Information and Learning in Oligopoly: An Experiment

Maria Bigoni
Margherita Fort

January 2013

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Information and Learning in Oligopoly:
An Experiment

Maria Bigoni
University of Bologna

Margherita Fort
University of Bologna
and IZA

ABSTRACT

Information and Learning in Oligopoly: An Experiment

This paper presents an experiment on learning in repeated games, which complements the analysis of players' actual choices with data on the information acquisition process they follow. Subjects play a repeated Cournot oligopoly, with limited *a priori* information. The econometrics hinges on a model built upon Experience Weighted Attraction learning, and the simultaneous analysis of data on the information gathered and on actions taken by the subjects. Results suggest that learning is a composite process, in which different components coexist. Adaptive learning emerges as the leading element, but when subjects look at the strategies individually adopted by their competitors they tend to imitate the most successful behavior, which makes markets more competitive. Reinforcement learning also plays a role, as subjects favor strategies that have yielded higher profits in the past.

Corresponding author:

Maria Bigoni
Department of Economics
University of Bologna
Piazza Scaravilli 2
40126 Bologna
Italy
E-mail: maria.bigoni@unibo.it

1. Introduction

What happens when the information required to find out and adopt the equilibrium behavior is concealed, and its acquisition and processing is difficult or costly? The present study aims at contributing to the quest for an answer to this long debated question, by tackling this problem through an experiment on a repeated Cournot oligopoly game with strict informational constraints. We investigate how subjects learn to play, by monitoring at the same time what they do and what information they look for.

The debate on learning in oligopoly was opened by Vega-Redondo (1997), who developed an original idea by Schaffer (1989) and proposed a theoretical model of behavior of Cournot oligopolists. His model shows that Walrasian behavior can emerge in the long run within any Cournot oligopoly with homogeneous goods, if firms tend to imitate the behavior that in the previous period proved most successful among their competitors (that is, they produce the level of output that yielded the highest profit) but with positive probability experiment with other strategies. In a number of subsequent works Vega Redondo's theory has been experimentally tested and compared with other learning models that make different assumptions about players' information and lead to different behaviors and market equilibria.[1] The results of these studies, however, are not completely unanimous, nor fully decisive.

---

[1]See for example Huck et al. (1999), Rassenti et al. (2000), Offerman et al. (2002) and Bosch-Domènech and Vriend (2003). More recently, further experiments on Vega-Redondo's imitation model have been conducted by Apesteguia et al. (2007, 2010), but these experiments were not framed as oligopolies.

The question of how learning works in this and similar settings is still open and deserves further investigation.

The main purpose of the present work is to tackle the problem of learning in experimental oligopolies with a novel approach, which consists of combining the study of how subjects actually play the game with an analysis of how they select the information they need when choosing their strategy. Instead of comparing subjects market behavior under different informational frameworks — which is the approach adopted in all the previous experiments about this topic — we provide players with a broad range of information, but force them to choose only some pieces of it. Subjects' process of information gathering is strictly (but not obtrusively) controlled, by means of a special software, originally called MouseLab and developed by Johnson et al. (1988). By 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 of the informational framework on their behavior. The data about subjects' information search pattern are then integrated in the econometric analysis, via a comprehensive model of learning built upon the Experience Weighted Attraction (EWA) learning model developed by Camerer and Ho (1998, 1999).

Our main result is that learning is a complex and faceted process in which different elements coexist. More specifically: the tendency to best reply to the strategies adopted by the opponents in previous periods seems to be the principal driver of players' behavior. Yet, players also display a tendency to follow the strategies adopted by the most successful among their opponents, and this form of "imitation" — or maybe simply learning from

peers' experience — plays a non negligible role, driving players' choices away from the best reply towards a more competitive behavior. Finally, a subject's own experience of the outcomes stemming from the different strategies is relevant too, and exerts on his/her future actions an influence which is at least as strong as the pull of imitation.

The paper is structured in the following way: an overview of the theoretical learning models that inform the analysis is presented in Section 2, where we also briefly survey the experimental literature on learning in oligopoly games, and on the study of cognition through information search. Section 3 presents the experimental design. Section 4 contains the results of the experiment and Section 5 concludes.

2. Related Literature

In this Section we review the theoretical and experimental literature on learning. We first draw a taxonomy of different theoretical models of learning, which represents the reference framework for our design and for the empirical analysis. Then, we discuss previous experimental evidence from studies aimed at testing the relative importance of different learning models. In this second part, we limit our attention to the papers whose design is most closely related to ours. This paper focuses on learning in a set-up where information on the strategic environment is scarce. For this reason, we do not consider here the growing literature on learning across supergames, where the goal is to study how behavior stabilizes when subjects have had

4

the time to understand all the main details of the strategic situation.[2]

2.1. Theoretical background

We draw a taxonomy of theoretical learning models, grouping them into three main categories: experiential learning, adaptive learning, and models based on imitation. *Experiential learning* models are characterized by the general assumption that agents learn exclusively from their own experience. Agents' beliefs about other players' strategies, as well as information about opponents' past actions or payoffs, do not play any role here. Conversely, all those models in which these elements are important can be classified as "observational". In contrast with experiential learning, *adaptive learning* presumes that agents are able to observe their rivals' past play and that their computational capabilities and their knowledge of the game structure are sufficient for them to compute a best response, given the strategy profile adopted by their opponents. Finally, models based on *imitation* prescribe that, in every period after the first one, each individual chooses an action among those which were actually taken by some player in the previous round. The main difference among the various imitation-based models consists in whom is to be imitated.

We now describe each class of models in more detail. This survey will show that different learning rules determine simultaneously (i) the amount and quality of information agents need in order to make a choice, and (ii) their actual behavior as a response to the information they may acquire. The

---

[2]See for example Dal Bó 2005; Engle-Warnick and Slonim 2006; Camera and Casari 2009; Dal Bó and Fréchette 2011; Fudenberg et al. 2012; Friedman et al. 2012.

main idea underlying the present work is that observing subjects' information search pattern should provide a clue about the learning process they (consciously or unconsciously) adopt. Comparing information search patterns with actual choices may support some existing learning theories and suggest the rejection of others.

Experiential Learning. We consider here two approaches to experiential learning: reinforcement learning, and learning based on trial-and-error. Reinforcement learning (Roth and Erev, 1995) rests on the basic hypothesis that players increase the probability of playing pure strategies that have met with success in previous periods.[3] Theoretical results on the convergence properties of reinforcement learning in games with a large action space and more than two players are scarce. By means of a simulation-based analysis, Waltman and Kaymak (2008) show that on average the sum of the quantities produced in a market is significantly lower than that predicted by the Nash equilibrium, yet higher than the joint profit maximizing quantity. However, it is not clear whether the quantities individually produced by each player converge in the long run.

Learning by trial-and-error (Huck et al., 2000, 2004) simply prescribes that subjects move in the direction that was, or would have been, more successful in the previous period. In a standard symmetric Cournot oligopoly with n firms, Huck et al. (2000) show that — by making a few assumptions on cost functions and market structure — this process converges to the joint

---

[3]A variation of the reinforcement learning model has been proposed by Stahl (1996). However, in his paper reinforcement learning is not applied to single strategies, but to "rules of thumb" that "map from a player's information to the set of feasible actions."

profit maximizing equilibrium.

Adaptive learning. According to Milgrom and Roberts (1991), a player's sequence of choices is consistent with adaptive learning if the player eventually chooses only strategies that can be justified in terms of the competitors' past play. This justification is based on choosing strategies that are undominated if rivals' strategies are restricted to the most recently observed strategies. Best response dynamic and fictitious play are two examples of adaptive learning processes.

Originally suggested by Cournot (1838), best response dynamics assume that in every period each subject sets his current output equal to the best (i.e., current period payoff maximizing) response to the last period outputs of his rivals. Cournot (1838) demonstrates that this adjustment process is stable and converges to the unique Nash equilibrium for a duopoly with linear demand and constant marginal cost. This result is not general: Theocharis (1960) shows that best reply dynamics does not converge in oligopolies with a linear setup and three or more firms. Huck et al. (1999), however, prove convergence in finite time if some inertia is introduced, namely, if it is assumed that with some positive probability in every period each player sticks to the strategy he chose in the previous period.

Fictitious play (Brown, 1951) assumes that each subject would take the empirical distribution of the actions chosen in past periods by each of his opponents to be his belief about that opponent's mixed strategy, and in every period he would play a best response to this belief when choosing his current strategy. Shapley (1964) proves that fictitious play does not necessarily converge to the equilibrium, yet Monderer and Shapley (1996)

7

show that for "finite weighted potential games" — a class of games that includes the Cournot game adopted in our paper — every fictitious play process converges in beliefs to equilibrium.

Models based on imitation. Imitation-based models all rely on the theory of perturbed Markov processes, but differ in whom is to be imitated.

Vega-Redondo (1997)'s model prescribes to "imitate the best", that is, to choose, in each period, the quantity produced in the previous period by the firm that obtained the highest profit. The paper considers a set-up in which all firms conform to this rule, but with some small probability, in every period each firm "mutates", and adopts a different strategy. The author shows that, as the probability of mutation tends to 0, in the limit all firms produce the Walrasian quantity.[4]

An alternative mode, based on imitation of the average, is proposed by Huck et al. (1999). They assume that subjects who are uncertain about what to do and observe that the quantity produced on average by the other firms differs from their own, imitate this average quantity — thinking along the lines of 'everyone else can't be wrong'. Without inertia the process converges to the average of all the starting values. With inertia the process depends on the realizations of the randomization device, hence it is path dependent.

Offerman et al. (2002) propose another model based on imitation of the exemplary firm. They assume that there is a tendency to follow the firm that sets the good example from the perspective of industry profits, i.e. the

---

[4]Alós-Ferrer (2004) expands Vega-Redondo's model by assuming that firms have longer memory, and shows that in this case all monomorphic states where firms produce a quantity between the Cournot and the Walrasian outcome are stochastically stable.

firm (or one of the firms) that has chosen the level of output that would have maximized joint profits in the market, had it been followed by all firms. The paper shows that if all firms follow the exemplary one(s), the unique stochastically stable state of the process is the collusive outcome.

Broadly speaking, models based on imitation could be classified as observational learning models. Note, though, that the amount of information at firms' disposal varies across them.[5] To *imitate the best*, firms must have the opportunity to observe or at least to infer the individual profits of each of the opponents, while to *imitate the average* they only need to know the average output. On the other hand, to *imitate the "exemplary firm"*, they must have a wider knowledge of the market because the firms should be able to evaluate what would be the sum of profits if all the firms were producing a given level of output. Indeed, the informational requirement for this model overlaps the one assumed by models of adaptive learning.

2.2. *Experimental Studies*

Several experiments investigate the role of information and learning in a Cournot oligopoly setting. The commonly adopted approach is to vary the quality and quantity of information provided to the subjects across treatments. The impact of the various informational frameworks on players' choices is then studied by comparing the subjects' behavior across treat-

---

[5]An extended taxonomy of models based on imitation is presented by Apesteguia et al. (2007), who consider the broader case in which players might observe not only the actions taken and the payoff realized by their opponents, but also the strategies and payoff of other agents playing the same game in different groups. We chose to restrict information to the outcome realized within a subject's group, to limit the degree of complexity of our experimental design.

ments. Results from those studies are not completely consistent; this can be due to the some relevant differences in the experimental design. Indeed, while the experiments performed by Huck et al. (1999) and by Offerman et al. (2002) provide a rather strong support for the theory proposed by Vega Redondo, other works — such as Bosch-Domènech and Vriend (2003) and Rassenti et al. (2000) — evidence no general trend towards the Walrasian equilibrium and do not find any strong indication that players tend to imitate the one who got the best performance in the previous period.

Huck et al. (1999) study a Cournot oligopoly with linear demand and cost, in which four symmetric firms produce a homogeneous good. The treatment variable is the information provided to the subjects, about the market and about other players' choices and profits. They find that increasing information about the market decreases total quantity and harshens competition, while providing additional information about individual quantities and profits increases total quantity. In a similar set-up, Offerman et al. (2002) study how players' behavior changes across three treatments, which differ with respect to the amount of information provided to the subjects about individual quantities and revenues of the other two competitors in their market. The Walrasian outcome tends only to be reached when players are informed about their opponents' profits, while the collusive outcome is a stable rest point only when players get information about the quantities individually produced by their opponents. On the contrary, when no information about others' individual quantities and profits is provided, the only rest-point is represented by the Cournot-Nash equilibrium.

