UNIVERSITÀ DEGLI STUDI DI FERRARA
DIPARTIMENTO DI ECONOMIA ISTITUZIONI TERRITORIO
Via Voltapaletto, 11 – 44100 Ferrara

Quaderno n. 23/2009

December 2009

What Do You Want to Know? Information Acquisition and Learning in Cournot Games

Maria Bigoni

What Do You Want To Know?

Information Acquisition and Learning in Experimental Cournot Games *

Maria Bigoni†

Abstract

This paper presents an experiment designed to study firms' behavior and market dynamics, when information about the market structure and opponents' actions is difficult to acquire and process. Our experimental software controls the information acquisition process of the subjects. Paying attention not only to what players *do*, but also to what they *know*, it is possible to better understand the cognitive processes guiding their choices and, consequently, the impact of the informational structure on their behavior. According to our results, Best-Response Dynamics is the main component of subjects' learning process. We also find that, when subjects look at the strategies individually adopted by their competitors, they tend to imitate the most successful behavior, which drives the market towards a more competitive outcome.

Keywords: Experiments, Learning, Information, Oligopoly, Mouselab.
JEL Classifications: L13, C91, C72

1 Introduction

The classical approach to the theory of tacit collusion entails a model of repeated interaction among firms, in which market demand and cost functions are common knowledge, and firms are able to predict what their profits will be indefinitely in the future.

However, there are situations in which such assumptions are not realistic. For instance, uncertainty about market demand, or the opponents' costs, can make it difficult to coordinate on the joint-profit maximizing outcome, and the lack of information about competitors' past actions can hinder the detection of defections. One may then wonder whether – if the context is stable enough – firms are able to "learn" from past experience and to get to collusion anyway.

*I am indebted to Sven Olof Fridolfsson and Chloè Le Coq for help in organizing and conducting the experiments. I am also grateful to Giancarlo Spagnolo, Ken Binmore, Magnus Johannesson, Andreas Ortmann, Davide Dragone, and to the participants of the 2007 AFSE-JEE meeting in Lyon for helpful comments and suggestions. A special thank goes to Giovanni Ponti, for his precious help in redacting the final version of the paper. The usual disclaimer applies.

†Department of Economics, University of Bologna. Piazza Scaravilli 2, 40126, Bologna, Italy. e-mail: maria.bigoni@unibo.it.
Telephone number: +39 051 2098890. Fax: +39 051 0544522.

This paper aims at studying firms' behavior and markets' dynamics in such situations. A repeated Cournot game is reproduced in a controlled experimental setting, in which subjects are provided with very limited *ex ante* information about the market structure. In each round they can acquire additional information, both on the market structure and on market outcomes in previous rounds. The information acquisition process is controlled in a strict but non-obtrusive way, via the software MouselabWEB, which allows the experimenter to verify which information the decision makers look at and for how long.[1] In this way, it is possible to make inference on the impact of different informational inputs on subjects' actual behavior, and on the learning mechanisms adopted by the players. We borrow this methodology from Johnson et al. (1988). A similar approach was also used in Costa-Gomes et al. (2001), Johnson et al. (2002), Costa-Gomes and Crawford (2006), Gabaix et al. (2006). However, to the best of our knowledge, so far these techniques have been applied to the study of one-shot games only.

While the framing of the experiment, described in Section 2, is pretty close to the one adopted in previous works on learning in oligopoly games, our experimental technique is novel, as in our set-up subjects are free to choose what information to acquire, and we can monitor this choice. In contrast, in previous papers the informational framework was varied exogenously across treatments.[2] A second element of novelty is that players face virtual opponents enacted by computerized automata. This guarantees control over the learning rule adopted by players' opponents and allows us to check if and how this learning rule affects the information search and the subjects' market behavior.[3]

Our results suggest that the information gathered by subjects affects their choices through a composite learning mechanism, in which different components coexist. *Best-Response Dynamics* seems to be the most important factor, as subjects try to form expectations about their opponents' future actions and to optimally react to them. However, when subjects also look at the strategies individually adopted by their competitors, they tend to *imitate the most successful behavior*, which makes market outcomes more competitive. Finally, our data suggest that the opponents' learning rule does not have a very strong impact on players' information acquisition and learning processes.

The paper proceeds as follows. Section 2 introduces the market setting and presents three alternative theoretical benchmarks. The experimental design and procedures are presented in Section 3. Experimental results are described in Section 4, while Section 5 concludes. Instructions for the experiment are reported in the Appendix.

2 Stage Game and Theoretical Predictions

In all treatments, the basic layout remains the same. Four identical firms compete *à la* Cournot in the same market for 40 consecutive rounds. Their product is perfectly homogeneous. In every round t each firm i chooses its own output q_i^t

[1] This software was developed by Martijn C. Willemsen and Eric J. Johnson, and is available from the website http://www.mouselabweb.org/index.html

[2] Previous articles on learning in oligopolistic markets that are most closely related to the present work are Huck et al. (1999), Rassenti et al. (2000), Offerman et al. (2002) and Bosch-Domènech and Vriend (2003).

[3] See Duersch et al. (2005) for a similar approach.

from the discrete set $\Gamma = \{0, 1, ..., 30\}$, which is the same for every firm. Choices are simultaneous.

Price p^t in round t is determined by the inverse demand function:

$$p^t = \max(0, 81 - \sum_i q_i^t)$$

Let $C_i(q_i^t) = q_i^t$ be the cost function for every firm i; firm i's profit in round t will be denoted by

$$\pi_i^t = p^t q_i^t - C_i(q_i^t).$$

These functions were chosen so that the three main theoretical outcomes – namely collusive (joint profit maximizing), Cournot and Walrasian – are sufficiently far one from the other and belong to the choice set Γ. More precisely, collusive equilibrium is denoted by $\omega^M = (10, 10, 10, 10)$, Cournot-Nash equilibrium is $\omega^N = (16, 16, 16, 16)$ and Walrasian equilibrium is $\omega^W = (20, 20, 20, 20)$.

2.1 Three Theoretical Benchmarks

We are interested in studying the market dynamics when the stage-game we just described is repeated, and firms do not have all the information (or the computational capabilities) to evaluate what the standard theory predicts to be their optimal behavior. A number of learning models that could be effectively adopted to study such a situation have been developed in the last 20 years. Table 1 reports the learning models which are relevant in the context of our research.

Rule	Required information	Predicted equilibrium
1. Best-Response Dynamics	Competitors' aggregate quantity and BR function	Nash ($q = 16$)
2. Imitate-the-Best	Last round individual profits and quantities	Walrasian ($q = 20$)
3. Trial-and-Error	Own past profits and quantities	Collusive ($q = 10$)

Table 1: Theoretical benchmarks

As Table 1 shows, we only focus on three learning models, namely Best-Response Dynamics (Cournot 1838; Huck et al. 1999), Imitate-the-Best (Vega-Redondo 1997) and Trial-and-Error (Huck et al. 2000; Huck et al. 2004). The choice of these three models is justified for the following reasons: first, they are particularly simple; second, they are based on very different informational requirements; third, they yield well distinct market outcomes in the long run, namely the Cournot, Walrasian and collusive outcomes, respectively. The three models – described in the remainder of this session – will play an important role in the design of the experiment, as illustrated in Section 3, and will also be used as benchmarks for the econometric analysis in Section 4.

