Information and Learning in Oligopoly: an Experiment

Maria Bigoni

May 18, 2007

1 Introduction

The experimental study I propose concerns the relation between the process of information search and players' behavior in a repeated Cournot oligopoly. The main question I will try to answer is what happens when information acquiring and processing is too difficult or too costly for the firm to behave according to the perfect rationality paradigm. I will investigate what pieces of information they look for and which heuristic they adopt, in order to understand how information may affect market behavior.
The topic is not particularly new: the interest on it reached its apex after 1997, when in an article appeared in Econometrica, Vega-Redondo [15] proposed a theoretical model of behavior of Cournot oligopolists which leads to surprising conclusions. According to the author's theory, if firms tend to imitate the behavior that proved most successful in the previous period (that is: they produce the level of output that yielded the highest profit) but with positive probability experiment other strategies, Walrasian behavior can emerge in the long run within any Cournot oligopoly with homogeneous goods. In a number of following works [1] Vega Redondo's theory has been experimentally tested and compared with other learning theories that make different assumptions about players' information and lead to different behaviors and market equilibria.
Despite the efforts made during those years, I do not believe that a final conclusion has been reached, rather, it seems that the attention has been averted from this topic without having provided a definitive answer to the questions it brings up.

The experiments I present here are not just replications of what has been done before, since I introduce a technical innovation – consisting in the use of the software "MouselabWEB" to design the experiments and to control for the information the agents really use. This software, developed by Martijn C. Willemsen and Eric J. Johnson[2], uses a common web interface to present the problem to the participants, allowing the experimenter to provide the players with information about the game and the previous moves of their opponents in a very straightforward and intuitive way. But what is really important about this new instrument is that it makes it possible for the experimenter to verify which information the decision makers look at and how long.
Paying attention not only to what players do but also to what they know, it is possible to better understand the mental mechanisms which guide their choices and consequently the impact that the informational framework has over their behavior.

I believe that the results of my experiments can contribute to the debate which has developed after the article by Vega-Redondo [15], because the technique I adopt enable me to investigate more deeply the relation between information and behavior.
Moreover, with my experiment I want to test MouselabWEB as an experimental device that to my knowledge has barely been used in experimental economics and that could be effectively adopted to investigate other interesting topics, such as the process of information acquisition in auctions and phenomena like informational cascades and herding behavior in financial markets.

[1] see for example Huck *et al.* 1999 [6], Rassenti *et al.* 2000 [12], Offerman *et al.* 2002 [11] and Bosh-Domènech and Vriend 2003 [1]

[2] See the website: http://www.mouselabweb.org/index.html

2 Related Literature

2.1 Information and Learning in Oligopoly Experiments

Between 1999 and 2002, four articles were published, which presented experiments regarding information and learning in a Cournot oligopoly setting. In these works the same experiment is repeated under different treatments, varying the quality and quantity of information provided to the subjects. The authors then compare the actual behavior observed in the different treatments and make inference about the impact that the various informational frameworks have on players' choices. Nonetheless a number of details changes from one experiment to the other, and maybe this is the first reason why the results obtained by the authors are not at all unanimous, nor they are conclusive: for example, the experiments performed by Huck, Normann and Oechssler [6] and by Offerman, Potters and Sonnemans [11] provide a rather strong support to the theory proposed by Vega Redondo, mentioned before, while the works presented by Rassenti et *al* [12] and by Bosch-Domènech and Vriend evidence no trend towards the Walrasian equilibrium and do not find any clear indication that players tend to imitate the one who got the best performance in the previous period.

Huck et *al.*'s experiments (HNO from now on) study a 40-periods Cournot market with linear demand and cost, in which four symmetric firms produce a homogeneous good. Across their five treatments, they vary the information they provide to the subjects, both about market and about what other players in the same market do. In particular, information about market can be:

Complete: the participants are informed about the symmetric demand and cost functions in plain words and they are provided with a 'profit calculator', which can compute market price and firm's profit when one enters the total output of other firms and his own output, and can also suggest to the subject the quantity which would yield him the highest payoff given the hypothetical total quantity produced by the competitors.

Absent: participants do not know anything about the demand and cost conditions in the market nor do the instructions explicitly state these would remain constant over time; all they know is that they would act on a market with four sellers and that their decisions represent quantities.

Partial: participants are just told that market conditions remain constant for all periods and coarsely informed about demand and profit functions.

In three of the treatments, participants are also informed about competitors' individual quantities and profits in the previous period, while in the remaining two treatments they are told only the total quantity the others have actually supplied. HNO find significant differences in individual and aggregate behavior across the treatments, and collect data suggesting that increasing information about the market decreases total quantity, while providing additional information about individual quantities and profits increases total quantity. HNO also test other learning theories besides the one proposed by Vega Redondo, and they find that when subjects know the true market structure, their quantity adjustments depend significantly on the myopic best reply to the quantity produced by their competitors in the very last period. In general, though, none of the theoretical learning models they consider, *per se*, seems to fully explain the observed behavior.

Offerman et *al.* [11] conducted a similar computerized experiment, obtaining results which are consistent and complementary to those presented by HNO. In their setting, a triopoly with non-linear demand and cost functions is repeated 100 times, with complete information about market. The authors study how players' behavior changes across three treatments, which differ for the amount of information provided to the subjects about individual quantities and revenues of the other two competitors in their market.

In one treatment ($Qq\pi$) firms were provided with individualized information about the quantities and the corresponding profits to the other two firms; in a second treatment (Qq) they were just told the quantities produced by the opponents, but not their profits, and in the last treatment (Q) firm were only informed of the total quantity produced it their market. As HNO, they observed a substantial difference between the treatments and the data they collected evidence that the feedback information provided to the subjects affects the behavioral rules they adopt. Moreover, in agreement with what reported by HNO, also in this study the Walrasian outcome is only reached quite often in treatment $Qq\pi$, where the players are informed about their opponents'profits. On the other hand, they observe that the collusive outcome seems to be a stable rest point only in treatment Qq and $Qq\pi$, but not in the treatment with no information about others' individual quantities and profits, in which the only rest-point is represented by the Cournot-Nash equilibrium.

The experiment performed by Bosch-Domènech and Vriend (BDV) differs from the previous two both for the setting and for the aims. While HNO and OPS compare the prognostic capability of different learning rules which lead to different theoretical outcomes, here the authors focus specifically on Vega-Redondo's behavioral rule and the main purpose is to investigate whether people are inclined to imitate successful behavior and, in particular, whether this behavior is more prevalent in a more demanding environment. The authors study a series of 22-periods Cournot duopolies and triopolies with homogeneous commodity and linear demand and cost functions. They examine six treatments altogether: for both duopolies and triopolies they consider three different treatments that differ in the way the information is provided and in the time pressure put on the players.
In the treatment denominated "easy", the players are given a profit table that conveniently summarizes all the information concerning the inverse demand curve and the cost function, and there is no time pressure on the players. After each period, each player gets information about the actions of each of the other players in the same market, but not about their profits.
In the 'hard' and 'hardest' treatments, players have just one minute to decide on their output level; after each period they receive feedback information both about the actions of all players and about the profits obtained by each of them, and the output decision which led to the highest profit is highlighted.
In the 'hard' version, the players get an inconveniently arranged enumeration of the market prices associated with all possible aggregate output levels and of all possible cost levels. The 'hardest' version differs from the 'hard' treatment in that the information about the demand side of the market is limited to the statement that 'the price level depends on aggregate output'.
The purpose of the 'hard' and 'hardest' treatments is to explore to what extent imitation is influenced by the bounds imposed on the subjects' choice capabilities and to check if it is actually more prevalent when the task of learning about the market becomes more difficult and at the same time the decision of the most successful firm is displayed more prominently, and the answer they give to this question is essentially negative. The data they collected show that as the learning-about-the-environment task becomes more complex, average output increases, but the Walrasian output does not seem to be a good description of the output levels observed in the experiment and if anything, imitation of successful behavior tends to decrease rather than to increase when moving to more complicated environments.