Bosch-Domènech and Vriend (2003) investigate whether people are more inclined to imitate successful behavior in a cognitively demanding environment. They study Cournot duopolies and triopolies, and vary the way information is provided and the time pressure put on the players. The data show that as the learning-about-the-environment task becomes more complex, average output increases, but remains far from the Walrasian output level. In duopolies, imitation of successful behavior tends to decrease rather than increase when moving to environments where less information is provided; only for triopolies, when subjects receive very limited information about the market structure, a weak trend towards the competitive outcome emerges.

Rassenti et al. (2000) study a repeated oligopoly game with homogeneous products, where cost functions are private information and differ across the firms. The demand function is linear, and is public information among the players. In one treatment, subjects were able to observe past output choices of each one of their rivals, in the other they were informed only about the past total output of rivals. The trend of total output over time is not consistent both with best response dynamic and with fictitious play. In addition, their experimental evidence supports neither learning models based on imitation, nor the more traditional hypothesis that information about competitors enhances the potential for collusion. The cost asymmetries introduced in this experiment, however, may partially explain the limited tendency to imitate.

The present work innovates the experimental approach adopted in the four papers described above, as it introduces the monitoring of the information acquisition process through a computer interface. To this purpose, we adopt a software — originally named MouseLab and developed by Johnson

et al. (1988) — whose main feature consists in hiding relevant pieces of information behind a number of boxes on the screen so that to access them the decision maker has to open the boxes and look at their content, with the restriction that he can open just one box at a time. To open a box, the player has to place his mouse cursor over it. By recording the number and the duration of the look-ups the program provides information about the decision makers' learning process. MouseLab has been used in several other studies (Johnson et al., 2002; Costa-Gomes et al., 2001; Costa-Gomes and Crawford, 2006; Gabaix et al., 2006), but to our knowledge, ours is the first experiment applying this technique to the analysis of learning processes in repeated strategic games.[6] In a companion paper, Bigoni (2010) adopts the same technique to study whether subjects' information acquisition process and behavior is affected by the learning rule adopted by the opponents, by letting participants play against computerized automata programmed to follow a specific rule. Results indicate that the learning rule followed by the opponents does not have a strong impact on the model of information acquisition and processing adopted by the subjects. With respect to Bigoni (2010), the present paper also presents a more in-depth analysis of subjects' learning beahvior, based on the EWA approach, as described in Section 4.4.

3. Experimental Design

The market environment of our experiment is similar to the one proposed in Huck et al. (1999). Four identical firms compete à la Cournot in the same

---

[6]A survey on the experimental study of cognition via information search can be found in Crawford (2008).

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. 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 $\omega^M = (10, 10, 10, 10)$, Cournot-Nash equilibrium is $\omega^N = (16, 16, 16, 16)$ and Walrasian equilibrium is $\omega^W = (20, 20, 20, 20)$.

A time limit of 30 seconds per round was introduced. If a subject failed to make his choice within the time limit, his quantity was automatically set equal to 0, granting him a profit of 0 for that period.[7] The aim of this constraint was to force subjects to choose the information they are really

---

[7]This happened 22 times (out of 1920), mostly in the first four periods. In Bigoni (2010), no time limit was imposed. Results indicate that the average response time in the first 5 periods ranges between 70 and 85 seconds, while it decreases to about 30-40 seconds after the tenth period. In light of these previous results, a time limit of 30 seconds appears to be a lapse long enough to let players take a reasoned decision, while still representing a binding constraint.

interested in, and – even more importantly – to reproduce an environment in which rationality is bounded because of external factors.

3.1. Information provided to the subjects

Participants knew how many competitors they had (anonymity was none-theless guaranteed). Instructions explained in plain words that there is an inverse relation between the overall quantity produced by the four firms and the market price and that a firm's production costs increase with the number of goods it sells. Besides, players were told that per-period profit is given by the market price times the number of goods sold by the firm, minus production costs (see the instructions in Appendix B).

Subjects were also provided with a calculator similar to the one proposed by Huck et al. (1999). This device had two functions. First, it could be used as a *Best Reply Calculator*, that computes the highest possible profit a subject could get, and the quantity that would yield him this profit, given an arbitrary quantity produced on the whole by the three competitors. Second, it could be used as a *Profit Calculator* to reckon the profit given both the quantity produced by the player himself and some arbitrary quantity produced by his competitors. The software recorded how many times each subject used the profit calculator and every trial he made.

We did not provide subjects with a complete payoff table, nor did we inform them about the details of the demand function. This feature of the design is crucial, as it allows us to investigate how subjects learn to play when in an environment where ex ante they have no means to adopt a fully rational profit maximizing behavior.

14

The number of rounds was common knowledge among subjects. According to game-theoretic predictions, cooperation should be sustainable only if our stage game were repeated in(de)finitely many times. However, Selten et al. (1997) point out that infinite supergames cannot be credibly replicated in the laboratory, because "a play cannot be continued beyond the maximum time available. The stopping probability cannot remain fixed but must become one eventually." In addition, Normann and Wallace (2012) show that the termination rule does not affect the aggregate subjects' behavior, except for inducing an end-game effect. In light of these considerations, a commonly known finite horizon was introduced, for sake of transparency and practicality.

In every period after the first one, the profits earned in the previous period by the player himself and by each of his opponents were displayed. Three distinct buttons — each corresponding to one of the player's competitors — allowed players to look at the strategy opponents employed in the previous period. Another button allowed the subject to open a window displaying, by means of a table and a couple of plots, the quantity chosen and the profits earned by the player himself in every previous period. It was also possible for the player to look at the aggregate quantity produced in each of the previous periods by his competitors. This information was conveyed through a table and a plot, if the subject pushed the corresponding button. A sixth button gave access to the calculator. Information was arranged so that there is a clear correspondence between each of the learning rules described in section 2.1 and subjects' potential choices, as displayed in Table 1.

It was not possible to access more than one piece of information at a

time, since opening a new window automatically closed the previous one. As subjects were asked to push a button in order to open a box and obtain the information hidden behind it, unintentional look-ups are avoided, ensuring a high level of control over players' information search pattern.

On the computer screen there was a counter showing the running cumulative profits earned by the player since the game began, and a timer displaying how many seconds remained before the end of the current period. Figure C.1 in Appendix C shows what subjects' computer screen looked like.

3.2. Experimental Procedures

The experiment was run on November 29 and 30, 2007, in the computer lab of the faculty of Economics, at the University of Bologna, in Italy. It involved 48 undergraduate students in Business Administration, Law and Economics, Commercial Studies, Economics, Marketing and Finance. Three identical sessions were organized, with 16 participants each. The length of the sessions ranged from one hour to one hour and fifteen minutes, including instructions and payment. The average payment was 13 Euro with a maximum of 17 and a minimum of 9, including a show-up fee of 4 Euro.

At the beginning of each session, subjects were welcomed into the computer room and sat in front of personal computers, and they were instructed not to communicate in any way with other players during the whole experiment. They received a printed copy of the instructions, which were read aloud so as to make them common knowledge.[8] Thereafter, they had the opportunity to ask questions, which were answered privately. Before starting

---

[8]A translation of the instructions can be found in Appendix B.

the experiment, subjects were also asked to complete a test on their computer, aimed at checking their understanding of the graphical interface they would have to use during the game. Only when all the players managed to answer correctly all the questions in the test, did the experiment begin. Each subject was randomly and anonymously matched with three other participants, who were to be his "opponents" throughout the whole game. At the end of the session, subjects were paid in cash, privately, in proportion to the profits they scored during the game. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

4. Results

In this Section, we first present some qualitative results about the output choices made by the subjects, and about their information search pattern. By means of a non-parametric approach, we then classify subjects according to the model of learning that best accounts for their information acquisition behavior, and we study whether subjects who display different information acquisition patterns also differ in terms of their approach to learning. We also consider how learning evolves as subjects gain experience, using this classificatory approach. To assess the prevalence of different approaches to learning in the population in a more general way, we consider an extension of the Camerer and Ho (1998)'s Experience Weighted Attraction (EWA) learning model.[9] We show how reduced form parameters of this generalized EWA model can be used to assess the relative weight of each learning approach on

---

[9]Appendix A discusses the identification of the structural parameters of this generalized EWA model.

the level of attraction that players attach to available strategies.

4.1. Overview of the results

Figure 1 and Table 2 present information about individual output.[10] First, we observe that the average output across all periods (17.6) is higher than the Cournot output (16), but lower than the Walrasian one (20) which, however, is the modal output in the first 10, and in the last 20 periods.[11] The Pareto-dominant collusive outcome of 10 is chosen only in 5.4% of the cases.

Table 2 shows an increase in the average output as the game proceeds. A non parametric Wilcoxon signed-rank test indicates that the quantity produced on average by each group of four players from period 1 to 20 is significantly lower than the quantity produced from period 21 to 40 (N=12, p-value: 0.0196).[12] Consequently, individual profits decrease across periods, and on average they remain substantially below the Cournot profit (512) throughout the whole game.

Table 2 also reports the standard deviation of the individual output, both within groups and across groups. The variability of individual output within groups does not decrease as subjects gain experience, as it would if all subjects adopted one of the four simple learning rules described in Section 2.1,

---

[10]In Figure 1 and in Table 2, the average is evaluated dropping the 22 observations (out of 1920) in which the outcome was zero because one of the subjects did not answer in time.

[11]We also observe peaks corresponding to multiples of five, revealing a tendency to simplify the game focusing only on some of the available strategies, which can probably explain why 15 is chosen more often than 16, representing the Nash equilibrium in the stage game.

[12]Two-sided test, groups of four subjects are taken as units of observation.

hence converging on the same output level. For this reason, it is difficult to draw conclusions on the learning rule adopted by the subjects, by considering only their actual behavior: observing how subjects distributed their attention among the different pieces of information gains even more interest. For this reason, we now turn to a more in depth analysis of subjects' information acquisition process. Our data suggest the following result:

Result 1. *The majority of subjects adopt an information acquisition pattern compatible with Adaptive Learning.*

A first piece of evidence in support of Result 1 comes from Table 3, which reports information on the fraction of subjects who accessed each piece of information, and on the average look-up times. It reveals that the vast majority of subjects look at least at one piece of information before making a choice. Only about one third of the subjects paid attention to their own past records, which suggests that experiential learning models find weak support in our data.[13] By contrast, aggregate competitors' output in past periods is an element that most subjects consult, throughout the whole game.

Another noticeable fact emerges from Table 3: the fraction of subjects looking at the output individually chosen by their competitors in the previous period increases from the first to the second half of the session, by about 10 percentage points. This shift in subjects' attention together with the previously observed increase in the average output level seems to be in line with

---

[13]Subjects who adopt such models should pay attention only to their own past actions and profits, and not to those of the others. Yet, it could well be that subjects do not look at these pieces of information because they remember it, and they prefer to acquire new elements rather than refreshing their memory of something they in fact already know.

Vega-Redondo's model. Notice also that on average, subjects do not only devote attention to the output chosen in the previous period by their best performing competitor, but also look at the output individually produced by other competitors, which is in partial contradiction with what is suggested by Vega-Redondo's theory of imitation. Subjects in this experiment seem to be concerned not only with the choice made by the competitor who performed best in the previous period, but also with the output chosen by each of the others.

Throughout the whole session, on average subjects spend about 20 seconds browsing the available information, before making a choice; no relevant trend across periods emerges under this respect. The figures on the average look-up times also reveal that using the Profit and Best Reply calculator takes longer than accessing the other pieces of information. About half of the subjects use of the calculator in this game. In particular, this device is mostly used to evaluate the myopic best reply to some aggregate quantity hypothetically produced by the player's opponents (see Table 4).

If a subject adopted some kind of adaptive learning, before using the calculator he should have gathered information about his competitors aggregate output in previous periods, by opening the appropriate boxes.[14] When this happens, we argue that the look-up sequence is consistent with adaptive learning (AL), and it turns out that this is the case in more than 90% of the times the calculator is used (Table 4).