Best-Response Dynamics. This adjustment process was originally proposed by Cournot (1838) in his duopoly analysis. Under the best-response dynamics each subject sets his current output equal to the best (i.e., current round payoff maximizing) response to the last round output of his rivals. Cournot proved that this adjustment process converges to the unique Nash equilibrium for a duopoly with linear demand and constant marginal cost. It is generally well known that best-response dynamics does not converge to a stable equilibrium in oligopolies with a linear setup and three or more firms, as proven by the general instability result found by Theocharis (1960). Yet, it has been shown by Huck et al. (1999) that this process converges to the static Nash equilibrium if some inertia is introduced, namely, if it is assumed that, in every round, with some positive probability each player sticks to the strategy he chose in the previous round.

As for the informational requirements, this model of learning presumes that agents are able to observe their rivals' past play and that their computational capabilities and knowledge of the market structure are sufficient for them to compute their best-response, given the strategy profile adopted by their opponents.

Imitate-the-Best Vega-Redondo (1997) proposed a theoretical model of behavior of Cournot oligopolists which leads to surprising conclusions. The behavioral rule described in the model essentially prescribes to "imitate the best", i.e. to produce, in each round, the quantity produced in the previous round by the firm that got the highest profit.[4] The author shows that in the long run a Walrasian behavior results, *within any quantity-setting oligopoly* producing an homogeneous good, provided that the market demand curve is downward sloping.

It can be shown that if all the firms active in the market conform to this rule, the induced market dynamics can be characterized as a Markov chain. If the learning process were to consist only of an imitation component, each monomorphic state (i.e. each state in which all firms produce the same quantity), would be an *absorbing state* of the Markov process. To investigate the relative robustness of these outcomes, Vega-Redondo (1997) adds to the learning dynamics a "noise" dynamics.[5] This implies that, with some common independent probability, in every round each firm may "mutate", so that all of the possible quantities can be chosen with a given positive probability. The dynamic market process then becomes ergodic, and one can find the unique invariant distribution to which the process converges in the long-run, and study its asymptotic behavior as the probability of mutation approaches 0. Vega-Redondo (1997)'s main result is that the whole mass of the limit invariant distribution is concentrated on the monomorphic state in which all the firms produce the Walrasian quantity.

To "imitate the best", firms must have the opportunity to observe the strategy individually adopted by each of the competitors, and they need to know the individual profits of each of the opponents, or at least they must be able to infer it from the information they have on the market structure.

[4] An alternative model of learning based on imitation is presented in Schlag (1998).

[5] A similar approach is adopted in Kandori et al. (1993) and Peyton Young (1993).

Trial-and-Error This model of learning has been proposed by Huck et al. (2000, 2004). It makes few assumptions about both the availability of information and firms' cognitive abilities, as it just requires that they know their own past actions and their own profits. Framed in a standard symmetric Cournot oligopoly with n firms, this learning rule simply says that a subject would not repeat a mistake, i.e. if profits last round have decreased after an increase in quantity, then one would not increase quantity again. On the other hand, if profits have increased following an increase in quantity, one would not decrease quantity next round. As in Imitate-the-Best, some degree of experimentation is introduced by assuming that, with some small probability, each firm chooses an arbitrary direction of change, instead of following the model's prescription. Trial-and-Error learning can be thought of as a particularly simple form of direction learning (Selten and Stoecker 1986, Selten et al. 2005). According to learning direction theory, agents should reconsider their past experiences to figure out what would have been a better choice, and adjust their next decision in this direction. Under Trial-and-Error it is not assumed that players know their payoff function. As a consequence, they may not be able to infer what could have been better in the past, and can only judge how successful a change in action was, on the basis of their own experience.

Huck et al. (2000) show that, if each firm can choose its outcome from a finite grid, Trial-and-Error rule defines a Markov process that converges to the joint profit maximizing equilibrium if the cost function is weakly convex and market conditions are such that there exists only one symmetric situation in which joint profits are maximized.

3 Experimental Design

To reduce strategic uncertainty and increase the number of independent observations, we let subjects play against three "virtual" players enacted by the computer and programmed to follow a specific learning rule.[6] In this way we can control for the effect of the opponents' behavior on players' choices. The treatment variable is then represented by the learning rule adopted by the virtual opponents. Note that the three opponents of the same subject are all programmed to follow *the same* learning rule.

To avoid deception, subjects are informed that their opponents are "automata", that is: they are enacted by the computer. Subjects also know that these automata do not play at random but choose according to some rule, nonetheless they do not necessarily choose the same output. No other information is provided about the way the automata play.

The experiment is repeated under three treatments. We will present first the elements which are common across treatments, then explain the differences.

3.1 Information Provided to the Subjects

Participants know how many competitors they have. Instructions explain in plain words that there is an inverse relation between the overall quantity produced by the four firms and market price, and that a firm's total production

[6] A related experiment in which subjects play against each other is presented in a companion paper (Bigoni 2009).

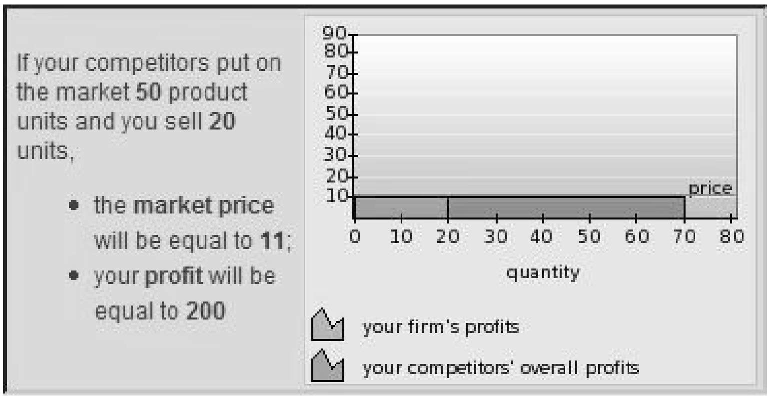


Figure 1: First function of the profit calculator.