The fourth experiment has been conducted by Rassenti *et al.* (RRSZ); it represents an oligopoly with homogeneous product, in which five firms interact repeatedly for 75 periods, with fixed payoff conditions. The setting exhibits a substantial difference from the previous three since in this case the cost functions – linear, with constant marginal costs and no fixed costs – are private information and are different among the firms. The demand function is linear, and is public information among the players.
The authors perform two different treatments: one in which subjects were able to observe the past output choices of each one of their rivals, the other in which they are informed only about the past total output of rivals.
They use their experimental results to test a number of learning models – such as best response dynamics, fictitious play and adaptive learning. None of these models receives strong support from the data they collected: the observation of actual movement of total output over time appears to be inconsistent with both the best response dynamic and with fictitious play, for most experiments. Moreover the authors show that their data do not provide any evidence neither in support for learning models based on imitation, nor for the more traditional hypothesis that information about competitors

enhance the potential for collusion, because the treatment conditions involving provision of information about rivals' outputs and prior experience do not seem to have a significant effect on total output levels. The evidence relative to individual behavior is mixed, and no predominant models of learning emerge; the most prominent result is that in general observed behavior for individual subject sellers is not converging to the static Nash equilibrium predictions for individual output choices in these experiments.

In light of these results, I would conclude that it is worthwhile going on investigating on information and learning in oligopolistic markets, because the topic is interesting from a theoretical point of view and it also has interesting practical implications, but a theory consistent with experimental data is still far from being definitely developed. For this reason I decided to design an experiment which is similar to the four previously mentioned under many respects but introduces the use of MouselabWEB (a recent development of MouseLab). This software allow the experimenter to monitor the information acquisition process through a computer interface: relevant pieces of information are hidden behind a number of boxes on the screen and to access them the decision maker has to open the boxes and look at their content. He can open just one box at a time, and by recording the number and the duration of the look-ups the program provides precious information about the decision makers' learning process. This technical innovation will hopefully provide a deeper insight into the problems under analysis.

2.2 Other Experiments using MouseLAB

One of the most famous experiments using MouseLab has been performed by Johnson, Camerer, Sen and Rymon(JCSR) [9]: in this work the information acquisition process is observed with the aim of testing the game theoretic assumption of backwards induction. The subjects were asked to play eight three-round alternating-offer bargaining games, with a different anonymous opponent each time. In the first round one of the two players makes an offer to his opponent about how to share a given amount of money; if the other player accepts, than the game is concluded, otherwise he will have to make a counteroffer about how to share a new pie, smaller than the first one. Again, if the first player accepts, the game is over and each of them gets his part as established in the agreement; on the other hand, if the first player rejects the offer the pie shrinks again and he will have the opportunity to make one last offer to his opponent. If even this offer is rejected, nobody gets anything. The sizes of the three pies are represented on the computer screen in front of each player, but they are hidden under three boxes that can be open only one at a time, simply by putting the mouse' cursor over the box itself. The box will stay open until the mouse is moved somewhere else.
The authors observe three measures of information search: the number of times each box is opened in a period, the total time each box stays opened in a period and the number of transitions from one specific box to another. They note that most of the looking time is spent looking at the first round pie size and contrary to the backward induction prediction there are always more forward predictions than backward ones. From the data collected through these experiments they conclude that people do not use backwards induction instinctively, even if an additional treatment in which players are previously trained to use backward induction shows that people are able to learn it when appropriately instructed.
Another interesting result they get is that there is a strong correlation between differences in information processing and differences in players' behavior. This and the other results presented in this paper testify that measuring attention directly can effectively contribute to the comprehension of both failure and successes of the game theoretic predictions and help to understand the role of information and learning in influencing the outcomes of different games.

Another seminal study on the information acquisition process has been done by Costa-Gomes, Crawford and Broseta [3](CGCB). They asked the subjects to play 18 two-players normal form games, with different anonymous partners. The payoff tables are hidden and MouseLAB is used to present them: for every combination of strategies, subjects can look up their own or their partner's payoff as many times as they want, but they can only see one of these numbers at a time. Till the end of the series of games, no feedback was provided to the agents, in order to suppress learning and repeated

game effects as much as possible.
In Johnson *et al.* the goal was to test a specific theory of behavior – namely backward induction. On the contrary, here the authors compare nine different decision rules (or types) and try to make inference about which one is more likely to inform players' behavior. As in JCSR, they assume that each decision rule determines both a player's information search and his decision once he gets the information he was looking for. Therefore, by observing both the information acquisition process performed by the agents and the choices they actually make when playing the games, it is possible to deduce what decision rule they adopt.
This study confirms the presence of a systematic relationship between subjects' deviations from search pattern associated with equilibrium analysis and their deviations from equilibrium decisions. Besides, according to Costa-Gomes *et al.*'s analysis most of the subjects are much less sophisticated than game theory assumes: between 67% and 89% of the population belong to two types, namely to the *Naïve* type, who best responds to beliefs that assign equal probabilities to each of their partner's possible strategies and to *L2* type, who best replies to *Naïve* subjects.

More recently, MouseLab has been used again in two experiments that provide further evidence about how the study of the information acquisition process can be useful to understand what behavioral rules and heuristics are adopted by subjects who display out of equilibrium choices.
One experiment has been conducted by Costa-Gomes and Crawford [4](CGC) and has the same theoretical and econometric framework of CGCB but it differs for the class of games submitted to the subjects. In this case, the participants are requested to play 16 different two-person guessing games, with anonymous partners and no feedback till the end of the series. The games have been designed so that the space of possible behaviors is wide and there is a strong separation of the guesses and searches implied by the different decision rules they analyze. The results they obtain are consistent with those presented in CGCB [3], but they are significantly sharper: many subjects can be easily attributed to a particular type only by their guesses, and most of the others can be identified via an econometric and specification analysis keeping into account also their information search pattern.

Another interesting application of MouseLab has been recently presented by Gabaix, Laibson, Moloche and Weinberg [5], who experimentally evaluate the *directed cognition model*: a bounded rationality model that assumes that at each decision point, agents act as if their next search operations were their last opportunity for search. Likewise the the other three experiments, the authors register the search pattern actually adopted by the subjects in two experiments and they compare it with with what is predicted by the *directed cognition model* and by the optimal search model (i.e. the Gittins-Weitzman algorithm), traditionally adopted in economics.
In the first experiment they ask the participants to choose among three projects whose outcome is uncertain, but can be discovered at a given cost. In the second experiment the subjects are requested to solve a highly complex choice problem in which the classical optimal choice model is analytically and computationally intractable: they have to choose one out of eight goods which each have nine attributes that could be discovered by opening different boxes on the computer screen. The players cannot collect all the information about the goods, because in this game time is a scarce resource. Individual information acquisition processes are recorded through the MouseLab interface, and the data collected this way reveal that the directed cognition model successfully predicts the empirical regularities observable in subjects' behavior.

The four experiments mentioned in this section evidence how the study of the information acquisition process is complementary to the observation of subjects' actual choices which traditionally constitutes the empirical basis for testing models of decision making or trying to develop new ones.

3 Experimental Design

The experiment will be repeated under three treatments. I shall present first the baseline treatment (BASE), then introduce the other two treatments emphasizing just the elements that make them different from the baseline.

3.1 Market Characteristics

The market environment I have chosen for my experiments is similar to the one proposed by HNO [6]; if possible it is even simpler. In all the sessions and treatments, the setting remains the same. Four identical firms compete à la Cournot in the same market for 40 consecutive periods. Their product is perfectly homogeneous. In every period t each firm i chooses its own output q_i^t from the discrete set $\Gamma = \{0, 1, ..., 30\}$, which is the same for every firm. The choice is simultaneous. Let q_{-i}^t denote the quantity produced in the same period by the other three firms.
Price p^t in period 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 period t will be denoted by

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

The shape of these functions has been chosen so that the three main theoretical outcomes – namely collusive, Cournot and Walrasian outcomes – are well separated one from the other and belong to the choice set Γ. More precisely, collusive equilibrium is denoted by $q^M = (10, 10, 10, 10)$, Cournot equilibrium is $q^C = (16, 16, 16, 16)$ and Walrasian equilibrium is $q^W = (20, 20, 20, 20)$.

3.2 Information Provided to the Subjects

Participants know how many competitors they have (anonymity is nonetheless guaranteed). 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 production costs increase with the number of goods it decides to sell. Besides, players are told that per-period profit is given by market price times the number of goods sold by the firm, minus production costs.
(see the instructions in Appendix B).

Subjects are also endowed with a *profit calculator* similar to the one proposed by Huck *et al.* [6]. 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 profits the subject would earn; if the player just fill 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 if he produced the suggested amount of good. The answers provided by the profit calculator are always displayed both graphically and textually (Figure 4 and 5). The software I developed for this experiment records how many times the subject uses the profit calculator and every trial he does.

