Learning in Cournot Oligopoly - An Experiment

Steffen Huck
Department of Economics
Humboldt University, Berlin

Hans–Theo Normann
Department of Economics
Humboldt University, Berlin

Jörg Oechssler*
Department of Economics
Humboldt University, Berlin

July 22, 1997

Abstract

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 behavior 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 behavior, while more information about the quantities and profits of other firms yields more competitive behavior.

JEL– classification numbers: C92, C72, L13.

Very preliminary – comments welcome

*Institut für Wirtschaftstheorie, Spandauer Str. 1, 10178 Berlin, Germany, FAX: +49–30–2093-5619, E–mail: oechsler@wiwi.hu-berlin.de

1 Introduction

The Cournot oligopoly model is one of the most widely used concepts in applied industrial organization. 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 (1838) who already suggested what is now known as the best reply dynamic. According to the best reply dynamic players adjust their quantities simultaneously to the best reply against other players' previous outputs.

Recently, it has been shown that adaptive learning dynamics in a broader class (see e.g. Milgrom and Roberts, 1991) converge to Nash equilibria if they converge. There are, however, dynamics which converge to different outcomes even in a simple Cournot oligopoly with a unique equilibrium. An example 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.

This paper reports on an experiment which was designed to test various learning theories in the context of a Cournot oligopoly. But since those learning theories are generally applicable, the experiment should contribute to the understanding of learning processes in all normal form games. In particular, we are interested in the role of imitation versus learning rules which are based on some form of best replies.

While it is generally known that best reply dynamics do not converge in oligopolies with three or more firms (Theocharis, 1960), we show that a best reply process with inertia (see e.g. Kandori *et al.*, 1993) does converge to the Nash equilibrium. Together with the 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 Stoecker, 1986), fictitious play, reinforcement learning (Roth and Erev, 1995), trial and error learning and 'imitate the average'. The latter two were not much discussed in the literature previously but they were suggested by the data of our experiment.

We attempt to test these theories experimentally by reproducing as closely as possible the conditions assumed in these processes. E.g. in order to match the theoretical setup, which requires inertia in the adjustment of strategies, we introduce a randomization 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 need not to know the *individual* quantities of other players.

Surprisingly given the standard use of the Cournot model in Industrial Organization, 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 (see e.g. Holt, 1985) and around the Nash outcome for experiments with three or more firms. Fouraker and Siegel (1963) conduct a number of experiments of which their tripoly experiments are most closely related to our study. More recently, Rassenti *et al.* (1996) ran several Cournot experiments with five firms. We will discuss these experiments in Section 4.

The paper is organized as follows. In the next section we describe the design of the experiment. We conducted five different treatments, which are partly nested in terms of the information provided to subjects. In Section 3 we derive theoretical predictions for several learning processes. Section 4 contains the experimental results for the aggregate (group) level as well as for the individual level. Section 5 concludes with a summary. The instructions for the experiment are printed in Appendix A, followed by screen shots of the computer program in Appendix B.

2 Experimental design

In a series of computerized[1] experiments we studied a homogenous 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 .01 as the smallest step. The demand side of the market was modelled with the computer buying all supplied units according

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

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

$$\pi_i^t = (p^t - 1)q_i^t. \tag{2}$$

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 we wanted to study learning with time series methods and 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 its 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.

There were five treatments which differed by the information provided to subjects (the design of treatments is summarized 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.[3] 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 she could either enter some amount as her own quantity in which case the calculator informed her about the resulting price and her resulting personal profit. Or, she could press a "Max"–button in which case she was informed about 'the quantity which would yield her

[2] Best reply dynamics would not converge without inertia and Vega–Redondo's (1997) imitation dynamic also assumes inertia. However, in another experiment (Huck, Normann, and Oechssler, 1997) we show that behavior is not significantly different with or without inertia.

[3] Since we recruited many non-economic students as subjects we were careful not to use any formulas or technical terms in the instructions.

Table 1: **Information in treatments**

Information about others	Market information complete	partial	absent
no	Best		NoIn
yes	Full	Imit+	Imit

the highest payoff given the total amount of others'.[4] Additionally, the calculator computed price and profit for this best response.[5] 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 treatments NoIn and Imit 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. In treatment NoIn subjects were informed after each period only about the profits they made with the quantity they had chosen.

In treatment Imit subjects were at the end of each period additionally informed about quantities and profits of the other three sellers. After running this treatment we suspected that subjects had problems understanding the situation as they made losses in almost all periods. We have added therefore another treatment, Imit+, in which some information about the market was given. In the instructions subjects learned that the market was symmetric, that demand and cost would remain constant over time and that

[4]While the profit calculator provides the same information as the usually used payoff tables, it certainly focuses attention on the best reply. By this we tried to give the Cournot equilibrium the best shot. If play deviates from Cournot nevertheless, this only strengthens our results.

[5]In the experiments we did not use the phrase 'best response.'

the price would be inversely related to the total quantity supplied.

The experiments were conducted in April and May 1997 in the computer lab 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 lab 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 as in (2) 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 the usual bankruptcy problems – that subjects would not be frustrated by low or negative payoffs.[6]

The average payoff was about DM 25 which is roughly $14.50. (The payoff for playing the static Cournot equilibrium every period would have been about DM 43). Experiments lasted between 45 (NOIN) and 90 (FULL) minutes including instruction time.

Instructions (see Appendix A) were written on paper and distributed in the beginning of each session. After the instructions were read, we conducted one trial round in which the different windows of the computer screen (see Appendix B) were introduced and could be tested. When subjects were familiar with both, the rules and the handling of the computer program, we started the first round.

3 Theoretical predictions

The experiment described in the previous section was designed to test whether and how subjects learn to play an oligopoly game. In this section we will derive theoretical predictions for several simple boundedly rational rules.

[6]See Holt (1985, p. 317) for the argument that the usual promises in the instructions that one can earn a "considerable amount of money" might bias subjects against zero–profit outcomes.

We are mainly interested in two well studied simple rules, namely the best reply dynamic and imitation dynamics. Those rules yield quite different predictions. Before we define those learning rules we will shortly discuss the static stage game solution and the dynamic solution of the game.

Consider again four firms $I = \{1, .., 4\}$ 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

$$q_i^N = \frac{100-1}{5} = 19.8,\ i \in I,$$

yielding a price of $p^N = 20.8$.

Of interest is also the symmetric *Walrasian* (or competitive) outcome where price equals marginal cost, $p^W = 1$,

$$q_i^W = \frac{100-1}{4} = 24.75,\ i \in I.$$

The *collusive* or joint profit maximizing outcome would be at

$$q_i^C = 12.375,\ i \in I,$$

with a corresponding price of 50.5.

We assume that firms exhibit *inertia* in the sense that each period each firm may revise its strategy only with (independent) probability $\theta = \frac{2}{3}$.

Fully rational subjects would realize that since they are matched with the same opponents for the entire game, they should analyze the game as a dynamic game. It is well known that in finitely repeated games for which the stage game has a unique equilibrium the unique subgame perfect equilibrium consists of repeating the stage game equilibrium. However, due to the inertia our game is not a repeated game. While it is true that playing q_i^N every period remains a Nash equilibrium, it is not sequentially rational to do so. It turns out that due to the inertia there is a slight tendency towards Stackelberg behavior. This can be seen by solving recursively a game with two periods.[7] Solving the game for 40 periods becomes extremely cumbersome as this entails taking account of 8^{39} endnodes in the extensive form game.

[7]The solutions for games up to 4 periods can be obtained from the authors upon request.

Two qualitative features of the dynamic solution can be stated, however. First, all subjects should start with the same quantity q_i^1 in period 1. Second, q_i^1 is somewhere between q_i^N and q_i^W. Third, quantities never increase above q_i^1 during the game. Whether one thinks that subjects in an experiment can perform such enormous calculations does not seem relevant given that actual play was hardly ever even close to those three qualitative features. It seems safe to assume that rather than solving the game right at the beginning, subjects learned to play during the 40 periods.

To trace out the implications of the learning processes we will at first allow for longer time horizons than 40 periods. In line with the experimental setup we require that outputs must be chosen from a finite grid $\Gamma := \{0, \delta, 2\delta, ..., v\delta\}$, for arbitrary $\delta > 0, v \in \mathbb{N}$. We assume that q^W, $q^N \in \Gamma$.

3.1 Best reply dynamics

First, we consider the best reply dynamic suggested already by Cournot (1838). 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

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

according to some probability distribution with full support.

