The Economic Journal, **109** (*March*), C80–C95. © Royal Economic Society 1999. Published by Blackwell Publishers, 108 Cowley Road, Oxford OX4 1JF, UK and 350 Main Street, Malden, MA 02148, USA.

LEARNING IN COURNOT OLIGOPOLY – AN EXPERIMENT*

Steffen Huck, Hans-Theo Normann and Jörg Oechssler

This experiment was designed to test various learning theories in the context of a Cournot oligopoly. We derive theoretical predictions for the learning theories and test these predictions by varying the information given to subjects. The results show that some subjects imitate successful behaviour if they have the necessary information, and if they imitate, markets are more competitive. Other subjects follow a best reply process. On the aggregate level we find that more information about demand and cost conditions yields less competitive behaviour, while more information about the quantities and profits of other firms yields more competitive behaviour.

The Cournot oligopoly model is one of the most widely used concepts in applied industrial organisation. While it is unlikely that inexperienced players would immediately coordinate on an equilibrium, there is a general intuition that over time players would learn to play according to the Cournot–Nash equilibrium. This dynamic story has a long tradition going back to Cournot who already suggested what is now known as the best reply dynamic. According to the best reply dynamic players adjust their quantities simultaneously by choosing best replies against other players' previous outputs. It is easy to see that this dynamic converges to a Nash equilibrium whenever it converges.

There are, however, dynamics which converge to different outcomes even in a simple Cournot oligopoly with a unique Nash equilibrium. Of particular interest is the imitation dynamic suggested by Vega-Redondo (1997). According to this dynamic players 'imitate the best', i.e. they choose the strategy of the player who had the highest profit last period. Vega-Redondo shows that this dynamic converges to the competitive outcome where price equals marginal cost.

While it is generally known that best reply dynamics do not converge in oligopolies with a linear setup and three or more firms (Theocharis, 1960), we show that a best reply process with inertia does converge to the Nash equilibrium. Together with Vega-Redondo's imitation result we have thus two theories making distinct predictions for convergence, which allows to test them experimentally. While we focus on best reply learning and imitation, we also consider alternative learning approaches like directional learning (Selten and Buchta, 1998), trial and error learning and 'imitate the average'.

We attempt to test these theories experimentally by reproducing as closely as possible the conditions assumed in these processes. For example, in order to match the theoretical setup, which requires inertia in the adjustment of

* We thank Werner Güth, Ulrich Kamecke, Manfred Königstein, Stan Reynolds, Susanne Wilpers, seminar audiences at the Stockholm School of Economics, the universities of Munich and Stockholm, and participants of the 1997 THEBORA workshop in Bonn, the 1998 meetings of the ESA and of the Royal Economic Society for very helpful comments. Two referees contributed valuable suggestions. Financial support by the DFG through SFB 373 is gratefully acknowledged.

strategies, we introduce a randomisation device which determines whether players can change their quantities from the previous period.

To differentiate between the different learning theories we vary the information provided to subjects. For example, if subjects follow the imitation strategy, all information they need consists of the quantities and profits of all players. For a best reply process to work subjects need to know the demand and cost conditions in addition to the *total* quantity of the other firms last period, but they do not need to know the *individual* quantities of other players.

Surprisingly, given the standard use of the Cournot model in Industrial Organisation, there are relatively few experimental studies on oligopoly with three or more quantity setting firms. Previous experiments found average quantities that lie between Nash and collusive outcomes for duopoly experiments and around the Nash outcome for experiments with three or more firms. Fouraker and Siegel (1963) conducted a number of experiments of which their triopoly experiments are most closely related to our study. More recently, Rassenti *et al.* (1996) and Offerman *et al.* (1997) ran several similar Cournot experiments. We will discuss these experiments in Section 3.

1. Experimental Design

In a series of computerised[1] experiments we studied a homogeneous multi-period Cournot market with linear demand and cost. There were four symmetric firms in each market. Quantities could be chosen from a finite grid between 0 and 100 with 0.01 as the smallest step. The demand side of the market was modelled with the computer buying all supplied units according to the inverse demand function

$$p^t = \max\{100 - Q^t, 0\}, \tag{1}$$

with $Q^t = \sum_{i=1}^{4} q_i^t$ denoting total quantity in period t. The cost function for each seller was simply $C(q_i^t) = q_i^t$. Hence, profits were $\pi_i^t = (p^t - 1)q_i^t$. The number of periods was 40 in all sessions and this was commonly known. We chose 40 periods as a compromise. On the one hand there is a need for a relatively long time horizon as some learning processes may take quite some time to converge (if at all). On the other hand there is the danger that if there are too many periods, subjects might get bored and take nonsensical decisions only to make something happen.

For theoretical reasons we introduced some inertia.[2] After round one chance moves, which were independent across individuals, determined in each period whether a subject was allowed to revise his quantity. This was done by a 'one–armed bandit' which appeared on the screen showing three equiprobable numbers '0', '1', and '2'. If '0' occurred, no adjustment was allowed. Hence, the probability for allowing revision was 2/3.

[1] We thank Abbink and Sadrieh (1995) for letting us use their software toolbox 'RatImage'.

[2] Best reply dynamics would not converge without inertia and Vega-Redondo's (1997) imitation dynamic also assumes inertia. However, we have also run sessions without inertia, and behaviour was not significantly different.

