IMT Institute for Advanced Studies, Lucca

Lucca, Italy

Theoretical and Experimental Economic Analysis of Collusion

PhD Program in Economics, Markets, Institutions

XX Cycle

By

Maria Bigoni

2008

The dissertation of Maria Bigoni is approved.

Program Coordinator: Prof. Fabio Pammolli, University of Florence

Supervisor: Giancarlo Spagnolo, University of Rome Tor Vergata, SITE Stockholm School of Economics

Tutor: Giancarlo Spagnolo, University of Rome Tor Vergata, SITE Stockholm School of Economics

The dissertation of Maria Bigoni has been reviewed by:

Magnus Johannesson, Stockholm School of Economics

Andreas Ortmann, CERGE-EI, Charles University and Academy of Sciences of the Czech Republic

IMT Institute for Advanced Studies, Lucca

2008

Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my deep appreciation to all people who provided me with advice and support during the work on this dissertation.

I am indebted to my supervisor Giancarlo Spagnolo, for his guidance, for his patience and for the enthusiasm for research he managed to pass on me. I am also grateful to Carlo Cambini, who has encouraged and motivated me throughout the three years of my doctoral studies.
I thankfully acknowledge all useful insights from other professors at IMT, as well as at the Stockholm School of Economics and at the University of Trento, where I had the possibility to spend some time as a visiting student. In particular I thank Riccardo Cambini, Andreas Ortmann and Magnus Johannesson, whose teachings and comments have been precious for the development of my work. Also special thanks to Chloè Le Coq and Sven-Olof Fridolfsson, who accepted to work with an unexperienced Ph.D student, guided my first steps into lab experiments, and made my research possible. My sincere gratitude goes to Giacomo Calzolari, for his indispensable help in organizing my experiment in Bologna.
Of course, all errors are mine.

I would not have completed my first year coursework, without the friendship and support of my friends in Lucca. Thanks for your camaraderie in the course of more serious pursuits.

Finally, this dissertation would not have been possible without my family and my friends at home, the best support network one could hope for. They endured and excelled through my frequent physical and mental absences, gave me a life to return to and anchored me in reality whenever I went a little too far. Thank you to my parents, who are my main mentors, and to Lorenzo, whose love, understanding and help have been as determinant for the completion of my work as they are for my life.

Vita

September 3, 1980	Born, Bologna, Italy
2002	Degree in Economics and Finance Final mark: 110/110 cum laude University of Bologna, Italy
2004	Laurea Magistralis in Economics Final mark: 110/110 cum laude University of Bologna, Italy
2004-2005	Paid intern Prometeia S.R.L. Bologna, Italy
2005-2008	Ph.D Program in Economics, Markets and Institutions I.M.T. Institute for Advance Studies Lucca, Italy
2006	Visiting Ph.D Student University of Trento Trento, Italy
2006	Visiting Ph.D Student Stockholm School of Economics Stockholm, Sweden
2006	Visiting Ph.D Student Université Toulouse 1 Toulouse, France
2007	Visiting Ph.D Student Stockholm School of Economics Stockholm, Sweden

2007-2008	Teaching Assistant in Microeconomics University of Bologna, Laurea Magistralis in Economics Bologna, Italy
2008-2009	Research Fellow University of Padua Padua, Italy

Publications

1. M. Bigoni, "Informazione a priori ed insorgenza dell'equilibrio: un modello di simulazione" *(A priori information and the onset of equilibrium: a simulation model)* in *Sistema Economics*, vol. 3, 2004.

Presentations

1. M. Bigoni, "Information and Learning in Oligopoly: an Experiment" at the *Conference of the French Economic Association on Behavioral and Experimental Economics*, Lyon, France, May 23-25, 2007.
2. M. Bigoni, S.O. Fridolfsson, C. Le Coq, G. Spagnolo, "Prospect Theory and Strategic Risk Considerations in Optimal Law Enforcement: Evidence from an Antitrust Experiment" at the *ESA 2007 World Meeting*, Rome, Italy, June 28- July 1, 2007.
3. M. Bigoni, S.O. Fridolfsson, C. Le Coq, G. Spagnolo, "Strategic Risk, Prospect Theory and Deterrence: Experimental Evidence" at the *Second Nordic Workshop on Behavioral and Experimental Economics*, Gothenburg, Sweden, November 16-17, 2007.

Abstract

This research project consists in a theoretical and experimental study of oligopolistic markets. I believe that an experimental approach to this subject might help to understand the interplay of the many different factors that affect firms behavior in this context, and to analyze problems which theory does not provide a clearcut answer to.
My project develops into two main parts: the first concern how firms can learn to collude tacitly, the second studies deterrence of explicit collusion.

Learning to collude The first part of this research concerns the relation between the process of information search and players' behavior in a repeated Cournot oligopoly. The main question I try to answer is what happens when information acquiring and processing is too difficult or too costly for the firm to behave according to the perfect rationality paradigm.

First, I present a review of the literature on bounded rationality in general – to provide a broad theoretical framework – and in particular on learning, which is the aspect I will mainly focus on.
The core of this part of my research consists in two experiments designed to study the relation between the process of information search and learning in a Cournot oligopoly, with limited *a priori* information. As the review of the literature will show, different theories of learning have been applied to this setting, each yielding a specific market outcome in the long run, and postulating specific informational requirements.
By allowing players to choose the information they wish to acquire, and controlling for these choices, I study the features of the learning

model actually followed by the subjects, and the relation between the information they gather and the market behavior they adopt.
According to my results, learning appears to be a composite process, in which different components coexist. Belief learning seems to be the leading element, as subjects try to form expectations about their opponents' future actions and to best reply to them. When subjects also look at the strategies individually adopted by their competitors, though, they tend to imitate the most successful behavior, which makes markets more competitive. Finally, reinforcement learning also plays a non-negligible role, as subjects tend to favor strategies that have yielded higher profits in the past. I show that these different elements may be usefully incorporated into a more sophisticated learning model, shaped after self tuning EWA learning model.

Deterring collusion The second experimental study is more policy oriented, and concerns the optimal law enforcement against explicit cartels. This part of the project is a joint work with Giancarlo Spagnolo, Chloé Le Coq and Sven Olof Fridolfsson, and has been financed by the Swedish Competition Authority.
In this study, we examine the effects of fines, leniency programs and reward schemes for whistleblowers on firms' decision to form cartels (cartel deterrence) and on their price choices.
Leniency policies and rewards for whistleblowers are being introduced in ever more fields of law enforcement, though their deterrence effects are often hard to observe, and the likely effect of changes in the specific features of these schemes can only be observed experimentally.

Chapter 7 reports results from an experiment designed to examine the effects of fines, leniency programs, and reward schemes for whistleblowers on firms' decision to form cartels (cartel deterrence) and on their price choices. Our subjects play a repeated Bertrand price game with differentiated goods and uncertain duration, and we run several treatments different in the probability of cartels being caught, the level of fine, the possibility of self-reporting (and not paying a fine), the existence of a reward for reporting, the option to communicate, and cartel leaders access to leniency. We find that fines following successful investigations but without leniency have a deterrence effect (reduce the number of cartels formed) but also a pro-collusive effect (increase collusive prices in surviving cartels). Leniency programs might not be more efficient than standard antitrust enforcement, since in our experiment they do deter a significantly higher fraction of cartels from forming, but they also induce even higher prices in those cartels that are not reported, pushing average market price significantly up relative to treatments without antitrust enforcement. With rewards for whistle blowing, instead, cartels are systematically reported, which completely disrupts subjects' ability to form cartels and sustain high prices, and almost complete deterrence is achieved. If the ringleader is excluded from the leniency program the deterrence effect of leniency falls and prices are higher than otherwise. As for tacit collusion, under standard anti-trust enforcement or leniency programs subjects who do not communicate (do not go for explicit cartels) tend to choose weakly higher prices than where there is no anti-trust enforcement. We also analyze post-conviction behavior, finding that after convictions caused by a report under the leniency program much fewer cartels form and prices are much lower than when conviction is due to an independent antitrust investigation. Finally, we find a strong cultural effect comparing treatments in Stockholm with those in Rome, suggesting that optimal law enforcement institutions differ with culture.

In chapter 8 we investigate the effects of risk preferences and attitudes towards risk on optimal antitrust enforcement policies. First, we observe that risk aversion is negatively correlated with players' proclivity to form a cartel, and that increasing the level of fines while reducing the probability of detection enhance deterrence. This confirms that the design of an optimal law enforcement scheme must keep risk attitudes into account, as suggested by Polinsky and Shavell [83].

We also notice that players' propensity towards communication drops right after detection even if the collusive agreement was successful, and it declines as the sum of the fines paid by a subject increases. This effect could be explained by availability heuristic [66] – a cognitive bias, where people's perception of a risk is based on its vividness and emotional

impact rather than on its actual probability.
Our results also confirm the crucial role of strategic risk considerations [10] (analogous to risk dominance for one shot games) in determining the effects of leniency programs. Indeed, we show that the effectiveness of leniency programs in deterring cartels is mostly due to the increased risk of a cartel member being cheated upon when entering a collusive agreement, while the risk of a cartel being detected by an autonomous investigation of the Authority seems to play a less important role.

Part I

Learning to collude

Chapter 1

Introduction

> *Recent developments in economics [...] have raised great doubts as to whether this schematized model of economic man provides a suitable foundation on which to erect a theory – whether it be a theory of how firms do behave or of how they 'should' rationally behave.*
> Herbert Simon, 1955

The debate about bounded rationality in economics begun with the widely known works written in the mid-1950s by Herbert Simon [92; 93], who first extensively investigated the problems connected to those complex models of strict rationality which are so pervasive in economic theory. Since then, though, several decades passed before bounded rationality were given the formal approach which was necessary for it to really affect mainstream economic theory. It was only in the 1980s that this topic bloomed and gave birth to a wide and diversified literature.

On the one hand, as noticed by Robert Aumann [5]:

> an important factor in making this possible was the development of computer science, complexity theory, and so on, areas of inquiry that created an intellectual climate conducive to the development of the theory of bounded rationality.

On the other hand, in 1979 Amos Tversky and Daniel Kahneman published their work on prospect theory in Econometrica [67] drawing economists attention onto the "fascinating compilation of experimental data demonstrating the circumstances under which rationality breaks down and other patterns of behavior emerge" (Rubinstein, [87]) and giving birth to a new research field, known as "psychology and economics".

During the last twenty years these topics have been approached by scholars in ways that were increasingly divergent – as I will try to show more in detail in the first part of Chapter 2 – and had a considerable impact on various fields of economics: from finance to marketing, to the economic analysis of law.

What I plan to do here is to investigate whether they also affected the theory of industrial organization and if so, how. A rather complete and up-to-dated work has been recently done by Ellison [33] who wrote a critical review of the literature focused on consumer irrationalities that rational firms might exploit. I would like to follow his steps but exploring the other side of the problem; that is, studying if models of bounded rationality can help us to understand what happens when firms

(i) do not have complete information about the environment they act in and have to learn which "game" they are playing and how to play it;

(ii) face costs of getting and processing information and/or

(iii) suffer for computational limitations and bounds.

More specifically, I will focus on the consequences that this kind of hypothesis can have on the traditional setting of the theory of oligopoly and tacit collusion.

The classical approach to the theory of tacit collusion entails a model of repeated interaction between firms which are active on the same market and compete in at least one dimension. In the basic model market demand and cost functions are common knowledge, and firms are able to predict what their profits will be indefinitely in the future, under different conditions.

In practice, firms should be able to determine which is the optimal *joint profits maximizing* price/quantity that each of them should set. If only one firm deviates from the collusive action, for the first period it will get higher profits, possibly close to monopoly ones, but if the other firms can *detect* this defection and are able to *punish* the deviating firm, they will surely retaliate in the following periods, so future profits for the deviating firm will be lower than those it would have got if it had not broken the tacit agreement. So every competitor in the market knows that each firm *i* is willing to stick to the collusive action as long as the discounted stream of collusive profits $\pi^C_{t,i}$ is higher than or at most equal to a threshold determined by the deviation payoff $\pi^D_{t,i}$ plus the discounted flow of payoffs the firm *i* will get during the "punishment" phase $\pi^P_{t,i}$.
Supposing that the punishment phase will start immediately after the defection and last forever, this means that:

$$\sum_{t=0}^{\infty} \delta^t_{t,i} \pi^C_{t,i} \geq \pi^D_{0,i} + \sum_{t=1}^{\infty} \delta^t_{t,i} \pi^P_{t,i}, \quad \pi^D_{t,i} > \pi^C_{t,i} > \pi^P_{t,i}, \quad \forall t, i$$

where $\delta_{t,i}$ represents the discount factor for firm *i* in period *t*.

The simplest situation, which can be considered as a baseline model, is the one in which two symmetric firms sell a homogeneous product and compete *à la Bertand* in a "frozen" market where demand and cost functions never change and where both competitors have the same constant discount factor δ. If the firms collude, both of them will be able to get half the monopoly profit π^M while the optimal deviation from the collusive price would ensure to the defecting firm the monopoly profit for the first period. In the subsequent period, though, the other firm would detect the deviation and trigger a never-ending price war, that would lead both firms to revert to the competitive price $p = c$ getting zero profits.
In this case collusion is sustainable if

$$\sum_{t=0}^{\infty} \delta^t \frac{\pi^M}{2} \geq \pi^M$$

that is if

$$\delta \geq \delta^* = \frac{1}{2}$$

The threshold δ^* appears to be a crucial element in this theoretical approach; as Ivaldi *et al.* notice [55]:

> Collusion is easier to sustain when this threshold is lower (then, even "impatient" firms with a lower discount factor could sustain collusion), and more difficult to sustain if this threshold is higher (in that case, even firms that place a substantial weight on future profits might not be able to sustain collusion). The determination of this critical threshold thus provides a natural way for assessing the scope for collusion.

In the baseline model it is very easy to determine the value of δ^*; in more complex environments the evaluation can be much less straightforward. To evaluate the potential for collusion in a specific market, one should look at all the industry characteristics that could raise this critical threshold and to the *facilitating factors* that could reduce it, and try to establish which force will predominate in the end[1]. A significant theoretical work has been done to provide the analytical tools necessary to make this kind of evaluations, and a number of different elements has been considered:

- from *basic structural variables* – such as the number of competitors, entry barriers, frequency of firms' interaction, and market transparency
- to the characteristics of the *demand side* – is the market growing, stagnating, or declining? Are there significant fluctuations or business cycles?
- and of the *supply side* – Is the market driven by technology and innovation, or is it a mature industry with stable technologies? Are firms in a symmetric situation, with similar costs and production capacities, or are there significant differences across firms? Do firms offer similar products, or is there substantial vertical or horizontal differentiation?

[1] For a wide discussion about this topic see Ivaldi *et al.* [55], and Motta, chapter 4 [75]

Each of the mentioned aspects is important to understand the mechanics of tacit collusion; but besides the industry characteristics, it could be useful to study also the features of the firms' decision process and the way in which they could affect the sustainability of tacit collusive agreements. This alternative approach to the problem is to be seen as a complement – not as an antagonist – of the previous one, as it could offer an insight on other factors that the regulator or the antitrust authorities might want to keep in mind when evaluating a specific market context. In particular looking at imperfect competition in a bounded rationality setting becomes more relevant when one wants to investigate:

- what is the minimal information requirement for the firms to reach a collusive agreement,
- how the "optimal" choice made by the firms can be affected by the costs related to strategic complexity or to the activity of processing information.

Suppose for instance that firms don't have the information required in classical models to sustain collusion. There are at least two ways in which this can be true: uncertainty about the market demand or about the opponents' costs makes it difficult to coordinate on the joint-profit maximizing outcome and the lack of transparency about the competitors' past actions constitutes an obstacle to the detection of defections. One may then wonder whether firms are able to "learn" from past experience and to get to collusion anyway – if the context is stable enough. If so, it would be also interesting to know how long this "learning phase" is and which factors could affect its length.

On the other hand there are also situations in which it could be "rational to be irrational": the costs a firm would have to face for the complete computation of each possible future scenario could be high enough to force it to a less precise forecast or to use a simpler strategy. Suppose, for instance that a firm has to face some costs for forecasting. Could it "rationally" decide to be "boundedly rational" in the sense of making predictions just till a certain period in the future? And if so: what kind of consequences would it have on market dynamic and equilibrium?

These are the main reasons that drew my interest on the topics I study in this firs part of the thesis. There is though a second issue that I want to stress here and which is connected to the previous ones: as Chapter 2 will show, many of the models that incorporate bounded rationality assumptions into oligopolistic settings have a truly *dynamic approach* which is usually absent in the traditional approach to the problem[2] . This provide us with a useful tool to study the building up phase of a collusive agreement, to understand more deeply what happens off the equilibrium path in order to predict in which direction the system will move when it is not in a steady state.

Among the many models of strategic interaction in oligopoly, I decided to focus on the Cournot model. This choice has been driven by a number of considerations: first of all Cournot model "is one of the simplest of a large number of market models whose predictions rest crucially on the notion of equilibrium. Many analyses of the effects of horizontal mergers are based upon the Cournot framework, and much of applied industrial organization uses the Cournot model as a benchmark." (Rassenti et al. [84]); but the model is interesting also because, how Offerman et al. [81] observe,

> although quantity-setting oligopoly is one of the "workhorse models" of industrial organization (Martin, 1993), empirically there is much ambiguity about its outcome. A recent survey by Slade (1994) indicates that most empirical studies reject the hypothesis that the outcome is in line with the Cournot–Nash equilibrium of the corresponding one-shot game. Interestingly, however, outcomes on both sides of the Cournot–Nash outcome are found. In the experimental literature, a similar state of affairs obtains. Many experimental oligopoly games result in higher than Cournot–Nash production levels, some result in lower production levels (Holt, 1995).

[2] see, for example, Aumann [5] "The difficulty is that ordinary rational players have foresight, so they can contemplate all of time from the beginning of play. Thus the situation can be seen as a one-shot game, each play of which is actually a long sequence of 'stage games', and then one has lost the dynamic character of the situation."

To pursue the goal of studying firms' behavior in oligopoly under the hypothesis of bounded rationality, I first did a broad research on the existing – both theoretical and experimental – literature about this subject, which is presented in chapter 2. In particular, I focused on several models of learning, which are obvious candidates to represent firms' behavior in a repeated game with imperfect and incomplete information, and limited or constrained computational capabilities, such as the one we are analyzing.
From this survey, it emerged that the relation between the process of information search and the model of learning adopted by the firms has a determinant impact on firms' actual market choices in a repeated Cournot oligopoly, and even if some experiments have already been run to investigate this topic, the results reached so far are not clearcut nor fully conclusive. The last part of the literature review concerns a new stream of experiments investigating the process of information acquisition as a way to understanding the heuristics and mental processes underlying peoples' decision making. These works are particularly important for my research, whose main novel contribution consists precisely in combining the study of learning with an experimental analysis of the way subjects select the information they need before choosing their strategy.

More specifically, I adopted this approach in two experiments about learning in a Cournot oligopoly setting. Instead of comparing subjects market behavior under different informational frameworks – which is the approach adopted in all the previous experiments about this topic – I provide the players with a broad range of information, but force them to choose only some pieces of it. The players' process of information gathering is strictly (but non obtrusively) controlled, by means of a special software, originally called MouseLab and developed by Eric J. Johnson et al. (1988) [63].
Paying attention not only to what players do but also to what they know, it is possible to better understand the mental mechanisms which guide their choices and consequently the impact that the informational framework has over their behavior.

The framing of the two experiments, described in Chapter 3, is the same, and is pretty close to the one adopted in previous experiments on the same topic. In the first experiment, presented in chapter 4, players face virtual opponents enacted by computer programs. This allows me to control for the learning rule adopted by players' opponents and to check if and how this affects the information search and market behavior of the subjects. In the second experiment (chapter 5) subjects interact with each other.

My results confirm that information gathered by the subjects affect their choices through a mechanism of learning, which though appears to be a composite process, in which different components coexist. Belief learning seems to be the leading element, as subjects try to form expectations about their opponents' future actions and to best reply to them. When subjects also look at the strategies individually adopted by their competitors, though, they tend to imitate the most successful behavior, which makes markets more competitive. Finally, reinforcement learning also plays a non-negligible role, as subjects tend to favor strategies that have yielded higher profits in the past. I show that these different elements may be usefully incorporated into a more sophisticated learning model, shaped after self tuning EWA learning model.

My conclusions are presented in chapter 6, where I also mention possible future developments of this research.

Chapter 2

Literature Review

This chapter will be devoted to a roundup of the many theories developed to represent and explain the behavior of boundedly rational players in games, which can possibly be used also to design models of oligopoly of the kind I am interested in. First, though, it can be worthwhile to dwell for a while on the main theoretical approaches developed during the last twenty years to incorporate into the economic theory the relatively new hypothesis that agents are not always to be represented as fully rational, or even "hyperrational", as they used to be till then. Time and space constraints do not allow this to be an exhaustive description, while each of the theories I will mention deserves a much wider coverage. Anyhow, following Rubinstein [88], Aumann [5] and In-Koo Cho [24], I try at least to draw a rough distinction between the main theoretical approaches to bounded rationality, and to sketch some of the most important elements that characterize each of them.

I will then focus on the principal models of learning, devoting special attention to those having been already applied to the oligopoly setting I am interested in. Finally, I will present some experiments which are related to my research, either because they study a similar situation (section 2.3), or because they adopt a similar technique (section 2.4).

2.1 Main Theoretical Approaches

Behavioral Economics This approach sprang from a project initially launched by Daniel Kahneman and Amos Tversky who not only refuted the standard use of the economic man paradigm but also identified psychological elements which are systematically used by decision makers. Their findings demonstrated that *emotions and procedural elements* ("heuristics and biases") – which were missing from the standard application of rationality in economics – are indeed involved with the human decision process. Researchers in this field usually preserve the assumption that the economic agent is rational in the economic sense of maximizing a target function; however they do not feel obliged to define the targets as material rewards. On the one hand, agents in these models maximize a utility function which also reflects agents' culture and psychological motives like fairness, envy and reciprocity, on the other hand there have been many attempt to design an objective function which reflects some behavioral regularities observed in the lab, such as loss aversion, reflection effect and changing attitudes towards risk depending on the size of the stakes.
Experimental methods are widely used, not only to test but also to develop new theories[1]. Note also that, in contrast to the "bounded rationality" approach, for the most part behavioral economics does not relate to the procedural elements of decision making.

I decided not to follow this approach in this first part of the project basically because I do not think it fits with the aim of my research: in fact while such an interest in psychological motives and in behavioral biases can have an important role also for IO when one focuses on how consumers behave and how firms – that are modeled as rational – can play on them, I think that it could not be of much use when one wants to represent simple strategic interactions between firms.[2] Besides,

[1] see for example the way in which Camerer and Ho set the parameters for their Experience-weighted Attraction (EWA) Learning model, described in Camerer, chapter 6 [18]

[2] To my knowledge the most interesting work on oligopolistic models that is somehow close to this approach has been developed by Friedman and Mezzetti, who propose a

even if it is true that many lab experiments confirmed the presence of "anomalies", that is of attitudes or conducts which are inconsistent with what traditional theory predicts, it has been also shown that these anomalies tend to disappear when there are large amounts of money at stake and when players are experienced or well trained (and I think that this is the case, when players are firms).

An Evolutionary Approach This second approach, in contrast to the other two, treats agents as automata, merely responding to changing environments and lacking any power of deliberating about their decisions. It imagines that the game is played over and over again by biologically or socially conditioned players who are randomly drawn from a large population. Each player is programmed to follow a particular behavior, representing its "genetic endowment", and there is no sense in which an individual can "choose" its own genetic endowment. Note, though, that single agents in this environment are not to be considered as player in the classic game-theoretical acceptation, because here, actually, the "players" are the populations – i.e., the species. Individuals indeed play no explicit role in the mathematical model; they are swallowed up in the proportions of the various pure strategies.
This theory studies the link between the classical concepts of equilibria of the games and the aggregate behavior that emerges from an *evolutionary process* that operates over time on the population distribution of behaviors.

This approach works completely out the problem of representing limited computational capabilities of agents – because actually here they do not have any – but of course it has not been developed with the aim of investigating the characteristics of decision processes and their impact on market outcomes; rather, the purpose seems to be to provide an insight into the dynamic aspects of game-like situations and to demonstrate that

dynamic model of n-firm oligopoly in which each firm solves a dynamic optimization problem believing that the other firms will alter their future choices in proportion to its own current change [39]. In the second part of the project, and in particular in chapter 8, we will see that a behavioral approach turns out to be useful in understanding the motives that induce subjects to comply or not to antitrust law.

there exist equilibrium concepts that do not strictly depend on those assumptions of full rationality and common knowledge that have been widely questioned. The situation that I intend to study, namely the presence of imperfect competition in oligopolistic markets, is hardly representable under this framing, because it requires to investigate the repeated interaction between a small number of firms who maybe are not perfectly rational in the traditional economic sense, but for sure follow an intentional strategy. Nonetheless I will not completely abandon this approach, because it shares some features with the learning models that will play a considerable role in my research project.

Bounded Rationality The term "bounded rationality" has a sort of double meaning, and the time has come to cope with this ambiguity: in what precedes the term has been adopted with a broad acceptation to describe all those theories that remove the hypothesis of full rationality of economic agents. Here instead it is used in a more restrictive way and refers to an approach which is not based on experimental evidence, like behavioral economics tends to be, but more on casual observations of the way in which people deliberate, and more generally of decision making processes. Models here are mainly intended to increase our understanding of the effect of *decision-procedural elements* on the outcome of an economic interaction. The border between "behavioral economics" and "bounded rationality" sometimes appears to be evanescent and the categorization of models into one or the other framework is not always straightforward. Probably, the main difference lays in that the first approach usually preserves the hypothesis of optimizing agents, while in the second one models generally do not even present a well defined objective function.
The pioneering works pertaining to this approach have been done since the mid-eighties on automata and Turing machines playing repeated games. These works were essentially intended to evaluate the effects of bounds on the complexity of strategies that players can use. In one strand, pioneered by Neyman [77], the players of a repeated game are limited to using mixtures of pure strategies, each of which can be programmed on

a finite automaton with an exogenously fixed number of states. This is reminiscent of Axelrod's [7] famous contest on the prisoner's dilemma game, who required the entrants in his experiment to write the strategies in a Fortran program not exceeding a stated limit in length. In another strand, pioneered by Rubinstein, the size of the automaton is endogenous; computer capacity is considered costly, and any capacity that is not actually used in equilibrium play is discarded.
Later on other, rather different models have been developed to study how decision processes are affected by diverse cognitive limitations – such as bounded recall[3] and limited capability of making forecast[4] – or to investigate how agents can possibly approach and simplify problems when they turn out to be too complicated for their cognitive abilities. Consider, for example the "case based decision theory" developed by Gilboa and Schmeidler [42] who propose a decision rule that chooses a "best" act based on its past performance in similar cases, and Jehiel's "analogy based expectation equilibrium" [61], based on the assumption that agents bundle nodes at which other agents must move into analogy classes, and only try to learn the average behavior in every class; this model predicts that in equilibrium at every node players choose best-responses to their analogy-based expectations, and expectations correctly represent the average behavior in every class.

In my opinion this field of research is quite fertile for further developments on the topics I am investigating, and for this reason in what follows I will focus mainly on models that pertain to it.

2.2 Learning Models

I now turn to examine more in detail the literature on learning, dedicating particular attention to those model that have been already applied to oligopoly, that in many cases have been also tested in the lab. Learning models are important in this setting for at least two reasons.

First they allow to formalize the behavior of subjects who do not have

[3]See Aumann and Sorin [6]

[4]See Jehiel [60; 58; 59]

full information on the game they are playing. The amount of information required by the learning rule on which firms base their behavior is not the same in all the models: in some extreme cases – like models of learning based on trial and error – it is hypothesized that they do not even know their own payoff matrix and have to learn the "rules of the game" just playing it repeatedly and observing the relation between the action they choose and the result they get. On the other hand, in "belief learning" models demand and cost functions are common knowledge and each firm can observe the opponents' moves. Because of limited cognitive abilities, though, they are not able to infer immediately the equilibrium strategy and they just play a best reply to a belief about other players' future moves that in turn depends on what they did in the past. The models also make different predictions about the equilibrium outcome of the learning processes, so if one could establish a relation between the amount and the nature of information hypothesized and the result in terms of collusion/competition in the market, then it would be possible to develop a theory which could provide some important policy implications in terms of optimal information disclosure. As we will see in section 2.3, some experimental economists have attempted to follow this path, but the results they got are somehow contrasting and not so clearcut, and some work is still to be done in this direction.
Second, learning models provide an insight also into the dynamics of market behavior. This could be useful for better understanding what happens in the "building up" phase of collusion, and which market and institutional features can accelerate or slow down the process.

In what follows, I try to draw a taxonomy of learning models, grouping them into four main categories: experiential learning, adaptive learning, aspiration based models and models based on imitation. It is difficult to classify these categories under the theoretical approaches to bounded rationality mentioned above. I would say that imitation-based and experiential learning models are more close to a "bounded rationality" approach, aspiration based models are influenced by the behavioral approach while adaptive learning processes does not represent strong departure from the classical approach.

2.2.1 Experiential Learning

Models of learning belonging to this class 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 in these models. In contrast, models in which this information is important can be classified ad "observational".

Reinforcement Learning This is probably the most famous model of this class. Originally proposed by Roth and Erev (1995) [86], it rests on the basic hypothesis that players increase the probability of playing pure strategies that have met with success in previous periods. More specifically, the model predicts that at the beginning of the game, before any experience has been acquired, each player has a certain – exogenously given – initial propensity to play each of his pure strategies. Then, after each period t, if pure strategy k is played and yields a payoff of π_k^t, then the propensity to play strategy k is increased by π_k^t while the propensity of playing all other pure strategies remains unchanged. The probability that a player plays pure strategy k at time t is given by the propensity of playing that strategy over the sum of the propensities attached to each the pure strategies in the player's choice set. Notice that the learning curve will be steeper in early periods and flatter later, as the impact of new experiences decreases over time.

Theoretical results on the convergence properties of reinforcement learning in games with a large action space and more than two players are scarce. A simulation based analysis of the long run dynamics produced by this learning model in a repeated Cournot game has been recently presented by Waltman and Kaymak (forthcoming) [103], who show that the mean of firms' joint quantity produced was significantly lower than what is predicted by Nash equilibrium, yet higher than the joint profit maximizing quantity. From their results, though, it is not clear whether the quantities individually produced by each player converge in the long run.

Trial and Error This model of learning has been proposed by Huck et al. [54; 53]. Like reinforcement learning models, the trial-and-error model makes few assumptions about both the availability of information and the cognitive abilities of an agent, as it just requires that the firms know their own past actions and their own profits. The underlying idea of the model, though, is fairly different. Framed in a standard symmetric Cournot oligopoly with *n* firms, this learning rule simply says that a subject would not repeat a mistake, i.e. if profits last period have decreased due to an increase in quantity, then one would not increase quantity again. On the other hand, if profits had increased following an increase in quantity, one would not decrease quantity next period.
Formally, in discrete time this means that each firm *i* sets its quantity q_i^t in period *t* equal to

$$q_i^t = q_i^{t-1} + \delta \,\text{sign}(q_i^{t-1} - q_i^{t-2}) \times \,\text{sign}(\pi_i^{t-1} - \pi_i^{t-2})$$

where π_i^t are the profits of firm *i* at period *t* and δ represents some strictly positive quantity. With some small probability $\epsilon > 0$ each firm chooses an arbitrary direction of change s_i^t. If each firm can choose its outcome from a finite grid $\Gamma = \{0, \delta, 2\delta, ..., v\delta\}$, this defines a Markov process. Huck *et al.* show that this process converges to the joint profit maximizing equilibrium for the two-firms symmetric case, under the assumptions that the cost function is weakly convex and market conditions are such that there exists only one symmetric situation in which joint profits are maximized.
The intuition is rather clear for when all firms start from an identical level of output. The question arises why firms that start from arbitrary initial quantities could become perfectly aligned. Suppose that two firms with different quantities move downwards. They will continue to do so until at least one firm's profit decreases and it will always be the firm with the smaller output to be the first. This is so because the firm selling the higher quantity gains more from the increase in price. Thus, while the smaller firm already moves upward, the other firm continues to move downward thereby decreasing the distance between the firms. Similarly, when moving upward the firm with higher output will be the first to

experience losses and to change direction. Roughly speaking, there is a general tendency to equalize quantities.
By means of simulations they extend this result to situations in which more than two firms – non necessarily symmetric – compete in the market.

In the continuous time the learning process is determined by the following equation:

$$\dot{q}_i(t) = \alpha \, \mathrm{sign}(\dot{q}_i(t)) \, \mathrm{sign}(\dot{\pi}_i(q(t)))$$

and the mathematical approach is more complicated, but it allows the authors to demonstrate that the joint profit maximum is the only stationary and asymptotically stable state of the process.

2.2.2 Adaptive Learning

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 reply, given the strategy profile adopted by their opponents. More specifically, according to the definition given by Milgrom and Roberts (1991) [73], 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. The *best response dynamic* and *fictitious play* are two examples of adaptive learning processes.

Best response dynamics This adjustment process has been analyzed for many models and was in fact the adjustment process originally suggested by Cournot (Cournot, 1960) in his duopoly analysis. Under the best response dynamic 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 demonstrated that this adjustment process was stable and converged to the unique NE for a duopoly with linear demand and constant marginal cost. It is generally well known that best reply dynamics do not converge in oligopolies with a linear setup and three or more firms, as proven by the general instability result found by Theocharis (1960) [101]. Yet, it has been shown by Huck *et al.* [52] that this process converges in finite time to the static Nash equilibrium 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 Under fictitious play (Brown, 1951 [13]) each subject would take the empirical distribution of the actions taken 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 reply to this belief when choosing his current strategy. It has been proven that in general, fictitious play does not necessarily converge to the equilibrium (see Shapley, 1964 [91]), yet Monderer and Shapley (1996) [74] showed that for a particular class of games, that they call "finite weighted potential games", every fictitious play process converges in beliefs to equilibrium. They also showed that a Cournot game with linear inverse demand function and arbitrary differentiable cost function belongs to this class of games.

2.2.3 Aspiration Based Models

Broadly speaking, this class of models in general suggests that firms (or agents) at any time adopt a pure-strategy. If they are achieving their aspiration level, then they are likely to continue with the same strategy. If however, they are below their aspiration level then they are likely to experiment and try something new. Aspiration based learning models in general justify cooperation, that means in our case collusion, as the only equilibrium that can emerge in the long run, irrespective of the initial conditions.

The first and simpler model of this kind is by Karandikar *et al.* [68] and

	C	**D**
C	σ, σ	$0, \theta$
D	$\theta, 0$	δ, δ

Table 2.1: *General normal form game*

has been developed for 2×2 repeated game whose structure is depicted in table 2.1, where it must be that $\sigma > \delta > 0 \quad \wedge \quad 0 \leq \theta \neq \delta$. Depending on the parameters this game can be a prisoner's dilemma game ($\theta > \sigma$), a game of common interest ($\sigma > \theta$) or a coordination game ($\delta > \theta$).
The state of the system at period *t* is

$$s_t = (A_t^1, \alpha_t^1, A_t^2, \alpha_t^2)$$

where A_t^i is the action chosen by player *i* in period *t*, and α_t^i is his *aspiration level*. The behavior rule for both players prescribes that

$$A_{t+1}^i \begin{cases} = A_t^i & \text{if } \pi_t^i \geq \alpha_t^i \\ \\ = A_t^i \text{ with probability } p & \text{if } \pi_t^i < \alpha_t^i \\ \neq A_t^i \text{ with probability } (1-p) & \end{cases}$$

where π_t^i is the payoff of player *i* at time *t*.

If the process is not perturbed the aspiration levels are updated according to the following rule:

$$\alpha_{t+1}^i = \lambda \alpha_t^i + (1-\lambda)\pi_t^i \quad \lambda \in [0,1]$$

This defines a Markov process on the space state $E \equiv \{C, D\}^2 \times \mathbb{R}^2$ which converges almost surely to a "pure strategy space" (pss) in which the action pair is played with payoffs exactly identical to the aspiration levels: so every pss is an absorbing state of the process.

The authors then assume that with some probability η the aspiration levels are subject to a "tremble", and they show that this perturbed process has only one limit distribution which weakly converges to the unique limit invariant distribution of the process μ^*. Finally, it is possible to demonstrate that $lim_{\lambda \to 1} \mu^*(C, C) = 1$, that is: the limit invariant distribution places almost all weight on the cooperative outcome, provided that the persistence parameter λ is sufficiently close to unity.

Aspiration based models have been applied to the Cournot oligopoly problem by Dixon [31]. The structure of the economy envisaged by Dixon's model consists of a large number of identical duopolies. The aspiration level is the same for all the firms, and depends on the average level of profits among all the markets. If firms are achieving their current aspiration level, then they do not experiment, whilst if they are below aspiration they do. The main result of the paper is that the collusive (joint-profit maximizing) outcome is the (almost) global attractor for this economic system: it is as if an economy-wide social pressure of the capital markets forced firms to earn at least average profits in the long-run, thus enforcing collusion.
Oechssler [79] simplifies and generalize Dixon's model, introducing a stochastic stability analysis which assumes that in every period each player experiments new strategies with a small positive probability and studies what happens when this probability tends to zero. It is shown that for games whose joint payoff function has a unique local maximum the unique stochastically stable state is the state in which every player chooses the joint profit maximizing action.

Aspiration based models can be classified either as "observational" or as "experiential", depending on the specified aspiration updating rule. If, as in the original model proposed by Karandikar *et al.*, aspiration levels are updated only on the basis of player's own past payoffs, then the model is experiential, since it does not require the players to observe the actions taken by their opponents, or the payoffs they got in past periods. On the contrary, Dixon's and Oechssler's models assume that players can observe not only the payoff of their own competitors, but also those achieved by other firms in different markets.

2.2.4 Models Based on Imitation

This class includes all the models prescribing that in every period after the first one each individual chooses an action among those which were actually chosen by some player in the previous round. Several models of learning through imitation have been developed, and the main difference between them consists in who is to be imitated, while the analytical tools adopted are substantially similar, and all based on the theory of perturbed Markov processes. All the models mentioned hereafter were specifically designed to be applied in a Cournot oligopoly setting.