---

[14] We consider a look-up sequence to be consistent with a best reply when − before using the profit calculator − the subject looks at the aggregate competitors' output, or at the individual output of each of the other three competitors.

When using the calculator, subjects have to enter a number corresponding to the hypothetical aggregate quantity produced by their opponents. This quantity can be seen as a proxy for their expectations about their competitors' future strategies. Table 5 displays the results obtained regressing this variable on the sum of the quantities produced by the player's opponents in the previous six periods.[15] According to these results, players' expectations are explained to a large extent by opponents' output observed in the previous two periods, in line with models of adaptive learning with short memory.[16]

4.2. Learning

The novel aspect of our dataset as compared to the previous experiments on learning in oligopolistic settings is that it allows to verify whether the learning model adopted by a subject drives both his information acquisition process, *and* his actual choices in the Cournot game. To test this hypothesis, here we consider the four broad categories of learning summarized in Table 1 (Experiential Learning, Adaptive Learning, Imitate the Best and Imitate the Average), and classify each subject according to the category that better predicts his information acquisition process. If our hypothesis is correct, we should observe that a specific learning model better predicts the choices of those subjects whose information acquisition process is compatible with that learning approach. We find the following result.

---

[15]The model includes individual fixed effects, and standard errors are clustered at the group level.

[16]In 44.14% of cases, the quantity used in the calculator is exactly identical to the quantity produced by competitors in the previous period.

Result 2. *A relation emerges between subjects' information acquisition process and their actual behavior: those who focus their attention on the pieces of information that are crucial for a specific learning rule also adjust their choices consistently with that learning rule.*

Our empirical evidence is based on an approach similar to the one proposed by Huck et al. (1999), which we apply first over the whole sample, and then separately for each of the categories of subjects identified on the basis of their information acquisition process. To classify subjects into the four learning categories, we consider the total look-up time spent collecting the information necessary to apply each of the four learning approaches.[17] Based on this measure, for each subject and each period we identify the category of learning that captures the longest look-up time. We also consider a residual category, which includes all the cases in which most of the look-up time is spent looking at pieces of information that are not compatible with any of the four categories we consider, so that the classification represents a complete partition. Subjects are then classified into the category, or the categories, that capture their information acquisition behavior in most of the periods. If all the four categories perform equally well, the subject is not classified. To compare the explanatory power of each category of models of learning, subject by subject, we perform pairwise comparisons between categories, using

---

[17]We exclude the look-up time dedicated to the Profit and Best-Reply Calculator, because it is disproportionately longer than the look-up time necessary to acquire the other pieces of information (see Table 3). Whenever a subject acquires the information necessary either to Imitate the Average, or to adopt some form of Adaptive Learning, to discriminate between the two we consider whether the look-up sequence is consistent with a Best-Reply process.

a z-test (Suissa and Shuster, 1985).[18]

This procedure classifies all subjects but three, and no subjects is classified under more than one category. It yields the following classification: Adaptive Learning includes 29 subjects, Imitate the Average includes 14 subjects, and Reinforcement Learning includes 2 subjects. The category of Imitate the Best does not include any subject. Table 6 reports the average output and the average profit, by category of subjects. It reveals that subjects classified under "Adaptive Learning" on average choose output levels that are significantly lower than those chosen by subjects classified under "Imitate the Average". Differences in Profits are not statistically significant.[19]

To explore whether subjects who differ in terms of their information acquisition process in fact also adopt different learning models, we follow the approach adopted in Huck et al. (1999), and estimate the following equation:

$$q_i^t - q_i^{t-1} = \beta_0 + \beta_{BR}(q_{BR}^{t-1} - q_i^{t-1}) + \beta_{ImitB}(q_{ImitB}^{t-1} - q_i^{t-1}) +$$
$$+ \beta_{ImitA}(q_{ImitA}^{t-1} - q_i^{t-1}) + \beta_{T\&E}(q_{T\&E}^{t-1} - q_i^{t-1}) \tag{1}$$

where q_{BR}^{t-1} denotes subject i's best reply to the opponents' aggregate output in period $t-1$, q_{ImitB} denotes the output chosen by the most successful subject(s) in the group, in period $t-1$,[20] q_{ImitA} denotes the average output chosen by the three opponents in period $t-1$, and $q_{T\&E}$ denotes the sign

---

[18]The R routines used for this part of the analysis have been kindly provided by Karl Schlag, whom we gratefully thank for suggesting this approach.

[19]Significance levels are obtained by means of regressions with standard errors clustered at the group level, which are available from the authors upon request.

[20]Here and throughout the whole paper, the most successful player may be the subject himself. In this case $q_{ImitB}^{t-1} = q_i^{t-1}$.

of the variation in the output predicted by a Trial and Error learning model (Huck et al., 2000, 2004). This last element is not included in the version of this regression adopted by Huck et al. (1999). We chose to add it in order to incorporate exactly one model for each of the categories of learning we consider.

Regression results are reported in Table 7. We first estimate the model over the full sample, then we distinguish between subjects classified under "Adaptive learning" and all the others, and finally we isolate those classified under "Imitate the Average".[21] In line with what already found by Huck et al. (1999), all coefficients have the expected sign, but adjustment is only partial, as all coefficients are far away from 1. Among all subjects, the learning rules based on the Best Response Dynamics and on Imitation of the Average seem to play a substantial role. However, consistently with findings from a related experiment in Bigoni (2010), a strong link emerges between the information acquisition pattern adopted by the subjects, and their learning approach. Among subjects classified under "Adaptive Learning", the learning rule based on Best-Response Dynamics is the most important factor, even if the estimated coefficients β_{BR} and β_{ImitA} are not significantly different. By contrast, the relative weight of the two learning rules based on imitation is markedly higher for subjects classified under "Imitate the Average". Finally, the learning rule based on Trial-and-Error does not find strong support in our data: the coefficient is not significant for any of the categories of subjects we

---

[21]In principle, one would like to estimate the model separately also for subjects classified under "Reinforcement learning", but in our case the sample size would be too small for meaningful elaborations.

identified. This confirms what already suggested by data in Table 3, which shows that on average subjects paid little attention to their own past records.

4.3. Learning to learn

To check whether subjects' approach to learning evolves with experience, we repeated the classificatory procedure presented in section 4.2 separately for the first 20 and the last 20 periods of play. Results are reported in Table 8. Most of the subjects (69%) are classified under the same category in the first and second half of the session. Of the remaining subjects, 3 are never classified, 5 switch from Adaptive Learning to Imitate the Average, 4 switch to Adaptive Learning from other categories, and 3 are multiply classified in the last 20 periods.

To the same end, we also estimate equation (1) separately for the first and the second half of the session, using the classification of subjects obtained separately for these two blocks of periods. Results are reported in Tables 9 and 10.

Two relevant findings emerge from Tables 9 and 10. First, in the first 20 periods of play, the learning model based on Trial and Error plays a non negligible role, especially among the subjects classified under Adaptive Learning; in the last 20 periods instead, the weight of this model becomes negligible. Second, for all subjects, the relative importance of models based on imitation increases over time.

4.4. Experience Weighted Attraction learning model.

As the purpose of this study is not to test one or more specific models of learning, but rather to explore how subjects choose between different pieces

of information available, and how they react to these elements intentionally collected, we develop an approach flexible enough to incorporate a broader spectrum of models, and appropriate for assessing which among these is more suitable to represent agents' behavior. To this aim, we start from the Experience-Weighted Attraction (EWA) learning model.

The EWA model was first proposed by Camerer and Ho (1998) and Ho et al. (2008).[22] It is a model that hybridizes features of other well known learning rules, such as reinforcement learning and adaptive learning, and that thanks to its flexibility has proven to fit data better than other models. This model is based on the idea that every player assigns to each strategy a given level of attraction, which can be represented by a number. Attractions are updated after every period, according to the players' experiences, and determine every player's probability distribution over his or her choice set. In the original EWA model, attractions are updated using the payoff that a strategy yielded in a period, and also the foregone profits a player would have earned had he chosen a different strategy. In the experiment presented here, foregone payoffs from unused strategies are not known to the players. Subjects, though, can use the calculator to discover the profit a particular strategy would yield, given the strategies chosen by the other players. As explained in section 3.1, the calculator has two functions: (i) it can be used by a player for evaluating the quantity that would yield him the highest profit given the aggregate quantity produced by his competitors, and for calculating

---

[22]The version of EWA learning we adopt as the basis of our analysis is a simpler and in many cases more workable version of the model — named "self tuning EWA learning" — developed by Ho et al. (2007).

26

the profit he would earn in the event that he produced the suggested amount of good (Best Reply calculator); *(ii)* it can also be used to determine the profit given both the quantity produced by the player and the sum of the quantities produced by his opponents (Profit calculator). By checking how a player used the calculator, we know precisely which information he uses to evaluate each strategy in every period.

If they wish, players can also access information about the strategies individually adopted in the previous period by their competitors. If they wanted to imitate the strategy chosen by the player who got the highest profit in the previous period — as suggested by Vega-Redondo — they would attach a higher attraction to that strategy, while virtually no relevance should be attributed to other elements.

Keeping these peculiar characteristics of the design in mind, we modify the attraction updating rule, so that in every period t attractions evolve depending on three elements: *(i)* the profit $\pi_i(s_i^j, s_{-i}(t-1))$ actually obtained by the player in period $t-1$; *(ii)* the profits $\pi_{i,imit}^j(t-1)$ obtained by each of the player's opponents playing strategy s^j in the previous period; *(iii)* the profits $\pi_{i,BRC}^j(t)$ and $\pi_{i,PC}^j(t)$ evaluated by the player using the Best Reply and the Profit calculator respectively, given his or her expectations about the competitors' choices.[23]

While the player always knows the strategy he played in the previous round and the profit he obtained, π_{imit}, π_{BRC} and π_{PC} may be known or

---

[23]If the Profit calculator is used more than once by player i in period t for the same strategy j, the profit $\pi_{i,PC}^j(t)$ is calculated as an average of the various profits associated to strategy s_i^j by the device (different profits correspond to different hypotheses about the other players' behavior).

unknown to the player, depending on the pieces of information he or she looks up. To control for the information the subject is aware of, we define four dummy variables. $d_{i,BRC}^j(t)$ takes value 1 if in period t player i uses the Best Reply calculator, and this device indicates strategy s^j as the best reply given the strategies played by the three opponents; it takes value 0 otherwise. $d_{i,PC}^j(t)$ takes value 1 if in period t player i uses the Profit calculator to check the profit that strategy s^j would yield, given the opponents' strategies; it takes value 0 otherwise. $d_{i,h}^j(t)$ takes value 1 if player i in period t knows that his opponent h played strategy s^j in the previous period, and 0 otherwise. $b_{i,h}(t-1)$ takes value 1 if player h had the highest profit in period $t-1$ among the opponents of player i, and 0 otherwise.

As a result, the updating rule for attractions is:

$$
\begin{aligned}
A_i^j(t) = {} & \frac{\phi_i(t)N(t-1)A_i^j(t-1) + \alpha\pi_i(s_i^j, s_{-i}(t-1))I(s_i^j, s_i(t-1))}{N(t)} + \\
& + \frac{\beta d_{i,BRC}^j(t)\pi_{i,BRC}^j(t) + \gamma d_{i,PC}^j(t)\pi_{i,PC}^j(t)}{N(t)} + \\
& + \frac{\epsilon \sum_{h \neq i} d_{i,h}^j(t)b_{i,h}(t-1)\pi_{i,imit}^j(t-1)}{N(t)} \\
& + \frac{\zeta \sum_{h \neq i} d_{i,h}^j(t)(1 - b_{i,h}(t-1))\pi_{i,imit}^j(t-1)}{N(t)}
\end{aligned}
\tag{2}
$$

where s_i^j denotes strategy j of player i, $s_{-i}(t)$ the strategy vector played by player i's opponents in period t. $I(x, y)$ is an indicator function which takes value 1 if $x = y$ and value 0 otherwise, and $\pi_i(s_i^j, s_{-i}(t))$ is the payoff i choosing strategy j when his or her opponents play $s_{-i}(t)$.