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

$$\Pi(r(q_{-i}) + \Delta, q_{-i}) = \left(\frac{99 - q_{-i}}{2}\right)^2 - \Delta^2.$$

Since the slope of $r(q_{-i})$ is $-\frac{1}{2}$ and $q^N \in \Gamma$, the grid points closest to $r(q_{-i})$ are either $r(q_{-i})$ itself or $r(q_{-i})+\delta/2$ and $r(q_{-i})-\delta/2$. That is, BR_i is either the singleton $\{r(q_{-i})\}$ or the set $\{r(q_{-i}) + \delta/2, r(q_{-i}) - \delta/2\}$.

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 know (see Theocharis, 1960) that the best reply dynamic is generally unstable for oligopolies with four firms. However, the next proposition shows that with inertia the best reply dynamic is stable and converges to the static Nash equilibrium.

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

Proof 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 zero.

From any state with $q_i = 0$, some i, 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 $(\omega^0, \omega^1, \omega^2, ...)$, such that for each $t \geq 1$ there is a unique player i who by choosing quantity q^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 the value of the potential function strictly increases. Hence, the process cannot return to a state it has visited before. Thus, all states other than ω^N are transient and the Proposition follows.$\square$

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 $\theta = \frac{2}{3}$. For randomized starting values (uniformly between 0 and 100) the process converges to a 1%–neighborhood of the Nash equilibrium on average in 26.5 periods. For starting values closer to the equilibrium convergence is even faster.

3.2 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, i.e. a strategy from the set

$$IB^{t-1} := \left\{ \begin{array}{c} q \in \Gamma : \exists i \in N, \text{ s.t. } q = q_i^{t-1} \text{ and} \\ \forall k \in N, \, \Pi(q_i^{t-1}, q_{-i}^{t-1})) \geq \Pi(q_k^{t-1}, q_{-k}^{t-1}) \end{array} \right\},$$

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 for each firm. No information about market or cost conditions is needed. Though, of course, the rule is more sensible if one knows that market conditions are constant and the same for all firms. 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 proved by Vega–Redondo (1997).

Proposition 2 *In the long run for $\varepsilon \to 0$ the Walrasian outcome (q_1^W, q_2^W, q_3^W, q_4^W) 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.

Simulation results on the convergence time of this process are not very meaningful since they depend too much on the assumed mutation process. A high rate of mutations clearly shortens the time until the competitive outcome is first reached.

3.3 Other learning processes

There are many other learning processes which could be relevant in our experiment. Some of those (e.g. fictitious play) yield similar predictions as the best reply process, i.e. convergence to Nash equilibrium, and are therefore difficult to differentiate from each other.[8] One qualitative theory that can be tested within our setup is Selten's *learning direction theory* (see Selten and Stoecker, 1986). The theory assumes that players have enough knowledge of the game to determine myopically the direction in which better actions can be found. In our context better actions are always 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 $\lessgtr$), the process has a unique restpoint which is the Cournot–Nash equilibrium. The information required is as in treatments BEST and FULL.

Even though subjects in the BEST treatment were not able to "imitate the best", they can still use some form of imitation. Maybe driven out of a desire for conformity they can *imitate the average* as they know the total quantity of the three other firms. 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 converge simply to the average of all starting values as the following equation shows,

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

With inertia the process depends on the realizations of the randomization device and is therefore path dependent.

Finally, we consider a learning process which works even if subject have as little information as in treatment NOIN.[9] We call it *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.

[8] While we did not provide subjects with a history of play, some subjects took notes and could have played fictitious play.

[9] Another learning process which works with this little information is reinforcement learning (Roth and Erev, 1995). However, since the grid in quantities is so fine, the number of actions in our game is about 10000 which makes reinforcement learning extremely slow.

Interestingly this process yields outcomes close to the collusive outcome. We have simulated a process in which firms follow with probability 0.95 the following equation

$$q_i^t = q_i^{t-1} + \text{sign}(q_i^{t-1} - q_i^{t-2}) \times \text{sign}(\pi_i^{t-1} - \pi_i^{t-2}) \times 0.01 \tag{3}$$

and with probability .05 they experiment and choose a quantity from the set $\{q_i^{t-1} - 0.01, q_i^{t-1}, q_i^{t-1} + 0.01\}$ each with equal probability. The simulations show that after an initial phase the process fluctuates around an average quantity of about 14.35.[10] The intuition is simple: if all firms happen to lower their quantity simultaneously, profits increase as long as Q^t is higher than the collusive outcome, and so firms will lower their quantity further. However, if Q^t is close to the collusive outcome, an increase in q_i^t by a single firm will be profitable driving Q^t back up again.

4 Experimental results

Probably the best summary of our results can be obtained from looking at Figures 1 through 5 which show the time path of total quantities in all 6 groups and for all 5 treatments. In each graph the upper line corresponds to the Walrasian total output of $Q^W = 99$ and the lower line corresponds to the Cournot–Nash output of $Q^N = 79.2$. As can be seen in Figure 1, quantities in treatment BEST converge fairly well to Q^N.[11] Total quantities in treatment FULL are clearly higher than in BEST, somewhere between Q^W and Q^N.

[place Figures 1-5 about here]

In the remaining three treatments subjects started with very high quantities as they did not have any information about demand or cost. However, in IMIT+ and NOIN subjects learn to adjust to reasonable quantities.[12] As

[10] If (7) is chosen with probability 1, the process converges to the collusive outcome.

[11] The reason that group BEST3 looks different is that there was one subject who played a limit pricing strategy. She would raise periodically her output to 100 and reap almost monopoly profits after the other players had reduced their quantities.

[12] In treatment IMIT+ group 2 (see the upper right graph in Figure 4) has to be considered a clear outlier. What happened was the following. In the last 20 periods one subject chose the maximum quantity $q_i = 100$ or $q_i = 99$ in 18 out of 20 rounds. Two other subjects reacted to this by supplying zero in the last 17 rounds. We chose to exclude this group from all further analysis. However, our main results are unchanged even if this group is included.

Table 2: **Summary Statistics**

Treatment	Mean35	Mean20	SDev35	SDev20	Gini35	Gini20
Best	84.32	82.56	13.11	10.00	.23	.20
	(4.56)	(2.48)	(11.94)	(10.64)	(.15)	(.13)
Full	90.68	91.60	9.69	9.58	.16	.17
	(5.77)	(6.48)	(4.30)	(4.46)	(.07)	(.08)
NoIn	102.14	93.55	30.07	23.20	.43	.39
	(19.79)	(14.73)	(13.72)	(10.83)	(.11)	(.13)
Imit	146.70	138.85	48.24	40.94	.49	.48
	(34.17)	(31.62)	(9.35)	(16.70)	(.11)	(.13)
Imit+	105.12	96.43	26.47	14.90	.23	.19
	(13.64)	(5.35)	(14.57)	(11.03)	(.13)	(.11)

Note: Standard deviations in parentheses.

mentioned above, we suspect that some subjects in treatment Imit might have had problems understanding the situation, which resulted in very high quantities. In fact, subjects in treatment Imit made losses in all but a few periods.

Table 2 gives the corresponding summary statistics. We report summary statistics for the last 35 and the last 20 rounds. The first 5 rounds cannot be seen as representative as subjects in treatments Imit, Imit+ and NoIn were not given enough information to find a reasonable starting value.

We have computed the following six measures:

- Mean35 (Mean20) is the average total quantity over the last 35 (20) periods and over all six groups in one treatment. E.g.

$$\text{Mean35} = \frac{1}{6}\sum_{j=1}^{6}\frac{1}{35}\sum_{t=6}^{40} Q_j^t.$$

- SDev35 (SDev20) is the average standard deviation of total quantities in the last 35 (20) periods over all six groups in one treatment.

$$\text{SDev35} = \frac{1}{6}\sum_{j=1}^{6}\sqrt{\frac{1}{35}\sum_{t=6}^{40}\left(Q_j^t - \frac{\sum_{t=6}^{40} Q_j^t}{35}\right)^2}.$$

- Gini35 (Gini20) is the average Gini coefficient of individual quantities within a period over the last 35 (20) periods and over all six groups in one treatment.

$$\text{Gini } 35 = \frac{1}{6}\sum_{j=1}^{6}\frac{1}{35}\sum_{t=6}^{40}\frac{1}{6Q_j^t}\sum_{i=1}^{4}\sum_{h=1}^{4}\left|q_{ji}^t - q_{jh}^t\right|.$$

Mean35 and Mean20 may be compared to the theoretical predictions about total quantities derived in Section 3. We report on non–parametric tests in the next section. But two observations can be made already. First, in treatments Best and Full, that is, in treatments in which subjects had complete information about the market, there does not seem to be a noticeable trend in total quantities as Mean35 and Mean20 are virtually the same in both treatments. Second, we have also computed means for periods 21 through 37 and compared them with the respective means for periods 38 through 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 between the last 3 rounds and the 17 rounds preceding them.