Imitate the best In an article appeared in Econometrica in 1997, Vega-Redondo [102] proposed a theoretical model of behavior of Cournot oligopolists which leads to surprising conclusions: the author in fact shows that in the long run a Walrasian behavior evolves *within any quantity-setting oligopoly* producing an homogeneous good, provided that the market demand curve is downward sloping.
The behavioral rule described in the model essentially prescribes to "imitate the best", that is to produce, in each period, the quantity produced in the previous period by the firm that got the highest profits. It can be shown that if all the firms present in the market conform to this rule, the market dynamic which comes forth can be characterized as a discrete time Markov chain.

If the learning process were to consist only of an imitation component, each monomorphic state (i.e. each state in which all the firms produce the same quantity), would be an *absorbing state* of the Markov process. To investigate the relative robustness of these outcomes, Vega-Redondo follows the standard approach and introduces a *mutation dynamic*. This implies that, with some common independent probability $\epsilon > 0$, in every period each firm "mutates" so that all of the possible outcomes can be chosen with a given positive probability. The dynamic market process then becomes ergodic, and one can find the unique invariant distribution μ_ϵ to which the process converges in the long run and study its behavior as $\epsilon \rightarrow 0$.

The main result of Vega-Redondo's model is that the limit invariant distribution $\mu^* \equiv lim_{\epsilon \to 0} \mu_\epsilon$ puts all the probability on the monomorphic state ω^*, where all the firms produce the Walrasian quantity.

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

Imitate the average Huck, Normann, Oechssler, 1999 [52] developed an alternative model of learning, based on the assumption that it seems reasonable that subjects who are uncertain about what to do and observe that the average quantity of the other firms deviates from their own quantity, imitate this average quantity – thinking along the line of 'everyone else can't be wrong'. A preference for cautious behavior and a taste for conformity could be further reasons for imitating the average. If all subjects were to follow this rule, clearly the process is bounded above and below by the highest and lowest initial quantities. Without inertia the process converges to the average of all the starting values. With inertia the process depends on the realizations of the randomization device and is therefore path dependent.

Imitate the "exemplary" firm Offermann, Potters and Sonnemans [81]) proposed another model based on imitation, which assumes that firms follow the firm that sets the good example from the perspective of industry profit. The exemplary firm is the firm (one of the firms) with the quantity that would give the highest sum of profits if it were followed by all firms. Adopting a stochastic stability approach, the authors also introduce a tremble in the model, hypothesizing that with some positive probability, in every period each firm experiments a new strategy from a distribution with full support on its choice set.

If firms adopt the predicted rule and follow the exemplary firm, the unique stochastically stable state of the process is the collusive outcome. The intuition for the result is easy. After one round of following, the non-exemplary firms will follow the firm that was exemplary in the previous period. From this symmetric outcome, they will only move away if one of the firms experiments and chooses a quantity which is closer to the collusive outcome. Eventually, this process will lead to 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. 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 which 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 adaptive learning models.

Apesteguia et al. (2007) [3] 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. They extend the taxonomy of learning models based on imitation, considering both whom the agents are able to imitate and whether they imitate the strategy which on average has yielded the best payoff in the reference group in the previous period (average rule), or the strategy which has yielded the best payoff ever in the reference group, in the previous period (max rule). They show that if one imitates one's own opponents, the Walrasian outcome predominates in the long run, while if on the other hand one imitates other players who face the same problem as oneself but play against different opponents, Nash equilibrium play is obtained, regardless of the imitation rule adopted. Finally, if players can observe both their own competitors and other players in different groups, the process converges to the Cournot equilibrium if the average rule is adopted, while it converges to the perfectly competitive equilibrium if players follow the max rule.

2.3 Information and Learning in Oligopoly Experiments

After Vega-Redondo's article about imitation-based learning [102], several experiments were run with the aim of testing and comparing this and other learning model, specifically in a Cournot oligopoly setting. In what follows, I shall summarize the main characteristics and the principal findings of four among the most representative of these experimental studies. In these works the same experiment is repeated under different treatments, varying the quality and quantity of information provided to the subjects. The authors then compare the actual behavior observed in the different treatments and make inference about the impact that the various informational frameworks have on players' choices. Nonetheless a number of details changes from one experiment to the other, and maybe this is the first reason why the results obtained by the authors are not at all unanimous, nor they are conclusive: for example, the experiments performed by Huck, Normann and Oechssler [52] and by Offerman, Potters and Sonnemans [81] provide a rather strong support to the theory proposed by Vega-Redondo, mentioned before, while the works presented by Rassenti et *al* [84] and by Bosch-Domènech and Vriend evidence no trend towards the Walrasian equilibrium and do not find any clear indication that players tend to imitate the one who got the best performance in the previous period.

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

When information is complete, participants are informed about the symmetric demand and cost functions in plain words and they are provided

with a ‘profit calculator’, which can compute market price and firm’s profit when one enters the total output of other firms and his own output, and can also suggest to the subject the quantity which would yield him the highest payoff given the hypothetical total quantity produced by the competitors.
Information is said to be absent when participants do not know anything about the demand and cost conditions in the market nor do the instructions explicitly state that these would remain constant over time; in these treatments the only thing subjects know is that they would act on a market with four sellers and that their decisions represent quantities. Finally, in treatments with partial information, participants are just told that market conditions remain constant for all periods and coarsely informed about demand and profit functions.

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

Offerman et *al.* [81] conducted a similar computerized experiment, obtaining results which are consistent and complementary to those presented by HNO. In their setting, a triopoly with non-linear demand and cost functions is repeated 100 times, with complete information about market. The authors study how players’ behavior changes across three treatments, which differ for the amount of information provided to the subjects about individual quantities and revenues of the other two

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

The experiment performed by Bosch-Domènech and Vriend (BDV) differs from the previous two both for the setting and for the aims. While HNO and OPS compare the prognostic capability of different learning rules which lead to different theoretical outcomes, here the authors focus specifically on Vega-Redondo's behavioral rule and the main purpose is to investigate whether people are inclined to imitate successful behavior and, in particular, whether this behavior is more prevalent in a more demanding environment. The authors study a series of 22-periods Cournot duopolies and triopolies with homogeneous commodity and linear demand and cost functions. They examine six treatments altogether: for both duopolies and triopolies they consider three different treatments that differ in the way information is provided and in the time pressure put on the players.
In the treatment denominated "easy", the players are given a profit table that conveniently summarizes all the information concerning the inverse demand curve and the cost function, and there is no time pressure on them. After every period, each player gets information about the actions of each of the other players in the same market, but not about their profits.

In the ‘hard’ and ‘hardest’ treatments, players have just one minute to decide on their output level; after each period they receive feedback information both about the actions of all players and about the profits obtained by each of them, and the output decision which led to the highest profit is highlighted.
In the ‘hard’ version, the players get an inconveniently arranged enumeration of the market prices associated with all possible aggregate output levels and of all possible cost levels. The ‘hardest’ version differs from the ‘hard’ treatment in that the information about the demand side of the market is limited to the statement that ‘the price level depends on aggregate output’.
The purpose of the ‘hard’ and ‘hardest’ treatments is to explore to what extent imitation is influenced by the bounds imposed on the subjects’ choice capabilities and to check if it is actually more prevalent when the task of learning about the market becomes more difficult while the decision of the most successful firm is displayed more prominently. The answer the authors give to this question is essentially negative. The data they collected show that as the learning-about-the-environment task becomes more complex, average output increases, but the Walrasian output does not seem to be a good description of the output levels observed in the experiment and if anything, imitation of successful behavior tends to decrease rather than to increase when moving to more complicated environments.

The fourth experiment has been conducted by Rassenti *et al.* (RRSZ); it represents an oligopoly with homogeneous product, in which five firms interact repeatedly for 75 periods, with fixed payoff conditions. The setting exhibits a substantial difference from the previous three since in this case the cost functions – linear, with constant marginal costs and no fixed costs – are private information and are different among the firms. The demand function is linear, and is public information among the players.
The authors perform two different treatments: one in which subjects were able to observe past output choices of each one of their rivals, the other in which they are informed only about the past total output of rivals.

They use their experimental results to test a number of learning models – such as best response dynamics, fictitious play and more general types of adaptive learning. None of these models receives strong support from the data they collected: the observation of actual movement of total output over time appears to be inconsistent with both best response dynamic and with fictitious play, for most experiments. Moreover the authors show that their data do not provide any evidence neither in support for learning models based on imitation, nor for the more traditional hypothesis that information about competitors enhance the potential for collusion, because the treatment conditions involving provision of information about rivals' outputs and prior experience do not seem to have a significant effect on total output levels. The evidence relative to individual behavior is mixed, and no predominant models of learning emerge; the most prominent result is that in general observed behavior for individual subject sellers is not converging to the static Nash equilibrium predictions for individual output choices in these experiments.

In light of these results, I would conclude that it is worthwhile going on investigating on information and learning in oligopolistic markets, because the topic is interesting from a theoretical point of view and it also has interesting practical implications, but a theory consistent with experimental data is still far from being definitely developed. For this reason I decided to design an experiment which is similar to the four previously mentioned under many respects but introduces the use of an experimental technique that allowed me to monitor the information acquisition process through a computer interface. The main idea underlying this software – originally developed by Johnson *et al.* (1988) [63] – 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. He can open just one box at a time, and by recording the number and the duration of the look-ups the program provides precious information about the decision makers' learning process. To my knowledge, this technique has never been applied to the analysis of learning processes in repeated strategic games. Next section summarizes four of the most famous experiments using

Mouselab, with the aim of explaining how this program works in practice and of pointing out its strength points. A more detailed survey on the experimental study of cognition via information search can be found in Crawford (in press) [28].

2.4 Experiments Controlling the Information Acquisition Process

One of the most famous experiments using MouseLab has been performed by Johnson, Camerer, Sen and Rymon(JCSR) [62]: in this work the information acquisition process is observed with the aim of testing the game theoretic assumption of backwards induction. The subjects were asked to play eight three-round alternating-offer bargaining games, with a different anonymous opponent each time. In the first round one of the two players makes an offer to his opponent about how to share a given amount of money; if the other player accepts, than the game is concluded, otherwise he will have to make a counteroffer about how to share a new pie, smaller than the first one. Again, if the first player accepts, the game is over and each of them gets his part as established in the agreement; on the other hand, if the first player rejects the offer the pie shrinks again and he will have the opportunity to make one last offer to his opponent. If even this offer is rejected, nobody gets anything. The sizes of the three pies are represented on the computer screen in front of each player, but they are hidden under three boxes that can be opened only one at a time, simply by putting the mouse' cursor over the box itself. The box will stay open until the mouse is moved somewhere else.
The authors observe three measures of information search: the number of times each box is opened in a period, the total time each box stays opened in a period and the number of transitions from one specific box to another. They note that most of the looking time is spent looking at the first round pie size and contrary to the backward induction prediction there are always more forward predictions than backward ones. From the data collected through these experiments they conclude that people do not use backwards induction instinctively, even if an additional treatment

in which players are previously trained to use backward induction shows that people are able to learn it when appropriately instructed.
They also find that there is a strong correlation between differences in information processing and differences in players' behavior. This and the other results presented in this paper testify that measuring attention directly can effectively contribute to the comprehension of both failure and successes of the game theoretic predictions and help to understand how information and learning can affect the outcomes of different games.

Another seminal study on information acquisition processes has been done by Costa-Gomes, Crawford and Broseta [26](CGCB). They asked the subjects to play 18 two-players normal form games, with different anonymous partners. The payoff tables are hidden and MouseLab is used to present them: for every combination of strategies, subjects could look up their own or their partner's payoff as many times as they wanted, but they could only see one of these numbers at a time. Till the end of the series of games, no feedback was provided to the agents, in order to suppress learning and repeated game effects as much as possible.
In Johnson *et al.* the goal was to test a specific theory of behavior – namely backward induction. On the contrary, here the authors compare nine different decision rules (or types) and try to make inference about which one is more likely to inform players' behavior. As in JCSR, they assume that each decision rule determines both a player's information search and his decision once he gets the information he was looking for. Therefore, by observing both the information acquisition process performed by the agents and the choices they actually make when playing the games, it is possible to deduce what decision rule they adopt.
This study confirms the presence of a systematic relationship between subjects' deviations from search pattern associated with equilibrium analysis and their deviations from equilibrium decisions. Besides, according to Costa-Gomes *et al.*'s analysis most of the subjects are much less sophisticated than game theory assumes: between 67% and 89% of the population belong to two types, namely to the *Naïve* type, who best responds to beliefs that assign equal probabilities to each of their partner's possible strategies and to *L2* type, who best replies to *Naïve*

subjects.

More recently, MouseLab has been used again in two experiments that provide further evidence about how the study of the information acquisition process can be useful to understand what behavioral rules and heuristics are adopted by subjects who display out of equilibrium choices.
One experiment has been conducted by Costa-Gomes and Crawford [25](CGC) and has the same theoretical and econometric framework of CGCB but it differs for the class of games submitted to the subjects. In this case, participants were requested to play 16 different two-person guessing games, with anonymous partners and no feedback till the end of the series. The games have been designed so that the space of possible behaviors is wide and there is a strong separation of the guesses and searches implied by the different decision rules analyzed in the article. Results are consistent with those presented in CGCB [26], but they are significantly sharper: many subjects can be easily attributed to a particular type only by their guesses, and most of the others can be identified via an econometric and specification analysis keeping into account also their information search pattern.

Another interesting application of MouseLab has been recently presented by Gabaix, Laibson, Moloche and Weinberg [40], who experimentally evaluate the *directed cognition model*: a bounded rationality model that assumes that at each decision point, agents act as if their next search operations were their last opportunity for search. As in the other three aforementioned experiments, the authors register the search pattern actually adopted by the subjects in two experiments and they compare it with with what is predicted by the *directed cognition model* and by the optimal search model (i.e. the Gittins-Weitzman algorithm), traditionally adopted in economics.
In the first experiment they ask the participants to choose among three projects whose outcome is uncertain, but can be discovered at a given cost. In the second experiment the subjects are requested to solve a highly complex choice problem in which the classical optimal choice model is analytically and computationally intractable: they have to choose one out

of eight goods which each have nine attributes that could be discovered by opening different boxes on the computer screen. The players cannot collect all the information about the goods, because in this game time is a scarce resource. Individual information acquisition processes are recorded through the MouseLab interface, and the data collected this way reveal that the directed cognition model successfully predicts the empirical regularities observable in subjects' behavior.

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

Chapter 3

Market Setting and Theoretical Benchmarks

In this chapter I describe the market environment characterizing both the experiments I have run, and derive some theoretical results applying three of the learning models described in section 2.2 to this specific setting.

3.1 Experimental Environment

The market setting I have chosen for my experiments is similar to the one proposed by HNO [52]; if possible it is even simpler. In all the sessions and treatments, the setting remains the same. Four identical firms compete à la Cournot in the same market for 40 consecutive periods. Their product is perfectly homogeneous. In every period t each firm i chooses its own output q_i^t from the discrete set $\Gamma = \{0, 1, ..., 30\}$, which is the same for every firm. The choice is simultaneous.
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 $\omega^W = (20, 20, 20, 20)$.

3.2 Three Theoretical Benchmarks

We are interested in studying the market dynamics when the stage game so defined is repeated several times, and the firms do not have all the information (or the computational capabilities) to evaluate what the standard theory predicts is an optimal behavior for them. As a

	Required information	**Predicted equilibrium**
Best Reply Dinamics	Competitors' aggregate quantity and BR function	Nash ($q = 16$)
Imitate the Best	Last period individual profits and quantities	Walrasian ($q = 20$)
Trial and Error	Own past profits and quantities	Collusive ($q = 10$)

Table 3.1: *Theoretical benchmarks*

first benchmark to evaluate the experimental results, we individuate three among the theoretical learning models introduced in section 2.2. The choice has been driven by three main motives: first, these three models are particularly simple; second, they are based on very different informational requirements; third, they yield well distinct market outcomes in the long run, namely the Cournot, Walrasian and joint profit maximizing outcomes, respectively. These models are summed up in table 3.1 and presented more in detail in the following paragraphs, where they are applied to the specific market environment characterizing my experiments.

3.2.1 Best Response Dynamics

Following Huck, Normann and Oechssler [52] we consider here the simplest model of best reply dynamic. This model theorize that in every period each player myopically chooses his output as a best reply to the sum of the quantities produced by the other three in the previous period. More precisely, the best reply correspondence for player *i* maps $\sum_{j \neq i} q_j^{t-1}$ to the set

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

Under the hypotheses I made on market structure, due to the discreteness and finiteness of the choice set, we have:

$$BR_i^t = \begin{cases} \{0\}, \text{ if } \sum_{j \neq i} q_j^{t-1} \geq 80 \\ \{30\}, \text{ if } \sum_{j \neq i} q_j^{t-1} < 20 \\ \left\{ \frac{80 - \sum_{j \neq i} q_j^{t-1}}{2} \right\} \text{ if } 20 \leq \sum_{j \neq i} q_j^{t-1} < 80 \text{ and } \sum_{j \neq i} q_j^{t-1} \text{ is even} \\ \left\{ \frac{80 - \sum_{j \neq i} q_j^{t-1}}{2} - 0.5, \frac{80 - \sum_{j \neq i} q_j^{t-1}}{2} + 0.5 \right\} \text{ otherwise.} \end{cases}$$

In this last case, it is assumed that the player chooses q_i^t from BR_i^t according to some probability distribution with full support.

The best reply dynamic defined this way yields a Markov chain over the state space $\Omega = \Gamma^4$ which does not necessarily converge to a stable equilibrium, consistently with what has been shown by Theocharis [101] for the case in which quantities are chosen in a continuous space.
To catch an intuition of this result, suppose for example that the system

reaches one of the two states of the absorbing set $s = \{(30,30,30,30),\ (0,0,0,0)\}$: once this has happened, the system will keep on oscillating between this two states and will never be able to escape the set.
HNO state the following theorem:

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

Within the framework considered here, this implies that the learning process brings the system to converge to the state $\omega^N = (16,16,16,16)$.

For sake of completeness I replicate here HNO's demonstration, applying it to the specific context under exam.
To introduce inertia into the learning model, HNO simply hypothesize that in every period each player i chooses q_i^t from the set BR_i^t with some fixed probability $(1-\theta)$, while with probability θ he sticks to the quantity he chose in the previous period, so $q_i^t = q_i^{t-1}$. We will see that the best reply dynamic with inertia can be represented by an ergodic Markov chain having only one recurrent set, therefore the probability distribution over the state space approximate the unique invariant distribution of the process, regardless of the initial state, and that this invariant distribution puts probability one over the state ω^N, which is the only recurrent state in Ω.

PROOF OF THEOREM 1: It is clear that ω^N is an absorbing state, since the process we have defined can never escape from it; in order to prove the result, it is necessary to demonstrate that no other state in Ω is recurrent, namely ω^N is accessible from any other state:

$$\omega' \to \omega^N \;\; \forall\, \omega' \in \Omega,\ \omega' \neq \omega^N$$

which means that there exist a $\tau \in \mathbb{N} \setminus \{0\}$ such that the probability $p^{(\tau)}_{\omega',\omega^N}$ of reaching state ω^N from ω' in τ periods is positive.

To prove this result we shall first show that the state ω^N is accessible from any $\omega^+ \in \Omega^+$, where $\Omega^+ = \{(q_1,q_2,q_3,q_4) \in \Omega : q_i > 0,\ i = 1,2,3,4\}$, then we shall conclude by verifying that for any $\omega^0 \in \Omega \setminus \Omega^+$ there exists an $\omega^+ \in \Omega^+$ such that ω^+ is accessible from ω^0 $(\omega^0 \to \omega^+)$, therefore, by the Chapman-Kolmogorov equation, it follows that $\omega^0 \to \omega^N$. $\forall \omega^0 \in \Omega$.

The first part of the proof requires the preliminary definition of the concepts of "ordinal potential", "ordinal potential game" and "improvement path", introduced by Monderer and Shapley [74].
Let $N = \{1, 2, ..., n\}$ be the set of players, Y_i denote the set of strategies of player i and $u_i : Y \rightarrow \mathbb{R}$ the payoff function of player i, where $Y = Y_1 \times Y_2 \times ... \times Y_n$ is the set of the strategy profiles.
A function $P : Y \rightarrow \mathbb{R}$ is an *ordinal potential* for the game $G = (N, Y, u)$ if, for every $i \in N$ and for every $y_{-i} \in Y_{-i}$

$$u_i(x, y_{-i}) - u_i(z, y_{-i}) > 0 \Leftrightarrow P(x, y_{-i}) - P(z, y_{-i}) > 0 \quad \forall x, z \in Y_i$$

An *ordinal potential game* is a game that admits an ordinal potential.
An *improvement path* in Y is a sequence $\gamma = (y^0, y^1, ...)$ of elements of Y such that, for every $k \geq 1$, there exists a unique player – say player i – such that the following conditions are simultaneously satisfied:

- $y^k = (x, y_{-i}^{k-1});\ x \in Y_i,\ x \neq y_i^{k-1}$
- $u_i(y^k) > u_i(y^{k-1})$

The proof of Theorem 1 relies on the following Lemma by Monderer and Shapley [74]:

Lemma 1.1 Every improvement path of a finite potential game is finite.

PROOF OF LEMMA 1.1: For every improvement path $\gamma = (y^0, y^1, ...)$, by the definition of ordinal potential we have:

$$P(y^0) < P(y^1) < P(y^2) < \dots$$

As Y is a finite set, this sequence must be finite.

This result can be applied to our model, since – as shown by HNO – the function $P(\omega) = (p(\omega) - 1) \prod_{j=1}^{4} q_j$ is an ordinal potential for our game if the strategy set of each player is restricted to $\Gamma \setminus \{0\}$. Therefore, there is a finite improvement path departing from every state $\omega^+ \in \Omega^+$ where $\Omega^+ = \{(q_1, q_2, q_3, q_4) \in \Omega : q_j > 0,\ j = 1, 2, 3, 4\}$.
By definition this improvement path ends in a state ω^k such that no player

can improve his own payoff by changing his strategy if the quantities chosen by the other players remain the same, i.e.

$$\nexists i \text{ s.t. } \pi_i(\omega^{k+1}) > \pi_i(\omega^k) \text{ where } \omega^{k+1} = (q', q^k_{-i}) \text{ for some } q' \neq q^k_i,\ q' \in \Gamma \backslash \{0\}$$

This condition is clearly satisfied only by $\omega^C = (16, 16, 16, 16)$, representing the unique Nash equilibrium of the stage game.

Finally, note that the Best Reply process with inertia can give rise to an improvement path over Ω^+, since with positive probability in every period only one player changes his strategy while the others stick to the quantity they previously chose. So, under the Best Reply process, ω^N is accessible from every state in Ω^+.

As mentioned before, to complete the proof of Theorem 1 it is enough to show that

$$\forall \omega^0 \in \Omega \setminus \Omega^+ \ \exists \omega^+ \in \Omega^+ \text{ s.t. } \omega^0 \rightarrow \omega^+$$

Let $\omega^0 = (q_1^0, q_2^0, q_3^0, q_4^0) \in \Omega \setminus \Omega^+$ and ω^{BR} be a state in which every player i chooses $q_i \in BR_i^0$, giving a best reply to ω^0. By definition, ω^{BR} is accessible from ω^0 under the process defined by the Best Response Dynamics ($\omega^0 \rightarrow \omega^{BR}$).

If $\sum_i q_i^0 < 79$ it is straightforward to see that $\omega^{BR} \in \Omega^+$ since the best reply to a quantity strictly smaller than 79 is always positive.

If $\sum_i q_i^0 \geq 79$ it can be shown that the sum $\sum_i BR_i^0 \leq 162 - \frac{3}{2}\sum_i q_i^0 \leq 43.5$, since for every player i $BR_i^0 \leq \frac{80 - \sum_{j \neq i} q_j^0}{2} + \frac{1}{2}$. Therefore $\exists \omega^+ \in \Omega^+$ s.t. $\omega^{BR} \rightarrow \omega^+$, thus the system moves from ω^0 to ω^+ with positive probability in at most 2 steps.

This concludes the proof.

The problem with the analysis I have just concluded is that it makes prediction for the long run outcomes, and for a positive, but undefined degree of inertia. Since it might be interesting also to check how the process behaves in the short run, for various degrees of inertia, I complete this and the following sections by presenting the results of simulations I have done for this and the other two models, under the setting presented in section 3.1. My simulation then reproduces a market with four firms, facing the demand and cost functions characterizing my experiments, and interacting for 40 consecutive periods. The quantities produced in the first period are randomly drawn from a uniform distribution over the set $\{0, 1, \ldots, 29, 30\}$, while firms' behavior in the following periods – in this case – is determined by the best response dynamics with a degree of inertia equal to θ, taking values 0.05, 0.1, 0.15. I ran 10000 cycles of simulation per each value of θ. The frequency distributions of individually chosen quantities in period 40 are reported in figure 3.1. According to my simulations, the higher is the degree of inertia, the

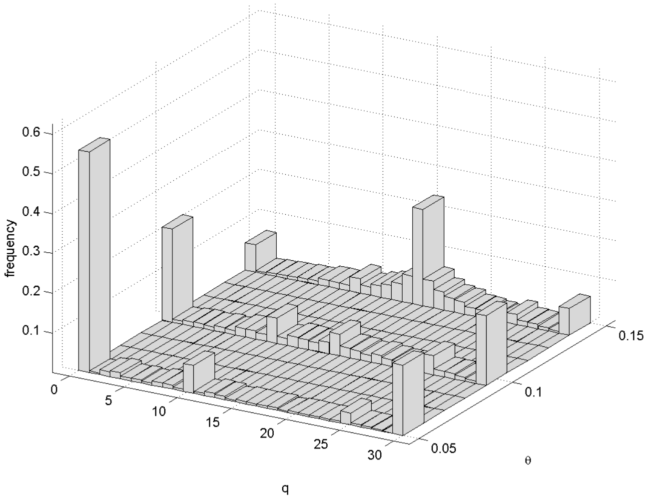


Figure 3.1: *Frequency distribution of individual choices in the 40th round; 10000 cycles of simulation.*

faster the convergence to the static Cournot-Nash equilibrium. More specifically, we notice that a degree of inertia below 10% is not sufficient to obatin some convergence within 40 periods.

3.2.2 Imitate the Best

The learning model presented here has been originally proposed by Vega-Redondo [102]. The core of the model is represented by the *imitation* *dynamic*: a discrete time dynamic which assumes that at every time *t* each firm chooses its output q_i^t from the set:

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

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

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

$$\lim_{\tau \to \infty} P^{(\tau)} = \mathbf{u}\mu$$

where $\mathbf{u}$ is the unit vector: namely, the Markov chain converges to its stationary distribution, regardless where it began.

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

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

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

This implies that under the "imitate the best" dynamic, the only stochastically stable outcome in our setting is the Walrasian outcome $\omega^W = (20, 20, 20, 20)$, in which all the firms get zero profits.

PROOF OF THEOREM 2:
The proof of this theorem relies on the graph-theoretic techniques developed by Freidlin and Wentzell [38], therefore some basic concepts should be introduced in order to expose it.
A *directed graph G* is an ordered pair $G := (V, A)$ with

- V, a set of vertices or nodes,
- A, a set of ordered pairs of vertices, called directed edges, arcs, or arrows.

An edge $e = (x, y), x, y \in V$ is considered to be directed from *x* to *y*, so *y* is said to be a direct successor of *x*, and *x* is said to be a direct predecessor of *y*. More generally, if there exists a path leading from *x* to *y*, then *y* is said to be a successor of *x*, and *x* is said to be a predecessor of *y*.
To apply this idea to the situation under analysis, first let *O* be the directed graph having Ω as the vertex set and in which for every vertex $\omega \in \Omega$ there exists an edge $e(\omega, \omega'),\ \forall \omega' \in \Omega,\ \omega' \neq \omega$.
A *resistance* $r(\omega', \omega'')$ can be associated to every edge $e(\omega', \omega'')$, where

$$r \text{ s.t. } 0 < \lim_{\epsilon\to 0} \epsilon^{-r} P^{\epsilon}_{\omega',\omega''} < \infty$$

and $P^{\epsilon}_{\omega',\omega''}$ denotes the (one step) transition probability from ω' to ω'' according to the perturbed process defined by the Imitation Dynamics

with a probability of error equal to ϵ. The resistance simply measures the total number of mistakes (or experiments) involved in the transition from state ω' to ω''.
An *ω-tree* H for any vertex ω of O is a tree spanning O so that for every $\omega' \neq \omega$, $\omega' \in \Omega$ there exists a unique directed path from ω' to ω. Let $r(H) = \sum_{(\omega',\omega'')\in H} r(\omega', \omega'')$ denote the resistance of the ω-tree H, and $\mathcal{H}_\omega$ be the set of all the ω-trees in O.
The *stochastic potential* of a state $\omega \in \Omega$ is:

$$\gamma(\omega) = \min_{H\in\mathcal{H}_\omega} r(H)$$

Let $\gamma^* = \min_{\omega\in\Omega} \gamma(\omega)$.

Now we can provide the proof for theorem 2, which follows directly from the following three lemmata:

Lemma 2.1 Let P^ϵ denote the Markov chain defined by the perturbed process, and μ^ϵ be its unique stationary distribution.
Then $\lim_{\epsilon\to 0} \mu^\epsilon = \mu^0$ *exists* and μ^0 is a stationary distribution of P^0 – the Markov chain defined by the unperturbed process. Moreover, the probability μ^0_ω associated to the state ω by the limit stationary distribution μ^0 is strictly positive if and only if $\gamma(\omega) = \gamma^*$.

Lemma 2.2 The stochastic potential $\gamma(\omega^W)$ equals the cardinality of Γ minus 1.

Lemma 2.3 For all $q \neq q^W$, the monomorphic state ω_q has a stochastic potential

$$\gamma(\omega_q) \geq |\Gamma|$$

PROOF OF LEMMA 2.1:. We shall follow the demonstration provided by Peyton Young [82] (Appendix), that we report here for sake of completeness.
First we can apply to P^ϵ a result established by Freidlin and Wentzel [38](Chapter 6, Lemma3.1) for every aperiodic, irreducible stationary Markov processes the unique stationary distribution is given by the formula:

$$\mu^\epsilon_\omega = p^\epsilon_\omega / \sum_{\omega'\in\Omega} {p^\epsilon}_{\omega'}$$

where

$$p_{\omega}^{\epsilon} = \sum_{H \in \mathcal{H}_{\omega}} \prod_{(\omega', \omega'') \in H} P_{\omega', \omega''}^{\epsilon}$$

Choose the ω-tree H with minimum resistance and consider the identity:

$$\epsilon^{-\gamma^*} \prod_{(\omega', \omega'') \in H} P_{\omega', \omega''}^{\epsilon} = \epsilon^{r(H) - \gamma^*} \prod_{(\omega', \omega'') \in H} \epsilon^{-r(\omega', \omega'')} P_{\omega', \omega''}^{\epsilon} \tag{3.1}$$

By the definition of *r*,

$$\lim_{\epsilon \to 0} \epsilon^{-r(\omega', \omega'')} P_{\omega', \omega''}^{\epsilon} > 0, \forall (\omega', \omega'') \in H \tag{3.2}$$

If $r(H) = \gamma(\omega) > \gamma^*$ it follows from (3.1) and (3.2) that

$$\lim_{\epsilon \to 0} \epsilon^{-\gamma^*} \prod_{(\omega', \omega'') \in H} P_{\omega', \omega''}^{\epsilon} = 0$$

therefore

$$\lim_{\epsilon \to 0} \epsilon^{-\gamma^*} p_{\omega}^{\epsilon} = 0.$$

Similarly, if $r(H) = \gamma(\omega) = \gamma^*$ we obtain

$$\lim_{\epsilon \to 0} \epsilon^{-\gamma^*} p_{\omega}^{\epsilon} > 0.$$

Since

$$\mu_{\omega}^{\epsilon} = \epsilon^{-\gamma^*} p_{\omega}^{\epsilon} / \sum_{\omega' \in \Omega} \epsilon^{-\gamma^*} p^{\epsilon}{}_{\omega'}$$

it follows that

$$\lim_{\epsilon \to 0} \mu_{\omega}^{\epsilon} = \begin{cases} = 0 \text{ if } \gamma(\omega) > \gamma^* \\ > 0 \text{ if } \gamma(\omega) = \gamma^* \end{cases}$$

Finally, since $\lim_{\epsilon \to 0} P_{\omega', \omega''}^{\epsilon} = P_{\omega', \omega''}^{0} \; \forall \omega', \omega'' \in \Omega$ and μ^{ϵ} satisfies the equation $\mu^{\epsilon} P^{\epsilon} = \mu^{\epsilon} \; \forall \epsilon > 0$, then $\mu^0 P^0 = \mu^0$. μ^0 is therefore a stationary distribution of P^0, hence it puts probability 1 over one of the monomorphic states. This concludes the proof for Lemma 1.
We shall now show that the monomorphic state having positive probability under the stationary distribution is precisely ω^W.

PROOF OF LEMMA 2.2: here we shall apply the proof provided by Vega Redondo [102] to the specific context we are analyzing, to show that:

$$\gamma(\omega^W) = |\Gamma| - 1 = 30.$$

Note that

$$\forall \omega \in \Omega,\ \exists \omega_q \text{ s.t. } P^0_{\omega,\omega_q} > 0 \text{ and therefore } r(\omega, \omega_q) = 0.$$

Consider any monomorphic state $\omega_q,\ q \in \Gamma$, and a state $\tilde{\omega}_q = (q_1, q_2, q_3, q_4)$ such that $\exists i \in I :\ q_i = q^W$ and $q_j = q \forall j \neq i,\ j \in I$. By the definition of r it is easy to check that $r(\omega_q, \tilde{\omega}_q) = 1$ since $P^\epsilon_{\omega_q,\tilde{\omega}_q} > 0$ if $\epsilon > 0$, according to the previously stated definition of the perturbed process.
To conclude the proof, it is enough to verify that, for any $q \neq q^W$, the resistance $r(\tilde{\omega}_q, \omega^W)$ is equal to zero. Indeed, it is straightforward to check that $P^0_{\tilde{\omega}_q,\omega^W} > 0$ under our hypotheses, because the firm producing q^W gets always the highest profit, regardless of the output produced by the others:

$$\pi(q^W, 3q) > \pi(q, 2q + q^W)\ \forall q \in \Gamma,\ q \neq q^W.$$

PROOF OF LEMMA 2.3: to prove that $\gamma(\omega) > |\Gamma|\ \ \forall \omega_q \neq \omega^W$ it suffices to show that at least two mistakes are necessary to escape from the basin of attraction of ω^W, namely: there is no

$$\tilde{\omega}^W = (q_1, q_2, q_3, q_4) : \exists i : q_i \neq q^W,\ q_j = q^W\ \forall j \neq i,\ j \in I$$

such that $r(\tilde{\omega}^W, \omega_q) = 0$, since $P^0_{\tilde{\omega}^W,\omega_q} = 0$ because the profit of the firm producing $q \neq q^W$ is always lower than the profit earned by each of the other three firms, producing q^W:

$$\pi(q, 3q^W) < \pi(q^W, 2q^W + q)\,\forall q \in \Gamma,\ q \neq q^W.$$

As for the best response dynamics, I ran a simulation to depict the behavior of this model in the short run, with different levels of experimentation. Again, I ran 10000 cycles of simulation per each value of ϵ, taking values 0.01, 0.05, 0.1. The frequency distributions of individually

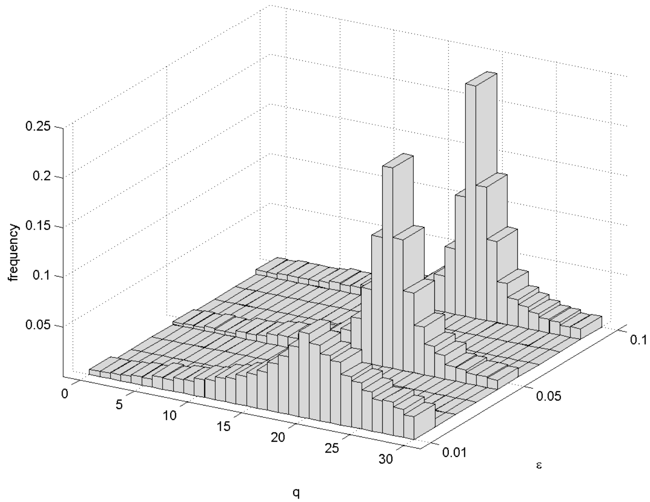


Figure 3.2: *Frequency distribution of individual choices in the 40th round; 10000 cycles of simulation.*

chosen quantities in period 40 are reported in figure 3.2.
Simulation results, show that with any of the considered levels of experimentation, the process converges pretty quickly to the predicted outcome, but convergence is faster as experimentation becomes more probable.

3.2.3 Trial and Error

This model of learning has been firstly proposed by Huck, Normann and Oechssler in 2000 [53] then revised by the same authors in a subsequent article [54] where they present a continuous time version of it.
Both versions of the learning model are, in principle, very simple. Assuming that the strategy set of the player is ordered, the model predicts that every time a player changes the strategy he adopts, he will check whether his payoff has consequently increased or decreased. If he observes a raise, in the following period he will keep on changing his strategy in the same direction as before. On the contrary, if the payoff declines the player will

change his strategy in the reverse direction. This is the model with the most lax hypotheses about information: it just requires that the firms know their own past actions and their own profits.
Huck *et al.* show that Trial and Error learning yields a collusive outcome. They proove this result analytically for the continuous time version, and for the discrete version with only two firms, and they extend it to the discrete case with more than two firms by means of simulations.

Since in the case we are analyzing both time and the strategy set are discrete, we will consider the discrete version of the model, applying it to the oligopoly setting described above.
Given the quantity $q_1^i \in \Gamma$ chosen by any firm in period one, in every following period $t > 1$ each firm will set how much to produce according to the following rule:

$$q_t^i = \begin{cases} 0 & \text{if } q_{t-1}^i + s_{t-1}^i < 0 \\ 30 & \text{if } q_{t-1}^i + s_{t-1}^i > 30 \\ q_{t-1}^i + s_{t-1}^i & \text{otherwise.} \end{cases}$$

where the direction of change is given by

$$s_t^i = \text{sign}(q_t^i - q_{t-1}^i)\text{sign}(\pi_t^i - \pi_{t-1}^i)$$

if $(q_t^i - q_{t-1}^i)(\pi_t^i - \pi_{t-1}^i) \neq 0$; otherwise s is randomly chosen among the values $-1, 0, 1$, each having positive probability.
This defines a Markov chain over the state space $\Omega = \Gamma^4 \times \{-1, 0, 1\}^4$. As for the previous model, we assume the possibility of experimentation or mistakes, thus defining a perturbed process, in which with some small probability $\epsilon > 0$ each firm chooses an arbitrary direction of change s_i^t. This defines a Markov process which is irreducible and aperiodic, therefore has a unique stationary stable distribution. By contrast, in principle the unperturbed process may have many stationary distributions.