The parameter α measures the impact of reinforcement learning on at-

tractions; parameters β and γ capture the effect of the information obtained through the profit calculator, hence the role of different forms of adaptive learning; finally, parameters ϵ and ζ allow to assess the importance of alternative models of learning through imitation, of the best and of other opponents, respectively.

$N(t)$ is a measure of the weight players put on past attractions relative to present ones, and is defined as: $N(t) = \phi_i(t)N(t-1) + 1 \quad t \geq 1$. It can be interpreted as the number of "observation-equivalents" of past experience relative to one period of current experience. The initial value $N(0)$ is set equal to 1. Function $\phi_i(t)$ has the role of a discount factor, that depreciates previous attractions. The hypothesis made in Ho et al. (2007) is that the weight put on previous experiences should be lower when the player senses that the environment is unstable or that the strategies adopted by her opponents are changing. The authors then build a "surprise index" $S_i(t)$ measuring the difference between opponents' most recently chosen strategies and the strategies they adopted in all previous periods, and let $\phi_i(t) = 1 - \frac{1}{2}S_i(t)$. The surprise index is made up of two main elements: a cumulative history vector $h_i(t)$ and a recent history vector $r_i(t)$. The vector element $h_i^j(t) = \frac{\sum_{\tau=1}^{t} I(s_{-i}^j, s_{-i}(\tau))}{t}$ measures the frequency with which strategy s_{-i}^j was adopted by player i's opponents in period t and in all previous ones. Vector $r_i(t)$ has the k-th element equal to 1 and all other elements equal to 0, where $s_{-i}^k = s_{-i}(t)$. The surprise index $S_i(t)$ simply sums up the squared deviations between the cumulative history vector $h_i(t)$ and the immediate history vector $r_i(t)$:

$$S_i(t) = \sum_j (h_i^j(t) - r_i^j(t))^2.$$

Since in the game considered here the strategy space is rather big, we use the sum of the quantities produced by player's opponents instead of $s_{-i}(t)$ – representing the strategy adopted by the three opponents in period t – when calculating the surprise index $S_i(t)$.[24]

Attractions determine probabilities. More specifically: the probability P_{ijt} that player i chooses strategy j in period t is assumed to be monotonically increasing in $A_i^j(t)$ and decreasing in $A_i^\kappa(t)$, $\kappa \neq j$. The relation between attractions and choice probabilities is represented by a logistic stochastic response function:

$$P_{ijt} = \frac{e^{\lambda A_i^j(t)}}{\sum_k e^{\lambda A_i^k(t)}} \tag{3}$$

where the parameter λ measures sensitivity of players to attractions.[25]

Appendix A discusses the identification of the model parameters and provides information about the variables initialization. The Generalized EWA model can be written as a conditional logit model with *reduced form* parameters $\tau_1 \equiv \lambda\alpha$, $\tau_2 \equiv \lambda\beta$, $\tau_3 \equiv \lambda\gamma$, $\tau_4 \equiv \lambda\epsilon$, $\tau_5 \equiv \lambda\zeta$ that completely characterize it. While the scale factor λ is not identified, provided that $\lambda \neq 0$,[26] the *reduced form* parameters τ_k, allow us to assess the relative strength of each factor on the attractions $A_i^j(t)$.[27] Therefore, in what follows we will focus on

---

[24]By definition, $S_i(t)$ is bounded between 0 and 2, and as a consequence, $\phi_i(t)$ is bounded between 0 and 1. Since $N(0) = 1$, it also follows that $N(t) \geq 1 \, \forall t$.

[25]Probabilities are monotonically increasing in attractions under the assumption that $\lambda > 0$.

[26]Note that if $\lambda = 0$, all reduced form parameters should also be zero. One can check whether the assumption is plausible in each specific empirical application.

[27]Properties of the maximum likelihood estimator for the conditional logit parameters τ_k, $k = 1, 2, \ldots, 5$ follow from McFadden (1974) who presents necessary and sufficient conditions for the existence of a maximum and its uniqueness.

the ratios between structural parameters (α/β, α/γ, α/ϵ, α/ζ, ..., ϵ/ζ).

We estimate the reduced form conditional logit model by maximum likelihood on four samples: first, we estimate the model on the full sample, including all individuals and all periods. Then we split the sample into groups of individuals following the classification introduced in Section 4.2 and estimate the model in each subsample, as in Table 7. We expect that the relative weight of each factor in determining the attraction changes across subsamples: information coming from the profit calculator (parameters β and γ) should have the strongest impact on attractions for subjects classified under "Adaptive learning"; conversely, attractions should be mainly driven by information about the opponents' profits (parameters ϵ and ζ) for subjects classified under "Imitate the Average". We can summarize our results as follows:

Result 3. *Adaptive learning is the main driving force for all types of subjects. Imitation is also an important factor, especially for subjects who pay relatively more attention to their opponents' individual choices.*

Support for Result 3 comes from Table 11 and Table 12. Reduced form parameters (in Table 11) are precisely estimated in all cases, and they are significantly different from zero.[28] Based on these results, we can rule out

---

[28]The estimated information matrix is non-singular supporting empirical identification of the model (Skrondal and Rabe-Hesketh, 2004, chpt. 5). The condition numbers − defined as the square root of the ratio of the largest to the smallest eigenvalue of the empirical information matrix at the maximum likelihood solution − associated to the estimated models reported in Table 11 range between 1.6 and 3.2, i.e. are small. The estimated correlation between parameters is also small (not exceeding 0.00001 in all cases, for all subsamples) further supporting the fact that in our sample we observe enough variability to identify all the reduced form parameters.

that λ is zero in all subsamples; as a consequence, all the relative weights of the different pieces of information on attractions are well defined. Their estimates are presented in Table 12. Standard errors and confidence intervals are computed *via* delta method.

Results indicate that Adaptive learning is the main driving force, over the full sample, and in the three specific subsamples: in most cases, parameters β and γ are significantly larger than α, ϵ and ζ.[29] Estimates in Table 11 suggest that Reinforcement learning and Imitation also play a role, as the coefficients for $\lambda\alpha$, $\lambda\epsilon$, and $\lambda\zeta$ are all significantly different from zero. However, one can hardly detect any difference about the magnitude of role of the two (Table 12, 95% C.I. for α/ϵ and α/ζ include 1 in almost all samples). In addition, the evidence supports the idea that subjects, when evaluating a strategy, do take into account the profits realized by other players choosing that strategy, but do not attach more weight to the profit realized by the best among their competitors. Indeed, the 95% confidence interval for ϵ/ζ always includes 1.

The comparison of results between subjects classified under "Imitate the average" and under "Adaptive learning" confirms that, for the first subsample, Imitation becomes relatively more important than Adaptive and Reinforcement learning, consistently with what we find in Table 7. The difference across sub-samples, however, is not statistically significant.[30] Finally we observe that, according to our estimates, subjects tend to attribute similar

---

[29]To see this, note that the bounds of the 95% confidence intervals are both below 1 for the ratios α/β and α/γ, while for the ratios β/ϵ, β/ζ, γ/ϵ, and γ/ζ they are both above 1, in all samples but the one where we group together all subjects that are not classified under "Adaptive learning".

[30]We tested this by estimating the fully interacted model. Results are available from the authors upon request.

importance to the results from the Best Reply calculator and from the Profit Calculator: the 95% confidence interval for β/γ includes 1, in the full sample (however, it does not include it for the sub-sample of subjects classified as "Adaptive learning", where $\beta < \gamma$ at 5% level, since the extremes of 95% confidence interval for the parameter β/γ are both below 1).

5. Conclusion

In this paper we presented an experiment in which subjects were asked to play a repeated Cournot game with incomplete information. The first aim of the experiment was to check what feedback information subjects are really interested in, and to test how information is linked to the learning model adopted and in turn to the market outcome.

According to our data, learning appears to be a composite process, in which different components coexist. Adaptive learning seems to be the leading element, as subjects form expectations about their opponents' future actions and try to best reply to them. It is also noticeable that in most of the cases the opponents' output inputted in the profit calculator — a proxy for players' expectations — is explained to a large extent by opponents' output observed in the previous two periods. This means that either subjects expected their opponents not to change their strategy much or that they decided to use the profit calculator only when the opponents' strategy was stable enough to let them make predictions about the future. A considerable amount of look-up time is dedicated to the strategies individually adopted by competitors. This could be compatible with imitation of the best, as in Vega-Redondo's theory. Yet, our results suggest that players are not only

interested in output produced by the most successful competitor, but by all of their opponents. These results are confirmed by the estimates obtained via our generalized version of Experience Weighted Attraction learning model, suggesting that there is no difference between the weight attached to the profits collected by the most successful opponent and by the other competitors and that imitation is not the main driving force in the observed learning process.

The results we obtain via our generalized EWA model stress the importance played by reinforcement learning in this setting: even though subjects do not pay much attention to information concerning their own past history of play, when assessing the strength of a strategy they seem to take into greater consideration their own experience than what they know about other players' results. This lead us to think that probably they do not need to look at the relative information, because they have a clear memory at least of the most recent periods.

This experiment aims at contributing to the understanding of learning mechanisms in game-like situations. It adopts a new experimental approach, based on the "Mouselab" technique, which allows us to complement the observation of the choices actually made by subjects with detailed data on their information acquisition patterns. Our results suggest that this tool could 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, or in which reputation is an asset. 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 experiments on these issues, it could be useful to have a strict control on the pieces of information subjects focus their attention on. Indeed, a better understanding of the relation between the inputs provided to economic agents and their choices could lead to improvements in the regulation of information disclosure, bringing the market outcome toward a more efficient equilibrium.

6. Tables & Figures
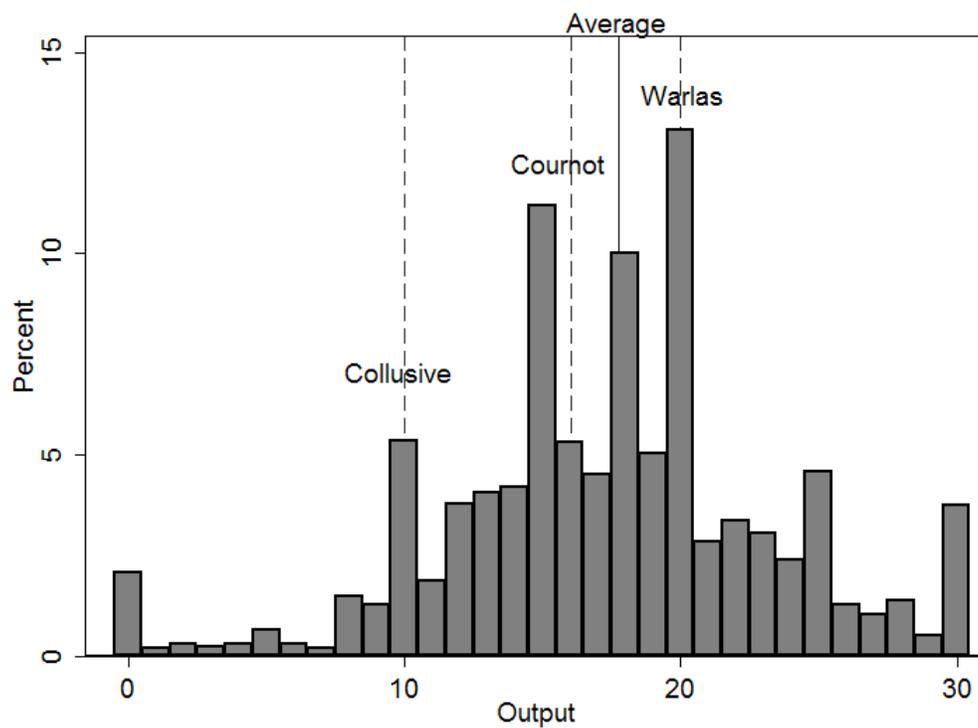
Figure 1: Frequency distributions of individual output levels, across periods.

Table 1:   information required by different models of learning.

piece of information	Experiential learning	Adaptive learning	Imitate the best	Imitate the average	Imitate the exemplary firm
aggregate competitors' quantity in previous rounds		x		x	
player's own profits and quantities in previous rounds	x				
last round individual profit and quantity of all competitors					x
last round individual profit and quantity of the best competitor			x		
profit/best reply calculator		x			x*

* Only at the beginning of the game, when players do not know the structure of the market.