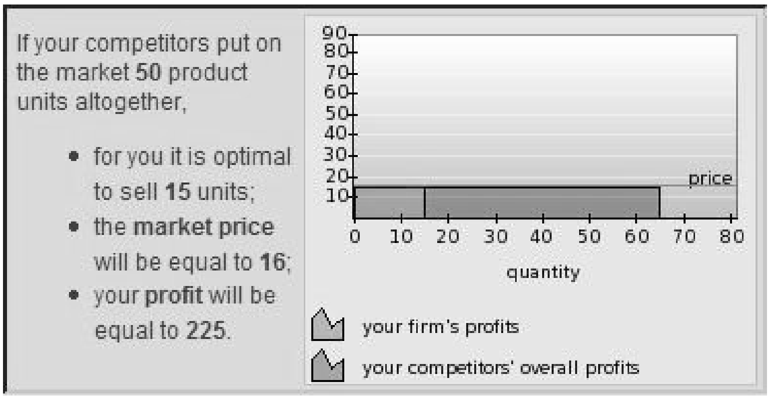


Figure 2: Second function of the profit calculator.

cost increases with the number of goods it sells. Besides, players are told that a firm's per-round profit is given by market price times the number of goods sold, minus production costs.[7]

Subjects are also endowed with a *profit calculator* similar to the one proposed by Huck et al. (1999). This device has two input fields that the subject can fill in: one for the total quantity produced by the other three firms in the market, one for the quantity produced by his own firm. If the player enters two (arbitrary) values, one for each of these fields, the profit calculator evaluates market price and the profit the subject would earn (Figure 1); if the subject just fills in the field pertaining to competitors' quantity and leaves the other one blank, the profit calculator computes the quantity that would yield him the highest profit and inform him about market price and profits he would earn, should he produce the suggested amount of good (Figure 2). The software we developed for this experiment records how many times subjects used the profit calculator and every trial they did.

[7]A copy of the instructions is provided in Appendix.

The number of rounds is equal to 40, and it is common knowledge among subjects.[8] After the first round, each player has the opportunity to look at three plots summing up information about what happened in the previous rounds (Figure 3). The graph in the top-left corner shows the quantity and the profit obtained by the player's firm in each of the previous rounds. The graph in the top-right corner is a bar-plot showing the quantity produced by each of the four firms in the market in the previous round, and the relative profit. The third graph displays the quantity produced by the player's firm compared with the aggregate quantity produced by his three competitors in each of the previous rounds, since the game began.[9] Subjects, however, are not able to look at all the three plots at the same time, since these plots are hidden behind three boxes on the computer screen and the player can open just one box at a time. Behind a fourth box is hidden the answer provided by the profit calculator. A box can be opened just putting the mouse cursor over it, and its content will be displayed on the screen until the cursor moves out of the box's borders. The software automatically records subjects' look-ups sequences and look-ups durations.

Besides these four boxes, on the computer screen there is a counter showing the cumulative profits earned by the player since the beginning of the experiment, and a timer displaying the current round's length, in seconds. After the last round, participants are shown their overall profit, compared with those of their three opponents.

3.2 Treatments

This experiment has been run under three conditions, which differ only by the learning rule adopted by the computer (Rule 1 to 3 in Table 1). In what follows we briefly recapitulate the rule adopted by the automata in each of the treatments.

Trial-and-Error (T&E) We programmed the computer so that in this treatment each automaton i sets its quantity q_i^t in round t equal to

$$q_i^t = q_i^{t-1} + s_i^{t-1}$$

where the quantity variation $s_i^t \in \{-1, 0, 1\}$ is given by

$$s_i^t = 1 \cdot \mathrm{sign}(q_i^t - q_i^{t-1}) \cdot \mathrm{sign}(\pi_i^t - \pi_i^{t-1})$$

if $(q_i^t - q_i^{t-1})(\pi_i^t - \pi_i^{t-1}) \neq 0$, where π_i^t are the profits of firm i at round t. If instead $(q_i^t - q_i^{t-1})(\pi_i^t - \pi_i^{t-1}) = 0$, the quantity variation is chosen at random from a uniform distribution over the set $\{-1, 0, 1\}$.

We also introduced a positive probability of error in the algorithm, so that in every round, with probability $\epsilon = 0.05$ each automaton randomly draws the quantity variation s_i^t from a uniform distribution over $\{-1, 0, 1\}$.

[8] This is in line with the procedure adopted in related works (Huck et al. 1999; Rassenti et al. 2000; Offerman et al. 2002; Bosch-Domènech and Vriend 2003), and with considerations made in Selten et al. (1997), and Normann and Wallace (2006).

[9] The order in which thee plots are presented varies across subjects, but it remains fixed across rounds.

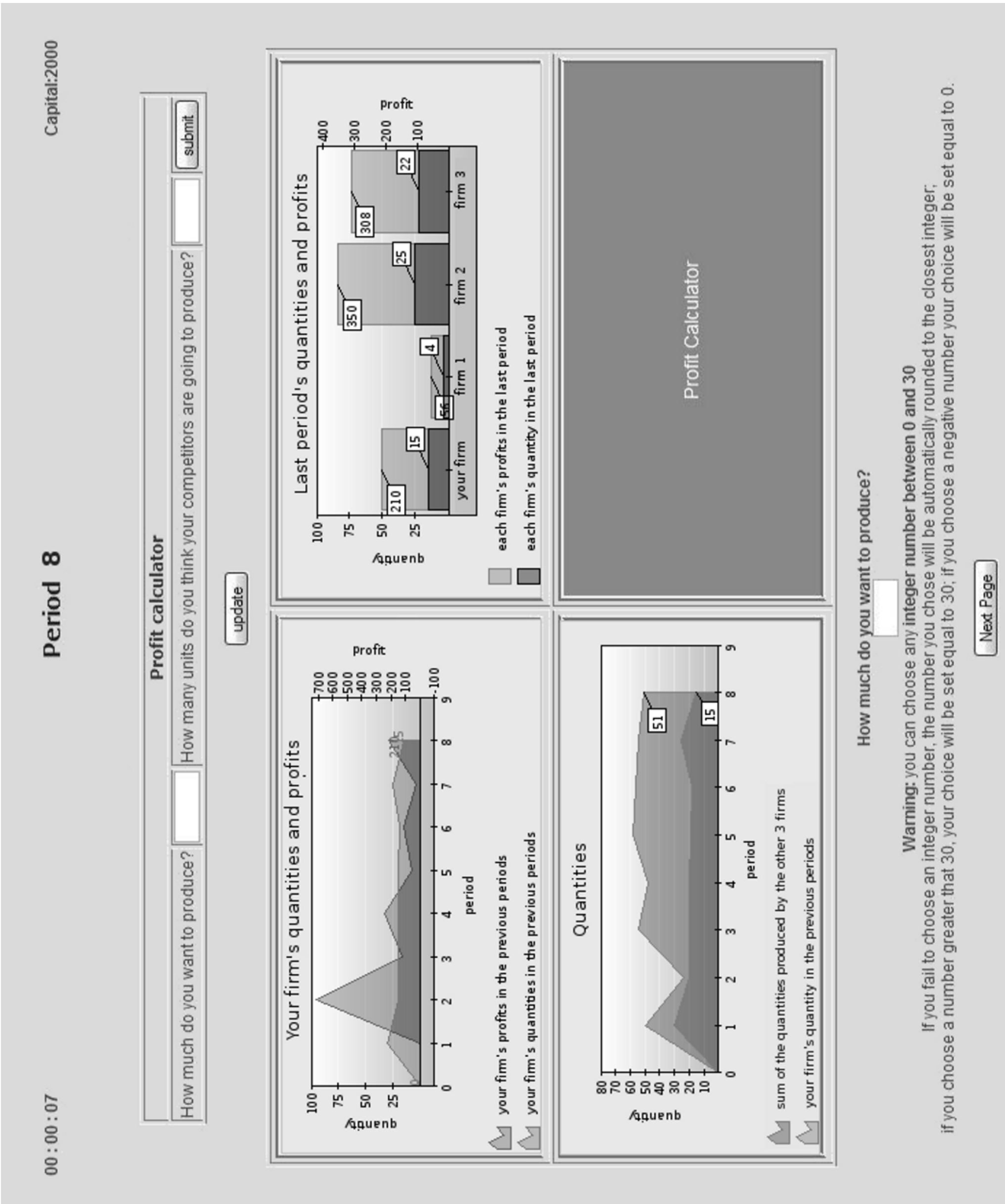


Figure 3: Screenshot of the graphical interface.