There were five treatments which differed by the information provided to subjects (the design of treatments is summarised in Table 1). In treatment *BEST* subjects possessed all essential information about the market, i.e. they were informed about the symmetric demand and cost functions in plain words. Furthermore, the software was equipped with a 'profit calculator', which served two functions. A subject could enter some arbitrary 'total quantity of other firms'. Then he could either enter some amount as his own quantity in which case the calculator informed him about the resulting price and his resulting personal profit. Or, he could press a 'Max'–button in which case he was informed about the quantity which would yield him the highest payoff given the hypothetical total amount of others.[3] The usage of the profit calculator was recorded.

After each market period subjects were informed about the total quantity the others had actually supplied, about the resulting price and their personal profits. Additionally, they were reminded of their own quantity. When deciding in the next period this information remained present on the screen. Results of earlier periods were, however, not available, but subjects were allowed to take notes and a few did.

Treatment *FULL* was essentially the same as *BEST*, with the important difference that subjects were additionally informed after each period about *individual* quantities and profits. This information also remained present on the screen while subjects decided in the next period.

In treatment *NOIN* subjects did not know anything about the demand and cost conditions in the market nor did the instructions explicitly state these would remain constant over time. All they knew was that they would act on a market with four sellers and that their decisions represented quantities. Subjects were informed after each period only about the profits they made with the quantity they had chosen.

As in *NOIN*, subjects in treatment *IMIT* did not know anything about demand or cost conditions. They were, however, additionally informed at the end of each period about quantities and profits of the other three sellers.

The fifth treatment, *IMIT+*, was like *IMIT* but in the instructions subjects were given the following additional information about the market: 'The price can

Table 1
Information in Treatments

Information about others	Market information complete	partial	absent
no	*BEST*		*NOIN*
yes	*FULL*	*IMIT+*	*IMIT*

[3] The profit calculator provides essentially the same information as the usually used payoff tables. Furthermore, it helps to avoid a possible bias due to limited computational abilities of subjects.

be between $100T$ and $0T$. The more is sold on the market in total, the *lower* is the price one obtains per unit. Your profit in a given round is then your revenues minus your production costs. Market conditions are constant for all periods and the same for each firm.'

The experiments were conducted in April and May 1997 in the computer laboratory of the economics department of Humboldt University. All subjects were recruited via posters from all over the campus. Almost half of the subjects studied fields other than economics or business and had no training in economics at all. Among the economics and business students almost none had any prior knowledge in oligopoly theory.

In each session eight subjects participated. Subjects were randomly allocated to computer terminals in the laboratory such that they could not infer with whom they would interact in a group of four. For each treatment we had six groups of subjects – making a total of 120 subjects who participated in the experiments.

Subjects were paid according to their total profits. Profits where denominated in 'Taler', the exchange rate for German Marks (500:1) was known. Additionally, subjects earned a fixed payoff of Taler 150 each round. This ensured that no losses could be made. Since we expected the Walrasian output (in which profits are zero) as a possible outcome in some treatments, we wanted to make sure – besides avoiding the usual bankruptcy problems – that subjects would not be frustrated by low or negative payoffs. The average payoff was about DM 25.[4] Experiments lasted between 45 (*NOIN*) and 90 (*FULL*) minutes including instruction time.

Instructions were written on paper and distributed in the beginning of each session.[5] After the instructions were read, we conducted one trial round in which the different windows of the computer screen were introduced and could be tested. When subjects were familiar with both, the rules and the handling of the computer programme, we started the first round.

2. Theoretical Predictions

Consider again four firms $I = \{1, \ldots, 4\}$ which play repeatedly an oligopoly game with quantity setting. For the demand function (1) and constant marginal cost of 1, the unique *Cournot Nash* equilibrium of the stage game is given by $q_i^N = (100 - 1)/5 = 19.8,\ i \in I$, yielding a price of $p^N = 20.8$.[6] Of interest is also the symmetric *Walrasian* (or competitive) outcome where price equals marginal cost, $p^W = 1$, $q_i^W = (100 - 1)/4 = 24.75,\ i \in I$. The symmetric *collu-*

[4] It might be interesting that economics and business students did marginally worse than other students, though the difference is not significant (24.71 DM for economics and business students vs. 25.03 DM for other students).

[5] Instructions and screen shots are available in the working paper version of this paper Huck *et al.* (1998*a*) or on http://www.wiwi.hu-berlin.de/~oechsler.

[6] Strictly speaking, due to the inertia the subgame perfect equilibrium of the dynamic game is not the repetition of the stage game Nash equilibrium. Instead, there is a slight tendency towards Stackelberg behaviour (calculations are available in the working paper version, 1998*a*). However, play in the experiment was never even close to this dynamic solution.

sive or joint profit maximising outcome would be at $q_i^C = 12.375,\ i \in I$, with a corresponding price of $p^C = 50.5$.

We assume that firms exhibit *inertia* in the sense that each period each firm may revise its strategy only with (independent) probability $\theta = 2/3$. In line with the experimental setup we require that outputs must be chosen from a finite grid $\Gamma := \{0, \delta, 2\delta, \ldots, v\delta\}$, for arbitrary $\delta > 0$, and some $v \in \mathbb{N}$ large enough. We assume that $q^W,\ q^N \in \Gamma$.

2.1. *Best Reply Dynamic*

First, we consider the best reply dynamic. Players myopically choose every period a best reply to the other players' total output from *last* period. Let $\Pi(q_i^t, q_{-i}^t)$ denote firm i's profit in period t given its quantity q_i^t and the total quantity of its opponents q_{-i}^t.

