

IMITATION OF SUCCESSFUL BEHAVIOUR IN COURNOT MARKETS*

Antoni Bosch-Domènech and Nicolaas J. Vriend

In an experimental standard Cournot oligopoly we test the importance of models of behaviour characterised by imitation of successful behaviour, in particular when the environment becomes more complex. We find that the players do not rely more on imitation in more demanding environments. We explain that the different pattern of output decisions in such environments seems predominantly related to a general disorientation of the players, and more specifically to a significant decrease of best-responses.

Imitation has played an important role in the recent literature on learning and adaptive behaviour in economics.¹ Imitation seems easy and straightforward, it does not use particularly many cognitive skills and it does not require much information. Imitation certainly qualifies as a ‘fast and frugal heuristics for making decisions’ (Gigerenzer and Goldstein, 1996). Not surprisingly, conventional wisdom asserts that imitation of the successful behaviour of others is likely in situations with little information or understanding. Yet imitation, although deemed easy, need not be a *realistic* mode of behaviour. To assess the extent of imitative behaviour, we set up an experimental *symmetric Cournot* market under different information treatments, some of them very conducive to imitation. The choice of a symmetric Cournot market is not accidental. First, recent theoretical results, e.g., Vega-Redondo (1997) underpin the basic qualitative insight dating back to Schaffer (1988*a*, 1988*b*, 1989) that in Cournot markets in which firms (loosely) imitate the more successful firms, there is a tendency for output levels to drift to the Walrasian competitive equilibrium. This implies that we can use the market outcomes as a first proxy for the possible presence of imitation. Second, any

* We wish to thank Xavier Herrero for his indispensable assistance in organising the experiments, and Tilman Börgers, Gary Charness, Richard Disney, Ido Erev, Jonathan Haskel, Charles Holt, Steffen Huck, Ulrich Kamecke, Francesco Luna, Paul Madden, Paola Manzini, Charles Noussair, Martin Sefton, Tilman Slembeck, Fernando Vega-Redondo, Nir Vulkan, and seminar and conference participants at Venice, Siena, Exeter, Manchester, London (QM), the London Business School, ESA (Mannheim), EARIE (Copenhagen), ESEM and EEA (Berlin), the ESRC Research Seminar in Game Theory (Kenilworth), RES (Nottingham), the Max Planck Institute (Jena), and SAET (Ischia) for discussions or comments concerning earlier versions of this paper. Two anonymous referees and the editor provided many very valuable suggestions. Financial support from the Spanish Ministry of Education and Science under contract PB98-1079 (AB-D), and from the Commission of the European Union under contracts FMRX-CT98-0238 (AB-D) and FMBICT950277 (NJV) is gratefully acknowledged. The usual disclaimer applies.

¹ For example, in evolutionary economics or industrial organisation in the Schumpeter tradition (Nelson and Winter, 1982), in evolutionary game theory with replicator dynamics and related processes (Weibull, 1995), and in computational economics with, e.g., Genetic Algorithms (Sargent, 1993). To quote Bikhchandani *et al.* (1998): ‘Social observers have long recognised imitation as important in human society. Machiavelli (1514) wrote: “Men nearly always follow the tracks made by others and proceed in their affairs by imitation.”’ (p. 152). Some people go even so far as to assert that imitation is the *sine qua non* of human culture, and that it is the systematic imitation of behaviour that sets humans apart from their ancestors; see, e.g., Blackmore (2000), or *The Economist* (1999).

symmetric game facilitates imitation of successful behaviour. Absence of imitation in such games would be damaging evidence against the prevalence of imitation.

Our experimental results show that even in complicated environments in which imitation seems straightforward, subjects are reluctant to resort to it. This suggests that subjects need not be naive learners, and that even in a difficult Cournot environment, they recognise that imitation of successful behaviour may reduce their profits. To infer from our results that imitation is not a good description of individual behaviour in markets may be stretching our observations excessively. After all, in real markets, forces may be at play that are kept out of our experiments. Nevertheless, our results certainly question any uncritical acceptance of imitation as a prevalent mode of behaviour in markets.

The remainder of this paper is organised as follows. In Section 1, we discuss Cournot oligopoly models. In Section 2, we present our experimental design. Section 3 contains the analysis of the data, and Section 4 compares our results with some other experiments, while Section 5 concludes.

1. Competitive Output in Cournot Markets

As is well known, output levels in Cournot markets may approach competitive output levels as the number of firms increases, because the Cournot–Nash equilibrium converges to the Walrasian competitive output. But there are two forms of behaviour that can sustain the competitive (Walrasian) equilibrium, even in Cournot markets with few players.

First, agents may have a preference for beating their opponents. This can be identified with what has been called *spiteful behaviour*, i.e., choosing an action that hurts oneself, but others even more. As explained in Appendix A,² in a Cournot market, a firm that stays closer to its Walrasian output tends to realise higher profits than its competitors. The classic reference in the experimental economics literature is Fouraker and Siegel (1963), which is a detailed study of (rivalistic) behaviour in oligopolies. In later experimental work (Holt, 1995, Davis, 1995), spiteful behaviour has been observed occasionally. It is asserted, then, that some players simply like to beat their opponents, even at the cost of spoiling the party; see also Levine (1998).

The second way to reach the Walrasian equilibrium in a Cournot market with few firms is through the adaptive behaviour of boundedly rational players. This approach goes back to Hamilton (1970), who introduced the dynamic implications of a spite effect in an evolutionary context, showing how a species impairing itself might increase its likelihood of surviving selection pressure if it harms other species even more. Schaffer (1988*a, b*, 1989) followed this line of argument in an economic context, showing that it is not necessarily the optimising firms that dominate markets if economic natural selection does its work, and explaining how the Walrasian competitive output would be a Symmetric Evolutionary Equilibrium in a Cournot oligopoly. Schaffer (1989) discussed two levels at which economic natural selection can take place: the level of the firms, and the level of individual

² The Appendices are to be found on the Royal Economic Society Website www.res.org

agents (managers). One could also think of such a selection process operating at the level of actions or strategies chosen by these agents (as a process of learning or adaptive behaviour). Vega-Redondo (1997) provides a dynamic analysis of this, showing that a process in which the action that led to the highest payoff in the previous period is chosen with a certain probability eventually leads to the Walrasian equilibrium in Cournot markets, provided there is some noise that eventually goes to zero; see also Rhode and Stegeman (1995). As Vega-Redondo (1997) argues, this is not because of the specifics of a particular example or of the narrow imitation process assumed. However, it is true with great generality in symmetric Cournot oligopoly games and holds for a more general class of imitation dynamics. The bottom line of this literature is the following qualitative insight. If players are more likely to imitate recently more successful behaviour, output levels tend to move to the Walrasian competitive equilibrium.

2. Experimental Design

We conducted seven experimental sessions in the computerised experimental laboratory LeeX at the Universitat Pompeu Fabra in Barcelona in Winter/Spring 1997. The subjects were undergraduate students in Humanities, Commercial Studies, Business Administration, and Economics, in approximately equal numbers.³ The experiment was based on the classic Cournot duopoly and triopoly experiments designed by Fouraker and Siegel (1963), and subsequently modified; see Holt (1985). For both the duopolies and triopolies we used three treatments based on different information setups.⁴ For each treatment there were 18 players simultaneously in the laboratory. Players sat in front of personal computers, and could not observe the screens of other players nor communicate with them. As will be explained below, the time that a session lasted depended on the treatment. It ranged from about one to about two hours. The average payment over the 126 players was 2,200 pesetas (\approx US\$ 15 at that time). Examples of the instructions given to the players can be found in Appendix B (see www.res.org).

2.1. *The Standard Cournot Model*

We consider a standard symmetric Cournot oligopoly. There is a number n of firms producing the same homogeneous commodity. The only decision variable for firm i is the quantity q_i to be produced. We assume that the inverse demand function is $P(Q) = a + bQ$, where $Q = \sum q_i$, $a > 0$, and $b < 0$, with the exact values of a and b differing from treatment to treatment (see below). The n players are firms competing in the same market for 22 consecutive periods. The total costs for a firm are given by $C(q) = kq + K$, with $k > 0$, and $K < 0$, ensuring positive profits at Walrasian output levels. Given the market price P , the profit V for an individual player is computed as follows: $V = Pq - C$.

³ Only students in Economics could have had some notions of game theory, but most of them did not, since only a few were third or fourth year students.

⁴ In addition to these six treatments we ran one control treatment, in which we rewarded the players for their relative performance (see below).

2.2. *Treatments: Duopolies and Triopolies*

Since there were always 18 players simultaneously in the laboratory, 9 duopoly or 6 triopoly markets were going on at the same time. The players were matched randomly and anonymously to form markets. They played for consecutive 22 periods in the same market. In our experiment, unlike in Fouraker and Siegel (1963), subjects knew the length of the experimental sessions. Players also knew whether it was a duopoly or triopoly they were in, but they did not know who was in their market. In the duopolies, the inverse demand curve was $P = 414 - 4Q$, and the cost function $C = 174q - 146$. In the triopoly these were: $P = 530 - 4Q$, and $C = 174q - 266$.

Three symmetric pure strategy profiles stand out in the static Cournot game: the joint-monopoly Pareto (P), the Cournot-Nash (N), and the competitive Walrasian output (W). Given the specifications of the demand and cost functions, these three output levels in the duopoly are $Q^P = 30$ (with $q^P = 15$), $Q^N = 40$ (with $q^N = 20$), and $Q^W = 60$ (with $q^W = 30$). In the triopoly they are $Q^P = 45$ (with $q^P = 15$), $Q^N = 66$ (with $q^N = 22$), and $Q^W = 90$ (with $q^W = 30$).

Players can choose a quantity from 8 to 32 and are allowed to enter only integer values for their output levels.⁵ Notice that the Walrasian individual output level of 30 is near the upper bound of the output range. Had we allowed for a higher range of output, random output choice or choice in the middle of the output range might have resulted in market output levels close to Walras, confounding the effect of imitation of successful firms.⁶

2.3. *Treatments: 'Bounded Rationality'*

We consider three different treatments that differ in the way the information is provided, and the time pressure put on the players. Yet, the underlying market is exactly the same in the three treatments, and in fact, the objectively 'available' information is essentially the same too in each treatment. Consequently, the theoretical benchmarks computed above are appropriate for all treatments. If some important pieces of information were objectively missing in one treatment, even fully rational agents might choose different output levels.

We call these three treatments 'easy', 'hard', and 'hardest' since, as we move from 'easy' to 'hardest', subjects have less time to make their decisions and the

⁵ Appendix A (on www.res.org) presents a formal analysis of the Cournot oligopoly. Due to the integer restriction, there are some asymmetric strategy profiles that are closely related as well. In the duopoly, one firm producing 19, and the other 21 would be an asymmetric Cournot-Nash equilibrium. In the triopoly $Q^P = 45$ gives the only symmetric Pareto output, but $Q = 44$ would give the same total profits. Also the Cournot-Nash equilibrium leading to $Q^N = 66$ is not strict since one player could deviate to $q^N = 23$ and be equally well-off. In fact, $Q^N = 67$ is an asymmetric Cournot-Nash equilibrium. The Walrasian equilibrium occurs at $Q^W = 89$. Hence, when all firms produce 29, two want to deviate to 30, and when all produce 30, one wants to deviate to 29. Nevertheless, $q^W = 30$ stands out in the sense that it is the only symmetric output where no player would realise a higher profit than the other two by deviating unilaterally.

⁶ One could likewise argue against placing the Walrasian output close to the end of the output interval, on the basis that players might be reluctant to make choices that seem extreme to them. But, as we will see in Section 2.4, the experimental evidence shows that the players in our Cournot markets are not reluctant to go towards the boundary if they think it appropriate to do so.

information is provided in a more confusing format. In the 'easy' treatment there is no time pressure on the players. They can spend as much time as they want. In the 'hard' and 'hardest' treatments, players have just one minute to decide on their output level. The differences in information format of the three treatments are more involved. In neither of them did the players get the demand and cost functions as such.

In the 'easy' version, the players were given a profit table that *conveniently* summarised all the information concerning the inverse demand curve and the cost function (see Appendix B). This profit table contains for each combination of outputs the profits for the firms. The column entry shows the output of firm X , and the row entry the output of the other firm (duopoly) or the average output level of the other firms (triopoly). The cells show the profits of firm X and the profits of the other firm (duopoly) or the average of the profits of the other firms (triopoly). In this treatment, after each period, each player gets information about the actions of each of the other players in the same market, but not about their profits. Notice, however, that if they wanted, they could look up the profits of the other players in the profit table. In addition, in the 'easy' version a player always gets a complete history of his own past actions and profits (see Figure B1 in Appendix B on www.res.org).

In the 'hard' version, the players did not get the convenient profit table. Instead, they got an *inconveniently* arranged enumeration of the market prices associated with all possible aggregate output levels. They got a similarly arranged enumeration of all possible cost levels. Since the players knew that their profits were simply their revenues minus their costs and that all firms were identical, they had exactly the same information as in the 'easy' version, although less discernible. In this treatment, after each period, both the actions of all players in the same market in that period, and the *profits* obtained by each player were given on the screen (see Figure B2 in Appendix B). In addition, to make sure that subjects did not miss what output decision had led to the highest profit, it was explicitly stated in the instructions that this output level would appear on the screen, as it does, with a *border of **s. Finally, no history of past individual decisions was provided in this and the next treatment, to have subjects concentrate on the decisions taken in the previous period.

The 'hardest' version differed from the 'hard' treatment in that the information about the demand side of the market was limited to the statement that '*the price level depends on aggregate output*'.⁷

The differences between the 'easy', and the 'hard' and 'hardest' versions can be summarised as follows. In the 'easy' version, by providing subjects with a convenient profit table, we minimise the need for further learning about the environment, thus facilitating best-replies. As we move into the 'hard' and 'hardest' versions, we scramble the information about the market, increasing the subjects' need to learn about the environment. This, combined with the one-minute time pressure, becomes an obstacle to 'doing the best one can'. In addition, we provide

⁷ Notice that the objectively available information is essentially still identical. Since the underlying demand function is deterministic and linear, it can be reconstructed from the observed quantities and profits, given the cost function.

flashing information about the most successful decision in each period, facilitating imitation of it.

Conventional wisdom claims that imitation should be more prevalent the more complex the subjects' decision task is. Therefore, we will use the 'easy' treatment as a benchmark. The question, then, is to what extent is imitation more frequent in more difficult environments, i.e., when time and information constraints are tighter. The purpose of the second and third treatments is to explore this question.

2.4. *Monetary Incentives*

In all treatments, each player gets a 'show-up' fee of 250 pesetas (\approx US\$ 1.65 at that time). If a player realised losses, these would be subtracted from the 'show-up' fee. This was known to the players. No player realised cumulative losses. In addition, players were paid depending on their performance, the details of the payoff scheme applied being known by the players. In the duopolies, the monetary payoff was 0.035 pesetas per profit point realised during periods 1 to 20, and 0.35 pesetas in the last two periods 21 and 22. In the triopolies the coefficients were 0.025 and 0.25. These coefficients varied between duopoly and triopoly in order to compensate for the fact that profit opportunities are different in these markets. The result was identical monetary incentives across the different information treatments.