SDev35 and SDev20 index the variability of total quantities over time and can therefore be interpreted as a measure for the stability of aggregate behavior. Furthermore, a comparison of the two numbers can be used to get a rough impression about the rate of convergence. In all treatments the variability of total quantities seems to decrease over time, though the differences are not significant.

Finally, the average Gini coefficients measure the intra–group inequality in the distribution of quantities and, because profits are linear in quantities, they also measure the income distribution. The lower the Gini coefficients, the more symmetric is firms' behavior. The Gini coefficients will be analyzed further in Section 4.2.

4.1 Information and competition

Two important results about the relationship between information and quantities can be obtained from reading Table 2. The treatments can be ordered in a partly nested way as displayed in Figure 6 in terms of the information available to subjects.

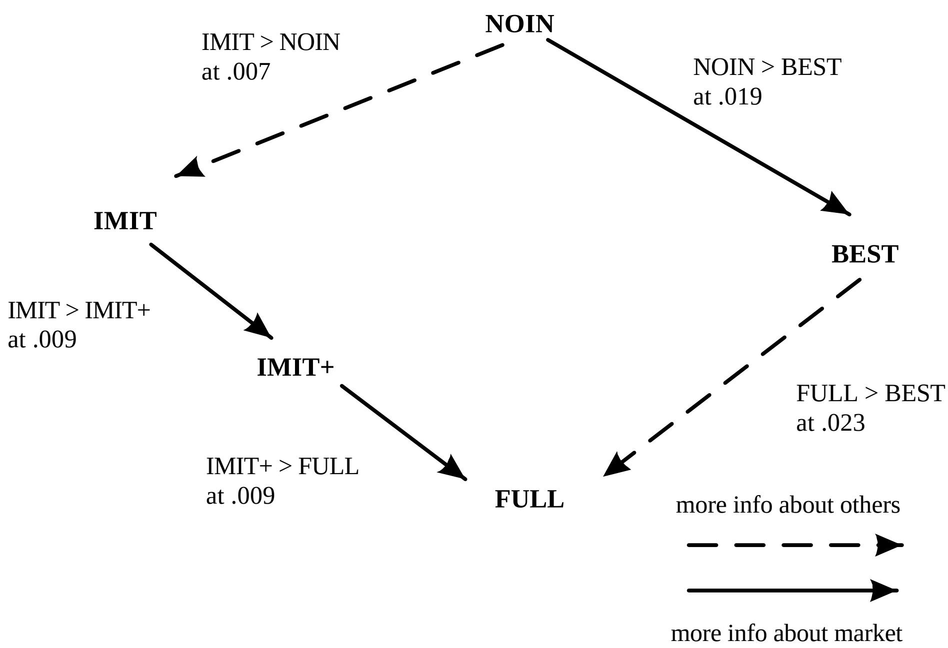

Figure 6: Significance levels of Mann–Whitney–U tests for differences in means based on last 35 periods, one–tailed.

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 to test for differences in means (based on the last 35 periods, one–tailed).[13] In

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

each case total quantities are significantly lower in the treatment with more information about the market. The significance levels are .019 for NOIN vs. BEST, .009 for IMIT vs. IMIT+, and .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. FULL and NOIN vs. IMIT. It shows that total quantities are significantly higher in the presence of information about others. The significance levels are .023 for BEST vs. FULL and .007 for NOIN vs. IMIT. Thus, providing additional information about individual quantities and profits *increases* total quantities.

The latter results are especially interesting with respect to the theoretical predictions about imitation. We have shown in Section 3 that imitation yields more competition than behavior based on myopic best replies or other rules discussed. The data reveal that if the information which is necessary to imitate successful behavior is available, competition indeed becomes more intense. Of course, the data do not coincide with the exact prediction obtained by studying imitation dynamics. For the above mentioned reasons treatment IMIT yields quantities in excess of the theoretical predictions. The IMIT+ results, however, 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.[14] Furthermore, some groups (see Figure 4) converge nearly perfectly to the competitive outcome.

In treatment FULL average total quantities are somewhere in between the competitive output and the Cournot output. But quantities in FULL are significantly higher than those in BEST, which indicates that imitation plays the predicted role. When looking at other players who receive higher payoffs due to higher quantities, the temptation to match the higher quantities apparently is hard to resist even if own profits are reduced by doing so. On the other hand, quantities in FULL are significantly lower than in IMIT+ which shows that at least some individuals follows best reply considerations.

In general, one has to take into account that Vega–Redondo's (1997) theorem is based on the notion of stochastic stability and, therefore, makes a prediction only for the very long run. Obviously, one should not expect to observe the stochastically stable state as the outcome of an experiment

[14] The appropriate Kolmogorov–Smirnov test shows that rejection is not even possible at a 20% level of significance. The corresponding tests for treatments FULL and IMIT lead to rejections at the 5% level. The same is true for testing whether the total quatities in BEST could have been drawn from a normal distribution with mean 79.2.

lasting only 40 periods. However, qualitatively our aggregate data support Vega–Redondo's result quite well. Whether this is also true on the individual level will be addressed in Section 4.3.

The fact that competition is more intense when firms know more about the individual quantities and profits of their rivals provokes traditional views on competition policy. For example, Green and Porter (1984) show that firms behave less competitive when they can observe their rivals' quantities immediately. For a similar reason anti–trust authorities like the European Commission often allow trade associations to publish only aggregate industry data.[15] Our result, however, indicates that it would be better to inform all firms about their rival's quantities and profits.

In this context it is also interesting to note that in none of the groups there has been any successful attempt to establish collusion. The averages of total quantities are above the static Cournot solution (Q^N) in all groups, and only occasionally total quantities fell below Q^N. The collusive price of 50.5 was exceeded in exactly 12 out of 1200 observations.[16]

The main observations can be summarized as follows.

Result 1 More information about the market yields less competitive outcomes.

Result 2 More information about behavior and profits of others yields more competitive outcomes.

Result 3 There were no successful attempts of collusion.

Result 2 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 with our Best treatment. One should note, however, that Fouraker and Siegel only report quantity choices of the second to last period of each session, which makes it impossible to say anything about the average quantities in their experiments. Hence, nothing can be concluded concerning learning of players.

[15] See e.g. the "Seventh Report on Competition Policy" by the Commission of the European Community, Luxemburg, 1978.

[16] Occasionally there have been 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.

Rassenti *et al.* (1996) ran several oligopoly experiments with five firms. Their central issue is whether repeated play in such a setting converges to the unique static Nash equilibrium. Surprisingly, and in contrast to previous studies and our own, they observed no convergence at all, neither to the Cournot equilibrium, nor to any other rest point. Moreover, this result does not depend on the amount of information subjects had about the other firms in the market.

Several differences between their design and ours could account for this. Most importantly, they introduce large differences in cost. The marginal cost parameter of the least efficient firm was nine times higher than that of the most efficient firm. This implied that the lowest Nash equilibrium output was only a little larger than the half of the highest Nash output. In addition, firms were not told how big the differences in costs were. As we show in the next section, one reason for the high volatility in Rassenti *et al.* (1996) could be the unequal distribution of quantities (and profits) in equilibrium.[17]

4.2 Inequality and stability

Table 2 suggests that there is a correlation between the standard deviations of total quantities and the Gini coefficients. In fact, when observing the experiments we had the impression that the volatilities of both, individual and total quantities, were related to the differences in individual quantities and profits. To test this in a more rigorous fashion we estimated whether the total changes of quantities from one round to another can be explained by inequality as measured by the Gini coefficients. Table 3 shows the results of an OLS regression for the following equation.

$$TVar^t = \beta_0 + \beta_1 Gini^{t-1} + \beta_2 Q^{t-1} + \beta_3 Z^t. \tag{4}$$

Total variation in a given period t, $TVar^t = \sum_{i=1}^{4} \left| q_i^t - q_i^{t-1} \right|$, is regressed on last period's Gini coefficient, $Gini^{t-1}$, on last periods total quantity, and on Z^t, the number of subjects who were allowed to revise their quantity in period t. Clearly, the last variable should increase $TVar^t$ and, unsurprisingly, β_3 is positive and significantly different from zero in all treatments. Neither is it surprising that β_2 is positive and (with the exception of IMIT) significantly so in all treatments.

[17]Another difference was that Rassenti *et al.* did not have inertia in their experiment. However, in Huck, Normann, and Oechssler (1997) we show that inertia does not have a noticable influence.