Assumption 1 (myopic best reply): If a firm has the opportunity to revise its strategy, it chooses a best reply against the profile of the other firms' previous output, i.e. a strategy from the set

$$BR_i^{t-1} := \{q \in \Gamma : \forall q' \in \Gamma,\ \Pi(q, q_{-i}^{t-1}) \geqslant \Pi(q', q_{-i}^{t-1})\},$$

according to some probability distribution with full support.

Note the informational requirements to play myopic best replies. (1) One needs to know the demand and cost functions. (2) One needs to know q_{-i}^{t-1}, i.e. last period's total output of the remaining players. And (3) one needs to know how to calculate a best reply. All three requirements are met in the experimental treatments *Best* and *Full*. In the three remaining treatments players do not have sufficient knowledge to calculate best replies.

It is well known (see Theocharis, 1960) that given a linear setup the best reply dynamic is unstable for oligopolies with four firms. However, in the Appendix we prove the following proposition.

Proposition 1 *The best reply dynamic with inertia converges globally in finite time to the static Nash equilibrium.*

The question of course arises how long the process will take to converge. To answer this question we have conducted a series of simulations for $\theta = 2/3$. For randomised starting values (uniformly between 0 and 100) the process converges to a 1%–neighbourhood of the Nash equilibrium on average in 26.5 periods.

2.2. *Imitate the Best*

An alternative learning procedure is to simply imitate the quantity of the player with the highest profit last period.

Assumption 2 (imitate the best) If a firm has the opportunity to revise its strategy, it chooses one of those strategies which received the highest payoff last period according to some probability distribution with full support. Furthermore, every period each firm 'mutates' (makes a mistake) with independent probability $\varepsilon > 0$ and chooses an arbitrary $q \in \Gamma$ (all q are chosen with some strictly positive probability).

The information required for 'imitate the best' is a list of last period's quantities and profits of all firms. No information about market or cost conditions is needed. Though, of course, the rule is more sensible if one knows that market conditions are symmetric and time invariant. The list of quantities and profits was provided in treatments *FULL*, *IMIT* and *IMIT+*. Additionally, players in treatments *FULL* and *IMIT+* were explicitly told that the market conditions are constant and symmetric.

The following proposition was proven by Vega-Redondo (1997).

PROPOSITION 2 *In the long run for $\varepsilon \to 0$ the Walrasian outcome will be observed almost always if players 'imitate the best'.*

The intuition for this result is straightforward. Whenever price is higher than marginal cost, the firm with the highest quantity makes the largest profit and vice versa if profits are negative. Hence, as long as profits are positive, the largest output gets imitated which drives up total output until price equals marginal cost. Note that this also explains why the Cournot–Nash equilibrium is not a stable rest point of 'imitate the best'. If one firm deviates to a higher quantity, profits of all firms decrease but profit of the deviator decreases by less.

2.3. *Other Learning Processes*

One learning theory that can be tested within our setup is Selten's *learning direction theory* (see Selten and Buchta, 1998). The theory assumes that players have a model which allows them to conclude in which direction better actions can be found. In our context it seems reasonable to suppose that if subjects have the information as in treatments *BEST* and *FULL*, better actions should be interpreted as lying in the direction of the best reply $r(q_{-i}^{t-1})$. In its weak form the theory leads to qualitative predictions of the form: if $q_i^{t-1} \gtrless r(q_{-i}^{t-1})$ then $q_i^t \lesseqqgtr q_i^{t-1}$, i.e., players do not move in a direction away from the best reply. In its stronger form (replacing $\lesseqqgtr$ by $\gtrless$), the process has a unique restpoint which is the Cournot–Nash equilibrium.

An alternative interpretation of directional learning and one which works even if subjects have as little information as in treatment *NOIN* is a process which we call *trial and error learning*. It simply says that a subject would not repeat a mistake, i.e. if profits last period have decreased due to an increase in quantity, then one would not increase quantity again. On the other hand, if profits had increased following an increase in quantity, one would not decrease quantity next period. Interestingly, this process yields collusion. In Huck *et al.*, (1998*b*) we show that the following process with some added noise converges globally to a neighbourhood of the collusive outcome.

$$q_i^t = q_i^{t-1} + \operatorname{sign}(q_i^{t-1} - q_i^{t-2}) \times \operatorname{sign}(\pi_i^{t-1} - \pi_i^{t-2}) \times \delta. \qquad (2)$$

The intuition is that rather quickly firms begin to move in step, i.e. they increase or decrease quantities together most of the time. Roughly speaking, when all firms lower their quantity simultaneously, profits increase as long as Q^t is above the collusive outcome, and so firms will lower their quantity further (and vice versa when Q^t is below Q^C). However, the firm with the lowest (highest) quantity will be the first to reverse the downward (upward) movement in quantities, which results in convergence across firms as well.

Finally, even though subjects in the *Best* treatment were not able to 'imitate the best', they can still use some form of imitation. In particular, they can *imitate the average* as they know the total quantity of the three other firms. It seems reasonable that subjects who are uncertain about what to do and observe that the average quantity of the other firms deviates from their own quantity, imitate this average quantity – thinking along the line of 'everyone else can't be wrong'. A preference for cautious behaviour and a taste for conformity could be further reasons for imitating the average. If all subjects were to follow this rule, clearly the process is bounded above and below by the highest and lowest initial quantities. Without inertia the process would follow the difference equation $q_i^t = (Q^{t-1} - q_i^{t-1})/3$. The solution

$$q_i^t = \frac{Q^1}{4} + \frac{3q_i^1 - q_{-i}^1}{4}\left(-\frac{1}{3}\right)^t \tag{3}$$

shows that the process converges simply to the average of all starting values. With inertia the process depends on the realisations of the randomisation device and is therefore path dependent.