The number of rounds is common knowledge among the subjects. According to game-theoretic predictions, cooperation should be sustainable only if our stage game were repeated in(de)finitely many times, but according to Selten *et al.* [13]

> Infinite supergames cannot be played in the laboratory. Attempts to approximate the strategic situation of an infinite game by the device of a supposedly fixed stopping probability are unsatisfactory since a play cannot be continued beyond the maximum time available. The stopping probability cannot remain fixed but must become one eventually.

In light of this consideration and of the results obtained by Normann and Wallace [10] – who show that the termination rule does not have a significant effect on players' behavior except for an end effect – I decided to adopt a commonly known finite horizon, for sake of transparency and practicality.

After the first round, each player has the opportunity to look at three plots summing up information about what has happened in the previous periods (Figure 6). The first graph is a bar-plot showing

the quantity produced by each of the four firms in the market in the previous period, and the relative profit. The second 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 periods, since the game began. The last plot shows the quantity and the profit obtained by the player's firm in each of the previous periods.
The 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. As mentioned before, MouselabWEB automatically records subjects' look-ups sequences and look-ups durations.
Besides these four boxes, on the computer screen there is a counter showing the running cumulative profits earned by the player since the game began, and a timer displaying how long it is since the current round started, that is how long has the subject been thinking about what choice to make next. Figure 8 shows how subjects' computer screen looks like.
After the last round the participants are shown their overall profit, compared with those of their three opponents (Figure 7).

3.3 Treatments and Sessions

The same experiment will be repeated under other two treatments: treatment TIME in which subjects are forced to make their choice under time-pressure, and treatment COSTLY in which subjects have to pay in order to access the different pieces of information they may be interested in. More specifically, treatment TIME differs from BASE only in that the players must decide how much to produce in at most 20 seconds, otherwise their output will be automatically set equal to 0. In treatment COSTLY, the player has to pay a certain amount of money every time he opens a box to discover the information it hides.

I expect that the increase in time pressure on the players' decision process (treatment TIME) or the introduction of a price for the access to information (treatment COSTLY) induce the subjects to reveal more clearly the heuristic they adopt and the information they look for.

Depending on the size of the lab, I will ask 16 or 20 subjects to take part to each session, so to have 4 or 5 markets per session, and I would like to do at least one session per treatment.
At the beginning of each session the participants are disposed in the lab so that they cannot communicate with each other. Instructions are written on a page that appears on the computer screen of each subject, but common knowledge of the information they contain is ensured by telling the participants that the pages they are reading are perfectly identical. Instructions are divided in several parts and at the end of each of them an understanding test is submitted to the reader, who has to answer correctly to go on to the next page.
When a player finishes reading the instructions he can start the game with the first round. The groups of four persons that represent the markets are made up at random, so it is virtually impossible for the players to know the identity of their opponents. After the last round of the game, subjects will have to answer a short questionnaire in which they are asked some questions about their strategies; then participants are called one by one in private and paid according to their total profits.

4 A Preliminary Experiment

The experiment described above has been designed to study the subjects' learning mechanisms and the relation between the observed information and the action taken. A first problem might be that both the way people behave and the way they learn might be affected by what their peers (namely the opponents, in our experiment) do. For this reason, the analysis of the results could be very complicated, because the individual characteristics of the players could affect the dynamics that emerge in every single group in different ways. Therefore, I decided to start my inquiry with a simpler experiment in which subjects play against three "virtual" players enacted by the computer and programmed to follow a specific learning rule. This way I can control for the effect of the opponents' behavior on the players' choices.

4.1 Experimental Design and Procedures

The experimental design is exactly the same as for the experiment presented in the previous section, except for the fact that subjects are informed that their opponents are "robots", that is: they are enacted by the computer. Subjects also know that these "robots" do not play at random but choose according to some rule, but nonetheless they do not necessarily choose the same output.

4.1.1 Treatments

This experiment has been run under three different treatments (T&E, BRD and ItheB), which differ only by the learning rule adopted by the computer. Since this experiment has to be a benchmark, I have chosen three extremely simple learning rules, that will be shortly described in what follows.

Trial and Error This model of learning has been proposed by Huck et al. [8, 7], and is the model with the most lax hypotheses about information: it just requires that the firms know their own past actions and their own profits. The learning rule 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.

More precisely, I programmed the computer so that in treatment T&E each "robot" player 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 direction of change s_i^t is given by

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

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 direction of change is randomly chosen among the values $-1, 0, 1$, each having equal probability.

The dynamic process defined by this learning rule converges only if some tremble is introduced: therefore, with probability $\epsilon = 0.05$ each "robot" chooses an arbitrary direction of change s_i^t.

Huck et al. [8, 7] prove that for the symmetric Cournot duopoly case this process yields collusion if the learning rule is adopted by both the firms, the cost function is weakly convex and market conditions are such that there exists only one symmetric situation in which joint profits are maximized.

I managed to prove (though not analytically, but only by means of simulations) that also within the framework adopted in my experiment, this learning rule defines a Markov process whose stationary stable distributions (which are more than one) assign positive probability exclusively to states which are in a neighborhood of the joint profit maximizing equilibrium.

The intuition is rather clear for when all firms start from an identical level of output. The question arises why firms that start from arbitrary initial quantities could become perfectly aligned. Suppose that two firms with different quantities move downwards. They will continue to do so until at least one firm's profit decreases and it will always be the firm with the smaller output to be the first. This is so because the firm selling the higher quantity gains more from the increase in price. Thus, while the smaller firm already moves upward, the other firm continues to move downward thereby decreasing the distance between the firms. Similarly, when moving upward the firm with higher output will be the first to experience losses and to change direction. Roughly speaking, there is a general tendency to equalize quantities.

Best Response Dynamics This model theorize that in every period each player myopically chooses his output as a best reply to the sum of the quantities produced by the other three in the previous period.
More precisely, the best reply correspondence for player i maps q_{-i}^{t-1} to the set

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

Under the market structure characterizing my experiment, due to the discreteness of the choice set, we have:

$$BR_i^t = \begin{cases} \{0\}, \text{ if } q_{-i}^{t-1} \geq 80 \\ \min\left\{30, \frac{80-q_{-i}^{t-1}}{2}\right\} \text{ if } q_{-i}^{t-1} < 80 \text{ and } q_{-i}^{t-1} \text{ is even} \\ 30 \text{ if } \frac{80-q_{-i}^{t-1}}{2} - 0.5 > 30 \text{ and } q_{-i}^{t-1}\text{is odd} \\ \left\{\frac{80-q_{-i}^{t-1}}{2} - 0.5, \frac{80-q_{-i}^{t-1}}{2} + 0.5\right\} \text{ otherwise.} \end{cases}$$

In this last case, I assume that the player chooses the closest integer to $\frac{80-q_{-i}^{t-1}}{2}$.

The best reply dynamic defined this way yields a Markov chain which does not necessarily converge to a stable equilibrium, consistently with what has been shown by Theocharis [14] for the case in which quantities are chosen in a continuous space.
To catch an intuition of this result, suppose for example that the system reaches one of the two states of the absorbing set $s = \{(30, 30, 30, 30), (0, 0, 0, 0)\}$: once this has happened, the system will keep on oscillating between this two states and will never be able to escape the set.
HNO[6] state the following theorem:

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

Within the framework considered here, this implies that the learning process brings the system to converge to the state $\omega^N = (16, 16, 16, 16)$. To apply this result to my experiment, I introduced some degree of inertia into the learning rule guiding the behavior of the "robot" players in treatment BRD: with independent probabilities equal to 0.05 in every round each of them chooses $q_i^t = q_i^{t-1}$, and otherwise follows the myopic best response dynamic.