In what follows, I will show that *(i)* the unperturbed process has several absorbing sets, *(ii)* all the states belonging to these absorbing sets have the same stochastic potential, therefore *(iii)* they are all stochastically stable, meaning that all the absorbing sets of the unperturbed process belong to the support of the limit distribution $\mu^0 = \lim_{\epsilon \to 0} \mu^\epsilon$.

first recurrent set	$11\downarrow$	$11\downarrow$	$11\downarrow$	$11\downarrow$	average quantity: 10
	$10\downarrow$	$10\downarrow$	$10\downarrow$	$10\downarrow$	
	$9\uparrow$	$9\uparrow$	$9\uparrow$	$9\uparrow$	
	$10\uparrow$	$10\uparrow$	$10\uparrow$	$10\uparrow$	
second recurrent set	$11\downarrow$	$11\downarrow$	$11\downarrow$	$10\uparrow$	average quantity: 10.25
	$10\uparrow$	$10\uparrow$	$10\uparrow$	$11\uparrow$	
	$11\downarrow$	$11\downarrow$	$11\downarrow$	$12\downarrow$	
	$10\downarrow$	$10\downarrow$	$10\downarrow$	$11\downarrow$	
	$9\uparrow$	$9\uparrow$	$9\uparrow$	$10\downarrow$	
	$10\uparrow$	$10\uparrow$	$10\uparrow$	$9\uparrow$	
third recurrent set	$11\downarrow$	$11\downarrow$	$10\uparrow$	$10\uparrow$	average quantity: 10.77
	$10\uparrow$	$10\uparrow$	$11\uparrow$	$11\uparrow$	
	$11\downarrow$	$11\downarrow$	$12\downarrow$	$12\downarrow$	
	$10\downarrow$	$10\downarrow$	$11\downarrow$	$11\downarrow$	
	$9\uparrow$	$9\uparrow$	$10\downarrow$	$10\downarrow$	
	$10\uparrow$	$10\uparrow$	$9\uparrow$	$9\uparrow$	
fourth recurrent set	$11\downarrow$	$10\uparrow$	$10\uparrow$	$10\uparrow$	average quantity: 10.47
	$10\uparrow$	$11\uparrow$	$11\uparrow$	$11\uparrow$	
	$11\downarrow$	$12\downarrow$	$12\downarrow$	$12\downarrow$	
	$10\downarrow$	$11\downarrow$	$11\downarrow$	$11\downarrow$	
	$9\uparrow$	$10\downarrow$	$10\downarrow$	$10\downarrow$	
	$10\uparrow$	$9\uparrow$	$9\uparrow$	$9\uparrow$	

Note: the numbers indicate the quantity produced by each of the four firms, the arrows the direction of change.

Table 3.2: *Recurrent sets of the unperturbed process*

Recall that under the maintained assumptions, the cardinality of the state space is $31^4 \times 3^4$. If we disregard the order of the players, then the number of possible states reduces to $\binom{96}{4}$. Among these states, we want to individuate those, if any, which are recurrent. This was done by means of simulations and numerical analysis. First, the evolution of the unperturbed process was simulated over 200 iterations of the stage game, replicating this cycle for 10000 times. Regardless of the initial state of each cycle, which was randomly chosen from a uniform distribution over the whole state space, it emerged that during the last ten iterations the process rests over only 22 of the possible states, and that these 22 states belong to four recurrent sets, as displayed in table 3.2. Through a following numerical analysis I checked that all the other states are transient under the unperturbed process, meaning that the t-steps transition probability from each of these states to at least one of the states belonging to the recurrent sets is positive, for a finite t [1].

It is easy to verify that one and only one mistake is sufficient to transit from a recurrent set to the following one. It follows that all the states in the four recurrent sets have the same stochastic potential, which equals the number of recurrent sets minus one. As a consequence following directly from the aforementioned Lemma 2.1, all the states in the recurrent sets are stochastically stable.

A quick look at the quantities produced by the four firms in all of the recurrent states confirms the result previously stated by Huck *et al.*, namely that the Trial and Error process converges to a neighborhood of the joint profit maximizing outcome, which in this case is obtained when all the firms produce a quantity equal to 10.

I conclude this chapter by presenting the results of the simulation I did to study the short run behavior of the trial and error model, with different probabilities of mistakes. As in the previous two models, I ran 10000 cycles of simulation per each level of the error probability ϵ, taking values 0.01, 0.05, 0.1. Alike the imitate the best process, also trial and error turns out to converge to states which are close to the stochastically stable one within the number of repetitions that will take place in my experiments, regardless of the probability of error. In particular, we observe that as the probability of error decreases, the variance of the distribution decreases too and convergence is more precise.

[1] The analysis was performed with Matlab ©. The code is available from the author upon request.

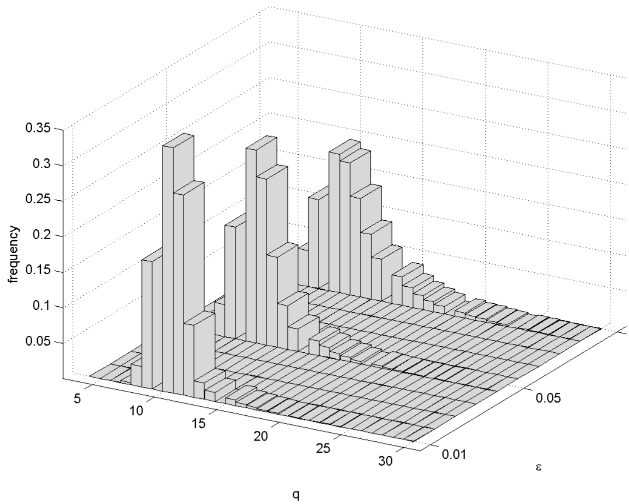


Figure 3.3: *Frequency distribution of individual choices in the 40th round; 10000 cycles of simulation.*

Chapter 4

First Experiment: playing against robots

4.1 Experimental Design

The way people behave and the way they learn might be affected by what their peers (namely their opponents, in our experiment) do. For this reason, the analysis of the results could be very complicated if the experiment included interaction between subjects, because the individual characteristics of the players could affect the dynamics that emerge in every single group in different ways[1]. Therefore, I decided to run this first version of the experiment in which I let subjects play against three "virtual" players enacted by the computer and programmed to follow a specific learning rule. This way I can control for the effect of the opponents' behavior on the players' choices. The treatment variable is then represented by the learning rule adopted by the "robot" opponents. Note that the three opponents of the same subject are all programmed to follow the same learning rule.

To avoid deception, subjects are informed that their opponents are "robots", that is: they are enacted by the computer. Subjects also know that these "robots" do not play at random but choose according to some

[1] I thank Magnus Johannesson for having pointed out this possible complication.

rule, nonetheless they do not necessarily choose the same output. No other information is provided about the way these "robots" play.

The experiment will be repeated under three treatments. I shall present first the elements which are common across the three treatments, then explain the differences between them.

4.1.1 Information Provided to the Subjects

Participants know how many competitors they have. Instructions explain in plain words that there is an inverse relation between the overall quantity produced by the four firms and market price and that a firm's production costs increase with the number of goods it decides to sell. Besides, players are told that per-period profit is given by market price times the number of goods sold by the firm, minus production costs.
(see the instructions in Appendix A.2).

Subjects are also endowed with a *profit calculator* similar to the one proposed by Huck *et al.* [52]. This device has two input fields that the subject can fill in: one for the total quantity produced by the other three firms in the market, one for the quantity produced by his own firm. If the player enters two (arbitrary) values, one for each of these fields, the profit calculator evaluates market price and the profit the subject would earn; if the player just fill in the field pertaining to competitors' quantity and leaves the other one blank, the profit calculator computes the quantity that would yield him the highest profit and inform him about market price and profits he would earn if he produced the suggested amount of good. The answers provided by the profit calculator are always displayed both graphically and textually (Figure A.1(a) and A.1(b)). The software I developed for this experiment records how many times the subject uses the profit calculator and every trial he does.

The number of rounds is common knowledge among the subjects. According to game-theoretic predictions, cooperation should be sustainable only if our stage game were repeated in(de)finitely many times, but according to Selten *et al.* [90]

> Infinite supergames cannot be played in the laboratory. At-

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

In light of this consideration and of the results obtained by Normann and Wallace [78] – who show that the termination rule does not have a significant effect on players' behavior except for an end effect – I decided to adopt a commonly known finite horizon, for sake of transparency and practicality.

After the first round, each player has the opportunity to look at three plots summing up information about what happened in the previous periods (Figure A.2). The first graph is a bar-plot showing the quantity produced by each of the four firms in the market in the previous period, and the relative profit. The second graph displays the quantity produced by the player's firm compared with the aggregate quantity produced by his three competitors in each of the previous periods, since the game began. The last plot shows the quantity and the profit obtained by the player's firm in each of the previous periods.

The subjects, however, are not able to look at all the three plots at the same time, since these plots are hidden behind three boxes on the computer screen and the player can open just one box at a time. Behind a fourth box is hidden the answer provided by the profit calculator. A box can be opened just putting the mouse cursor over it, and its content will be displayed on the screen until the cursor moves out of the box's borders. As mentioned before, the software automatically records subjects' look-ups sequences and look-ups durations.

Besides these four boxes, on the computer screen there is a counter showing the running cumulative profits earned by the player since the game began, and a timer displaying how long it is since the current round started, that is how long has the subject been thinking about what choice to make next. Figure A.4 shows how subjects' computer screen looks like.

After the last round the participants are shown their overall profit, compared with those of their three opponents (Figure A.3).

4.1.2 Treatments

This experiment has been run under three different treatments (T&E, BRD and ItheB), which differ only by the learning rule adopted by the computer. For sake of simplicity, I have chosen the three basic learning rules described in section 3.2. In what follows I briefly recapitulate the rule adopted by the robot players in each of the treatments.

Trial and Error (T&E) I programmed the computer so that in this treatment each "robot" player i sets its quantity q_i^t in round t equal to

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

where the direction of change s_i^t is given by

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

if $(q_i^t - q_i^{t-1})(\pi_i^t - \pi_i^{t-1}) \neq 0$, where π_i^t are the profits of firm i at round t. If instead $(q_i^t - q_i^{t-1})(\pi_i^t - \pi_i^{t-1}) = 0$, the direction of change is randomly chosen among the values $-1, 0, 1$, each having equal probability.
I also introduced a positive probability of error in the model, so in every period, with probability $\epsilon = 0.05$ each "robot" chooses an arbitrary direction of change s_i^t randomly drawn from a uniform distribution over $\{-1, 0, 1\}$.

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

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

Notice that, in the last case – when according to the original model the choice would have been ambiguous – I assumed that the players choose the smallest integer greater than $\frac{80-q_{-i}^{t-1}}{2}$.

I introduced some degree of inertia into the learning rule guiding the behavior of the "robot" players in treatment BRD: with independent probabilities equal to 0.05 in every round each of them chooses $q_i^t = q_i^{t-1}$, otherwise follows the myopic best response dynamic. The level of inertia is pretty low, and according to my simulations it should not be sufficient to guarantee convergence to the Nash equilibrium within 40 periods. Since in this experiment I am more interested in studying how the opponents' behavior affects human subjects' information acquisition and learning processes, than to observe whether the market outcomes indeed converge to some theoretically predicted level, I decided to keep the level of "noise" low in all the three treatments, and set the probability of trembles at 5% in each of them.

Imitate the Best (ItheB) In the last treatment (ItheB) the "robot" players behave according to the learning model firstly proposed by Vega-Redondo [102]. Robots are then programmed to choose their output q_i^t from the set:

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

where, in our case, $\Gamma = \{0, 1, \ldots, 29, 30\}$.
Notice that this set is always a singleton in the setting of my experiments.

To introduce a small degree of experimentation in the robot's behavior, I programmed them to choose their quantity according to the imitate the best rule, with a probability equal to 95%, while with a 5% probability they choose a quantity randomly drawn from the set $\{0, 1, \ldots, 29, 30\}$, according to a normal distribution centered on the quantity they have chosen in the previous round, whith a standard deviation equal to 10.

In all the three treatments, the quantity set by the robot players in the first period is randomly drawn from a uniform distribution over the set $\{0, 1, \ldots, 30\}$.

I adopted a within subjects design, so every subject played against a unique type of "robot" players. The program was designed using MouselabWEB (see `www.mouselabweb.org`), a process tracing tool developed by Martijn Willemsen and Eric Johnson and derived from the aforementioned software MouseLab.

Two sessions of this experiment were run at the Stockholm School of Economics, on April 2, 2007. Twelve undergraduate students took part in the first session and eleven took part in the second one. Sessions lasted about one hour and a half each, and the average payment (including the show up fee) was equal to 179 SEK [2]
In total, 7 subjects played under the ItheB treatment, and 8 under each of the other two treatments.

At the beginning of each session the participants were disposed in the lab so that they could not communicate with each other. Instructions were written on a page that appeared on the computer screen of each subject, but common knowledge of the information they contain is ensured by telling the participants that the pages they are reading are perfectly identical. Instructions were divided in several parts and at the end of each of them an understanding test was submitted to the reader, who had to answer correctly to proceed to the next page.
When a player finished reading the instructions he could start playing. After the last round of the game, subjects were asked to answer a short questionnaire including some questions about their individual characteristics (age, gender, education) and about the strategies they adopted in the game; then participants were called one by one in private and paid according to their total profits.

4.2 Results

In what follows I shall present the results from the two sessions of this experiment. I will first examine the quantities chosen by the players, then the way they used information, and finally I'll present an attempt to establish a relation between information search patterns and actual behavior by means of two learning models.

[2] 1 SEK was about 0.107 €at the time the experiment took place.

4.2.1 Quantities

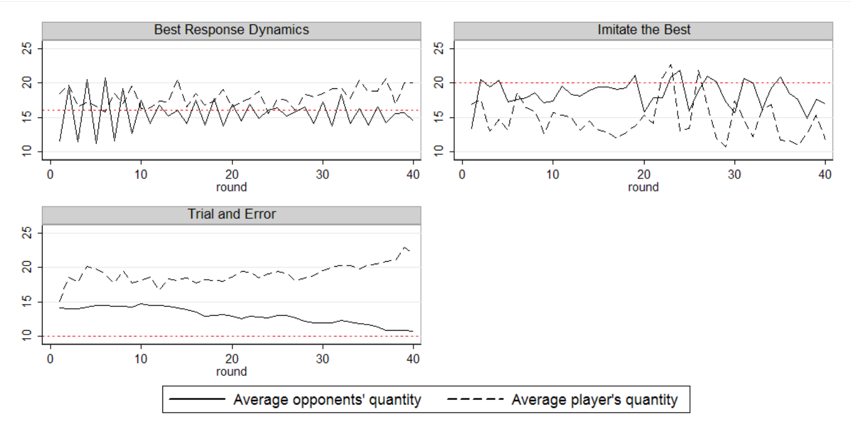
Table 4.1: Avearge quantities and prices across treatments

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

Table 4.1 displays the quantities produced on average by the subjects and by their "robot" competitors in the three treatments, first across all the 40 rounds, then just for the last 10 rounds.

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

Under the – somehow unrealistic – hypothesis that subjects follow exactly one of the three learning rules simulated by the computer, we should have observed at least in one of the three treatments a convergence towards the predicted equilibrium. In fact, the average choice of the robot players is not so far from what is predicted by the theory – still it is significantly different from it according to a Wilcoxon signed-rank[3] test under the three treatments – but the average quantity chosen by the players is even less close to the theoretically anticipated one and the distance increases if we look only at the last ten rounds: the observed values seem to depart from the predicted ones as the game proceeds (cfr. figure 4.1).



Note: the dotted line represents the theoretically predicted quantity.

Figure 4.1: *Average quantities chosen by subjects and by their "robot" competitors.*

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

4.2.2 Attention

Figure 4.2 shows the average share of the time dedicated by each subject to the four pieces of information they could look up during the game. The first noticeable fact is that most of the players'attention is devoted to the plot that represents profits earned and quantities produced in the previous period by the player himself and by each of his competitors.

[3]5% significance level

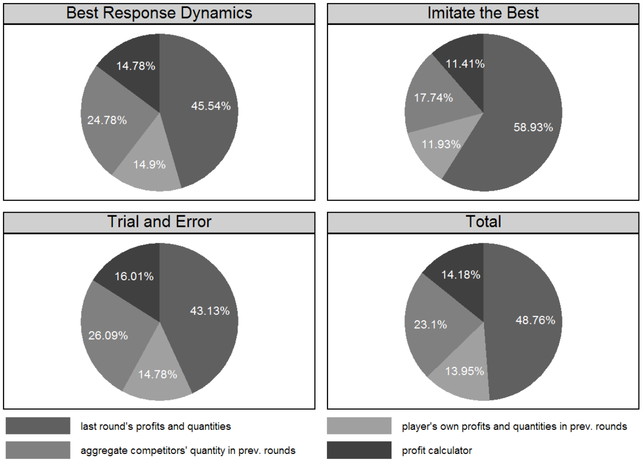


Figure 4.2: *Distribution of players' attention in the three treatments.*

This means for example that if they wanted to imitate the best performer in the previous period, as suggested by Vega Redondo, in general they know the information necessary to do it. On the other side, a theory of learning such as Trial and Error is less supported by our data, because subjects do not seem to be very interested in the graph representing the series of player's own profits and quantities, which includes the only information required to apply this learning model.

It also interesting to observe that the type of competitor the players face seems to affect not only their average choice but also the way they distribute their attention to the different pieces of information. Wilcoxon rank sum tests support the hypothesis that under treatment ItheB the share of time spent looking up last period profits and quantities is higher than in the other two treatments, while the time spent on the other three pieces of information is shorter. No significant difference in the allocation of attention emerges between the other two treatments. Given the small sample we are dealing with, though, it is not clear whether the observed difference among treatments is due to subjects' individual characteristics or if it is driven by the diversity in virtual competitors' behavior.

On the other hand, if we observe how the allocation of attention evolves

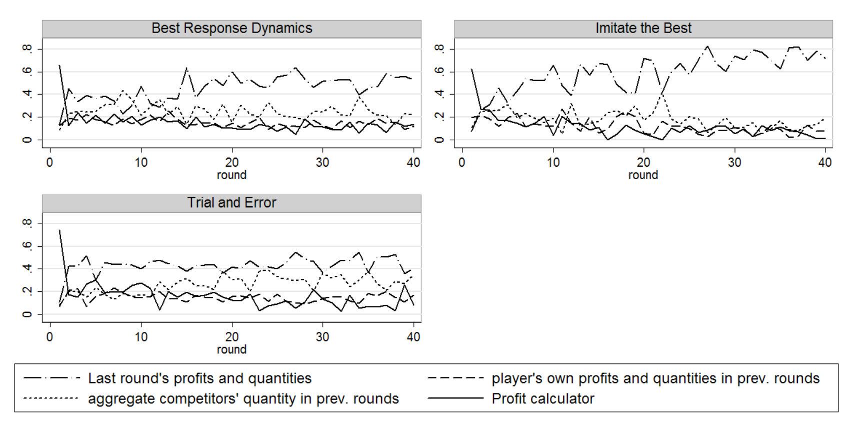


Figure 4.3: *Allocation of players' attention along the game.*

along the whole game (see figure 4.3), we notice the presence of a trend which is common across treatments. In the first round, obviously all the look up time is devoted to the profit calculator which is the only information available, probably in the attempt of figuring out how the market works. Then, from the second round on, players' attention seems to be mainly attracted by the plot displaying individual profits and quantities in the previous round, and, to a minor extent, from the graph showing the cumulated quantity chosen by the three virtual opponents in all the previous round. The attention dedicated by players to their own past remain scarce all over the game, reenforcing our skepticism about the trial and error learning model.

4.2.3 Learning

The sharp decrease in the decision time during the first 20 rounds (Figure 4.4) together with the decrease in the use of the profit calculator suggests that most of the learning about the market structure takes place during the first half of the game. Once they have clearly understood the relation between quantities, prices and profits, subjects focus their attention to what the other players do.

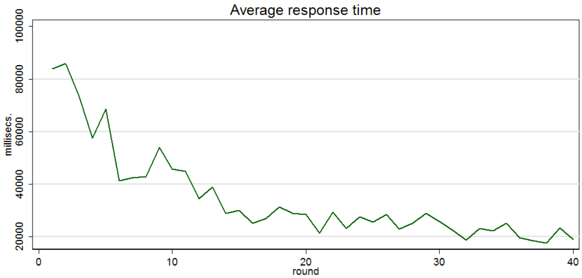


Figure 4.4: Average decision time across rounds, in milliseconds.

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

Simple learning models In a first attempt to have a picture of the learning model adopted by the players in this game, I adopted a measure proposed by HNO to assess to which extent the three simple learning rules presented in section 4.1.2 are able to predict each single choice the subjects made.

Let

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

where a_i^t is the quantity predicted, in turn, by Imitate the Best (IB), Trial and Error (TE) and Best Response Dynamics (BR)[4]. Clearly, $z_i^t = 1$ indicates that the rule perfectly predicted the move taken by the player. In general, $z_i^t > 0$ implies that the rule has correctly anticipated the direction

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

Table 4.2: *Hit ratios*

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

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

of the variation in the quantity produced, while the opposite is true when $z_i^t < 0$.

Table 4.2 shows that Best Response Dynamics is the rule that seems to better fit with our data, predicting the right direction of change in quantities in at least 70% of the cases and providing a rather precise forecast ($0.5 \leq z_i^t < 1.5$) in at least 19.41% of the observations. Contrary to what observed by HNO, Imitate the Best is the rule with the worst perfomance, since in two treatment it predicts the wrong direction of change in at least half of the cases.

Table 4.3 shows how many subjects report positive z_i^t values in at least 70% of the rounds and how many present hits close enough to 1 ($0.5 \leq z < 1.5$) at least 30% of their decisions. We observe that, again, Best Response Dynamics appears to be the only rule that is applied with a certain degree of consistency. No subjects seem to adopt the Imitate the Best rule, while two subjects behave in accordance with Trial and Error rule in the treatment in which this is the rule that inform the behavior of the vitual players.

Table 4.3: *Hit ratios at the individual level.*

	$z > 0$ in at least 70% of the rounds			$0.5 \leq z < 1.5$ in at least 30% of the rounds			Total
	TE	IB	BR	TE	IB	BR	
Best Response Dyn.	1	0	5	0	0	1	8
Imitate the Best	1	3	5	0	0	3	7
Trial and Error	2	1	5	2	0	3	8

Since from the analysis presented above it emerges that none of the three simple learning rules considered so far is able to exactly predict players' choices, I now shift focus to the direction of the players' output decisions. Following BDV, I shall consider a model in which the sign of a player's output change, Δq, is a function of the direction, x, indicated by the target output levels according to each of the three learning rules. This way, I should be able to determine to what extent each behavioral rule affects the way players adjust their output in every round of the game. Let:

$$\Delta q_i^t = sign(q_i^t - q_i^{t-1})$$

and, for every learning rule *r*, let

$$x_{r,i}^t = sign(a_{r,i}^t - q_i^{t-1})$$

where $a_{r,i}^t$ denotes the quantity predicted for player *i* at round *t* by rule *r*, as above. In this experiment, I also observe which different pieces of information each subject looked at in every single period. This data are incorporated into the model, to see whether and how information affects the way players adapt their choices as they gain experience of the game. Remember that in this game information is hidden behind four boxes on the computer screen. Denoting with *b* the box, I then created four dummy variables, $d_{b,i}^t$, indicating whether in the generic period t subject i opened the box *b* containing (*i*) the results provided by the Profit Calculator (ProfCalc), (*ii*) information about quantities individually produced by each of the players in the last period and corresponding profits (LastRound), (*iii*) quantities produced and profits obtained by the player herself in all the previous rounds (HistPl) or (*iv*) the sum of quantities produced by the player's three opponents in each of the previous rounds (HistOpp).
The ordered-probit model, then, assumes a latent response variable, z, that is a linear function of the independent variables plus a normally distributed error term, u [5]:

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

Altogether, the model includes 19 explanatory variables. Data from the three treatments were pooled, to obtain a sufficiently large number of observations. I first estimated the full model, then I progressively obtained a more compact model using Likelihood-Ratio tests with a

[5] For simplicity, I omit here subscripts for round and individual players.

significance level of 5%. For sake of simplicity, I present only the last estimate in table 4.4. To take into account the possible correlation between observations pertaining to the same subject, the variance-covariance matrix is corrected by clustering at the subject level.

Table 4.4: *Ordered Probit Model: Estimations*

	Coefficient	**Standard Error**
βs		
TE	0.069	0.078
BR	0.917***	0.300
IB	0.592***	0.086
γs		
LastRound	0.448*	0.240
δs		
HistPlxBR	-0.229**	0.096
ProfCalcxBR	0.449***	0.151
LastRoundxBR	-0.615**	0.310
cut1	0.184	0.232
cut2	1.065	0.249
N	874	
logL	-863.324	
Likelihood-Ratio test	$\chi^2(7) = 162.22$	

Note: In this table and in the following ones, the symbols ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 4.4 evidences some noticeable facts. First, regardless of the information observed by subjects, the learning rule based on myopic best reply seems to inform their choices to a great extent, which confirms what already highlighted through the hit ratios. Second, this attitude towards myopically best replying is counterbalanced by a tendency to imitate the best performer. Third, the information observed by the subjects significantly affect the way they behave. In particular, when a subject observes the quantities produced and the profits individually earned by each of her competitors in the previous round, her choices tend to be driven away from what predicted by the Best Response rule, and in general her output is likely to increase. Similarly, when subjects look at their own history of play – in terms of output produced and profits earned – their choices are generally less consistent with a myopic best reply, whose predictive power appears instead to be enhanced when subjects use the profit calculator. Finally, the learning rule based on Trial and Error does not find strong support in our data: the coefficient is never significant, no matter what information a subject looked at.

The first impression is that the hypothesis that subjects follow some very simple heuristic to choose their strategy in our game should be rejected. Learning through trial and error does not seem to be a plausible explanation of subjects behavior, both because the players pay too little attention the their own past profits and quantities, which is the only information required to apply this learning rule, and because their choices are not in line with what is theoretically predicted according to this model. On the other hand, imitating the best performer – per se – is not able to forecast the observed choices correctly, even if subjects' look up patterns are consistent with this learning model. Myopic best reply seem to drive players' choices, at least partially, and is supported by the information they acquire, on average. In fact, to apply this learning rule the subjects need to know the sum of the quantities produced by their competitors in the last period – an information they almost always look at – and they must be able to compute a best reply, which means that either they use the profit calculator or they have used it extensively in the past and already know what the best reply is. Still, this model does not fully explain the observed variations in players' behavior. In fact, myopic best reply cannot be the only driver of players' choices. If it was, under treatment BRD we should have observed convergence towards the Cournot equilibrium, but our data do not provide any evidence in this direction.

According to my ordered probit regression, players' behavior is rather driven by the interplay of best reply and imitation, which in a sense confirms Vega-Redondo's idea: even if subject are incline to adopt the best reply when they know the market structure sufficiently well, if they are provided with information about their rivals' strategies and choices, they will be tempted to imitate those who are more successful, which yields more competitive outcomes.

Still, the model I estimated is essentially based on three extremely simple learning rules, and could therefore be too rigid to encompass all the facets of players learning behavior. For this reason, I decided to estimate a second, more complicated learning model, based on the self-tuning experience weighted attraction learning model proposed by Ho, Camerer and Chong (2007) [48].

EWA Learning Model The parametric version of the Experience-Weighted Attraction (EWA) model was first proposed by Camerer and Ho [19] and [49]. It is a model that hybridizes features of other well known learning rules, such reinforcement learning and belief 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 model, attractions are updated using the payoff that a strategy either yielded, or would have yielded, in a period: the rule for updating attraction $A_i^j(t)$ attached by player *i* to strategy *j* in period *t* is

$$A_i^j(t) = \frac{\phi N(t-1) A_i^j(t-1)}{N(t)} + \frac{[\delta + (1-\delta) I(s_i^j, s_i(t))]\pi_i(s_i^j, s_{-i}(t))}{N(t)} \quad (4.1)$$

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 and $N(t)$ is a measure of the weight players put on past attractions relative to present ones; it can be interpreted as the number of "observation-equivalents" of past experience relative to one period of current experience.
$I(x, y)$ is an indicator function which takes value 1 if $x = y$ and value 0 otherwise. So, according to this model, we assume that players are able to evaluate the foregone payoffs they would have earned in period t had they chosen a different strategy s_i^j.

The parameter δ measures the relative weight given to hypothetical payoffs, compared to actual payoff $\pi_i(s_i(t), s_{-i}(t))$. The second parameter to be estimated is ϕ: a discount factor that depreciates previous attractions.

The variable $N(t)$ is also updated after every period according to the rule:

$$N(t) = \phi(1-\kappa)N(t-1) + 1 \quad t \geq 1 \tag{4.2}$$

where parameter κ determines the growth rate of attractions, which reflects how quickly players lock into a strategy: a third parameter that has to be estimated. When $\kappa = 0$ attractions are weighted averages of lagged attractions and past payoffs, so that attractions cannot grow outside the bounds of the payoffs in the game. When $\kappa = 1$ attractions cumulate, so they can be much larger than stage-game payoffs.

Attractions determine probabilities. More specifically: the probability $P_i^j(t+1)$ that player i chooses strategy j in period $t+1$ is monotonically increasing in $A_i^j(t)$ and decreasing in $A_i^k(t),\ k \neq j$. The relation between attractions and choice probabilities is represented by a logistic stochastic response function:

$$P_i^j(t+1) = \frac{e^{\lambda A_i^j(t)}}{\sum_k e^{\lambda A_i^k(t)}} \tag{4.3}$$

where the parameter λ measures sensitivity of players to attractions.

One of the main criticisms on parametric EWA concerns the number of parameters to be estimated. To solve this issue, Ho, Camerer and Chong (2007) [48] developed a simpler version on the model in which some parameters are fixed at plausible values, while others are replaced with functions of experience, that no longer need to be estimated. More specifically: parameter k is set equal to 0, because this capture almost all familiar learning models, and also because according the previous work by the same authors this parameter does not affect fit much. The initial experience $N(0)$ is restricted to be equal to 1, which is not a crucial assumption since in general subjects come to the experiment with weak priors, whose influence gradually disappears as the experiment proceeds. Finally, parameters δ and ϕ are replaced with functions of players' experience.

Parameter δ – representing weight of foregone payoffs – is substituted with the function:

$$\delta_i^j(t) = \begin{cases} 1 & \text{if } \pi_i(s_i^j, s_{-i}(t)) \geq \pi_i(s_i(t), s_{-i}(t)) \\ 0 & \text{otherwise.} \end{cases}$$

meaning that subjects reenforce, by a weight of one, only chosen strategies and all the other strategies that would have yielded a weakly higher payoff.
The discount factor ϕ is instead replaced by the "change detector" function $\phi_i(t)$ varying across time within the same game. The hypothesis made by the authors, here, 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. They 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 by 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 the previous ones. Vector $r_i(t)$ instead has all the elements equal to 0 but the k-th, 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 $r_i(t)$:

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

Self-tuning EWA model has two important advantages: first, it is particularly flexible since the functions $\delta_i^j(t)$ and $\phi_i(t)$ naturally vary across time, people, games and strategies; second, it can shift from a learning model to a different one as the game proceeds.

A modified version of self-tuning EWA model According to EWA learning model, attractions are updated keeping into account also foregone profits, but in the experiment I present here foregone payoffs from unused strategies are not known by the players. Subjects, though, can use the profit calculator to discover the profit a particular strategy would yield, given the strategies chosen by the other players. As explained in section 4.1.1, the profit calculator can be used in two different ways:

1. the profit calculator can be used by the players to evaluate the quantity that would yield them the highest profit given the aggregated quantity produced by their competitors, and inform them about the profit they would earn if they produced the suggested amount of good.

2. it can be also used to know the profit given both the quantity produced by the player and the sum of the quantities produced by his opponents.

By checking how a player used the profit calculator in each period, I know precisely which information he used to evaluate each strategy in every period.

If they wish, players can also access information about the profits earned 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.

Keeping this peculiar characteristics of the game in mind, I decided to change the attraction updating rule, so that attractions in every period t are modified considering three elements:

- the profit $\pi_i(s_i^j, s_{-i}(t-1))$ actually obtained by the player in period $t-1$;

- the profits $\pi^j_{i,imit}(t-1)$ obtained by each of the player's opponents playing strategy s^j in the previous period;

- the profits $\pi^j_{i,PC1}(t)$ and $\pi^j_{i,PC2}(t)$ evaluated by the player using the first and the second function of the profit calculator respectively, given his or her expectations about the competitors' choices [6].

[6]If the second function of the profit calculator is used more than once by player i in

While the player always knew the strategy he played in the previous round and the profit he obtained, π_{imit}, π_{PC1} and π_{PC2} may be known or unknown to the player, depending on the pieces of information he or she decided to look up.

To check for the information the subject is aware of, we define three dummy variables:

$$d^j_{i,imit}(t) = \begin{cases} 1 & \text{if in period } t \text{ player } i \text{ opened the box containing information about quantities produced and profits earned by each of his opponents in period } t-1 \text{ and in that period one of the opponents had played strategy } s^j. \\ 0 & \text{otherwise.} \end{cases}$$

$$d^j_{i,PC1}(t) = \begin{cases} 1 & \text{if in period } t \text{ player } i \text{ used the first function of the profit calculator, and this device indicates strategy } s^j \text{ as the best reply to the strategies played by the three opponents, and associates it to some profit } \pi^j_{i,PC1}(t) \\ 0 & \text{otherwise.} \end{cases}$$

$$d^j_{i,PC2}(t) = \begin{cases} 1 & \text{if in period } t \text{ player } i \text{ used the second function of the profit calculator, and this device associates strategy } s^j \text{ to some profit } \pi^j_{i,PC1}(t) \text{ given the opponents' strategies.} \\ 0 & \text{otherwise.} \end{cases}$$

These dummy variables, in a sense, replace the function $\delta^j_i(t)$ representing the weight of foregone profits in the original version of self-tuning EWA learning.

Now, it is possible to state our modified updating rule for attractions:

period t, the profit $\pi^j_{i,PC2}(t)$ is calculated as an average of the various profits associated to strategy s^j_i by the device (different profits correspond to different hypotheses about the other players' behavior).

$$
\begin{aligned}
A_i^j(t) = {} & \frac{\phi_i(t)N(t-1)A_i^j(t-1) + \alpha\pi_i(s_i^j, s_{-i}(t-1))}{N(t)} + \\
& + \frac{\beta d_{i,PC2}^j(t)\pi_{i,PC2}^j(t) + \gamma d_{i,PC1}^j(t)\pi_{i,PC1}^j(t)}{N(t)} + \\
& + \frac{\delta d_{i,imit}^j(t)\pi_{i,imit}^j(t-1)}{N(t)}
\end{aligned}
\tag{4.4}
$$

In this new formula, attractions in period t are updated according to a weighted average between (some of) the information the player has about the profit each strategy yielded in the previous period and the profit it may yield in the future.
Parameters α, β, γ and δ then measure, respectively, the relative weight given to player's own experience, to the results potentially provided by the the profit calculator and to the profits earned by player's opponents, if observed.

Note that, in our model, $A_i^j(t)$ depends on the profits actually earned and (possibly) on the profits the opponents achieved *in the previous period*, and may be updated in period t with the profits evaluated by the profit calculator. So, the probabilities $P_i^j(t)$ depend on $A_i^j(t)$, and not on $A_i^j(t-1)$, as in the original model. The updating rule is then:

$$
P_i^j(t) = \frac{e^{\lambda A_i^j(t)}}{\sum_k e^{\lambda A_i^k(t)}} \tag{4.5}
$$

Estimation procedure Model's parameters are estimated via maximum likelihood, using our small samples of experimental data for the three treatments we have.
I estimated my modified version of the EWA learning model, first assuming all players have the same parameters values regardless of the treatment – then keeping the three treatments separated.

The log-likelihood function is given by:

$$\begin{aligned} LL(\alpha,\beta,\gamma,\delta,\lambda) &= \sum_{t=1}^{40}\sum_{i=1}^{n} ln\left(\sum_{j=1}^{m} I(s_i^j, s_i(t))P_i^j(t)\right) \\ &= \sum_{t=1}^{40}\sum_{i=1}^{n} ln\left(\sum_{j=1}^{m} I(s_i^j, s_i(t))\frac{e^{\lambda A_i^j(t)}}{\sum_k e^{\lambda A_i^k(t)}}\right) \end{aligned} \tag{4.6}$$

where n is the number of players, and m the number of strategies.

As Ho et *al.* [49] and Camerer et *al.* [21], I assumed that initial attractions are the same for all players:

$$A_1^j(1) = A_2^j(1) = \cdots = A_n^j(1) = A^j(1), \;\; \forall j$$

and I used first period data to initialize attractions. Estimating initial attractions as parameters of the model, in fact, would introduce too many degrees of freedom into the model.

Unfortunately, it is not possible to apply here the technique adopted by Ho et *al.* [49] and Camerer et *al.* [21] to initialize attractions, and simply set:

$$\frac{e^{\lambda A^j(1)}}{\sum_k e^{\lambda A^k(1)}} = f^j \tag{4.7}$$

where f^j is the frequency of strategy j in the first period, because in our data-set a number of strategies are never played in the first round.

I then had to follow a somehow more complicated procedure and set the initial attractions so to maximize the likelihood of the first-period data, separately from the rest of the data, for a value of λ derived from the overall likelihood-maximization.

Our objective function is:

$$\begin{aligned} O(A^1(1), A^2(1), ..., A^m(1)) &= \sum_{i=1}^{n} ln\left(\sum_{j=1}^{m} I_i^j(1)\frac{e^{\lambda A^j(1)}}{\sum_{k=1}^{m} e^{\lambda A^k(1)}}\right) \\ &= n\left(\sum_{j=1}^{m} f^j \lambda A^j(1) - ln\left(\sum_{k=1}^{m} e^{\lambda A^k(1)}\right)\right) \end{aligned}$$

Clearly, the objective function is always decreasing in $A^h(1)$, for all the strategies h having $f^h = 0$. So, $A^h(1)$ should be arbitrarily set for these strategies. I chose to set $A^h(1)$ so that

$$P^h(1) = \frac{e^{\lambda A_i^h(1)}}{\sum_k e^{\lambda A_i^k(1)}} = \frac{1}{1000}, \quad \forall h : f^h = 0$$

For the other strategies, we have to solve the system given by the equations:

$$\frac{e^{\lambda A^j(1)}}{\sum_{f^k>0} e^{\lambda A^k(1)}} = f^j \quad j : f^j > 0 \tag{4.8}$$

I let the strategy j with the lowest frequency to have $A^j(1) = 0$ (which is necessary for identification) and solve for the other attractions as a function of λ and the frequencies f^j.