Table 3: **Factors influencing the variability: OLS results**

$$TVar^t = \beta_0 + \beta_1 Gini^{t-1} + \beta_2 Q^{t-1} + \beta_3 Z^t$$

	BEST	FULL	NOIN	IMIT	IMIT+
β_1	32.91	46.52	53.66	30.31	71.91
	(6.50)	(7.22)	(10.52)	(13.30)	(9.11)
β_2	.22	.25	.39	.05	.17
	(.06)	(.05)	(.04)	(.04)	(.05)
β_3	4.37	2.86	17.08	20.12	8.03
	(1.13)	(.66)	(1.99)	(2.68)	(1.78)
β_0	−24.64	−25.67	−74.98	−25.81	−33.58
	(5.42)	(4.91)	(8.91)	(14.08)	(7.44)
R^2	.267	.327	.434	.210	.374

Note: Standard deviations in parentheses. Z denotes the number of subjects who were able to adjust their strategies.

Our initial hypothesis concerning the impact of inequality on total changes is clearly confirmed. The values of β_1 are not only highly significant but also very substantial. This may not be surprising for the treatments in which subjects were informed about others' behavior. Naturally, subjects tried to avoid being the market's sucker and, therefore, revealed a tendency to imitate successful strategies. Such behavior obviously leads to greater adjustments when the Gini coefficients are large as compared to a situation in which quantities and profits are evenly distributed. But this cannot be the only reason for the impact of inequality since β_1 is also an important explanatory variable in treatments BEST and NOIN, in which subjects could not observe individual quantities and profits of others. In BEST subjects were able to compute best replies and average payoffs of others, and movements toward either quantity could explain why total adjustments are larger when quantities are more unevenly distributed. The real puzzle is the question why the inequality also drives adjustments in NOIN.

We suspect that the reason is more a matter of correlation rather than causation. Recall that the unique Nash equilibrium is symmetric. When quantities are all roughly the same, whether above or below q^N, subjects could move together in the direction of q^N and find their profit increased every time. Once the Cournot output is reached, no subject has individually success with changing his quantity, and total variability should be low. In Figure 5 groups 2 and 3, which converged almost perfectly to the Cournot equilibrium, have an average Gini coefficient which is clearly below the treatment average (.29 versus .39). On the other hand, if quantities are very unequally distributed, subjects can never be close to the Nash equilibrium and adjustments will be large.

Result 4 The more unequal quantities are distributed, the higher is the volatility of the adjustment process.

4.3 Individual learning behavior

While the analysis of group level behavior gives some insight as to the relative performance of the different learning theories, a closer look at individual behavior seems warranted. Since we suspected considerable heterogeneity in learning behavior between subjects we first analyzed each subject's time series data separately.

We have estimated with OLS the following equation

$$q_i^t - q_i^{t-1} = \beta_0 + \beta_1 \left(r_i^{t-1} - q_i^{t-1}\right) + \beta_2 \left(ib^{t-1} - q_i^{t-1}\right) + \beta_3 \left(i_i^{t-1} - q_i^{t-1}\right) \quad (5)$$

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 $\beta_k = 0, k \neq 1$.[18] Similarly, 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 explaining 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 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 Tables 4 and

[18]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.

Table 4: **Subject specific OLS regressions**

H_0	Number of subjects out of 24 (for IMIT+: 20) for which H_0 is rejected at the 5% level			
	BEST	FULL	IMIT	IMIT+
$\beta_1 = 0^*$	16	16	-	-
$\beta_2 = 0^*$	-	6	18	9
$\beta_3 = 0^*$	18	11	9	11
$\beta_0 = 0^*$	10	4	10	3
$\beta_1 = \beta_2 = 0$	-	21	-	-
$\beta_1 = \beta_3 = 0$	21	20	-	-
$\beta_2 = \beta_3 = 0$	-	16	24	18
$\beta_1 = \beta_2 = \beta_3 = 0$	-	24	-	-
$\beta_1 = 1$	23	22	-	-
$\beta_2 = 1$	-	23	21	18
$\beta_3 = 1$	22	23	22	14

Note: Only periods in which subjects were allowed to adjust their quantities are included. * The alternative hypothesis was that $\beta_i > 0$.

Table 5: **OLS Regressions with pooled data**

	BEST	FULL	IMIT	IMIT+
β_1	.430 (.038)	.366 (.044)	-	-
β_2	-	.110 (.038)	.465 (.046)	.435 (.040)
β_3	.340 (.038)	.344 (.038)	.151 (.048)	.273 (.047)
β_0	1.42 (.377)	2.14 (.447)	8.70 (1.33)	1.08 (.643)
R^2	.410	.507	.356	.439
DW	2.05	2.06	2.20	2.17
Obs.	610	631	620	533

Note: Standard deviations in parentheses. DW = Durbin Watson statistic. Subject dummies are used with the restriction that their coefficients sum to zero. Only periods in which subjects were allowed to adjust their quantities are included.

5 we have included only variables in the regression which were observable to the subjects.[19]

Running (5) for each subject separately we were able to perform hypothesis tests which are shown in Table 4. For almost all subjects at least one variable helped to explain the data, which is demonstrated by F–tests for the null–hypothesis of all β_k being zero. In all treatments "imitate the average" seems to play a substantial role. Furthermore, two thirds of the subjects in BEST and FULL adjust at least partially in the direction of the best reply. However, only one or two, respectively, do so completely (i.e. with $\beta_1 = 1$). In FULL there are 7 subjects for which both, best reply and imitate the average are significant. But only one subject of those for which imitate the best is important also takes best replies or the average quantity into account. It seems that there is a minority of subjects who are pure imitators of the best action and they may be responsible for the higher quantities in FULL vs. BEST.

A similar picture emerged when we estimated equation (5) with pooled data of all subjects in each treatment (see Table 5). The coefficients yield an indication about the relative importance of the explanatory variables. The coefficients of all three variables are highly significant and have the expected sign. 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 responsible for the difference in outcomes between those two treatments. The records about the use of the profit calculator have the same tendencies. While the profit calculator was used on average in 17.8 rounds in treatment BEST, it was used on average only in 15.4 rounds in FULL. Given the inertia subjects had the possibility to use the calculator on average in 27 rounds. Thus, in more than half the rounds subjects consulted the calculator.

In both, IMIT and IMIT+ the "imitate the best" variable is the major factor. In IMIT+ the experimental results nicely match the theoretical prediction, namely convergence to the Walrasian output. We suspect that responsibility for the high quantities in IMIT lies with the very high constant term, which might hint that other, hidden factors are relevant.

Result 5 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. In all

[19] Since none of the variables was observable in NOIN we did not run the regressions for this treatment.

treatments in which this was possible subjects were partly influenced by the average quantity of the other players.

Including other variables in equation (5) 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 lagged versions of the variables in (5). 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 or regret type learning processes (see e.g. Hart and Mas–Colell, 1997). However, lagged variables were not significant in any treatment.

Another learning theory which has been used much recently is reinforcement learning (Roth and Erev, 1995). Reinforcement learning assumes that strategies are played with probabilities proportional to their accumulated payoffs in the past. While Erev and Roth (1997) have generalized their model by adding recency parameters, reference points and experimentation, the model is not able to explain why factors matter which we found to be of key importance. In particular, information about the market and information about the individual play of others turned out to be very important treatment effects (see Figure 6). Neither effect can be explained by reinforcement learning.

We have also tested the other two learning theories mentioned in Section 3.3. Directional learning requires a knowledge of the direction in which payoff improvements can be found. Hence, it can be applied only in treatment Best and Full. Table 6 shows the resulting hypotheses and the percentages of cases for which each hypothesis was correct. Three observations are apparent: (1) Directional learning does better in treatment Best than in Full. (2) It performs better for upward adjustment than for downward adjustments. And (3) the performance with respect to the strong hypotheses is very poor (note that the weak hypotheses count non–movement as success). We take those relatively low success rates as evidence against directional learning.