Best-Response Dynamics (BRD) In this treatment, the three automata behave according to the best-response dynamics. Therefore, in every round after the first one, they set their quantity q_i^t according to the following rule:

$$q_i^t = \begin{cases} 0, & \text{if } \sum_{j \neq i} q_j^{t-1} \geq 80 \\ \min\left\{30, \frac{80-\sum_{j\neq i} q_j^{t-1}}{2}\right\} & \text{if } \sum_{j \neq i} q_j^{t-1} < 80 \text{ and } \sum_{j \neq i} q_j^{t-1} \text{ is even} \\ \min\left\{30, \frac{80-\sum_{j\neq i} q_j^{t-1}}{2} + 0.5\right\} & \text{if } \sum_{j \neq i} q_j^{t-1} < 80 \text{ and } \sum_{j \neq i} q_j^{t-1} \text{ is odd} \end{cases}$$

We introduced some inertia into the algorithm governing the automata in treatment BRD: with independent probabilities equal to 0.05, in every round, each of them chooses $q_i^t = q_i^{t-1}$, otherwise it follows the myopic best-response dynamics.

Imitate-the-Best (ItheB) In the last treatment (ItheB) the automata behave according to Vega-Redondo (1997)'s algorithm. Automata are programmed to choose their output q_i^t equal to:

$$q_i^t = q_j^{t-1} \text{ where } j \in I \text{ and } \pi_j^{t-1} \geq \pi_h^{t-1} \; \forall h \in I,\, h \neq j$$

To introduce a small degree of noise in the automata's behavior, we programmed them to choose their quantity according to the Imitate-the-Best rule, with a probability equal to 0.95, while with a 0.05 probability they choose a random quantity.[10]

In all the three treatments, the quantity set by the automata in the first round is randomly drawn from a uniform distribution over the set Γ.

We adopted a within-subjects design, so every subject played against a single type of automata. The program was designed using MouselabWEB.[11] Two sessions of this experiment were run at the Stockholm School of Economics, on April 2, 2007. Twelve undergraduate students took part in the first session and eleven took part in the second one. Sessions lasted about one hour and a half, and the average payment (including the show up fee) was equal to 19.15 €.[12] In total, 7 subjects played under the ItheB treatment, and 8 under each of the other two treatments.

At the beginning of each session, participants were welcomed and seated in the lab. Tables and computers were arranged so to avoid any form of communication between participants. As soon as all subjects were ready to start, instructions were displayed on their computer screens, and each subject could read them at his own pace. Instructions were split into several parts and at the end of each of them an understanding test was submitted to the reader, who had to answer correctly before proceeding to the next page. When a subject finished reading the instructions he could start playing. At the end of the session, participants were called one by one in private and paid according to their total profits.

[10] This random quantity is equal $\lfloor x \rfloor$, where x is drawn from a normal distribution over the set $\Gamma = \{0, 1, \ldots, 29, 30\}$, with mean equal to the quantity the automaton chose in the previous round, and a standard deviation equal to 10.

[11] Martijn C. Willemsen and Eric J. Johnson, http://www.mouselabweb.org

[12] Payment was done in Swedish Kronors (SEK); 1 SEK was about €0.107 at the time the experiment took place, hence the average payment was 179 SEK.

4 Results

In what follows, we first examine the quantities chosen by the players, and show that they differ substantially across treatments, and are generally far from the theoretical benchmarks derived from the three models introduced in Section 2.1. Then, we investigate how information was used, and observe that in all treatments subjects pay much attention to opponents' actions and profits in the previous round, which is the information required by the Imitate-the-Best learning rule. Finally, we try to establish a relation between information search patterns and actual behavior by means of a learning model. Results suggest that subjects' learning process is mainly driven by a Best-Response-Dynamics, but it is also affected by a tendency to imitate the best performer.

4.1 Quantities

Table 2: Avearge quantities and prices across treatments

treatment	**player's quantity**	**competitors' quantity**	*predicted quantity*	**price**	*predicted price*
		40 rounds			
BRD	17.98	15.56	*16*	17.19	*17*
ItheB	14.72	18.46	*20*	12.69	*1*
T&E	19.05	13.01	*10*	22.93	*41*
Total	17.36	15.56		17.82	
		last 10 rounds			
BRD	19.16	15.30	*16*	16.64	*17*
ItheB	13.39	18.25	*20*	14.60	*1*
T&E	20.79	11.45	*10*	25.88	*41*
Total	17.97	14.86		19.23	

Table 2 displays the average quantities produced by the subjects and by their virtual competitors in the three treatments, first across all the 40 rounds, then just for the last 10 rounds.

First, we notice that subjects react differently when faced with different opponents. The average quantity chosen by the subjects under T&E and BRD is higher than under ItheB.[13]

If all subjects followed exactly one of the three learning rules adopted by the automata, we would observe a convergence towards the predicted equilibrium in at least one of the three treatments. Figure 4 shows that, in fact, the quantity chosen on average by the automata (represented by a solid line in the graph) is

[13] Wilcoxon rank-sum tests reject the hypothesis that observed quantities under ItheB and under BRD come from the same distribution at the 5% significance level, and the hypothesis that observed quantities under T&E and under ItheB come from the same distribution at 1% significance level. In contrast, the difference between BRD and T&E is not significant. In these and in the following tests, averages at subject level are taken as independent observations.

