x#3, 1x#4, 3x#9	1005
24	666	2x#5, 4x#6	945
25	769	3x#7, 3x#11	1023

Table 10 influence environment on cooperation

7. Conclusions

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. There were, however, important differences between the players; both with respect

---

[31] With average signaling of 359 per firm, 95% of the customers will be reached, and with 552 signals per firm this is 99%. The symmetric stationary equilibrium strategy implies sending 927 signals per firm. The aggregate effect of the last 375 of these signals per firm is that three additional customers will be attracted to the market.

[32] We omit here session 22 because 4 players came from sessions 91 and 92, in which production and signaling had an upper limit of 999 instead of 4999. That is, they had got their experience in a different environment anyway, independent from their own attitude towards aggressiveness or cooperation.

to their actions and to their payoffs. In particular, we distinguished three categories of players: 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. 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. Based on the game-theoretic analysis, we proposed a 2-step model, 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 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 reverse the direction their signaling was moving into, they often continued walking downhill. This might be a more general psychological feature. 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, like e.g. reinforcement learning (see Roth & Erev [1995]).[33] Using the hill climbing hypothesis, we analyzed how far the players were inclined to go myopically for immediate profits, or looked further ahead. 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 model seems to offer a reasonable explanation for the average market outcomes.

We showed that the group of players that came back as experienced players had a clear selection bias; with the successful players being almost twice as likely to play again. The experienced players' initial period actions had a different distribution than the unexperienced one, which suggests that they have learned something, but they did not simply start with those values they had learned as unexperienced players. With respect to our 2-step model, the experienced players' behavior was not significantly different from that of the unexperienced. We showed that competitiveness of the environment in which the players got their experience had a significant influence on their behavior as experienced players.

---

[33] 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.

Appendix A. Instructions, and Computer Screen

Table A.1 contains the English version of the instructions given 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.
* **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.

Table A.1 instructions to the players

Figure A.1 shows the computer screen as viewed by a player acting as firm 'X' in a given period.

```
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
----------------------------------------------------------------------------
```

Figure A.1 computer screen firm 'X'

Appendix B. Notation, Variables, and Parameter Values

Table B.1 gives the notation used throughout the paper. Superscripts will be used for the time index, and subscripts for the identity of a firm. In addition, the table gives an overview of the parameter values used. The last column indicates whether the parameter value was known or not to the players. Recall that in addition to these parameter values, the players did not know the functional form of the demand.

symbol	meaning	unexperienced (1-10, and 91-92)	experienced (21-25)	known
c	*'marginal'* cost production	0.25	0.27	yes
f	patronage rate satisfied consumers	0.56	0.63	no
g	price minus *'marginal'* cost production	0.75	0.73	yes
k	*'marginal'* cost signaling	0.08	0.21	yes
m	# firms	6	6	no
n	# consumers	712	1422	no
N	total # agents	718	1428	no
p	price of the commodity	1.00	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[34]	4999	yes
S	aggregate # signals all firms	-	-	no
S_	aggregate # signals other firms	-	-	no
V	value	-	-	no
x	sales	-	-	own
z	production	-	-	own
-	maximum value for z	4999[35]	4999	yes
-	# periods	± 150	± 300	no

Table B.1 notation and parameter values

Appendix C. Game-theoretic Analysis

The profit function is 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: $q_i^t = round(trunc[f \cdot x_i^{t-1}] + \dfrac{s_i^t}{S^t} \cdot [1 - \exp(-\dfrac{S}{N})] \cdot [n - \sum_{j=1}^{m} trunc(f \cdot x_j^{t-1})])$.

The demand function is deterministic. Hence, z=q=x. 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 in period T-t'-1 and the value V in period T-t': $V_i^{T-t'-1} = \Pi_i^{T-t'-1} + V_i^{T-t'}$. Hence, the first-order condition is:

$$\frac{\partial V_i^{T-t'-1}}{\partial s_i^{T-t'-1}} = \frac{\partial \Pi_i^{T-t'-1}}{\partial s_i^{T-t'-1}} + \frac{\partial V_i^{T-t'}}{\partial s_i^{T-t'-1}} = 0$$. We will now consider these two terms separately.