I made one last modification to the model proposed by Ho, Camerer and Chong (2007) [48], in the way I computed the "change-detector" function $\phi_i(t)$. Since in this game the strategy space is rather big, I used 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 $S_i(t)$.
Finally, as Camerer and Ho ([20]), I restricted λ to be positive.

Estimation Results Table 4.5 displays the results of the four regressions based on my modified version of self-tuning EWA learning model. From these estimates it emerges that the learning process taking place in this game is complex and results form the interplay of different components. First, we can observe that the use of profit calculator exerts a strong influence on players' choices: parameters β and γ – measuring the weight attached to the answers provided by the first and second function of the profit calculator, respectively – are both highly significant in all the treatments. Moreover, both β and γ are also always higher in value than the estimate for δ, the parameter sizing the importance attributed to competitors' profits, when a player observes them. On the other hand, it is worth noting that the estimate for δ is positive under all the treatments, and significant in two out of three of them. This seems to

Table 4.5: *Results of the EWA learning model*

	Treatment			
	pooled data	Best Resp. Dyn.	Imit. the Best	Trial and Error
	b/se	b/se	b/se	b/se
α	0.582***	0.685***	1.002***	0.373***
	(0.066)	(0.128)	(0.229)	(0.063)
β	0.948***	0.873***	1.971***	0.782***
	(0.112)	(0.174)	(0.452)	(0.138)
γ	0.919***	0.850***	0.732**	1.016***
	(0.132)	(0.256)	(0.368)	(0.185)
δ	0.260***	0.187**	0.227**	0.115
	(0.051)	(0.094)	(0.091)	(0.079)
$\ln(\lambda)$	0.443***	0.465**	0.501**	0.516***
	(0.108)	(0.183)	(0.214)	(0.155)
LL	-2310.844	-838.855	-750.956	-702.016
Sample size	920	320	280	320

suggest that even if imitation is not the main driving force of learning, it does play a non negligible role. Finally, we notice that also the estimates for α are positive and highly significant everywhere, which means that reinforcement learning based on player' own past experience, rather than on imitation or on beliefs about what the other players are going to do, is also important in this game.

We observed that in general the strongest weight is attributed to foregone and expected profits, measured by means of the profit calculator. Does this lead players towards forms of learning close to myopic best reply, or do they adopt more sophisticated forms of belief learning? Remember that players could use the profit calculator to find out the best reply to a given quantity produced by the opponents (PC1), or simply to evaluate the profit they would earn given the quantity chosen by the opponents and their own output choice (PC2). In both cases, the profit evaluated depends on player's expectations about the strategy adopted by the opponents. So we can use the opponents' quantity imputed into the profit calculator as a measure for players' expectations.

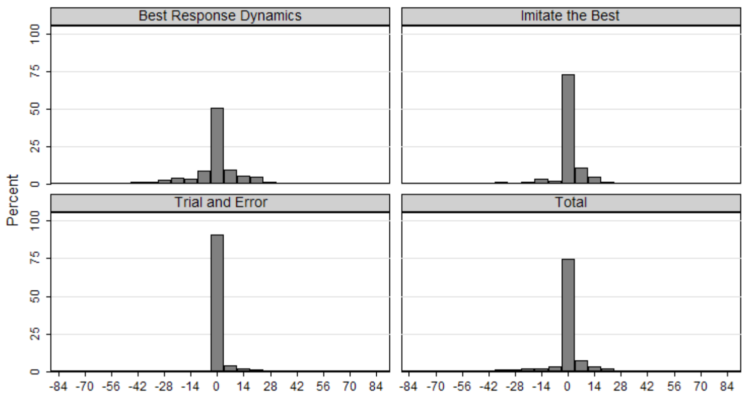


Figure 4.5: *Frequency distribution of the distance between the opponents' quantity entered in the profit calculator and the one observed in the previous period.*

Figure 4.5 presents the frequency distribution of the difference between this quantity and the aggregate quantity actually chosen by competitors in the previous period. According to these data, in most of the cases in which players used the profit calculator, they expected their opponents not to change their strategy much; but this can simply mean that they decided to use this device only when the opponents' strategy was stable enough to let them make predictions about the future.

In fact, I showed in section 4.2.2 that the information provided by the profit calculator is not the element to which players dedicate most of their attention. Table 4.6 shows also that on average only half of the players used the profit calculator in each round. So, even if Best Response Dynamics seems to inform players' behavior to a great extent, it is not the only force underlying the learning model adopted by the subjects: first because they are not always able to best reply since they do not use the profit calculator in every period, and because other models of learning – based on imitation or reinforcement – play a role.

Table 4.6: *Average usage of the first and the second function of the profit calculator, by treatment*

treatment	**PC1 (best reply)**	**PC2**	**overall**
BRD	0.497	0.078	0.575
ItheB	0.214	0.129	0.343
TandE	0.297	0.259	0.556
Total	0.341	0.157	0.498

4.3 Remarks

From the data collected through this first experiment emerges that players behavior cannot be encompassed by the simple models of learning I shortly described in section 4.1.2, namely learning through trial and error, imitation of the best and best response dynamics . My results confirm that information provided to the subject has an important effect on the way they behave. I observe that players tend to best reply to the strategy adopted by their opponents in the last period, when they have the necessary information to do so. Still, this is not the information they are most interested in: they dedicate most of their attention to the strategies individually adopted by their opponents and the payoffs each of them earned in the previous round. This piece of information seems to drive them away from the best reply, and possibly leads them to a more "imitative" behavior. Even if imitation is not the driving force of subjects' learning – so the market outcomes we observe are far away from those predicted by Vega-Redondo (1997) [102] – it still leads to a more aggressive competition than the one that would emerge if all players adopted a learning model only based on myopic best reply. According to the data I collected, finally, it also seems that the learning rule adopted by the opponent – extreme as it was – does not have a very strong impact on the model of information acquisition and processing adopted by the players.

Chapter 5

Second Experiment: human opponents

The first experiment, in which subjects played against computers, provided us with an interesting overview of how information and learning interact in determining subjects choices. Yet, it presents some drawbacks: first, the rather rigid behavior of the robot opponents does not allow us to draw sound conclusions about the market outcomes that may arise in the virtual market reproduced in the lab. Second, the absence of a time limit for taking decisions encouraged subjects to look at all the information they were provided, instead of forcing them to choose only the elements being really salient for their choices. Third, the experimental design does not allow us to verify if players really want to imitate the best performer or if they are influenced in some other way by the strategies adopted by the opponents, since a player could discover the quantities produced by each of his competitors by opening a single box on the screen. One last possible problem with the design of the first experiment concerns the way subjects opened the boxes hiding pieces of information. Since it was possible to open these boxes simply putting the mouse cursor over them, we cannot be sure that some of the lookups were unintentional.

I collected more data and fixed these problems by running an interactive version of the experiment, in which subjects played against each other. In this second variant of the experiment I introduced a time limit of thirty seconds for subjects to make their choice in each round, I increased the number of boxes on the subjects' computer screens to hide only a single piece of information behind each of them, so to have a more detailed control on players' information search pattern, and I let the subjects open a box by clicking on it to avoid unintentional look-ups.

5.1 Experimental Design

As I mentioned above, this second experiment was designed and run to improve and complete the results of the preceding one. The structure of the game – including market characteristics, number of competitors and number of repetitions – remained substantially unchanged. The major difference is in that this time each player's opponents are represented by other participants to the experiment, not enacted by the computer. Three other important innovations are introduced.
First, the graphical interface was radically changed; figure A.5 in the appendix shows how subjects' computer screen looks like [1] .

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 – served to display the strategy they chose in the previous period, that is the quantity they decided to produce. 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. As in the first experiment, 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. Finally, subjects could use the profit calculator, which had the same two functions described for the previous experiment (see section 4.1.1). As before, it was not possible to access various pieces of information at the same time, since opening a new window automatically

[1] A translation of the instructions is available in Appendix A.4

closed the previous one. This new interface granted me a greater control over the subjects' information search behavior, allowing me to check whether subjects are in fact more interested in the strategy adopted by the competitor who got the highest profit in the previous period, and making sure that every look-up is intentional.
Second, a time limit of 30 seconds per round was introduced, so to force subjects to choose the information they are really interested in, and to reproduce an environment in which rationality is bounded because of external factors. 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.

The experiment was programmed and conducted with the software z-Tree (Fischbacher 2007 [35]), more suitable than MouselabWEB to manage interaction between subjects.

5.2 Results

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 1 hour and 1 hour and 15 minutes, including instructions and payment. The average payment was 13 €with a maximum of 17 and a minimum of 9, including a show-up fee of 4 €.

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 [2], which were read aloud so to make them common knowledge. Thereafter, they had the opportunity to ask questions, which were answered privately. Before starting the real game, subjects were also asked to complete a test on their computer, aimed at checking their understanding of the

[2] A copy of the instructions, in Italian, can be found in Appendix

graphical interface they would have had to use during the game.
Only when all the players managed to answer correctly to all the questions in the test, the real game began. Each subject was randomly and anonymously matched with other three participants, who were to be his "opponents" throughout the whole game. At the and of the game, subjects were paid in cash, privately, in proportion to the profits they scored during the game.

In what follows I will first present some qualitative results about the output choices made by the subjects, and about their information search pattern. I will then try to establish a relation between the information acquired and the choices made by the subjects by means of the same two learning models adopted for the previous experiment. We will notice that, notwithstanding the differences between the two experiments, the results are surprisingly consistent.
Finally, I will briefly comment on the effects of education on the learning model adopted by the subjects.

Quantities

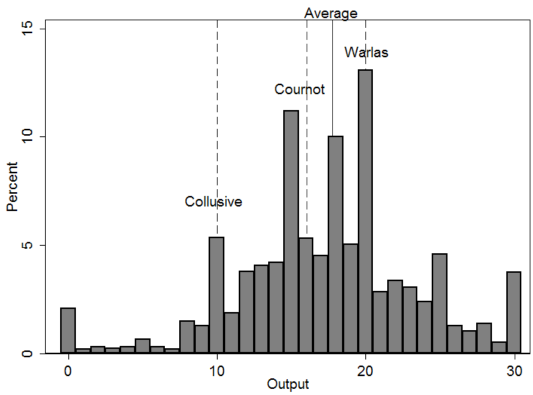


Figure 5.1: *Frequency Distributions of Individual Output Levels.*

Figure 5.1 displays the frequency distribution of the individual output choices in all the periods and sessions of the experiment. First, we observe that the average (17.76)[3] is higher than the Cournot output (16), but lower than the Walrasian one (20), which is instead the modal output. 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. The Pareto-dominant collusive outcome of 10 is chosen only in 5.36% of the cases.

Table 5.1: *Average individual output choice*

	Session 1	Session 2	Session 3
Periods 1-10	17.238	17.013	16.968
Periods 11-20	17.101	17.385	18.038
Periods 21-30	17.585	17.623	19.815
Periods 31-40	17.478	17.912	18.911
Total	17.353	17.487	18.436

Looking at table 5.1, we observe an increase in the average output as the game proceeds. For all the three sessions, though, a non parametric Wilcoxon rank-sum test fails to reject (at the 1% significance level) the hypothesis that observations for the first and the last ten periods are drown from the same distribution. We also notice that the average quantity produced in the third session is significantly higher than in the other two[4].

Figure 5.2 presents the aggregate quantity produced in each group across all the periods of a game. We notice that the variability in total outcome remains high even towards the end of the game, with fluctuations between the Cournot and Walrasian equilibrium outcomes. In general, therefore, we cannot speak of convergence.

[3] In this figure and in the following ones, the average is evaluated dropping the 40 observations in which the outcome was zero because a subject did not answered in time.

[4] According to a Wilcoxon rank-sum test, at 1% significance level.

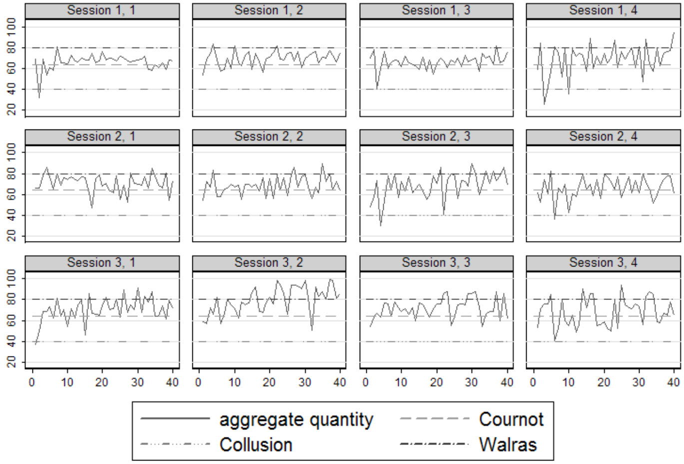


Figure 5.2: *Aggregate quantity produced within each group.*

5.2.1 Attention

Figure 5.3 shows that in this case most of players' attention is devoted to the profit calculator, contrary to what we observed in the previous experiment where subjects played against robot and were found to dedicate most of their look-up time to check the outcomes chosen and the profits obtained by each competitor in the previous period. Indeed, in this second experiment, the share of look-up time spent using the profit calculator is on average significantly higher than the time spent looking at any other piece of information, according to a Wilcoxon matched-pairs signed-ranks test (1% significance level). In line with what observed for the first experiment, in all the sessions players paid limited attention to their own past, so the trial and error learning model – which only requires this sort of information to be applied – finds weak support in these data. On the other hand, if we observe the results for the three sessions, we notice that some differences emerge. In particular, in the third session the time spent using the profit calculator is significantly less that in the other two sessions (according to a Wilcoxon rank-sum test, at 1% significance level), while more attention is paid to the outcome individually chosen

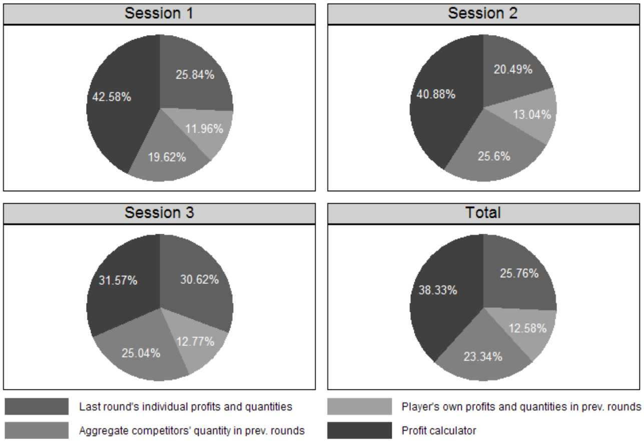


Figure 5.3: *Distribution of players' attention in the three sessions.*

by each of the player's opponents in the previous period. We have already observed above that the average outcome in the third session was significantly higher than in the other two sessions. So, these data provide some support to Vega-Redondo's idea that information about the strategies chosen by the opponents yields a more aggressive competition between players. In the following, we will see that this impression is supported also by a deeper econometric analysis.

Another noticeable fact emerges from figure 5.4: similarly to what already observed in the first experiment, the fraction of look-up time dedicated to the profit calculator decreases along the game, and on average is significantly higher in the first than in the last 20 periods (see table 5.2), while the opposite is true for the time spent looking at the output individually chosen by the player's competitors in the previous period. Again, this shift in players attention together with the previously observed increase in the average output level seems to be in line with Vega-Redondo's model, even if the evidence is weak due to the lack of significance of the increase in quantities.

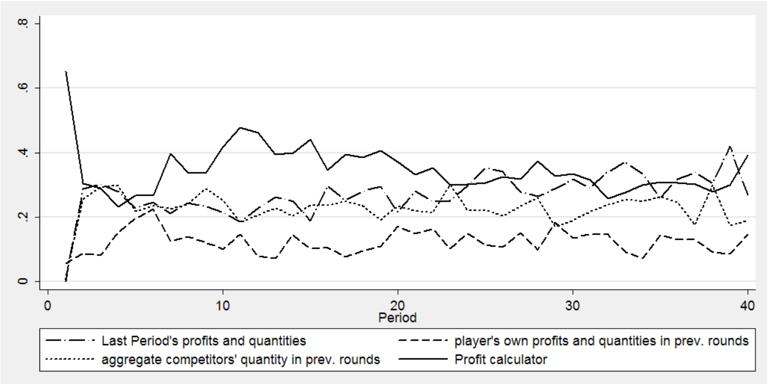


Figure 5.4: Allocation of players' attention along the game.

Use of the Profit Calculator We have already noted that players made wide use of the profit calculator in this game. In particular, this device has been used in almost half of the observations collected through the three sessions, mostly to evaluate the myopic best reply to some aggregate quantity hypothetically produced by the player's opponents (see table 5.3), which is in line with what we observed in experiment one (table 4.6).

Suppose a subject followed the aforementioned Best Response Dynamics rule and best replied to the aggregate output chosen by his opponents in the previous period: before using the profit calculator he should have gathered information about his competitors aggregate output in previous periods, by opening the appropriate box. When this happened, I claim that the look-up sequence is consistent with best response dynamics (BRD), and it turns out that this is the case in more than 80% of the times the profit calculator was used.

When using the profit calculator, subjects had 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. Figure 5.5 presents the frequency distribution of the difference between this quantity and the aggregate quantity actually chosen by competitors in the previous period.

Table 5.2: *Fraction of look-up time dedicated to the profit calculator and to competitors' output choices in the first and in the second half of the game.*

	Periods	Session 1	Session 2	Session 3	Total
profit calculator	*1-20*	0.384	0.414	0.319	0.371
		$\sim$	$>^{**}$	$>^{***}$	$>^{***}$
	21-40	0.370	0.309	0.261	0.315
competitors' output choice	*1-20*	0.276	0.172	0.271	0.241
		$\sim$	$<^{***}$	$<^{***}$	$<^{***}$
	21-40	0.293	0.295	0.338	0.308

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

Table 5.3: *Use of the two functions of the profit calculator, and percentage of observations in which the look-up sequence is consistent with a myopic best reply.*

Use of the p.c.	N. obs (%)	% of L.U. sequences consistent with BRD
both functions	130 (13.56%)	91.54%
1st function (best reply) only	584 (60.90%)	82.86%
2nd function only	245 (25.55%)	79.62%
total	959	82.06%
Profit calculator not used	961	–

According to these data, more than half of the times the profit calculator was used the quantity inputed belonged to the interval $[Q_{-1}(t-1) - 3, Q_{-1}(t-1) + 3]$, where $Q_{-1}(t-1)$ represents the sum of the quantities produced by the player's opponents in the previous period. This is consistent with what previously observed for the first experiment, and provides further support to the best response dynamics as a model of learning in this setting.

Interest for the strategies adopted by opponents. As we have seen above, a considerable amount of attention is dedicated to the boxes show-

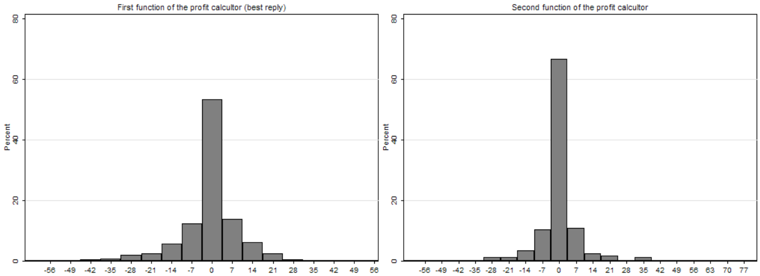


Figure 5.5: *Frequency distribution of the distance between the opponents' quantity entered in the profit calculator and the one observed in the previous period.*

ing the output individually chosen by each of the player's opponents in the previous period. Table 5.4 shows that on average the look-up time

Table 5.4: *Average look-up time*

	Best		Not Best
Opponent 1	2.07	$>^{***}$	1.53
Opponent 2	1.44	$>^{***}$	0.78
Opponent 3	2.45	$>^{***}$	1.69

Note: The statistical test is a two sample Wilcoxon rank-sum test. The symbol *** indicates significance at the 1% level

dedicated to each of the opponents was greater when he had gotten the highest profit in the previous period. Still, if we compare the time spent looking at the strategy adopted by "the best" with the total time dedicated to the strategies adopted by the other opponents, we notice that the latter is significantly higher (see table 5.5).

So, in partial contradiction with what suggested by Vega-Redondo's theory of imitation, players in this experiment seem to be concerned not only with the choice made by the competitor who best performed in the previous period, but also with the output chosen by each of the others.

Table 5.5: Average look-up time

Session	Best opponent		Other opponents
Session 1	1.700	$<^{***}$	3.088
Session 2	1.317	$<^{***}$	2.146
Session 3	2.309	$<^{***}$	3.211
Total	1.775	$<^{***}$	2.815

Note: The statistical test is a Wilcoxon signed-rank test. The symbol *** indicates significance at the 1% level

5.2.2 Learning

As for the previous experiment, I start the analysis of the learning mechanisms adopted by the players by comparing the explanatory power of the three aforementioned simple learning models: trial and error, imitation of the best and best response dynamics.

First, I use the same measure adopted in section 4.2.3 to assess to which extent these three simple learning rules are able to predict each single choice the subjects made. So again I let

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

where a_i^t is the quantity predicted, in turn, by Imitate the Best (IB), Trial and Error (TE) and Best Response Dynamics (BR).

Table 5.6 confirms what we already observed for the first experiment, namely that best response dynamics is the rule that provides the most precise forecast ($0.5 <= z_{BR}^t < 1.5$) and predicts the right direction of change ($z_{BR} > 0$) with the highest frequency. It is also worth noting that "imitate the best" in general overshots, when it predicts the right direction of change, meaning that it forecasts a variation in quantity that is more than twice as big as the actual one ($0 <= z_{IB} < 0.5$); the reverse is true for "trial and error" model. Table 5.7 shows how many subjects report positive z_i^t values in at least 70% of the rounds and how many present hits close enough to 1 ($0.5 \leq z < 1.5$) at least 30% of their

__Table 5.6:__ Hit ratios (%)

	$z < 0$	$0 <= z < 0.5$	$0.5 <= z < 1.5$	$z >= 1.5$
IB	45.45	26.70	21.38	6.47
BR	29.61	27.74	28.29	14.36
TE	47.09	13.21	6.14	33.55

decisions. None of the subjects seems to adopt the "trial and error" rule, while nine players behave according to what predicted by "imitate the best" in at least 12 periods. Yet, "best response dynamics" is the model that seems to inform the behavior of most of the players.

__Table 5.7:__ Hit ratios at the individual level

	$0.5 <= z < 1.5$	$z > 0$	**total**
IB	9	3	48
BR	18	25	48
TE	0	0	48

Being interested also in how information affects the learning model players adopt, I grouped observations by the type of information subjects spent most of their look up time on and measured the percentage of observations in which ($0.5 <= z_i^t < 1.5$), for each of the three basic learning rules (Table 5.8). We notice that in general "best response dynamics" prevails over the other two rules, but when players dedicated the greatest part of their attention to the output individually chosen by each of their competitors in the previous period, then is "imitate the best" rule that gets the highest score, and this effect is even more pronounced when the time spent looking at the strategy adopted by the best among the opponents is longer than the time dedicated to the others. This is one more element in favor of Vega-Redondo's theory claiming that when subjects have information about their opponents strategies and payoffs, they tend to become more competitive since they are tempted to imitate the one who got the best result.

Ordered probit estimation

Estimation procedure Analogously to what has been done above for the first experiment, I now consider a model in which the sign of a player's output change, Δq, is a function of the direction, x, indicated by the target output levels according to each of the three learning rules. This way, I should be able to determine to what extent each behavioral rule affects the way players adjust their output in every round of the game.
As above, let:

$$\Delta q_i^t = sign(q_i^t - q_i^{t-1})$$

and, for every learning rule *r*, let

$$x_{r,i}^t = sign(a_{r,i}^t - q_i^{t-1})$$

where $a_{r,i}^t$ denotes the quantity predicted for player *i* at round *t* by rule *r*, as above. In this experiment, I also control for the information gathered by players in every period. As mentioned above, each of the three basic learning models considered here requires a different type of information: more precisely, to imitate the best one must have looked at the output produced by the best competitor in the previous period, to play according to trial and error a player needs to remember the strategy he has chosen and the profits he has obtained in the last two periods, meaning that he probably would have to open the box containing information about his own history of play, while to best reply he must know the aggregate output of his competitors in the previous period and he must use the profit calculator. For this reason, I created three dummy variables $d_{b,i}^t$, which indicate respectively if player i in period t had the information necessary to best reply (InfoBR), to imitate the best (InfoIB) or to follow the "trial and error" model (InfoTE).
These dummy variables enter the model per se, but are also interacted with the variables x denoting the sign of the output variation predicted by the three learning models considered here.

Table 5.8: $0.5 <= z < 1.5$ (%)

	Longest L.U. Time	**IB**	**BR**	**TE**	**N. obs.**
opponents' individual output	mostly to the best	36.73	19.05	4.08	147
	mostly to the others	23.79	23.45	5.52	290
profit calculator	L.U. seq. consistent with BR	17.77	37.50	8.43	664
	L.U. seq. not consistent with BR	20.37	31.48	1.85	108
	opponents' aggregate past output	22.66	23.56	4.83	331
	player's past output and profits	20.32	23.53	6.95	187
	no information acquired	14.43	15.46	3.09	97
	Total	21.38	28.29	6.14	1824

Since subjects repeatedly interact with the same three opponents throughout the whole game, a critical point in this analysis is how to control for repeated observations of the same individuals or the same group. Moreover, I also wanted to check for possible correlation between data collected within each of the three sessions. For this purpose, I adopted a multilevel model with a random effect at the subject level nested within a random effect at the group level, which in turn is nested within a random effect at the session level.

More specifically, I assume the latent response variable, z, be a linear function of the independent variables plus a subject specific error term $\zeta_{i,g,s}$, a group specific error term $\eta_{g,s}$, a session specific error term θ_s and finally and i.i.d. error term $u_{t,i,g,s}$. Random intercepts are assumed to be independently normally distributed, with a variance that is estimated through the regression.

The full model is then:

$$z = \sum_r \beta_r x_r + \sum_b \gamma_b d_b + \sum_r \sum_b \delta_{b,r} d_b x_r + \zeta + \eta + \theta + u$$

where subscripts r and b both take values in $\{IB, TE, BR\}$. For simplicity I omitted the subscripts for individual, group, session and period.

The dependent variable is derived in the standard way for an ordered probit given the latent variable and cutoffs between categories. Maximum likelihood estimation is used to fit values for the cutoffs, β, γ and δ, and for the variances of the subject, group and sessions specific error terms. The model was estimated using GLLAMM [5], a software specifically designed to provide a maximum likelihood framework for models with unobserved components, such as multilevel models, certain latent variable models, panel data models, or models with common factors.

I first estimated the full model, then I progressively obtained a more compact model using Likelihood-Ratio tests with a significance level of 5%. First, I checked for significance of the variances at the session and

[5] see Rabe-Hesketh and Skrondal, 2004 [94] and http://www.gllamm.org

group levels; the null hypothesis that these were not significant was not rejected, so I adopted a more compact model with a random effect only at the subject level. Then, in steps, I eliminated the dependent variables that turned out not to be significant. For sake of simplicity, I present only the last estimate of this reduced model in table 5.9.

Results This regression evidences some results that confirm what we already observed for the first experiment, and some which instead go in a different direction. As for the first experiment (see table 4.4), we notice that the rule based on trial and error does not find strong support in these data even if it seems to guide, at least in part, players' behavior when they acquire the information necessary to apply it. In the previous experiment, we noticed that when players looked at the strategies adopted in the previous period by their opponents, and at the relative profits, they tended to increase their quantity. Here a similar effect emerges, since the estimated coefficient for InfoIB is positive and highly significant.

Remarkably, if the impact of information is not taken into consideration, the "imitate the best" model seems to account for the greatest part of output variations; the coefficient for "best response dynamics" is also positive and significant, but is smaller in magnitude. This relative weakness of the best response model disappears if we consider the effect of information: indeed, according to these statistics, if a subject acquires any of the three pieces of information considered here he will be more incline to move in the direction predicted to best reply, and particularly so if he uses the profit calculator after having looked at the aggregate output produced by his competitors in previous periods (namely, when InfoBR$=1$).

One possible reason why in this experiment the coefficient for x_{IB} is higher than the one for x_{BR} – contrary to what observed in the first experiment – is that here subjects were always informed about the profits individually obtained in the previous period by each of the players in their group. It is possible, then, that any time they realized that their profit was not the highest they tended to increase their output, then moving in

Table 5.9: Ordered Probit Model: Estimations

	Coefficient	**Standard Error**
βs		
TE	-0.022	0.041
BR	0.134**	0.066
IB	0.484***	0.053
γs		
InfoIB	0.327***	0.080
δs		
InfoTExTR	0.148**	0.068
InfoTExBR	0.151**	0.067
InfoBRxBR	0.520***	0.067
InfoIBxBR	0.258**	0.075
cut1	-0.041	0.065
cut2	0.369***	0.066
N	1824	
logL	-1576.005	

Note: In this table and in the following ones, the symbols ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

the direction predicted by "imitate the best", even if they did not know the exact output chosen by the player who had got the best profit.

EWA learning model I conclude my analysis of the learning process adopted by players in this game with the EWA learning model, originally proposed by Camerer and Ho and already described in section 4.2.3. Remember that in this model I hypothesized that after each period players update the attraction associated with every pure strategy according to the information they collected in that period, and these attractions are assumed to determine the choice probability distribution over the strategy set. More specifically, in the first experiment attractions were updated on the basis of the (*i*) the profit actually obtained by the player in the previous period, (*ii*) the profits obtained by each of the player's opponents in the previous period, if the player had this information and

(*iii*) the profits evaluated through the profit calculator, in case it was used by the player.
As mentioned above, I modified the design of this second experiment in order to be able to disentangle the effect of information about profits and strategies of the best among player's opponents, and of the same information concerning other opponents. I therefore introduced a slight modification to the way attractions were updated in equation (4.4):

$$
\begin{aligned}
A_i^j(t) = & \frac{\phi_i(t)N(t-1)A_i^j(t-1) + \alpha\pi_i(s_i^j, s_{-i}(t-1))}{N(t)} + \\
& + \frac{\beta d_{i,PC2}^j(t)\pi_{i,PC2}^j(t) + \gamma d_{i,PC1}^j(t)\pi_{i,PC1}^j(t)}{N(t)} + \\
& + \frac{\epsilon \sum_{h\neq i} d_{i,h}^j(t) b_{i,h}(t)\pi_{i,imit}^j(t-1)}{N(t)} + \\
& + \frac{\zeta \sum_{h\neq i} d_{i,h}^j(t)(1-b_{i,h}(t))\pi_{i,imit}^j(t-1)}{N(t)}
\end{aligned}
\tag{5.1}
$$

where the dummy $d_{i,h}^j(t)$ indicates whether player i in period t knew that his opponent h had played strategy s^j in the previous period, the dummy $b_{i,h}(t-1)$ takes value one if player h had the highest profit in period $t-1$ among the opponents of player i, and $\pi_{i,imit}^j(t-1)$ represents the profit obtained in period $t-1$ by player i's opponent choosing strategy s^j. All the other terms of the equation have exactly the same meaning as in equation (4.4), and the rest of the estimated EWA model is perfectly alike the one I estimated for the previous experiment.

In order to test whether the difference between the estimated weight attributed to the profits achieved by the best competitor and the profits earned by other competitors is significant, I also estimated a restricted version of EWA model, in which equation (5.1) is replaced by equation

(5.2)

$$A_i^j(t) = \frac{\phi_i(t)N(t-1)A_i^j(t-1) + \alpha\pi_i(s_i^j, s_{-i}(t-1))}{N(t)} + \\ + \frac{\beta d_{i,PC2}^j(t)\pi_{i,PC2}^j(t) + \gamma d_{i,PC1}^j(t)\pi_{i,PC1}^j(t)}{N(t)} + \\ + \frac{\delta \sum_{h \neq i} d_{i,h}^j(t)\pi_{i,imit}^j(t-1)}{N(t)} \tag{5.2}$$

which basically replicates equation (4.4) for this second experiment.

Results Table 5.10 displays estimation results for the unrestricted and restricted version of the EWA model. These results are surprisingly

Table 5.10: *Results of the EWA learning models*

	unrestricted		restricted	
	b	se	b	se
α	0.476***	0.097	0.477***	0.148
β	0.845***	0.171	0.848***	0.263
γ	1.187***	0.241	1.189***	0.373
δ	–	–	0.356***	0.114
ϵ	0.338***	0.077	–	–
ζ	0.383***	0.089	–	–
ln(λ)	0.499**	0.197	0.497	0.308
Log-Likelihood	-5680.350		-5680.603	
Sample size	1920		1920	

in line with those obtained for the first experiment and presented in table 4.5, notwithstanding the differences between the two experiments, and the two subject pools. Learning in this setting appears to be a blended process in which different components play an important role. The component related to belief learning seems to predominate: subjects attach the highest weight to strategies hypothetically tested by means of the profit calculator. A Wald test fails to reject the hypothesis that β and γ are equal, so according to my results subjects tended to attribute less importance to the results from the first function of the profit calculator

– evaluating best reply – than from the second one – which computes the profit given the output choice of the player and the aggregate output hypothetically produced by his opponents.
Both a Wald test on the results of the unrestricted model, and a likelihood ratio test between the restricted and the unrestricted model fail to reject the hypothesis that ϵ and ζ are equal. This means that subjects, when evaluating a strategy, do take into account the profits realized by other players choosing that strategy (the coefficient is always positive and highly significant), but do not attach more weight to the profit realized by the best among their competitors.
Finally, we notice that the estimates for α are positive, significant, and higher than those for parameters δ and ϵ and ζ, respectively for the restricted and unrestricted models. This suggest that reinforcement learning, based on player's own past experiences – plays a role, which seems to be even more important than the one played by "imitation".

5.2.3 Individual Characteristics

Finally, I would like to point out some interesting differences emerging among students having different educational backgrounds. Table 5.11 shows some facts about the composition of our subject pool.

Table 5.11: *Subjects' education*

	master	bachelor
Business Administration	8	3
Law and Economics	1	
Commercial Studies		6
Economics	6	3
Finance	2	7
Marketing		12

First, it is worthwhile noting that master students pay relatively less attention to the strategies individually adopted by their opponents, and more to the profit calculator. According to our previous results, this should make them less incline to imitate and more to best reply, which in theory should be viewed as a more “rational” behavior. In fact, master students turn out to be less aggressively competitive: the average quantity they choose is significantly lower, and so are their profits. They seem to adopt a follower behavior, best replying to opponents who tend to keep their own output high.

Table 5.12: *Comparison between bachelor and master students*

	bachelor		master
Average share of L.U time			
competitors’ individual output	0.314	$>^{***}$	0.204
player’s own history of play	0.119	$<^{***}$	0.133
competitors’ past aggregate output	0.213	$<^{***}$	0.259
profit calculator	0.329	$<^{**}$	0.368
Average output	17.76	$>^{***}$	16.72
Average profit	176.82	$>^{***}$	158.14

Second, sizable differences emerge between students with different curricula of studies. In particular, students in Finance, Marketing and Law and Economics show much less interest in the profit calculator, and much more in their competitors’ output. This is somehow surprising, since all curricula except Commercial Studies and Law and Economics envisage an introductory course in microeconomics – including elements of the theory of oligopoly – for first year bachelor students, while only master students in Economics have a specific training in Industrial Organization and Game Theory. So, it is not clear whether this different attitude towards information and learning – which as we have seen has an impact on the level of market competition – derives from a different approach to economics characterizing different curricula or from individual attitudes that in turn have affected also subjects’ choice for a certain course of studies.

Table 5.13: *Allocation of attention: a comparison between different curricula*

	Business Admin.	Comm. Studies	Economics	Finance	Marketing	Law and Econom.
Competitors' individual output	0.234	0.219	0.123	0.428	0.340	0.476
Player's own history of play	0.118	0.130	0.113	0.102	0.145	0.176
Competitors' past aggregate output	0.227	0.207	0.262	0.188	0.251	0.152
Profit calculator	0.383	0.409	0.476	0.258	0.244	0.136

Chapter 6

Conclusion

In this chapter I 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.

Two versions of this experiment were presented here: a first preliminary version was run in Stockholm, letting subjects play against virtual opponents, simulated via computer programs, and a second, more complete version run in Bologna, asking subjects to play against each other.

Notwithstanding the differences between the two experiments, some common results emerge.

First, learning appears to be a composite process, in which different components coexist. The leading element seems to be a sort of belief learning, in which 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 inputed in the profit calculator – a proxy for players'expectations – is pretty close to the aggregate opponents' output observed in the previous period, meaning 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. Second, a considerable amount of look-up time is dedicated to the strategies individually adopted by competitors, especially in the first experiment, when they are enacted by a computer program. As predicted by Vega-Redondo's theory, this piece of information generally boosts competition. Yet, the results of the second experiment 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 my modified version of EWA 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 by subjects assessing the "strength" of a particular strategy. Anyhow, all tests I have done agree on that imitation is not the main driving force in the observed learning process.
Third, the "trial and error" learning model which was found to perform quite well in HNO does not find strong support in my data. Subjects are not interested in their own past history of play, which is the only piece of information required by this learning rule, and the model often fails to predict even the direction of change in players' output.
Fourth, the model I derived from Camerer and Ho's EWA learning stresses the importance played by reinforcement learning in this setting: when assessing the strength of a strategy subjects seem to take into greater consideration their own experience than what they know about other players' results.

Finally, from the data collected in Italy with my second experiment, it emerges that subjects's specific training in economics might affect their behavior both in terms of information search pattern and in terms of actual choices. This aspect deserves further investigation and suggests that it could be interesting to repeat the experiment with market professionals, in order to see whether their experience in the field affects their approach to the game.

With my experiment I meant to contribute to the understanding of learning mechanisms in game-like situations. I also wanted to test experimental devices based on the "Mouselab" technique as scientific instruments that might be usefully adopted in other experiments on learning and to investigate other interesting situations in which imperfect information of some of the agents plays a crucial role, 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 situations like those, a better comprehension of the relation between the data and stimuli provided to economic agents and their choices might help the regulator to set rules of information disclosure that bring the market outcome towards a more efficient equilibrium.