3. Experimental Results

Table 2 reports average total quantities over the last 35 (*Mean*35) and the last 20 (*Mean*20) periods (averaged over all six groups in a treatment).[7] The first 5 rounds cannot be seen as representative as subjects in treatments *Imit*, *Imit+* and *NoIn* were not given enough information to find reasonable starting values.

As shown in Table 2 behaviour in all treatments was very competitive. Even in *Best* (where we observed the lowest average total quantities) the mean across all sessions is above the Cournot prediction. Five out of six sessions have an average total quantity above Q^N and the average quantity is more than one standard deviation above Q^N. Average total quantities of all sessions in *Full* and in five out of six *NoIn* sessions lie between the Cournot and the Walrasian prediction (Q^W).

The *Imit+* results are remarkably close to the theoretical prediction, which is $Q^W = 99$. One cannot reject the hypothesis that the average total quantities in treatment *Imit+* are drawn from a normal distribution with mean 99.[8] The

[7] In treatment *Imit+* one group had to be considered a clear outlier as one subject chose the maximum quantity in 18 of the last 20 periods. We chose to exclude this group from all further analysis. However, our main results are unchanged even if this group is included.

[8] The appropriate Kolmogorov–Smirnov test shows that rejection is not even possible at a 20% level of significance.

Table 2
Average Total Quantities

Treatment	BEST	FULL	NOIN	IMIT	IMIT+
Mean35	84.32	90.68	102.14	146.70	105.12
	(4.56)	(5.77)	(19.79)	(34.17)	(13.64)
Mean20	82.56	91.60	93.55	138.85	96.43
	(2.48)	(6.48)	(14.73)	(31.62)	(5.35)

Note: Standard deviations in parentheses. The theoretical benchmarks are: $Q^N = 79.2$, $Q^W = 99$, $Q^C = 49.5$

median of the distribution (98.8) almost exactly hits Q^W. In treatment *IMIT* behaviour was far in excess of Q^W. In fact, subjects made losses in most periods. Recall that the only difference between *IMIT+* and *IMIT* was the information given in the instructions that inverse demand is non–stochastic and decreasing in quantity. That such small differences in the framing of an experiment make such a difference is interesting by itself.

It is also noteworthy that in none of the groups there was any successful attempt to establish collusion. Only occasionally total quantities fell below Q^N. The collusive price of 50.5 was exceeded in just 12 out of 1,200 observations. There have been few individual attempts to establish cooperation by supplying quantities close to 12. But this was always exploited by other firms so that the cooperators eventually gave up. Thus we have

Result 1 In all treatments average behaviour was more competitive than the Cournot prediction. In *IMIT+* average behaviour matches almost perfectly the theoretical prediction, i.e. the Walrasian outcome. There were no successful attempts of collusion in any treatment.

We have also computed means for periods 21 to 37 and compared them with the respective means for periods 38 to 40 to check for possible end game effects. End game effects could be present since we announced the length of the game in advance. It turned out, however, that in none of the treatments there were any significant differences in average quantities between the last 3 rounds and the 17 rounds preceding them.

3.1. *Information and Competition*

Two main results about the relationship between information and quantities can be obtained from analysing the data. Recall, that the five treatments can be ordered in a partly nested way as displayed in Fig. 1 in terms of the information available to subjects.

To measure the effect of additional *information about the market* we test *NOIN* vs. *BEST*, *IMIT* vs. *IMIT+*, and *IMIT+* vs. *FULL*. Taking each group as a single observation we applied the Mann–Whitney–U statistic (see Siegel and Castellan, 1988) to test for differences in means (based on the last 35 periods, one–

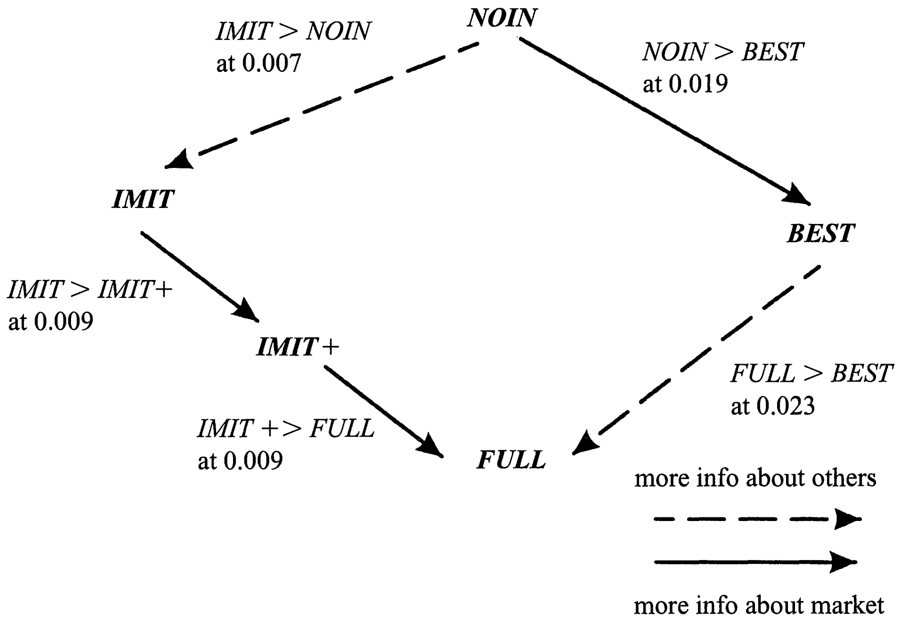

Fig. 1. *Significance levels of Mann–Whitney–U tests for differences in means based on last* 35 *periods, one–tailed.*

tailed).[9] In each case total quantities are significantly lower in the treatment with more information about the market. The significance levels are 0.019 for *Noin* vs. *Best*, 0.009 for *Imit* vs. *Imit+*, and 0.009 for *Imit+* vs. *Full*. Thus, increasing the information about the market *decreases* total quantities.