The decision to make the payoffs of periods 21 and 22 ten times higher reflects our interest, as will be explained in Section 3, in these last two periods. The higher payoffs were incentives for the players to stay concentrated till the very end. In addition, they were meant to reinforce the so-called 'end effect', making collusive and strategic behaviour more unlikely during the last periods,⁸ thereby encouraging higher output levels. In these circumstances, levels of output below Walras give *additional* credibility to a conclusion of little imitation.

In addition to the treatments mentioned above, we also ran a session with an 'easy' duopoly treatment in which players were rewarded by their *relative*, instead of their absolute, performance. After each period, only the player who made the largest profit received a fixed positive payoff (100 pesetas during each of the periods 1 to 20, and 1,000 pesetas in periods 21 and 22), while the other player(s) got nothing. When two players had the same profit, they shared the monetary reward of the winner. The main purpose of this treatment was to verify that the 'easy' treatment was very easy indeed, and that our experimental set-up did not contain any insurmountable obstacle to the attainment of the Walrasian output expected in such a treatment, given the induced preference to beat their opponents. Indeed, a Walrasian output level of 30 was chosen by 100% of the players in the last two periods in this control treatment.

⁸ Notice that since the game is a finitely repeated one with complete information, the unique subgame perfect equilibrium is playing the equilibrium of the stage game anyway.

3. Analysis

As explained in Section 2, the ‘*hard*’ and ‘*hardest*’ treatments should be particularly conducive to imitation. The intuition being that when the subjects are more intensely confronted with their bounded rationality, imitation of the successful firm becomes a more prevalent mode of behaviour. We therefore formulate the hypothesis to be tested in the following terms:

HYPOTHESIS. *In the ‘hard’ and ‘hardest’ treatments, imitation of successful behaviour will be more prevalent than in the ‘easy’ treatments.*

Our analysis of the experimental data proceeds in two stages. In Sections 3.1 and 3.2, based on the literature briefly outlined in Section 1, we take the point of view that imitation can be said to be ‘*more prevalent*’ only if it has a significant effect on output levels (approaching Walras). Now, output levels around the Walrasian equilibrium, may or may not be due to imitation. However, if output levels do not approach the competitive output and there is enough noise or experimenting, then we can conclude that imitation of the most profitable decision is *not prevalent* enough to affect significantly the output levels in the Cournot market. Taking the output levels as a first proxy for the occurrence of imitation of successful behaviour allows us to keep the definition of imitation of successful behaviour relatively open. Then, in Section 3.3, we analyse and explain individual output decisions as such. To do this, we use a more precise definition of the behavioural rule ‘*imitation of successful behaviour*’.⁹

These two stages of the analysis, which should be seen as complementary, both point to a rejection of the hypothesis. It appears that the bounded rationality factor introduced in the ‘*hard*’ and ‘*hardest*’ treatments did not induce players to imitate successful behaviour more often. We now explain in detail how we derive this conclusion.

3.1. *The Last Two Periods*

To start with, we focus on the last two periods, the standard approach when it comes to testing an equilibrium hypothesis; see Crawford (1998). Table 1 presents the *average* output level per player in the last two periods for each market in the various treatments. These output levels are the independent observations (9 for each duopoly treatment, and 6 for each triopoly) for all our statistical tests in this Section. For presentational reasons, we also present graphs of the *individual* output levels in the last two periods. Obviously these 36 individual observations neglect the possible interdependence between the individual output levels in a given market, and they will therefore not be used in our tests.

⁹ An alternative way of detecting imitation might be to ask the players. This should be done with great care since players need not be aware of fact that they imitate each other, even if they do. For example, a study based on a Fourier analysis of the voices in the Larry King Live talk-show revealed interesting convergence patterns in the frequencies used by the host and his interviewed guests. However, it seems unlikely that they were doing this consciously (Mirsky, 1996).

Table 1
Average Output per Player in the Last Two Periods for Each Market

Market	Duopolies			Triopolies		
	'Easy'	'Hard'	'Hardest'	'Easy'	'Hard'	'Hardest'
1	20.00	15.75	25.00	24.50	26.00	28.67
2	19.75	22.00	16.75	21.83	21.67	23.83
3	18.00	26.25	22.50	25.50	24.83	29.33
4	16.75	23.00	20.25	23.33	29.83	20.83
5	20.00	26.00	20.00	24.17	18.83	27.83
6	18.50	31.00	22.50	22.83	24.50	27.67
7	15.75	16.50	25.75			
8	15.00	27.75	23.75			
9	20.00	22.00	24.75			
average	18.19	23.36	22.36	23.69	24.28	26.36

Figure 1 shows the frequency distribution of actions in the last two periods for the 'easy' duopolies. Notice that a Pareto output of 15 is the most frequently chosen (14 out of 36 times), followed by a Cournot–Nash output (13 times). In period 21 there were ten colluding players, and four of them persisted in colluding in the last period. Since no output larger than 22 was chosen, no subject got close to the Walrasian output of 30. A sign test shows that the 95% confidence interval for the median output ranges from 16 to 19. These results are in agreement with

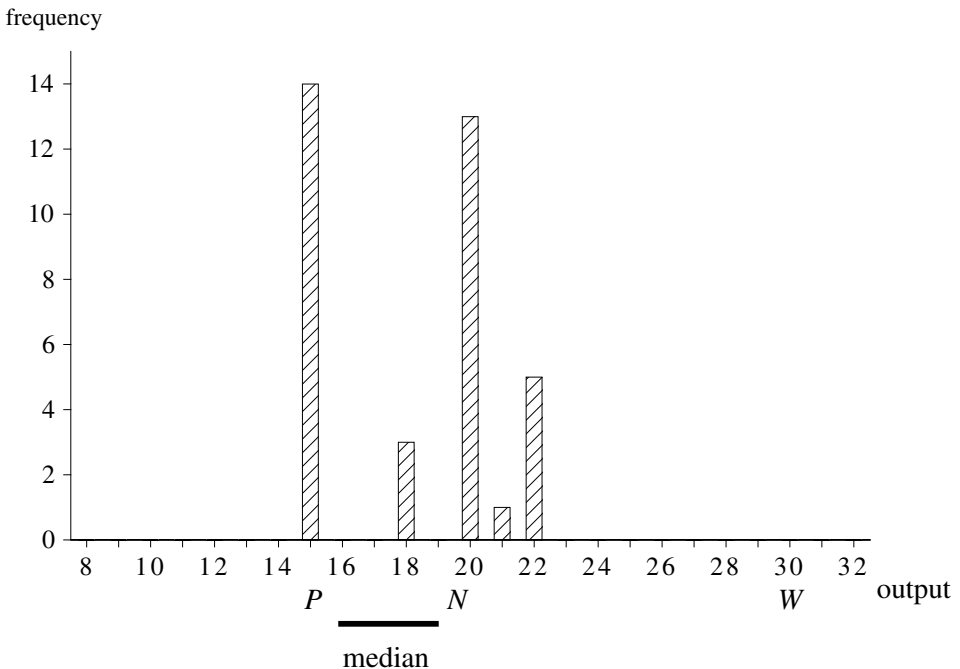


Fig. 1. Frequency Distribution of Individual Output Levels in the 'easy' Duopoly, last Two Periods

Fouraker and Siegel (1963), and other previous experiments. But whereas most of those experiments focused on possible explanations for the persistence or unravelling of collusion (see Holt (1995) for a survey), we will merely use the 'easy' markets as the benchmark treatment.

The data do not suggest a preference for beating the opponents. Assuming that preferences are constant across the treatments, any differences in the output levels in the other treatments can be attributed to a bounded rationality effect. Will the Walrasian output become more prevalent as the task of learning about the market becomes more and more difficult, at the same time that the decision of the most successful firm is displayed more prominently?

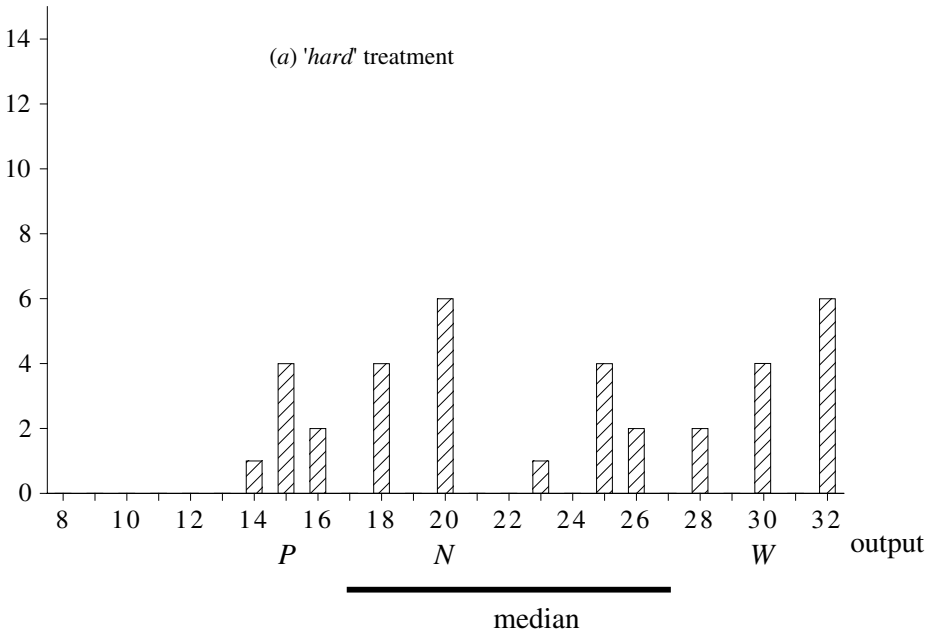
Figures 2(a) and 2(b) give the frequency distributions of the output levels in the last two periods of the 'hard' and 'hardest' duopoly treatments. In the 'hard' case, output is spread along the whole range, and has two peaks corresponding to Cournot and the highest possible output, 32, above Walras. The decrease of collusion with respect to the 'easy' version may be the result of the added difficulty in learning about the environment, making it more difficult to discover the Pareto output, and making it less useful to punish the other player, since punishment makes sense only if one believes the other understands what is expected from him when punished. The spreading out is equally notable in the 'hardest' case, but while the average output increased from 'easy' to 'hard', it slightly fell from 'hard' to 'hardest'. As can be seen from the graphs, this is not due to a fall in the relative frequency of outputs around the Walrasian equilibrium, but to a fall in the frequency of outputs in between Cournot–Nash and Walras, and to a frequency increase for outputs between Pareto and Cournot–Nash.

Do these output levels approach the Walrasian competitive equilibrium? First, a Wilcoxon–Mann–Whitney test (Wilcoxon test from here on) shows that the players produce significantly more in the 'hard' version than in the 'easy' version (at 2.5% significance level; 1-sided). Second, using the same test we find that firms in the 'hard' version produce significantly less than the Walrasian output level (significant at 0.5%, 1-sided).¹⁰ Third, we apply the sign test to determine the 95% confidence interval for the median output level. Thanks to the spread in output levels this ranges from 17 to 27, still away from the Walrasian equilibrium, but including the Cournot–Nash equilibrium of 20.

Applying the same tests to the 'hardest' duopolies, we find that they produce significantly more than in the 'easy' duopoly (0.5%, 1-sided Wilcoxon), but less than the Walrasian output of 30 (0.5%, 1-sided Wilcoxon), while there is no significant difference with the 'hard' duopolies. The 95% confidence interval for the median output (20 to 25, sign test) is even farther away from the Walrasian output than in the 'hard' version, including again the Cournot–Nash equilibrium. Hence, the hypothesis that in the 'hard' and 'hardest' treatments the Walrasian equilibrium would be a good description of the market must be rejected for the duopolies. The inference is that imitation was not prevalent.

¹⁰ We effectively compare the distribution of output levels with the distribution actually generated in our benchmark treatment in which the players were rewarded for their relative performance (see Section 2).

frequency



frequency

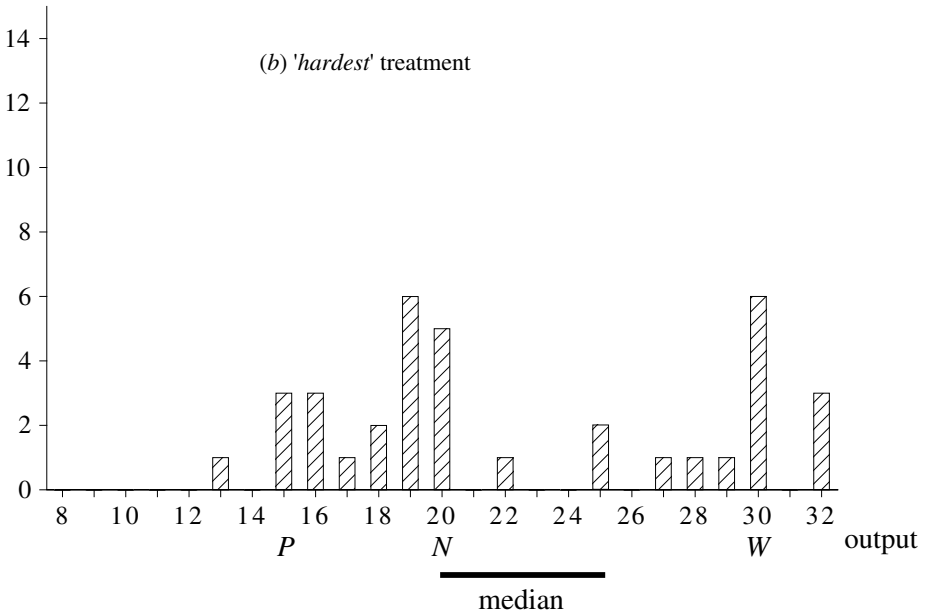


Fig. 2. Frequency Distributions of Individual Output Levels in the 'hard' and 'hardest' Duopoly, Last Two Periods

Figure 3 presents the frequency distribution of the individual output levels for the triopolies in the 'easy' treatment, the baseline treatment. We observe that there is a wider spread of output levels than in the baseline duopoly session, and that practically no collusion occurs. The latter is related, according to the usual explanation, to the increased difficulty, with three rather than two subjects, of learning what other players are up to and rewarding and punishing individual players. Since the only available instrument works through the market and does not discriminate between players. The most frequent output levels in the last two periods are 23 (10 times), 22 (6 times), and 21 (4 times). The 95% confidence interval for the median output goes from 22 to 25 (sign test), including the Cournot–Nash equilibrium at 22. This, again, confirms previous experimental observations. We want again to verify whether imitation comes to the fore as we make the players' learning-about-the-environment task more complicated.