Imitate the Best In the last treatment (ItheB) the "robot" players behave according to the learning model firstly proposed by Vega-Redondo [15]. The core of the model is represented by the *imitation dynamic*: a discrete time dynamic which assumes that at every time t each firm chooses its output q_i^t from the set:

$$B^{t-1} = \{q \in \Gamma : \exists j \in I \text{ s.t. } q_j^{t-1} = q \text{ and } \pi_j^{t-1} \geq \pi_i^{t-1} \ \forall i \in I, \, i \neq j\}$$

where, in our case, $\Gamma = \{0, 1, \ldots, 29, 30\}$.
This learning process, when applied to the specific context of our fictitious market, defines a Markov chain over the state space $\Omega = \Gamma^4$. Let ω_q stand for the *monomorphic state* $(q, q, ..., q)$ in which every firm chooses the same quantity $q \in \Gamma$. It is easy to verify that $\forall q \in \Gamma$ the *monomorphic state* ω_q is absorbing and that all the non-monomorphic states are transient. Therefore, the process has a number of recurrent sets equal to the cardinality of Γ, and there is a stationary distribution μ_q corresponding to each of them, which puts probability one over ω_q. Thus, the long run behavior of the evolutionary process consisting only in the imitation dynamic displays a large potential multiplicity, since it can rest forever in any monomorphic state.

To investigate the robustness of each of these multiple outcomes, Vega Redondo introduce a perturbation into the process, assuming that in every period t each firm sets its quantity according to the imitation rule with probability $1 - \epsilon$, while with probability ϵ it departs from the rule and chooses its quantity according to a distribution with full support over Γ. The interpretation here can be that with small probability every firm makes an error or it experiments a different strategy. This *perturbed process* defines a Markov chain irreducible and ergodic – since each state is accessible from any other one and all the states are aperiodic. As a consequence, the chain has only one stationary distribution μ_ϵ, which clearly depends on ϵ; the Markov chain converges to this stationary distribution, regardless where it began.

Recall that the perturbation has been introduced into the imitation process in order to test the robustness of the multiple outcomes of the unperturbed process. We are then interested in investigating the behavior of the perturbed process as $\epsilon \to 0$.

The crucial result for our application is is a straight consequence of the theorem stated by Vega-Redondo:

> THEOREM 2 The limit distribution $\mu^* = \lim_{\epsilon\to 0} \mu_\epsilon$ is a well defined element of the unit simplex $\Delta(\Omega)$. Moreover, μ^* puts probability one over the state $\omega^W = (q^W, q^W, q^W, q^W)$ where q^W s.t. $p(nq^W)q^W - C_i(q^W) \geq p(q^W)q - C_i(q)$.

.

Applying this result to the setting adopted in my experiment we get that under the "imitate the best" dynamic, the only stochastically stable outcome is the Walrasian outcome $\omega^W = (20, 20, 20, 20)$, in which all the firms get zero profits.

The "robot" players in treatment ItheB follow this learning rule, and the probability of a tremble is set equal to 0.05.

I adopted a within subjects design, so every subject played against a unique type of "robot" players. 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 each, and the average payment (including the show up fee) was equal to 179 SEK [3]
In total, 7 subjects played under the ItheB treatment, and 8 under each of the other two treatments.

4.2 Results

In what follows I shall present some very preliminary results from these first two sessions, being well aware that the number of observations I have does not allow me to draw any sound conclusion from them.

4.2.1 Differences across treatments

Table 1 displays the quantities produced on average by the subjects and by their "robot" competitors in the three treatments, first across all the 40 rounds, then just for the last 10 rounds.

The first important thing to notice is that, unsurprisingly, subjects react differently when faced with different opponents: the average quantity produced under T&E is significantly higher than the one produced under BRD, which in turn is higher than under ItheB (Wilcoxon rank-sum test rejects the hypothesis that observed quantities under ItheB and under BRD come form the same distribution at any significance level, and the hypothesis that observed quantities under T&E and under BRD come form the same distribution at 1% significance level).

Under the – somehow unrealistic – hypothesis that subjects follow exactly one of the three learning rules simulated by the computer, we could have observed at least in one of the three treatments a convergence towards the predicted equilibrium. Even though the behavior of the robot players is not so far from what is predicted by the theory, the average quantity chosen by the players is quite different from the theoretically anticipated one and the distance is even wider if we look only at the last ten rounds: the observed values seem to depart from the predicted ones as the game proceeds.

[3] 1 SEK was about 0.107 Euro at the time the experiment took place.

Table 1: **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	

Figure 4.2.1 shows the average share of the time dedicated by each subject to the four pieces of information they could look up during the game. The first noticeable fact is that most of the players'attention is devoted to the plot that represents profits earned and quantities produced in the previous period 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 period, as suggested by Vega Redondo, in general they know the information necessary to do it. On the other side, a theory of learning such as 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.

It also interesting to observe that the type of competitor the players face affect not only their average choice but also the way they distribute their attention to the different pieces of information. Wilcoxon rank sum tests support the hypothesis that under treatment ItheB the share of time spent looking up last period profits and quantities is higher than in the other two treatments, while the time spent on the other three pieces of information is shorter. No significant difference in the allocation of attention emerges between the other two treatments.

4.2.2 Evidence of Learning

I have built another measure of the allocation of the attention, by counting the number of periods in which the each subject spent on a particular box. Figure 4.2.2 evidence the presence of a trend [4]: the interest on the last period's results increases in time, while the profit calculator receives less and less attention. On the other hand, the attention dedicated by players to their own past remain scarce all over the game, reenforcing our skepticism about the trial and error learning model.

The sharp decrease in the decision time during the first 20 rounds (Figure 4.2.2) 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 game. Once they have clearly understood the relation between quantities, prices and profits, subjects focus their attention to what the other players do.

The question is: do they only react to the past – as suggested by the three simple learning rules I adopted for my "robot" players? or do they also try to predict their competitors future choices? To answer this question, I started to compare the explanatory power of the three simple learning models described above.

[4] since the pattern is similar for the three treatments I only report the aggregate data

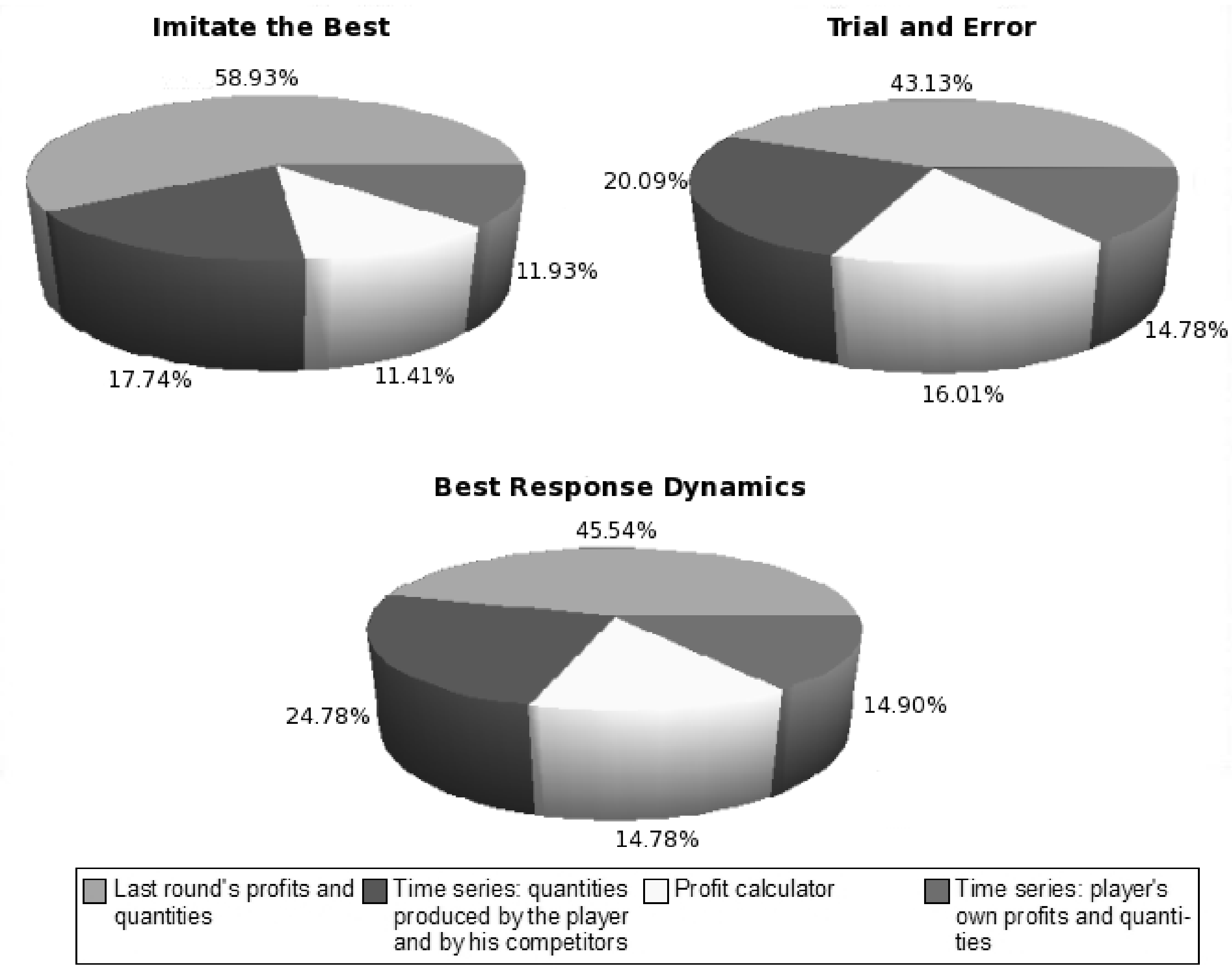


Figure 1: Distribution of players' attention in the three treatments.