Part II

Deterring Collusion

Chapter 7

Fines, Leniency, Rewards and Organized Crime [1]

joint with Sven-Olof Fridolfsson, Chloé Le Coq and Giancarlo Spagnolo

7.1 Introduction

The last decade has witnessed major innovations in the law enforcement against price cartels. Following the US example, leniency programs that reduce sanctions for cartel members that self-report to the competition authority have been introduced in most OECD countries and have become the main tool for cartel discovery and prosecution.[2] The European Competition Network, a forum including all European competition

[1]Many thanks to Tore Ellingsen, Magnus Johanneson, Dorothea Kuebler, Joe Harrington, Charles Plott, Patrick Rey, Jean Tirole, for discussions and advice related to this project, and to audiences in Berlin (ESMT, WZB), Copenhagen, Gerzensee (ESSET 2007), Mannheim (RNIC 2007), Toulouse, Stockholm (Ifn and Konkurrensverket) and Naples (Università Federico II) for comments and suggestions. We also gratefully acknowledge research funding from Konkurrensverket (the Swedish Competition Authority) that made this research possible.

[2]The Antitrust Division of the US DoJ has had a leniency programs for cartels since the seventies but reformed the program in 1993 and 1994, introducing the novel Corporate and

authorities also launched a "model" for the design of effective leniency programs.[3] In some jurisdictions (e.g. Korea) reward schemes for whistle-blowers that report a cartel have also been adopted, following their successful use in the fight of government fraud (US False Claim Act) and tax evasion.[4]

The introduction of leniency programs increased dramatically the number of cartels detected and convicted in the US and the EU, and this is why they are considered a tremendous success (see Spagnolo 2008 [99] for details). A higher number of detected and prosecuted cases, however, is not always a good indicator of the effectiveness of Anti-trust policies.[5] For example, an extremely lenient policy that reduces fines to almost all parties of a discovered cartels in exchange for information will enormously facilitate prosecution and generate many spontaneous reports. Such a policy could well make a competition authority famous as a successful agency, but is likely to heavily damage society by at the same time (a) encouraging cartel formation through the drastic reduction in expected fines that such a overly lenient policy generates, and (b) increasing the cost of prosecution (by the higher number of prosecuted cartels, given that prosecution costs are a pure dead weight losses for society). Law enforcement's main objective is crime deterrence, i.e. prevention. An efficient and successful antitrust policy against cartels should have tough ex ante deterrence effects that keep low both the costs of prosecution and those of price fixing activities.

The purpose of this paper is to examine the deterrence effects of antitrust fines, leniency programs, and reward schemes for whistle-blowers.

Individual Leniency Policy, and later on introduced the Amnesty Plus scheme. Analogously, the EC's DG Comp introduced a first Leniency Notice in 1996, and revised it in 2002.

[3]The ECN leniency "model" is substantially more lenient than the US program, as it allows to partially reduce fines practically to all members of a cartel, while the US leniency program is restricted to the first party that reports only.

[4]In December 2006 The US Congress strengthened the legislation that already allowed whistleblowers to cash as rewards 15 to 30% of taxes and fines recovered by the IRS thanks to their help, by making the payment of the reward almost automatic. A change that, by reducing the agency discretion on the payment of the reward, resembles the 1993 change in the US Corporate Leniency Program.

[5]This was often mentioned in the lively debate on the effectiveness of antitrust enforcement (see e.g. Crandall and Winston 2003 [27], Baker 2003 [8], Kwoka 2003 [70], and Werden 2003 [104], among others).

In particular we focus on how monetary fines, leniency programs, and reward schemes against price fixing cartels affect market participants' decision to form cartels (cartel deterrence) and their price choices. We also analyze how different design of antitrust policy may affect firms' ability to enter in tacit collusive agreements instead, since as Whinston (2006) [106] recently reminded us, the final objective of competition policy is to keep prices at competitive level, not to deter explicit horizontal agreements per se. The main questions raised in this paper are, therefore, the following:

1. What are the effects of traditional antitrust law enforcement, fines following successful investigations of the competition authority but no leniency, on cartel formation and on pricing behavior in and outside the formed cartels?

2. What are the effects of introducing a leniency program when reporting the cartel? Does it make a difference if the report is secret or not? Do things improve when expected fines are higher, or when the ringleader is banned from leniency as in the US (but not in the EU)? Is the possibility to report used as improved opportunity to undercut the cartel, or as a threat to punish defectors and thereby stabilize cartels, or both?

3. What are the effects of rewarding the first party that applies for leniency with a bonus equal to the fines paid by the co-conspirators in that cartel? Also, do agents exploit the reward scheme taking turn in reporting and cashing the reward when the scheme is too generous and makes this a profitable option?

4. What are the effects of these different law enforcement instruments on agents' choice of collusive price and on their ability to sustain *tacit* collusion, given the importance of this issue in the recent debate?

The big problem with optimizing law enforcement policy by looking at its real world performance is that for cartels and analogous forms of organized crime (fraud, corruption, earning management, etc.) there is precious little else to look at than discovered and prosecuted cases. Contrary to most other types of crimes, where there are conscious victims that denounce and thereby signal the frequency of the crime independently from the fraction of these crimes where the criminal is detected and convicted, victims of cartels and analogous forms of organized crime (corruption, fraud, etc.) are mostly not aware of them. This implies that we cannot directly observe the total population of cartels in society and how this changes with the introduction of new policies, though indirect methods offer partial indications (see e.g. Harrington, 2007 [44]; Miller, 2007).[6] This intrinsic lack of observability, accompanied by the fact that many design features of the proposed and theoretically analyzed schemes have never been implemented in reality, makes experimental investigation a crucial policy tool, an almost unique possibility to try to measure the likely change in deterrence, prices and welfare caused by the many different possible designs of law enforcement policies.

We consider an experimental framework, as close as possible to the strategic situation agents face in an oligopolistic industry subject to current antitrust laws, in which subjects play a repeated Bertrand price game with differentiated goods. Subjects can decide to coordinate on price (and thus they form a cartel). We consider several treatments different in the probability of cartels being caught, the level of fine, the possibility of self-reporting (and not paying a fine), the existence of a reward for reporting, the option to communicate, and cartel leaders access to leniency. We are not the first to look at these issues experimentally. Apesteguja, Dufwemberg and Selten (2006) [2] and Hinloopen and Soetevent (2006) [47], for example, have already produced instructive pieces of work in this direction. However, as will be explained in depth in the next section discussing the literature, we found that both those previous experiments could be further improved in one way or another, and that they do not cover most of the important policy issues we wanted to deal with in our experiments. In particular we find new results on secret reports, reward

[6] Harrington (2007) [44] develops a smart indirect method to estimate the likely changes in deterrence caused by the introduction of a new law enforcement instrument based on the observed changes in duration of the detected cartels. Miller (2007) further develops the approach and applies it to cartels detected in the US in the last decades, finding positive deterrence effects of cartel formation consistent with those we observe in our experiment. Unfortunately this work appears not to offer results or implication regarding price and welfare changes, which as our results show may not go in the expected direction (Sprouls 1993 [100] offers empirical evidence that prices weakly increase after antitrust conviction, which is also consistent with our experimental results).

schemes, the interaction between fines, leniency and deterrence and tacit collusion.[7]

Not all our results could be included in this first and very preliminary draft of the paper. Among the results in this paper, we found that traditional antitrust law enforcement, fines following successful investigations of the competition authority and no leniency, has a deterrence effect (reduces the number of cartels formed) but also has a pro-collusive effect (increases collusive prices). Leniency programs might not be more efficient than standard antitrust enforcement, since they do deter a significantly higher fraction of cartels from forming, but they also induce higher prices in cartels that are not reported. With rewards for whistle blowing, instead, cartels are systematically reported, disrupting completely subjects' ability to form cartels and to sustain high prices. Also, we find that when the reward scheme is 'wrongly designed' in the sense that can be exploited, in our case by completely eliminating the risk of being fined at no cost, subjects do not recognize the possibility to manipulate and gain from the scheme, a result in line with recent experiments in other fields (see e.g. Dal Bo, 2005 [29]). If the ringleader is excluded from the leniency program, as under the US leniency policy, the deterrence effect of leniency falls and prices are higher than otherwise. We also analyze tacit collusion, and find that under standard antitrust enforcement or leniency programs, subjects who do not communicate (do not go for explicit cartels) choose significantly higher prices than where there is no anti-trust enforcement whatsoever. This is not the case anymore when reward schemes are introduced.

The remainder of this paper is organized as follows. The next section discusses related literature, theoretical and experimental. Section 3

[7]We are aware of two other studies that deal with not exactly the same issues but are somewhat related. Hamaguchi and Kawagoe (2005) [43] design an experiment where subjects are forced to collude. Most obviously, such a setup cannot address the issue of how different policies perform in terms of cartel deterrence. Hamaguchi et al. (2005) adapt the setup of Hinloopen and Soetevent but to a repeated procurement auction with leniency programs. They consider a different game since there is only one winner at each period, so when players are colluding, they have to decide who will win the auction. They found evidence of deterrence effects with Leniency programs as well as higher prices under leniency and antitrust than under communication.

describes the underlying theoretical model and the experimental design, contrasting it with previous ones. Section 4 presents our results. Section 5 concludes, discussing implications for the theory and practice of designing deterrence mechanisms for cartels and similar forms of organized crime. An appendix contains the instructions for the experiment.

7.2 Literature review

7.2.1 Theory

Starting with the contributions of Motta and Polo (2003) [76], Rey (2003) [85], and Spagnolo (2000a [96],b [97]), a theoretical literature has blossomed in the last decade that analyzes the optimal design of anti-cartel policies based on the provision of incentives to breach trust and to self-report.[8] Different effects of leniency and rewards are considered in this literature. The focus here is on the deterrence effects of the first part of the leniency policies, restricted to firms that self-report before an investigation by the competition authority has invested them. The most important effects identified by the literature in this respect are:

1. *The protection from fines effect.* Spagnolo (2000a [96], 2004 [98]) and Rey (2003) [85] suggest that amnesty offered to the first firm reporting before an investigation is opened may have deterrence effects by ensuring that if a cartel member wants to undercut the cartel, it can report and avoid paying the fine.[9]

2. *The reward effect.* Spagnolo (2000a [96], 2004 [98]), Buccirossi and Spagnolo (2001 [14], 2006 [15]), Rey (2003) [85] and Aubert et al. (2006) [4] suggest that rewards could further increase deterrence by generating stronger temptations to undercut the cartel and cash the reward by reporting. Spagnolo (2000a [96], 2004 [98]) shows

[8]Other early pieces include Aubert et al. (2006) [4], Buccirossi and Spagnolo (2001 [14], 2006 [15]), Ellis and Wilson (2001), Harrington (2007) [44] and Harrington and Chen (2007) [23]. See Spagnolo (2008) [99] for a review of this growing theoretical literature.

[9]More recently Harrington (forthcoming) [45] coined a perhaps nicer acronym for the same effect, *deviator amnesty effect*, but here we stick to temporal priority.

that such a mechanism can for the first time deliver the first best in a model a la Becker (1968) [9], complete deterrence without investigation costs, provided that fines are sufficiently but finitely large, and that the reward is lower than total fines.

3. *The 'reporting as a threat' and 'what does not kill us makes us stronger' effects.* Spagnolo (2000b) [97] and Buccirossi and Spagnolo (2001 [14], 2006 [15]) show that when self-reporting becomes attractive thanks to leniency programs, the *threat* of self-reporting to punish an agent that did not behave as the cartel agreed upon may also become credible, and may be exploited to enforce cartels that would not be sustainable otherwise. Building on this idea, Ellis and Wilson (2001) obtain a related effect, showing that, for cartels that are not deterred, leniency programs have the effect of reinforce/stabilize collusion. The reason is that if a cartel is formed, then leniency induces cartel members to self-report after any defection from agreed collusive strategies, thereby strengthening the punishment for defections of an amount equal to antitrust fines.

4. *Tacit collusion and post-conviction pricing.* Antitrust doctrine agreed in the 50s that the focus should be restricted to 'explicit cartels', i.e. to conspiracies where firm managers meet or communicate with the explicit objective of coordinating on higher prices, and leave alone tacit collusion, i.e. cases where firms manage to coordinate on and sustain high prices without explicit communication. Whinston (2006) [106] reopened the debate, arguing that what is important for welfare are prices, so that we should reflect more on how antitrust enforcement may affect firms' ability to sustain prices, even when high prices are sustained by *tacit* collusion. On a different but related stance, Buccirossi and Spagnolo (2007) [16] suggest that antitrust fines might have the effect of inducing firms to increase collusive prices following conviction, either because they do not realize they are a 'sunk cost' and try to recover them through higher margins, or because paid fines may help firms coordinating on higher post-conviction prices sustained by tacit collusion.

We discuss such effects when we present the experimental results.

7.2.2 Experiments

Apesteguja, Dufwemberg and Selten (2006) [2], and Hinloopen and Soetevent (2006) [47] are the first to analyse experimentally the effects of leniency policies on cartel deterrence.[10]

Apesteguja, Dufwemberg and Selten (2006)

Apesteguja et al. (2006) [2] conducted the first experimental investigation of the effects of Leniency policies and rewards schemes on cartel deterrence. This elegant paper first develops a stylized but static theoretical framework that tries to capture the main points made in the theoretical literature mentioned earlier on the general deterrence effects of leniency policies, and then uses it to undertake an experimental analysis of these effects. The market game they focus on is a one-shot homogeneous discrete Bertrand oligopoly. This is embedded in four alternative legal frameworks: in *Ideal* there is no antitrust law, cartels are not possible (communication is not allowed), and colluding firms face neither full nor reduced fines; in *Standard* convicted firms face fines equal to 10% of their revenue and no reduction if they report; in *Leniency* firms that report a cartel they took part in receive a reduction in their fine; in *Bonus* reporting firms receive part of the fines paid by other firms as a reward. Strategically equivalent collusive subgame perfect equilibria, including one implementing the monopoly price, exist in both *Standard* and *Leniency*, sustained by the threat of reporting if a defection takes place. The reason is that if a firm defects in an homogeneous Bertrand game, its opponent will have no revenue, so even if there is no leniency self-reporting is costless for a party whose opponent defected and is

[10] We are also aware of a third experimental paper on leniency programs, by Hamaguchi and Kawagoe (2005) [43]. However, in their experimental design subjects are forced to collude, to look then at the effects of leniency on the likelihood of cartels' disruption. Such a setup focuses only on desistance, and has therefore nothing to say about relative efficiency of different policies in terms of general deterrence, i.e. of prevention of voluntary cartel formation, which is the object of the present study.

therefore a credible threat that can sustain collusion in the one-shot game.

The experimental setting allows for pre-play communication and let subjects play in groups of three, and for the rest it follows closely the theoretical model just described.

The experimental results confirm that agents understand and use the threat of reporting to sustain collusion in the one-shot Bertrand game: prices are substantially higher in *Standard* and in *Leniency*, where collusive equilibria exist sustained by the threat to self-report if cheated upon, than in *Ideal* where no such threat is available and the only equilibrium is the Bertrand one. *Leniency* has a significant deterrence effect relative to Standard, although prices are much higher than in *Ideal*, without any antitrust. Surprisingly, the experimental results are inconsistent with the theoretical predictions that rewarding reporting firms should reduce cartel formation: the *Bonus* treatment has non-significant effects on collusion.

As also argued by the authors, this paper can be seen as a first exploratory step in the experimental analysis of cartel deterrence mechanism. The reason is that both theory and experiment make a number of simplifications that may affect the result in a non trivial way.

First, the game and experiment allow for only one round of decisions, leaving experimental subjects no way to learn. This may be a problem for the interpretation of the experimental results. The equilibria agents are choosing among in *Standard* and *Leniency*, and the difference with *Bonus* are not that easy to understand. Most recent experimental studies of one-shot games allow for some repetition precisely because it is well known from earlier work on public goods games that the first decisions are typically mistaken. In fact, it is possible that the surprising result on the ineffectiveness of *Bonus* is driven by subjects not fully grasping the situation.

In our experiment we try to improve on this point by having both a repeated game, and five initial rounds for subjects to experiment the game.

Second, the theoretical framework used for the experiment resembles closely that in Spagnolo (2000b) [97], but for fines chosen equal 10% of firms' revenue, so that reporting if a partner-cartelist undercuts becomes a credible threat even without leniency programs in place (in *Standard*). However, the 10% revenue cap for EU fines that inspired the 10% of revenue fines is relative to firms' total yearly turnover in the last period the cartel was active. In an appropriately dynamic framework, therefore, fines would never be zero because of a defection. Moreover, it is hard to imagine a market where, if a firm undercuts the cartel, other firms have zero revenue for one full year. Firms are active in many markets and total business stealing appears impossible in reality, so that absent leniency policies, a firm that reports a cartel would be subject to a positive fine, the multiplicity of equilibria in *Standard* would disappear as after a defection reporting is dominated by not doing it (and avoiding the fine), and *Leniency* may then fare much worse than how depicted. Relatedly, homogeneous good Bertrand competition is a degenerate case of price competition with differentiated products, and the collusive equilibria in Standard would disappear with a little product differentiation.

To improve on these points, in our experiment we chose an infinitely repeated differentiated product Bertrand game, and fixed fines rather than fines that go to zero for some price choices.

Hinloopen and Soetevent (2006)

Hinloopen and Soetevent (2006) [47] study experimentally both the general deterrence and desistance effects (shortening non-deterred cartels' duration) of Leniency Programs. They use an infinite horizon set up that allows for communication before prices are chosen each period, and where the stage game is the same homogeneous Bertrand game used by Apesteguja et al. (2006) [2]. Subjects are matched in groups of three at the beginning of each treatment, and then play without re-matchings for at least 20 rounds, after which the continuation probability falls to 80%.

They embedded these oligopoly games in four different treatments: *Benchmark*, where subjects cannot communicate; *Communication*, where subjects can communicate before choosing prices; *Antitrust*, where subjects that communicated are exposed to a positive probability of being detected and fined; and *Leniency*, which differs from *Antitrust* by the possibility to self-report after the choice of price and before the random audit by the competition authority. In *Leniency*, therefore, subjects can only self-report after prices have been chosen and made public, so that subjects cannot both secretly report and secretly undercut the collusive price, as is possible in reality where competition authorities may keep the report secret to arrange for dawn raids allowing (or even asking to) the reporting firm the possibility to secretly undercut former cartel partners. Although they also use fines equal to 10% of revenue in the period of conviction, Hinloopen and Soetevent include a fixed cost of reporting under the leniency program that destroys Apesteguja et al.'s (2006) [2] one-shot collusive equilibria sustained by the threat of reporting after defections.

These authors' main findings are that leniency: (i) increases cartel deterrence (fewer cartels are formed); (ii) reduces the duration of cartels that are not deterred (agents that form a cartel defect more afterwards); and (iii) make agents defect more aggressively than in the absence of the Leniency Program. They do not find that the leniency program affects the likelihood that a detected and fined cartel forms again thereafter (no effects on recidivism).

Although Hinloopen and Soetevent (2006) [47] is the experimental work which is closest to what we do, our experiments differ from theirs in several respects.

First, in each stage-game we will allow subjects to both self-report and set prices before any of these choices is observed by other subjects. In our experiment, therefore, subjects have the possibility to simultaneously secretly report and defect/undercut cartel partners, much like in reality, and then they will have also the possibility to self-report after observing price choices, if nobody reported before price became public. We consider this a major improvement towards the realism of the experimental set up that allows agents to defect from the cartel and avoid fines, as possible in reality, and us to disentangle and quantify reports linked to defections (protection from fines effect) and reports that are made to punish defections from the cartel ('reporting as a threat' or 'what does not kill us makes us stronger').

Second, we will have fixed fines rather than Hinloopen and Soetevent's fines equal to 10% of the revenue of the period in which a cartel is detected, because in this second case fines vary a lot with the outcome of the stage-game in which a cartel happens to be detected, so that it is not clear what the expected fine perceived by subjects actually is, which makes it impossible to cleanly analyze the role of fines and their interaction with leniency.

Third we follow Apesteguja et al. (2006) [2] in framing the experiment explicitly as a cartel/antitrust game, rather than having a "neutral" frame as in Hinloopen and Soetevent, as we want to make sure that subjects do not misunderstand the situation, and we want to minimize the possible impact of social preferences on subjects' choices.

Fourth, we use a perhaps more realistic oligopoly model, a repeated differentiated product Bertrand game.

Fifth our subjects are rematched with positive and constant probability all along the treatments, so that each supergame has a constant continuation probability, like in Dal Bó (2005) [29], Dal Bó and Frechette (2007) [30], and Blonski et al. (2007) [10].

Sixth, we use duopolies rather than triopolies, to avoid that agents may be unwilling to punish defections too hard by the unwillingness to harm a third 'innocent' (non-defecting) party, as suggested by Holt (1995) [50].

In addition, and again differently from Hinlopen and Soetevent, we consider other important effects:

- we look at the effects of rewards for whistleblowers; indeed, verifying their deterrence effects in an appropriately dynamic environment to test the robustness of Apesteguja et al.'s (2006) [2] surprising mixed finding is one of the central research questions motivating our study;

7.3 Experimental Design

Our experimental design is most closely related to the one by Hinloopen and Soetevent, but we introduce a number of crucial modifications. Our innovations are mainly relative to the timing when subjects can self-report, the fines' structure, the oligopoly game, and the framing of the experiment. We also consider two extra treatments, REWARD and LENRING, which will be discussed in detail below when we describe our experiment.

In our experiment, each subject represented a firm and played in anonymous two-persons groups a repeated duopoly game. In every stage game, the subjects had to take three types of decisions. First, they had to choose a price in a discrete Bertrand price game with differentiated goods. Second, they had to decide whether or not to form a cartel by discussing prices. Third, the subjects could choose to self report cartels to a competition authority. The attractiveness of this latter opportunity depended on the details of the antitrust law enforcement institution - the treatment variables of our experiment.

7.3.1 The Bertrand game

In each period, the subjects had to choose a price from the choice set $\{0, 1, ..., 11, 12\}$. The resulting profits depended on their own price choice and on the price chosen by their competitor and were reported in a profit table distributed to the subjects (see Figure 7.1). This table was derived from the following standard linear Bertrand game. (The details of the Bertrand game were not described to the subjects.)

The demand function for each firm i was given by:

$$q_i(p_i, p_j) = \frac{a}{1+\gamma} - \frac{1}{1-\gamma^2}p_i + \frac{\gamma}{1-\gamma^2}p_j$$

where p_i (p_j) is the price chosen by firm i (competitor j), a is a parameter accounting for the market size and $\gamma \in [0, 1)$ denotes the degree of substitutability between the two firms' products. Each firm faced a constant marginal cost, c, and had no fixed costs. The profit function, $\pi_i(p_i, p_j)$, was thus given by

$$\pi_i(p_i, p_j) = (p_i - c)q_i.$$

In our experimental setup, we chose $a = 36$, $c = 0$ and $\gamma = 4/5$ and restricted the subjects' choice set to $\{0, 2, ..., 22, 24\}$. These parameters

		your competitor's price												
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
your price	**0**	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	29	38	47	56	64	68	68	68	68	68	68	68	68
	2	36	53	71	89	107	124	128	128	128	128	128	128	128
	3	20	47	73	100	127	153	180	180	180	180	180	180	180
	4	0	18	53	89	124	160	196	224	224	224	224	224	224
	5	0	0	11	56	100	144	189	233	260	260	260	260	260
	6	0	0	0	0	53	107	160	213	267	288	288	288	288
	7	0	0	0	0	0	47	109	171	233	296	308	308	308
	8	0	0	0	0	0	0	36	107	178	249	320	320	320
	9	0	0	0	0	0	0	0	20	100	180	260	324	324
	10	0	0	0	0	0	0	0	0	0	89	178	267	320
	11	0	0	0	0	0	0	0	0	0	0	73	171	269
	12	0	0	0	0	0	0	0	0	0	0	0	53	160

Figure 7.1: *Payoff table*

yield the payoff table distributed to each subject. To simplify the table we also relabeled each price by dividing it by 2 and rounded the payoffs to the closest integer. In the unique Bertrand equilibrium, both firms charge a price equal to 3 yielding per firm profits of 100. The monopoly price (charged by both firms) is 9, yielding profits of 180. Note also that a firm would earn 296 by unilaterally and optimally undercutting the monopoly price, i.e. by charging a price of 7. In this case the other (cheated upon) firm only earns a profit of 20. Similarly, there are gains from deviating unilaterally from other common prices than the monopoly price as well as associated losses for the cheated upon firm; in the range of prices in between the Bertrand price and the monopoly price, i.e. in the range $\{4, ..., 8\}$, these gains and losses are smaller than when a subject deviates unilaterally from the monopoly price.

7.3.2 Cartel formation

Throughout the experiment, the subjects could form cartels by discussing prices. At the beginning of every period, a communication window opened if and only if both subjects agreed to communicate. This communication stage, which is described in more detail below, was designed in such a way that it would result in a common price on which to cooperate. This agreed upon price was non-binding, however, and therefore each subject could cheat on the agreement by subsequently charging a price different from the agreed upon price.

Whenever two subjects chose to communicate, they were considered to have formed a cartel. In this case, the subjects risked to be fined as long as the competition authority had not yet detected the cartel. This implied that two subjects could be fined in a period even if no communication took place in that specific period; for example, two subjects could be fined in a period in which they did not communicate if they had communicated in the previous period and the competition authority had not detected the associated cartel in that period. Once a cartel was detected, however, it was considered to be dismantled and in subsequent periods, the former cartelists did not run any risk of being fined unless they communicated again.

7.3.3 Antitrust law enforcement (Treatments)

Whenever two subjects had formed a cartel, a competition authority could detect the cartel and convict its members for price fixing. Detection could happen in two ways. First, in every period, the competition authority detected cartels with an exogenous probability, α. If this happened, both cartel members had to pay an exogenous fine, F. Second, the cartel members could self-report the cartel, in which case the cartel members were convicted for price fixing with certainty. If this happened, the size of the fine depended on the details of the law enforcement institution.

We ran five types of treatments and we adopted a *between subjects* design, so that every subject only played the game under a single treatment. Each treatment corresponded to a specific type of antitrust law, that is our treatment variables were the different law enforcement institutions. The differences between the treatments are summarized in table 7.1.

Our baseline treatment corresponds to a *laisser faire* regime and is denoted COMMUNICATION: in this treatment, $\alpha = F = 0$ so that forming a cartel by discussing prices is legal. To simplify the instructions and to eliminate irrelevant alternatives, subjects were not allowed to report cartels. In the four other treatments, denoted ANTITRUST, LENIENCY, LENRING and REWARD, the expected fine (given that no reports took place) was strictly positive ($\alpha = 0.1$ and $F = 200$ yielding an expected fine $\alpha F = 20$) and cartel members were allowed to report cartels in which they participated. The ANTITRUST treatment corresponds to traditional antitrust laws without any leniency program: in case a report took place, both cartel members (including the reporting one) had to pay the full fine F. The LENIENCY treatment corresponds to current antitrust laws embedded with a leniency program: in case the cartel was reported by one of the cartel members only, the reporting member paid no fine while the other one paid the full fine, F; if instead both cartel members reported the cartel simultaneously, both paid a reduced fine equal to $F/2$. The treatment LENRING was identical to LENIENCY except that the first subject attempting to communicate was treated as the cartel's initiator - the so-called ringleader - and, as a result, was not eligible for the leniency program. (The way the ringleader was identified is described in more detail below). Finally, the REWARD treatment differed from LENIENCY in one respect only: if only one cartel member reported the cartel, his/her fine was not only reduced to 0; in addition, he was rewarded with the full fine, F, paid by the other cartel member.

In addition to these five treatments, we also ran a number of other treatments to check the robustness of our results to changes in α and F. First, we ran two additional antitrust and leniency treatments with higher expected fines equal to 60 ($\alpha = 0.2$ and $F = 300$). These treatments were denoted ANTIHIGH and LENHIGH respectively. Second, we ran two additional reward treatments, both with an expected fine equal to 0

Table 7.1: *Treatments*

Treatment	**fine (F)**	**prob. of detection (α)**	**report**	**report's effects**
COMMUNICATION	0	0	No	–
ANTITRUST	200	0.10	Yes	pay the full fine
LENIENCY	200	0.10	Yes	no fine (half the fine if both report)
LENRING	200	0.10	Yes*	no fine
REWARD	200	0.10	Yes	reward (half the fine if both report)

*Only for the player who's the last to decide to communicate.

($\alpha = 0$) but with different fines. The treatment denoted REWLOW had a relatively low fine ($F = 200$) while REWHIGH had a high fine ($F = 1000$).

7.3.4 The experiment's timing and the rematching procedure

At the end of each period, subjects were rematched with the same competitor with a probability of 85%. With the remaining probability of 15%, all subjects were randomly matched into new pairs. When this happened, the history in the previous match did no longer matter; for example, a subject could no longer be fined for a cartel formed in a previous match. The subjects were also informed that the experiment would end if more than 20 periods had passed and the 15% probability event took place or if the experiment lasted for more than 2 hours and 30 minutes. This latter possibility was so unlikely that it never happened.

This re-matching procedure had several advantages. First, the subjects were playing truly *infinitely repeated games* without problems associated with end effects. Second, each subject played several repeated games against different competitors. Thereby we observed the subjects' behavior in a larger number of repeated games.

Before the experiment started, the subjects were paired with the same competitor for five practice periods. During these practice periods, subjects were assigned to different competitors than those that they faced in the first period of the 'true' (i.e. remunerated) experiment. Participants were informed about this.

7.3.5 The timing of the stage game

With the exception of the COMMUNICATION treatment, a stage game consisted of 7 steps. In COMMUNICATION, steps 4,5 and 6 were skipped. An overview of the steps is given in Figure 7.2.

Step 1: Communication decision. Each subject was asked whether or not he wished to communicate with his competitor. If both subjects pushed the yes button within 15 seconds, the game proceeded to step 2. Otherwise the two subjects had to wait for additional 30 seconds before

Figure 7.2: *Timing of the stage game*

pricing decisions were taken in Step 3. In all periods, subjects were also informed whether they were matched with the same opponent as in the previous round or if a re-match had taken place.

In the treatment LENRING, the first subject to push the button within the time window of 15 seconds was treated as the ringleader. If instead only one of the subjects pushed the yes button, then this subject was treated as a ringleader even if the cartel was formed in later periods. In either case, both subjects were informed at the end of Step 1 about the identity of the (possibly only potential) ringleader.

Step 2: Communication. If both subjects decided to communicate in step 1, a window appeared on their computer screen asking them to simultaneously state a minimum acceptable price in the range $\{0, ..., 12\}$. When both of them had chosen a price, they entered a second round of price negotiations, in which they could choose a price from the new range $\{p_{min}, ..., 12\}$, where p_{min} was defined as the minimum among the two prices selected in the previous negotiation round. This procedure went on until 30 seconds had passed. The resulting minimum price p_{min} was referred to as the agreed upon price.

Step 3: Pricing. Each subject had to choose his price from the choice set $\{0, ..., 12\}$. Possible price agreements reached in step 2 were not binding. The subjects were informed that if they failed to choose a price within 30 seconds, then their default price would be so high that their profits became 0.

Step 4: First Reporting Decision. If communication took place in the current period or in one of the previous periods and had not yet been discovered by the competition authority, subjects had a first opportunity to report the cartel.

Step 5: Market prices and second reporting decision. Subjects learn the prices set by their opponent. If communication took place in the current period or in one of the previous periods and was not yet discovered by the competition authority and nobody has reported it in step 4, subjects have again the opportunity to report the cartel. The crucial difference between this second reporting opportunity and the first one is that the subjects knew the price chosen by the competitor. In addition the subjects were informed about their own profits and the profits of their competitor, gross of the possible fine.

Step 6: Detection. If communication took place in the current period or in one of the previous periods and had not yet been discovered or reported in steps 4 or 5, the competition authority discovered the cartel with probability α.

Step 7: Summary of the current period. At the end of each period, all the relevant information about the stage game was displayed: agreed upon price (if any), prices chosen by the two players, possible fines and net profits. In case players were fined, they were also told how many players reported. This step lasted 20 seconds

Note that with our experimental setup subjects have two opportunities to report the cartel: first at step 4, right after having set their price, then again at step 5, after having been informed about the price chosen by their opponent. In our design, reporting can thus be used for two different purposes: (i) deviating subjects may report to get protection against prosecution and (ii) cheated upon subjects may report to punish their opponents, if they have not reported before.

7.3.6 Experimental procedure

Our experiment took place in March, April and May 2007 at the Stockholm School of Economics (Sweden) and at Tor Vergata university (Rome, Italy). Some additional sessions were run in November 2007 in Stockhlm and in December 2007, in Rome. Session lasted on average 2 hours, including instructions and payment. The average payment was: (i) in Stockholm SEK 239[11], with a minimum of 112.5 SEK and a maximum of

[11] At the time of the experiment, 1 SEK=0.109 Euro

387 SEK and (ii) in Rome Euros 22.65 with a minimum of 11.5 Euros and a maximum of 31.5 Euros, including a show up fee equal to 50 SEK in Sweden, and to 7 Euros in Italy. We ran a treatment for every session; the number of subjects per session ranged from 16 to 32, and the total number of subjects was 444. No subject took part in more than one session.

Subjects were welcomed in the lab and seated, each in front of a computer. When all subjects were ready, a printed version of the instructions and the profit table was distributed to them. Instructions were read aloud to ensure common knowledge of the rules of the game. The subjects were then asked to read the instructions on their own and ask questions, which were answered privately. When everybody had read the instructions and there were no more questions (which always happened after about fifteen minutes), each subject was randomly matched with another subject for the five practice rounds. After the practice rounds, participants had again a last opportunity to ask questions about the rules of the game. Again, they were answered privately. Then they were randomly rematched into new pairs and the real play started.

At the end of each session, the subjects were paid privately in cash. The subjects started with an initial endowment of 1000 points in order to reduce the likelihood of bankruptcy. At the end of the experiment the subjects were paid an amount equal to their cumulated earnings (including the initial endowment) plus a show up fee of 50 SEK in Stockholm and 7 Euros in Rome. The conversion rates were 20 points for 1 SEK in Stockholm and 200 points for 1 Euro in Rome.

The experiment was programmed and conducted with the software z-Tree (Fischbacher 2007)[35].

7.3.7 Equilibrium set

In the games presented above, the equilibrium structure can be described in the following way. For each treatment and each price that the firms want to collude on, there exists a critical discount factor such that the firms can collude on the desired price level if and only if the firms' discount factor is larger than the critical discount factor. While it is not trivial to find these critical discount factors for each treatment, it is possible to rank them. Let $\delta_{Communication}$, $\delta_{Antitrust}$, $\delta_{Leniency}$, $\delta_{LenRing}$ and δ_{Reward} denote the critical discount factors for the, COMMUNICATION, ANTITRUST, LENIENCY, LENRING and REWARD treatments respectively. Provided that the probability of detection, α, and the size of the fine, F, are equal in all treatments, then it can be shown that

$$0 = \delta_{Antitrust} < \delta_{Communiction} < \delta_{Len} = \delta_{LenRing} < \delta_{Reward} < 1.$$

The only surprising feature of this ranking is that collusion can be sustained for any discount factor in the ANTITRUST treatment ($\delta_{Anti} = 0$). The reason is simple: in the stage game, it is a subgame perfect equilibrium to collude. Indeed, if *both* firms' strategies stipulate that one should report the cartel whenever a firm unilaterally deviates from the collusive price, then it is no longer profitable to deviate due to the reports. Furthermore, the reports are credible: since both firms (including the deviating one) report the cartel following a deviation, both firms are indifferent between reporting and not reporting, and thus reporting is an equilibrium in the reporting subgame. Of course, the weakness of this subgame perfect equilibrium is that it is sustained through weakly dominated strategies. When the stage game is infinitely repeated, however, it is easy to construct strategies with the same flavor, which are not weakly dominated.

The key to the above observation is that in the ANTITRUST treatment, reports can be used as punishments against deviators. This is not the case in the other treatments. In the COMMUNICATION treatment, it is trivial that reports cannot be used. To see that reports cannot be used as punishments in the remaining treatments, note that optimal deviations involve secret reports. Thus cartels are dismantled after unilateral deviations and therefore reports cannot be used as punishments. For this reason, the ranking of the remaining critical discount factor is the expected one.

7.3.8 Empirical Methodology

A critical point in our analysis is how to control for repeated observations of the same subject or the same duopoly, when testing the significance of the observed differences across treatments. Before explaining more in detail the procedure we adopted, it is useful to introduce here some terminology: we call "individual-level" data the data representing individual decisions of the subjects, e.g. the decision to communicate or not in a given period or the decision to unilaterally deviate from a collusive agreement; we call instead "duopoly-level" data the data that refer to variables that always have the same value for the two members of a duopoly. Thus, the presence of a cartel within a duopoly in a given period, or the fact that a given cartel is detected by the antitrust authority, are duopoly-level data.

Given the structure of our game, we need to account for correlation between two observations from the same individual, as well as correlation between two observations from different individuals who belong to the same duopoly. Moreover, since we have run the experiment in two different cities, we also have to control for the possible correlation among observations collected in the same city. To this purpose, we adopted multilevel random effect models.

Since in our experiment a subject may take part in more than one duopoly during the game, the random effects at the subject level and at the duopoly level are not nested, which makes it difficult to estimate a model with a random effect at the duopoly level and a random effect at the subject level at the same time.
To overcome this complicacy, when analyzing individual-level data, we hypothesized the presence of a random effect for every subject within any particular match (which accounts for the correlation among observations pertaining to the same match), nested with a random effect for every subject across different matches, which is in turn nested with a random effect at the city level.

To analyze duopoly-level data we make the assumption that correlation between observations belonging to the same subject but to different duopolies can be disregarded. We therefore hypothesize to have only a random effect at the duopoly level, nested with a random effect at the city level.

The only independent variable of our simple regressions is the treatment, as a dummy. To analyze individual-level data, we adopt a four-levels model of the following form:

$$y_{hijk} = \beta_0 + \beta_1 TREAT_{hijk} + \eta^{(2)}_{ijk} + \eta^{(3)}_{jk} + \eta^{(4)}_{k}$$

where *h*, *i* , *j* and *k* are indices for measurement occasions, subjects in matches, subjects across matches and cities, respectively. *TREAT* is the dummy variable for the treatment. Since we always compare only two treatments at a time, this variable takes value 1 in correspondence of one of the two treatments, and value 0 in correspondence of the other one. $\eta^{(2)}_{ijk}$ represents the random intercept for subject *j* in match *i*, and in city *k* (second level), $\eta^{(3)}_{jk}$ represents the random intercept for subject *j* in city *k* (third level) and $\eta^{(4)}_{k}$ represents the random intercept for city *k* (fourth level). Random intercepts are assumed to be independently normally distributed, with a variance that is estimated through our regression.