Table 2:   Individual output and profits.

	Output			Profit
	Average	s.d. across groups	s.d. within groups	
Periods 1-10	16.7	2.7	5.1	220.9
Periods 11-20	17.3	2.0	5.1	184.4
Periods 21-30	18.2	2.6	5.9	142.8
Periods 31-40	18.1	2.6	5.4	149.8
Total	17.6	2.5	5.4	174.4

Table 3:   Pattern of information acquisition.

	Total	Periods	
		1-20	21-40
Profit and best-reply calculator	49.9%*(13.7)*	51.8%*(14.3)*	48.1%*(13.0)*
Aggregate competitors' output	68.1% *(6.1)*	66.8% *(6.2)*	69.5% *(6.0)*
Player's past profits and output	34.4% *(6.5)*	34.6% *(6.3)*	34.3% *(6.8)*
Output of the best performer	41.3% *(4.3)*	36.8% *(4.2)*	45.7% *(4.4)*
Output of other competitors	45.3% *(6.2)*	39.8% *(6.4)*	50.8% *(6.1)*
Total information	93.6%*(19.6)*	92.7%*(19.9)*	94.5%*(19.3)*

Note: Average look-up times, in seconds, are reported in parentheses. Averages are computed considering only subjects who accessed each specific piece of information.

Table 4:   Use of the two functions of the calculator, and percentage of observations in which the look-up sequence is consistent with a best reply to the aggregate quantity observed in the previous period.

	N. obs (%)	% of Look-up sequences consistent with Adaptive Learning
both functions	130 (13.56%)	96.92%
1st function (best reply) only	584 (60.90%)	93.84%
2nd function only	245 (25.55%)	87.76%
total	959	92.70%
Profit calculator not used	961	—

38

Table 5: Quantity input into the Profit Calculator

	1st function (best reply)	2nd function
Competitors' aggregate output in period:		
$t-1$	0.466***	0.541***
	(0.06)	(0.06)
$t-2$	0.149***	0.192**
	(0.03)	(0.06)
$t-3$	0.031	-0.053
	(0.03)	(0.07)
$t-4$	0.072	0.134*
	(0.04)	(0.07)
$t-5$	0.024	-0.086
	(0.03)	(0.10)
$t-6$	-0.017	0.029
	(0.03)	(0.05)
Constant	14.184**	12.411
	(4.59)	(7.93)
Sample size	649	330

The symbols ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 6: Classification of subjects into categories of learning models.

Category	N.	Output		Profit	
		Mean	[C. I.]	Mean	[C. I.]
Others	5	18.8	[14.9,22.8]	170.0	[94.0,246.0]
		$\sim$		$\sim$	
Imitate the Average	14	18.7	[17.2,20.1]	187.6	[155.6,219.6]
		$\vee$**		$\sim$	
Adaptive Learning	29	16.9	[16.2,17.5]	164.9	[139.2,190.7]

Table 7: Regression results from equation (1).

Parameter	Samples			
	Full Sample	Adaptive Learning	Others	Imitate the Average
Period	0.049***	0.044**	0.062**	0.051**
	(0.016)	(0.019)	(0.029)	(0.022)
β_{BR}	0.401***	0.443***	0.322***	0.277***
	(0.024)	(0.018)	(0.049)	(0.086)
β_{ImitB}	0.058**	0.024	0.120***	0.154***
	(0.023)	(0.030)	(0.042)	(0.038)
β_{ImitA}	0.378***	0.330***	0.464***	0.471***
	(0.067)	(0.070)	(0.103)	(0.079)
$\beta_{T\&E}$	0.104	0.154	0.048	0.193
	(0.135)	(0.192)	(0.264)	(0.299)
β_0	0.301	0.122	0.607	0.530
	(0.344)	(0.468)	(0.487)	(0.658)
Sample size	1809	1093	716	528
Log-likelihood	-5504.5	-3304.3	-2184.4	-1565.5
Wald tests on the estimated coefficients (p-values).				
$\beta_{BR} = \beta_{ImitB}$	0.0000***	0.0000***	0.0064***	0.2817
$\beta_{BR} = \beta_{ImitA}$	0.7800	0.1502	0.2416	0.1948
$\beta_{ImitA} = \beta_{ImitB}$	0.0002***	0.0010***	0.0087***	0.0005***

The symbols ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The model includes random effects at the individual level, and standard errors are clustered at the group level. The regression is run using the Stata GLLAMM package (Skrondal and Rabe-Hesketh, 2004).

Table 8: Classification of subjects into categories of learning models, in the first and in the second half of the session.

First 20 periods		Last 20 periods				
	Multiply classified	Unclassified	Imitate the Average	Adaptive Learning	Experiential learning	Total
Multiply classified	0	0	0	1	0	1
Unclassified	0	3	0	0	0	3
Imitate the Average	1	0	11	2	0	14
Adaptive Learning	2	0	5	21	0	28
Experiential learning	0	0	0	1	1	2
Total	3	3	16	25	1	48

Table 9: Regression results from equation (1), for periods 1-20.

Parameter	Samples			
	Full Sample	Adaptive Learning	Others	Imitate the Average
Period	0.039	0.001	0.147***	0.056
	(0.030)	(0.050)	(0.041)	(0.137)
β_{BR}	0.415***	0.439***	0.304**	0.451***
	(0.038)	(0.024)	(0.136)	(0.093)
β_{ImitB}	0.056	-0.031	0.102	0.305***
	(0.039)	(0.049)	(0.102)	(0.095)
β_{ImitA}	0.370***	0.340***	0.564***	0.273**
	(0.087)	(0.094)	(0.089)	(0.130)
$\beta_{T\&E}$	0.403***	0.474*	0.174	0.654
	(0.152)	(0.262)	(0.294)	(0.444)
β_0	0.398	0.813	-0.442	0.241
	(0.557)	(0.769)	(0.525)	(2.106)
Sample size	855	499	249	107
Log-likelihood	-2604.3	-1505.3	-740.6	-333.9
Wald tests on the estimated coefficients (p-values).				
$\beta_{BR} = \beta_{ImitB}$	0.0000***	0.0000***	0.3831	0.4051
$\beta_{BR} = \beta_{ImitA}$	0.6772	0.2947	0.1967	0.3669
$\beta_{ImitA} = \beta_{ImitB}$	0.0012***	0.0053***	0.0000***	0.7990

The symbols ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The model includes random effects at the individual level, and standard errors are clustered at the group level. The regression is run using the Stata GLLAMM package (Skrondal and Rabe-Hesketh, 2004).

Table 10: Regression results from equation (1), for periods 21-40.

Parameter	Samples			
	Full Sample	Adaptive Learning	Others	Imitate the Average
Period	0.016	0.024	-0.035	0.173
	(0.021)	(0.027)	(0.029)	(0.127)
β_{BR}	0.415***	0.471***	0.331***	0.562**
	(0.031)	(0.040)	(0.062)	(0.237)
β_{ImitB}	0.067*	0.105***	0.073	-0.192
	(0.039)	(0.037)	(0.060)	(0.305)
β_{ImitA}	0.503***	0.438***	0.592***	0.700**
	(0.040)	(0.053)	(0.063)	(0.282)
$\beta_{T\&E}$	-0.179	-0.270	0.162	-2.175***
	(0.178)	(0.221)	(0.198)	(0.680)
β_0	1.349**	-0.378	4.253***	0.416
	(0.678)	(0.951)	(1.129)	(3.952)
Sample size	954	537	378	79
Log-likelihood	-2866.3	-1549.1	-1153.8	-256.3
Wald tests on the estimated coefficients (p-values).				
$\beta_{BR} = \beta_{ImitB}$	0.0000***	0.0000***	0.2367	0.0076***
$\beta_{BR} = \beta_{ImitA}$	0.0940*	0.4931	0.0142**	0.5040
$\beta_{ImitA} = \beta_{ImitB}$	0.0000***	0.0003***	0.0008***	0.0000***

The symbols ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The model includes random effects at the individual level, and standard errors are clustered at the group level. The regression is run using the Stata GLLAMM package (Skrondal and Rabe-Hesketh, 2004).

Table 11: Conditional logit model estimates. Full sample and subsample by information acquisition patterns.

Parameter	Samples			
	Full Sample	Adaptive Learning	Others	Imitate the Average
Reinforcement Learning				
$\lambda\alpha$	0.009***	0.006***	0.011***	0.011***
	(0.001)	(0.002)	(0.001)	(0.001)
Adaptive Learning				
$\lambda\beta$	0.017***	0.017***	0.018***	0.016***
	(0.001)	(0.001)	(0.003)	(0.003)
$\lambda\gamma$	0.020***	0.022***	0.015***	0.017***
	(0.002)	(0.003)	(0.005)	(0.006)
Imitation				
$\lambda\epsilon$	0.006***	0.006***	0.007***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
$\lambda\zeta$	0.007***	0.006***	0.008***	0.009***
	(0.001)	(0.001)	(0.002)	(0.002)
Obs	58032	35061	22971	16926
LogL	-5673.948	-3332.253	-2326.865	-1700.242
Condition number[+]	2.001	1.668	3.250	3.062
N.subjects[++]	48	29	19	14

[+] The condition number is defined as the square root of the ratio of the largest to the smallest eigenvalue of the empirical information matrix at the maximum likelihood solution.

[++] Each subjects contributes to the estimates with $J \times (T-1) = 31 \times 39$ observations. 3 subjects are not classified (see Section 4.2 in the text) but observations for these subjects are included in the full sample. Standard errors are clustered at the group level and allow for dependency within clusters.

Table 12: Maximum likelihood estimates of the relative weight of each factor on attractions calculated using the reduced form estimates in Table 11. Standard errors obtained by the delta method in parentheses. Full sample and subsample by information acquisition patterns.

| Structural Parameters Contrast | | Samples | | |
	Full Sample	Adaptive Learning	Others	Imitate the Average
α/β	0.512***	0.328***	0.624***	0.666***
s.e.	(0.063)	(0.120)	(0.057)	(0.105)
95% c.i.	[0.388,0.637]	[0.093,0.562]	[0.513,0.735]	[0.450,0.872]
α/γ	0.434***	0.255***	0.749***	0.643***
	(0.046)	(0.085)	(0.280)	(0.264)
95% c.i.	[0.343,0.524]	[0.088,0.442]	[0.201,1.297]	[0.125,1.161]
α/ϵ	1.349***	0.911***	1.638***	1.380***
	(0.224)	(0.357)	(0.303)	(0.237)
95% c.i.	[0.910,1.787]	[0.211,1.611]	[1.044,2.232]	[0.916,1.845]
α/ζ	1.201***	0.880***	1.321***	1.187***
	(0.208)	(0.375)	(0.295)	(0.198)
95% c.i.	[0.794,1.608]	[0.145,1.614]	[0.743,1.899]	[0.799,1.575]
β/γ	0.847***	0.778***	1.201***	0.966***
	(0.921)	(0.097)	(0.480)	(0.408)
95% c.i.	[0.667,1.027]	[0.589,0.967]	[0.261,2.141]	[0.167,1.766]
β/ϵ	2.633***	2.781***	2.627***	2.074***
	(0.333)	(0.406)	(0.352)	(0.386)
95% c.i.	[1.980,3.287]	[1.986,3.576]	[1.936,3.318]	[1.318,2.830]
β/ζ	2.345***	2.684***	2.118***	1.783***
	(0.419)	(0.512)	(0.489)	(0.462)
− 95% c.i.	[1.525,3.166]	[1.682,3.687]	[1.162,3.075]	[0.878,2.687]
γ/ϵ	3.109***	3.575 ***	2.187***	2.146***
	(0.627)	(0.818)	(0.932)	(0.903)
95% c.i.	[1.880,4.338]	[1.972,5.178]	[0.361,4.014]	[0.375,3.917]
γ/ζ	2.769***	3.451***	1.764***	1.845***
	(0.521)	(0.853)	(0.801)	(0.728)
95% c.i.	[1.749,3.790]	[1.778,5.123]	[0.194,3.334]	[0.419,3.271]
ϵ/ζ	0.890	0.967***	0.806***	0.860***
	(0.158)	(0.153)	(0.205)	(0.195)
95% c.i.	[0.582,1.200]	[0.665,1.266]	[0.404,1.209]	[0.477,1.242]

References

Alós-Ferrer, C., 2004. Cournot versus walras in dynamic oligopolies with memory. International Journal of Industrial Organization 22, 193 – 217.