To measure the effect of additional information about individual quantities and profits we test *Best* vs. *Noin* and *Noin* vs. *Imit*. It turns out that total quantities are significantly higher in the presence of information about others. The significance levels are 0.023 for *Best* vs. *Full* and 0.007 for *Noin* vs. *Imit*. Thus, providing additional information about individual quantities and profits *increases* total quantities.

The latter result is especially interesting with respect to the theoretical predictions about imitation. We have shown in Section 2 that imitation yields more competition than behaviour based on myopic best replies or other rules discussed. The data reveal that if the information, which is necessary to imitate successful behaviour, is available, competition indeed becomes more intense. Quantities in *Full* are significantly higher than those in *Best*, which indicates that imitation plays the predicted role. On the other hand, quantities in *Full* are significantly lower than in *Imit+* which shows that at least some individuals follow other considerations.

The fact that competition is more intense when firms know more about the individual quantities and profits of their rivals provokes traditional views on

[9] Results are essentially the same for the last 20 periods.

competition policy.[10] Also anti–trust authorities often allow trade associations to publish only aggregate industry data. Our result, however, indicates that it might increase competition if firms are informed about their rivals' quantities and profits.

The main observations can be summarised as follows.

Result 2 More information about the market yields less competitive outcomes.

Result 3 More information about behaviour and profits of others yields more competitive outcomes.

Result 3 confirms a result obtained earlier by Fouraker and Siegel (1963). They find that quantities in their 'complete information' treatment, which is comparable to our *FULL* treatment, are higher than in their 'incomplete' information treatment, which is roughly comparable to our *BEST* treatment.

Rassenti *et al.* (1996) and Offerman *et al.* (1997) ran several oligopoly experiments with five and three firms, respectively. Offerman *et al.*'s main treatment conditions are again similar to our *BEST* and *FULL* treatments. While they do observe higher quantities in full information treatments, they do not observe significant differences in average quantities. Rassenti *et al.* also confirm that markets with four or five firms are already too competitive to make collusion sustainable.

3.2. *Individual Learning Behaviour*

While the analysis of group level behaviour gives some insight as to the relative performance of the different learning theories, a closer look at individual behaviour seems warranted. For this we have estimated the following equation

$$q_i^t - q_i^{t-1} = \beta_0 + \beta_1(r_i^{t-1} - q_i^{t-1}) + \beta_2(ib^{t-1} - q_i^{t-1}) + \beta_3(i_i^{t-1} - q_i^{t-1}) \qquad (4)$$

where r_i^{t-1} denotes subject *i*'s best reply (i.e. reaction function) given the other firms' quantities in $t-1$; ib^{t-1} stands for 'imitate the best' and denotes the quantity of the firm which had the highest profit in period $t-1$; finally i_i^{t-1} denotes the average quantity of the other firms' output in $t-1$. Note, that a subject who strictly played a myopic best reply every period would have $\beta_1 = 1$ and $\beta_k = 0$, $k \neq 1$.[11] Analogously, for someone who follows the rule 'imitate the best' or 'imitate the average'.

The choice of r_i^{t-1} and ib^{t-1} as variables does not need an explanation given our emphasis on those two learning rules. We chose to include additionally i_i^{t-1} for a simple reason. In treatment *BEST* subjects are not able to observe the

[10] For example, Stigler (1964) argues that collusion is easier to sustain if price cuts of rivals can be detected immediately.

[11] Thus, one advantage of using the differences between the variables and q_i^{t-1} rather than the absolute values in the regression is that the coefficients have a nice interpretation. The other advantage is that it avoids problems of serial correlation.

quantities or profits of individual other firms. Hence, they cannot imitate the best firm. They may, however, imitate the average. In fact, i_i^{t-1} turns out to be an important variable. In the regressions we have included only variables which were observable to the subjects.[12]

Goldfeld–Quandt tests indicated that variances were significantly lower in later rounds for all treatments except *FULL*, which could have been expected if learning behaviour converges. To correct for heteroscedasticity we have therefore estimated (4) with weighted least squares (WLS) using t^{α} as a weight for all observations. The α's were chosen so as to maximise the log-likelihood function. Given the possibility that subjects changed their learning behaviour during the experiment we used Chow tests to determine possible structural breaks. Only in *FULL* we found a significant break. For *FULL* we have therefore estimated (4) with slope and intercept dummies for periods 1–10, 11–20, 21–30. To account for individual differences in learning behaviour we added subject intercept and slope dummies.