Finally, trial and error learning is a type of learning which is applicable even in treatment Noin. Table 7 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

Table 6: **Directional learning**

Treatment	hypothesis	% of cases correct
Best	$r_{-i}^{t-1} > q_i^{t-1} \Rightarrow q_i^t \geq q_i^{t-1}$	83.2
	$r_{-i}^{t-1} > q_i^{t-1} \Rightarrow q_i^t > q_i^{t-1}$	68.7
	$r_{-i}^{t-1} < q_i^{t-1} \Rightarrow q_i^t \leq q_i^{t-1}$	69.2
	$r_{-i}^{t-1} < q_i^{t-1} \Rightarrow q_i^t < q_i^{t-1}$	54.6
Full	$r_{-i}^{t-1} > q_i^{t-1} \Rightarrow q_i^t \geq q_i^{t-1}$	77.1
	$r_{-i}^{t-1} > q_i^{t-1} \Rightarrow q_i^t > q_i^{t-1}$	54.1
	$r_{-i}^{t-1} < q_i^{t-1} \Rightarrow q_i^t \leq q_i^{t-1}$	66.3
	$r_{-i}^{t-1} < q_i^{t-1} \Rightarrow q_i^t < q_i^{t-1}$	41.5

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

Table 7: **Trial and error learning**

		$\Delta q^{t-1} > 0$	$\Delta q^{t-1} = 0$	$\Delta q^{t-1} < 0$
	$\Delta\pi^{t-1} > 0$	$\Delta q^t \geq 0$	–	$\Delta q^t \leq 0$
Predictions	$\Delta\pi^{t-1} = 0$	–	–	–
	$\Delta\pi^{t-1} < 0$	$\Delta q^t \leq 0$	–	$\Delta q^t \geq 0$

	Best	Full	Noin	Imit	Imit+	Total
# of decisions	626	643	633	653	546	3093
Correct predictions	511	517	500	530	433	2488
Success in %	81.63	80.40	78.99	81.16	79.30	80.44

they applied. Note, however, that there is a theoretical problem. If subjects in NoIn 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 plays differently, then very different outcomes can result.

Result 6 Among the examined alternative learning theories only trial and error learning performs reasonably well. Both, reinforcement learning and directional learning, fail to explain important features of our results.

Since we have used many non–economic students as subjects, it might be interesting how they fared as compared to the economic students who should have (so one thinks) a superior understanding of the situation. In fact, 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 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 individuals' propensity to imitate successful strategies, while the latter is based on individuals' ability and willingness to adjust behavior towards best replies. Competition, however, is always strong enough to frustrate any attempts for collusion. Furthermore, we find that the stability of behavior over time depends on the distribution of quantities and profits with more symmetry inducing less adjustments. This is partly explained by imitation and partly by the observation that if firms adjust their quantities in the same direction it is more likely that they find the static equilibrium of the game fast.

The analysis of the individual data showed that none of the theoretical learning processes which are discussed in Section 3 can on its own explain the observed behavior. Focussing on myopic best reply dynamics and imitation dynamics we find, however, that both adjustment rules play a role for subjects' 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. 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. These results are obtained for pooled data as well as for purely individual data. Since individual quantity adjustments depend on the available information, the composition of the available information drives the path of aggregate quantities or the intensity of competition.

Concerning alternative learning hypotheses we find that 'learning direction theory' largely fails while 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, we found no support for fictitious play as lagged variables do not seem to matter. Likewise, reinforcement learning is not able to explain any of our significant treatment effects.

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. No examined learning theory is rich enough to account for all these factors. This leads us to the view that there is need for new theoretical learning models accounting for these phenomena.

References

[1] 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.

[2] Cournot, A. (1838) *Reserches sur les Principles Mathematiques de la Theorie des Richesses*, Paris: Hachette, translated as *Research into the Mathematical Principles of the Theory of Wealth*, New York: Kelley, 1960.

[3] Erev, I. and Roth, A. (1997), "Modeling how People Play Games: Reinforcement Learning in Experimental Games with Unique Mixed Strategy Equilibria", mimeo (University of Pittsburgh).

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

[5] Green, E. and Porter, R. (1984), "Non–cooperative Collusion Under Imperfect Price Information", *Econometrica,* 52, 87-100.

[6] Hart, S. and Mas–Colell, A. (1997), "A Simple Adaptive Procedure Leading to Correlated Equilibrium", mimeo (Hebrew University, Jerusalem).

[7] Holt, C. (1985), "An Experimental Test of the Consistent–Conjecture Hypothesis", *American Economic Review,* 75, 314-325.

[8] Huck, S., Normann, H., and Oechssler, J. (1997), "Stability of the Cournot Process – Experimental Evidence", mimeo (Humboldt University, Berlin).

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

[10] Milgrom, P. and Roberts, J. (1991), "Adaptive and Sophisticated Learning in Normal Form Games", *Games and Economic Behavior*, 3, 82-100.

[11] Monderer, D. and Shapley, L. (1996), "Potential Games", *Games and Economic Behavior*, 14, 124-143.

[12] Rassenti, S., Reynolds, S., Smith, V., and Szidarovszky, F. (1996), "Adaptation and Convergence of Behavior in Repeated Experimental Cournot Games", mimeo (University of Arizona).

[13] 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*, 8, 164-212.

[14] 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*, 7, 47-70.

[15] Theocharis, R. (1960), "On the Stability of the Cournot Solution on the Oligopoly Problem", *Review of Economic Studies*, 73, 133-134.

[16] Vega–Redondo, F. (1997), "The Evolution of Walrasian Behavior", *Econometrica,* 65, 375-384.

Appendix A: Translation of instructions

Welcome to our experiment. Please read these instructions carefully. In the next 1 or 2 hours you will have to make some decisions at the computer. You can earn some real money. But please be quite during the entire experiment and do not talk to your neighbors. Those who do not follow this rule will have to leave and will not get paid. If you have a question please raise your arm.

You will receive your payment discretely at the end of the experiment. We guarantee anonymity with respect to other participants and we do not record any information connecting your name with your performance.

You can operate the computer with the keyboard or the mouse. Before the experiment there is enough time to make yourself familiar with the computer in a trial round. Money in the experiment is denominated in "Taler". At the end we exchange your earnings into DM at a rate of 500 T = 1 DM. The experiment is divided into several rounds. As said we start with a trial round. The real experiment starts with round 1.

You represent a firm which produces and sells a certain product. Besides you there are 3 other firms which produce and sell the same product. Your task is to decide how much to produce of your good. The capacity of your factory allows you to produce between 0 and 100 units each round. Production cost are 1T per unit. All units (also those of the other firms) are sold on a market (like on a stock exchange or in an auction).

For this the following important rule holds: The price can be between 100T and 0T. The more is sold on the market in total, the *lower* is the price one obtains per unit. To be precise the price falls by 1T for each additional unit supplied. If – this is only an example – the other firms supply together 10 units and your firm supplies 3 units, then total quantity is 13. The resulting price is $100 - 13 = 87$. If the total quantity were 90, the price would be $100 - 90 = 10$. *Profit per unit* is the difference between the price and the cost per unit of 1T. Note that you make a *loss* if the price is lower than the per unit cost. Your profit in a given round results from multiplying the profit per unit with your supplied quantity.

This ¶ only for BEST and FULL.

For this the following important rule holds: The price can 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.

This ¶ only for IMIT+.

In each round the quantities of all firms are recorded and the resulting profits are calculated. In each round you will be told your profit. Profits from all periods are added and the sum is paid out to you in cash at the end. Additionally you receive a fixed payment of 150T each round. This will be added to your profit each round.

In the first round you decide on a quantity you want to produce and sell. In all further rounds *chance* decides whether you have the opportunity to revise your quantity. The computer has a mechanism which is comparable to a "one–armed bandit": If you draw a "1" or a "2", you may change your quantity. If you draw a "0", you may not. That is, you may change your quantity in 2 out of 3 cases.

With a "0" the quantity of last period is supplied automatically again. Note, that your quantity might be fixed for several rounds. Following a "1" or a "2" you may revise your quantity.

In this case you will receive the following information. You are told each firm's last period quantity, the total quantity of the other firms last period, last period's price, and the profit of each firm.

This ¶ only for FULL.

In this case you will receive the following information. You are told the total quantity of the other firms last period, and last period's price.

This ¶ only for BEST.

In this case you will receive the following information. You are told each firm's last period quantity and the profit of each firm.

This ¶ only for IMIT & IMIT+.

Additionally, you have access to a **profit calculator.** The profit calculator is shown on the last page of the instructions. It has two functions: 1. It calculates your profit for arbitrary quantity combinations. That is, you can enter two values, a total quantity for the others (button "A") and a quantity for yourself (button "I"), and the machine tells you how much you would earn. 2. You can let it calculate for arbitrary quantities of others (button "A") the quantity at which you would make the highest profit (button "M"). You can use the machine as much as you want before each decision. Before we start you will have enough time to get to know the profit calculator directly at the computer.

This ¶ only for BEST and FULL.

Everything we have explained to you holds for the other firms as well. In fact, you are all reading exactly identical instructions.

The experiment lasts for 40 periods in total. Afterwards you will receive your payments in DM. We want to reassure you again that all data will be treated confidentially.