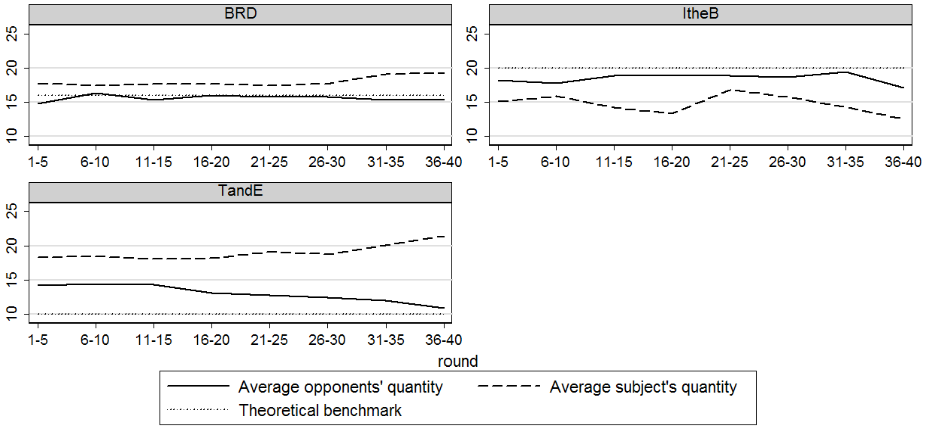


Figure 4: Average quantities chosen by subjects and by their virtual competitors.

relatively close to the theoretically predicted one (indicated with a dotted line). Yet, according to a Wilcoxon signed-rank test, only for treatment BRD the difference between actual and predicted quantity is not significant, while for the other two treatments it is significant at the 5% level. In contrast, the average quantity chosen by the subjects (represented by the dashed line in the graph) is far from the theoretical benchmark, and the distance between actual and predicted quantity increases if we look only at the last ten rounds.[14] Learning and experience seem to drive subjects away from the predicted equilibria.

Quantities chosen in this game seem to be substantially driven by the mechanical behavior adopted by the automata. What is really interesting here, instead, is the way players use the information they are provided, and how this affects their choices.

4.2 Attention

Figure 5 shows the average share of look-up time dedicated by each subject to the four pieces of information they could access during the game. The first noticeable fact is that in all treatments most of players' attention is dedicated to the plot that represents profits earned and quantities produced in the previous round by the player himself and by each of his competitors. This means, for example, that if they wanted to imitate the best performer in the previous round, as suggested by Vega-Redondo (1997), in general they knew the information necessary to do it. By contrast, Trial-and-Error is less supported by our data, because subjects do not seem to be very interested in the graph representing the series of player's own profits and quantities, which includes the only information required to apply this learning model.

From Figure 5 it also appears that in treatment ItheB the share of attention

[14] The difference is significant at the 5% level for treatments ItheB and T&E, and at the 10% level in tretament BRD, according to Wilcoxon signed rank tests.

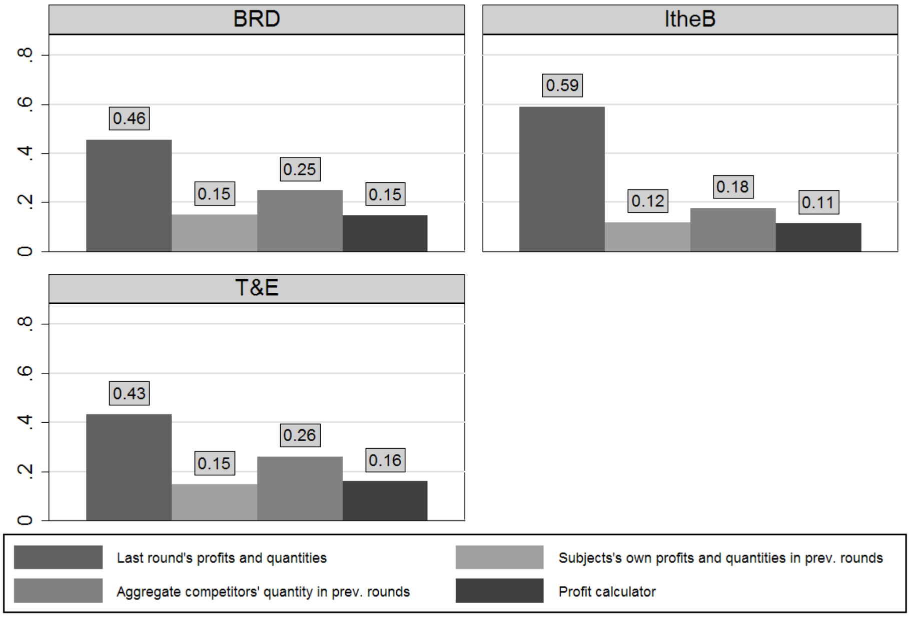


Figure 5: Distribution of players' attention in the three treatments.

allocated to what opponents have done and earned in the previous round is higher than in the other two treatments. Yet, Wilcoxon rank-sum tests do not reject the hypothesis that the share of attention allocated to the different pieces of information is the same across treatments.

Figure 6 looks at how the allocation of attention evolves along the whole experiment. There, we notice the presence of a trend which is common across treatments. In the first round, all the look up time is devoted to the profit calculator which is the only information available, probably in the attempt of figuring out how the market works. Then, from the second round on, players' attention seems to be mainly attracted by the plot displaying opponents' individual profits and quantities in the previous round, and, to a minor extent, from the graph showing the cumulated quantity chosen by the three virtual opponents in all the previous rounds. The attention dedicated by players to their own past experience remains scarce all over the game, reinforcing our skepticism about the relevance of the Trial-and-Error learning model.

4.3 Learning

Figure 7 represents the evolution of the average decision time across rounds. The sharp decrease in the decision time we observe in all treatments, together with the decrease in the use of the profit calculator, suggests that most of the learning about the market structure takes place during the first half of the experiment.

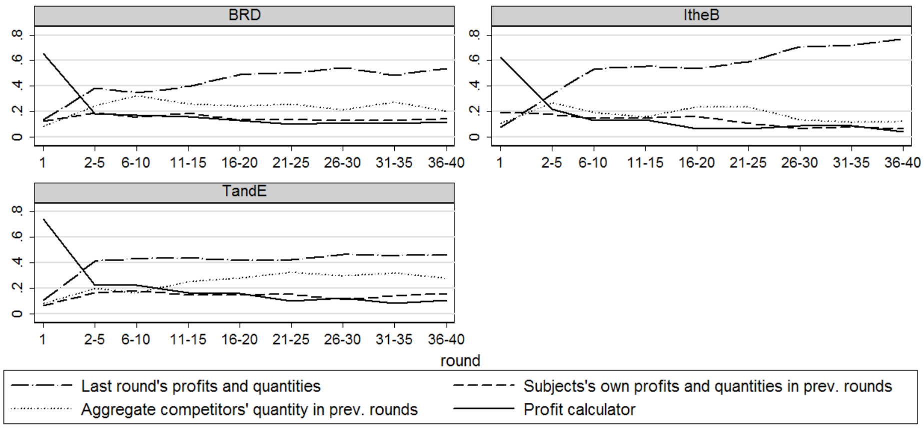


Figure 6: Allocation of players' attention along the game.

Once subjects have clearly understood the relation between quantities, prices and profits, they focus their attention on what the opponents do.

Now the question is: how do subjects use the information on opponents' behavior they collect? Do they imitate the best among their competitors, as suggested by Vega-Redondo (1997)'s model, or do they try to optimally respond to the opponents' past actions? Is any of the three learning models presented in Section 2.1 able to encompass the observed subjects' behavior? To answer these questions, we compare the explanatory power of our three learning models.

In a first attempt to have a picture of the learning model adopted by the players in this game, we used hit ratios – a measure proposed by Huck et al. (1999) – to assess to which extent the three learning rules presented in section 3.2 are able to predict each single choice subjects made.