4.2.3 Simple learning models

Following HNO [6], I estimated the following equation (as a panel with fixed effect):

$$q_i^t - q_i^{t-1} = \beta_0 + \beta_1 avg_i^{t-1} + \beta_2 best_i^{t-1} + \beta_3 imit_i^{t-1} + \beta_4 trial_i^{t-1}$$

where $imit_i^{t-1}$ represents the difference between the quantity that player i should have produced in period t according to the "Imitate the Best" learning rule and the quantity q_i^{t-1} he produced in the period before; $best_i^{t-1}$ is variation in quantity suggested by the "Best Reply" rule, and $trial_i^{t-1}$ the one predicted by "Trial and Error". avg_i^{t-1} is simply the difference between the average quantity produced in the previous period by the three competitors[5] and the quantity q_i^{t-1}.

Table 2[6] evidence some noticeable facts. First, under the three treatments, the learning rule based on myopic best reply seems to predict fairly well the observed variations in prices.
Second, the values assumed by the other coefficients are different under the three treatments[7]: that is, if we could rely on the few observations we have, we would say that the type of opponents faced by a player affects also the way he adapts his present behavior to what happened in the past.

[5]I have first tried to run the same regression without this variable, but the explanatory power of the model was too low, and it increased noticeably by introducing *avg* among the independent variables.

[6]the(*) symbol indicates the significance level of the coefficients in the usual way.

[7]To be precise, a Chow test reveals a structural break between BRD and ItheB and between BRD and T&E, but not between ItheB and T&E

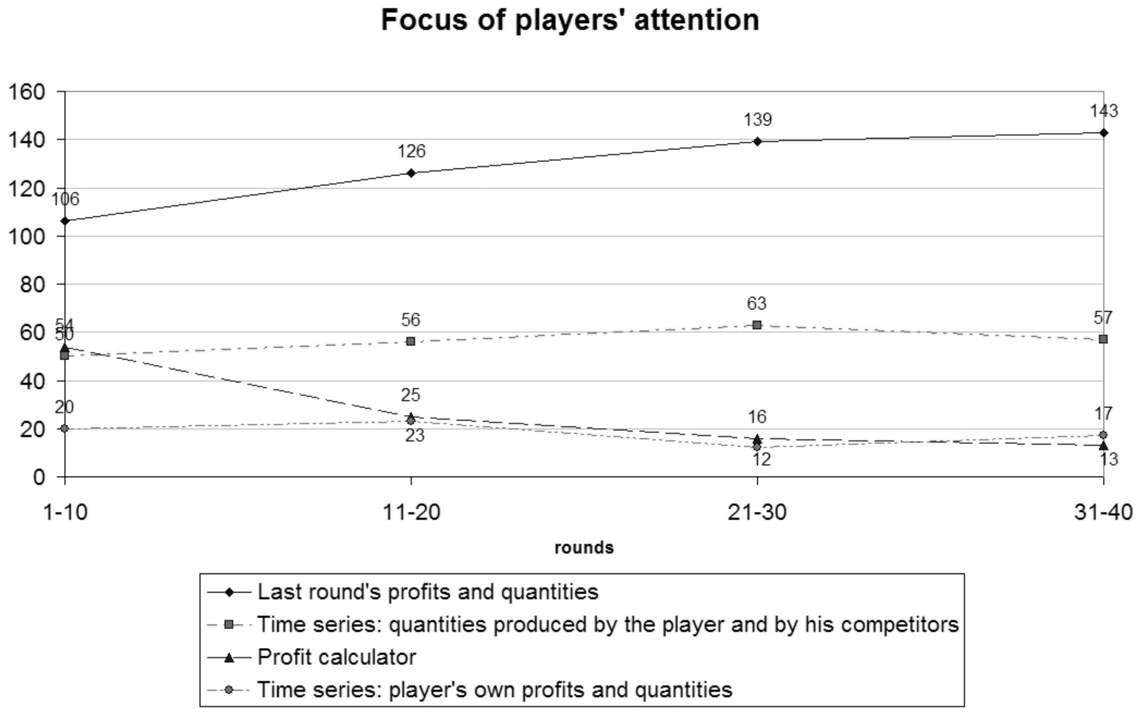


Figure 2: How the focus of players attention changes over time

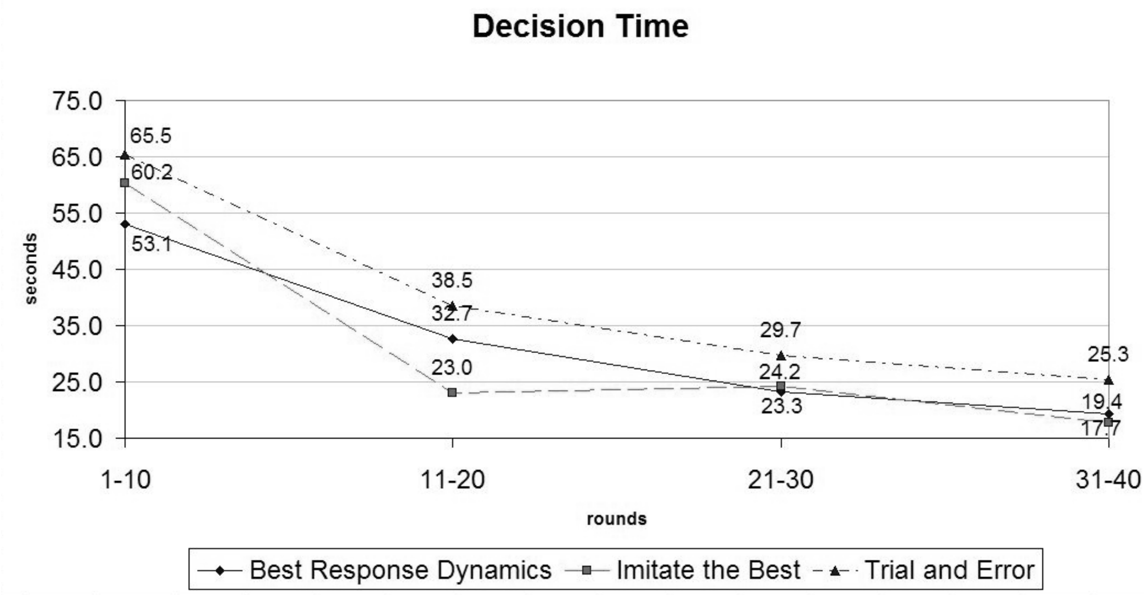


Figure 3: Average decision time across rounds, in seconds.

Third, the learning rule based on the imitation of the best performer in the last round does not find strong support in our data: the coefficient β_3 is significant only under T&E, and even in this treatment it is not the most important driver of players' behavior. Therefore, even if the subjects spent most of their decision time looking at the individual choices and profits in the last round, in general they did not use these informations to "imitate the best".

What I have presented here is just a short and superficial panoramic of the data collected during the first two sessions of my experiment. These data require a more accurate analysis, but the first impression is that the hypothesis that subjects follow some very simple heuristic 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 because the players pay to little attention the their own past profits and quantities, which is the only information required to apply this learning rule. On the other hand, imitating the best performer in the last period does not predict well the observed choices. The myopic best reply seem to drive players' choices, at least partially, and is consistent with the information they acquire, on average. In fact, to apply this learning rule the subjects need to know the sum of the quantities produced by their competitors in the last period – an information they almost always look at – and they must be able to compute a best reply, which means that either they use the profit calculator or they have used it extensively in the past and already know what the best reply is. Still, this model does not fully explain the observed variations in players behavior. Moreover, if it was the only driver of players' choices, under treatment BRD we should have observed convergence towards the Cournot equilibrium, which is not supported by the data.