---

[34] Two sessions (91 and 92) were played in which the players faced an upper limit of 999 instead of 4999 for their production and signaling decision variables. Unless otherwise noticed explicitly, these two sessions are excluded from the analyses.

[35] 999 in sessions 91 and 92.

Determination of first term: The profit function can be rewritten as: $\mathbb{II}_i^t = g \cdot q_i^t - k \cdot s_i^t \Rightarrow$

$$\frac{\partial \mathbb{II}_i^t}{\partial s_i^t} = g \cdot \frac{\partial q_i^t}{\partial s_i^t} - k$$. Neglecting the term [1-exp(-S/N)], and the roundings and truncations, demand is

given by: $q_i^t = f \cdot q_i^{t-1} + \dfrac{s_i^t}{S^t} \cdot (n - f \cdot Q^{t-1}) \Rightarrow \dfrac{\partial q_i^t}{\partial s_i^t} = \dfrac{S_{-i}^t}{(S^t)^2} \cdot (n - f \cdot Q^{t-1})$. Since all

consumers visit a firm: $Q^{t-1} = n$. Hence, we get: $\dfrac{\partial \mathbb{II}_i^t}{\partial s_i^t} = g \cdot \dfrac{S_{-i}^t}{(S^t)^2} \cdot n \cdot (1 - f) - k \Rightarrow$

$$\frac{\partial \mathbb{II}_i^{T-t'-1}}{\partial s_i^{T-t'-1}} = g \cdot \frac{S_{-i}^{T-t'-1}}{(S^{T-t'-1})^2} \cdot n \cdot (1 - f) - k$$.

Determination of the second term: $\dfrac{\partial \mathbb{II}_i^t}{\partial s_i^t} = g \cdot \dfrac{S_{-i}^t}{(S^t)^2} \cdot (n - f \cdot Q^{t-1}) - k$. First-order condition:

$$\frac{\partial \mathbb{II}_i^T}{\partial s_i^T} = 0 \Rightarrow \frac{k}{g} = \frac{m-1}{m} \cdot \frac{1}{S^T} \cdot (n - f \cdot Q^{T-1}) \Rightarrow$$

$$S^T = \frac{g}{k} \cdot \frac{m-1}{m} \cdot (n - f \cdot Q^{T-1}) \Rightarrow s_i^T = \frac{g}{k} \cdot \frac{m-1}{m^2} \cdot (n - f \cdot Q^{T-1})$$. We know:

$Q^{T-1} = n$. Hence: $s_i^T = \dfrac{g}{k} \cdot \dfrac{m-1}{m^2} \cdot n \cdot (1 - f)$. Since $\mathbb{II}_i^t = g \cdot q_i^t - k \cdot s_i^t$, we get

$$V_i^T = g \cdot [f \cdot q_i^{T-1} + \frac{1}{m} \cdot (n - f \cdot Q^{T-1})] - k \cdot [\frac{g}{k} \cdot (\frac{m-1}{m^2}) \cdot n \cdot (1 - f)]$$. Since

$Q^{T-1} = n$, we get: $V_i^T = g \cdot [f \cdot q_i^{T-1} + \dfrac{n}{m^2} \cdot (1 - f)] \Rightarrow V_i^T = g \cdot [A_0 \cdot q_i^{T-1} + B_0]$.