The hit ratio is defined as:

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

where a_i^t is the quantity predicted, in turn, by Imitate-the-Best (IB), Trial-and-Error (TE) and Best-Response Dynamics (BR)[15]. Clearly, $z_i^t = 1$ indicates that the rule perfectly predicted the move taken by the player. In general, $z_i^t > 0$ implies that the rule has correctly anticipated the direction of the variation in the quantity produced, while the opposite is true when $z_i^t < 0$.

Table 3 shows that Best-Response Dynamics is the rule that better fits the data, predicting the right direction of change in quantities in at least 70% of the cases and providing a rather precise forecast ($0.5 \leq z_i^t < 1.5$) in at least 19.41% of the observations. Contrary to what observed by Huck et al. (1999),

[15] z_i^t was set equal to 1 if both the numerator and the denominator were null, it was set equal to minus the absolute value of the numerator when only the denominator was null.

Table 3: Hit ratios

Treatment	**Learning Rule**	$z < 0$	$0 <= z < 0.5$	$0.5 <= z < 1.5$	$z >= 1.5$
	Trial-and-Error	43.42	22.04	6.91	27.63
BRD	Imitate-the-Best	54.28	28.62	10.53	6.58
	Best-Response Dyn.	30.26	36.51	19.41	13.82
	Trial-and-Error	40.98	38.35	3.01	17.67
ItheB	Imitate-the-Best	33.08	43.61	15.79	7.52
	Best-Response Dyn.	20.68	46.24	31.20	1.88
	Trial-and-Error	34.21	30.26	23.03	12.50
T&E	Imitate-the-Best	50.00	44.08	4.93	0.99
	Best-Response Dyn.	26.97	42.76	23.68	6.58

Note: only observations from round 3 on are used, because for rounds 1 and 2 the Trial and Error rule cannot make any prediction.

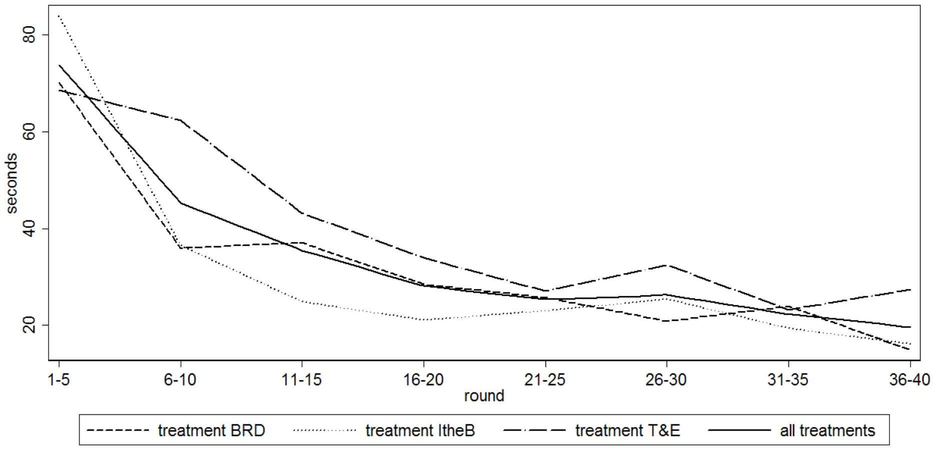


Figure 7: Average decision time across rounds, in seconds.

Imitate-the-Best is the rule with the worst performance, since in two treatment it predicts the wrong direction of change in at least half of the cases.

Table 4 shows how many subjects report positive z_i^t values in at least 70% of the rounds and how many present hits close enough to 1 ($0.5 \leq z < 1.5$) at least 30% of their decisions. Again, we observe that Best-Response Dynamics appears to be the only rule that is applied with a certain degree of consistency. No subjects seem to adopt the Imitate-the-Best rule, while two subjects behave in accordance with Trial-and-Error rule in the treatment in which this is the rule that informs the behavior of the automata.

Table 4: Hit ratios at the individual level.

Treatment	$z > 0$ in at least 70% of the rounds			$0.5 \leq z < 1.5$ in at least 30% of the rounds			N.obs.
	Trial-and-Error	Imitate-the-Best	Best-Resp.	Trial-and-Error	Imitate-the-Best	Best-Resp.	
BRD	1	0	5	0	0	1	8
ItheB	1	3	5	0	0	3	7
T&E	2	1	5	2	0	3	8

Since from the analysis presented above it emerges that none of the three learning rules is able to exactly predict players' choices, we now shift focus to the direction of the players' output decisions. Following Bosch-Domènech and Vriend (2003), we consider a model in which the sign of a player's output change, Δq, is a function of the direction, x, indicated by the target output levels according to each of the three learning rules. This way, we should be able

to determine to what extent each behavioral rule affects the way players adjust their output in every round of the game.
Let:

$$\Delta q_i^t = \begin{cases} -1 & \text{if } q_i^t - q_i^{t-1} < 0 \\ 0 & \text{if } q_i^t - q_i^{t-1} = 0 \\ 1 & \text{if } q_i^t - q_i^{t-1} > 0 \end{cases}$$

and, for every learning rule r, let

$$x_{r,i}^t = \begin{cases} -1 & \text{if } a_{r,i}^t - q_i^{t-1} < 0 \\ 0 & \text{if } a_{r,i}^t - q_i^{t-1} = 0 \\ 1 & \text{if } a_{r,i}^t - q_i^{t-1} > 0 \end{cases}$$

where $a_{r,i}^t$ denotes the quantity predicted for player i at round t by rule r, as above. In this experiment, we also observe which different pieces of information each subject looked at in every single round. These data are assembled into the model, to see whether and how information affects the way players adapt their choices as they gain experience during the experiment. Remember that in this experiment information is hidden behind four boxes on the computer screen. Denoting with b the box, we then created four dummy variables, $d_{b,i}^t$, indicating whether in round t subject i opened the box b containing (*i*) the results provided by the Profit Calculator (ProfCalc), (*ii*) information about quantities individually produced by each of the players in the last round and corresponding profits (LastRound), (*iii*) quantities produced and profits obtained by the subjects himself in all the previous rounds (HistPl) or (*iv*) the sum of quantities produced by the subjects's three opponents in each of the previous rounds (HistOpp).

We assume that the value of Δq_i^t depends on a latent response variable y. Then, we estimate an ordered-probit model, where the latent response variable, y, is a linear function of the independent variables plus a normally distributed error term, u (for simplicity, here we omit subscripts for round and individual players):

$$y = \sum_r \beta_r x_r + \sum_b \gamma_b d_b + \sum_r \sum_b \delta_{b,r} d_b x_r + u$$

Altogether, the model includes 19 explanatory variables. Data from the three treatments were pooled, to obtain a sufficiently large number of observations. We first estimated the full model, then we progressively obtained a more compact model using Likelihood-Ratio tests with a significance level of 5%. Here we only present the finale estimate (Table 5). To take into account possible effects of unobserved individual characteristics, we include a fixed effect at the individual level.