Table 3 gives the result from estimating (4) with pooled data of all subjects in each treatment. The coefficients yield an indication about the relative importance of the explanatory variables. All coefficients have the expected sign. However, subjects adjust on average only incompletely as coefficients are far away from 1. In all treatments (except *IMIT*) 'imitate the average' seems to play a substantial role. In *BEST* and *FULL* the best reply variable is the most important factor. However, in *FULL* imitation of both sorts becomes more important, which is probably responsible for the difference in outcomes

Table 3

WLS Regressions

	BEST	*FULL*	*IMIT*	*IMIT+*
β_1	0.373**	0.260**		
	(0.044)	(0.045)	–	–
β_2		0.078*	0.381*	0.414**
	–	(0.038)	(0.162)	(0.044)
β_3	0.289**	0.194**	0.061	0.263**
	(0.052)	(0.034)	(0.234)	(0.052)
β_0	0.731	1.33**	9.29**	0.764
	(0.387)	(0.442)	(1.48)	(0.522)
α	0.2	0	0.3	0.8
R^2	0.430	0.517	0.505	0.441
DW	2.08	2.17	2.11**	2.22
Obs.	626	643	633	546

Note: **(*) significant at 1% (5%) level. Standard deviations in parentheses. DW = Durbin Watson statistic. α = power of weight of function. Subject dummies for intercepts and slopes were used. Only periods in which subjects were allowed to adjust their quantities were included.

[12] Including *ib* in (4) for *BEST* did not change the results and β_2 was not significant. Including *r* for *IMIT* and *IMIT+* caused collinearity problems. We have therefore imposed the theoretical restriction $\beta_1 = 0$. Since none of the variables was observable in *NOIN*, we did not run the regressions for this treatment.

between those two treatments. The records about the use of the profit calculator have the same tendencies. Given the inertia subjects had the possibility to use the calculator on average in 27 rounds. In more than half the rounds subjects consulted the calculator. But it was used on average in 17.8 rounds in treatment *BEST*, whereas it was used only in 15.4 rounds in *FULL*.

In both, *IMIT* and *IMIT+* the 'imitate the best' variable is the major factor. When looking at other players who receive higher payoffs due to higher quantities, the temptation to match the higher quantities is apparently hard to resist even if own profits are reduced by doing so.

Including other variables in (4) did not prove successful. We have tried three other variables. The first was 'imitate the average if better', i.e. imitate the average only if average profits are higher than own profit. The second was 'imitate the highest quantity'. Both of these variables did not add any explanatory power. Finally, we have included higher order lags of the variables in (4). Even though we did not provide subjects with this kind of information, some were taking notes or might have had a good memory. The use of lagged variables would be evidence for learning theories like fictitious play. However, lagged variables were not significant in any treatment.[13]

Two hypotheses are compatible with the finding of significant coefficients for the imitation terms in our regressions. Either *all* subjects are to *some* extent imitators or *some* subjects primarily imitate and others follow different leaning rules. To decide this question we can look at the individual decisions more closely.

Let

$$z_i^t = \frac{q_i^t - q_i^{t-1}}{a_i^{t-1} - q_i^{t-1}} \tag{5}$$

where a^{t-1} is the point prediction implied by playing myopic best reply (BR), imitate the best (IB), or imitate the average (AV), respectively. Obviously, $z_i^t = 1$ follows in case of perfect adjustment, while $z_i^t < 0$ implies a severe qualitative violation.[14] Table 4 shows the overall frequency distribution of the z–values. Note first that generally perfect hits ($z = 1$) are quite rare. The only exception is *IB*, which produces perfect hits in up to 15.5% of relevant cases. Thus, it seems that 'imitate the best' is the one learning rule which, when applied, is applied most precisely. All three theories predict the direction of change ($z \geqslant 0$) accurately in most cases.

Table 5 shows how many subjects score positive z–values in at least 70% of their decisions and how many score nearly perfect hits ($0.8 \leqslant z \leqslant 1.2$) in more than 20% of their decisions. For example, 2/3 of the subjects in *BEST* adust in the direction of *BR* in at least 70% of their decisions and 1/4 do so almost perfectly in 20% of their decisions. In *IMIT+* (and to some lesser extent in *IMIT*) there seems to be a substantial number of committed imitators. More than half

[13] This is in contrast to a recent study by Boylan and El-Gamal (1993) who find that fictitious play explains their data better than a best reply process.

[14] In case of $a_i^{t-1} - q_i^{t-1} = 0$ and $q_i^t - q_i^{t-1} \neq 0$ we set $z_i^t < 0$. If also $q_i^t - q_i^{t-1} = 0$, then we set $z_i^t = 1$.

Table 4

Hit Ratios

		$z<0$	$0 \leqslant z < 0.8$	$0.8 \leqslant z < 1$	$z=1$	$1<z < 1.2$	$z>1.2$
FULL	*BR*	32.2	50.7	3.8	1.4	3.3	8.6
	AV	30.0	44.8	4.0	0.6	4.6	15.9
	IB	38.0	42.5	2.7	7.9	2.2	6.7
BEST	*BR*	26.4	43.0	5.4	5.4	3.1	16.6
	AV	35.7	44.6	4.1	1.0	2.3	12.3
IMIT+	*AV*	31.1	43.6	2.7	1.3	2.8	18.4
	IB	36.2	31.8	4.4	15.5	2.3	9.8
IMIT	*AV*	34.0	39.6	3.2	1.1	2.6	19.4
	IB	39.8	36.5	2.7	9.4	2.6	9.0

Note: Only rounds in which subjects were allowed to adjust their quantities are included.