The frequency distributions of output levels for the 'hard' and 'hardest' triopoly are given in Figures 4(a) and 4(b). One observation again stands out, which is the wider spreading of choices compared to the 'easy' version. In the 'hard' version, the frequency of the Walrasian output actually decreases, while Cournot–Nash loses in frequency in favour of lower values. The highest frequency in periods 21 and 22 corresponds to outputs of 18, 20, and 25 (4 times each), and the average output in the last two periods is 24.3. The spread is not surprising due to the added difficulty in learning about the environment. It is also an indication that

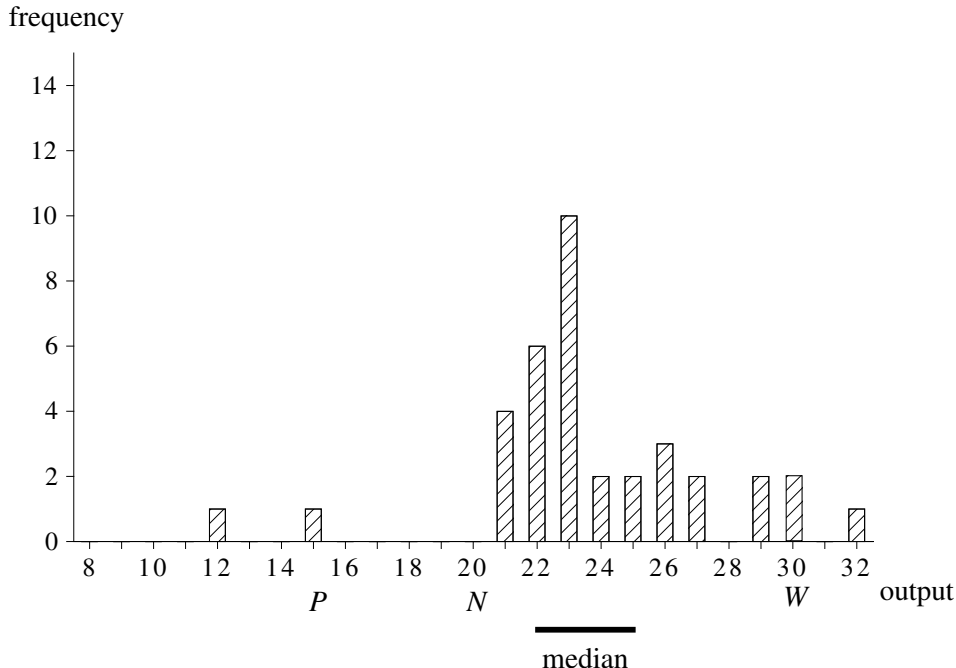
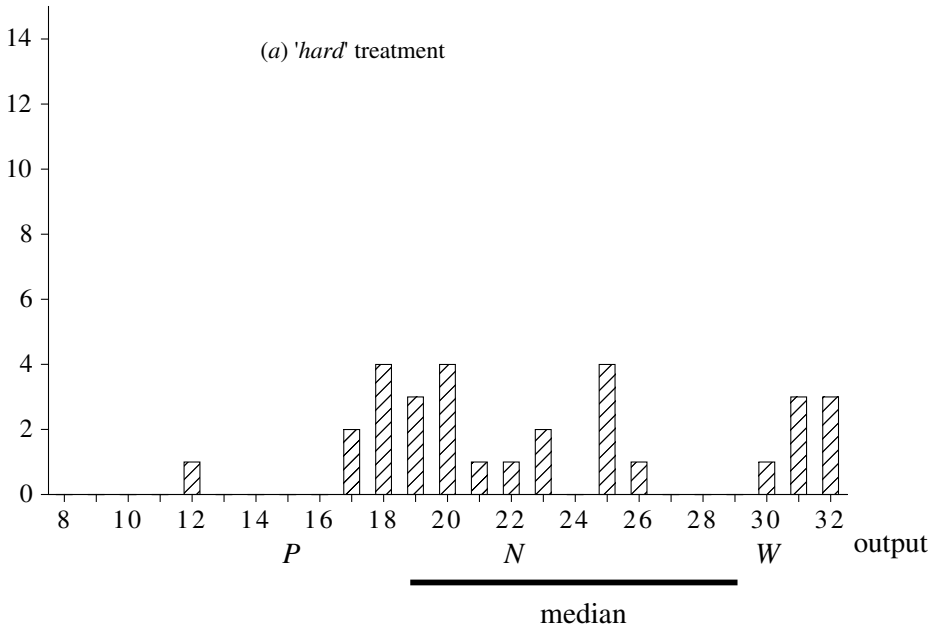


Fig. 3. Frequency Distribution of Individual Output Levels in the 'easy' Triopoly, Last Two Periods

frequency



frequency

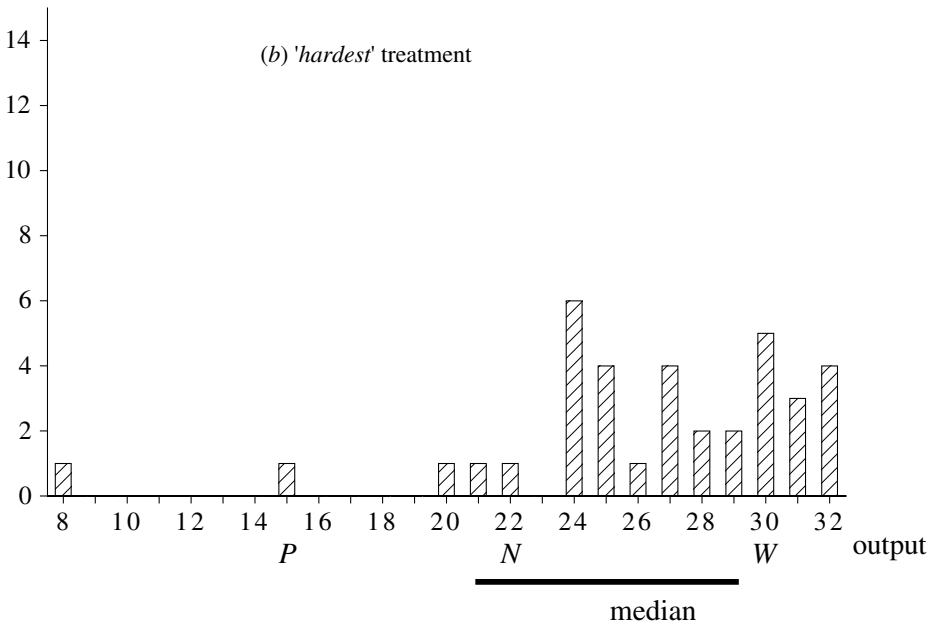


Fig. 4. Frequency Distributions of Individual Output Levels in the 'hard' and 'hardest' Tripoly, Last Two Periods

imitation of the successful firms was not frequent. Let us apply the same statistical tests as before. First, a Wilcoxon test cannot reject the null hypothesis that the output levels are from the same distribution against the alternative hypothesis that they are higher in the 'hard' triopoly than in the corresponding 'easy' treatment at any conventional significance level. The same test confirms that they are lower than the Walrasian output levels (significant at 0.5%, 1-sided). Second, a sign test provides us with a 95% confidence interval for the median output level ranging from 19 to 29, including the Cournot–Nash output of 22 but not the Walrasian output of 30.

In the 'hardest' triopolies, the modal output in the last two periods is 24 (6 times), followed by 30 (5 times), with an average of 26.4, and the spread is again considerable. We find that the output in the 'hardest' triopolies is *not* significantly higher than in the 'easy' triopolies, and there is also no significant difference with the 'hard' triopolies. Moreover, output in the 'hardest' triopolies is significantly lower than the Walrasian level of 30 (0.5%; 1-sided Wilcoxon). The sign test gives a 95% confidence interval for the median output from 21 to 29, including only the Cournot–Nash equilibrium.

Hence, we have to conclude that in the triopoly sessions, a case in favour of the hypothesis seems rather weak at best.

Combining the evidence for the duopolies and triopolies, on the basis of our analysis of the output levels in the last two periods, two conclusions can be drawn.

FACT 1. As the learning-about-the-environment task becomes more complex, output choices become more spread out.

The equation between environmental complexity and uncertainty of results seems to be confirmed in all cases. This seems an interesting finding. It confirms that the 'hard' and 'hardest' treatments differ in a significant way from the 'easy' one, that bounded rationality does matter. It also discredits the view that even the 'hardest' Cournot market is simply too easy a game to test for imitation. We will get back to this important difference in behaviour across the treatments in Section 3.3.

*FACT 2. As the learning-about-the-environment task becomes more complex, average output increases, but the Walrasian output is not a good description of the output levels observed in the experiment.*¹¹

3.2. The Trend

Since our interpretation of the hypothesis to be tested in this stage of the analysis concerns an equilibrium prediction, thus far we focused on the output levels in the last two periods. Nevertheless, data from previous periods can help decide whether there is anything in the dynamics which suggests that running the experiment *longer* would lead to output levels closer to the Walrasian output. As we will see, there is no such evidence.

¹¹ Notice in particular that the 95% confidence intervals of the median output levels never include the Walrasian output, whereas they always include Cournot–Nash.

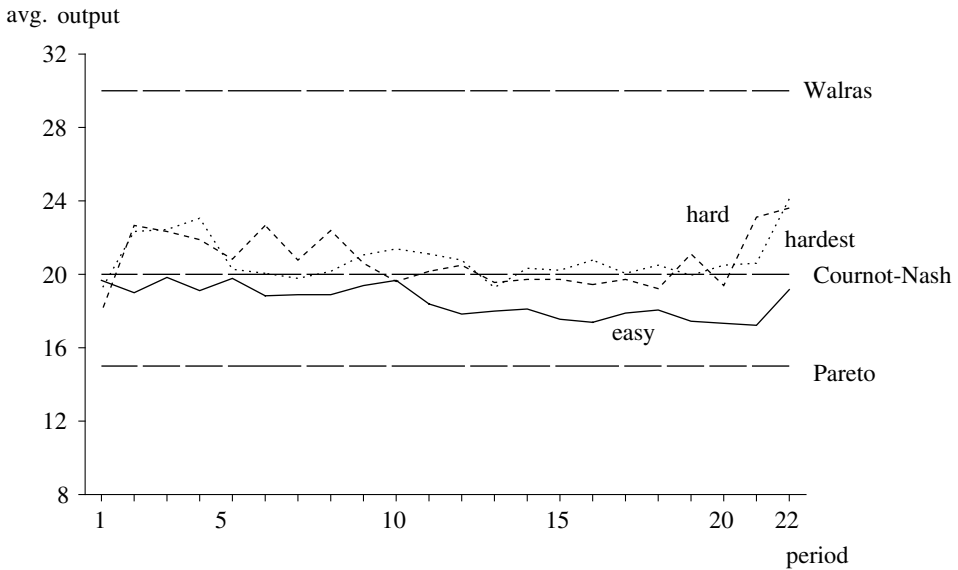


Fig. 5. Time Series of the Average Output per Player in the Duopolies

Figure 5 gives the time series of the average outputs in the duopolies. In the 'easy' duopolies, output starts just below 20, and slowly decreases as collusion builds up. With the collapse of collusion in the last two periods, output shoots up towards Cournot–Nash. In the 'hard' duopolies, output starts slightly below 20 as well, increases to 22.7 in period 2, and then very slowly goes down to 19.4 in period 20, followed by an end effect leading to 23.6 in period 22. Hence, it stays close to Cournot–Nash, and if a trend exists during the first 20 periods, it is downward sloping. We see a very similar pattern in the 'hardest' duopoly treatment. Average output starts at 19.3, reaches a peak in period 4 at 23.1, followed by fluctuations around Cournot–Nash until the last period, where we see a jump to 24.1. The conclusion is that in the 'hard' and 'hardest' duopolies nothing suggests a trend towards Walrasian output levels.

Figure 6 shows the average output levels for the triopolies. In the 'easy' triopolies, output starts slightly below 20, increases a little bit until period 5, and then stays relatively constant till the end. There is no end effect, as collusion was never established. In the 'hard' triopolies, output starts at 16.3, increases to 24.9 in period 4 and then slowly decreases to 21.8 in period 16, followed by a slight increase to 22.5 in period 20 (remember Cournot–Nash is at 22) and 24.6 in period 22. Notice that the average output level from period 11 to 20 is always lower than in the 'easy' baseline treatment. The 'hardest' triopolies are somewhat but not excessively different. The average output level in period 1 is 20.2, followed by a quick jump to 24.5 in the second period, and reaches a first peak at 26.8 in period 8. After this, the average output level fluctuates somewhat, with a dip of 23.8 in period 18 and a value in the final period of 25.9. As we see, the series is consistently above the 'easy'

avg. output

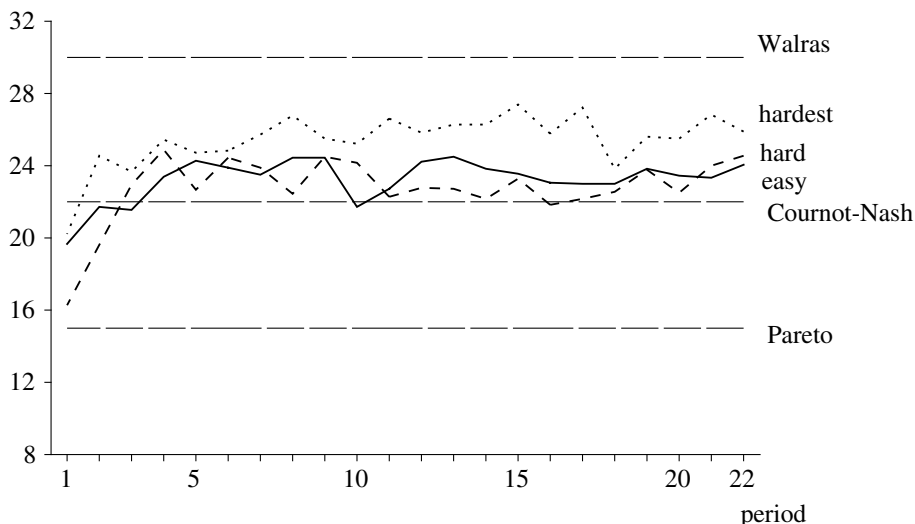


Fig. 6. *Time Series of the Average Output per Player in the Triopolies*

and ‘*hard*’ series but, apart from the increase in the first periods, there is no upward trend.

In conclusion, the analysis of all the 22 periods of the sessions does not reveal any trend towards the competitive equilibrium in any of the six treatments. Nothing in the observed decisions seems to indicate that the sessions were too short for convergence to the Walrasian equilibrium. Of course, on the basis of our experimental evidence reported here we cannot exclude that such an upward trend itself might start later. But our hypothesis concerned imitation based on bounded rationality; the longer the sessions go on, the less bounded rationality is an issue, because players learn more and more about the environment. Hence, even if a trend were to start later on, it would be hard to explain it in terms of bounded rationality leading to imitation.

However, apart from this logical point, we actually did gather some evidence as to whether an upward trend might start later. In the ‘*hard*’ and ‘*hardest*’ duopolies and triopolies, once the 22 periods of each session were finished, we informed subjects that since the session had run so smoothly and fast, we still had some time left, and that we would repeat the same experiment with the same matches for another 12 periods, the last two periods having again payoffs ten times the payoffs of the previous periods.¹² In these 12 periods again, there is no sign of a trend towards Walras.¹² Therefore we can state the following.

¹² We do not fully report these data since the procedure may be deemed controversial and the results spurious for the strict purpose of the experiment. Yet they throw some light on whether a trend towards Walrasian results was lurking in our sessions of 22 periods. The corresponding graphs are available from the authors upon request.

FACT 3. *There is no trend towards the Walrasian equilibrium.*

Two points concerning the relation between Sections 3.1 and 3.2 merit a brief discussion. First, one might argue that what happens in the periods 1 to 20 is in some sense ‘*cheap play*’, preparing the ground for the last two periods, which offer much higher payoffs. There is no denying that this could be the case. However, if it were so, then we would be dealing with some kind of strategic behaviour, which implies a reasoning process at a higher cognitive level than the learning-through-imitation hypothesis assumes. To the extent that some of this may be happening in the duopoly markets, it reinforces the conclusion that imitation is not a driving force in our experiments.¹³ That end effects are hardly noticeable in the triopoly markets indicates that in these markets the first twenty periods are not mere foreplay. The absence of any significant advance towards Walras should be taken at face value.