Table 5 shows some noticeable results. First, regardless of the information observed by subjects, the learning rule based on Best-Response Dynamics seems to inform their choices to a great extent, which confirms what already highlighted in Tables 3 and 4, using hit ratios. Second, this attitude towards adopting the myopic best response is counteracted by a tendency to imitate the best performer. Third, the information observed by the subjects significantly affects the way they behave. In particular, when a subject observes the quantities produced and the profits individually earned by each of her competitors in

Table 5: Fixed effects ordered probit model

	Coefficient	**Standard Error**
βs		
TE	-0.113	0.094
BR	0.938***	0.188
IB	0.749***	0.089
δs		
HistPlxTE	0.249**	0.112
HistPlxBR	-0.221**	0.104
ProfCalcxBR	0.451***	0.101
LastRoundxBR	-0.521***	0.178
cut1	-0.479	0.189
cut2	0.440	0.188
N	874	
logL	-835.568	
Wald test $\chi^2(29) =$	230.430	p-value: 0.0000

Note: symbols ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. Standard errors are robust for heteroskedasticity.

the previous round, her choices tend to be driven away from what predicted by the Best-Response rule. Similarly, when subjects look at their own history of play – in terms of output produced and profits earned – their choices are generally less consistent with a Best-Response Dynamics, whose predictive power appears instead to be enhanced when subjects use the profit calculator. Finally, the learning rule based on Trial-and-Error does not find strong support in our data: the coefficient is not significant, and the learning rule seems to drive subjects behavior only when they specifically look at their own past sequence of choices and payoffs. As we have seen in Section 4.2, though, this piece of information is often neglected by subjects.

Summing up, our results on quantities (Section 4.1), information search patterns (Section 4.2), and learning (Section 4.3) indicate that the hypothesis that subjects follow some very simple rule to choose their strategy in our game should be rejected. Learning through Trial-and-Error does not seem to be a plausible explanation of subjects' behavior, both because players pay too little attention to their own past profits and quantities, which is the only information required to apply this learning rule, and because their choices are not in line with what is theoretically predicted according to this model. On the other hand, the Imitate-the-Best rule – *per se* – is not able to forecast the observed choices correctly, even if subjects' look-up patterns are consistent with this learning model. Myopic best-response seem to drive players' choices, at least partially. This conjecture is supported by the information subjects acquire, on average. To apply this learning rule, subjects need to know the sum of the quantities produced by their competitors in the last round – an information they almost always look at – and they must be able to compute a best-response, which means that either they use the profit calculator or they have used it extensively in the past and already know their best-response function. Nonetheless, Best-Response-Dynamics does not fully explain the observed variations in players' behavior. Results from our ordered probit regression (Table 5) indicate that players' behavior is driven by the interplay of best-response and imitation. Even if subject are incline to adopt the best-response when they know the market structure sufficiently well, if they are provided with information about their rivals' strategies and choices, they are tempted to imitate those who are more successful, which yields more competitive outcomes.

5 Conclusion

From the data collected in this experiment, it emerges that players' behavior cannot be encompassed by any of the models of learning we shortly described in section 3.2, alone. Moreover, our results confirm that information provided to subjects has an important effect on the way they behave. We observe that players tend to best-respond to the action taken by their opponents in the last round, when they have the necessary information to do so. Still, this is not the information they are most interested in: they dedicate most of their attention to the quantities individually chosen by their opponents and the profits each of them earned in the previous round. This piece of information seems to drive them away from best-response, and possibly leads them to a more "imitative" behavior. Even if imitation is not the driving force of subjects' learning – so the market outcomes we observe are far away from those predicted by Vega-Redondo (1997) – it still leads to a more aggressive competition than the one that would emerge if all players adopted a learning model only based on myopic best-response. Finally, according to our data, it seems that the learning rule adopted by the opponent – extreme as it was – does not have a very strong impact on the model of information acquisition and processing adopted by the players.

Appendix

Instructions

Thank you for taking part in our experiment.

Read carefully all the instructions; if something is not clear, please let us know by raising your hand. From now on, you are requested not to communicate with other participants in any way.

Your task

During this experiment, you will be asked to act as the owner of a firm which produces and sells a given product: your task consists in deciding how many units of your product to supply to the market.

Your firm has **three competitors** that sell on the same market a product exactly identical to yours.

The experiment consists in **40 consecutive periods**. In every period, you will be asked to choose how many units to produce **(between 0 and 30)**, and the same will be done by your competitors. Your choices affect both your firm's profits and the profits of the other three firms.

Price, costs and profits

The **market price** at which you will sell your product will be higher, the smaller the total number of products your firm and your competitors put on the market; if the total number of units sold on the market is very high, the price will be equal to zero.
No unit remains unsold: the whole production will be purchased by consumers at the market price.

Your **total production costs** will be larger, the higher the number of units you supply to the market.

Your **profit** will be equal to the market price times the number of units you sell, minus production costs.

Earnings and Payment

You will receive an **initial capital** of 2000 points.
As the experiment proceeds, your per-period profits and losses will be added to your capital.
Your cumulated capital will be displayed in the top right corner of the screen.

Your goal is to **maximize your capital.**

100 points correspond to 1 SEK.

At the end of the game, your capital will be converted in SEK, and will be payed to you privately.
In addition, you will be payed a **show up fee** of 50 SEK.

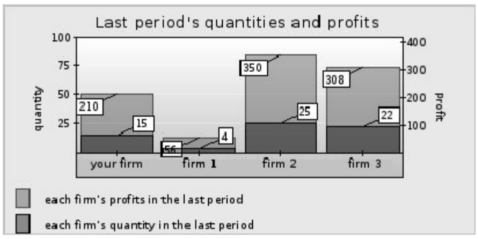
Next Page

Information at your disposal

Before choosing how much to produce, you will be given the opportunity to look at some plots providing information on market characteristics and on what happened in the previous periods.

Examples of these plots are presented in what follows.

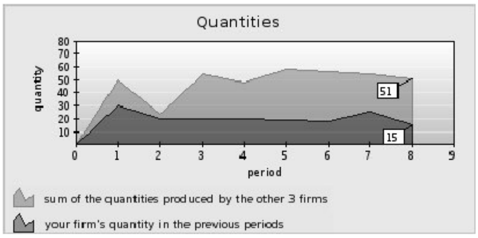
Profits and quantities of each firm in the last period



This bar plot represents profits and quantities of **each of the four firms in the last period**.

The plot in the example tells us that, in the last period, firm 2 produced 25 units, gaining a profit equal to 350.

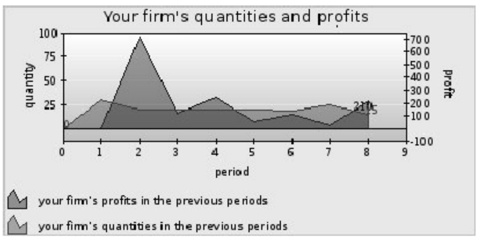
Quantities produced by your firm and by your competitors in all the previous periods



The second plot represents the **quantity** produced by your firm, and the **sum of the quantities** produced by the other three firms during **all the previous periods**.