5 Concluding Remarks

I have presented here an experimental project on the effects of information and learning in a Cournot oligopoly setting. Two versions of the experiment have been mentioned: one in which subjects play against each other, and a simplified version in which subjects play against the computer.
The first impression from the data collected through the first two sessions of this second experiment is that players behavior cannot be encompassed by the simple models of learning I shortly described in section 4. I plan to collect more data, both repeating the experiment in which players face "robots" and running the interactive version of the experiment. Then I want to compare the data with the predictions of other, more sophisticated learning models such as, for example, the Experience-Weighted Attraction Learning model proposed by Camerer, Ho and Chong [2]. As anticipated at the beginning, the data I will collect make it possible to test different learning models not only by comparing actual and predicted choices, but also actual and predicted search patterns.

First, I want to look for a model which fits the data, explaining not only the choices but also the patterns of information search adopted by the subjects. The question follows whether there is a single model that encompasses the behavior observed when subjects play against each other and when they face the three types of "robot" opponents, or instead the differences in the observed behavior depending on the type of opponent are so wide that a single model is not able to explain them all.
A second, related question is if, how and why players' behavior changes when they face real opponents. The impression I draw from the data collected so far, is that when they face simulated opponents, subjects spend most of their time trying to understand the logic followed by the computer in order to predict its choices and to exploit them. Human opponents' behavior is less homogeneous and less foreseeable, therefore I expect to observe a different search pattern when subjects are matched with each other.
The last question concerns whether subjects adopt less sophisticated learning models when they are under time pressure or when information is costly.

With my experiments I hope I'll be able to answer these questions and to contribute to the understanding of learning mechanisms in game-like situations. Moreover, my experiments will be also a way to test MouselabWEB as an experimental device that might be usefully adopted in other experiments on learning and to investigate other interesting situations in which imperfect information of some of the agents plays a crucial role. Examples might be auctions and financial markets, but also markets where hiding some attributes of the good being sold or the price of its add-ons may enable the sellers to get profits well above the competitive level.
In situations like those, a better comprehension of the relation between the data and stimuli provided to economic agents and their choices might help the regulator to set rules of information disclosure the bring the market outcome toward a more efficient equilibrium.

A Figures

A.1 Profit Calculator

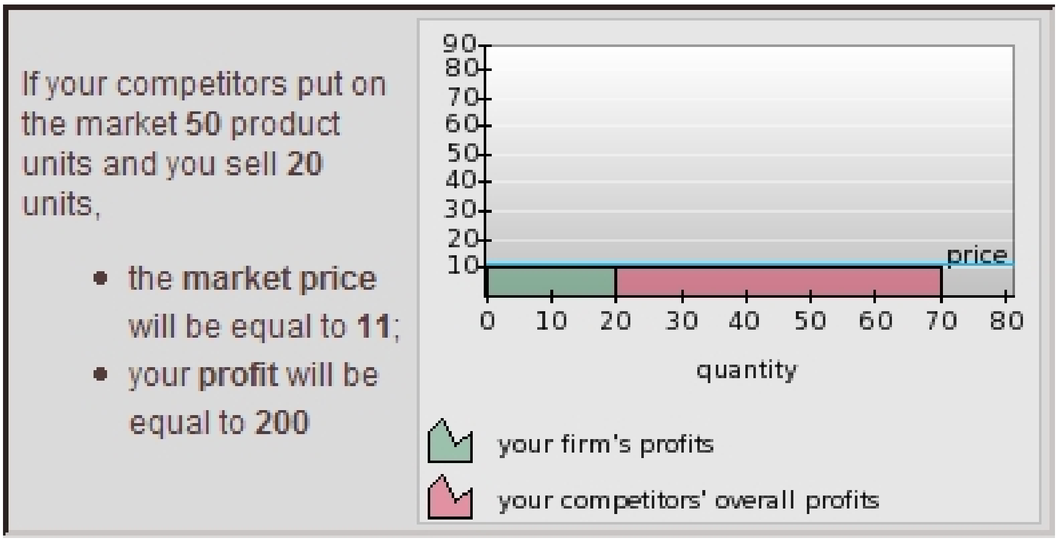


Figure 4: Profit calculator, example 1

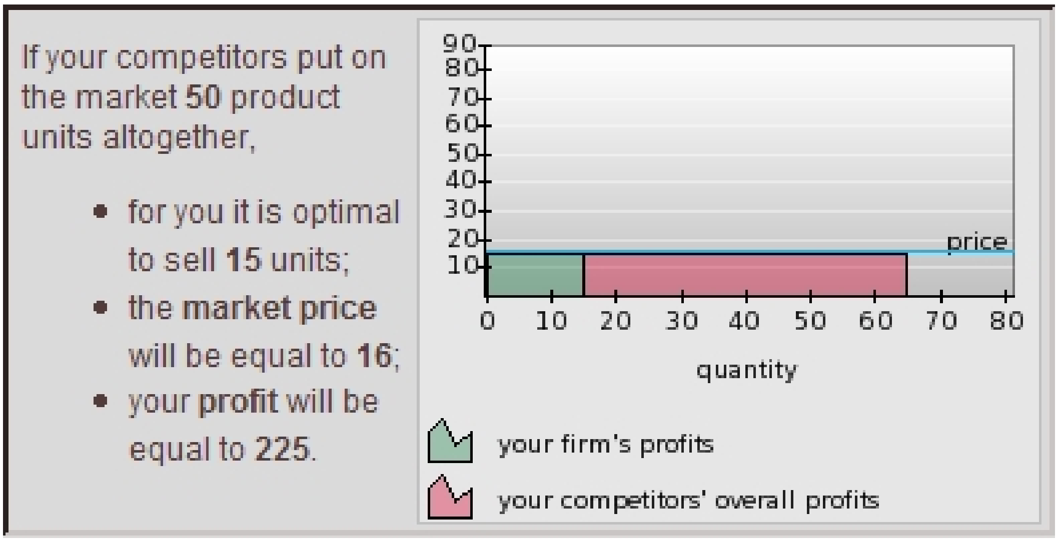


Figure 5: Profit calculator, example 2

A.2 Information about past rounds

The following three graphs are the only means through which information about what happened in the past round is displayed to the subjects.