The general three-levels model we adopt when looking at duopoly-level data has the following form:

$$y_{hlk} = \beta_0 + \beta_1 TREAT_{hlk} + \eta^{(2)}_{lk} + \eta^{(3)}_{k}$$

As above, *h* and *k* are indices for measurement occasions and cities, while *l* is the index for duopolies. $\eta^{(2)}_{lk}$ and $\eta^{(3)}_{k}$ represent random intercepts at the duopoly and city levels.

When comparing observations collected in a single city, we adopt a model which is analogous to the previously described ones, but without the last level.

We ran logit regressions to analyze the decision to communicate, the decision to deviate, and the rates of cartel formation and of cartel detection; we adopted instead linear regressions for prices and agreed upon prices. To estimate our model we used an ordinary panel regression with random effect, when the number of considered levels was equal to 2, while we used GLLAMM (see Rabe-Hesketh and Skrondal, 2004 and http://www.gllamm.org) when the number of considered levels was equal or higher than three.

7.4 Results

7.4.1 Traditional and modern law enforcement

In this section we report the subjects' behavior in the COMMUNICATION, ANTITRUST and LENIENCY treatments. The purpose is to assess how traditional antitrust law (ANTITRUST) and more modern law enforcement institutions embedded with a leniency program (LENIENCY) perform relative to a *laisser faire* regime (COMMUNICATION). Our primary interest is to document how these different policies perform in terms of ex ante deterrence and their implications for the subjects' price choices. In addition we also report post conviction deterrence and prices, that is whether cartelists, after having been convicted, are deterred from reforming the dismantled cartel. We postpone our analysis of the LENRING and REWARD treatments to two subsequent sections.

Cartel Deterrence

Table 7.2 reports the rates of communication attempts and of cartel formation provided that subjects are not currently cartel members. These rates are our main measures for evaluating the success of the different policies in terms of ex ante deterrence, that is the main objective of Antitrust policies.

Result 1 (Ex ante deterrence) [12]Traditional antitrust laws (ANTITRUST)

[12]As explained above, we used a multilevel random intercept model to compare the results across treatments. We ran a regression per each couple of treatments.
For the **Rate of communication** the single observation is the binary decision to communicate of a subject in every single period. We test the significance of the differences between treatments by modeling the binary outcome by a four-level random intercept logistic regression, since here we analyze individual-level data.
Similarly, for the **Rate of individual deviation** the single observation is represented by the individual decision to deviate from the last collusive agreement. Here we consider only the

are effective in deterring explicit cartel formation and modern antitrust laws (LENIENCY) even more so.

Result 1 stems from the fact that the rates of communication attempts and of cartel formation are significantly lower in ANTITRUST than in COMMUNICATION. Moreover LENIENCY was even more successful in terms of ex ante deterrence since the rates of communication attempts and of cartel formation were significantly lower in LENIENCY than in ANTITRUST. Relative to COMMUNICATION, the rates of individual communication attempts decreased by 31% in ANTITRUST and by 56% in LENIENCY. These differences were even more striking for the rates of cartel formation, with a 55% and 77% decrease in ANTITRUST and LENIENCY respectively.. This is line with Miller's (2007) estimate that leniency may be associated with a 52 percent decrease in the rate at which cartels form. These results are also (partly) consistent with previous experimental results. Apesteguja et al. do find a reduction in the percentage of formed cartels (from 67% to 50%) when Leniency is introduced (compared to the case when firms can report but there is no reduced fine). Hinloopen and Soetevent (2006) [47] find a similar pattern as we do concerning the rate of cartel formation although they observe no significant differences between their antitrust and leniency treatments with respect to the rate of communication attempts.[13]

Table 7.2 also reports the rates of detection due to self reporting by subjects - a first source of cartel instability. The rate of reporting is

cases in which the subjects had previously formed an agreement on prices. Again, these are individual-level data, so to evaluate the significance of the observed differences we adopted a four-level random intercept logistic regression.

For the **Rate of cartel formation** we have a single observation per duopoly per period, which indicates if in that period a cartel has been formed within that duopoly. We consider only the cases in which no cartel pre-existed. The analysis here concerns duopoly-level data, the binary response is therefore modeled by a three level random intercept logistic regression.

For the **Rate of reporting** we consider only the cases in which a cartel exists. We have a single observation per duopoly per period, which indicates whether the cartel has been discovered in that period *because at least one of the two cartel members reported it to the antitrust authority*. As for the rate of cartel formation, we adopted a three level random intercept logistic regression, because the analysis concerns duopoly-level data.

[13]Our results are not perfectly comparable with those of Hinloopen and Soetevent because they report the rate of communication for all periods while we report the rate of communication provided that no cartel has been formed previously.

Table 7.2: *Deterrence effects*

	Communication		Antitrust		Leniency
Rate of communication attempts	0.835	$>^{**}$	0.566	$>^{***}$	0.377
Rate of cartel formation	0.716	$>^{***}$	0.315	$>^{***}$	0.178
Rate of individual deviation	0.564	$>^{***}$	0.424	$\approx$	0.373
Rate of reporting	–	–	0.092	$<^{***}$	0.507

Note: In this table and in the following ones, the symbols ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

small in ANTITRUST while it increases substantially and significantly in LENIENCY. Hence:

Result 2 (Cartel stability and self-reporting) Modern antitrust laws (LENIENCY) reduce cartel stability due to self reporting.

Result 2 is in line with Miller's (2007) conclusion that leniency programs are associated with a 62% increase in the rate of detection, even though we observe an even higher increase. This is also consistent with Aspeteguja et al. who find an increase of 50% of the detection rate.

Our experimental design also allows us to distinguish between different motives behind reporting behavior. As already mentioned, subjects can either report in order to protect themselves against fines using the first reporting opportunity or they can report and punish their competitor after having observed the competitor's price choice. Clearly, we should expect to observe the former type of reports in LENIENCY only. By contrast, the latter opportunity to report in order to punish deviators may be observed both in ANTITRUST and LENIENCY, although one may argue that reports to punish should be rare in both treatments. In ANTITRUST, subjects may find it too costly to report and in LENIENCY, an optimal deviation involves a simultaneous secret report, implying that a cheated upon subject should not be able to punish by reporting.

Table 7.3 reports the number of first and second reports in ANTITRUST and LENIENCY. There are almost no first reports in ANTITRUST (as expected) and only few second reports. It is interesting to note, however, that the few subjects who used the second report did so systematically in order to punish an undercutting rival. Thus some subjects were willing to take a quite large cost in order to punish deviators. Whether this reflected that these few subjects used optimal punishments or altruistic punishment as described by Fehr and Gächter (2002) [34] is an open question.[14]

[14] Fehr and Gächter (2002) [34] analyze a repeated one shot public good game experiment and argue that subjects are willing to bear the cost of punishing free riders. They explain that "Free riding may cause strong negative emotions among the cooperators and these emotions, in turn, may trigger their willingness to punish the free riders". Their experimental evidence gives stronger support to the hypothesis of altruistic punishments.

Table 7.3: Self reporting

	Treatment	Antitrust1	Leniency1
# of reports (% of possibilities to report)		60 (4.7%)	195 (35.6%)
# of First Reports	In total	4	168
	Simultaneous deviation	1	140
# of Second Reports	In total	56	27
	At least one deviated	54	27
	Only rival deviated	46	13
	Rival deviated more	7	1

In LENIENCY, most reports took place during the first secret reporting stage and were combined with a simultaneous undercutting of the agreed upon price. This is consistent with optimal deviations and the ”protection gains fines motive”. Still there were a non-negligible number of first reports where the subjects did not simultaneously deviate. Furthermore, there were also some second reports and these were typically carried out as a punishment against an undercutting rival who for some reason did not simultaneously report the cartel during the first reporting step.

Finally, Table 7.2 reports the rates of deviation from agreed upon prices - a second source of cartel instability.

Result 3 (Cartel stability and price deviations) Both modern antitrust laws (LENIENCY) and traditional ones (ANTITRUST) increase cartel stability by reducing the rate of deviations from agreed upon prices.

Result 3 stems from the fact that the rate of deviation in both the ANTITRUST and LENIENCY treatments are significantly lower than in the COMMUNICATION treatment. This suggests that antitrust polices may generate trust among subjects, provided that none of the subjects have previously reported the cartel. As we will see when we next comment on the subjects' price choices, this observation implies that current antitrust policies are not unambiguously positive despite the fact

In their experiment the individual punisher never meets the same subject again and thus the observed patten cannot be explained by optimal punishments.

they increase ex ante cartel deterrence (and, in the LENIENCY treatment also the probability of detection due to higher rates of self reporting).

Prices

The ultimate objective of antitrust law enforcement is to generate low prices. Table 3 presents for our first three treatments the average price, the average price within cartels, the agreed upon price and the average price given that subjects do not communicate. The first lesson to be drawn from this table is that cartel deterrence is desirable, since it reduces prices. Indeed, for each treatment in Table 3, prices within cartels are higher than the prices without communication. (Although not reported, these differences are statistically significant.) Combined with our earlier finding that ANTITRUST reduces the rate of cartel formation relative to COMMUNICATION and that LENIENCY further reduces that rate, it suggests that average prices should be highest in COMMUNICATION followed by ANTITRUST and lowest in LENIENCY. If anything, our data suggests the reverse:

Result 4 (Average prices) Both traditional (ANTITRUST) and modern (LENIENCY) antitrust laws appear ineffective in reducing average prices.

Result 4 stems from the fact that average prices in ANTITRUST and LENIENCY are higher (although not significantly so) than in COMMUNICATION. This pattern thus suggests that both traditional and modern policies embedded with a leniency program are counter productive by increasing prices. The main driving force behind this result is that these policies appear to increase cartel stability (as noted in Result 3) and naturally this translates into higher prices within cartels (see Table 7.4).

Result 5 (Prices within cartels - what does not kill us makes us stronger) Both traditional (ANTITRUST) and modern (LENIENCY) antitrust laws increase cartel prices significantly.

Spagnolo (2000b) [97], Buccirossi and Spagnolo (2001 [14], 2006 [15]) and Ellis and Wilson (2001) suggested that antitrust policies embedded

Table 7.4: Price levels

	Communication		Antitrust		Leniency
Average price	4.913	$\approx$	5.348	$>^{***}$	4.844
Price within cartels	4.971	$<^{***}$	6.114	$<^{**}$	7.024
Agreed upon price	7.689	$<^{***}$	8.242	$\approx$	8.218
Price without communication	3.227	$\approx$	3.890	$\approx$	4.013

Note: For the **Average price**, a single observation is represented by the average among the prices chosen in a period by the two members of a duopoly. The same is true for **Price within cartels**, but here we only consider the cases in which the members of a duopoly have formed a cartel which has not been detected or reported yet. For **Agreed upon prices** we only consider the cases in which the subjects have communicated and found an agreement on the price to set. A single observation is given by the agreed upon price per duopoly, per period. For **Price without communication** we restrict our analysis to the cases in which no communication has taken place in the present period, and any possible previous agreement on prices has already been broken. A single observation is represented by the average among the prices chosen by the two members of a duopoly, in every single period.
As explained above, we used a three-level random intercept linear model to compare the results across treatments since the analysis here concerns duopoly-level data. We ran a regression per each couple of treatments.

with a leniency program could have the effect of stabilizing those cartels that are not deterred. Their idea was that reporting could be used as a punishment against deviators, since reporting is less costly with a leniency program. This potential explanation for the high cartel prices in LENIENCY is not completely convincing in the context of our experiment; since we allowed for secret reports, deviators could in effect protect themselves against such punishments. In fact, one may argue that reports as threats against deviators should be more relevant in ANTITRUST, since optimal punishments in that treatment involve reports. However, since we observed very few reports in ANTITRUST (although most of these few reports were used as punishments against deviators), it seems unlikely that reports as a threat against deviators were the main explanation for the high cartel prices in ANTITRUST.

In our view it seems more reasonable that antitrust policies generate trust among cartel members provided that the cartels are not reported. It is also interesting to note that cartel prices are significantly higher in LENIENCY than in ANTITRUST. This pattern suggests that the tougher the policy, the larger is the potential for generating trust among cartel members.

It is also interesting to note that the price levels for non cartel members appear to be higher (although insignificantly so) in the Antitrust and Leniency treatments than in the Communication treatments. One possible interpretation of this pattern is that a refusal to communicate when it is costly to do so, does not signal as clearly an unwillingness to cooperate. As a result, current antitrust policies may also facilitate tacit collusion. It should be emphasized, however, that because of higher deterrence, average prices overall are not significantly higher in the Antitrust and Leniency treatments.

High expected fines

To test the robustness of our findings to changes in α and F, we ran the two additional treatments, ANTIHIGH and LENHIGH with higher expected fines of 60 ($\alpha = 0.2$ and $F = 300$). Table 7.5 reports the rates of communication attempts and of cartel formation as well as average prices and prices within and outside cartels. These figures are compared with those for our original treatments, ANTITRUST and LENIENCY.

The first lesson from this table is that higher expected fines increase deterrence and reduce average prices under traditional antitrust laws but not under modern laws embedded with a leniency program. The reason is probably that the expected fine mostly increased through an increase in the probability of detection - this probability was doubled while the size of the fine increased by 50 % only - and that under leniency, many cartels are reported irrespective of the probability of detection, thereby reducing subjects sensitivity to changes in that probability. Note also that the prices within cartels increased in the leniency treatment but not in the antitrust treatment.

Post-conviction behavior

In this section we analyze agents' behavior after they are convicted and fined for a cartel they had formed before. This is interesting for at least two important and related reasons. The first reason is that of course there is not only general, ex ante deterrence. A second form of deterrence that any law enforcement policy should presumably generate is ex post *specific* deterrence, some times called *desistance* in the antitrust literature: the policy should ensure that convictions stop the convicted wrongdoer(s) from committing the crime again. The crucial question here is, therefore, how do convictions, in general and in particular when generated by different law enforcement policies (presence of leniency, size of fines), affect agents' following decision whether to form another cartel and - whether or not a new cartel is formed - their price choices? The topic is particularly interesting for antitrust in light of Sproul's (1993) [100] empirical finding that for a sample of US antitrust indictments prices often rose after antitrust conviction (see also the discussion in Whinston 2006 [106]).

The second related reason why post-conviction behavior is important is that a number of recent studies, theoretical and experimental, suggest that in oligopolistic industries the payment of a large sunk cost by competitors may lead to an increase in prices, either because the sunk cost

Table 7.5: *High expected fines*

	Antitrust		AntiHigh		LenHigh		Leniency
Rate of communication decision	0.590	$>^{**}$	0.452	$\approx$	0.435	$\approx$	0.344
Rate of cartel formation	0.316	$>^{***}$	0.195	$\approx$	0.163	$\approx$	0.146
Average price	4.34	$>^{*}$	4.00	$>^{*}$	3.65	$\approx$	3.93
Price within cartels	5.03	$\approx$	5.22	$<^{**}$	6.21	$>^{*}$	5.49
Price without communication	3.32	$\approx$	3.20	$\approx$	3.17	$\approx$	3.46

acts as a coordination device for explicit or tacit collusion (e.g. Offerman and Potters 2006 [80]; Janssen 2006 [57]), or because agents are subject to a 'sunk-cost bias', that is, they use simple mark-up pricing rules of thumb to try to recover the costs sunk by charging a higher markup (see e.g. Baliga et al. 2006, who also describe how the best business administration textbooks in fact suggest these pricing rules based on average cost and mark ups as optimal ones). Buccirossi and Spagnolo (2007) [16] noted that if these effects were present and significant in oligopolies, then the existing theory of optimal fines could not be applied to cartels as commonly done in the antitrust debate (it would be misleading): it should first be extended to incorporate these effects in the evaluation of the costs and benefits trade offs that lead to the optimal fines.

Specific deterrence (or 'desistance') Figure 7.3 shows the cumulative number of new cartels (vertical axis) formed by convicted agents in the five periods following the conviction (horizontal axis), separately for our Antitrust and Leniency treatments. The plots are slightly 'optimistic', in the sense that some of the matches end before the five periods after conviction considered, possibly leading to a slight underestimation of the number of cartels that form again after conviction. Still, the data tell us quite a lot.

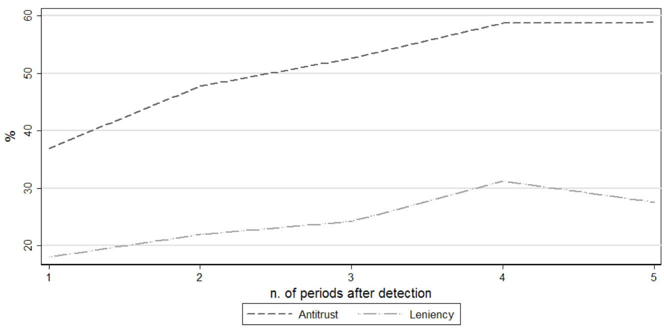


Figure 7.3: *% of cartels re-established*

First, there is a large fraction of agents that do not form new cartels after having been convicted and fined for a first cartel, even though our treatments are designed so that agents' situation in terms of expected fines, discount factors, available actions and payoff functions after a conviction is exactly identical to that before the first convicted cartel was formed. What differs after a conviction is only the history of play, as agents now have played several rounds, formed one or more cartels, and were convicted and fined. If history or agents' experience did not matter, so that 'bygones are bygones', in our stationary framework we would expect all 'rational' agents that chose to form a cartel a first time and were convicted and fined to form a new cartel the period after conviction. Instead, more than half of former cartelist did not form a new cartel periods days after the conviction. This suggests that history and experience matter a lot in our experiment.

Second, there is a strong difference between the specific deterrence effects of Antitrust and Leniency: close to 40% of convicted cartels come to life again in Antitrust treatments, but not in Leniency treatments. In other words, in our experiment *the introduction of leniency policies produces a strong increase in desistance*. Leniency policies appear much better at reducing recidivism than standard antitrust policies without leniency. This result is in stark contrast with Hinloopen and Soetevent (2006), who in their experiment find no improvement in desistance linked to the introduction of leniency policies. The reason behind their opposite result, in our view, is most likely due to their experimental set up not allowing for secretly reporting and simultaneously deviating from cartel agreement, as is possible in our experiment and in reality. As we have seen before, in our experiment most of convictions in Leniency treatments are linked to agents simultaneously undercutting cartel price and self-reporting. This joint action is likely to generate substantial more distrust among agents than a discovery by the competition authority, and thereby to make substantially more difficult for convicted agents to trust each other again.

Post-conviction prices In his paper on the effects of antitrust indictments on prices charged after the indictment - in the absence of an effective leniency program - Sproul (1993) [100] finds that:

1. On average prices rise gradually after an indictment for price fixing.

2. The largest immediate drops in price after conviction are about 9-10 percent.

3. Post-conviction prices are negatively correlated with the severity of penalties.

Sproul suggests that some of the cartels he analyzes could involve efficiencies, and imputes the increase in average prices to a loss of these efficiencies. A comparison between his results and ours, particularly under Antitrust treatments (there was no serious leniency program at the time of the cartels studied by Sproul) might help to understand some aspects of the phenomenon under analysis.

Figure 7.4 shows price choices in cartels before conviction (conviction takes place at time 0) and after conviction, separately for convicted agents that have formed again a new cartel and by those that did not do it, and for Antitrust and Leniency treatments. The stylized facts that emerge from our experiment are the following:

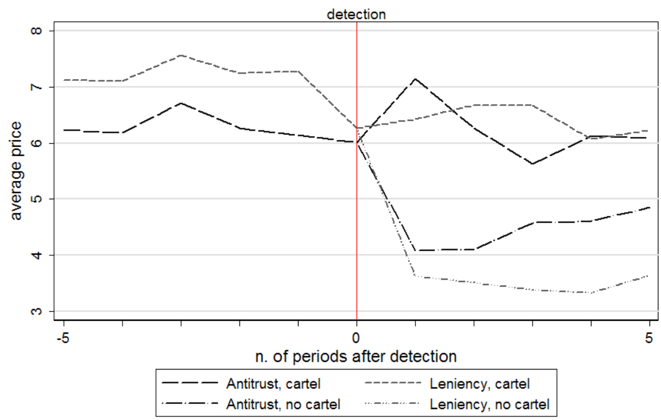


Figure 7.4: *Price before and after detection*

a) Prices after conviction are on average lower than in cartels before conviction.

b) When cartels are re-established after conviction, prices stabilize at levels close to that prevailing in the cartel the period when the cartel was convicted.

c) When cartels are not re-established, prices fall substantially with respect to the prevailing cartel price at the time of conviction, and remain low.

d) Post-conviction prices are higher in Leniency than in Antitrust treatments when a new cartel is formed after conviction, while the opposite happens when a new cartel is not formed after conviction.

The fact that average prices within cartels that are restored after conviction remain close to the level observed in the period in which the previous cartel was detected, whether leniency is granted to the whistle-blowers or not, appears consistent with Sproul's (1993) [100] findings given that in our framework there are no efficiencies linked to collusion. Somewhat in contrast to Sproul (1993) [100] we also find that, in the large number of cases where a novel cartel is not formed after conviction, prices fall much below the level reached in the period in which detection took place, which drives down average post-conviction prices. True, this happens much more often in Leniency treatments than in Antitrust ones, and prices when new cartels do not form are much lower in Leniency than in Antitrust treatments, while at the time of Sproul's cartels an effective leniency policy was not in place. Still, even focusing only on Antitrust treatments, it appears that prices fall on average after conviction. On the other hand, to explain why in his sample prices do not fall after indictment, Sproul hypothesizes that "the government mainly prosecutes cost-reducing cartels". Such an interpretation is not questioned by our data, since in our experiment cartels have no effects on costs.

As for the effects of Leniency, it appears to have the novel effect of strongly increasing desistance thereby reducing average prices, even though prices are substantially higher in cartels that manage to form again after a conviction caused by a leniency application.

The difference that arises between Leniency treatments and Antitrust treatments when players decide not to form a new cartel after being detected is also interesting (stylized fact d)). While under Leniency the average price remains close to Bertrand and to the level observed before the (detected) cartel arose, under Antitrust average non-collusive prices after detection rise as if – after having formed an explicit cartel and having experienced the fine – some of the subjected try to reach a tacit agreement on prices. A possible interpretation of this effect is that under Antitrust detection does not affect trust between cartelists, while under Leniency detection and defection are often simultaneous, and the cartel is discovered because it is reported by the deviating player; therefore, post-conviction tacit collusion is more difficult to achieve under Leniency.

Size of the fine and post-conviction prices: 'sunk cost bias' and co-ordination effects As mentioned before, there are studies pointing at possible coordination role of sunk costs (e.g. Offerman and Potters 2006 [80]), while other studies point at possible 'sunk cost bias' in decision making, where agents try to recover sunk fixed cost by increasing a mark up on the average cost chosen when setting the unit price (Baliga et al. 2004). To distinguish the two effects, we hypothesized that the first effect should imply improved coordination in general, and therefore also in newly formed cartels.

Table 7.6 reports post conviction prices from our experiment, in newly formed cartels and outside, and the level of the fines levied on convicted agents. Consistently with Sproul's finding number 1, we observe a negative (though not always significant) correlation between the size of the fine and post-conviction prices. In our experiment this effect is somehow puzzling, since even before getting fined our subjects were informed about the size of the fine and the probability of detection, so if they were fully rational they should not change their behavior after detection. A deeper analysis is required to understand the reasons that lead to this finding [15].

We observe that post conviction prices are generally lower when the fine (and the expected fine) is higher, both within cartels and outside

[15] To investigate this matter, we ran some other related experiments' whose results will be presented in a companion paper.

Table 7.6: Size of the fine and post-conviction pricing

Treatment	Fine	Prices outside cartels	Prices within cartels
Antitrust	200	4.418	7.297
		$\approx$	$\vee^*$
AntiHigh	300	3.310	5.750
Leniency	200	3.776	6.732
		$\approx$	$\approx$
LenHigh	300	3.181	4.700

cartels, whether leniency is granted to the "whistleblowers" or not. Consequently, our evidence seems to contradict the hypothesis of a sunk cost bias, which would affect prices of firms that choose not to re-establish a cartel after being fined; our results are also against the hypothesis of a coordination effect of the fine for cartels restored right after their detection.

To test the significance of the observed difference in post conviction prices between Antitrust and AntiHigh, and between Leniency and LenHigh, we estimated a three level random effect linear model using GLLAMM, following the procedure explained in section 7.3.6. As mentioned above, this procedure allows us to keep into account the correlation between observations from the same duopoly, and also the correlation between observations from the same city. We notice that the differences we observe are economically, but not statistically significant in most of the cases. According to our results, the difference in post conviction prices between Leniency1 and Leniency3 is not significant, neither within cartels nor outside cartels. On the other hand, the difference between prices observed in Antitrust1 and Antitrust3 is significant, but only outside cartels. This lack of statistical significance may be due to the sample size, which is very small since we restrict our analysis only to the cases in which a cartel was discovered and dismantled in the previous period.

7.4.2 Ineligibility for Cartel Ringleader[16]

Under the US Corporate Leniency Policy, a firm is ineligible for amnesty if it is the instigator of the cartel - the so called ringleader. In order to qualify for amnesty, the policy requires that the "corporation did not coerce another party to participate in the illegal activity and clearly was not the leader in, or the originator of, the activity" (Corporate Leniency Policy, *supra* note 58). By contrast, and following the revision of the EU Leniency Notice in 2002, also the ringleader is eligible for amnesty in the EU. Whether or not ineligibility of the ringleader has desirable consequences in terms of deterrence and prices is not clear cut. Excluding the ringleader from the leniency program may increase deterrence if each firm wait for some other firm to take the initiative of forming the cartel. As noted by Leslie (2006) [71], however, extending amnesty to the ringleader may increase deterrence as well by ensuring that even the ringleader cannot be completely trusted, as it may also loose confidence and rush to report under the leniency program.[17]

Table 7.7: *Deterrence effects*

	Leniency		LenRing
Rate of communication attempts	0.344	$\approx$	0.290
Rate of cartel formation	0.146	$\approx$	0.135
Rate of individual deviation	0.472	$>^{***}$	0.230
Rate of detection	0.646	$>^{***}$	0.289

To evaluate the pros and cons of ringleader ineligibility, we ran the additional treatment, LenRing. Tables 7.7 and 7.8 compare the effects on deterrence and on price levels of eliminating the possibility of amnesty for the ringleader. Three features are striking in these tables. First, the LenRing treatment has no significant effect on cartel deterrence relative

[16]We thank Joe Harrington for suggesting this treatment.

[17]There are other arguments for and against the ineligibility of ringleaders. Extending leniency to the ringleader may be important to elicit self-reporting, as it may not be that clear to a firm considering whether to apply for leniency if it risks being regarded as ringleader. On the other hand, in an adversarial system, where testimony is crucial to persuade juries, testimony by a ringleader may not be convincing.

to the Leniency treatment. Second, cartels appear to become more stable and third the LenRing treatment increase prices significantly according to all our price measure. These findings are summarized in the next result[18].

Result 6 (Ringleader) If the ringleader is excluded from the leniency program, the deterrence effect of leniency falls and prices are higher than otherwise.

Result 7 thus suggests that the US practice of excluding the ringleader from the leniency program is unambiguously bad in our set up. While we find this result an interesting first step, that confirms some observers' concerns that excluding ringleaders may reduce the effectiveness of the leniency program, we should also emphasize one important caveat. In our experiment subjects were matched pairwise into duopolies to avoid social preferences effects towards non-defecting third parties.

Table 7.8: *Price levels*

	Leniency		**LenRing**
Average price	3.926	$<^{***}$	4.847
Price within cartels	5.494	$<^{***}$	7.284
Agreed upon price	7.099	$\approx$	7.833
Price without communication	3.457	$<^{***}$	3.912

This, however, is the worst conceivable situation for the US policy of excluding ringleaders, as the ban leaves only one cartel member with the option to self-report obtaining leniency, eliminating the incentives to "race to report" generated by the risk that another cartel member could do it before. With more than two firms, therefore, it is likely that LenRing treatment will show more desirable properties. Therefore, further experimental research with many cartel members is needed to attempt any policy conclusion on this feature.

[18] Treatment LENRING was run only in Rome. For sake of consistency, in tables 7.7 and 7.8 we only consider observations gathered in Rome for treatment Leniency as well.

7.4.3 Rewards

So far we have only considered policies that have been extensively implemented in reality. Given that none of these policies yielded fully satisfactory results, it is natural to turn attention to policies that have been advocated in the literature on optimal law enforcement. The type of policy that we consider here is one where the reporting subject gets rewarded by an amount equal to the fine paid by its rival.[19] Tables 7.9 and 7.10 compare the effects on deterrence and on price levels of introducing such reward schemes.

Result 7 (Ex post deterrence) Cartels are systematically reported in the Reward treatment.

This result is corroborated by Table 6 showing that the rate of detection due to reporting is almost equal to one in the Reward treatment. In fact, a simple inspection of the data in the Reward treatment reveals that almost every time a cartel was formed, at least one of the subjects reported it: out of the 120 times a cartel was formed, the cartel was reported during the first period in 118 cases. In one of the remaining cases, it was reported in the subsequent period, while there was only a single duopoly in which the players resisted the temptation of reporting and managed to sustain the collusive agreement for 7 consecutive periods. This cartel ended because a re-matching took place.

There are two potential explanations to this phenomenon. First the subjects could in principle exploit the reward system by taking turns in reporting and cashing in the reward. The second hypothesis, first proposed by Apesteguja et al, is that subjects form a cartel with the hope of fooling their competitor by undercutting the agreed-upon price and by reporting the cartel in order to cash in the reward. The next result confirms this latter hypothesis.

Result 8 (Cartel stability) The antitrust policy with rewards significantly reduces cartel stability.

[19]Korea is the only country we are aware of that adopts this kind of reward schemes for whistleblowers in antitrust; analogous schemes are however used in other fields of law enforcement, particularly in the US.

Table 7.9: Deterrence effects

	Antitrust		Reward		Leniency
Rate of communication decision	0.566	$>^{***}$	0.484	$>^{***}$	0.377
Rate of cartel formation	0.315	$>^{***}$	0.220	$\approx$	0.178
Rate of individual deviation	0.424	$<^{***}$	0.781	$>^{***}$	0.373
Rate of detection	0.092	$<^{***}$	0.937	$>^{***}$	0.507

This result is reflected by the fact that the rate of individual deviation increased substantially in the Reward treatment. Note also that at least one subject undercut the agreed upon price in 111 out of the 118 cartels that only lasted one period.

Table 6 also suggests that the antitrust policy with rewards is not more efficient in deterring cartels ex ante than the traditional policies. Indeed the rates of communication attempts and of cartel formation in the REWARD treatment are not significantly different from corresponding rates in the ANTITRUST and LENIENCY treatments. Nevertheless:

Result 9 (Ex ante deterrence) The antitrust policy with reward strongly deters explicit cartel formation, the more the longer subjects play.

Result 7 is explained by the fact that the subjects eventually learned that it was not possible to form cartels for the purpose of cashing in the reward and, as a result, the number of formed cartels was reduced drastically. This appears clearly in Figure 7.5 showing that the number of cartels formed were reduced as subjects were re-matched.

Result 10 (Prices) The antitrust policy with rewards strongly reduces both prices in explicit cartels and subjects' ability to collude tacitly.

The systematic reports when subjects took part in cartels probably undermined trust among the subjects and, as a result, also prices dropped

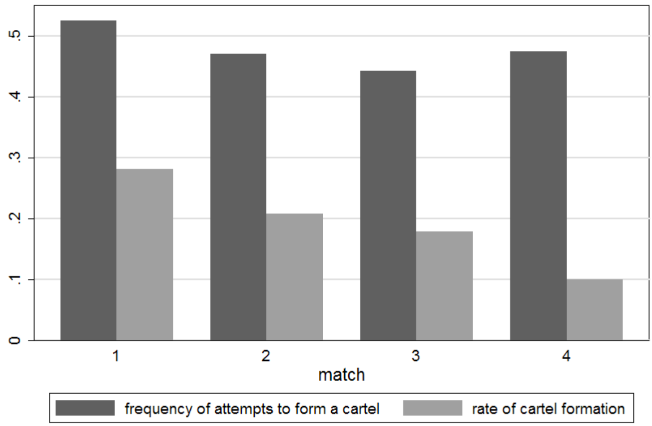


Figure 7.5: *Rates of cartel formation and communication attempts in the four matches of the Reward treatment.*

Table 7.10: *Price levels*

	Antitrust		Reward		Leniency
Average price	5.348	$\approx$	3.975	$\approx$	4.844
Price within cartels	6.144	$>^{**}$	5.339	$<^{***}$	7.024
Agreed upon price	8.242	$\approx$	8.512	$\approx$	8.218
Price without communication	3.890	$\approx$	3.565	$\approx$	4.013

(though not significantly) in the Reward treatment. This is most striking in Table 3 where all measures of prices are the lowest in the Reward treatment. In particular, prices also dropped when subjects did not communicate. Thus prices were not only low because cartels were deterred from forming, giving further support for the claim that the systematic reports undermined trust. Rewarding whistleblowers appears therefore the only antitrust policy able to reduce price and increase welfare in our experimental set up.

Also, note that we designed the reward scheme in such a way that it could be exploited by cartel members: the scheme is such that if cartel members took turns in self-reporting and cashing the reward, the expected fine would be zero (Spagnolo 2004 [98], 2008 [99] makes it clear that - for this reason - in reality the whistleblower's reward should always be strictly smaller than the sum of fines paid by the other wrongdoers, so that there exist no possibility to manipulate/exploit the scheme). Still, none of our subject appear to have realized this opportunity, a result that confirm that some legal scholar's claims that reward schemes could be manipulated are unfounded empirically, besides being incorrect theoretically when the scheme is appropriately designed (by an economist). The result is consistent with Dal Bo's (2005) [29] finding that asymmetric (alternating) actions cooperative equilibria in the repeated Prisoner's Dilemma are never played by experimental agents even when they are way more efficient than standard, stationary cooperative equilibria.

In theory, and contrary to any previous result in the economic literature on law enforcement, it is possible to achieve the first best of full deterrence with finite fines and no inspection probability using reward schemes (Spagnolo, 2000 [96], 2004 [98]). To verify this we ran a further treatment REWLOW, identical to REWARD but for the probability of detection that we set at zero ($\alpha = 0$). The results showed that some agents still need to try a couple of times to induce others to form a cartel just to report and cash the bonus. After a couple of attempts they learn that with the whistleblower reward scheme everybody reports immediately once a cartel forms, so that entering a cartel is never profitable, and cartels disappear. The first best is therefore only achieved in asymptotic form. Indeed, after enough learning our agents appear to converge to the first best. We also run a second treatment REWHIGH, exactly equal to REWLOW but for the fine, which was higher ($F = 1000$). In this last treatment we observed a further increase in deterrence (the rate of communication attempts dropped to 25.9%, and the rate of cartel formation was 7.8%). Yet, full deterrence was not achieved even in this treatment, because at least some of the players still tried to form a cartel and to fool their opponent by deviating from the agreement and simultaneously reporting it.

As a final remark we stress that that the results of our treatments with rewards forcefully confirm the importance of learning in experimental settings with complicated/realistic games if we compare to those of Apesteguja et al. (2006) [2]. In particular, the results show that the preliminary and strange conclusion by Apesteguja et al. (2006) [2] - that Rewards are not more effective than Leniency - was premature and due to the subjects' inability to learn in that set up, as these authors already conjectured.

7.4.4 Culture, Trust and Antitrust

We run our experiments in Stockholm and in Rome, two towns with quite distinct cultures. It is not obvious that one or the other culture should lead to more cartels given the legal framework, as our experiment was framed, cartel formation was presented as illegal, and Swedes are thought to be more law abiding than Italians. This would point at Italians colluding more. On the other hand, if one reads the World Values Survey (1999), finds other important differences between Sweden and Italy that may point in a different direction. In particular, here are two relevant questions with relative answers:

(i)"Give authorities information to help justice": Strongly agree

[Italy: 40.2% vs. Sweden: 26%]

(ii) "Most people can be trusted" [Italy: 31.8% vs. Sweden: 63.7%]

The difference in the answers to the first question suggests that leniency programs could be more effective in Italy. The difference in the answers to the second question suggest that Swedes are more confident in the cooperation of partners, so that they are more likely to coordinate on collusive/cooperative equilibria.

Separating treatments according to location we found results consistent with the differences in answers to the World Value Surveys: Swedes collude more often, coordinate on higher prices, and deviates much less often than Italians. In all treatments prices lower and cartels less frequent in Italy than in Sweden, and defection and applications to leniency are much more frequent in Italy. According to our results, nordic countries may be in more need of antitrust enforcement because of their 'cooperative' culture than southern ones.

7.5 Conclusions

This paper reports results from an experiment designed to examine the effects of fines, leniency programs, and reward schemes for whistle-blowers on firms' decision to form cartels (cartel deterrence) and on their price choices. We consider an experiment in which subjects play a repeated Bertrand price game with differentiated goods, running several treatments different in the probability of cartels being caught, the level of fine, the possibility of self-reporting (and not paying a fine), the existence of a reward for reporting, the option to communicate, and cartel leaders access to leniency. We find that fines following successful investigations but without leniency have a deterrence effect (reduce the number of cartels formed) but also a pro-collusive effect (increase collusive prices in surviving cartels). Leniency programs might not be more efficient than standard antitrust enforcement, since in our experiment they do deter a significantly higher fraction of cartels from forming, but they induce even higher prices in those cartels that are not reported, pushing average market price higher than without antitrust enforcement. With rewards for whistle blowing, instead, cartels are systematically reported, which completely disrupts subjects' ability to form cartels and sustain high prices. If the ringleader is excluded from the leniency program the deterrence effect of leniency does not increase while prices are higher than otherwise. As for tacit collusion, we find that under standard anti-trust enforcement or leniency programs, subjects who do not communicate (do not go for explicit cartels) choose weakly higher prices than where there is no anti-trust enforcement. We also analyze and post-conviction behavior, finding that after convictions caused by a report under the leniency program much fewer cartels form and prices are lower than when conviction takes place under standard antitrust policies without leniency. Finally, we find a strong cultural effect in the deterrence power of the various law enforcement regime when comparing treatments in Stockholm and Rome.