Table 5

Number of subjects with	$z \geqslant 0$ in 70% of cases			$0.8 \leqslant z \leqslant 1.2$ in 20% of cases			
	BR	*AV*	*IB*	*BR*	*AV*	*IB*	Total
FULL	9	12	7	2	1	4	24
BEST	16	6	–	6	1	–	24
IMIT+	–	9(11)	5(6)	–	1	11(12)	18(24)
IMIT	–	9	8	–	0	5	24

Note: Only rounds in which subjects were allowed to adjust their quantities are included. For *IMIT+* the number in brackets also includes subjects from the otherwise excluded outlier group.

of the subjects follow *IB* almost perfectly in 20% of their decisions. Again, we see that this rule is applied with the highest consistency if applied at all.

Result 4 If subjects have the necessary information to play best replies, most do so, though adjustment to the best reply is almost always incomplete. If subjects additionally have the necessary information to 'imitate the best', at least a few subjects become pure imitators.

We have also tested the alternative learning theories mentioned in Section 2.3. Table 4 shows that in all treatments in which this was possible subjects were partly influenced by the average quantity of the other players. Directional learning interpreted as a movement towards best replies ($z \geqslant 0$) can also be evaluated by looking at Table 4. Three observations are apparent: (1) Directional learning does better in treatment *BEST* than in *FULL*. (2) Overall, directional learning predicts correctly in more than 70% of the time. And (3) (not shown in the Table) it performs better for upward adjustment than for downward adjustments.

Finally, trial and error learning is a type of learning which is applicable even in treatment *NOIN*. Table 6 shows the theoretical predictions and the number of times those predictions were correct. In its weak form the hypotheses are correct for all 5 treatments in about 80% of the cases to which they applied. Note, however, that there is a theoretical problem. If subjects had really been playing according to trail and error, they would have ended up near the collusive outcome, which is the theoretical prediction. But they converged either to Nash or to even higher quantities. The theoretical prediction, however, holds only if *all* subjects behave according to trial and error. If just one subject in each group deviates, very different outcomes can result. Furthermore, it is possible that modifications of the trail and error process (e.g. varying step sizes) yield different results.

Result 5 All examined alternative learning theories have some predictive power, but trial and error learning performs slightly better than directional learning and 'imitate the average'.

4. Summary

In a series of experiments we investigated multi–period Cournot markets under various information conditions. On an aggregate level the two main results are that providing more information about quantities and profits of the competing firms increases competition whereas additional information about the market structure decreases competition. The former result is explained by some individuals' propensity to imitate successful strategies, while the latter is based on individuals' ability and willingness to adjust behaviour towards best replies. Competition, however, is always strong enough to frustrate any attempts to collude.

The analysis of the individual data showed that none of the theoretical learning processes which are discussed in Section 2 can on its own explain the observed behaviour. Focusing on myopic best reply dynamics and imitation dynamics we find, however, that both adjustment rules play a role for subjects'

Table 6

Trial and Error Learning

Predictions		$\Delta q^{t-1} > 0$	$\Delta q^{t-1} = 0$		$\Delta q^{t-1} < 0$
	$\Delta\pi^{t-1} > 0$	$\Delta q^{t} \geqslant 0$	–		$\Delta q^{t} \leqslant 0$
	$\Delta\pi^{t-1} = 0$	–	–		–
	$\Delta\pi^{t-1} < 0$	$\Delta q^{t} \leqslant 0$	–		$\Delta q^{t} \geqslant 0$
	BEST	*FULL*	*NOIN*	*IMIT*	*IMIT+*
No. of decisions	626	643	633	653	546
Correct predictions	511	517	500	530	433
Success in %	81.63	80.40	78.99	81.16	79.30

decisions provided that they possess the necessary information to apply these rules. When subjects know the true market structure, their quantity adjustments depend significantly on the myopic best reply. When subjects know individual profits, their adjustments go significantly towards the most successful strategy of the previous period.

Concerning alternative learning hypotheses we find that a simple learning rule, which we called 'trial-and-error learning' and which demands that subjects do not make the same mistake twice, performs quite well. Furthermore, the data indicate that whenever subjects can calculate the average quantities of their competitors, their adjustments also depend on these which hints at a taste for conformity.

Overall we find that learning plays an important role in our experiments. However, learning takes place in a delicate manner and is highly information sensitive.[15] No examined learning theory is rich enough to account for all these factors.

Humboldt University, Berlin

Appendix

We can specify BR_i (suppressing the time index t) in more detail by noting that $\Pi(q_i, q_{-i})$ is symmetric around $r(q_{-i})$, where $r(q_{-i}) := \arg\max_{q_i \in \mathbb{R}_+} \Pi(q_i, q_{-i})$ would be i's reaction function if he could choose q_i continuously. Symmetry follows since $\Pi[r(q_{-i}) + \Delta, q_{-i}] = (99 - q_{-i})^2/4 - \Delta^2$. Since the slope of $r(q_{-i})$ is $-1/2$ and $q^N \in \Gamma$, the grid points closest to $r(q_{-i})$ are either $r(q_{-i})$ itself or both, $r(q_{-i}) + \delta/2$ and $r(q_{-i}) - \delta/2$. That is, BR_i is either the singleton $\{r(q_{-i})\}$ or the set $\{r(q_{-i}) + \delta/2, r(q_{-i}) - \delta/2\}$.

Proof of Proposition 1 The best reply process defined by Assumption 1 yields a finite Markov process on the state space Γ^4 with a unique absorbing state $\omega^N = (q_1^N, q_2^N, q_3^N, q_4^N)$. To prove convergence it suffices to show that all remaining states are transient, i.e. the probability of the process returning to such a state is strictly less than one.