More in general: $V_i^{T-t'} = g \cdot [A_{t'} \cdot q_i^{T-t'-1} + B_{t'}]$, where: $A_0 = f$, $B_0 = \dfrac{1}{m^2} \cdot (1 - f) \cdot n$,

and: $A_{t'+1} = f + f \cdot A_{t'} \Rightarrow \dfrac{\partial V_i^{T-t'}}{\partial s_i^{T-t'-1}} = g \cdot A_{t'} \cdot \dfrac{\partial q_i^{T-t'-1}}{\partial s_i^{T-t'-1}} \Rightarrow$

$$\frac{\partial V_i^{T-t'}}{\partial s_i^{T-t'-1}} = g \cdot A_{t'} \cdot \frac{S_{-i}^{T}}{(S^{T})^2} \cdot n \cdot (1-f) \ .$$

Combining the two terms we get:

$$\frac{\partial V_i^{T-t'-1}}{\partial s_i^{T-t'-1}} = [g \cdot \frac{S_{-i}^{T-t'-1}}{(S^{T-t'-1})^2} \cdot n \cdot (1-f) - k] + [g \cdot A_{t'} \cdot \frac{S_{-i}^{T-t'-1}}{(S^{T-t'-1})^2} \cdot n \cdot (1-f)] = 0 \Rightarrow$$

$$g \cdot (1 + A_{t'}) \cdot (\frac{m-1}{m}) \cdot \frac{1}{S^{T-t'-1}} \cdot (1-f) \cdot n = k \Rightarrow$$

$$s_i^{T-t'} = \frac{g}{k} \cdot (1 + A_{t'+1}) \cdot (\frac{m-1}{m^2}) \cdot (1-f) \cdot n \quad . \text{ Now consider the difference equation}$$

$A_{t'+1} = f + f \cdot A_{t'}$, which can be solved as: $A_{t'+1} = [f - \dfrac{f}{1-f}] \cdot f^{t'+1} + \dfrac{f}{1-f}$. Hence,

assuming the game is played an infinite number of periods, the optimal action in a given period is:

$$\lim_{t'\to\infty} \lim_{T\to\infty} s_i^{T-t'} = \frac{g}{k} \cdot (1 + \frac{f}{1-f}) \cdot (\frac{m-1}{m^2}) \cdot (1-f) \cdot n \qquad = \frac{g}{k} \cdot (\frac{m-1}{m^2}) \cdot n \quad .$$

Appendix D. Some Additional Data

Figures D.1.a to D.1.c present the time series for signaling, production, and profits for period 1 to 131 averaged over the 66 unexperienced players. Figures D.2.a to D.2.c do the same for the 30 experienced players for periods 1 to 301. 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 graphs for signaling (D.1.a and D.2.a) we added two other benchmarks. The first one is called 'losses', and corresponds to the line drawn in figure 4. It is the signaling level beyond which positive profits are impossible, given the equilibrium production 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.
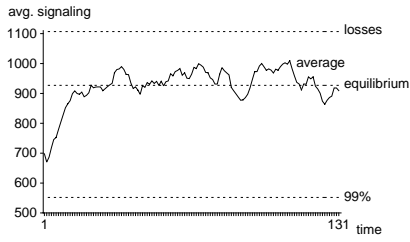
Figure D.1.a average signaling
unexperienced players, periods 1-131
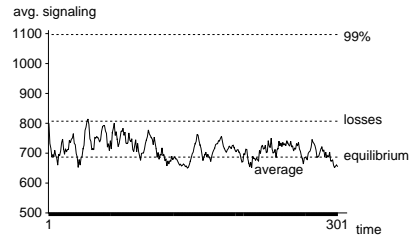
Figure D.2.a average signaling
experienced players, periods 1-301
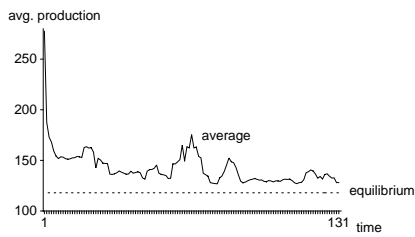
Figure D.1.b average production
unexperienced players, periods 1-131
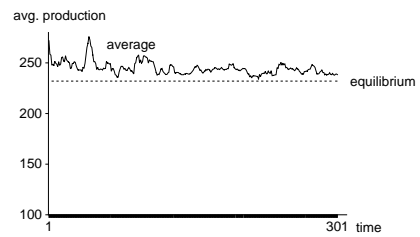
Figure D.2.b average production
experienced players, periods 1-301
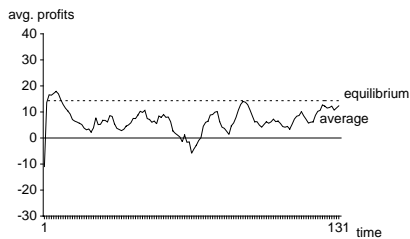
Figure D.1.c average profits
unexperienced players, periods 1-131
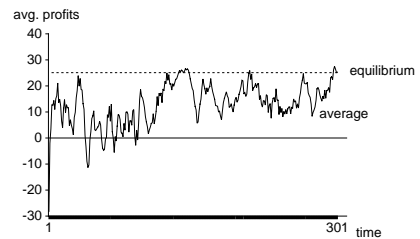
Figure D.2.c average profits
experienced players, periods 1-301