In our example, the plot shows that in the last period, the sum of the quantities produced by the other 3 firms is equal to 51 and that it slightly diminishes during the last four periods.

Profits and quantities of your firm in all the previous periods



The third chart presents profits earned and quantities produced by **your firm** in **all the previous periods**.

The chart in the example reveals that in the second period you earned more than 700 and that from that period on your profits were always positive, but never higher than 300.

Next Page

The profit calculator - 1

A profit calculator is also provided to you, to help you making your choices.

The profit calculator is an instrument you can use to better understand how the market works. It has two functions.

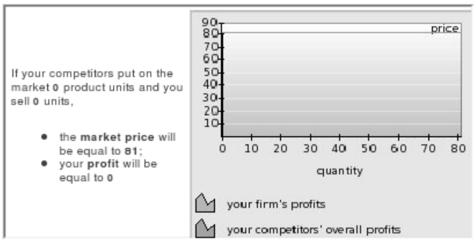
First function of the profit calculator

The profit calculator can tell you how much you would earn if you produced a number *x* of units and the sum of the units produced by your competitors were equal to a given number *y*. To activate this function you have to answer both the questions it asks, then you have to press first "Submit", then "Update" and look at the result in the box.

Here is an **example** of the profit calculator.

Profit Calculator				
How much do you want to produce?		How many units do you think your competitors are going to produce?		Submit

Update



Before going on with the instructions, try to use the profit calculator and answer the following question.

What would your profit be if you produced 23 units and the sum of the units produced by the other three firms is equal to 51? (You have to enter an integer number)

OK

The profit calculator - 2

Second function of the profit calculator

The profit calculator can also evaluate for you the number of units you should produce to maximize your profit *in the present period*, given the sum of the units *y* produced by the other 3 firms. To activate this second function you have to enter *only* the number of units you think your competitors supply to the market *but not the number of units produced by your firm,* then you have to press "Submit" and "Update".

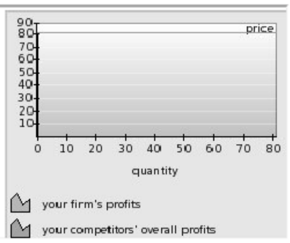
Here is, again, an **example** of the profit calculator.

Profit Calculator				
How much do you want to produce?		How many units do you think your competitors are going to produce?		Submit

Update

If your competitors put on the market **0** product units and you sell **0** units,

- the **market price** will be equal to **81**;
- your **profit** will be equal to **0**



Before going on with the instructions, try to use the profit calculator and answer the following question.

How many units should you produce to maximize your profits in this period, if the sum of the units produced by the other three firms is equal to 51? (You have to enter an integer number)

OK

How to consult the plots

You will not be able to look at the four plots at the same time..
In fact, they are hidden behind four windows like the ones displayed in this page; to open each window and see its content you just have to put your mouse cursor over it.
Now you can test this mechanism.

Profit calculator				
How much do you want to produce?		How many units do you think your competitors are going to produce?		Submit

Update

Profit Calculator	Quantities produced by your firm and by your competitors in all the previous periods
Profits and quantities of each firm in the last period	Profits and quantities of your firm in all the previous periods

To go on you have to correctly answer the following question.

In which **period** the sum of the quantities of the other 3 firms has reached its maximum? (You have to enter an integer number)

OK

References

Bigoni, M. (2009, July). Information and Learning in Oligopoly: An Experiment. Available at SSRN: http://ssrn.com/abstract=1134750.

Bosch-Domènech, A. and N. J. Vriend (2003). Imitation of Successful Behaviour in Cournot Markets. *Economic Journal 113*(487), 495–524.

Costa-Gomes, M. and V. Crawford (2006). Cognition and Behavior in Two-Person Guessing Games: An Experimental Study. *The American Economic Review 96*(5), 1737–1768.

Costa-Gomes, M., V. P. Crawford, and B. Broseta (2001, September). Cognition and Behavior in Normal-Form Games: An Experimental Study. *Econometrica 69*(5), 1193–1235.

Cournot, A. (1838). *Recherches sur les principes mathématiques de la théorie des richesses*. Engl. transl. by N. T. Bacon.

Duersch, P., A. Kolb, J. Oechssler, and B. Schipper (2005). Rage against the machines: How subjects learn to play against computers.

Gabaix, X., D. Laibson, G. Moloche, and S. Weinberg (2006, September). Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model. *American Economic Review 96*(4), 1043–1068.

Huck, S., H.-T. Normann, and J. Oechssler (1999). Learning in Cournot Oligopoly–An Experiment. *Economic Journal 109*(454), C80–95.

Huck, S., H.-T. Normann, and J. Oechssler (2000, May). Trial & Error to Collusion - The Discrete Case. Bonn Econ. Discussion Papers.

Huck, S., H.-T. Normann, and J. Oechssler (2004, Feb.). Through Trial and Error to Collusion. *International Economic Review 45*(1), 205–224.

Johnson, E. J., C. Camerer, S. Sen, and T. Rymon (2002, May). Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining. *Journal of Economic Theory 104*(1), 16–47.

Johnson, E. J., D. A. Schkade, and J. R. Bettman (1988). Monitoring Information Processing and Decision: The Mouselab System. working paper, Graduate School of Industrial Administration, Carnegie-Mellon University.

Kandori, M., G. J. Mailath, and R. Rob (1993, Jan.). Learning, mutation, and long run equilibria in games. *Econometrica 61*(1), 29–56.

Normann, H.-T. and B. Wallace (2006, Dec.). The Impact of the Termination Rule on Cooperation in a Prisoner's Dilemma Experiment. SSRN eLibrary. http://ssrn.com/paper=952953.

Offerman, T., J. Potters, and J. Sonnemans (2002, October). Imitation and Belief Learning in an Oligopoly Experiment. *Review of Economic Studies 69*(4), 973–97.

Peyton Young, H. (1993, January). The Evolution of Conventions. *Econometrica 61*(1), 57–84.

Rassenti, S., S. S. Reynolds, V. L. Smith, and F. Szidarovszky (2000, February). Adaptation and convergence of behavior in repeated experimental

Cournot games. *Journal of Economic Behavior & Organization 41*(2), 117–146.

Schlag, K. H. (1998, Jan.). Why imitate, and if so, how?, : A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory 78*(1), 130 – 156.

Selten, R., K. Abbink, and R. Cox (2005, Apr.). Learning direction theory and the winner's curse. *Experimental Economics 8*(1), 5–20.

Selten, R., M. Mitzkewitz, and G. R. Uhlich (1997). Duopoly Strategies Programmed by Experienced Players. *Econometrica 65*(3), 517–556.

Selten, R. and R. Stoecker (1986). End behavior in sequences of finite prisoner's dilemma supergames a learning theory approach. *Journal of Economic Behavior & Organization 7*(1), 47 – 70.

Theocharis, R. (1960, February). On the Stability of the Cournot Solution on the Oligopoly Problem. *Review of Economic Studies 73*(2), 133–134.

Vega-Redondo, F. (1997). The Evolution of Walrasian Behavior. *Econometrica 65*(2), 375–384.