Apesteguia, J., Huck, S., Oechssler, J., 2007. Imitation – theory and experimental evidence. Journal of Economic Theory 136, 217–235.

Apesteguia, J., Huck, S., Oechssler, J., Weidenholzer, S., 2010. Imitation and the Evolution of Walrasian Behavior: Theoretically Fragile but Behaviorally Robust. Journal of Economic Theory 145, 1603–1617.

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

Bigoni, M., 2010. What do you want to know? Information acquisition and learning in experimental Cournot games. Research in Economics 64, 1 – 17.

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

Brown, G., 1951. Iterative solution of games by fictitious play. Activity Analysis of Production and Allocation 13, 374–376.

Camera, G., Casari, M., 2009. Cooperation among Strangers under the Shadow of the Future. The American Economic Review 99, pp. 979–1005.

Camerer, C., Ho, T.H., 1998. Experience-Weighted Attraction Learning in Coordination Games: Probability Rules, Heterogeneity, and Time-Variation. Journal of Mathematical Psychology 42, 305–326.

Camerer, C., Ho, T.H., 1999. Experienced-Weighted Attraction Learning in Normal Form Games. Econometrica 67, 827–874.

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

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

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

Crawford, V., 2008. Look-ups as the Windows of the Strategic Soul: Studying Cognition via Information Search in Game Experiments, in: Caplin, A., Schotter, A. (Eds.), Perspectives on the Future of Economics: Positive and Normative Foundations. Oxford University Press. volume 1 of *Handbook of Economic Methodologies*.

Dal Bó, P., 2005. Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games. American Economic Review 95, 1591–1604.

Dal Bó, P., Fréchette, G., 2011. The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence. American Economic Review 101, 411–429.

Engle-Warnick, J., Slonim, R.L., 2006. Learning to trust in indefinitely repeated games. Games and Economic Behavior 54, 95 – 114.

Fischbacher, U., 2007. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. Experimental Economics 10, 171–178.

Friedman, D., Huck, S., Oprea, R., Weidenholzer, S., 2012. From imitation to collusion: Long-run learning in a low-information environment. Discussion Paper, Wissenschaftszentrum Berlin für Sozialforschung, Abteilung Ökonomik des Wandels, Schwerpunkt Märkte und Politik.

Fudenberg, D., Rand, D.G., Dreber, A., 2012. Slow to anger and fast to forgive: Cooperation in an uncertain world. American Economic Review 102, 720–49.

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

Ho, T., Camerer, C., Chong, J., 2007. Self-tuning experience weighted attraction learning in games. Journal of Economic Theory 133, 177–198.

Ho, T.H., Wang, X., Camerer, C., 2008. Individual Differences in EWA Learning with Partial Payoff Information. Economic Journal 118, 37–59.

Huck, S., Normann, H.T., Oechssler, J., 1999. Learning in Cournot Oligopoly–An Experiment. Economic Journal 109, C80–95.

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

Huck, S., Normann, H.T., Oechssler, J., 2004. Through Trial and Error to Collusion. International Economic Review 45, 205–224.

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

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

McFadden, D.L., 1974. Conditional Logit Analysis of Qualitative Choice Analysis, in: Zarembka, P. (Ed.), Frontiers in Econometrics. New York: Academic Press. chapter 4, pp. 105–142.

Milgrom, P., Roberts, J., 1991. Adaptive and sophisticated learning in normal form games. Games and Economic Behavior 3, 82–100.

Monderer, D., Shapley, L., 1996. Potential Games. Games and Economic Behavior 14, 124–143.

Normann, H.T., Wallace, B., 2012. The impact of the termination rule on cooperation in a prisoner's dilemma experiment. International Journal of Game Theory 41, 707–718.

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

Rassenti, S., Reynolds, S.S., Smith, V.L., Szidarovszky, F., 2000. Adaptation and convergence of behavior in repeated experimental Cournot games. Journal of Economic Behavior & Organization 41, 117–146.

Roth, A., Erev, I., 1995. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. Games and Economic Behavior 8, 164–212.

Salmon, T.C., 2001. An Evaluation of Econometric Models of Adaptive Learning. Econometrica 69, 1597–1628.

Schaffer, M.E., 1989. Are profit-maximisers the best survivors? : A darwinian model of economic natural selection. Journal of Economic Behavior & Organization 12, 29–45.

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

Shapley, L.S., 1964. Some topics in two-person games, in: Dresher, M., Shapley, L.S., Tucker, A.W. (Eds.), Advances in Game Theory. Princeton, NJ: Princeton Univ. Press., pp. 1–28.

Skrondal, A., Rabe-Hesketh, S., 2004. Generalized Latent Variable Modeling: multilevel, longitudinal, and structural equation models. CRC Press.

Stahl, D.O., 1996. Boundedly Rational Rule Learning in a Guessing Game. Games and Economic Behavior 16, 303 – 330.

Suissa, S., Shuster, J.J., 1985. Exact unconditional sample sizes for the 2×2 binomial trial. Journal of the Royal Statistical Society. Series A (General) 148, pp. 317–327.

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

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

Waltman, L., Kaymak, U., 2008. Q-learning agents in a Cournot oligopoly model. Journal of Economic Dynamics and Control 32, 3275–3293.

Appendix A. Identification of the Generalized EWA Model

To show that the model parameters are identified we follow three steps: we first show that the generalized EWA model can be written as a conditional logit model; then we discuss whether the parameters of the conditional logit model are identified. The statistical properties of the maximum likelihood estimator for the parameters of the conditional logit model are discussed in McFadden (1974) who presents necessary and sufficient conditions for the existence of a maximum (Lemma 3) and its uniqueness. As in McFadden (1974), we derive the first order conditions (FOC) for maximum likelihood (ML) estimation and compute the second order derivatives. We discuss what type of variation in the data is required for the ML-FOC not to be trivially met. Finally we discuss the link between the conditional logit models parameters and the parameters of the generalized EWA model.

In the main text, we also discuss the *empirical identification* by checking non-singularity of the estimated information matrix (Skrondal and Rabe-Hesketh, 2004, chpt. 5). Empirical identification (EI) is based on properties of the estimated parameters rather than the theoretical reduced form parameters. Skrondal and Rabe-Hesketh (2004) stress that EI has some advantages compared to the analytical method: i) it can be based on the estimated information matrix, a natural byproduct of maximum likelihood estimation; ii) it assesses identification where it matters, i.e. at parameter estimates; iii) it addresses problems inherent in the sample on which inferences must be based. In our sample, the estimated information matrix is non-singular in both the full sample and in all the sub-samples considered.

There are two main differences between the standard EWA model and

our generalized EWA model: first, we use the version of the generalized EWA model that incorporates self-tuning (Ho et al., 2007); second we use a different definition of the attraction updating rule. As discussed in Section 4.4 in the main text, the updating rule in the generalized EWA is

$$A_i^j(t) = N(t)^{-1}[\phi_i(t)N(t-1)A_i^j(t-1) + \alpha I(s_i^j, s_i(t))\pi_i(s_i^j, s_{-i}(t))+$$
$$+ \beta d_{i,BRC}^j(t)\pi_{i,BRC}^j(t) + \gamma d_{i,PC}^j(t)\pi_{i,PC}^j(t)+$$
$$+ \epsilon \sum_{h \neq i} d_{i,h}^j(t)b_{i,h}(t-1)\pi_{i,imit}^j(t)+$$
$$+ \zeta \sum_{h \neq i} d_{i,h}^j(t)(1 - b_{i,h}(t-1))\pi_{i,imit}^j(t)]$$

This generalized version of the updating rule differs from standard EWA in the following: i) hypothesized payoffs do not play any role but only observed forgone payoffs matter; ii) terms associated with the information acquisition behavior are added.

To discuss the identification of the model parameters it is useful to re-write the updating rule as function of the initial attractions A_{i1}^j. To simplify notation, consider the re-parametrization below

$$A_i^j(t) = x_{it}^0 A_i^j(t-1) + \sum_{k=1}^{5} \theta_k x_{it}^{kj}$$

where

$x_{it}^0 = \phi_i(t)\frac{N(t-1)}{N(t)}$

$x_{it}^{1j} = \frac{I(s_i^j, s_i(t))\pi_i(s_i^j, s_{-i}(t))}{N(t)}$ and $\theta_1 \equiv \alpha$ using Section 4.4 notation

$x_{it}^{2j} = \frac{d_{i,BRC}^j(t)\pi_{i,BRC}^j(t)}{N(t)}$ and $\theta_2 \equiv \beta$ using Section 4.4 notation

$x_{it}^{3j} = \frac{d_{i,PC}^j(t)\pi_{i,PC}^j(t)}{N(t)}$ and $\theta_3 \equiv \gamma$ using Section 4.4 notation

$x_{it}^{4j} = \frac{\sum_{h \neq i} d_{i,h}^j(t)b_{i,h}(t-1)\pi_{i,imit}^j(t)}{N(t)}$ and $\theta_4 \equiv \epsilon$ using Section 4.4 notation

$x_{it}^{5j} = \frac{\sum_{h \neq i} d_{i,h}^j(t)(1-b_{i,h}(t-1))\pi_{i,imit}^j(t)}{N(t)}$ and $\theta_5 = \zeta$ using Section 4.4 notation.

By backward substitution, for $t \geq 2$, we get

$$A_i^j(t) = A_i^j(1)\prod_{s=2}^{t} x_{is}^0 + \sum_{k=1}^{5} \theta_k \left(\sum_{s=2}^{t-1} (x_{is}^{kj} \prod_{l=s+1}^{t} x_{il}^0) + x_{it}^{kj} \right) \qquad \text{(A.1)}$$

where A_{i1}^j is the initial attraction for strategy j for individual i.[1]

We set $N(0) = 1$, $S_i(1) = 1 \, \forall i$; as a consequence, $\phi_i(1) = \frac{1}{2}$, $N(1) = \frac{3}{2}$, as explained in Section 4.4. We also need to describe the equation for first period attractions $A_i^j(1)$: since before the experiment starts no participant has any information on the game and the available strategies, we set $A_i^j(0) = 0$. In period 1, no information on past actions and profits of group members is available, hence attractions for period one are only determined by how each subject used the profit and best reply calculator. Therefore, first period attractions are described as in equation (A.2) below.

$$A_i^j(1) = \beta \frac{d_{i,BRC}^j(1)\pi_{i,BRC}^j(1)}{N(1)} + \gamma \frac{d_{i,PC}^j(1)\pi_{i,PC}^j(1)}{N(1)} \equiv \theta_2 x_{i1}^{2j} + \theta_3 x_{i1}^{3j} \qquad \text{(A.2)}$$

---

[1] Notice that the factor $\prod_{s=2}^{t} x_{is}^0$ is always strictly smaller than 1. To see this, remember that $N(t) = \phi_i(t)N(t-1) + 1$ — hence x_{it}^0 can be rewritten as $x_{it}^0 = 1 - \frac{1}{N(t)}$ — and $N(t) \geq 1 \, \forall t$ (see footnote 24 in Section 4.4). This implies that the process determining the attraction is not explosive and initial attractions do not have an increasingly large influence through time even if there is some persistence. Indeed, empirical results from the generalized EWA model are robust to different choices for the initial attractions, as the comparison between the results reported in Section 4.4 and those reported in Bigoni (2009) shows.