[For BEST and FULL a screen shot followed as shown in Appendix B]

Appendix B: Screen shots

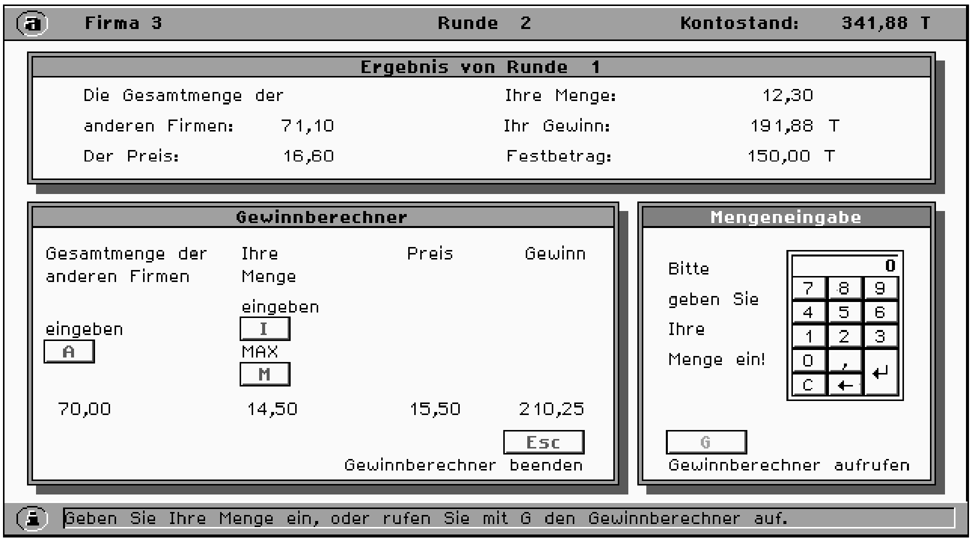

Figure 7: Screenshot of treatment Best

Translation (from top to button, left to right):

Bar at top: Firm 3, Round 2, Balance: 341.88 T

Window at top: Result of round 1, Total quantity of other firms: 71.10, The price: 16.60, Your quantity: 12.30, Your profit: 191.88 T, Fixed payment: 150.00 T.

Lower left window: Profit calculator, Enter total quantity of other firms, Enter your quantity, Price, Profit, Exit profit calculator: Esc.

Lower right window: Enter quantity, Please enter your quantity, open profit calculator.