The following table D.1 presents the individual averages and first period actions for unexperienced players. The averages are taken over the periods that a player actually played. Table D.2 present analogous data for experienced players. Table D.3 gives the session averages. Those averages are taken over the periods that the session lasted.

session	player	period 1 prod.	sign.	sales	all periods prod.	sign.	sales	profits	number periods	also session	player
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	-	-
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 D.1 summary data unexperienced players

session	player	period 1 prod.	sign.	sales	all periods prod.	sign.	sales	profits	number periods	also session	player
91	1	0	67	0	180	396	84	7.4	151	-	-
91	2	0	1	0	77	82	19	-7	151	-	-
91	3	10	30	10	197	827	184	68.7	151	-	-
91	4	300	100	84	209	892	193	69.4	151	22	3
91	5	50	50	42	164	750	152	51.3	151	22	2
91	6	25	1	1	78	281	64	21.6	151	-	-
92	1	10	40	10	215	900	186	60.4	151	22	1
92	2	100	120	64	224	942	202	70.6	151	-	-
92	3	14	10	5	121	581	109	32.4	151	-	-
92	4	100	150	79	170	713	149	49.8	151	-	-
92	5	500	500	265	29	30	11	1.5	151	22	4
92	6	100	200	100	62	184	43	13.3	151	-	-

Table D.1 (cont.) summary data unexperienced players

session	player	period 1 prod.	sign.	sales	all periods prod.	sign.	sales	profits	number periods	also session	player
21	1	300	300	134	153	364	141	23	301	1	2
21	2	100	50	22	232	587	224	38.4	301	1	1
21	3	250	300	134	111	282	107	17.8	301	2	5
21	4	850	1800	805	800	2010	784	145.6	301	1	3
21	5	130	150	67	124	298	111	15.1	301	1	4
21	6	25	100	25	4	9	3	-0.5	301	2	6
22	1	120	300	55	199	631	182	-4.1	301	92	1
22	2	50	100	18	146	431	125	-5.2	301	91	5
22	3	1200	4999	919	416	1410	385	-23	92	91	4
22	4	1000	1250	230	236	268	77	-43	49	92	5
22	5	100	1000	100	937	2805	865	23.2	301	10	5
22	6	50	50	9	357	1245	329	-29.3	71	10	2
23	1	200	500	90	726	2320	718	34.9	301	4	1
23	2	30	3000	30	416	1644	377	-80.2	25	3	5
23	3	1100	3000	537	665	2058	594	-18.1	113	3	2
23	4	100	350	63	205	665	199	3.6	301	9	6
23	5	20	60	11	66	215	63	-0.3	301	9	3
23	6	50	1000	50	162	510	154	3.6	301	9	4
24	1	11	17	4	69	126	45	0.1	301	6	6
24	2	400	4000	400	1012	2697	911	72	301	6	3
24	3	800	600	158	25	46	17	0.4	301	6	1
24	4	60	50	13	16	38	12	0.2	301	5	6
24	5	30	90	24	112	323	99	0.8	301	6	5
24	6	300	500	132	271	765	253	19.1	301	5	4
25	1	300	500	133	75	218	66	-0.5	301	7	3
25	2	100	1000	100	285	951	280	3.8	301	11	5
25	3	400	1500	399	429	1398	420	11.3	301	7	2
25	4	700	1500	399	272	834	257	8.1	301	11	1
25	5	200	660	175	173	548	166	4.3	301	11	2
25	6	200	50	13	212	667	201	3.4	301	7	6