Our results have policy implications for general deterrence of organized crime similar to cartels, and as a test for the theoretical results mentioned in Section 2 (the protection from fines effect, the reward effect, the reporting as a threat and tacit collusion). Our results have only marginal implications for most of the many other theoretical papers, including cornerstone contributions to this literature.

For example, Motta and Polo (2003) [76], the first economic analysis on leniency programs, focuses mainly on the effects of the second part of the leniency programs opened to firms already under prosecution. The only implication of our experimental results for that paper is that they do not support its two policy conclusions that (i) to have deterrence effects leniency programs must be opened to firms already under investigation (in our experiment they aren't), and that (ii) introducing a leniency program is a second best choice relative to standard antitrust law enforcement if there is a large enough budget. Analogous, Harrington (forthcoming) [45] does not consider general deterrence but focuses on desistance effects, i.e. the ability of law enforcement mechanism to shorten the duration of cartels that were not deterred by the mechanism. It also introduces a stochastic movement in a law enforcement parameters to generate equilibrium applications of cartel members to the leniency schemes. The only implication of our experimental results for that paper is that our real world agents did form cartels and then apply for leniency in our fully deterministic, stationary oligopolistic environment. This suggests that deterministic theoretical analyses are perfectly OK, they are not at odd with the evidence that people form cartels and then report them.

Aubert et al. (2006) [4] do focus on the first part of leniency programs, on general deterrence and on rewards, but their contribution is about the costs and benefits of providing leniency and rewards to the individual employees of colluding firms. In our experiment we only have single decision makers, so we have no evidence relevant to that issue. This sounds, however, as an exciting topic for future experimental work.

Chapter 8

Risk Aversion, Prospect Theory, and Strategic Risk in Law Enforcement[1]

joint with Sven-Olof Fridolfsson, Chloé Le Coq and Giancarlo Spagnolo

8.1 Introduction

This paper reports results from an experiment designed to examine the effects of fines and of leniency programs on firms' decision to form cartels (cartel deterrence). We consider an experiment in which subjects play a repeated Bertrand price game with differentiated goods, running several treatments, which differ in the probability of cartels being caught, in the level of fine and the possibility of self-reporting and getting leniency.

[1]Many thanks to Magnus Johannesson for discussions and advice related to this project, and to audiences in Rome (World ESA Meeting 2007) and in Gothenburg (Second Nordic Workshop in Behavioral and Experimental Economics, 2007). We also gratefully acknowledge research funding from Konkurrensverket (the Swedish Competition Authority) that made this research possible.

Leniency policies, or programs, grant full or partial reductions of the sanctions to firms that report hard information about their cartel to the Antitrust Authority and cooperate with it along the prosecution phase, helping to convict their former partners. These policies have been introduced in most OECD countries and have become the main tool for cartel discovery and prosecution; their validity and effectiveness, though, is hard to asses since in fact the number of undetected cartels is not observable. Therefore, it is only possible to compare the number of detected cartels with and without leniency programs, but not the total number of existing cartels, meaning that in principle an observed increase in convicted cartels could even be due to an increase in cartel activity. For this reason, we think that an experimental approach is needed to collect more evidence about leniency's effects.

The results presented in this paper appear to be also relevant to the analysis of many other forms of multi-agent organized crime – corruption, auditor-manager collusion, corporate crime in general – which share with cartels some crucial features that well designed law enforcement programs may exploit [2]. A first important characteristic of these category of crimes is that cooperation among several agents is required to perform the illegal activity, so that free riding, "hold-up", and "moral hazard" issues become relevant. Moreover, the criminal activity takes the form of an ongoing relation, meaning that the membership of the criminal organization produces flows of present and expected future benefits and damages, instead of isolated gains or losses. Finally, cooperating wrongdoers, by acting together, inevitably end up having information on each others' misbehavior that could be reported to third parties, which is the main characteristic that could be exploited by leniency programs.

In this paper we don't only investigate if, but also how leniency programs work. In particular, we try to figure out which are the determinants of deterrence.

From a theoretical point of view, there are three conditions that have to be satisfied for a cartel to be formed. First, the individual incentive to commit the crime must be strong enough, which means that the expected

[2] Spagnolo (2004) [98]

utility it provides overcomes the expected disutility from the uncertain punishment. Second, the incentive compatibility constraint must be satisfied; so the long run gains from sticking to the collusive agreement must be higher than the short run gains from deviation plus its long run negative consequences. Third, the level of trust among cartel's members must be high enough: indeed we claim that the stability of a cartel does not only depend on each member's incentives to deviate, but also on the perceived risk of being cheated upon by other members, or strategic risk, which increases as trust among cartel's members decrease, and also as the "sucker's" payoff worsen. Strategic risk also affects cartelists' ability to coordinate on the joint profit maximizing equilibrium. Indeed, if the level of perceived risk associated to it is too high, then they could select a different equilibrium, with lower incentives to deviate or better outcomes for the cheated upon players in case a deviation takes place.
According to these considerations, the perceived risk of detection and trust – intended as the perceived risk of being cheated upon – are two important drivers of individuals' proclivity to collusion. In our study, we analyze how different legal frameworks impact on these two types of risk. In doing this, we also keep into consideration the many findings in psychology and in behavioral and experimental economics which show that the way people react to probabilities attached to risky outcomes departs in many ways from the standard model of full rationality.

Our main results are that *(i)* deterrence is generally higher when leniency is granted to the whistleblowers,*(ii)* a negative relation emerges between the sum of the fines paid and participants' willingness to cooperate, and *(iii)* communication rates drop after conviction. Our analysis shows that strategic risk and availability heuristic are among the main drivers of these three outcomes, even if they are generally disregarded in the traditional approach to the study of law enforcement. In particular, strategic risk is determinative both in explaining deterrence under Leniency treatments, and in justifying the drop in communication rates after detection when a deviation took place. Availability heuristic and the salience of fines, on the other hand, are the most plausible reasons why players willingness to communicate decreases after detection even if no deviation has previously taken place, as a fresh memory of the punishment increases the perceived probability of detection. These behavioral biases also seem to motivate the negative effect exerted by the sum of the fines paid on participants' willingness to cooperate.

The paper proceeds as follows: Section 2 presents the literature related with crime deterrence and with the behavioral effects that might affect it. The experimental design and procedure are described in Section 3. Section 4 reports the results and explain the empirical methodology adopted to analyze our data, and Section 5 concludes, also mentioning possible perspectives for further research.

8.2 Related Literature

8.2.1 Rational agents

Individual benefits and costs from criminal activities. Public enforcement of law is a widely investigated subject since 1968, when Gary Becker [9] published an article which was to become one of the cornerstones of economic analysis of law.
Under Becker's approach, then followed by many scholars such as Polinsky and Shavell [83], individuals are fully rational utility maximizers, and so are offenders. It is assumed that

> a person commits an offense if the expected utility to him exceeds the utility he could get by using his time and other resources at other activities. Some persons become "criminals", therefore, not because their basic motivation differs from that of other persons, but because their benefits and costs differ.

According to this approach, the number of offenses committed by a person can be related, by means of a function, to the probability of conviction, to the punishment imposed in case of detection and to the gains associated with the illegal activity: in practice, the potential offender will commit a crime only if the expected utility attached to the crime's outcome exceeds the expected (dis)utility of the possible sanction.

Organized Crime. As noticed by Spagnolo (2004) [98] and by Motta and Polo (2003) [76], when we consider organized crime as opposed to individual crime, the balance between private benefits and costs from the offense cannot be the only determinant of the decision to commit it. Spagnolo states that "criminal organizations suffer of an intrinsic 'governance problem' since to curb moral hazard and ensure internal cooperation they cannot rely on explicit contracts enforced by the legal system. For this reason, many forms of organized crime must take the form of – or be conducted within – long-term dynamic criminal relationships". Such relationships can be modeled as repeated Prisoner's dilemma like games, in which each player is always tempted to "take the money and run", and does not do that as long as the expected long run gains from sticking to the illegal agreement overcome the short run gains from defection plus the long run consequences that might result from it.
It is precisely on this balance that Leniency programs exert their deterrent effect: in absence of leniency, reporting the crime to the authority is always a dominated action, while with a Leniency Program agents may in fact find it convenient to report when they deviate from the criminal agreement, which can make defection more desirable than adherence to the illegal organization.

Strategic risk According to the prevailing theory, an individual's decision to cooperate – i.e. to take part in a criminal agreement – is determined by the trade off between the consequences he expects from sticking to the agreement and the anticipated outcome from deviation, but not by the consequences possibly yielded by a deviation by some other member of the agreement.
In a recent paper Blonski and Spagnolo [11] critique this approach, suggesting that

> real world agents do care about what would happen if other agents defected from the agreed strategy profile, and that these considerations should not be left out of our models.

They formalize the consequences of a variation in the sucker's payoff in a Prisoner's Dilemma game, and argue that a decrease in the cheated upon player's payoff increase the "strategic risk" associated with playing cooperatively, reducing the sustainability of cooperative equilibria in the long run. To sketch the idea of strategic risk, consider an infinitely repeated PD game, whose stage game is represented in table 8.1: According to the

	C	**D**
C	r	s
D	t	p

Table 8.1: *Prisoner's dilemma*

standard approach, the discounting factor required to sustain collusion must be higher than a threshold value $\underline{\delta}$, identified by the constraint:

$$\frac{r}{1-\delta} \geq t + \delta \frac{p}{1-\delta}$$

which can be rearranged as

$$\underbrace{\delta \frac{r-p}{1-\delta}}_{\text{LR inc. to coop.}} \quad \geq \quad \underbrace{t-r}_{\text{SR inc. to def.}} \quad .$$

where the left-hand side of the inequality represents the long-run incentive to cooperate, and the right-hand side stands for the short-run incentive to deviate.

Strategic risk approach suggests that also the *short run disincentive to cooperate* $(p - s)$ matters, so the appropriate threshold value for the discount factor is $\delta^* > \underline{\delta}$, determined by:

$$\underbrace{\delta \frac{r-p}{1-\delta}}_{\text{LR inc. to coop.}} \quad \geq \underbrace{t-r+p-s}_{\text{total SR inc. to def.}}$$

Bolnski and Spagnolo's theoretical hypothesis has found some support in the experimental evidence recently provided by Dal Bó and Frechette [30], who study how the evolution of cooperation in infinitely repeated prisoners dilemmas is affected by changes in the difference between the reward from cooperation and the sucker's payoff.

8.2.2 Behavioral Law and Economics

The models mentioned above proceed with the hypotheses of neoclassical economics, assuming individuals to be perfectly rational expected utility maximizers. Empirical evidence collected by psychologists and by experimental economists, though, casts some doubts about these assumptions: people's behavior has been proved to violate the classical paradigm of *homo oeconomicus* in many ways. More specifically, according to Jolls, Sustein and Thaler (1998) [65], human behavior departs from the standard conception of *homo oeconomicus* under three main respects: they state that people display

- bounded rationality, in that they suffer from certain biases – such as overoptimism, mis-perception of probabilities or self serving biases – and they adopt heuristics that lead to mistakes;
- bounded will power, which sometimes reflects into myopic behavior;
- bounded self interest, meaning that they care for other people's well being.

In light of these findings, the authors suggest that models of economic analysis of law that do not keep this factors into account may lead to erroneous conclusions, therefore they develop and propose a new approach to this branch of studies, "informed by a more accurate conception of choice, one that reflects a better understanding of human behavior and its wellsprings."

Since their seminal article, Behavioral Law and Economics has developed and has been applied to several specific topics in economic analysis of law (Jolls 2007, [64], and Garoupa 2003 [41] for a critical review).

In what follows, we will focus only on one aspect of bounded rationality: in particular we will consider how a biased perception of risk may affect law enforcement.

Risk attitude Of the various behavioral aspect that might affect law enforcement, risk attitude and risk perception are among the most important ones. Indeed, the effects of risk aversion were already mentioned by Becker (1968) [9], who argues that, if players were risk neutral it would be possible to minimize the costs of apprehension and conviction by lowering the probability of detection arbitrarily close to zero while rising the severity of the punishment, and states that this should be a fortiori true if offenders were risk avoiders. Notice that this is not only a theoretical matter, but it has strong policy implications which are currently under debate within the Competition Authority of a European country.

Considerations about agents' risk attitude are generally well integrated into the traditional approach to economic analysis of law [3]. But other aspects seem to be important too.

Heuristics and biases. Tversky and Kahneman (1979) [67] already observed that the way people react to probabilities attached to risky outcomes departs in many ways from the basic tenets of expected utility theory: they notice for example that individuals tend to overweight outcomes that are considered certain, relative to outcomes which are merely probable (certainty effect); they also observe that agents round probabilities or outcomes in order to simplify the analysis of risky prospects, and that "a particularly important form of simplification involves the discarding of extremely unlikely outcomes." On the other hand, they also suggest that low probabilities are overweighted, meaning that people overreact to rare events but may underreact to common ones. The interplay of the two last mentioned effects implies that "highly unlikely events are either ignored or overweighted, and the difference between high probability and certainty is either neglected or exaggerated." This consideration clearly plays a role for the analysis of optimal law enforcement: indeed, if a very small probability of liability is approximated to zero, then Becker's argument claiming that it is possible to reduce prosecution costs without affecting deterrence, by lowering the probability of detection and harshening the penalties, does not hold anymore.

[3] see, for example, Polinsky and Shavell (2000) [83]

The weight attached by individuals to the risk of conviction may be also affected by another behavioral bias: the salience effect. As highlighted by Akerlof (1998) [1], outstanding events and vivid information may exert undue influence on decisions: he refers to this principle to explain time inconsistent decisions – arguing that present costs and benefits are salient if compared to future ones – and to provide a possible reason for the "undue" obedience to authority – which can emerge when disobedience is perceived as more salient than compliance because it implies a deviation from the status quo or from a previous course of actions, and when some degree of disutility is attached disobedience.
Similarly, one could argue that an exacerbation of punishments may increase deterrence since extremely harsh penalties are more salient, than overweighted.
A close but different behavioral effect concerning probability perception, foregrounded by Tversky and Kahneman (1982) [66], is "availability heuristic": a mechanism by which occurrences of events associated with extremely high utilities or disutilities are perceived as being more frequent than they actually are. The main difference between availability heuristic and salience is that according to the first one risk perception is driven by memory-dependent mechanisms, while the second one states that attention is guided by the most vivid present stimuli.
Availability heuristic has been tested and confirmed by Folkes (1988)'s studies on the risk perceived by consumers when purchasing a product [37], and by a recent study by Keller et al. (2006) [69] on perception of flood risk, which testifies that past experience of flooding increases risk perception independent on the information exogenously provided about this risk. This piece of evidence supports the hypothesis that people who experienced past flooding events and have of them images that are tagged with affect perceive the same probability information differently from people without such memories.
Availability heuristic and salience can be interpreted as a result of the interplay of two fundamental ways in which human beings comprehend risk: "risk as analysis and risk as feelings" (Slovic et al., 2004 [95]), the first one being based on the brain "analytical system" – which encodes reality

in abstract and symbolic terms, builds logical connections between events and requires logical or empirical justification for actions – the second one being related to the "experiential system" – which on the contrary is associated with the experience of affects, motivate actions on the basis of the emotional memory of related events and encodes reality in concrete images, metaphors and narratives. The authors suggest that availability heuristic may work because images and events that are tagged with affect are more easily recalled or imagined. Events that are more sensational or salient are also more affectively charged, which might explain the overestimation of their frequency or probability, both ex ante, before they are actually experienced by the subjects, and ex post, when individuals have memories associated with these events.

Empirical evidence As we have just seen, there are many psychological effects that might affect law enforcement through the way people perceive the risk of being liable. Nonetheless, to our knowledge, the empirical and experimental evidence collected to test the different theoretical predictions in the specific context of crime deterrence is not very rich.
Levitt performed some interesting field studies about the actual relation between punishment and deterrence, also keeping into account possible bias like criminals' myopia or overoptimism (Levitt, 1998 [72]), but experimental approach seems to be more suitable to analyze psychological motives behind peoples responses to legal sanctions. In this second field of research, we should mention a study performed by Cason and Gangadharan (2006) [22], who experimentally analyze a model of compliance developed by Harrington (1988) [46] in which the enforcement agency modifies the inspection frequency and severities of the penalties depending on the firms past compliance. They find that violation rate does not change as sharply as predicted by the model when the probability of detection and size of the fine change, and show that the observed behavior might be captured by a quantal choice model, which accounts for boundedly rational decision making by allowing individuals to make errors, assuming though that errors that are more costly are also less probable.

A second experimental work testing predictions of behavioral economics in the context of law enforcement has been carried on by Jaquemet et al. (in press) [56], who study the role of optimism bias on the monitoring of illegal activities. They show that subjects exhibit a strong tendency to under-evaluate their own likelihood of experiencing an unfavorable event as compared to the one of others, which leads to a lower level of deterrence.
A third empirical work on punishment and deterrence has been recently presented by Fishman and Pope (2006) [36]: they study punishment induced deterrence, i.e. "the subsequent deterrent effect [...] that actually experiencing punishment for a crime has on the specific individual who was punished, conditioning for changes in expected benefits and costs of future criminal activity". Using field data from the movie-rental market, they explore the effect of having to pay a late fee on costumers' movie-rental and movie-return decisions, and show that:(*i*) experiencing punishment decreases the offender's crime rate (in the short run); (*ii*) salience (size and temporal proximity) of punishment is positively related with deterrence and (*iii*) the effect does not vanish with experience.
Their results confirm that the experience of a penalty affects the weight individuals attach to punishment when they have to decide whether to comply or not to a prescription, an effect that to us could be attribute to the aforementioned availability heuristic.

Organized Crime and Trust. Assessing and evaluating the risk of liability is an important problem for every potential offender, regardless the nature of his crime. When we focus on organized crime, though, another significant element has to be taken into account, namely trust between members of the "criminal organization". Reciprocal trust is important for two reasons: first, as mentioned before, organized crime can be modeled as a repeated prisoner's dilemma like situation, in which each player will choose to stick to the criminal agreement only if he has a strong enough belief that the other will do the same. Second, cooperating wrongdoers, by acting together, inevitably end up having information on each others' misbehavior that could be reported to third parties, and each of the members of the organization has to be sufficiently confident that this will not happen.

Behavioral and experimental economics offer a rich literature about trust; for sake of conciseness, we cannot mention it all, and we will only cite one recent paper by Sapienza, Toldra and Zingales (2007) [89], whose results are particularly interesting to us. They study a standard trust game and, among other things, they find that a trusting behavior is determined by three main factors: beliefs about others trustworthiness, risk aversion and other regarding preferences. Considerations about the third element are outside the scope of our work, while the first two factors play a crucial role for the analysis of organized crime. The relation between risk aversion and trust reveals that offenders' risk attitudes affect deterrence not only via the risk of liability, but also because of the risk of betrayal on behalf of some other member of the criminal organization; beliefs about others' trustworthiness, on the other hand, appears to be even more important in repeated prisoner's dilemma like games that they are in one shot trust games, since they affect the level of perceived strategic risk, modifying the critical discount factor required to sustain the illegal agreement.

8.3 Experimental Design

In our experiment, each subject represented a firm and played in anonymous two-persons group a repeated duopoly game. In every stage game, the subjects had to take three types of decisions. First, the subjects had to choose whether or not they wanted to form a cartel by discussing prices. Second, they had to choose a price in a discrete Bertrand price game with differentiated goods. Third, the subjects could choose to self report cartels to a competition authority. The attractiveness of this latter opportunity depended on the details of the antitrust law enforcement institution - the treatment variables of our experiment.

8.3.1 The Bertrand game

In each period, the subjects had to choose a price from the choice set $\{0, 1, ..., 11, 12\}$. Their payoff depended on their own price choice and on the price chosen by their competitor and were reported in a payoff table distributed to the subjects. This table indicated a subjects' profits depending on its own price choice and the price chosen by its competitor (see figure 8.1) and was derived from the following standard linear Bertrand game. (The details of the Bertrand game were not described to the subjects.)

		your competitor's price												
		0	1	2	3	4	5	6	7	8	9	10	11	12
your price	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	29	38	47	56	64	68	68	68	68	68	68	68	68
	2	36	53	71	89	107	124	128	128	128	128	128	128	128
	3	20	47	73	100	127	153	180	180	180	180	180	180	180
	4	0	18	53	89	124	160	196	224	224	224	224	224	224
	5	0	0	11	56	100	144	189	233	260	260	260	260	260
	6	0	0	0	0	53	107	160	213	267	288	288	288	288
	7	0	0	0	0	0	47	109	171	233	296	308	308	308
	8	0	0	0	0	0	0	36	107	178	249	320	320	320
	9	0	0	0	0	0	0	0	20	100	180	260	324	324
	10	0	0	0	0	0	0	0	0	0	89	178	267	320
	11	0	0	0	0	0	0	0	0	0	0	73	171	269
	12	0	0	0	0	0	0	0	0	0	0	0	53	160

Figure 8.1: *Payoff table*

The demand function for each firm i was given by:

$$q_i(p_i, p_j) = \frac{a}{1+\gamma} - \frac{1}{1-\gamma^2}p_i + \frac{\gamma}{1-\gamma^2}p_j$$

where p_i (p_j) is the price chosen by firm i (competitor j), a is a parameter accounting for the market size and $\gamma \in [0,1)$ denotes the degree of substitutability between the two firms' products. Each firm faced a constant marginal cost, c, and had no fixed costs. The profit function, $\pi_i(p_i, p_j)$, was thus given by

$$\pi_i(p_i, p_j) = (p_i - c)q_i.$$

In our experimental setup, we chose $a = 36$, $c = 0$ and $\gamma = 4/5$ and restricted the subjects' choice set to $\{0, 2, ..., 22, 24\}$. These parameters yield the payoff table distributed to each subject. To simplify the table we also relabeled each price by dividing it by 2 and rounded the payoffs to the closest integer. In the unique Bertrand equilibrium, both firms charge a price equal to 3 yielding per firm profits of 100. The monopoly price (charged by both firms) is 9, yielding profits of 180. Note also that a firm would earn 296 by unilaterally and optimally undercutting the monopoly price, i.e. by charging a price of 7. In this case the other (cheated upon) firm only earns a profit of 20. Similarly, there are gains from deviating unilaterally from other common prices than the monopoly price as well as associated losses for the cheated upon firm; in the range of prices in between the Bertrand price and the monopoly price, ie in the range $\{4, ..., 8\}$, these gains and losses are smaller than when a subject deviates unilaterally from the monopoly price.

8.3.2 Cartel formation

Throughout the experiment, the subjects could form cartels by discussing prices. At the beginning of every period, a communication window opened if and only if both subjects agreed to communicate. This communication stage, which is described in more detail below, was designed in such a way that it would result in a common price on which to cooperate. This agreed upon price was non-binding, however, and therefore each subject could cheat on the agreement by subsequently charging a price different from the agreed upon price.

Whenever two subjects chose to communicate, they were considered to have formed a cartel. In this case, the subjects risked to be fined as long as their cartel was not detected by the competition authority. This implied that two subjects could be fined in a period even if no communication took place in that specific period; for example, two subjects could be fined in a period in which they did not communicate if they communicated in the previous period and the competition authority did not detect the associated cartel in that period. Once a cartel was detected, however, it was considered to be dismantled and in subsequent periods, the former cartelists did not run any risk of being fined unless they communicated again.

8.3.3 Treatment variables

Whenever two subjects had formed a cartel, a competition authority could detect the cartel and convict its members for price fixing. Detection could happen in two ways. First, in every period, the competition authority detected cartels with an exogenous probability, α. If this happened, both cartel members had to pay an exogenous fine, F. Second, the cartel members could self-report the cartel, in which case the cartel members were convicted for price fixing with certainty. If this happened, the size of the fine depended on the details of the law enforcement institution.

We ran nine treatments of our game, adopting a *between subjects* design, so that every subject only played the game under a single treatment. The nine treatments differ in the specific type of antitrust law adopted (with or without leniency for those who report the cartel), in the probability of detection and in the size of the fine imposed to the detected cartels' members. The differences between the treatments are summarized in table 8.2.

Antitrust Policy. Our baseline treatment corresponds to a *laisser faire* regime and is denoted COMMUNICATION: in this treatment, $\alpha = F = 0$ so that forming a cartel by discussing prices is legal. To simplify the instructions and to eliminate irrelevant alternatives, subjects were not allowed to report cartels. In the five other treatments cartel members were

Antitrust Policy	**fine (F)**	**probability of detection (α)**	**report**	**report's effects**
ANTITRUST	200	0.10	Yes	pay the full fine
	1000	0.02		
	300	0.2		
	1000	0		
LENIENCY	200	0.10	Yes	no fine (half the fine if both report)
	1000	0.02		
	300	0.2		
	1000	0		
COMMUNICATION	0	0	No	–

Table 8.2: *Treatments*

allowed to report cartels in which they participated. The ANTITRUST treatments corresponds to traditional antitrust laws without any leniency program: in case a report took place, both cartel members (including the reporting one) had to pay the full fine F. The LENIENCY treatments corresponds to current antitrust laws embedded with a leniency program: in case the cartel was reported by one of the cartel members only, the reporting member paid no fine while the other one paid the full fine, F; if instead both cartel members reported the cartel simultaneously, both paid a reduced fine equal to $F/2$. Note that under Leniency treatments a player who decides to deviate from the agreement is always better off if he simultaneously reports the cartel. So, in principle, the introduction of Leniency Programs should tighten the incentive compatibility constraint, since deviating becomes less risky, thus more attractive. Leniency should also harshen strategic risk, because the cheated upon firm not only suffers for the exploitation, but also has to pay the fine for sure.

Probability of Detection and Size of the Fine We also vary the probability of detection and the size of the fine across treatments: in particular, per each of the two considered antitrust policies, we have two treatments with an expected fine of 20 – one with a high probability of detection ($\alpha = 0.10$) and a low fine ($F = 200$), the other in which, vice versa, the probability of detection is low ($\alpha = 0.02$) and the fine is high ($F = 1000$) – and one treatment in which the expected fine is higher: $\alpha = 0.2$ and $F = 300$.

A different mix of magnitude and probability of the fine affects the riskiness of the collusive outcomes, but, as discussed above, it is not obvious what kind of effect this could generate in terms of deterrence. For example, if agents are perfectly rational and risk neutral, and they do not react to strategic risk, their preferences should be only marginally affected by a change in the determinants of the fine which leaves the expected fine constant. Such a change, in fact, has no impact on the expected collusive profits and has at most a marginal effect on the profitability of a deviation from collusion. As suggested by Becker, if on the contrary agents are risk averse we should observe higher deterrence when the size of the fine is higher and the probability lower, whether leniency programs are present or not. In addition, under Leniency Programs an increase of the magnitude of the fine dramatically reduces the profit a firms obtains when "cheated upon", that is when their opponent deviates from the collusive agreement. As mentioned above, these profits play no role in the standard theory, since they do not affect the conditions for an agreement being supportable in equilibrium, but they do matter for strategic risk, because they enter the definition of the short run disincentive to cooperate. Moreover, all the behavioral biases affecting risk perception we enumerated above might play a role in determining the outcome of such changes in the components of the expected fine.

The experiment we present here was not specifically designed to test any of these theoretical predictions, but to investigate the effects of different legal settings in light of them. With the same exploratory aim, we designed two additional treatments –one with Leniency, one without it – in which the fine is high ($F = 1000$), but can be inflicted only in case of reporting because the probability of detection is set to be null ($\alpha = 0$): that is, the antitrust authority is not able to discover any cartel that is not reported by at least one of its members. Comparing the results of these treatments with those we get from the corresponding treatments where the size of the fine is the same ($F = 1000$) but the probability of detection is positive ($\alpha = 0.02$), we can study if a very small probability of detection is overweighted or underestimated, and we can also check for the role played by strategic risk in this setting. Indeed, if strategic risk did not affect players' decision, we should not observe any deterrence in the treatment with $\alpha = 0$.

8.3.4 Experiment's timing and rematching procedure

At the end of each period, subjects were rematched with the same competitor with a probability of 85%. With the remaining probability of 15%, all subjects were randomly matched into new pairs. When this happened, cartels formed within the previous match could not be fined anymore. The experiment lasted at least 20 rounds. From the 21st round on, we introduced a termination probability of 15%, while the probability of rematching was reduced to 0. Subjects were also informed that the game would have been stopped in case the experiment lasted for more than 2 hours and 30 minutes. This latter eventuality never took place.

This re-matching procedure had several advantages. First, the subjects were playing truly *infinitely repeated games* without problems associated with end effects. Second, each subject played several repeated games against different competitors, which allowed us to observe the subjects' behavior in a larger number of repeated games.

Before the experiment started, the subjects were paired with the same competitor for five practice periods. Participants were informed that during these practice periods, they were paired with different competitors than those that they faced in the first period of the 'true' (i.e. remunerated) experiment. They were also told that profits realized during the trial periods were not to affect their earnings from the experiment.

8.3.5 The timing of the stage game

In the ANTITRUST and LENIENCY treatments, a stage game consisted of 7 steps (see figure 8.2). In the COMMUNICATION treatment steps 4,5 and 6 were skipped.

Figure 8.2: *Stage game*

We will now describe more in details each single step.

Step 1: Communication decision. Each subject was asked whether or not he wished to communicate with his competitor. If both subjects pushed on the yes button within 15 seconds, the game proceeded to step 2. Otherwise the two subjects had to wait for an additional 30 seconds before pricing decisions were taken in Step 3. In all periods, subjects were also informed whether they were matched with the same opponent as in the previous round or if a re-match had taken place.

Step 2: Communication. If both subjects decided to communicate in step 1, a window appeared on their computer screen asking them to simultaneously state a minimum acceptable price in the range $\{0, ..., 12\}$. When both of them had chosen a price, they entered a second round of price negotiations, in which they could choose a price from the new range $\{p_{min}, ..., 12\}$, where p_{min} was defined as the minimum among the two prices selected in the previous negotiation round. This procedure went on until 30 seconds had passed. The resulting minimum price p_{min} was referred to as the agreed upon price.

Step 3: Pricing. Each subject had to choose his price from the choice set $\{0, ..., 12\}$. Possible price agreements reached in step 2 were not binding. The subjects were informed that if they failed to choose a price within 30 seconds, then their default price would be so high that their profits became 0.

Step 4: First Reporting Decision. If communication took place in the current period or in one of the previous periods and had not yet been discovered by the competition authority, subjects had a first opportunity to report the cartel.

Step 5: Market prices and second reporting decision. Subjects were informed about the prices set by their opponent, their own profits and the profits of their competitor, gross of the possible fine. If communication had taken place in the current period or in one of the previous periods and had not not yet discovered by the competition authority and nobody had reported it in step 4, subjects had again the opportunity to report the cartel. The crucial difference between this second reporting opportunity and the first one is that the subjects knew the price chosen by their competitors.

Step 6: Detection. If communication took place in the current period or in one of the previous periods and had not yet been discovered or reported in steps 4 or 5, the competition authority discovered the cartel with probability α.

Step 7: Summary of the current period. At the end of each period, all the relevant information about the stage game are displayed: agreed upon price (if any), prices chosen by the two players, possible fines and net profits. In case players were fined, they were also told how many players reported.

Note that with our experimental setup subjects have two opportunities to report the cartel: first at step 4, right after having set their price, then again at step 5, after having been informed about the price chosen by their opponent. In our design, reporting can thus be used for two different purposes: (i) deviating subjects may report to get protection against prosecution and (ii) cheated upon subjects may report to punish their opponents, if they have not reported before.

8.3.6 Measuring risk aversion

We needed also a measure of risk aversion, to check for the effects it has on subjects' decision to communicate; due to the length of our main game, though, we could not adopt the – now standard – ten paired lotteries choice proposed by Holt and Laury (2002) [51], which is too time consuming, and we chose a shorter procedure, which provided us with a less precise but still reliable proxy.

At the end of the main game, each of the subjects was presented the following situation: given an initial endowment of 25 Euro they were asked to choose how much to keep and how much to invest into a risky project, yielding a return equal to 2.5 with 50% probability, and a return equal to 0 otherwise. After all the answers had been collected, a coin was tossed to determine the outcome of the risky project and only one of the subjects was randomly drawn to be paid according to his choice. It was made clear to the subjects that their choice and earnings in this second game could not affect in any way the profit they had made in the previous game.

Note that the initial endowment was chosen so that the amount of money at stake had approximately the same magnitude than the average cumulated profit in the main game: the amount of money invested should then be a reliable proxy of the degree of risk aversion displayed by the subjects when playing the Bertrand game.

8.3.7 Experimental procedure

Our experiment took place in May 2007 at Tor Vergata University (Rome, Italy) [4]. Session lasted on average 2 hours, including instructions and payment. We ran all the eight treatments and the investment game to check for risk attitude, involving 282 students in total. The average payment in the main game was equal to 23.60 €, with a maximum of 34€and a minimum of 11€, while the average payout for the investment game was 30.33€, with a minimum of 0 and a maximum of 62.5€.
For some of the results, we will present also data collected within the same experimental project, in March and April 2007 at the Stockholm School of Economics (Sweden). In Stockholm we did not run the investment game, while the Bertrand game was exactly alike the one we did in Rome. We ran only 5 treatments in Stockholm, namely: Communication plus the two Leniency and two Antitrust treatments in which the expected fine is equal to 20. 78 students were involved, in all. The average payment in Stockholm SEK 248[5], with a minimum of 130 SEK and a maximum of 330 SEK.

The experiment was computerized, and the programs were written with z-tree [35]. At the beginning of each session, subjects were welcomed in the lab and seated, each in front of a computer. When all subjects were ready, a printed version of the instructions and the profit table was distributed to them. Instructions were read aloud to ensure common knowledge of the rules of the game. The subjects were then asked to read the instructions on their own and ask questions, which were answered privately. When everybody had read the instructions and there were no more questions (which always happened after about fifteen minutes), each subject was randomly matched with another subject for the five practice rounds. After the practice rounds, participants had a last opportunity to ask questions about the rules of the game. Again, they were answered privately. Then they were randomly rematched into new

[4] Treatment Antitrust with $\alpha = 0$ and $F = 1000$ was run in an additional session, taking place at Tor Vergata University in December 2007. Students having taken part to previous sessions were not admitted.

[5] At the time of the experiment, 1 SEK=0.109 Euro

pairs and the real game started.

At the end of each session, the subjects were paid privately in cash. The subjects started with an initial endowment of 1000 points in order to reduce the likelihood of bankruptcy. At the end of the experiment the subjects were paid an amount equal to their cumulated earnings (including the initial endowment) plus a show up fee of 50 SEK in Stockholm and 7 Euros in Rome. The conversion rates were 20 points for 1 SEK in Stockholm and 200 points for 1 Euro in Rome.

8.4 Results

In this section, we will first present some aggregate results: we will briefly analyze the data collected through the investment game, and we will compare the average rate of communication in the different treatments.
We will then study what are the drivers of the subjects' decision to communicate, according to the results we got from a logit regression.
In the last part we will focus on the effects of conviction on players' behavior, analyzing how their willingness to communicate changes after they had got fined.

8.4.1 A proxy for risk aversion

Figure 8.3 displays the distribution of choices in the investment game: we find that more than 20% of the players are risk neutral or risk lover, which is in line with Holt and Laury (2002)'s findings; consistently with most empirical and experimental findings (see Eckel and Grossman, in press [32]) we also observe that women invested significantly less than men: the correlation between gender and investment is 21.65%, (significant at the 0.1%).

8.4.2 The decision to communicate under different treatments

Here we present an overview of our results about how the legal framework affects the individual decision of taking part in a cartel, when this

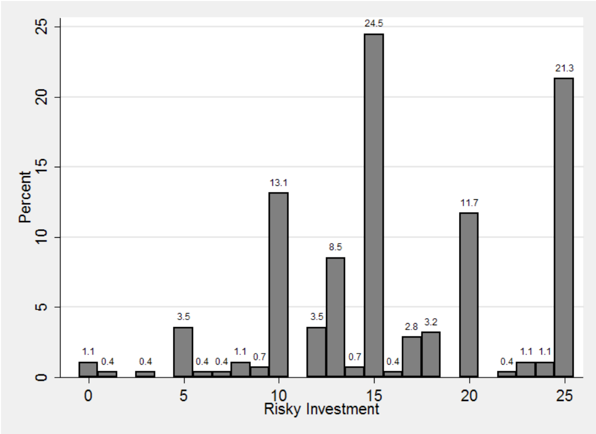


Figure 8.3: *Choice distribution in the investment game.*

Table 8.3: *Communication decision rates under different treatments*

α	F	**Antitrust**	**Leniency**
0.1	200	0.59	0.34
0.02	1000	0.38	0.25
0.2	300	0.45	0.43
0	1000	0.54	0.28
Communication		0.78	

Data collected in Rome. 5026 observations in total.

choice is illegal and risky. Notice that in our setting, communication is risky only when subjects are not currently cartel members, because the decision to communicate again when a cartel has already been established does not affect the probability of detection, nor the punishment imposed in case of liability. For this reason, for Antitrust and Leniency treatments, we restrict our attention to the attempts of communicate made by subjects that are not already members of a cartel. Comparing the rate of communication decision observed under the six Antitrust and Leniency treatments with those obtained in the benchmark treatment, Communication, we can evaluate the success of different legal frameworks in terms of ex ante deterrence, that is the main objective of Antitrust policies. A first look at these data leads to some preliminary observations:

- A large increment in the *actual* fine increases deterrence
- A large increment in the *expected* fine does not increase deterrence
- Given α and F, deterrence is higher under leniency
- When $\alpha = 0$ deterrence does not drop under leniency, but it does under antitrust.

To assess the significance of these results and to individuate the drivers of the communication decision, we need to study our data more in detail, taking into account some technical aspects of our dataset that make the econometric analysis less straightforward, as we shall explain in next paragraph.

8.4.3 Empirical methodology

A critical point in our analysis is how to control for repeated observations of the same subject or the same duopoly, when testing the significance of the observed differences across treatments. Given the rematching procedure we adopted, we need to account for correlation between two observations from the same individual, as well as correlation between two observations from different individuals who belong to the same duopoly. Moreover, since the experiment was run in two different cities, when we pool together the data gathered in Rome and Stockholm we also have to control for the possible correlation among observations collected in the same city. To this purpose, we adopted multilevel random effect models.

Since in our experiment a subject may take part in more than one duopoly during the game, the random effects at the subject level and at the duopoly level are not nested, which makes it difficult to estimate a model with a random effect at the duopoly level and a random effect at the subject level at the same time.

To overcome this complicacy, we hypothesized the presence of a random effect for every subject within any particular match (which accounts for the correlation among observations pertaining to the same match), nested with a random effect for every subject across different matches, which is in turn nested with a random effect at the city level.