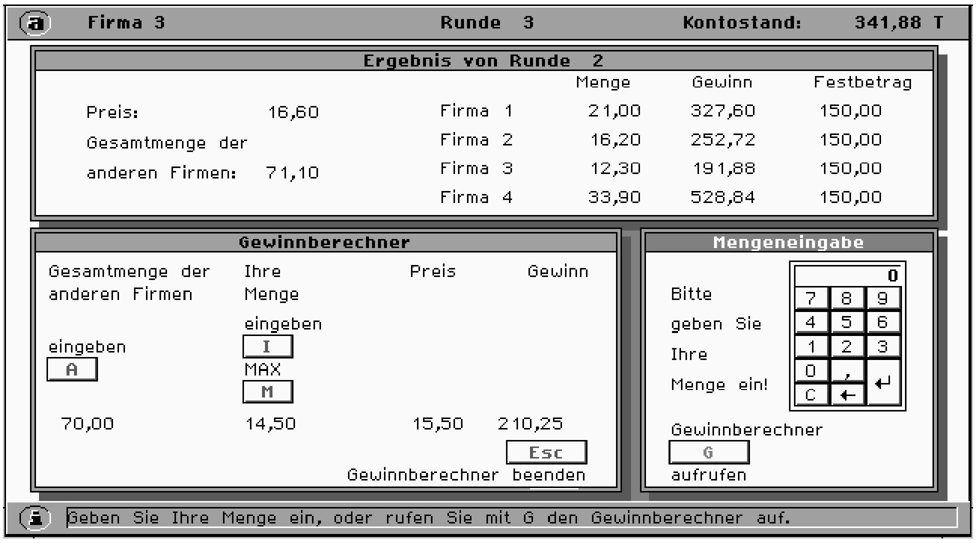

Figure 8: Screenshot of treatment FULL

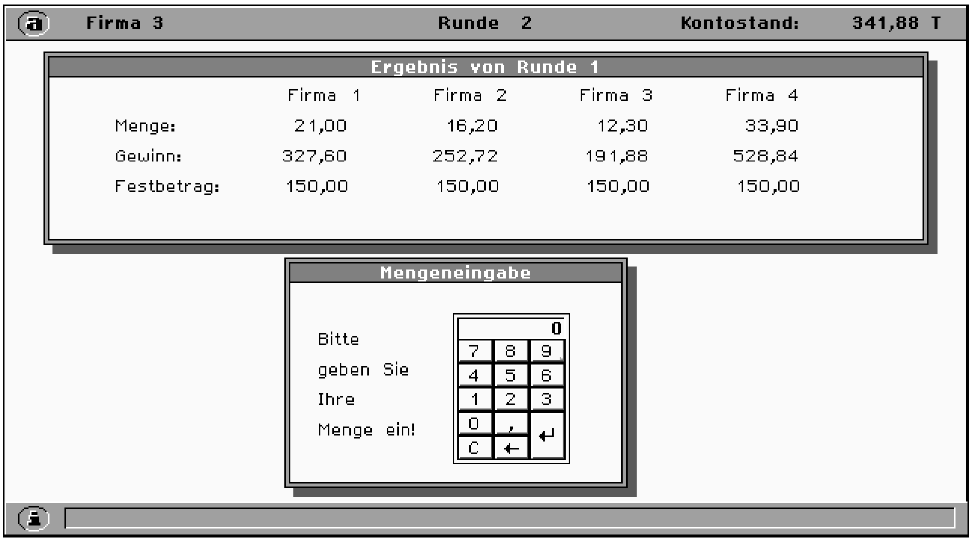

Figure 9: Sreenshot for IMIT and IMIT+

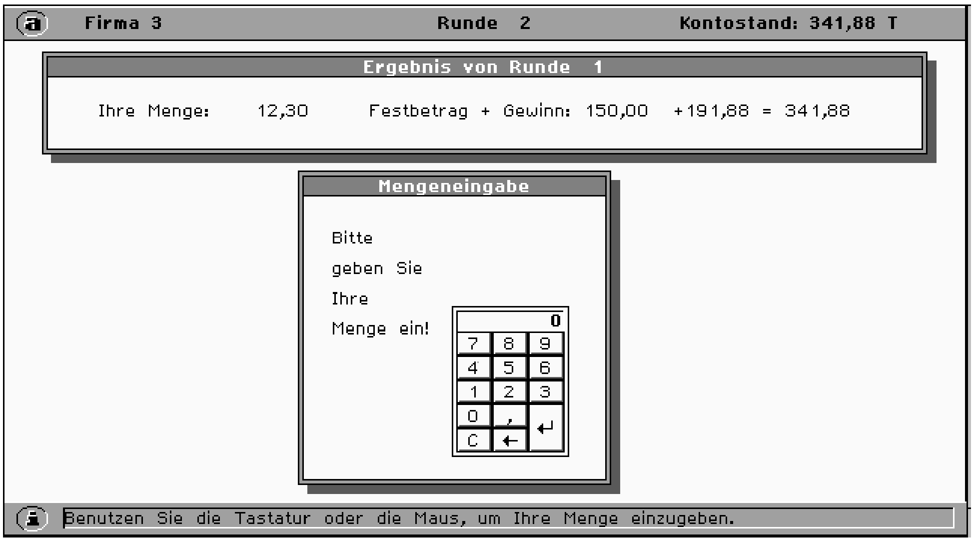

Figure 10: Screenshot for NOIN

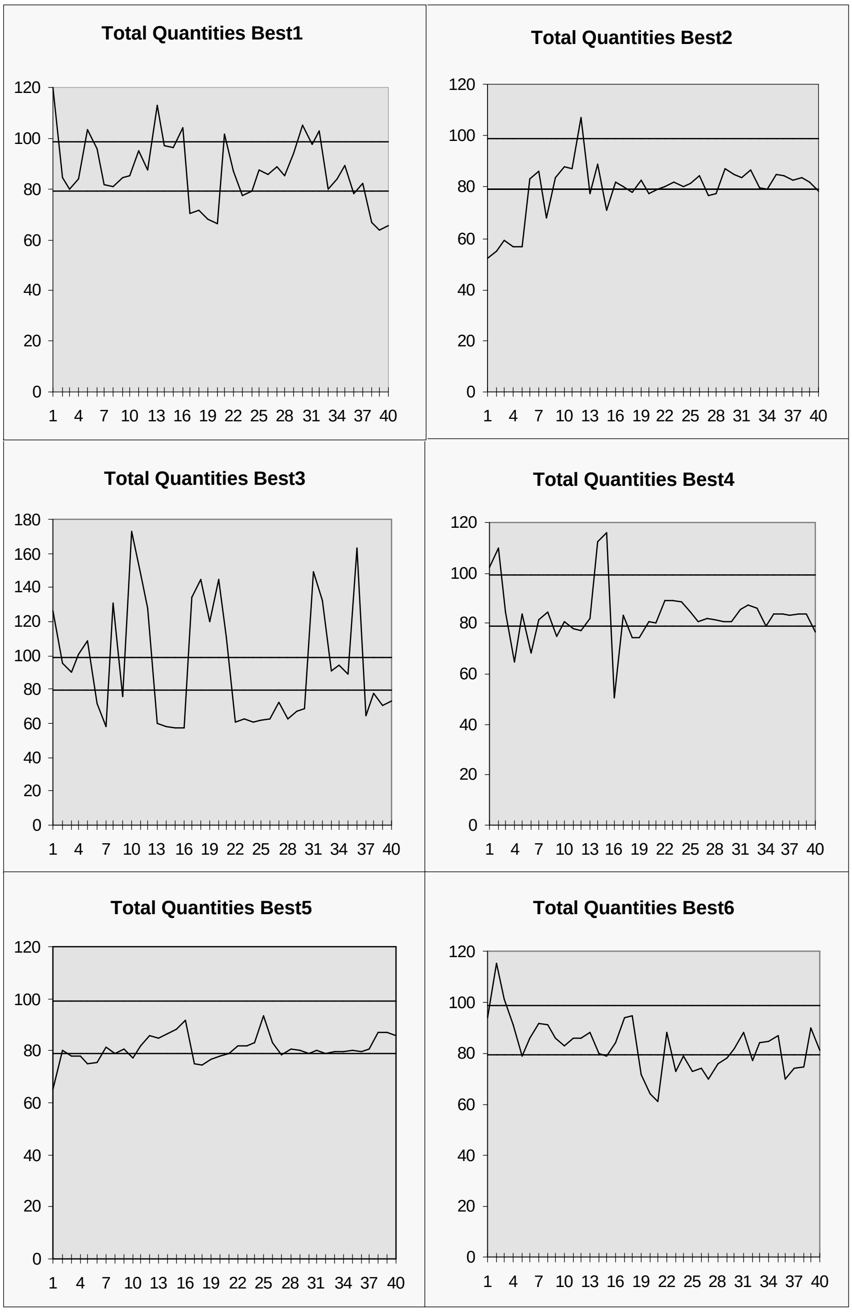

Figure 1: Treatment Best. Upper line = Q^W, lower line = Q^N.