Second, in Section 3.1 we focused on the last two periods, whereas in this Section, where we search for a trend, we seem to discard the output levels in these last two periods as mere ‘*end effects*’. While this might seem inconsistent at first glance, what we showed in Section 3.1 is that *even with an end effect* pushing the average output levels up, we do not see the output levels approach Walras in any significant sense. Had we focused on the outputs of all periods, we would have been even farther away from the Walrasian output levels.

3.3. *Analysis of Individual Decisions*

The research question underlying the first stage of the analysis was whether raising the complexity of the environment would lead to such an increase in imitation that the output levels would approach Walras. We concluded that the output levels in the two treatments most conducive to imitation did not get close to Walras in any significant sense. Nevertheless, we observe differences in output patterns between the ‘*easy*’ treatment and the other two. Consequently, in this Section we analyse how these output patterns are related to the individual period-to-period decisions and, in particular, what the role is of imitation and other related behavioural rules. Specifically, we pay attention to the set of behavioural rules presented in Table 2. Each of these rules determines a target output level, possibly more than one, on the basis of previous actions and outcomes for each period after the first.¹⁴ These rules fall broadly into three classes. Rules 1 to 4 belong to the class of so-called belief-based rules. They all explicitly choose a best-response to some belief. The rules from 5 to 8 could be considered some form of imitation. Rules 9 to 16 are versions of reinforcement learning and rule 17 is closely related to this, in the sense that it is also directly and exclusively based on observed profits. For a detailed description of the rules we refer to Appendix C (www.res.org).

We now describe how accurately each behavioural rule predicts the players’ actions in each of the treatments. We distinguish two forms of predictive success. First, a rule predicts correctly if a player chooses exactly the target output level.

¹³ The same argument applies to any other form of dynamically strategic behaviour.

¹⁴ The only exception, as we will explain below, is hill climbing.

Table 2
Set of Individual Behavioural Rules

1	<i>BESTREP</i>	myopic best-reply
2	<i>ADPTEXP</i>	adaptive expectations
3	<i>FICT</i>	fictitious play
4	<i>GEOFICT</i>	geometric fictitious play
5	<i>IMITATE</i>	imitate-the-best
6	<i>IMAVG</i>	imitate-the-average
7	<i>EXMPL</i>	imitate-the-exemplary
8	<i>STAY</i>	stay put (imitate oneself)
9	<i>AVGRL</i>	average reinforcement learning
10	<i>AVGRLD</i>	average reinforcement learning with discounting
11	<i>CUMRL</i>	cumulative reinforcement learning
12	<i>CUMRLD</i>	cumulative reinforcement learning with discounting
13	<i>AVGXRL</i>	extended average reinforcement learning
14	<i>AVGXRLD</i>	extended average reinforcement learning with discounting
15	<i>CUMXRL</i>	extended cumulative reinforcement learning
16	<i>CUMXRLD</i>	extended cumulative reinforcement learning with discounting
17	<i>HILL</i>	hill climbing

Second, we also count as correct predictions those cases in which a player moves into the direction of the target without overshooting the target. The quantitative measure of predictive success we use is Selten's (1991) '*Difference measure m*', which takes into account both the relative frequency of correct predictions, and the relative size of the set of predicted outcomes. It is defined as follows:

$$m = r - a, \quad (1)$$

where $r = \text{'hit rate'}$ (relative frequency of correct predictions) and $a = \text{'area'}$ (relative size of the set of predicted outcomes, i.e., the number of predicted outcomes divided by the number of possible outcomes).¹⁵

In Table 3(a) we present the average '*Difference measure m*' for each treatment in the duopoly and triopoly settings, focussing on players hitting the targets exactly.¹⁶ To test for differences in the levels of predictive success for given rules across treatments as presented in Table 3(a), we apply a 2-sided Wilcoxon test, using the Difference measures m for each rule at the market level as independent observations.¹⁷ Applying conventional significance levels, Table 3(b) indicates which differences are significant. For example, the difference between the '*easy*' and the '*hard*' duopoly for the adaptive-expectations rule is given as '+ (2%)', meaning that this rule performs better in the '*easy*' treatment at a 2% significance level. A blank indicates no significant difference.

¹⁵ Notice that a high 'hit rate' means high accuracy, whereas a small 'area' means high precision. Observe that $-1 < m < 1$, with $m = 0$ indicating a rule which predicts just as well as uniform random behaviour over output levels (as the expected hit rate r equals the area a when play was in fact uniform random). In other words, m will be positive only if actual behaviour is better described by a given rule than by uniform random behaviour. We refer to Selten (1991) for an analysis of the theoretical properties of this measure.

¹⁶ This and subsequent tables have also been produced for players moving in the direction of these targets. As they do not appear substantially different, they are not reproduced here. They are available from the authors upon request.

¹⁷ Hence, we have 9 observations for each duopoly treatment, and 6 for each triopoly.

Table 3(a)
Difference Measure m' for Exact Hits (average per treatment)

		Duopolies			Triopolies		
		'Easy'	'Hard'	'Hardest'	'Easy'	'Hard'	'Hardest'
1	BESTREP	0.14	0.03	0.03	0.13	0.01	-0.01
2	ADPTEXP	0.11	0.01	0.04	0.07	0.00	0.00
3	FICT	0.10	-0.01	0.03	0.13	-0.01	0.00
4	GEOFICT	0.12	0.00	0.03	0.08	0.01	-0.02
5	IMITATE	0.47	0.18	0.24	0.13	0.16	0.18
6	IMAVG	0.44	0.17	0.21	0.14	0.09	0.05
7	EXMPL	0.47	0.19	0.29	0.09	0.13	0.10
8	STAY	0.50	0.22	0.31	0.20	0.23	0.23
9	AVGRL	0.19	0.09	0.14	0.08	0.08	0.07
10	AVGRLD	0.24	0.08	0.15	0.08	0.07	0.07
11	CUMRL	0.44	0.18	0.21	0.16	0.13	0.20
12	CUMRLD	0.48	0.21	0.27	0.22	0.17	0.23
13	AVGXRL	0.09	0.03	0.04	0.01	0.05	0.02
14	AVGXRLD	0.14	0.02	0.05	0.02	0.02	0.03
15	CUMXRL	0.42	0.17	0.21	0.14	0.13	0.14
16	CUMXRLD	0.45	0.18	0.25	0.17	0.16	0.20
17	HILL	0.17	0.05	0.09	0.09	0.06	0.06

Table 3(b)
Significant Differences in 'Difference Measure m' Between Treatments
 (significance levels in parenthesis)

		Duopolies			Triopolies		
		'Easy'- 'Hard'	'Easy'- 'Hardest'	'Hard'- 'Hardest'	'Easy'- 'Hard'	'Easy'- 'Hardest'	'Hard'- 'Hardest'
1	BESTREP				+ (5%)	+ (1%)	
2	ADPTEXP	+ (2%)		- (10%)	+ (10%)		
3	FICT				+ (2%)	+ (5%)	
4	GEOFICT	+ (2%)			+ (10%)	+ (2%)	+ (10%)
5	IMITATE	+ (1%)	+ (5%)				
6	IMAVG	+ (5%)	+ (10%)			+ (10%)	
7	EXMPL	+ (2%)					
8	STAY	+ (1%)	+ (5%)				
9	AVGRL						
10	AVGRLD	+ (10%)					
11	CUMRL	+ (1%)	+ (1%)				
12	CUMRLD	+ (1%)	+ (1%)				
13	AVGXRL						
14	AVGXRLD	+ (5%)					
15	CUMXRL	+ (2%)	+ (5%)				
16	CUMXRLD	+ (1%)	+ (2%)				
17	HILL						

A number of observations can be made about Tables 3(a) and (b). First, if there is a significant difference then this is generally a decrease in performance from 'easy' to 'hard' or from 'easy' to 'hardest'. Second, both in the duopolies and the triopolies, the belief-based rules, rules 1 to 4, largely disappear in the 'hard' and 'hardest' treatments. Third, all imitation type of rules and the cumulative

reinforcement rules have high scores in the 'easy' duopolies, whereas these rules perform worse in the other treatments. Fourth, in particular, there seems no evidence of an increase for the imitate-the-best rule in the 'hard' and 'hardest' treatments. Fifth, the most successful rule throughout is the stay-put rule, with the only exception being the 'easy' triopoly, where it is closely beaten by one other rule. Sixth, the cumulative reinforcement learning rules seem to do much better than those based on average payoffs, and the extended reinforcement learning rules do systematically worse than the standard reinforcement learning rules.

FACT 4. *As the environment becomes more complex, simple rules of behaviour have less predictive success.*

In Table 3(a) we showed, for presentational reasons, the data as averages per treatment, whereas the tests in Table 3(b) took place at the level of market averages. But of course it might be interesting to take an even closer look at the data, as there might be heterogeneity even below the market level, at the level of the individual players. Table 4 presents for each rule a classification of the individual players, distinguishing five different classes for the 'Difference measure m '. Table 4 confirms that the number of players using frequently one of the simple decision rules decreases with the complexity of the environment.¹⁸ On the basis of the previous observations we can state:

FACT 5. *In general there is no increase of imitation as the environment gets more difficult for the players.*

One of the things that is clear from the data is that there is an overlap between the targets of various rules. Therefore, we will estimate a model of individual behaviour taking into account these target output levels *simultaneously*. Whereas the analysis we presented above was centred on the players hitting their targets exactly, we now shift focus to the *direction* of the players' output decisions. In the model we consider next, the sign of a player's output change, Δq , is a function of the direction, x , indicated by the target output levels according to the various behavioural rules. In other words, the question considered is: If different behavioural rules guide a player to adjust her output into different directions, then to which extent does a player follow each of these rules?

Denoting the individual output level in period t as $q(t)$, distinguishing behavioural rules with the subscript i , and omitting subscripts for the individual players, we define the following variables:

$$\begin{aligned} \Delta q(t) = \text{'down'} & \quad \text{if} \quad q(t) < q(t-1) \\ \text{'same'} & \quad \text{if} \quad q(t) = q(t-1) \\ \text{'up'} & \quad \text{if} \quad q(t) > q(t-1) \end{aligned} \quad (2)$$

¹⁸ If we look at the imitate-the-best rule in the triopolies, we see the appearance of two serious imitators in the 'hardest' treatment (with $0.50 < m \leq 0.75$) and four players with $0.25 < m \leq 0.50$. As it turns out, two of these latter four plus the two most serious imitators were playing in the markets 1 and 3, in which average output levels in the last two periods were 28.67 and 29.33 (see Table 1), closer to Walrasian output than in any other market in the 'hardest' treatments.

Table 4
Frequency Distribution Individual Difference Measures m' (exact hits)

			Duopolies					Triopolies				
			$m \leq 0$	$0 < m \leq 0.25$	$0.25 < m \leq 0.50$	$0.50 < m \leq 0.75$	$0.75 < m \leq 1.00$	$m \leq 0$	$0 < m \leq 0.25$	$0.25 < m \leq 0.50$	$0.50 < m \leq 0.75$	$0.75 < m \leq 1.00$
1	<i>BESTREP</i>	'easy'	5	8	4	1	0	5	10	3	0	0
		'hard'	10	6	2	0	0	11	7	0	0	0
		'hardest'	10	7	1	0	0	11	7	0	0	0
2	<i>ADPTEXP</i>	'easy'	3	12	2	1	0	4	13	1	0	0
		'hard'	8	10	0	0	0	9	9	0	0	0
		'hardest'	5	12	1	0	0	10	8	0	0	0
3	<i>FICT</i>	'easy'	3	13	1	1	0	2	12	4	0	0
		'hard'	10	8	0	0	0	11	7	0	0	0
		'hardest'	7	11	0	0	0	13	5	0	0	0
4	<i>GEOFICT</i>	'easy'	2	13	2	1	0	4	13	1	0	0
		'hard'	10	8	0	0	0	7	11	0	0	0
		'hardest'	5	12	1	0	0	14	4	0	0	0
5	<i>IMITATE</i>	'easy'	0	4	5	7	2	2	14	2	0	0
		'hard'	1	10	7	0	0	2	13	3	0	0
		'hardest'	0	9	9	0	0	3	9	4	2	0
6	<i>IMAVG</i>	'easy'	0	5	3	8	2	0	15	3	0	0
		'hard'	4	8	5	1	0	3	14	1	0	0
		'hardest'	1	12	2	3	0	1	17	0	0	0
7	<i>EXMPL</i>	'easy'	0	6	3	7	2	3	13	2	0	0
		'hard'	1	9	8	0	0	2	14	1	1	0
		'hardest'	2	8	4	4	0	4	11	3	0	0
8	<i>STAY</i>	'easy'	0	2	7	7	2	0	12	6	0	0
		'hard'	0	12	6	0	0	1	12	4	1	0
		'hardest'	0	7	8	3	0	0	12	4	1	1
9	<i>AVGRL</i>	'easy'	0	12	5	1	0	5	11	2	0	0
		'hard'	5	11	2	0	0	2	15	0	1	0
		'hardest'	5	10	1	2	0	6	10	2	0	0
10	<i>AVGRLD</i>	'easy'	0	10	6	2	0	5	11	2	0	0
		'hard'	5	12	0	1	0	3	14	0	1	0
		'hardest'	4	10	2	2	0	6	10	2	0	0

Table 4
Continued

		Duopolies					Triopolies					
		$m \leq 0$	$0 < m \leq 0.25$	$0.25 < m \leq 0.50$	$0.50 < m \leq 0.75$	$0.75 < m \leq 1.00$	$m \leq 0$	$0 < m \leq 0.25$	$0.25 < m \leq 0.50$	$0.50 < m \leq 0.75$	$0.75 < m \leq 1.00$	
11	<i>CUMRL</i>	'easy'	0	4	6	6	2	1	14	3	0	0
		'hard'	0	14	3	1	0	0	16	2	0	0
		'hardest'	0	11	6	1	0	1	11	3	3	0
12	<i>CUMRLD</i>	'easy'	0	3	6	7	2	0	12	6	0	0
		'hard'	0	12	6	0	0	0	14	2	2	0
		'hardest'	1	10	6	1	0	1	9	6	1	1
13	<i>AVGXRL</i>	'easy'	4	12	1	1	0	9	9	0	0	0
		'hard'	5	13	0	0	0	4	13	1	0	0
		'hardest'	6	11	1	0	0	8	10	0	0	0
14	<i>AVGXRLD</i>	'easy'	2	12	4	0	0	7	11	0	0	0
		'hard'	7	11	0	0	0	4	14	0	0	0
		'hardest'	4	14	0	0	0	8	9	1	0	0
15	<i>CUMXRL</i>	'easy'	0	6	5	5	2	1	14	3	0	0
		'hard'	1	12	5	0	0	3	13	2	0	0
		'hardest'	1	11	5	1	0	1	13	3	1	0
16	<i>CUMXRLD</i>	'easy'	0	3	7	6	2	1	12	5	0	0
		'hard'	1	12	5	0	0	2	14	1	1	0
		'hardest'	0	10	6	2	0	1	11	5	1	0
17	<i>HILL</i>	'easy'	4	8	5	1	0	3	13	2	0	0
		'hard'	9	8	1	0	0	5	13	0	0	0
		'hardest'	7	10	1	0	0	5	12	1	0	0

and

$$x_i(t) = \begin{cases} -1 & \text{if } target_i(t) < q(t-1) \\ 0 & \text{if } target_i(t) = q(t-1) \\ 1 & \text{if } target_i(t) > q(t-1). \end{cases}$$