From any state with some $q_i = 0$, there exists a transition to another state with $q_i > 0$, $\forall i$. Now consider the game with the restriction that $q_i > 0$, $\forall i$. Let $P(q_1, q_2, q_3, q_4) := (p - 1)\prod_{i=1}^4 q_i$, where p is defined as in (1). Since $P(q_1, q_2, q_3, q_4)$ is an ordinary potential function, by Lemma 2.3 of Monderer and Shapley (1996) every improvement path is finite. An improvement path is a sequence $(\omega^0, \omega^1, \omega^2, \ldots)$, such that for each $t \geqslant 1$ there is a unique player i who by choosing quantity q_i^t strictly improves his payoff, i.e. $\omega^t = (q_i^t, q_{-i}^{t-1})$ and $\Pi_i(q_i^t, q_{-i}^{t-1}) > \Pi_i(q_i^{t-1}, q_{-i}^{t-1})$. Note that the best reply process gives rise to an improvement path if in each period exactly one player gets to adjust his strategy, which occurs with positive probability. Along an improvement path

[15] Note that this is evidence against learning theories of the reinforcement type (see e.g. Roth and Erev, 1995) since those are insensitive to information conditions.

the value of the potential function strictly increases. Hence, with positive probability the process moves along an improvement path to the equilibrium, which is an absorbing state. Thus, all states other than ω^N are transient and the Proposition follows. □

References

Abbink, K. and Sadrieh, A. (1995). 'RatImage—research assistance toolbox for computer-aided human behaviour experiments.' SFB 303 Discussion Paper B-325, University of Bonn.

Boylan, R. and El-Gamal, M. (1993). 'Fictitious play: a statistical study of multiple economic experiments.' *Games and Economic Behavior*, vol. 5, (March), pp. 205–22.

Fouraker, L. and Siegel, S. (1963). *Bargaining Behavior.* New York: McGraw–Hill.

Huck, S., Normann, H. and Oechssler, J. (1998*a*), 'Learning in Cournot oligopoly—an experiment.' Discussion Paper No. 119, Humboldt University, Berlin.

Huck, S., Normann, H. and Oechssler, J. (1998*b*), 'Through trial and error to collusion.' Mimeo, Humboldt University, Berlin.

Monderer, D. and Shapley, L. (1996). 'Potential games.' *Games and Economic Behavior*, vol. 14, (May), pp. 124–43.

Offerman, T., Potters, J. and Sonnemans, J. (1997). 'Imitation and adaptation in an oligopoly experiment.' Mimeo, Amsterdam University.

Rassenti, S., Reynolds, S., Smith, V. and Szidarovszky, F. (1996). 'Adaptation and convergence of behavior in repeated experimental Cournot games.' Mimeo, University of Arizona.

Roth, A. and Erev, I. (1995). 'Learning in extensive form games: experimental data and simple dynamic models in the intermediate term.' *Games and Economic Behavior*, vol. 8, (January), pp. 164–212.

Selten, R. and Buchta, J. (1998). 'Experimental sealed bid first price auctions with directly observed bid functions.' In (D. Budescu, I. Erev, and R. Zwick, eds.) *Games and Human Behavior: Essays in Honor of Amnon Rapoport*, Mahwak, NJ: Lawrenz Erlaaum Association.

Stigler, G. J. (1964). 'A theory of oligopoly.' *Journal of Political Economy*, vol. 72, (February), pp. 44–61.

Siegel, S. and Castellan, N. J. (1988). *Nonparametric Statistics for the Behavioral Sciences*, second edition, New-York: McGraw-Hill.

Theocharis, R. (1960). 'On the stability of the Cournot solution on the oligopoly problem.' *Review of Economic Studies*, vol. 73, (February), pp. 133–4.

Vega-Redondo, F. (1997). 'The evolution of Walrasian behavior.' *Econometrica*, vol. 65, (March), pp. 375–84.

LINKED CITATIONS

You have printed the following article:

Learning in Cournot Oligopoly-An Experiment
Steffen Huck; Hans-Theo Normann; Jörg Oechssler
The Economic Journal, Vol. 109, No. 454, Conference Papers. (Mar., 1999), pp. C80-C95.
Stable URL:
http://links.jstor.org/sici?sici=0013-0133%28199903%29109%3A454%3CC80%3ALICOE%3E2.0.CO%3B2-B

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

[Footnotes]

[2] **The Evolution of Walrasian Behavior**
Fernando Vega-Redondo
Econometrica, Vol. 65, No. 2. (Mar., 1997), pp. 375-384.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199703%2965%3A2%3C375%3ATEOWB%3E2.0.CO%3B2-4

[10] **A Theory of Oligopoly**
George J. Stigler
The Journal of Political Economy, Vol. 72, No. 1. (Feb., 1964), pp. 44-61.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28196402%2972%3A1%3C44%3AATOO%3E2.0.CO%3B2-3

References

A Theory of Oligopoly
George J. Stigler
The Journal of Political Economy, Vol. 72, No. 1. (Feb., 1964), pp. 44-61.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28196402%2972%3A1%3C44%3AATOO%3E2.0.CO%3B2-3

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

On the Stability of the Cournot Solution on the Oligopoly Problem
R. D. Theocharis
The Review of Economic Studies, Vol. 27, No. 2. (Feb., 1960), pp. 133-134.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28196002%2927%3A2%3C133%3AOTSOTC%3E2.0.CO%3B2-4

The Evolution of Walrasian Behavior
Fernando Vega-Redondo
Econometrica, Vol. 65, No. 2. (Mar., 1997), pp. 375-384.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199703%2965%3A2%3C375%3ATEOWB%3E2.0.CO%3B2-4