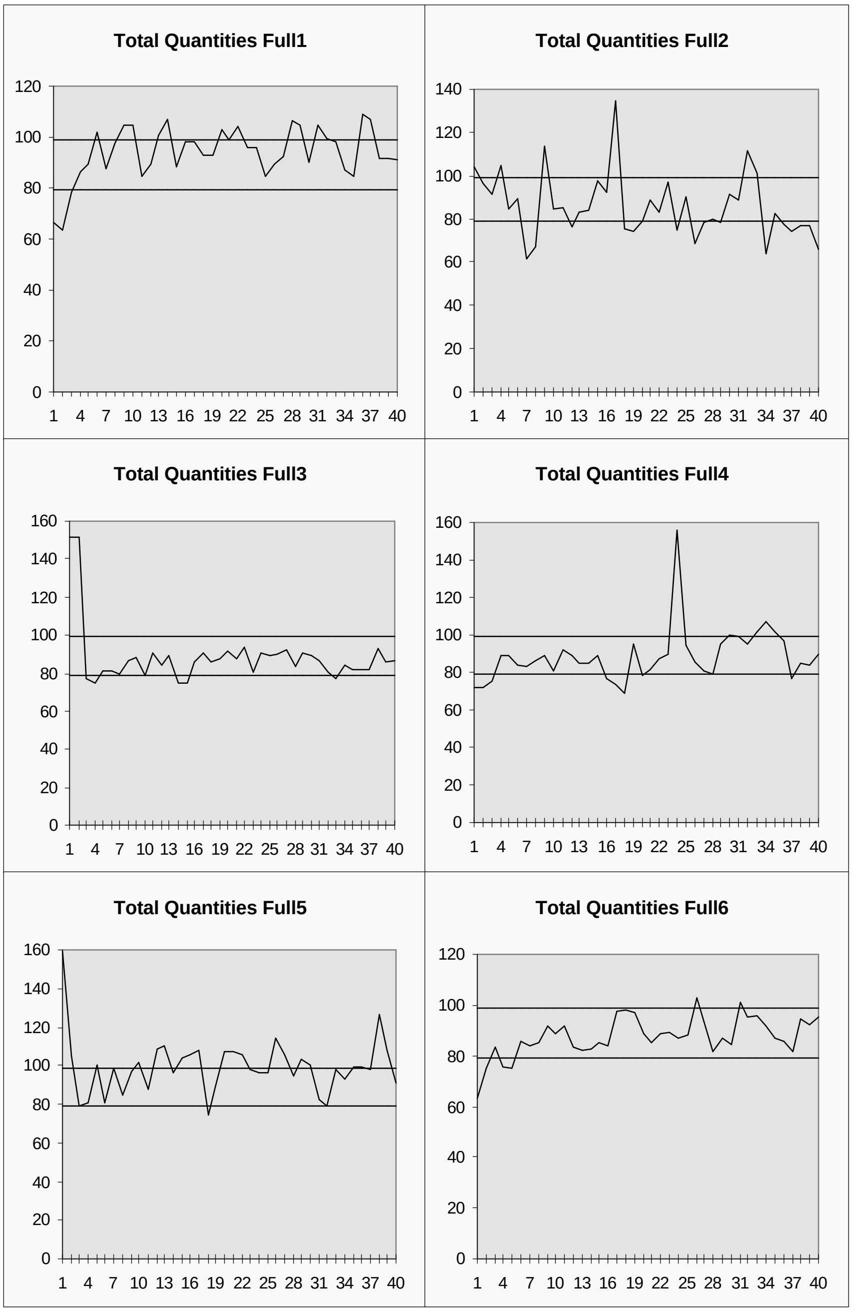

Figure 2: Treatment Full

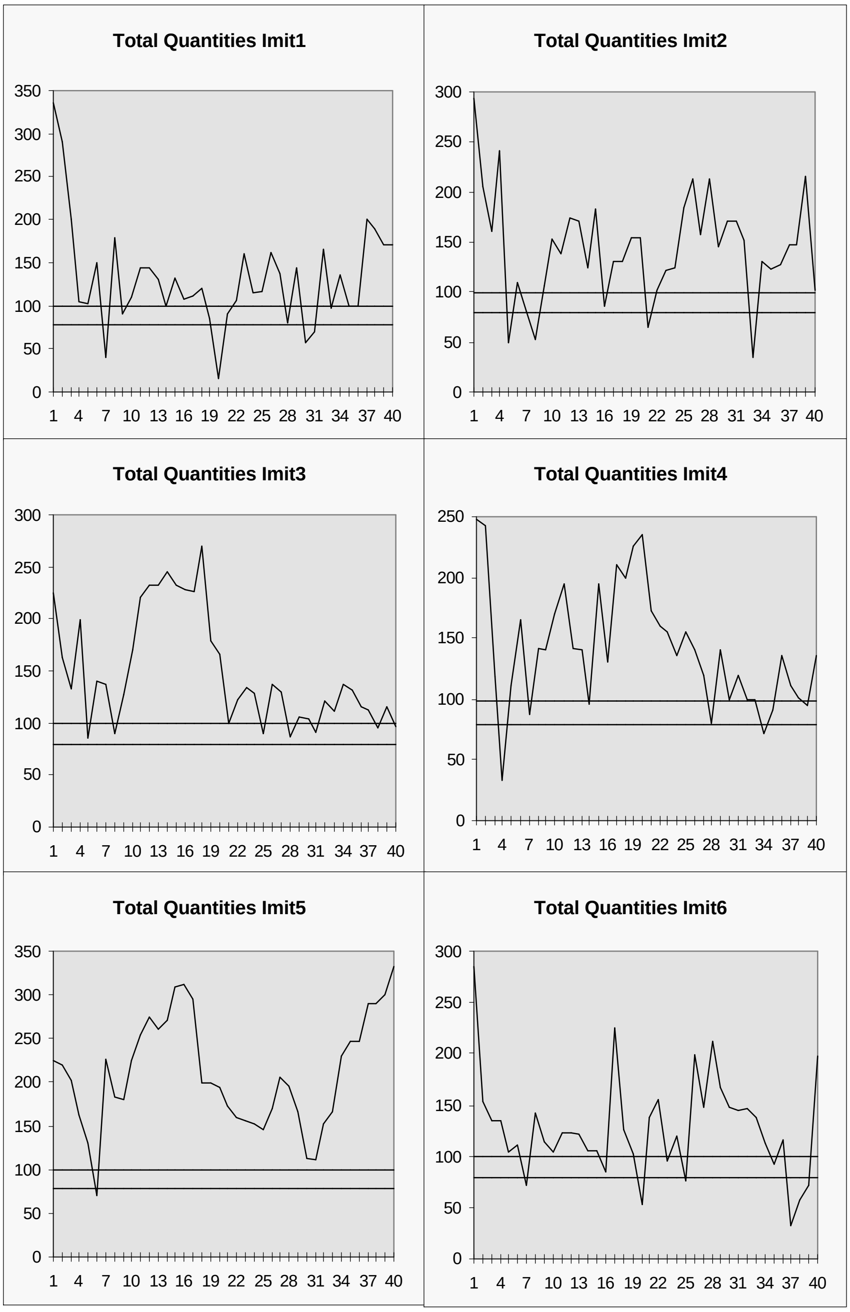

Figure 3: Treatment Imit.

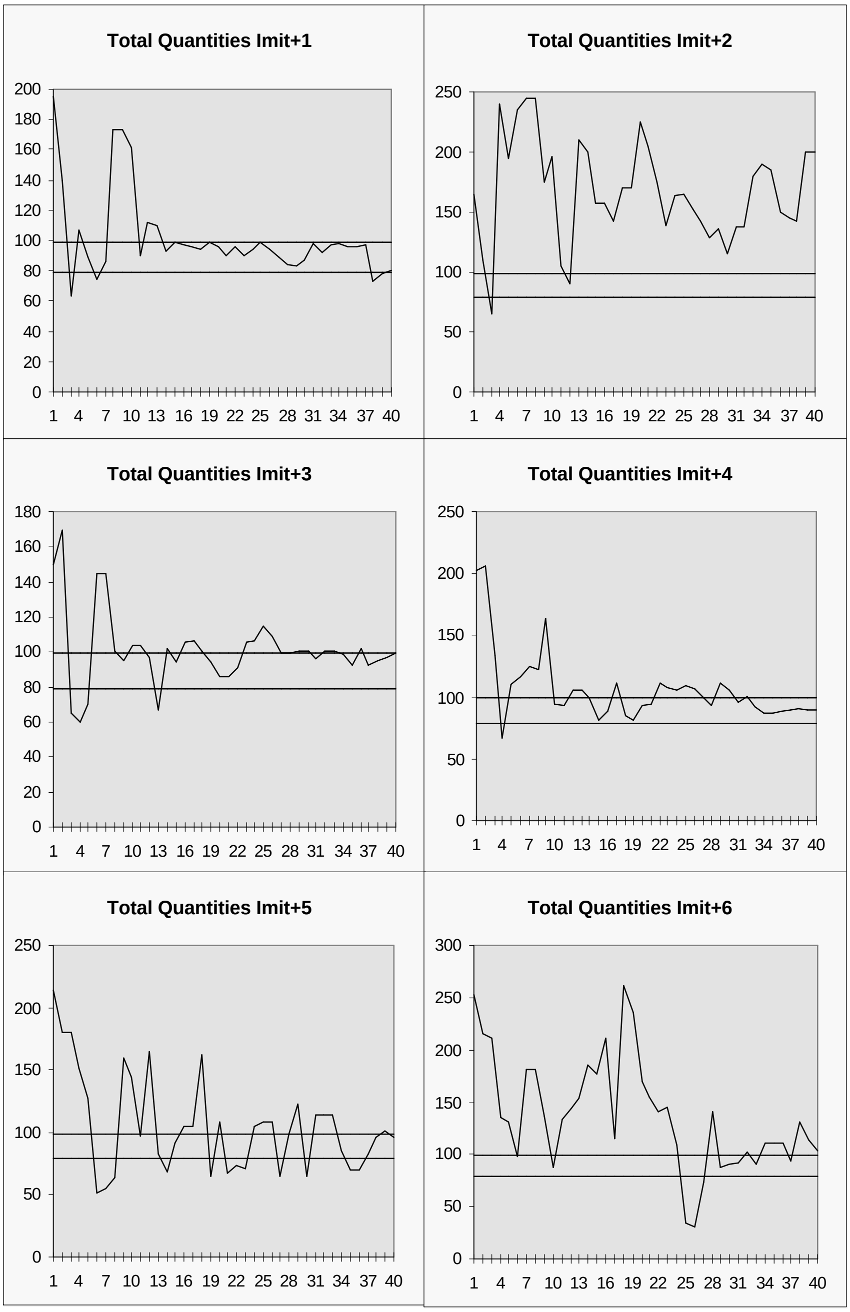

Figure 4: Treatment IMIT+.

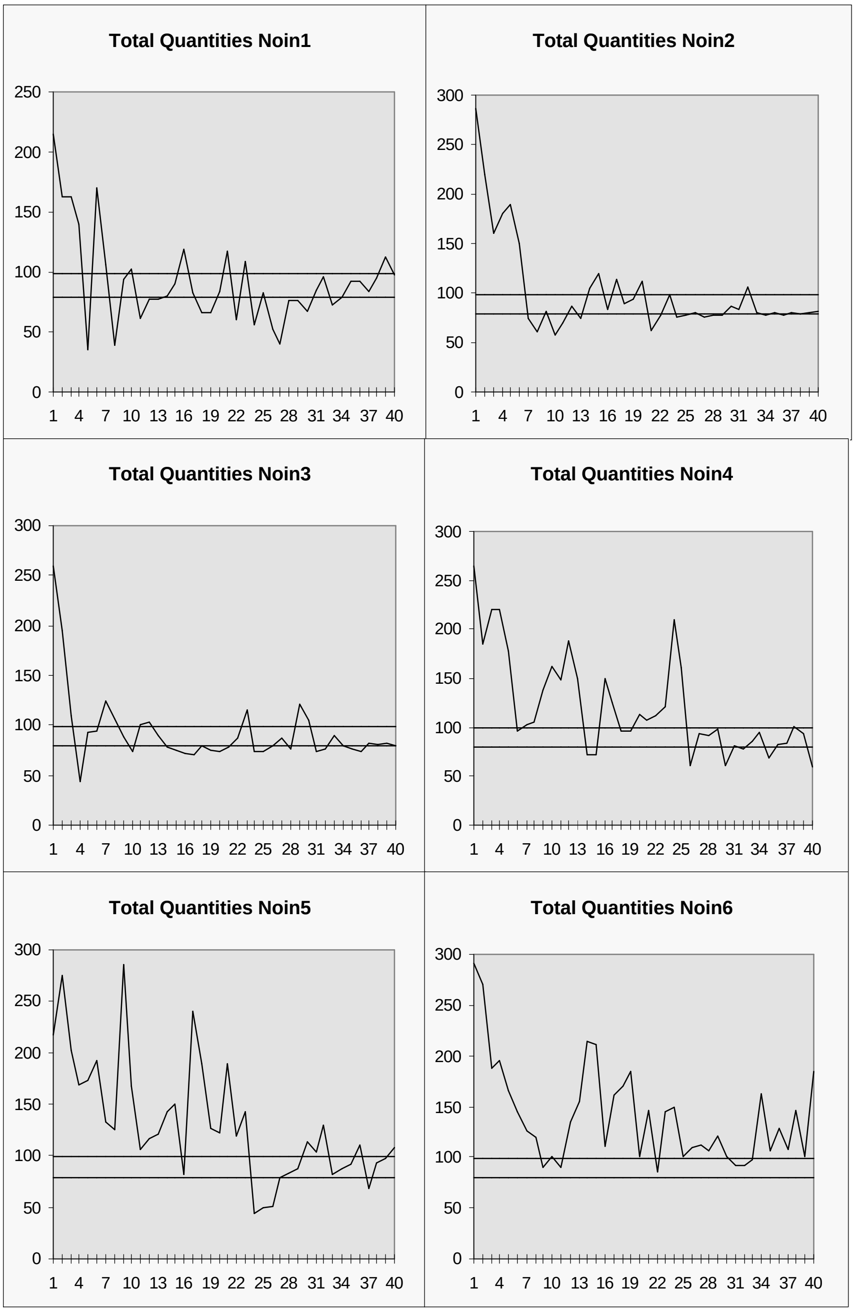

Figure 5: Treatment Noin.