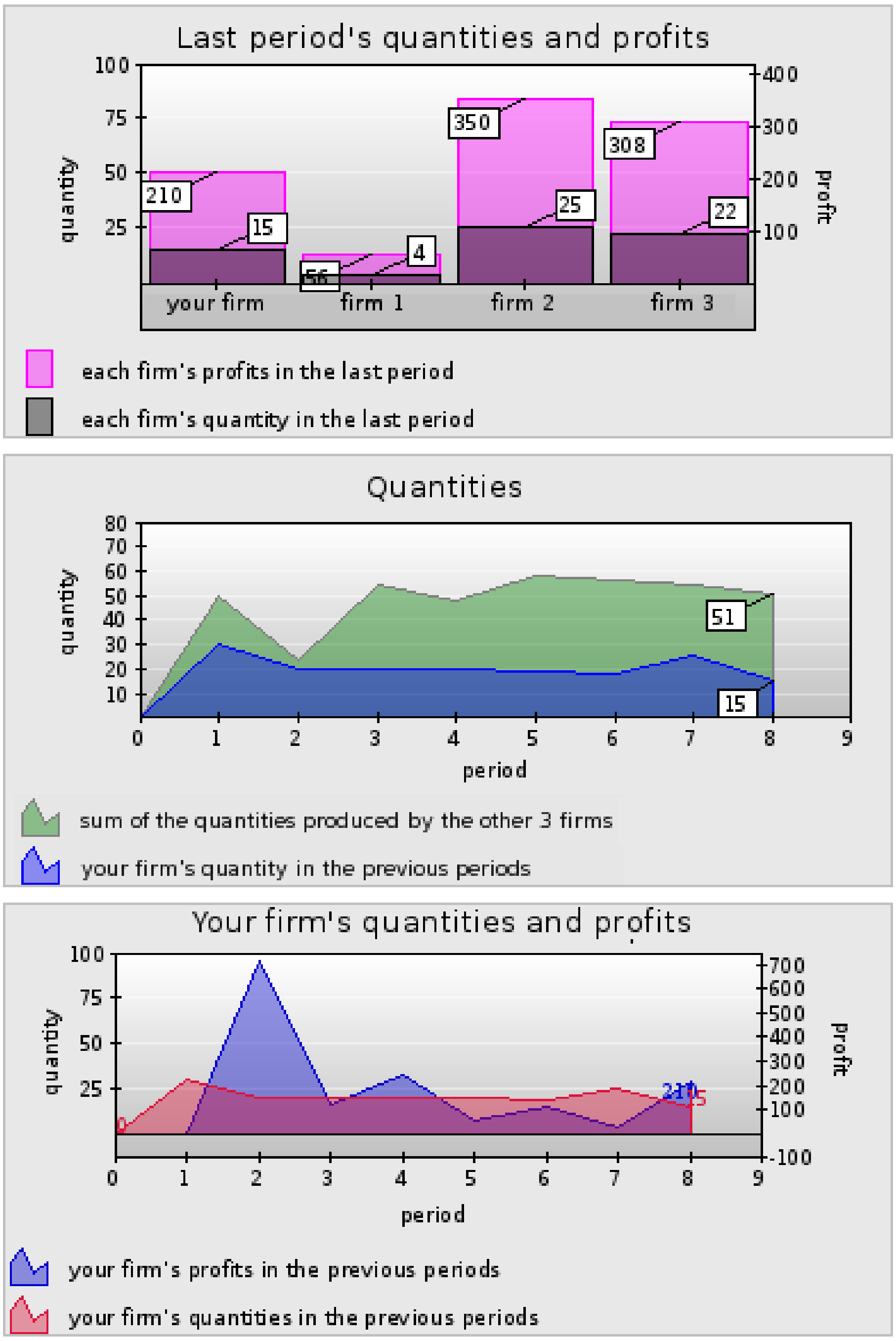


Figure 6: Information about the past periods displayed to the subjects at each round.

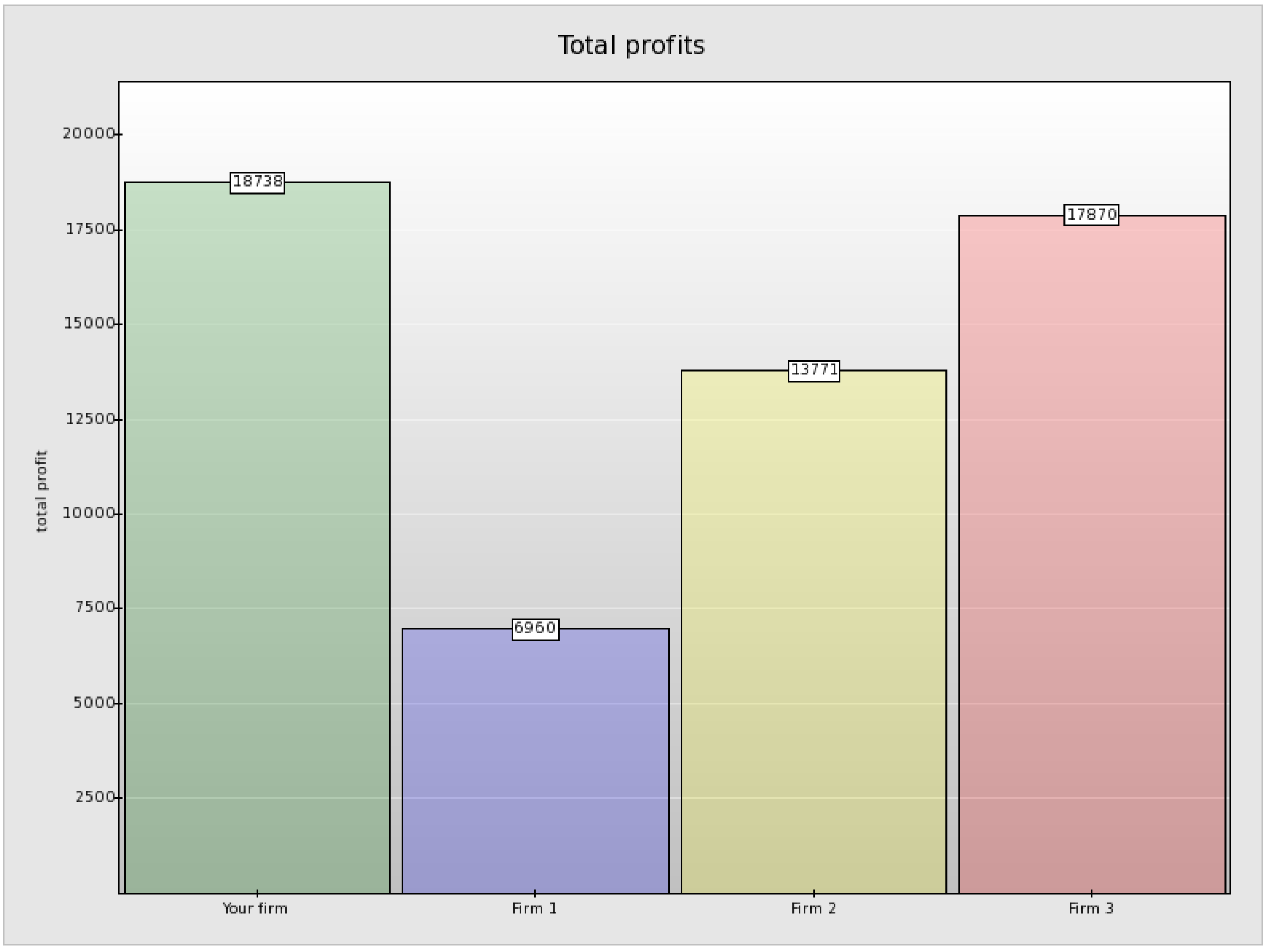


Figure 7: Plot displayed to each subject after the last period, at the end of the game.