To analyse whether the behaviour of the individual players *differs* across the ‘easy’, ‘hard’ and ‘hardest’ treatments, we estimate ordered-probit models in the following way. We pool all data for the duopolies, and the same for all data for the triopolies. We define two dummy variables. One, d_i^{hh} , takes value 1 for the ‘hard’ and ‘hardest’ treatments, and another, d_i^h , takes value 1 for the ‘hardest’ treatment, while both dummies are 0 otherwise. The ordered-probit model, then, assumes a latent response variable, say z , that is a linear function of the independent variables x_i plus a normally distributed error term, $u(t)$:

$$z(t) = \sum_{i=1} \beta_i x_i(t) + \gamma_0 d_0^{hh} + \sum_{i=1} \gamma_i d_i^{hh} x_i(t) + \delta_0 d_0^h + \sum_{i=1} \delta_i d_i^h x_i(t) + u(t). \quad (3)$$

We estimate, by maximum-likelihood, the β , γ , and δ coefficients together with two cut points, α_1 and α_2 , which determine the probability that the estimated linear function, with random error, implies the three possible outcomes (‘down’, ‘same’, or ‘up’) as follows:

$$\begin{aligned} \text{Prob}[\Delta q(t) = \text{‘down’}] &= \text{Prob}[z(t) \leq \alpha_1] \\ \text{Prob}[\Delta q(t) = \text{‘same’}] &= \text{Prob}[\alpha_1 < z(t) \leq \alpha_2] \\ \text{Prob}[\Delta q(t) = \text{‘up’}] &= \text{Prob}[\alpha_2 < z(t)]. \end{aligned} \quad (4)$$

The target output levels we consider are determined as follows. Recall (see Table 3(a)) that for the class of belief-based rules, the best-response rule outperformed the other rules, while the cumulative-reinforcement-learning-with-discounting rule performed best among the reinforcement learning rules. Therefore, in the following analysis, we only include these two rules from these two classes. In addition, we include the imitate-the-best, the imitate-the-average, and the imitate-the-exemplary rules.¹⁹ Combined with the dummies, this implies that we have a total of 17 explanatory variables. Starting from the model that includes all variables, we use Likelihood-Ratio tests (applying a 5% significance level) to get a more parsimonious model. The estimated values of the simplified models for the duopolies and triopolies are presented in Table 5.²⁰ The estimated β s concern all treatments, the γ s show to which extent the ‘hard’ and ‘hardest’ treatments *differs* from the ‘easy’ one, whereas the δ s in turn indicate to what extent the ‘hardest’ treatment *differs* from the ‘hard’ treatment. To find the resulting coefficients for the ‘easy’ treatments, one simply takes the estimated β values. For the ‘hard’ treatments, one sums the corresponding values for the β and the γ coefficients; for the ‘hardest’ treatments one adds on top of that the corresponding δ values as well. For

¹⁹ For each rule i we had to determine the value of the variable x_i . As explained above, some behavioural rules sometimes indicate multiple targets. In these cases, we selected the target closest to the output level subsequently chosen by the player. If this still left multiple targets, we picked one of these randomly.

²⁰ $\chi^2(13) = 10.95$ for the restrictions in the duopolies, and $\chi^2(10) = 17.70$ for the triopolies.

Table 5
Estimates of Ordered-probit Models (equation (3))

	Duopolies		Triopolies	
	Coef.	Std. err.	Coef.	Std. err.
β s (for all treatments):				
<i>BESTREP</i>	0.352	0.066	0.357	0.072
<i>IMITATE</i>	0.930	0.073	0.726	0.078
<i>EXMPL</i>			0.245	0.080
<i>CUMRLD</i>			0.634	0.124
γ s (to be added for 'hard' and 'hardest'):				
<i>BESTREP</i>	-0.250	0.080	-0.299	0.083
<i>CUMRLD</i>	0.216	0.080	-0.552	0.161
δ s (to be added for 'hardest'):				
<i>CUMRLD</i>			0.335	0.149
α s (cut points):				
α_1	-0.208	0.047	-0.289	0.086
α_2	0.921	0.052	0.510	0.087
Observations:	$n = 1,134$		$n = 1,134$	
Likelihood-Ratio test:	$\chi^2(4) = 215.80$		$\chi^2(7) = 298.78$	

example, if we consider the third column of Table 5, with the coefficients for the triopolies, the estimated values pertaining to the cumulative-reinforcement-learning-with-discounting rule are 0.634 for the 'easy' treatment, 0.082(0.634 - 0.552) for the 'hard' treatment, and 0.417 (0.082 + 0.335) for the 'hardest' treatment.

We draw the following conclusions from these estimated models. First, best-responses are chosen significantly less in the 'hard' and 'hardest' treatments than in the 'easy' treatments. Second, there is no significant difference for the imitate-the-best rule across treatments. Third, in the duopolies, cumulative-reinforcement-learning-with-discounting is significantly positive only in the 'hard' and 'hardest' treatment. In the triopolies, however, cumulative-reinforcement-learning-with-discounting is already significant in the 'easy' treatment. It is significantly less so in the 'hard' triopolies, although on balance still positive. It becomes more significant again in the 'hardest' triopolies.

If we focus on the effect that the best-response rule has on the probabilities of reducing, maintaining or increasing the output level, we observe differences between the 'easy' treatments on the one hand and the 'hard' and 'hardest' treatments on the other. The probability of changing output in the direction *opposite* to the one suggested by the best-response rule is larger in the 'hard' and 'hardest' treatments than in the 'easy' treatments. If we now look at the ratio of cases in which the best-response rule suggests an output reduction to those cases where the rule suggests an increase in output in the 'hard' and 'hardest' treatments, we find that it is 1.21 in the duopolies and 2.03 in the triopolies. This may contribute to explain why output levels are slightly higher in the 'hard' and 'hardest' treatments, in particular in the triopolies.

FACT 6. *As the environment becomes more complex, the players' output decisions are significantly affected by a decline in the level of best-responses.*

The overall picture coming out of this analysis of individual period-to-period decisions is that, if anything, imitation of successful behaviour tends to decrease rather than to increase when moving to more complicated environments. Hence, in as far as there is a change in the pattern of output levels in the 'hard' and 'hardest' treatments it is not because of an increase in imitation. Apart from the vanishing of collusion when moving to the 'hard' and 'hardest' treatments (most notably in the duopolies), our experimental evidence suggests the main cause of it is the greater confusion and disorientation of the players in the 'hard' and 'hardest' treatments. This showed not only in the greater spread of output levels but also in the lower predictive success of the behavioural rules considered in this paper and, in particular, of the best-response rule (in both duopolies and triopolies). The latter also explains the slightly higher output levels in the more complicated treatments (most notably in the triopolies).

4. Comparison with Other Experiments

After we had run our experiments, we learned of some related experiments on Cournot markets; see Huck *et al.* (1999, 2000), and Offerman *et al.* (1997).²¹ These papers differ in two essential ways from ours. First, the prime questions they ask, and hence their main results, concern the relation between the information feedback made available to the players and the resulting output levels. In our paper, the main question is whether people are inclined to imitate successful behaviour and, more particularly, whether this behaviour would be more prevalent in a more demanding environment. Second, and obviously related to the previous point, their experimental designs are different from ours, as they deliberately restrict the information available to the players in the various treatments. Consequently, in some of their treatments players have no information concerning the underlying Cournot market, while in other they have no information about the other individual firms. As a result of the different information provided in their treatments, best-reply behaviour was ruled out *a priori* in some of them, whereas in others imitation of successful behaviour or collusion were made impossible *a priori*. Hence, observed changes in behaviour are to a large extent the direct result of the exclusions implied by their experimental designs. In our paper, on the other hand, being interested in the rules that the players *themselves* would actually *choose* to use in the various circumstances, we essentially maintained the same information in all treatments. In particular, the option of imitation was available to the players in each of our treatments.

Although the experiment in Huck *et al.* (1999) (HNO henceforth) had been designed to focus on different research questions, there are some treatments that deserve a careful comparison with our findings. In particular, their treatment 'full' (in which the players are fully informed about the market conditions as well as the

²¹ Since all three projects were undertaken simultaneously and independently, none of them is especially geared to test or elaborate on specific findings of the others.

actions and outcomes of all individual players in the market) is somewhat similar to our 'easy' treatment, while their 'imit' and 'imit+' treatments bear some resemblance to our 'hard' and 'hardest' treatments. In 'imit', the players were informed about the quantities and profits of all individual players but they did not know anything about the market conditions. Whereas in 'imit+' they also received some minimal qualitative information about the market conditions. The average output level observed in their 'full' treatment is considerably above Cournot–Nash and, in fact, closer to Walras than to Cournot–Nash. The average 'imit+' output choices are close to Walras and, in the 'imit' treatment, output levels are far above Walras. The main difference, then, between HNO and our findings that needs to be explained is that the output levels in all three treatments of HNO considered here are higher than in the corresponding treatments in our experiment.

There are several differences in the experimental design that might play a role in this respect. First, we consider duopolies and triopolies, while there are four firms in HNO. Second, the maximum output for an individual firm is 32 in our experiment, whereas it is 121 in HNO.²² Third, in HNO there is inertia in the sense that with probability $\frac{1}{3}$ an individual player is not allowed to change her output in each given period, while there is no such inertia in our experiment. Fourth, we consider 22 periods, and HNO 40 periods. Below we will offer some conjectures as to how these four factors might explain the general difference in output levels between the two experiments. Some of these conjectures will be based on a computational analysis but we will also offer some logical considerations.

In our computational analysis, we consider a behavioural model based on imitation of successful behaviour with different forms of experimentation as represented by the pseudo-code in Table D1 of Appendix D (on www.res.org). We consider two types of experimentation: local experimentation, and global experimentation. The first is based on Rhode and Stegeman (1995), the latter on Vega-Redondo (1997) (see Appendix D for details). Combining these two types of experimentation with a range of possible noise levels, with various levels of imitation, and taking into account all combinations of parameter values used either in our experiment or in HNO, this implies 3,600 different parameter configurations. For each of these parameter setups we realise 1,000 repetitions, computing the average output levels in the last two periods. Detailed tables concerning all results are available from the authors upon request. We summarise the findings of our analysis in the following computational results.

RESULT 1. For any type of experimentation, for any level of noise, for any level of imitation, for any number of firms, for any level of inertia, for any number of periods, the average production level is higher with a high production cap than with a low production cap.

With local experimentation and a high production cap, two things matter. First, for very low noise levels the randomly chosen initial output levels play a relevant role. As these are expected to be relatively high, with very low noise levels, output

²² This corresponds to the individual output cap of 100 in HNO, multiplied by a scaling factor of 1.21.

might still be relatively high after a given number of periods and the convergence rate might be relatively low, where for this computational analysis we define convergence to Walras as the average output during the last two periods of an iteration being between 28 and 32. As the noise levels increase beyond the lowest levels this effect soon disappears. With a low production cap, the initial output levels do not play such a role. Second, with a high production cap, as noise levels further increase, the rate of convergence to Walras goes down, but the expected output level stays rather close to Walras. With a low production cap, however, as the noise levels increase and local experimentation becomes more important relative to imitation, the expected output level is pushed down *below* Walras, because the low cap prevents players from experimenting far above Walras. That is, output levels fluctuating widely around the Walrasian output cannot occur with a low production cap. In other words, with a low production cap, the expected output level can be close to Walras only if imitation is not dominated by local experimentation, while with a high production cap, output can be close to Walras even with relatively little imitation.

If experimentation is global, the story for the low output cap is similar to the local experimentation case. However, with a high production cap, the expected output level increases considerably *above* Walras as the noise level increases and experimentation starts to dominate imitation.

What is more, with both local and global experimentation, we see a similar effect when the probability of imitation decreases. With a low output cap the expected output level goes *down* but with a high cap it goes *up*. That is, we see again that as noise becomes more important, relative to imitation, it has an opposite effect for the high cap version and the low cap setup.

Result 1, therefore, suggests a possible explanation for the fact that the output levels in our experiment are lower than in HNO, as we use a lower output cap. Notice that result 1 takes the behaviour of the players as *given*. If, in addition, a higher output cap (by making the environment more complicated) happens to lead to higher noise levels, then, in the case of global experimentation, this would reinforce the average output increase with a high output cap.²³

RESULT 2. *For any type of experimentation, for any level of noise, for any level of imitation, for any production cap, for any level of inertia, for any number of periods, there is more convergence towards Walras the higher the number of firms.*²⁴

Result 2 suggests that output in our experiment would have been higher with four instead of two or three firms (and hence closer to HNO), and that output in the 'full' treatment in HNO would have been lower with only two or three firms (and hence, again, closer to our results).²⁵

²³ For local experimentation this effect does not occur.

²⁴ This difference comes down to zero as the noise level becomes totally dominant (i.e., the support for local experimentation approaching the complete feasible output range or the tremble probability approaching 1.00).

²⁵ The findings of Huck *et al.* (2001) lend further credibility to this conjecture. Notice, however, that result 2 also suggests that output in the 'imit' and 'imit+' treatments in HNO would have been even higher if they had used two or three players instead of four.

RESULT 3. *The inertia parameter and the numbers of periods considered have hardly any effect in most parameter configurations.*²⁶

Result 3 shows that the numbers of periods used in the two experiments (either 22 or 40) and the inertia parameter do not help to explain the difference between our experiment and HNO. Both the number of periods and the inertia are not an issue with four firms (as in HNO), nor are they an issue for two or three firms with a low production cap (as in our experiment).²⁷

In addition to these computational results, which apply for *given* behaviour of the players, there is reason to believe that the different designs of our experiment and of HNO's might have led to *different behaviour* of the players. This is first of all related to the number of firms. In the easier treatments we observe the well-known phenomenon that collusive behaviour may occur with two firms but not with three or four firms. But the different production caps used in the two experiments might have led to different behaviour of the players as well.

In the treatments in which information about the market conditions is not readily available, the initial output choices tend to cluster around the middle of the feasible output range. In HNO, where individual output levels between 0 and 100 are allowed, this implies an expected initial output of around 50 per player in these treatments. Since the individual Walrasian output level in the Cournot markets studied in HNO is at 24.75, this implies huge expected losses for all players. What is important in this respect is that subjects in economic experiments *expect* to earn a positive amount of money as a reward for sensible behaviour. Therefore, when players in HNO start out making losses, they just *know* this cannot be right. Hence, they start seeking to reduce their losses and, sooner or later, they discover they need to reduce their output levels to achieve this. Now, when the players are changing their outputs downwards (whether it is through imitation or through some other algorithm), the first theoretical benchmark they encounter is Walras. This is an important landmark, as it heralds positive profits. In other words, in the experimental design of HNO, Walras is by far the most conspicuous of the three benchmarks considered.