Substituting equation (A.2) in equation (A.1), we get

$$
\begin{aligned}
A_i^j(t) =& \theta_2 x_{i1}^{2j} \prod_{s=2}^{t} x_{is}^0 + \theta_3 x_{i1}^{3j} \prod_{s=2}^{t} x_{is}^0 + \sum_{k=1}^{5} \theta_k \left(\sum_{s=2}^{t-1} (x_{is}^{kj} \prod_{l=s+1}^{t} x_{il}^0) + x_{it}^{kj} \right) \\
=& \sum_{k=2}^{3} \theta_k (x_{i1}^{kj} \prod_{s=2}^{t} x_{is}^0 + \sum_{s=2}^{t-1} (x_{is}^{kj} \prod_{l=s+1}^{t} x_{il}^0) + x_{it}^{kj}) \\
& + \sum_{k \in (1,4,5)} \theta_k (\sum_{s=2}^{t-1} (x_{is}^{kj} \prod_{l=s+1}^{t} x_{il}^0) + x_{it}^{kj}) \qquad t \geq 2
\end{aligned}
$$

Following Camerer and Ho (1999) and Salmon (2001), the probability P_{ijt} that individual i chooses strategy j at time t can be generally written as

$$
P_{ijt} = \frac{exp(f(A_i^j(t))}{\sum_j exp(f(A_i^j(t))} \tag{A.3}
$$

with $f(\cdot)$ monotone increasing, i.e. each individual chooses the strategy that has the highest attraction. The formulation typically adopted is

$$
P_{ijt} = \frac{exp(\lambda A_i^j(t))}{\sum_{j=0}^{30} exp(\lambda A_i^j(t))} \tag{A.4}
$$

with the restriction $\lambda > 0$.

Following the statistical literature on models for categorical responses we refer to $\lambda A_i^j(t)$ as the *index* or *linear predictor*. In our generalized EWA the index is

$$
f(A_i^j(t)) = \lambda A_i^j(t) = \lambda A_{i1}^j \prod_{s=2}^{t} x_{is}^0 + \sum_{k=1}^{5} \lambda \theta_k \left(\sum_{s=2}^{t-1} x_{is}^{kj} (\prod_{l=s+1}^{t} x_{il}^0) + x_{it}^{kj} \right) \qquad t \geq 2 \tag{A.5}
$$

4

and

$$\lambda A_i^j(1) \prod_{s=2}^{t} x_{is}^0 = \lambda\beta \frac{d_{i,BRC}^j(1)\pi_{i,BRC}^j(1)}{N(1)} \prod_{s=2}^{t} x_{is}^0 + \lambda\gamma \frac{d_{i,PC}^j(1)\pi_{i,PC}^j(1)}{N(1)} \prod_{s=2}^{t} x_{is}^0$$

$$= \lambda\beta x_{i1}^{2j} \prod_{s=2}^{t} x_{is}^0 + \lambda\gamma x_{i1}^{3j} \prod_{s=2}^{t} x_{is}^0$$

$$= \lambda\theta_2 x_{i1}^{2j} \prod_{s=2}^{t} x_{is}^0 + \lambda\theta_3 x_{i1}^{3j} \prod_{s=2}^{t} x_{is}^0$$

$$\text{(A.6)}$$

The likelihood for sample $(y_{ijt})_{i=1,\dots,N;t=2,\dots,40;j=0,\dots,30}$ is:

$$L(\lambda, \alpha, \beta, \gamma, \epsilon, \zeta) \propto \prod_{i=1}^{48} \prod_{t=2}^{40} \prod_{j=0}^{30} P_{ijt}^{y_{ijt}} \tag{A.7}$$

where y_{ijt} are binary indicators that take the value 1 if individual i at time t chooses strategy j and 0 otherwise. A convenient re-parametrization of the model to discuss parameters identification is $\tau_0 = \lambda$ and $\tau_k = \lambda\theta_k, k = 1, 2, \dots, 5$.[2] The log-likelihood corresponding to equation (A.7) becomes

$$l(\tau) = l(\tau_1, \dots \tau_5) \propto \sum_{i=1}^{48} \sum_{t=2}^{40} \sum_{j=0}^{30} y_{ijt} log(P_{ijt}) \tag{A.8}$$

Incorporating the expression for the probabilities P_{ijt} in equation (A.4), we

---

[2]The model log-likelihood function can be written as a function of $\tau_k, k = 1, 2, \dots, 5$.

get

$$l(\tau) \propto \sum_{i=1}^{48} \sum_{t=2}^{40} \sum_{j=0}^{30} y_{ijt} log\left(\frac{exp(\lambda A_i^j(t))}{\sum_{l=0}^{30} exp(\lambda A_i^l(t))} \right)$$

$$= \sum_{i=1}^{48} \sum_{t=2}^{40} \sum_{j=0}^{30} y_{ijt} \left(\lambda A_i^j(t) - log\left(\sum_{l=0}^{30} exp(\lambda A_i^l(t)) \right) \right)$$

$$= \sum_{i=1}^{48} \sum_{t=2}^{40} \sum_{j=0}^{30} y_{ijt} \lambda A_i^j(t) - \sum_{i=1}^{48} \sum_{t=2}^{40} (\sum_{j=0}^{30} y_{ijt}) log\left(\sum_{l=0}^{30} exp(\lambda A_i^l(t)) \right)$$

$$= \sum_{i=1}^{48} \sum_{t=2}^{40} \sum_{j=0}^{30} y_{ijt} \lambda A_i^j(t) - \sum_{i=1}^{48} \sum_{t=2}^{40} 1 \cdot log\left(\sum_{j=0}^{30} exp(\lambda A_i^l(t)) \right)$$

where the linear predictors $\lambda A_i^j(t)$ are given by equations (A.5) and (A.6).

The statistical properties of the maximum likelihood estimator for the parameters $\tau_k, k = 1, 2, \ldots, 5$ are discussed in McFadden (1974) who presents necessary and sufficient conditions for the existence of a maximum (Lemma 3) and its uniqueness. We now show that differentiation of the log-likelihood for our model yields conditions analogous to equation (19) and (20) in McFadden (1974).

Substituting the expression of the linear predictors $\lambda A_i^j(t)$ in the log-likelihood, we get

$$l(\tau) \propto \sum_{i=1}^{48} \sum_{t=2}^{40} \sum_{j=0}^{30} y_{ijt} \left((\tau_2 x_{i1}^{2j} + \tau_3 x_{i1}^{3j}) \prod_{s=2}^{t} x_{is}^0 + \sum_{k=1}^{5} \tau_k \left(\sum_{s=2}^{t-1} x_{is}^{kj}(\prod_{l=s+1}^{t} x_{il}^0) + x_{it}^{kj} \right) \right)$$

$$- \sum_{i=1}^{48} \sum_{t=2}^{40} log\left(\sum_{j=0}^{30} exp\left((\tau_2 x_{i1}^{2j} + \tau_3 x_{i1}^{3j}) \prod_{s=2}^{t} x_{is}^0 + \sum_{k=1}^{5} \tau_k \left(\sum_{s=2}^{t-1} x_{is}^{kj}(\prod_{l=s+1}^{t} x_{il}^0) + x_{it}^{kj} \right) \right) \right)$$

6

Let

- $\frac{\partial \lambda A_i^j(t)}{\partial \tau_k} = (\sum_{s=2}^{t-1} x_{is}^{kj} \prod_{l=s+1}^{t} x_{il}^0 + x_{it}^{kj}) \equiv w_{it}^{kj}, \quad k = 1, 4, 5$

- $\frac{\partial \lambda A_i^j(t)}{\partial \tau_k} = \sum_{s=1}^{t-1} x_{is}^{kj} \prod_{l=s+1}^{t} x_{il}^0 + x_{it}^{kj} \equiv z_{it}^{kj}, \quad k = 2, 3$

- $\bar{w}_{it}^k \equiv \sum_{j=0}^{30} P_{ijt} w_{it}^{kj}$

- $\bar{z}_{it}^k \equiv \sum_{j=0}^{30} P_{ijt} z_{it}^{kj}$

As a consequence, we can write

- $\frac{\partial P_{ijt}}{\partial \tau_k} = P_{ijt}(w_{it}^{kj} - \bar{w}_{it}^k), \quad k = 1, 4, 5$

- $\frac{\partial P_{ijt}}{\partial \tau_k} = P_{ijt}(z_{it}^{kj} - \bar{z}_{it}^k), \quad k = 2, 3$

The first derivatives of the log-likelihood function are:

$$\frac{\partial l(\tau)}{\partial \tau_k} = \sum_{i=1}^{48} \sum_{t=2}^{40} \sum_{j=0}^{30} y_{ijt} z_{it}^{kj} - \sum_{i=1}^{48} \sum_{t=2}^{40} \frac{1}{\left(\sum_{j=0}^{30} exp\left((\tau_2 x_{i1}^{2j} + \tau_3 x_{i1}^{3j}) \prod_{s=2}^{t} x_{is}^0 + \sum_{k=1}^{5} \tau_k w_{it}^{kj} \right) \right)}$$

$$\cdot \left(\sum_{j=0}^{30} exp\left((\tau_2 x_{i1}^{2j} + \tau_3 x_{i1}^{3j}) \prod_{s=2}^{t} x_{is}^0 + \sum_{k=1}^{5} \tau_5 w_{it}^{kj} \right) z_{ij}^{kt} \right)$$

$$= \sum_{i=1}^{48} \sum_{t=2}^{40} \sum_{j=0}^{30} y_{ijt} z_{it}^{kj} - \sum_{i=1}^{48} \sum_{t=2}^{40} \sum_{j=0}^{30} P_{ijt} z_{it}^{kj} = \sum_{i=1}^{48} \sum_{t=2}^{40} \sum_{j=0}^{30} \left((y_{ijt} - P_{ijt}) z_{it}^{kj} \right) \qquad k = 2, 3$$

exploiting the definition of P_{ijt}; similarly

$$\frac{\partial l(\tau)}{\partial \tau_k} = \sum_{i=1}^{48} \sum_{t=2}^{40} \sum_{j=0}^{30} (y_{ijt} - P_{ijt}) w_{it}^{kj} \quad k = 1, 4, 5$$

The derivatives above are analogous to those in equation (19) in McFadden (1974).

7

The second order derivatives are

$$\frac{\partial^2 l(\tau)}{\partial \tau_k \partial \tau_k} = -\sum_{i=1}^{48}\sum_{t=2}^{40}\sum_{j=0}^{30} w_{it}^{kj}\left(P_{ijt}\left(w_{it}^{kj} - \bar{w}_{it}^k\right)\right)$$

$$= -\sum_{i=1}^{48}\sum_{t=2}^{40}\sum_{j=0}^{30} P_{ijt}\left(w_{it}^{kj} - \bar{w}_{it}^k\right)^2 \quad k = 1, 4, 5$$

where the last equality follows from the fact that $\sum_j P_{ijt}(\bar{w}_{it}^k)^2 = \sum_j P_{ijt}w_{it}^{kj}\bar{w}_{it}^k$.

$$\frac{\partial^2 l(\tau)}{\partial \tau_k \partial \tau_k} = -\sum_{i=1}^{48}\sum_{t=2}^{40}\sum_{j=0}^{30} z_{it}^{kj}\left(P_{ijt}\left(z_{it}^{kj} - \bar{z}_{it}^k\right)\right)$$

$$= -\sum_{i=1}^{48}\sum_{t=2}^{40}\sum_{j=0}^{30} P_{ijt}\left(z_{it}^{kj} - \bar{z}_{it}^k\right)^2 \quad k = 2, 3$$

where the last equality follows from the fact that $\sum_j P_{ijt}(\bar{z}_{it}^k)^2 = \sum_j P_{ijt}z_{it}^{kj}\bar{z}_{it}^k$.

$$\frac{\partial^2 l(\tau)}{\partial \tau_l \partial \tau_k} = \sum_{i=1}^{48}\sum_{t=2}^{40}\left(\sum_{j=0}^{30} w_{it}^{kj} P_{ijt}(z_{il}^{lj} - \bar{z}_{it}^l)\right)$$

$$= -\sum_{i=1}^{48}\sum_{t=2}^{40}\left(\left(\sum_{j=0}^{30} z_{it}^{lj} P_{ijt}w_{it}^{kj}\right) - \bar{w}_{it}^k\bar{z}_{it}^l\right)$$

and

$$\frac{\partial^2 l(\tau)}{\partial \tau_k \partial \tau_l} = \sum_{i=1}^{48}\sum_{t=2}^{40}\left(\sum_{j=0}^{30} z_{it}^{lj} P_{ijt}(w_{it}^{kj} - \bar{w}_{it}^k)\right)$$

$$= -\sum_{i=1}^{48}\sum_{t=2}^{40}\left(\left(\sum_{j=0}^{30} z_{it}^{lj} P_{ijt}w_{it}^{kj}\right) - \bar{w}_{it}^k\bar{z}_{it}^l\right).$$