A.3 Graphic Interface

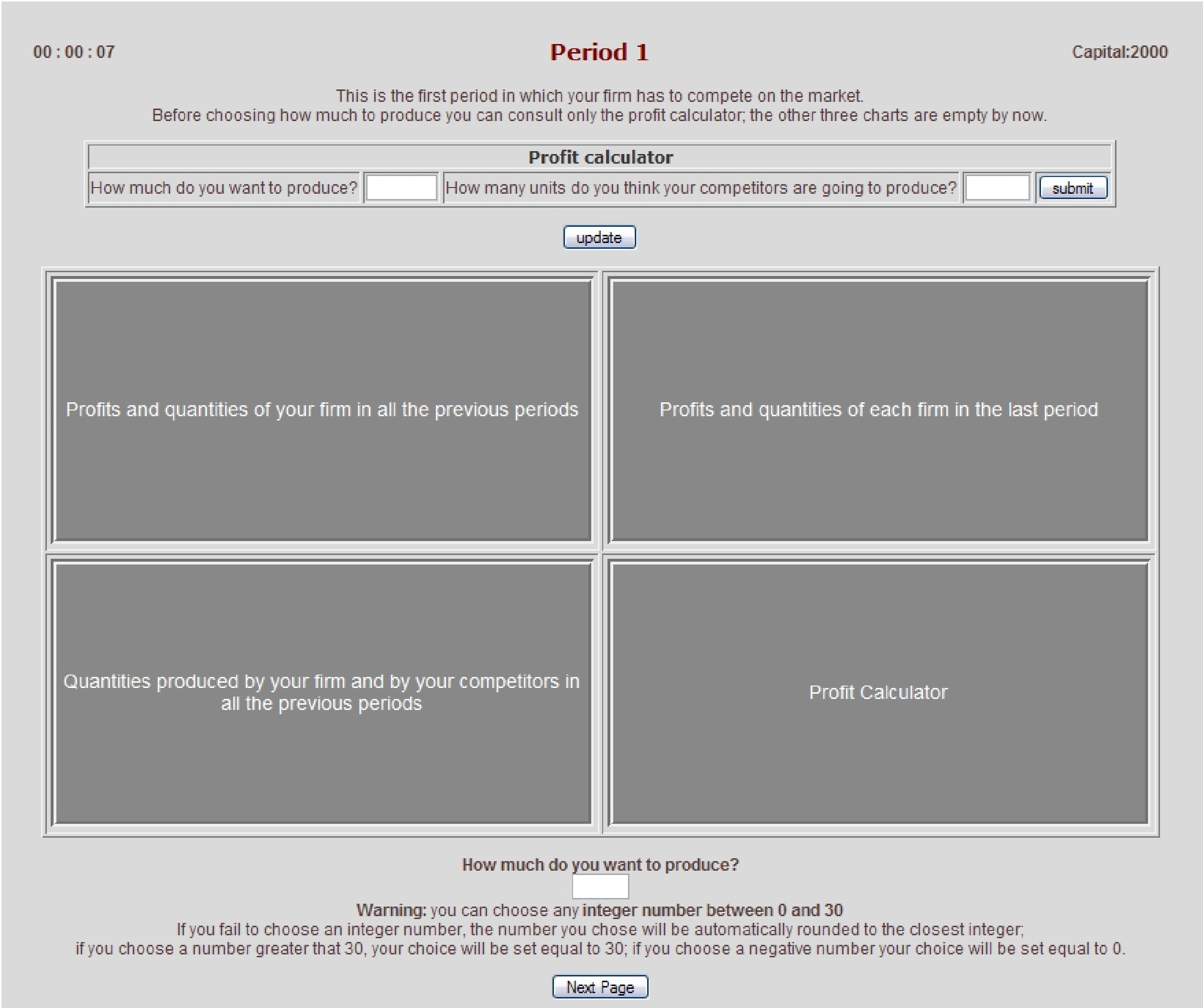


Figure 8: This image represents the graphic interface through which the game is presented to the subjects.

B Instructions

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.

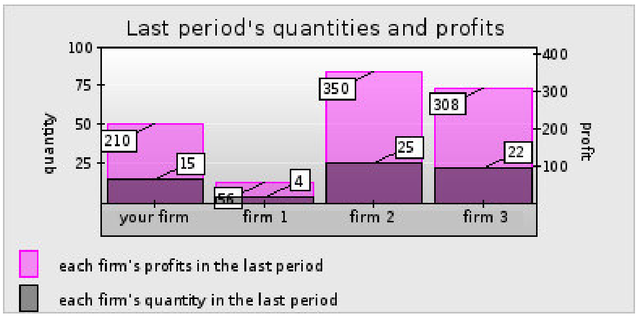
At the end of the game, your capital will be converted in SEK, and will be payed to you privately.
In addition, you will be payed a **show up fee** of 50 SEK.

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.

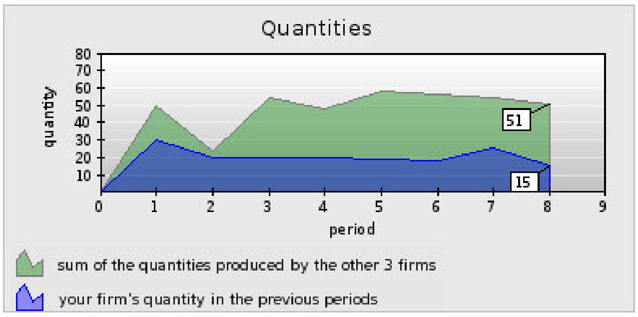
Profits and quantities of each firm in the last period



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.

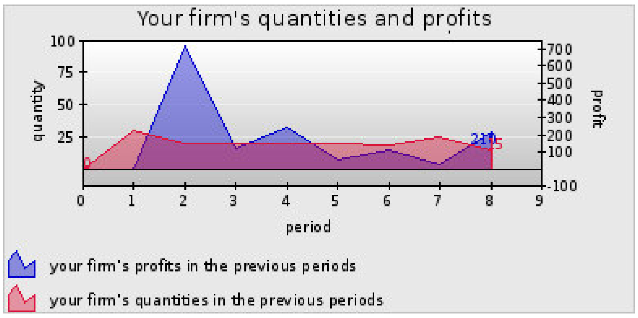
Quantities produced by your firm and by your competitors in all the previous periods



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.

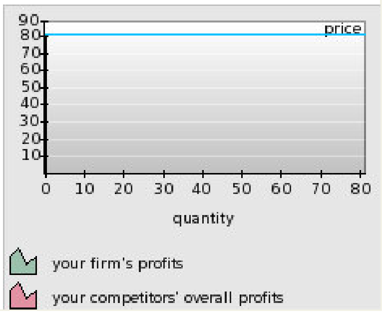
Here is an **example** of the profit calculator.

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

Update

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.

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

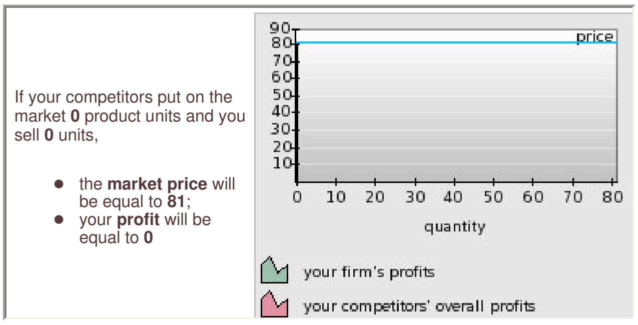
Second function of the profit calculator

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



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

[1] Bosch-Domènech, A., and Vriend, N. J. Imitation of successful behaviour in cournot markets. *Economic Journal 113*, 487 (2003), 495–524.

[2] Camerer Colin, Ho Teck-Hua, C. J.-K. Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. *Journal of Economic Theory 104* (2002), 137–188.

[3] Costa-Gomes, M., Crawford, V. P., and Broseta, B. Cognition and behavior in normal-form games: An experimental study. *Econometrica 69*, 5 (September 2001), 1193–1235.

[4] Costa-Gomes, M. A., and Crawford, V. P. Cognition and behavior in two-person guessing games: An experimental study. *American Economic Review* (Dec. 2006).

[5] Gabaix, X., Laibson, D., Moloche, G., and Weinberg, S. Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review 96*, 4 (Sept. 2006), 1043–1068.

[6] Huck, S., Normann, H.-T., and Oechssler, J. Learning in cournot oligopoly–an experiment. *Economic Journal 109*, 454 (1999), C80–95.

[7] Huck, S., Normann, H.-T., and Oechssler, J. Trial & error to collusion - the discrete case. Bonn Econ Discussion Papers bgse6_2000, University of Bonn, Germany, May 2000.

[8] Huck, S., Normann, H.-T., and Oechssler, J. Through trial and error to collusion. *International Economic Review 45*, 1 (02 2004), 205–224.

[9] Johnson, E. J., Camerer, C., Sen, S., and Rymon, T. Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of Economic Theory 104*, 1 (May 2002), 16–47.

[10] Normann, H.-T., and Wallace, B. The impact of the termination rule on cooperation in a prisoner?s dilemma experiment. May 2005.

[11] Offerman, T., Potters, J., and Sonnemans, J. Imitation and belief learning in an oligopoly experiment. *Review of Economic Studies 69*, 4 (October 2002), 973–97.

[12] Rassenti, S., Reynolds, S. S., Smith, V. L., and Szidarovszky, F. Adaptation and convergence of behavior in repeated experimental cournot games. *Journal of Economic Behavior & Organization 41*, 2 (February 2000), 117–146.

[13] Selten, R., Mitzkewitz, M., and Uhlich, G. R. Duopoly strategies programmed by experienced players. *Econometrica 65*, 3 (1997), 517–556.

[14] Theocharis, R. On the stability of the cournot solution on the oligopoly problem. *Review of Economic Studies 73*, 2 (February 1960), 133–134.

[15] Vega-Redondo, F. The evolution of walrasian behavior. *Econometrica 65*, 2 (1997), 375–384.

Table 2: Dep = $q^t - q^{t-1}$

Variable	**BRD** (Std. Err.)	**Imit. the Best** (Std. Err.)	**Trial & Error** (Std. Err.)
avg	0.753***	0.188	0.076*
	(0.080)	(0.096)	(0.033)
best	0.360***	0.504***	0.404***
	(0.047)	(0.071)	(0.058)
imit	0.019	0.081	0.236***
	(0.098)	(0.120)	(0.063)
trial	-0.397	-0.392	0.479***
	(0.273)	(0.425)	(0.126)
Intercept	2.117***	-0.032	-0.472
	(0.390)	(0.561)	(0.293)
R-squared	0.594	0.495	0.401
N	312	273	312