If we compare this with our experiment, we see an important difference. In our experiment, the players expected initial output level is 20. This is right at Cournot–Nash, in the middle of the three theoretical benchmarks. This implies that the players are expected to start with positive profits. They do not, however, know whether these profits are reasonable or not. They will have to find out. Now, if the players start moving towards Walras (e.g., through imitation of successful behaviour), they notice that things are not going well, as they have accumulated in the meantime lots of evidence, through their very own experience, about where they

²⁶ This effect decreases with the number of firms and is smaller for global than for local experimentation. The only case in which output levels are substantially affected occurs with a high production cap and just two firms. If these firms experiment locally, then average output comes down with a higher number of periods and is higher with inertia than without.

²⁷ That the number of periods does not really matter confirms the view of HNO (p. C87), who also conclude that the inertia did not seem to matter in their experiment (p. C81).

could realise higher profits. Therefore, although the players may not fully understand why, they become reluctant to proceed further in the direction of Walras.²⁸

5. Concluding Remarks

Conventional wisdom asserts that in more demanding environments, imitative behaviour is increasingly likely (provided that subjects are '*boundedly rational*' and know whom to imitate), as more sophisticated forms of behaviour become harder to apply. Our experiment shows that this view is unwarranted. In spite of our efforts at scrambling information and at flashing the most successful decision on the screen in each period, the amount of imitation of successful behaviour did not increase in the more difficult treatments.

Why, then, does bounded rationality not lead to imitation in Cournot markets? On the one hand, it could be that subjects in the more demanding treatments in our experiment were mainly disoriented and just noted that imitation was not a good idea. After all, in the Cournot game, a player would worsen her own payoffs by systematically imitating more successful players. As Bandura (1986) asserts, people tend to imitate only when it is useful.²⁹ On the other hand, it could be that the players rely on more sophisticated forms of behaviour in more demanding environments than covered by our set of behaviour rules. After all, boundedly rational as subjects can be, they still possess a whole arsenal of responses to somebody else's actions, starting with the use of their imagination; see Selten (1978).³⁰

Our conclusion that players are not very impressed by information external to themselves conforms to experimental evidence in a variety of different settings. For example, Allsopp and Hey (1997) test for herd behaviour in an experimental setup and find that players rely much more on their private signals than some theoretical work had suggested; see Banerjee (1992). Kraemer *et al.* (2000) report that, in experiments on information aggregation, the players do not aggregate enough and rely too much on private information signals. A similar result was found by Huck and Oechssler (1999) in experiments on information cascades.

Economic theory also provides some clues as to why the more complicated environments might not lead the players to move towards Walrasian output levels in Cournot games. Dixon *et al.* (1995) show that there are evolutionary processes that lead to Pareto outcomes and Eichberger and Kelsey (1999) demonstrate how Knightian uncertainty leads to an equilibrium with output levels *below* Cournot–Nash in a one-shot Cournot game.

²⁸ Notice that the experimental design of HNO implies that the relevant piece of information inducing this reluctance is much less apparent, as the players in HNO have much less experience of profits around Cournot–Nash.

²⁹ In Björnerstedt and Schlag (1996) terminology, imitating the best is not '*strictly improving*' in Cournot markets. That is, when used by everyone, the rule does not enable to learn which action maximises expected payoff. Björnerstedt and Schlag show that such rules might not be evolutionary robust.

³⁰ A Cournot game allows for subtle behaviour. It would be surprising if subjects got stuck in a strategy of plain imitation. Consider, as an example in another context, the richness of reciprocal interactions among boundedly rational beings, in this case guillemots (*Uriae algae*), as reported by Roberts and Sherratt (1998).

We have concluded that confronting players more intensely with their bounded rationality does not lead to the prevalence of imitation of successful behaviour in Cournot games. One should bear in mind, however, that the purpose of our paper was not to accept or reject a finished theory of boundedly rational behaviour. Our understanding of boundedly rational behaviour, including imitation, is still limited, and our paper should be seen as a contribution to the long-term goal of understanding this behaviour.

Universitat Pompeu Fabra
Queen Mary, University of London

Date of receipt of first submission: May 1999

Date of receipt of final typescript: March 2002

References

- Allsopp, L. and Hey, J.D. (1997). 'An experiment to test a model of herd behaviour', mimeo.
- Bandura, A. (1986). *Social Foundations of Thought and Action. A Social Cognitive Theory*, Englewood Cliffs, NJ: Prentice-Hall.
- Banerjee, A.V. (1992). 'A simple model of herd behavior', *Quarterly Journal of Economics*, vol. 107, pp. 797–817.
- Bikhchandani, S., Hirshleifer, D. and Welch, I. (1998). 'Learning from behavior of others: conformity, fads, and informational cascades', *Journal of Economic Perspectives*, vol. 12(3), pp. 151–70.
- Björnerstedt, J. and Schlag, K.H. (1996). 'On the evolution of imitative behavior', mimeo, University of Bonn.
- Blackmore, S. (2000). *The Meme Machine*, Oxford: Oxford University Press.
- Crawford, V. (1998). 'A survey of experiments on communication via cheap talk', *Journal of Economic Theory*, vol. 78, pp. 286–98.
- Davis, D.D. (1995). 'Advance production, Cournot outcomes and trigger strategies: an experimental investigation', mimeo.
- Dixon, H., Wallis, S. and Moss, S. (1995). 'Axelrod meets Cournot: oligopoly and the evolutionary metaphor part 1', Discussion Papers in Economics No. 95/8, University of York.
- The Economist* (1999). 'Chimpanzee behaviour. Culture club', June 19th, p. 120.
- Eichberger, J. and Kelsey, D. (1999). 'Uncertainty and strategic interaction in economics', mimeo.
- Fouraker, L.E. and Siegel, S. (1963). *Bargaining Behavior*, New York: McGraw-Hill.
- Gigerenzer, G. and Goldstein, D.G. (1996). 'Reasoning the fast and frugal way: models of bounded rationality', *Psychological Review*, vol. 103, pp. 650–69.
- Hamilton, W. (1970). 'Selfish and spiteful behavior in an evolutionary model', *Nature*, vol. 228, pp. 1218–25.
- Holt, C.A. (1995). 'Industrial organization: a survey of laboratory research', in (J. Kagel and A.E. Roth eds.) *The Handbook of Experimental Economics* (pp. 349–443), Princeton, NJ: Princeton University Press.
- Huck, S., Normann, H.-T. and Oechssler, J. (1999). 'Learning in Cournot oligopoly – an experiment', *Economic Journal*, vol. 109, pp. C80–95.
- Huck, S., Normann, H.-T. and Oechssler, J. (2000). 'Does information about competitors' actions increase or decrease competition in experimental oligopoly markets?', *International Journal of Industrial Organization*, vol. 18(1), pp. 39–57.
- Huck, S., Normann, H.T. and Oechssler, J. (2001). 'Two are few and four are many: number effects in experimental oligopolies', Economic Discussion Paper 12/2001, University of Bonn.
- Huck, S. and Oechssler, J. (1999). 'informational cascades in the laboratory: do they occur for the right reasons?', mimeo.
- Kraemer, C., Nöth, M. and Weber, M. (2000). 'Information aggregation with costly information and random ordering: experimental evidence', SFB 504 discussion paper 00–35, University of Mannheim.
- Levine, D.K. (1998). 'Modeling altruism and spitefulness in experiments', *Review of Economic Dynamics*, vol. 1(3).
- Mirsky, S. (1996). 'On presidents and king', *Scientific American*, vol. 275(5), November, pp. 20–1.

- Nelson, R.R. and Winter, S.G. (1982). *An Evolutionary Theory of Economic Change*, Cambridge, MA: Harvard University Press.
- Offerman, T., Potters, J. and Sonnemans, J. (1997). 'Imitation and belief learning in an oligopoly experiment', mimeo, University of Amsterdam, Department of Economics.
- Rhode, P. and Stegeman, M. (1995). 'Evolution through imitation (with applications to duopoly)', Working Paper E95-43, Virginia Polytechnic Institute and State University, Department of Economics.
- Roberts, G. and Sherratt, T.N. (1998). 'Development of cooperative relationships through increasing investment', *Nature*, vol. 394, pp. 175-9.
- Sargent, T.J. (1993). *Bounded Rationality in Macroeconomics*, Oxford: Oxford University Press.
- Schaffer, M.E. (1988a). 'Economic natural selection and "relative" versus "absolute" profit maximisation', mimeo, London School of Economics.
- Schaffer, M.E. (1988b). 'Evolutionary stable strategies for a finite population and a variable contest size', *Journal of Theoretical Biology*, vol. 132, pp. 469-78.
- Schaffer, M.E. (1989). 'Are profit maximizers the best survivors?', *Journal of Economic Behavior and Organization*, vol. 12, pp. 29-45.
- Selten, R. (1978). 'Chain-store paradox', *Theory and Decision*, vol. 9, pp. 127-59.
- Selten, R. (1991). 'Properties of a measure of predictive success', *Mathematical Social Sciences*, vol. 21(2), pp. 153-67.
- Vega-Redondo, F. (1997). 'The evolution of Walrasian behavior', *Econometrica*, vol. 65, pp. 375-84.
- Weibull, J. (1995). *Evolutionary Game Theory*, Cambridge, MA: MIT Press.

Technical Appendix to IMITATION OF SUCCESSFUL BEHAVIOUR IN COURNOT MARKETS*

Antoni Bosch-Domènech and Nicolaas J. Vriend

ECONOMIC JOURNAL, vol. 113 (April), pp. 495–524

Appendix A. The Spite Effect and the Cournot Oligopoly

The essence of the spite effect is illustrated by the bimatrix game in Figure A1 (Palomino, 1995), where T and B are the two possible strategies, and the lowercase letters are the payoffs to the row and column player, with $a > b > c > d$. Clearly, (T, T) is the only Nash equilibrium since no player can improve by deviating from it, and this is the only combination for which this holds. Now, consider the strategy pair (B, T) , leading to the payoffs (b, c) . Remember that $a > b > c > d$. Hence, by deviating from the Nash equilibrium, the row player hurts her own payoff, but she hurts the column player's payoff even more.

Let us now focus on a standard symmetric Cournot oligopoly. There are several symmetrical firms producing the same homogeneous commodity. The only decision variable for firm i is the quantity q_i to be produced. Once production has taken place, for all firms simultaneously, the firms bring their output to the market, where the market price P is determined such that demand equals supply. To give the intuition behind the spite effect in this Cournot game, let us consider a simple symmetric Cournot market in which the inverse demand function is $P(Q) = a + bQ$, where $Q = \sum q_i$, and in which the cost function for the individual firm is $TC(q) = K + kq$. Making the appropriate assumptions on the parameters a and b ensures that the demand curve is downward-sloping. We can distinguish three symmetric output levels of the static Cournot oligopoly game specified above for the case in which the

	T	B
T	a, a	c, b
B	b, c	d, d

Fig. A1. *Bimatrix Game with Payoffs $a > b > c > d$*

* We wish to thank Xavier Herrero for his indispensable assistance in organising the experiments, and Tilman Börgers, Gary Charness, Richard Disney, Ido Erev, Jonathan Haskel, Charles Holt, Steffen Huck, Ulrich Kamecke, Francesco Luna, Paul Madden, Paola Manzini, Charles Noussair, Martin Sefton, Tilman Slembeck, Fernando Vega-Redondo, Nir Vulkan, and seminar and conference participants at Venice, Siena, Exeter, Manchester, London (QM), the London Business School, ESA (Mannheim), EARIE (Copenhagen), ESEM and EEA (Berlin), the ESRC Research Seminar in Game Theory (Kenilworth), RES (Nottingham), the Max Planck Institute (Jena), and SAET (Ischia) for discussions or comments concerning earlier versions of this paper. Two anonymous referees and the editor provided many very valuable suggestions. Financial support from the Spanish Ministry of Education and Science under contract PB98-1079 (AB-D), and from the Commission of the European Union under contracts FMRX-CT98-0238 (AB-D) and FMBICT950277 (NJV) is gratefully acknowledged. The usual disclaimer applies.

players have complete information. First, suppose that the two firms collude, maximising their joint-profits. This leads to an aggregate output level called Pareto $Q^P = (k - a)/(2b)$. Second, if the firms behave as price-takers in a competitive market, they simply produce up to the point where their marginal costs are equal to the market price P . Given the specification of the oligopoly model above, this implies an aggregate competitive, or Walrasian, output level of $Q^W = (k - a)/b$. If, instead, the firms realise that they influence the market price through their own output, they produce up to the point where their marginal costs are equal to their marginal revenue. Taking the output level of the other firm as given, this leads to an aggregate Cournot–Nash equilibrium output of $Q^N = (k - a)/\{b[(1/n) + 1]\}$.

To see how a spite effect might influence the outcomes of a Cournot market game, suppose, to simplify for illustrative convenience, that there are only two firms, that fixed and marginal costs are zero (Schaffer, 1989) and let us concentrate on the Walrasian equilibrium. Observe that there are two alternative ways to look at it, based on different behavioural assumptions. In both cases it is the spite effect that makes it an equilibrium. First, suppose that the firms' preferences are such that they do not care about absolute payoffs but only about relative payoffs. Any utility function assigning a higher value to an outcome in which the firm beats the other firm and a lower value to an outcome in which it gets beaten will, after elimination of all weakly dominated strategies, leave only one strategy: producing its equal share of Q^W .

To see why this is the only strategy where a firm is sure it can never be beaten, look at Figure A2 and focus on the Walrasian output Q^W . Suppose firm i produces its equal share of the Walrasian output: $q_i = Q^W/2$. If firm j does the same, aggregate output is Q^W , the market price P is zero and both make a zero profit. What happens when firm j produces more than $Q^W/2$? The price P will become negative, and both firms will make losses. But it is firm i that makes less losses, because it has a lower output level sold at the same market price P . What happens instead if firm j produces less than $Q^W/2$? The price P will be positive and, hence, this will increase firm j 's profits. But again it is firm i that makes a greater profit, because it

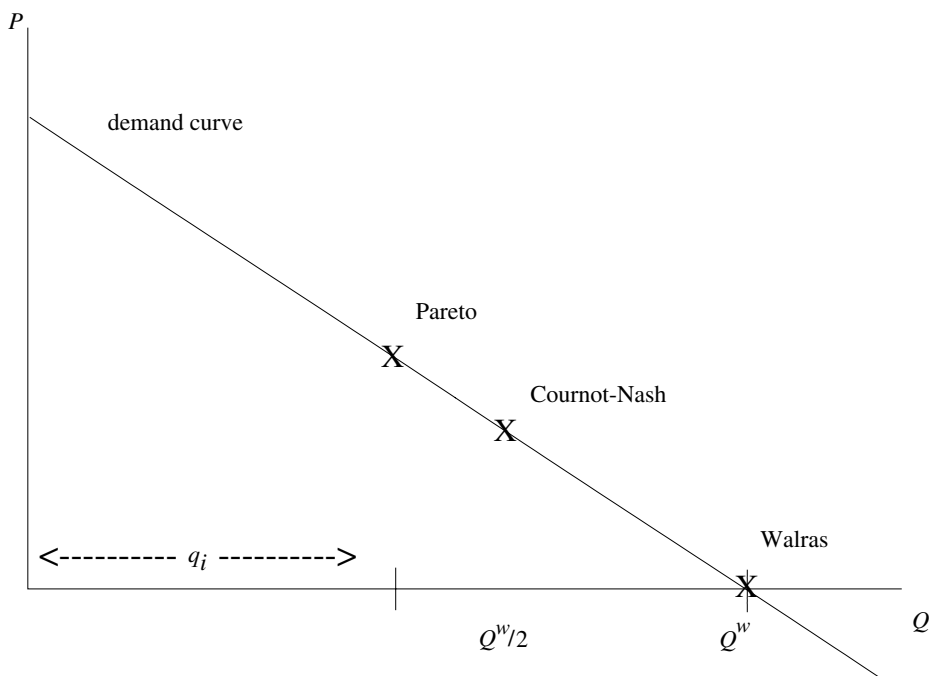


Fig. A2. Example Cournot Duopoly

has a higher output level sold at the same market price P . In some sense, firm i is free riding on firm j 's production restraint. Hence, the firm that produces its equal share of Q^W will have the highest relative payoff in this Cournot duopoly. Note that this implies in particular the following. If firm i produces its share of the symmetric Walrasian output, while firm j naively chooses the symmetric output level to maximise its absolute payoffs (i.e., its equal share of the Cournot-Nash output), it is firm i that realises the highest profits. Moreover, even if firm j is aware of the fact that firm i is producing at the Walrasian output level, and maximises its profits taking this into account, it is firm i that realises the highest payoffs. If we consider more than two firms, matters become slightly more complicated, but the following holds. Whenever the aggregate output level is below Walras, i.e., on average an individual firm produces less than its share of the Walrasian output level, the price will be positive, and it is the firms with the higher output levels that generate the higher profits. Exactly the reverse holds when aggregate output exceeds the Walrasian output level: the lower a firm's output level, the higher its profits will be.