Table D.2 summary data experienced players

	period 1			all periods				number
session	prod.	sign.	sales	prod.	sign.	sales	profits	periods
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
21	276	450	198	237	592	228	40	301
22	420	1283	222	255	773	230	-1	301
23	250	1318	130	241	770	231	5	301
24	267	876	122	251	666	223	15	301
25	317	868	203	241	769	232	5	301

Table D.3 session averages

References

Arthur, W.B. (1992). *On Learning and Adaptation in the Economy* (Working Paper 92-07-038) Santa Fe Institute.

Atkinson, R.C., & Suppes, P. (1958). An Analysis of Two-person Game Situations in Terms of Statistical Learning Theory. *Journal of Experimental Psychology*, *55*, 369-378.

Case, K.E., & Fair, R.C. (1996). *Principles of Microeconomics* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Davis, D.D., & Holt. C.A. (1992). *Experimental Economics*. Princeton, NJ: Princeton University Press.

Ellison, G. (1993). Learning, Local Interaction, and Coordination. *Econometrica*, *61*, 1047-1071.

Garvin, S., & Kagel, J.H. (1994). Learning in Common Value Auctions: Some Initial Observations. *Journal of Economic Behavior and Organization*, *25*, 351-372.

Green, J.R. (1983). Comment on "A. Kirman, On Mistaken Beliefs and Resultant Equilibria". In R. Frydman & E.S. Phelps (Eds.), *Individual Forecasting and Aggregate Outcomes: 'Rational Expectations Examined'*. Cambridge: Cambridge University Press.

Holland, J.H., Holyoak, K.J., Nisbett, R.E., & Thagard, P.R. (1986). *Induction: Processes of Inference, Learning, and Discovery*. Cambridge, MA: MIT Press.

Kampmann, C., & Sterman, J.D. (1995). *Feedback Complexity, Bounded Rationality, and Market Dynamics* (mimeo).

Kandori, M., Mailath, G.J., & Rob, R. (1993). Learning, Mutation, and Long Run Equilibria in Games. *Econometrica*, *61*, 29-56.

Kiefer, N.M., & Nyarko, Y. (1989). Optimal Control of an Unknown Linear Process with Learning. *International Economic Review*, *30*, 571-586.

Kirman, A.P. (1983). On Mistaken Beliefs and Resultant Equilibria. In R. Frydman & E.S. Phelps (Eds.), *Individual Forecasting and Aggregate Outcomes: 'Rational Expectations Examined'*. Cambridge: Cambridge University Press.

Kirman, A.P. (1993). Learning in Oligopoly: Theory, Simulation, and Experimental Evidence. In A.P. Kirman & M. Salmon (Eds.), *Learning and Rationality in Economics*. Oxford: Blackwell.

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.

Nagel, R. (1995). Unraveling in Guessing Games: An Experimental Study. *American Economic Review*, *85*, 1313-1326.

Ochs, J. (1995). Coordination Problem. In J. Kagel & A.E. Roth (Eds.), *The Handbook of Experimental Economics* (pp. 195-252). Princeton, NJ: Princeton University Press.

Roth, A.E., & 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.

Savage, L.J. (1954). *The Foundations of Statistics*. New York: Wiley.

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., & 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.

Vriend, N.J. (1994). Artificial Intelligence and Economic Theory. In E. Hillebrand & J. Stender (Eds.), *Many-Agent Simulation and Artificial Life* (pp. 31-47). Amsterdam: IOS.

Vriend, N.J. (1996a). Rational Behavior and Economic Theory. *Journal of Economic Behavior and Organization*, *29, No. 2*, 263-285.

Vriend, N.J. (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, H.P. (1993). The Evolution of Conventions. *Econometrica*, *61*, 57-84.