Hence, the matrix of the second order derivatives is symmetric.

$$\frac{\partial^2 l(\tau)}{\partial \tau_l \partial \tau_k} = \frac{\partial^2 l(\tau)}{\partial \tau_k \partial \tau_l}$$

$$= -\sum_{i=1}^{48}\sum_{t=2}^{40}\left(\left(\sum_{j=0}^{30} z_{it}^{lj} P_{ijt} w_{it}^{kj}\right) - \bar{w}_{it}^{k}\bar{z}_{it}^{l}\right)$$

$$= -\sum_{i=1}^{48}\sum_{t=2}^{40}\left(\left(\sum_{j=0}^{30} z_{it}^{lj} P_{ijt} w_{it}^{kj}\right) - \left(\left(\sum_{j=0}^{30} z_{it}^{lj} P_{ijt}\right)\left(\sum_{j=0}^{30} w_{it}^{kj} P_{ijt}\right)\right)\cdot 1\right.$$

$$\left. + \left(\sum_{j=0}^{30} z_{it}^{lj} P_{ijt}\right)\left(\sum_{j=0}^{30} w_{it}^{kj} P_{ijt}\right) - \left(\sum_{j=0}^{30} z_{it}^{lj} P_{ijt}\right)\left(\sum_{j=0}^{30} w_{it}^{kj} P_{ijt}\right)\right)$$

$$= -\sum_{i=1}^{48}\sum_{t=2}^{40}\left(\left(\sum_{j=0}^{30} z_{it}^{lj} P_{ijt} w_{it}^{kj}\right) - \left(\left(\sum_{j=0}^{30} z_{it}^{lj} P_{ijt}\right)\left(\sum_{j=0}^{30} w_{it}^{kj} P_{ijt}\right)\right)\cdot \sum_{j=0}^{30} P_{ijt}\right.$$

$$\left. + \left(\sum_{j=0}^{30} z_{it}^{lj} P_{ijt}\right)\left(\sum_{j=0}^{30} w_{it}^{kj} P_{ijt}\right) - \left(\sum_{j=0}^{30} z_{it}^{lj} P_{ijt}\right)\left(\sum_{j=0}^{30} w_{it}^{kj} P_{ijt}\right)\right)$$

$$= -\sum_{i=1}^{48}\sum_{t=2}^{40}\sum_{j=0}^{30} P_{ijt}\left(z_{i1}^{lj} w_{it}^{kj} - z_{it}^{lj}\bar{w}_{it}^{k} - \bar{z}_{it}^{l} w_{it}^{jk} + \bar{w}_{it}^{k}\bar{z}_{it}^{l}\right)$$

$$= -\sum_{i=1}^{48}\sum_{t=2}^{40}\sum_{j=0}^{30} (z_{it}^{lj} - \bar{z}_{it}^{l}) P_{ijt}(w_{it}^{kj} - \bar{w}_{it}^{k})$$

These derivatives are analogous to equation (20) in McFadden (1974). The matrix of the second order derivatives is the negative of a weighted moment matrix of the independent variables and thus is negative semidefinite. It follows that the log-likelihood function is concave in τ. The necessary and sufficient condition for the matrix of the second order derivatives to be negative definite (Axiom 5 in McFadden, 1974) is expected to hold in our case because the order condition is satisfied (i.e. number of observations larger to the number of parameters) and the data vary across alternative sets and are

not collinear.[3]

The properties of the estimator thus follow from McFadden (1974). More specifically, in our case, existence is guaranteed by the fact that: i) the order condition is satisfied; ii) the data vary across alternative sets and are not collinear; iii) the sample size is large, thus the probability that Axiom 6 (McFadden, 1974) fails approaches zero.

While λ is not identified, provided that there is enough variability in the data, we can identify all the parameters that appear in the model likelihood, namely $\tau_1 = \lambda\alpha$ $\tau_2 = \lambda\beta$, ... $\tau_5 = \lambda\zeta$. Under the assumption that $\lambda \neq 0$, the ratio between these parameters is informative on the relative strength of each component in determining the attraction of strategy j.

---

[3]This is confirmed by the fact that we can show non-singularity of the estimated information matrix (Skrondal and Rabe-Hesketh, 2004, chpt. 5). We also changed the initial conditions for maximum likelihood estimation over the interval [-5,5] for all τ parameters: maximum likelihood estimates are unaffected. Descriptive statistics on the variables z_{it}^{kj}, w_{it}^{kj} suggest that the within-subject, within-period variability is substantial, even if it decreases between period 2 and period 40.

Appendix B. Instructions

Welcome to this experiment about decision making in a market. The experiment is expected to last for about 1 hour and 15 minutes. You will be paid a minimum of 4 Euro for your participation. In addition to this, you can earn up to 20 Euro if you make good decisions.

We will first read the instructions aloud. Then you will have time to read them on your own. If you have questions, raise your hand and you will be helped privately. 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 manager of a firm which produces and sells a given product: your task consists of deciding how many product units to put on the market in every period.

Your firm has **three competitors** that sell on the same market a product which is exactly identical to yours. Your competitors are three from among the participants in the experiment taking place today in this room, but you will not have the opportunity to discover who they are, not even at the end of the game. Your identity will be kept secret as well.

The experiment consists of **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 those of your three competitors.

Every period lasts **30 seconds**: if in a period you fail to make your choice within the time limit, the computer will automatically set the number of units produced by your firm in that period equal to 0, and your profit in that period will be equal to 0 too.

Price, costs and profits. The market **price** at which you will be able to sell your product will be the higher, the smaller the total number of product units your firm and your competitors put on the market; if the total number of product units sold on the market is sufficiently high, the price will be equal to zero.

No product unit remains unsold: all the product units you put on the market will be purchased by consumers at the market price.

To produce, you will have to bear a **production cost** which will be the higher, the more product units you put on 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 endowment of 2000 points. At the end of each period, your per-period profits or your possible losses will be added to your total profit, which will always be displayed in the top-right corner of the screen. Notice that your total profit cannot become negative.

At the end of the game, your total profit will be converted in Euros, according to the rate:

1000 points = 1 Euro

The corresponding amount of money will be payed to you in cash, privately, at the end of the session. Remember that, in addition, you will be payed 4 Euros for your participation.

Information at your disposal. At the top of your computer screen you can read:

1. the number of periods elapsed since the game began (top left corner)

2. your total profit (top right corner)

3. the number of seconds (top, center) you still have at your disposal to take a decision. Remember that every period lasts 30 seconds, and if you do not take a decision in time it will be as if you decided to produce 0 units and in that period your profit will be equal to 0.

Before choosing how many units to produce, you will have the opportunity to look at some information on market characteristics and on what happened in the previous periods.

In particular, in every period following the first one, you will be informed about the profits obtained in the previous period by your firm and by your competitors. Moreover, you will be able to get more information about:

1. the quantity produced in the previous period by each of your competitors;

2. the quantities produced and the profits obtained by your firm in each of the previous periods: this information will be displayed both by means of a plot and in a table;

3. the quantity produced on the whole by each of your three competitors in the previous periods: this information will also be presented both by means of a plot and in a table.

In addition, you will have the opportunity to use a **profit calculator**, a device you can use to better understand how the market works. The profit calculator has two functions:

1. to evaluate your profit, given the number of units produced by your firm and the number of units produced on the whole by your competitors.

2. to evaluate the maximum profit you could earn — and the number of units your firm should produce in order to get such profit — given the number of units produced on the whole by your competitors.

Progress of the experiment. When the reading of these instructions is over, you will have the opportunity to ask for clarifications about the aspects of the experiments which are unclear.

When we have answered all possible questions you will be asked to complete a test on your computer, which will allow us to check that you have fully understood the instructions, and give you the opportunity to get to grips with the software used in this experiment. The answers you give in this test will not affect your earnings in any way, nor they will influence any other aspect of the experiment. During the test, you will still have the opportunity to ask questions, as always by raising your hand.

When all the participants have completed their test, the real experiment will begin. The computer will randomly generate groups of four people; every participant in the experiment will belong to one and only one group during the whole experiment. The other three members of the group you belong to are your competitors, who then remain the same over all the 40 periods of the game.

Every period lasts at most 30 seconds. The maximum length of the game therefore is approximately 20 minutes.

At the end of the fortieth period the game will end, and the points scored by each of the participants will be converted into Euros.

Before being paid privately, you will be asked to answer a short questionnaire about the experiment, and you will be obliged to hand back the

instructions.

THANK YOU VERY MUCH FOR PARTICIPATING IN THIS EXPERIMENT AND GOOD LUCK!

Appendix C. Graphical Interface
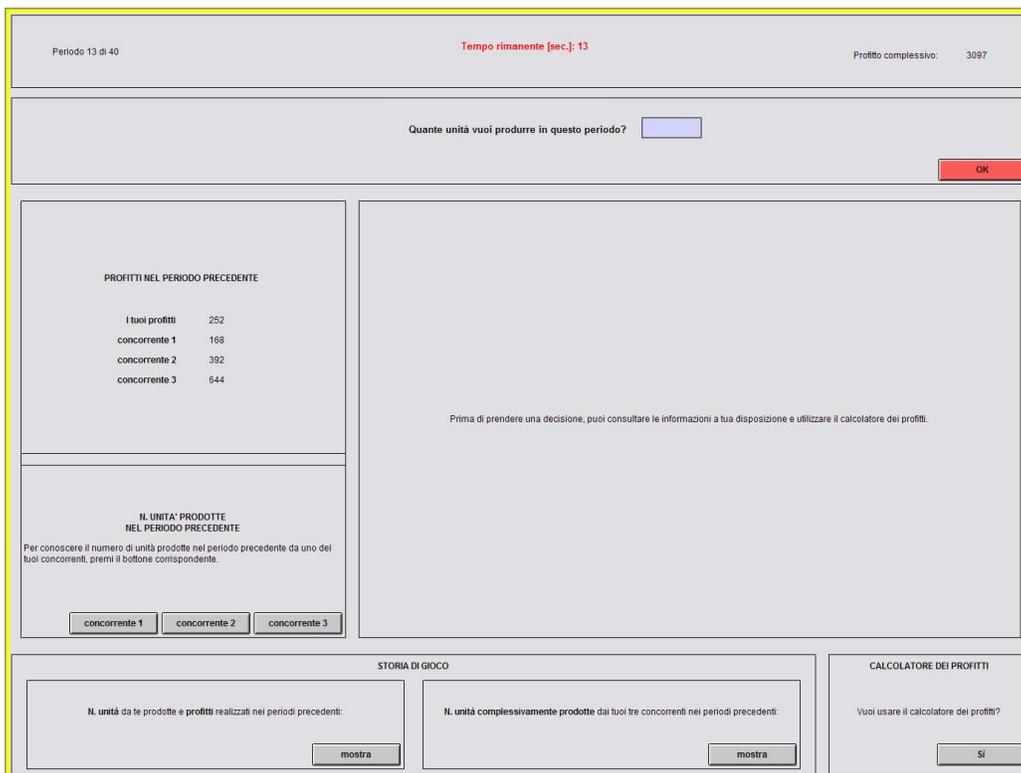
Figure C.1: Graphical interface

Translation. From top to button, left to right. [ ] indicate a button.

bar at the top: period 13 out of 40, remaining time [sec.]: 13, total profit: 3097

box at the top: how many units do you want to produce in this period? [OK]

first box on the left: Profits in the previous period

your profit: 252

16

competitor 1: 168

competitor 2: 392

competitor 3: 644

second box on the left: # of units produced in the previous period.

To know the number of units produced in the previous period by one of your competitors, push the corresponding button. [competitor 1] [competitor 2] [competitor 3]

center-right box: before taking a decision, you can look at the information at your disposal and use the profit calculator.

bottom-left box: history of play

# of units you produced and profits you obtained in the previous periods [show]

# of units produced on the whole by your three competitors in the previous periods [show]

bottom-right box: profit calculator

do you want to use the profit calculator? [yes]