Now, suppose that the firms do not have a preference for beating their competitors but that they are boundedly rational and tend to imitate successful behaviour in the sense that, with a certain probability, they choose the output levels that led to the highest profits in the past. Whenever the average output is below Walras, it is the highest output firm, realising the highest profit, that is most likely to be imitated, and the other way round. As a result, the market will converge to the Walrasian equilibrium.

Appendix B. Instructions to the Players

Table B1 gives the English translation of the Spanish instructions to the players in the 'easy' duopoly.

Table B1
Instructions 'easy' Duopoly

Instructions
<p><i>Introduction</i></p> <ul style="list-style-type: none"> • This is a decision experiment. The instructions are simple and, if you pay attention, you can gain a reasonable amount of money that will be paid to you at the end of the experiment. From now on till the end of the experiment you are not allowed to communicate with each other. If you have a question, please raise your hand. • Each of you will play a firm that produces a fictitious good that is sold in a fictitious market. • Within each market there will be only 2 firms that sell the same good. One is your firm and the other is a firm that is identical to yours. • Who will be this other firm will be decided randomly. • The other people in the laboratory participate in other markets that have nothing to do with yours. In other words, various markets will operate simultaneously but independently in the laboratory. • You will never know the identity of the person you are matched with, nor will he be aware of yours. • The experiment will last 22 consecutive periods and the other firm that participates in your market will be the same during all periods of the experiment. <p><i>Decisions and Outcomes</i></p> <ul style="list-style-type: none"> • Each period all firms simultaneously make only 1 decision: the quantity to be produced and supplied to the market. Only integer values from 8 to 32 can be chosen. • You will get a Table showing the various levels of profit or loss you and the other firms can attain depending upon the quantities chosen by you and the other firm. The quantities one firm (firm X) may produce are listed across the top of the Table, while the quantities produced by the other firm are listed down the left-hand margin. The profits for firm X and for the other firm are given within the body of the Table by the intersection of the quantities produced. The top number in bold gives the profit for firm X, whereas the bottom number in <i>italic</i> gives the profit of the other firm. Since the two firms are identical, at any moment you can identify either yourself or the other firm with firm X. We will do some exercises with the Table in a moment.

Table B1

Continued

Instructions
<ul style="list-style-type: none"> • After each period, you will get some information on your screen. At the top of the screen, you will see your output level and that of the other firm in the previous period. At the bottom, you will see the history of your own output levels and profits realised. • There is no time limit for your period to period decisions. Decisions will ordinarily be made every few minutes or so. <p><i>Payment</i></p> <ul style="list-style-type: none"> • Each player gets a fixed fee of 250 Pesetas just for participating in the experiment. • In addition, each player will be paid according to the total profits realised by his firm. • During periods 1 to 20, the monetary reward will be 0.035 Pesetas for every profit point realised. • During periods 21 and 22 (the last 2 periods), the monetary reward will be 0.35 Pesetas for each profit point realised. You will receive a reminder of this higher payoff (10 times as high) at the start of period 21. • Note that losses realised will be subtracted from the 250 Pesetas. • At the end of the experiment, we will add up your profits and calculate your monetary rewards. This will be done such that you will not see what other players earned. <p><i>Keyboard</i></p> <ul style="list-style-type: none"> • To make your choice of output level, please enter a number. Remember that only integer values from 8 to 32 can be chosen. • To confirm (or not) your choices, enter Y (or N) with your keyboard. • Please, before confirming your choices, always make sure that you did not make a typing error.

Table B2 shows the instructions given to the players in the ‘hard’ duopoly. We only list the subsection ‘Decisions and Outcomes’, which replaces the corresponding subsection in the ‘easy’ duopoly. The remainder of the instructions was identical to the ‘easy’ version.

The only change made in the instructions of the ‘hardest’ duopoly with respect to the ‘hard’ version was that the information concerning the market demand was removed, that is, the item marked (•) in Table B2.

Table B2

Instructions ‘hard’ Duopoly

Instructions
<p><i>Decisions and Outcomes</i></p> <ul style="list-style-type: none"> • Each period all firms simultaneously make only 1 decision: the quantity to be produced and supplied to the market. Only integer values from 8 to 32 can be chosen. (•) Given the TOTAL quantity supplied to the market by you and the other firm in a given period, the price is determined by the market. For total output levels from 16 to 64, taking steps of 1, the market prices will be 350 (with total output equal to 16), 346, 342, 338, 334, 330, 326, 322, 318, 314, 310, 306, 302, 298, 294, 290, 286, 282, 278, 274, 270, 266, 262, 258, 254, 250, 246, 242, 238, 234, 230, 226, 222, 218, 214, 210, 206, 202, 198, 194, 190, 186, 182, 178, 174, 170, 166, 162, 158 (64). This market price implies the revenue a firm gets for EACH UNIT it supplied to the market. Assume that all units produced are actually sold. • For a given period, the costs to a firm producing a certain quantity in that period are as follows, starting with the minimum output of 8, and going in unit steps to the maximum output of 32: 1246 (with output equal to 8), 1420, 1594, 1768, 1942, 2116, 2290, 2464, 2638, 2812, 2986, 3160, 3334, 3508, 3682, 3856, 4030, 4204, 4378, 4552, 4726, 4900, 5074, 5248, 5422 (32). • The profits to a firm for a given period are simply its revenues minus its costs. • After each period, you will get some information on your screen. You will see your output level and that of the other firm in the previous period, plus the profits realised by you and by the other firms in the that period. We also indicate (with *****) which firm realised the highest profit in the previous period. • There is a 1 minute time limit for your period to period decisions. The experimenter will give a warning after 30 seconds, after 50 seconds and after 60 seconds.

The following Table was given to the players in the 'easy' duopolies, with the **bold** faced numbers indicating the profits for firm X, and the numbers in *italic* the profits for the other firm.

		Output Firm X											
PROFITS		8	9	10	11	12	13	14	15	16	17	18	19
<i>Output</i>		1554	1694	1826	1950	2066	2174	2274	2366	2450	2526	2594	2654
<i>other firm</i>	8	<i>1554</i>	<i>1522</i>	<i>1490</i>	<i>1458</i>	<i>1426</i>	<i>1394</i>	<i>1362</i>	<i>1330</i>	<i>1298</i>	<i>1266</i>	<i>1234</i>	<i>1202</i>
		1522	1658	1786	1906	2018	2122	2218	2306	2386	2458	2522	2578
	9	<i>1694</i>	<i>1658</i>	<i>1622</i>	<i>1586</i>	<i>1550</i>	<i>1514</i>	<i>1478</i>	<i>1442</i>	<i>1406</i>	<i>1370</i>	<i>1334</i>	<i>1298</i>
		1490	1622	1746	1862	1970	2070	2162	2246	2322	2390	2450	2502
	10	<i>1826</i>	<i>1786</i>	<i>1746</i>	<i>1706</i>	<i>1666</i>	<i>1626</i>	<i>1586</i>	<i>1546</i>	<i>1506</i>	<i>1466</i>	<i>1426</i>	<i>1386</i>
		1458	1586	1706	1818	1922	2018	2106	2186	2258	2322	2378	2426
	11	<i>1950</i>	<i>1906</i>	<i>1862</i>	<i>1818</i>	<i>1774</i>	<i>1730</i>	<i>1686</i>	<i>1642</i>	<i>1598</i>	<i>1554</i>	<i>1510</i>	<i>1466</i>
		1426	1550	1666	1774	1874	1966	2050	2126	2194	2254	2306	2350
	12	<i>2066</i>	<i>2018</i>	<i>1970</i>	<i>1922</i>	<i>1874</i>	<i>1826</i>	<i>1778</i>	<i>1730</i>	<i>1682</i>	<i>1634</i>	<i>1586</i>	<i>1538</i>
		1394	1514	1626	1730	1826	1914	1994	2066	2130	2186	2234	2274
	13	<i>2174</i>	<i>2122</i>	<i>2070</i>	<i>2018</i>	<i>1966</i>	<i>1914</i>	<i>1862</i>	<i>1810</i>	<i>1758</i>	<i>1706</i>	<i>1654</i>	<i>1602</i>
		1362	1478	1586	1686	1778	1862	1938	2006	2066	2118	2162	2198
	14	<i>2274</i>	<i>2218</i>	<i>2162</i>	<i>2106</i>	<i>2050</i>	<i>1994</i>	<i>1938</i>	<i>1882</i>	<i>1826</i>	<i>1770</i>	<i>1714</i>	<i>1658</i>
		1330	1442	1546	1642	1730	1810	1882	1946	2002	2050	2090	2122
	15	<i>2366</i>	<i>2306</i>	<i>2246</i>	<i>2186</i>	<i>2126</i>	<i>2066</i>	<i>2006</i>	<i>1946</i>	<i>1886</i>	<i>1826</i>	<i>1766</i>	<i>1706</i>
		1298	1406	1506	1598	1682	1758	1826	1886	1938	1982	2018	2046
	16	<i>2450</i>	<i>2386</i>	<i>2322</i>	<i>2258</i>	<i>2194</i>	<i>2130</i>	<i>2066</i>	<i>2002</i>	<i>1938</i>	<i>1874</i>	<i>1810</i>	<i>1746</i>
		1266	1370	1466	1554	1634	1706	1770	1826	1874	1914	1946	1970
	17	<i>2526</i>	<i>2458</i>	<i>2390</i>	<i>2322</i>	<i>2254</i>	<i>2186</i>	<i>2118</i>	<i>2050</i>	<i>1982</i>	<i>1914</i>	<i>1846</i>	<i>1778</i>
		1234	1334	1426	1510	1586	1654	1714	1766	1810	1846	1874	1894
	18	<i>2594</i>	<i>2522</i>	<i>2450</i>	<i>2378</i>	<i>2306</i>	<i>2234</i>	<i>2162</i>	<i>2090</i>	<i>2018</i>	<i>1946</i>	<i>1874</i>	<i>1802</i>
		1202	1298	1386	1466	1538	1602	1658	1706	1746	1778	1802	1818
	19	<i>2654</i>	<i>2578</i>	<i>2502</i>	<i>2426</i>	<i>2350</i>	<i>2274</i>	<i>2198</i>	<i>2122</i>	<i>2046</i>	<i>1970</i>	<i>1894</i>	<i>1818</i>
		1170	1262	1346	1422	1490	1550	1602	1646	1682	1710	1730	1742
	20	<i>2706</i>	<i>2626</i>	<i>2546</i>	<i>2466</i>	<i>2386</i>	<i>2306</i>	<i>2226</i>	<i>2146</i>	<i>2066</i>	<i>1986</i>	<i>1906</i>	<i>1826</i>
		1138	1226	1306	1378	1442	1498	1546	1586	1618	1642	1658	1666
	21	<i>2750</i>	<i>2666</i>	<i>2582</i>	<i>2498</i>	<i>2414</i>	<i>2330</i>	<i>2246</i>	<i>2162</i>	<i>2078</i>	<i>1994</i>	<i>1910</i>	<i>1826</i>
		1106	1190	1266	1334	1394	1446	1490	1526	1554	1574	1586	1590
	22	<i>2786</i>	<i>2698</i>	<i>2610</i>	<i>2522</i>	<i>2434</i>	<i>2346</i>	<i>2258</i>	<i>2170</i>	<i>2082</i>	<i>1994</i>	<i>1906</i>	<i>1818</i>
		1074	1154	1226	1290	1346	1394	1434	1466	1490	1506	1514	1514
	23	<i>2814</i>	<i>2722</i>	<i>2630</i>	<i>2538</i>	<i>2446</i>	<i>2354</i>	<i>2262</i>	<i>2170</i>	<i>2078</i>	<i>1986</i>	<i>1894</i>	<i>1802</i>
		1042	1118	1186	1246	1298	1342	1378	1406	1426	1438	1442	1438
	24	<i>2834</i>	<i>2738</i>	<i>2642</i>	<i>2546</i>	<i>2450</i>	<i>2354</i>	<i>2258</i>	<i>2162</i>	<i>2066</i>	<i>1970</i>	<i>1874</i>	<i>1778</i>
		1010	1082	1146	1202	1250	1290	1322	1346	1362	1370	1370	1362
	25	<i>2846</i>	<i>2746</i>	<i>2646</i>	<i>2546</i>	<i>2446</i>	<i>2346</i>	<i>2246</i>	<i>2146</i>	<i>2046</i>	<i>1946</i>	<i>1846</i>	<i>1746</i>
		978	1046	1106	1158	1202	1238	1266	1286	1298	1302	1298	1286
	26	<i>2850</i>	<i>2746</i>	<i>2642</i>	<i>2538</i>	<i>2434</i>	<i>2330</i>	<i>2226</i>	<i>2122</i>	<i>2018</i>	<i>1914</i>	<i>1810</i>	<i>1706</i>
		946	1010	1066	1114	1154	1186	1210	1226	1234	1234	1226	1210
	27	<i>2846</i>	<i>2738</i>	<i>2630</i>	<i>2522</i>	<i>2414</i>	<i>2306</i>	<i>2198</i>	<i>2090</i>	<i>1982</i>	<i>1874</i>	<i>1766</i>	<i>1658</i>
		914	974	1026	1070	1106	1134	1154	1166	1170	1166	1154	1134
	28	<i>2834</i>	<i>2722</i>	<i>2610</i>	<i>2498</i>	<i>2386</i>	<i>2274</i>	<i>2162</i>	<i>2050</i>	<i>1938</i>	<i>1826</i>	<i>1714</i>	<i>1602</i>
		882	938	986	1026	1058	1082	1098	1106	1106	1098	1082	1058
	29	<i>2814</i>	<i>2698</i>	<i>2582</i>	<i>2466</i>	<i>2350</i>	<i>2234</i>	<i>2118</i>	<i>2002</i>	<i>1886</i>	<i>1770</i>	<i>1654</i>	<i>1538</i>
		850	902	946	982	1010	1030	1042	1046	1042	1030	1010	982
	30	<i>2786</i>	<i>2666</i>	<i>2546</i>	<i>2426</i>	<i>2306</i>	<i>2186</i>	<i>2066</i>	<i>1946</i>	<i>1826</i>	<i>1706</i>	<i>1586</i>	<i>1466</i>
		818	866	906	938	962	978	986	986	978	962	938	906
	31	<i>2750</i>	<i>2626</i>	<i>2502</i>	<i>2378</i>	<i>2254</i>	<i>2130</i>	<i>2006</i>	<i>1882</i>	<i>1758</i>	<i>1634</i>	<i>1510</i>	<i>1386</i>
		786	830	866	894	914	926	930	926	914	894	866	830
	32	<i>2706</i>	<i>2578</i>	<i>2450</i>	<i>2322</i>	<i>2194</i>	<i>2066</i>	<i>1938</i>	<i>1810</i>	<i>1682</i>	<i>1554</i>	<i>1426</i>	<i>1298</i>

		Output firm X												
PROFITS		20	21	22	23	24	25	26	27	28	29	30	31	32
<i>Output</i>		2706	2750	2786	2814	2834	2846	2850	2846	2834	2814	2786	2750	2706
<i>other firm</i>	8	1170	1138	1106	1074	1042	1010	978	946	914	882	850	818	786
		2626	2666	2698	2722	2738	2746	2746	2738	2722	2698	2666	2626	2578
	9	1262	1226	1190	1154	1118	1082	1046	1010	974	938	902	866	830
		2546	2582	2610	2630	2642	2646	2642	2630	2610	2582	2546	2502	2450
	10	1346	1306	1266	1226	1186	1146	1106	1066	1026	986	946	906	866
		2466	2498	2522	2538	2546	2546	2538	2522	2498	2466	2426	2378	2322
	11	1422	1378	1334	1290	1246	1202	1158	1114	1070	1026	982	938	894
		2386	2414	2434	2446	2450	2446	2434	2414	2386	2350	2306	2254	2194
	12	1490	1442	1394	1346	1298	1250	1202	1154	1106	1058	1010	962	914
		2306	2330	2346	2354	2354	2346	2330	2306	2274	2234	2186	2130	2066
	13	1550	1498	1446	1394	1342	1290	1238	1186	1134	1082	1030	978	926
		2226	2246	2258	2262	2258	2246	2226	2198	2162	2118	2066	2006	1938
	14	1602	1546	1490	1434	1378	1322	1266	1210	1154	1098	1042	986	930
		2146	2162	2170	2170	2162	2146	2122	2090	2050	2002	1946	1882	1810
	15	1646	1586	1526	1466	1406	1346	1286	1226	1166	1106	1046	986	926
		2066	2078	2082	2078	2066	2046	2018	1982	1938	1886	1826	1758	1682
	16	1682	1618	1554	1490	1426	1362	1298	1234	1170	1106	1042	978	914
		1986	1994	1994	1986	1970	1946	1914	1874	1826	1770	1706	1634	1554
	17	1710	1642	1574	1506	1438	1370	1302	1234	1166	1098	1030	962	894
		1906	1910	1906	1894	1874	1846	1810	1766	1714	1654	1586	1510	1426
	18	1730	1658	1586	1514	1442	1370	1298	1226	1154	1082	1010	938	866
		1826	1826	1818	1802	1778	1746	1706	1658	1602	1538	1466	1386	1298
	19	1742	1666	1590	1514	1438	1362	1286	1210	1134	1058	982	906	830
		1746	1742	1730	1710	1682	1646	1602	1550	1490	1422	1346	1262	1170
	20	1746	1666	1586	1506	1426	1346	1266	1186	1106	1026	946	866	786
		1666	1658	1642	1618	1586	1546	1498	1442	1378	1306	1226	1138	1042
	21	1742	1658	1574	1490	1406	1322	1238	1154	1070	986	902	818	734
		1586	1574	1554	1526	1490	1446	1394	1334	1266	1190	1106	1014	914
	22	1730	1642	1554	1466	1378	1290	1202	1114	1026	938	850	762	674
		1506	1490	1466	1434	1394	1346	1290	1226	1154	1074	986	890	786
	23	1710	1618	1526	1434	1342	1250	1158	1066	974	882	790	698	606
		1426	1406	1378	1342	1298	1246	1186	1118	1042	958	866	766	658
	24	1682	1586	1490	1394	1298	1202	1106	1010	914	818	722	626	530
		1346	1322	1290	1250	1202	1146	1082	1010	930	842	746	642	530
	25	1646	1546	1446	1346	1246	1146	1046	946	846	746	646	546	446
		1266	1238	1202	1158	1106	1046	978	902	818	726	626	518	402
	26	1602	1498	1394	1290	1186	1082	978	874	770	666	562	458	354
		1186	1154	1114	1066	1010	946	874	794	706	610	506	394	274
	27	1550	1442	1334	1226	1118	1010	902	794	686	578	470	362	254
		1106	1070	1026	974	914	846	770	686	594	494	386	270	146
	28	1490	1378	1266	1154	1042	930	818	706	594	482	370	258	146
		1026	986	938	882	818	746	666	578	482	378	266	146	18
	29	1422	1306	1190	1074	958	842	726	610	494	378	262	146	30
		946	902	850	790	722	646	562	470	370	262	146	22	-110
	30	1346	1226	1106	986	866	746	626	506	386	266	146	26	-94
		866	818	762	698	626	546	458	362	258	146	26	-102	-238
	31	1262	1138	1014	890	766	642	518	394	270	146	22	-102	-226
		786	734	674	606	530	446	354	254	146	30	-94	-226	-366
	32	1170	1042	914	786	658	530	402	274	146	18	-110	-238	-366

Figures B1 and B2 present two examples of the screens faced by the players. First, the 'easy' version of the triopolies in Figure B1.

previous period (period 2):		your production:	20
		production firm X:	8
		production firm Y:	17
next period (period 3):		your production:	... (please Enter)
history:			
period	your production	your profit	
1	23	2658	
2	20	3786	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
.....			
21	
22	

Fig. B1. Example Screen 'easy' Triopoly

The screen faced by the players in the 'hard' and 'hardest' triopolies is shown in Figure B2.

previous period (period 2):			
	production	profit	best
you	20	3786	*****
firm X	8	1674	
firm Y	17	3258	
next period (period 3):			
	your production:	...	(please Enter)

Fig. B2. Example Screen 'hard' and 'hardest' Triopoly

Appendix C. Description of the Individual Behavioural Rules

1 *BESTREP* myopic best-reply

The myopic-best-reply rule chooses a best-reply against the most recent output choice(s) by the other player(s). In the case where we check only for the direction of a player's choice, this corresponds to what is known as learning direction theory; Selten and Stoecker (1986).

2 *ADPTEXP* adaptive expectations

The adaptive-expectations rule adapts the belief about the (aggregate) output of the other player(s) by taking a convex combination of the existing belief and the most recent (aggregate) output of the other player(s), with a weight of 0.75 on the former and 0.25 on the latter. Then, the rule chooses a best-response against this updated belief.

3 *FICT* fictitious play

The fictitious-play rule chooses a best-response against the relative frequency distribution of past output levels by the other firm(s).¹

4 *GEOFICT* geometric fictitious play

The geometric-fictitious-play rule is similar to rule 3. But now, in each period, previous frequency observations are discounted, multiplying them by 0.75.² Notice that the case of extreme discounting would correspond to myopic best-reply. Also notice that, given the linear Cournot model, for any discount factor between zero and one, this rule is very similar to the adaptive-expectations rule 2.

5 *IMITATE* imitate-the-best

The imitate-the-best rule is a restricted representation of the general class of imitation of successful behaviour on which we focus in this paper. The rule chooses to imitate the output level of the firm with the highest profits in the previous period.

6 *IMAVG* imitate-the-average

The imitate-the-average rule simply imitates the previous average output of the other player(s) in the market. When we consider only the direction of the players' output decisions, this rule is very similar to a rule Fouraker and Siegel (1963) analyse related to aspiration levels. Their '*rivalistic*' hypothesis says that if a firm's output in the previous period was below that of its rival, then it will increase its output in the next period and the other way round. There are two, very different, reasons to look at the imitate-the-average rule. First, it is a dynamic, reciprocating strategy of the type '*give-as-good-as-you-get*' (GGG). Notice that in the duopoly this corresponds exactly to the Tit-for-Tat strategy (see, e.g. Axelrod (1984)). Second, Huck *et al.* (1999) argue that the imitate-the-average rule makes sense because players might reason that whereas individual players might go wrong, the average opponent cannot be too wrong.³

7 *EXMPL* imitate-the-exemplary

The imitate-the-exemplary rule is due to Offerman *et al.* (1997). The rule chooses the output level from among all those chosen in the previous period, such that if all firms would imitate that exemplary firm, it would lead to the highest profits.

8 *STAY* stay put (imitate oneself)

The stay-put rule simply tells a firm to repeat the same output level as the previous period.

¹ Since our Cournot model is linear, this is equivalent to choosing a best-response against the expected output level for the other firm(s). Hence, our rule also corresponds to what Offerman *et al.* (1997) call '*adapted fictitious play*'.

² This parameter value corresponds to Cason and Friedman (2000), who refer to some experimental evidence in this respect.

³ In addition, they consider this rule because the imitate-the-best rule could not be applied in some of their treatments due to information restrictions.

9 *AVGRL* average reinforcement learning

The average-reinforcement-learning rule selects an output level on the basis of a player's average realised payoffs. For all reinforcement learning rules we use a deterministic version with initial strengths zero, and the target is always the most reinforced output level.

10 *AVGRDL* average reinforcement learning with discounting

The average-reinforcement-learning-with-discounting rule is similar to the previous rule, but now the average reinforcement is a weighted average of previous performances. Each period, previous performance observations are discounted, multiplying their weight by 0.75.

11 *CUMRL* cumulative reinforcement learning

The cumulative-reinforcement-learning rule corresponds to the average-reinforcement-learning rule with average reinforcement substituted by reinforcement on the basis of a player's cumulative realised payoffs.

12 *CUMRLD* cumulative reinforcement learning with discounting

The cumulative-reinforcement-learning-with-discounting rule adds the discounting (as above) to the previous rule.

13 *AVGXRL* extended average reinforcement learning

The extended reinforcement learning rules differ from the standard reinforcement learning rules in that they also consider the payoffs realised by the other player(s) in the market. That is, apart from experiencing payoffs itself, a firm now also reasons about the actions and payoffs it observes of the other firm(s), and reinforces those actions as if tried by itself. Notice that the extended reinforcement learning rules can be seen as a generalisation of the imitate-the-best rule as they consider not only the most recent period but also earlier payoffs. The extended-average-reinforcement-learning rule is based on average payoffs.

14 *AVGXRLD* extended average reinforcement learning with discounting

The extended-average-reinforcement-learning-with-discounting rule is doing the same, but this time earlier payoff observations are discounted (as above) every period. Notice that in case of extreme discounting we would be basically back at the imitate-the-best rule.⁴

15 *CUMXRL* extended cumulative reinforcement learning

The extended-cumulative-reinforcement-learning rule is similar to the extended-average-reinforcement-learning rule, but now based on cumulative payoffs.

16 *CUMXRLD* extended cumulative reinforcement learning with discounting

The extended-cumulative-reinforcement-learning-with-discounting rule adds the discounting of earlier payoff observations (as above) to the extended-cumulative-reinforcement-learning rule.

17 *HILL* hill climbing

The hill-climbing rule works as follows. If a firm had increased its output in the previous period, and its profits went up (down), then the next round its target is the latest output level plus (minus) one.⁵ And the opposite if it had decreased its output in the previous period. Obviously this rule cannot be applied in the first two periods. Also, whenever a firm's output or profits had remained unchanged in the previous two periods the rule does not apply (as no gradient can be observed).

⁴ The only difference being that initial strengths are zero.

⁵ In the case where we only look for the direction of a player's output decision, we also consider a generalisation of this rule, allowing other step sizes as well, without significantly changing the results.

Appendix D. Pseudo-code of the Behavioural Model

Table D1 presents the pseudo-code of the behavioural model that we used for the computational analysis in Section 5. With local experimentation, the parameters are those defining the support on which the output experimentation takes place (X_1 and X_2). In principle they vary in synchronisation from 1 to 25 in steps of 1, but where necessary they are adjusted separately such that the new output cannot exceed the lower and upper output limits. With global experimentation, the parameter is the probability with which a player experiments, using the entire output range. This probability p_{tremble} ranges from 0.04 to 1.00 in steps of 0.04. Common to both types of experimentation, we vary the probability of imitation (p_{imit}), considering the values 0.20, 0.50, and 1.00. Notice that as the probability of imitation on the one hand, and either the probability of trembling or the support of the local experimentation on the other hand vary, imitation becomes either more or less prevalent in this behavioural model.

Table D1
Pseudo-code Computational Analysis

```

for all players do
begin
  with probability  $p_{\text{inert}}$  do stay put
  else
  begin
    with probability  $p_{\text{imit}}$  do
    begin
      planned output:=imitation of most successful;
      if experimentation=local then new output:=planned output +  $x$  {with  $x \sim U(-X_1, X_2)$ }
      else if experimentation=global then
      begin
        with probability  $p_{\text{tremble}}$  do new output:= $x$  {with  $x \sim U(\text{min\_prod}, \text{max\_prod})$ }
        else new output:=planned output;
      end;
    end
  end
  else
  begin
    planned output:=previous output;
    if experimentation=local then new output:=planned output +  $x$  {with  $x \sim U(-X_1, X_2)$ }
    else if experimentation=global then
    begin
      with probability  $p_{\text{tremble}}$  do new output:= $x$  {with  $x \sim U(\text{min\_prod}, \text{max\_prod})$ }
      else new output:=planned output;
    end;
  end;
end;
end;
end;

```

References

- Axelrod, R. (1984). *The Evolution of Cooperation*, New York: Basic Books.
- Cason, T. and Friedman, D. (2000). 'Buyer search and price dispersion: a laboratory study', mimeo.
- Fouraker, L.E. and Siegel, S. (1963). *Bargaining Behavior*, New York: McGraw-Hill.
- Huck, S., Normann, H.-T. and Oechssler, J. (1999). 'Learning in Cournot oligopoly - an experiment', *ECONOMIC JOURNAL*, vol. 109, pp. C80-95.
- Offerman, T., Potters, J. and Sonnemans, J. (1997). 'Imitation and belief learning in an oligopoly experiment', mimeo, University of Amsterdam, Department of Economics.

- Palomino, F. (1995). 'On the survival of strictly dominated strategies in large populations', *Institute for Economic Analysis* WP 308.95, Universitat Autònoma de Barcelona.
- Schaffer, M.E. (1989). 'Are profit maximizers the best survivors?', *Journal of Economic Behavior and Organization*, vol. 12, pp. 29–45.
- Selten, R. and Stoecker, R. (1986). 'End behavior in sequences of finite prisoner's dilemma supergames. a learning theory approach', *Journal of Economic Behavior and Organization*, vol. 7, pp. 47–70.