To analyze data collected both in Rome and in Stockholm, we adopt a four-levels random intercept logit model of the following form:

$$CommDec_{hijk} = \mathbf{x}_{hijk}\beta + \eta^{(2)}_{ijk} + \eta^{(3)}_{jk} + \eta^{(4)}_{k}$$

where *h*, *i* , *j* and *k* are indices for measurement occasions, subjects in matches, subjects across matches and cities, respectively. $CommDec_{hijk}$ represents the $h-$th communication decision of subject *j* in match *i*, and in city *k*. $\mathbf{x}_{hijk}$ is a vector of explanatory variables (including the constant), with fixed regression coefficients β; $\eta^{(2)}_{ijk}$ represents the random intercept for subject *j* in match *i*, and in city *k* (second level), $\eta^{(3)}_{jk}$ represents the random intercept for subject *j* in city *k* (third level) and $\eta^{(4)}_{k}$ represents the random intercept for city *k* (fourth level). Random intercepts are assumed to be independently normally distributed, with a variance that is estimated through our regression.

When comparing observations collected in a single city, we adopt a model which is analogous to the previously described one, but without the last level.

To estimate our model used GLLAMM [6], a software specifically designed to provide a maximum likelihood framework for models with unobserved components, such as multilevel models, certain latent variable models, panel data models, or models with common factors.

8.4.4 Drivers of the decision to communicate

In this section we present the results of a logit regression we ran to assess which are the most important factors affecting subjects' decision to communicate. For these results, we only consider the data we collected in Rome, therefore we will adopt a three levels random intercept logit

[6] see Rabe-Hesketh and Skrondal, 2004 [94] and http://www.gllamm.org

model of the form presented above. Table 8.4 presents the results of this regression[7].

__Table 8.4:__ Results of the logit regression.

	Coefficient	Std. Err.
$A_{0.1,200}$	-0.425	0.488
$A_{0.02,1000}$	-1.890***	0.566
$A_{0.2,300}$	-1.704***	0.500
$A_{0,1000}$	-0.377	0.506
$L_{0.1,200}$	-2.340***	0.387
$L_{0.02,1000}$	-3.198***	0.529
$L_{0.2,300}$	-1.764***	0.475
$L_{0,1000}$	-2.631***	0.491
Paid fine (/1000)	-1.048***	0.216
Frequency of detection	-0.073	0.598
Cumulated earning (/1000) -0.002	0.083	
Investment (/25)	1.216**	0.557
Constant	0.927*	0.517
LogLikelihood	-2666.716	
#obs.	5398	

Data collected in Rome. 5026 observations in total.

The dependent variable is the decision to communicate: as mentioned before, since we are interested in deterrence of cartel formation, we do not consider in this regression the decisions taken by subjects already members of an existing cartel.
The independent variables are:

- 8 dummy variables, one for each treatment (communication is the benchmark)

- the total fine paid by the subject up to the period in which he takes the decision.

[7] Note: In this table as well as in the following results, the symbols ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively.

- the frequency of detection observed by the subject, measured as the ratio between the number of times a cartel he belonged to was detected by the Authority (without it being reported) and the number of periods in which the subject had taken part in a cartel.
- the subject's cumulated earnings.
- the amount of money put into the risky asset in the "investment game"

Note that the cumulated fine and cumulated earnings have a much higher magnitude than the other regressors; for this reason, we divided those numbers by 1000, so that all the variables had approximately the same scale. For similar reasons, the sum chosen by the subject in the investment game enters the equation in terms of ratios of the total amount of money available, namely 25.

The regression's results substantially confirm our preliminary observations.

Result 1: the size of the *actual* fine matters. Deterrence is significantly higher when *actual fine* is higher, the expected fine being the same. According to one-sided z-tests, we have:

$$A_{0.02,1000} <^{**} A_{0.1,200} \text{ and } L_{0.02,1000} <^{*} L_{0.1,200}$$

Result 2: the size of the *expected* fine has no direct effect on deterrence. A higher expected fine does not necessarily imply higher deterrence:

$$A_{0.1,200} >^{**} A_{0.2,300} \text{ and } L_{0.1,200} \approx L_{0.2,300}$$

but

$$A_{0.02,1000} \approx A_{0.2,300} \text{ and } L_{0.02,1000} <^{***} L_{0.2,300}$$

These first two results seem to confirm Becker's suggestion that it is possible to achieve higher deterrence while decreasing prosecution costs by increasing the size of the fine and reducing the effort spent in investigation.

Result 3: deterrence is generally higher under Leniency. Only for the two treatments with higher expected fine ($\alpha = 0.2$ and $F = 300$) the effect does not seem to be significant:

$$L_{0.1,200} <^{***} A_{0.1,200}$$

$$L_{0.02,1000} <^{**} A_{0.02,1000}$$

$$L_{0.2,300} \approx A_{0.2,300} L_{0,1000} <^{***} A_{0,1000}$$

Result 4: deterrence when $\alpha = 0$. As mentioned before, according to the standard theory we should not observe any deterrence when the probability of detection is null. Indeed, we notice that the coefficient for $A_{0,1000}$ is not significantly different from zero. Remarkably, a different result holds for Leniency treatments: a small probability of detection seems to play no role in deterrence under Leniency, when the fine is high enough:

$$L_{0.02,1000} \approx L_{0,1000}$$

Currently, some concern has been expressed about the contingency that the many leniency applications keep the agency busy with prosecution, to the detriment of investigation; thus, the probability that a cartel is detected because of the autonomous investigation by the authority would decrease (lower α.)
According to our results, this should not be a serious problem: we observe that deterrence remains high even if $\alpha = 0$, provided that the fine is high enough.

Now we would like to examine these results in light of the theories about players' behavior mentioned in section 8.2, to see if and to which extent these theories are supported by the evidence we collected.

Risk aversion. Our regression shows that subjects who chose to put more money in the risky lottery of the investment game are also more incline to communicate, which is in line with the ideas discussed in section 8.2.2. There, we have seen that risk aversion can affect deterrence in at least two ways: first, it increases the perceived dis-utility connected to the risk of conviction; second, it can also worsen the perceived risk of being betrayed by the other player, in case a cartel is established. A certain degree of risk aversion in some of the players is also a possible reason why deterrence is higher in treatments $A_{0.02,1000}$ and $L_{0.02,1000}$ than in treatments $A_{0.1,200}$ and $L_{0.1,200}$, respectively, even if the expected fine remains constant.

Strategic risk. We do not observe any significant difference between the levels of deterrence under Leniency when the fine is high ($F = 1000$) and the probability of detection is low or null. This fact supports the idea that in presence of leniency programs it must be the risk of being cheated upon by the other cartelist – i.e. strategic risk – and not the risk of being liable that determines deterrence. In treatments Antitrust, on the other hand, we observe that deterrence when the probability of detection is low but positive is significantly higher than when this probability is null. Moreover, the estimated coefficient for treatment dummy $A_{0,1000}$ is not significantly different from zero. This is not in contrast with the theory of strategic risk. In fact, both findings suggest that it is the risk of detection and punishment – and not the risk of being betrayed by other cartelists – that discourage players from colluding, when no leniency is granted to those who report the cartel. In fact, in this case reporting the cartel when deviating from the collusive agreement is not a dominant strategy, so it is possible that the perceived strategic risk is lower and may even be negligible when there is no risk of detection.

Perception of small probabilities. We mentioned above that according to the research developed by Tversky and Kahneman [67], the perception of very small probabilities may have ambiguous outcomes: it is possible that they are overemphasized or even approximated to zero, depending on the context and on the individual characteristics of the subject. The significant difference between the estimated coefficients for $A_{0,1000}$ and $A_{0.02,1000}$ seems to imply that in general, in our game a very small probability of detection is not disregarded. In a sense, this is another element supporting the importance of strategic risk in the situation we depicted. In fact, the difference in deterrence disappears under the two Leniency treatments with fine equal to 1000. If players do not approximate a probability of 2% to zero, then this proability must be disregarded because other factors predominate, and among them strategic risk appears to be one of the most plausible.

Availability heuristic. The hypothesis that people's perception of a risk is based not only on its actual probability, but also on its vividness and emotional impact is validated by our data. According to our regression, the sum of the fines paid by a subject in previous periods has a significant and substantial negative effect on his willingness to communicate, meaning that subjects who have paid a very high fine, but also those who have repeatedly paid a lower fine, are less incline to collude again. This is in line with the findings of Fishman and Pope (2006) [36] about punishment-induced deterrence and with the idea that the experience of the penalty affects subjects' willingness to commit the crime, the more the harsher is the penalty, or more generally the stronger is its memory.
The sum of the fines paid appears to be the only factor affecting players behavior throughout the game, thus introducing some dynamics in their choice pattern. Communication decision does not seem to be affected by the player's cumulated earnings, which is obviously highly correlated with the number of periods elapsed since the beginning of the game and of the match. This seems to rule out an endowment effect, and any learning effects other than the one deriving from the experience of punishment.

8.4.5 Post-conviction Behavior

In this section we will describe how players modify their decision to communicate in the periods following conviction. First, we introduce a distinction between two categories of conviction, according to the outcome they generate for the players in terms of payoffs: we will say that conviction has a

- *symmetric outcome*, if it hurts both players (approximately) in the same way. For example, this is the case when a cartel is detected by the Authority, but also when reporting is used as a punishment device under Leniency. In this last case, only one of the players deviated from the collusive agreement, thus getting higher profits, but the cartel was reported only by the other player: both players then obtain low profits, because one of them was cheated upon, the other one got fined.

- *asymmetric outcome*, if one of the cartel's members got hurt more than the other, as when only one player deviate and simultaneously reports the cartel.

In this section we use data collected both in Rome and in Stockholm, since we are not going to use the information about subjects risk aversion. We will also pool together the data across treatments.
First, we observe that conviction has a symmetric outcome 94.12% of the time under Antitrust treatments, and only 37.90% of the times under Leniency. This is mainly due to the fact that deviators often report the cartel to protect themselves from fines under Leniency, but not under Antitrust: indeed, under Leniency treatments, players who undercut the agreed upon price also reported the cartel in 62.38% of the cases, while under Antitrust treatments this percentage drops to 4.95%. So, if conviction with asymmetric outcomes discourages communication more, then we would have at least a partial explanation of why we generally observe more deterrence under Leniency treatments.

Figure 8.4 displays the percentage of convicted agents (vertical axis) who chose to communicate again in the five periods following conviction (horizontal axis), separately for symmetric and asymmetric conviction outcomes. We observe that, as a consequence of asymmetric conviction, most of the subjects decide not to communicate anymore with the same competitor, while the about one subject out of two chose to communicate again when conviction'outcome was symmetric.
To check whether this finding is significant and robust to other factors,

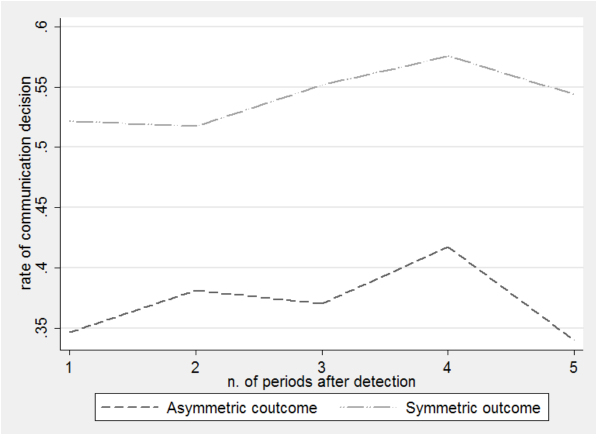


Figure 8.4: *Post-conviction communication decision.*

we ran the following four levels random intercept logit regression:

$$CommDec_{hijk} = \beta_0 + \beta_1 CumEarnings_{hijk} + \beta_2 Fine_{hijk} + \\ + \beta_3 Symm_{hijk} + \beta_4 SymmXFine_{hijk} + \eta^{(2)}_{ijk} + \eta^{(3)}_{jk} + \eta^{(4)}_{k}$$

where, the dependent variable is communication decision in the period immediately following conviction, $CumEarnings$ represents the cumulated earnings of the subject, $Symm$ is a dummy variable equal to 1 when the outcome of conviction was symmetric, $Fine$ measures the fine actually paid by the convicted subject, and $SymmXFine$ represents the interaction between these two factors. Finally, as explained before, $\eta^{(2)}_{ijk}$ represents the random intercept for subject j in match i, and in city k (second level), $\eta^{(3)}_{jk}$ represents the random intercept for subject j in city k (third level) and $\eta^{(4)}_{k}$ represents the random intercept for city k (fourth level).

Results from this regression, displayed in table 8.5, show that symmetry of conviction outcomes positively and significantly affects subjects' decision to communicate after conviction. Columns three, four and five of the table present the estimation results for the reduced models obtained progressively deleting the factors the turned out not to be significant,

Table 8.5: *Regression's results*

	Coefficient (s.e.)	Coefficient (s.e.)	Coefficient (s.e.)	Coefficient (s.e.)
Cumulated Earnings	0.144	0.140		
	(0.127)	(0.126)		
Fine	1.017*	0.778*	0.650	
	(0.548)	(0.431)	(0.412)	
Symmetry	0.780***	0.655***	0.668***	0.684***
	(0.265)	(0.197)	(0.196)	(0.193)
SymmXFine	-0.545			
	(0.768)			
Constant	-1.241***	-1.182***	-0.886***	-0.723***
	(0.347)	(0.337)	(0.200)	(0.165)
LogLikelihood	-452.397	-452.648	-453.268	-454.551
#obs.	692	692	692	692

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

according to loglikelihood-ratio tests at the standard 5% significance level.

Noticeably, even when the outcome of conviction is symmetric, only half of the convicted subjects decide to communicate again. Part of this effect is due to the fact that most of the collusive agreements had been broken before they were detected: more precisely, if we restrict our attention to the cases in which conviction had a symmetric outcome, we still observe that in 76.22% of the detected cartels at least one member undercut the agreed upon price before detection took place. Some players therefore probabily decided non to communicate again because of a lack of trust that had already emerged before conviction.
On the other hand, table 8.6 shows that even if the cartel had not been previously broken, still about 30% of the subjects chose not to communicate again. The number of observations is too small to make sound inference; nonetheless we believe that the behavior of these players could be explained with reference to availability heuristic: indeed, they behave as if the punishment recently experienced affected their perception of the level of risk of liability.

Table 8.6: *Communication rate when conviction outcome is symmetric*

	rate of comm.	N. obs.
broken	46.13%	310
not broken	70.59%	102

8.5 Conclusion

Our experiment shows that strategic risk, availability heuristic and saliency bias have important effects on cartel deterrence, though in general they are not taken into account in most of the theoretical analyses of law enforcement. This result, obtained through an exploratory experiment – probably the first one in this field – calls for further, more specific experimental tests of the observed effects. A deeper and wider experimental evidence could then support the development of a theoretical analysis of cartel deterrence which incorporates strategic risk and the behavioral effects that have non-negligible effects according to our findings.

It would be also interesting to check whether our results are robust to a change in the framing of the experiment, so to see if our conclusions can be extended to other kinds of organized crime, such as corruption, fraud, auditor-manager collusion, and corporate crime in general.

The interplay of rational considerations and behavioral biases in shaping deterrence of criminal activities deserves our attention and is a promising area for future research. While substantially more evidence is needed before drawing definitive theoretical conclusions, the glimpse that our study offers hopefully is a useful first step which will open the way for a rather new branch of experimental studies.

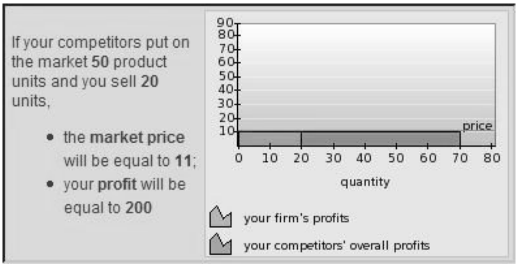
Part III

Appendices and Bibliography

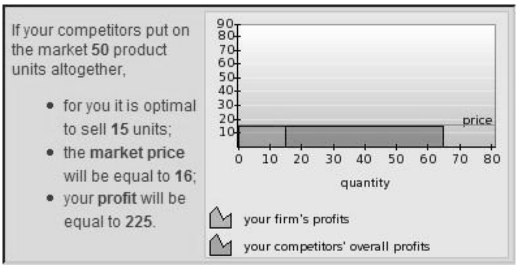
Appendix A

A.1 Experiment 1: Figures

Profit Calculator



(a) example 1



(b) example 2

Figure A.1: *Profit calculator*

Information about past rounds

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

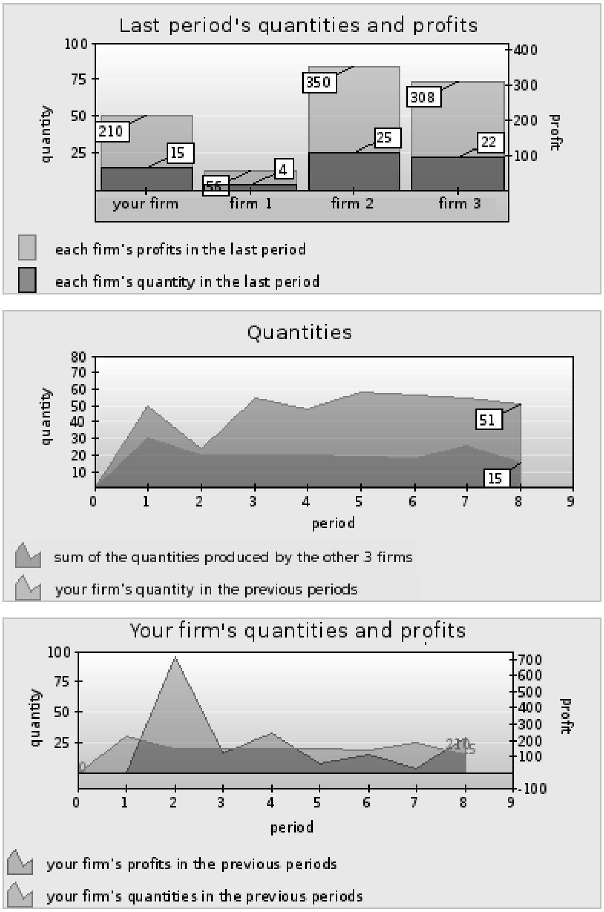


Figure A.2: *Information about the past periods displayed to the subjects at each round.*

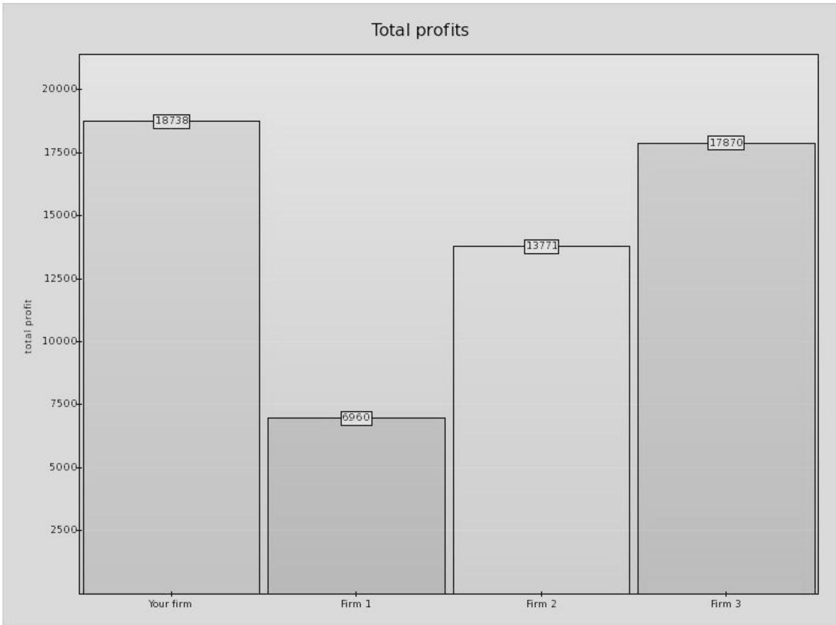


Figure A.3: *Plot displayed to each subject after the last period, at the end of the game.*

Graphical Interface

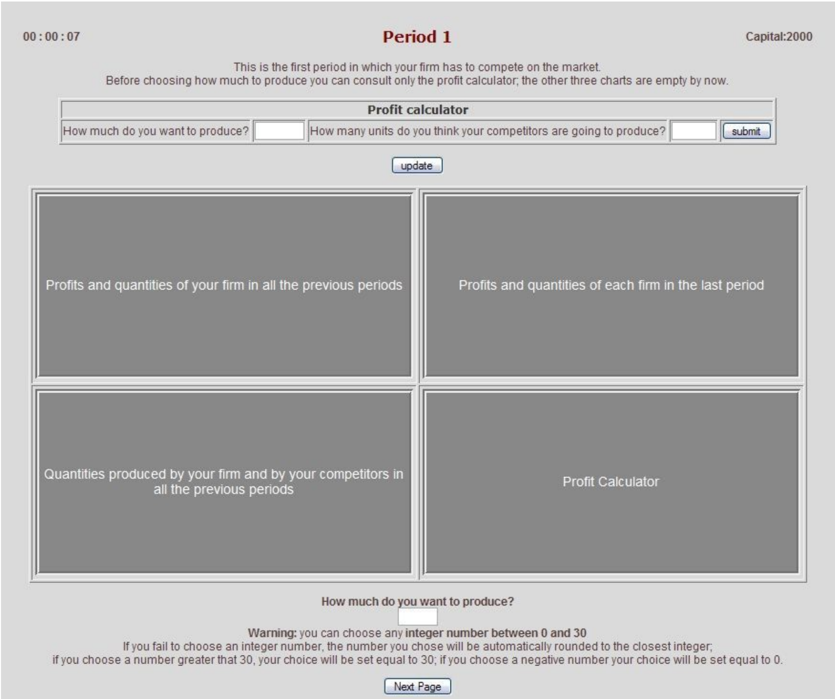


Figure A.4: *Graphical interface through which the game is presented to the subjects.*

A.2 Experiment 1: Instructions

Thank you for taking part in our experiment.

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

Your task

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

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

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

Price, costs and profits

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

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

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

Earnings and Payment

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

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

100 points correspond to 1 SEK.

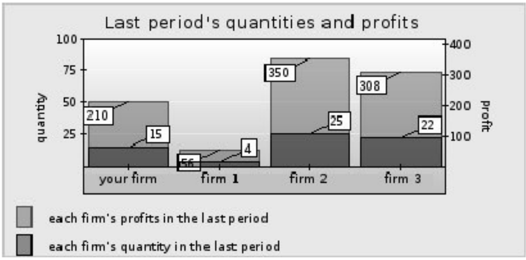
At the end of the game, your capital will be converted in SEK, and will be payed to you privately.
In addition, you will be payed a **show up fee** of 50 SEK.

Information at your disposal

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

Examples of these plots are presented in what follows.

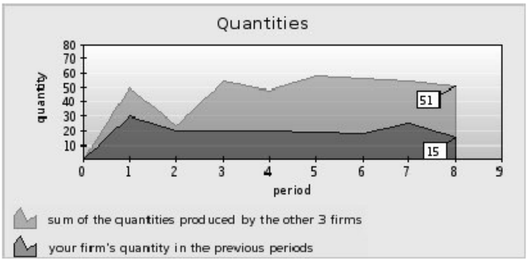
Profits and quantities of each firm in the last period



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

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

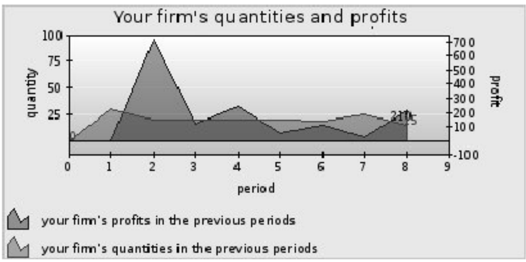
Quantities produced by your firm and by your competitors in all the previous periods



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

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

Profits and quantities of your firm in all the previous periods



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

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

The profit calculator - 1

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

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

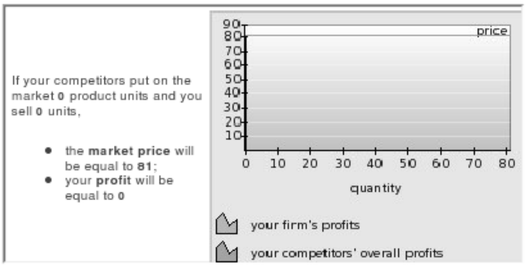
First function of the profit calculator

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

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

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

Update



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

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

OK

The profit calculator - 2

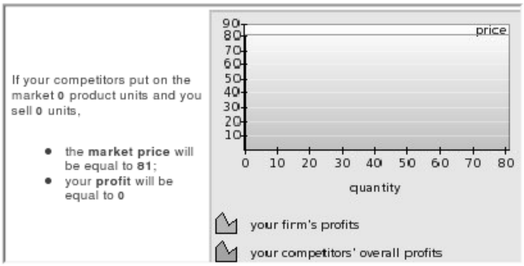
Second function of the profit calculator

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

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

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

Update



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

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

OK

How to consult the plots

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

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

Update

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

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

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

OK

A.3 Experiment 2: Graphical Interface

Figure A.5 represents the graphical interface of the second experiment.

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?

first box on the left: Profits in the previous period
your profit: 252
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. [OK]

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]

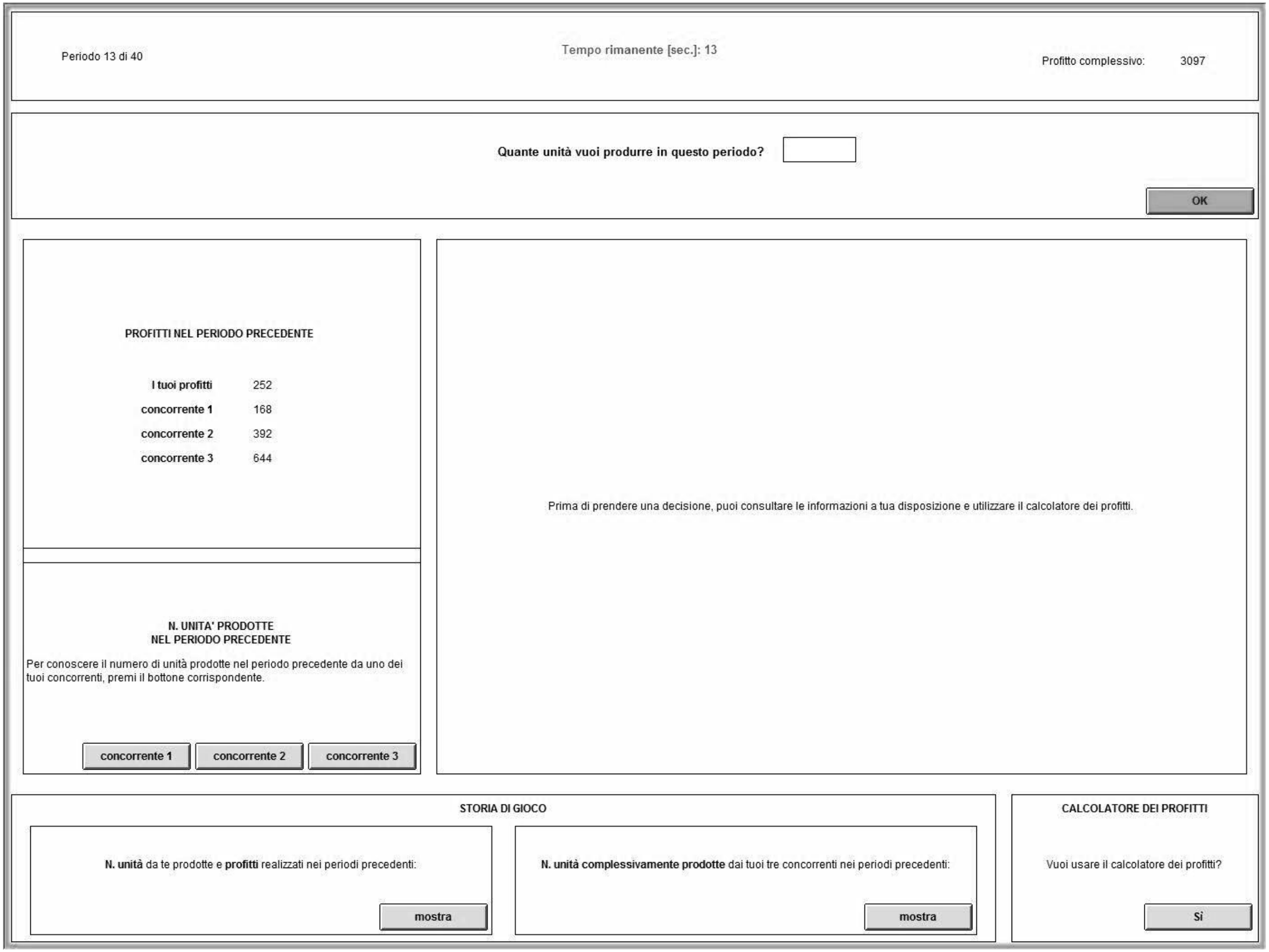


Figure A.5: *Graphical interface of the second experiment.*

A.4 Experiment 2: 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€for your participation. On top of that you can earn up to 20 €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 then 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 in 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 among the participants to 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 in **40** consecutive **periods**. In every period, you will be asked to choose how many units to produce (between 0 and 30), and the same will be done by your competitors. Your choices affect both your firm's profits and the ones 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 be always 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 €for your participation.

Information at your disposal. At the top of your computer screen you will 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. 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. 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 the 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 you 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 possibility of asking questions, always raising your hand.

When all the participants have completed their test, the real experiment will begin. The computer will randomly generate groups of four persons; every participant to 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 have to hand back the instructions.

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

Appendix B

Instructions for the Leniency1 treatment

Welcome to this experiment about decision making in a market. The experiment is expected to last for about 1 hour and 45 minutes. You will be paid a minimum of 50 SEK for your participation. On top of that you can earn more than 300 SEK 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 then have questions, raise your hand and you will be helped privately.

In summary, the situation you will face is the following. You and one other participant referred to as your competitor produce similar goods and sell them in a common market. As in most markets, the higher the price you charge, the more you earn on each sold good, but the fewer goods you sell. And, as in many markets, the lower the price charged by your competitor, the more customers he or she will take away from you and the less you will sell and earn. It is possible, however, to form a cartel with your competitor, that is, you will have the possibility to communicate and try to agree on prices at which to sell the goods. In reality, cartels are illegal and if the government discovers the cartel, cartel members are fined. In addition members of a cartel can always report it to the government. The same happens in this experiment. If you communicate to discuss prices, even if both of you do not report, there is still a chance that the 'government' discovers it and if this happens, you will have to pay a 'fine'. If you report, and if you are the only one to report, you will not pay any fine but your competitor will pay the full fine. Conversely, if only your competitor reports the cartel, you will pay the full fine and your competitor will not pay any fine. If instead both of you report the cartel you will both pay 50% of the fine.

Timing of the experiment

In this experiment you will be asked to make decisions in several periods. You will be paired with another participant for a sequence of periods. Such a sequence of periods is referred to as a match. You will never know with whom you have been matched in this experiment.

The length of a match is random. After each period, there is a probability of 85% that the match will continue for at least another period. So, for instance, if you have been paired with the same competitor for 2 periods, the probability that you will be paired with him or her a third period is 85%. If you have been paired with the same competitor for 9 periods, the probability that you will be paired with him or her a tenth period is also 85%.

Once a match ends, you will be paired with another participant for a new

match, unless 20 periods or more have passed. In this case the experiment ends. So, for instance, if 19 periods have passed, with a probability of 15% you are re-matched, that is you are paired with another participant. If 21 periods have passed, with a probability of 15% the experiment ends.

When you are re-matched you cannot be fined anymore for a cartel formed in your previous match with your previous competitor.

The experimental session is expected to last for about 1 hour and 45 minutes but its actual duration is uncertain; that depends on the realization of probabilities. For this reason, we will end the experimental session if it lasts more than 2 hours and 30 minutes.

Before the experiment starts, there will be 5 trial periods during which you will be paired with the same competitor. These trial periods will not affect your earnings. When the experiment starts, you will be paired with a new competitor.

Prices and Profits

In each period you choose the price of your product. Your price as well as the price chosen by your competitor determines the quantity that you will sell.

The higher your price, the more you earn on each sold good, but the fewer goods you sell. Therefore your price has two opposing effects on your profit. On the one hand, an increase in your price may increase your profit, since each good that you sell will earn you more money. On the other hand, an increase in your price may decrease your profit, since you will sell less.

Furthermore, the higher the price of your competitor, the more you will sell. As a result, your profits increase if your competitor chooses a higher price.

To make things easy, we have constructed a profit table. This table is added to the instructions. Have a look at this table now. Your own prices are indicated next to the rows and the prices of your competitor are indicated above the columns. If you want to know your profit if, for example, your competitor's price is 5 and your price is 4, then you first move to the right until you find the column with 5 above it, and then you move down until you reach the row which has 4 on the left of it. You can read that your profit is 160 points in that case.

Your competitor has received an identical table. Therefore you can also use the table to learn your competitor's profit by inverting your roles. That is, read the price of your competitor next to the rows and your price above the columns. In the previous example where your price is equal to 4 and your competitor's price is equal to 5, it follows that your competitor's profit is 100 points.

Note that if your and your competitor's prices are equal, then your profits are also equal and are indicated in one of the cells along the table's diagonal. For example, if your price and the price of your competitor are equal to 1, then your profit and the profit of your competitor is equal to 38 points. If both you and your competitor increase your price by 1 point to 2, then your profit and the profit of your competitor becomes equal to 71.

Note also that if your competitor's price is sufficiently low relative to your price, then your profit is equal to 0. The reason is that no consumer buys your good, since it is too expensive relative to your competitor's good.

Fines

In every period, you and your competitor will be given the opportunity to communicate and discuss prices. If both of you agree to communicate, you will be considered to have formed a cartel, and then you might have to pay a fine F. This fine is given by:

F = 200 points

You can be fined in two ways. First, you and your competitor will have the opportunity to report the cartel. If you are the only one to report the cartel, you will not pay any fine but your competitor will pay the full fine, that is 200 points. Conversely, if only your competitor reports the cartel and you do not, then you will have to pay the full fine equal to 200 points and your competitor will not pay any fine. Finally, if both of you report the cartel, you will both pay 50% of the fine, that is 100 points.

Second, if neither you nor your competitor reports the cartel, the government discovers it with the following probability.

Probability of detection = 10%.

Note that you will run the risk of paying a fine as long as the cartel has not yet been discovered or reported. Thus you may pay a fine in a period even if no meeting takes place in that period. This happens if you had a meeting in some previous period which has not yet been discovered or reported.

Once a cartel is discovered or reported, you do not anymore run the risk of paying a fine in future periods, unless you and your competitor agree to communicate again.

Earnings

The number of points you earn in a period will be equal to your profit minus an eventual fine or plus an eventual reward. Note that because of the fine, your earnings may be negative in some periods. Your cumulated earnings, however, will never be allowed to become negative.

You will receive an initial endowment of 1000 points and, as the experiment proceeds, your and your competitor's decisions will determine your cumulated earnings. Note that 20 points are equal to 1 SEK. Your cumulated earnings will be privately paid to you in cash at the end of the session.

Decision making in a period

Next we describe in more detail how you make decisions in each period. A period is divided into 7 steps. Some steps will inform you about decisions that you and your competitor have made. In the other steps you and your competitor will have to make decisions. In these steps, there will be a counter indicating how many seconds are left before the experiment proceeds to the next step. If you fail to make a decision within the time limit, the computer will make a decision for you.

Step 1: Pairing information and price communication decision

Every period starts by informing you whether or not you will play against the same competitor as in the previous period.

Remember that if you are paired with a new competitor, you cannot be fined anymore for cartels that you formed with your previous competitors.

In this step you will also be asked if you want to communicate with your competitor to discuss prices. A communication screen will open only if BOTH you and your competitor choose the "YES" button within 15 seconds. Otherwise you will have to wait for an additional 30 seconds until pricing decisions starts in Step 3.

Step 2: Price communication

After the communication screen has opened, you can "discuss" prices by choosing a price out of the range $\{ 0, 1, 2, \dots , 12 \}$. In this way you can indicate to your competitor the minimum price that you find acceptable for both of you. When both of you have chosen a price, these two prices are displayed on the computer screen. You can then choose a new price but now this price should be greater or equal to the smaller of the two previously chosen prices. This procedure is repeated until 30 seconds have passed. The screen then displays the smaller of the two last chosen prices, which is referred to as the agreed-upon price. Note, however, that in the next step, neither you nor your competitor is forced to choose the agreed-upon price.

Step 3: Pricing decision

You and your competitor must choose one of the following prices: 0, 1, 2, . . . , 12. When you choose your price, your competitor will not observe your choice nor will you observe his or her price choice. This information is only revealed in Step 5. The experiment proceeds after 30 seconds have passed. If you fail to choose a price within 30 seconds, then your price is chosen so high that your profits will be 0.

The experiment proceeds to the first reporting decision in Step 4 if you communicated in Step 2 or if in previous periods you formed a cartel not yet discovered or reported. Otherwise you have to wait for 10 seconds until market prices are revealed in Step 5.

Step 4: First (secret) reporting decision

By choosing to push the "REPORT" button, you can report that you have been communicating in the past. As described above, if you are the only one to report, you will not pay the fine; the opposite happens if only your competitor reports; and if both of you report, you will both pay 50% of the fine.

If you do not wish to report, push instead the "DO NOT REPORT" button.

When you decide whether or not to report, your competitor will not observe your choice, nor will you observe his or her choice. This information is only revealed when market prices are revealed in Step 5.

If you do not reach a decision within 10 seconds, your default decision will be "DO NOT REPORT".

Step 5: Market prices and second reporting decision

In this step your and your competitor's prices and profits are displayed.

In case you have formed a cartel not yet discovered or reported, the screen will also display whether or not you or your competitor reported it in the first reporting step (Step 4). If not, you will get a new opportunity to report.

If you wish to report, push the "REPORT" button. If you do not wish to report, push instead the "DO NOT REPORT" button.

Figure B.1: *Timing of the stage game*

Again, if you are the only one to report, you will not pay the fine. On the contrary, f your competitor reports and you don't you will have to pay the fine and he will not. If both you and your competitor report, you will both pay 50% of the fine, that is 100 points.

Step 6: Detection probability

If this step is reached, you formed a cartel either in the current period or in previous periods. Furthermore the cartel has not yet been discovered or reported. The cartel can nevertheless be discovered. This happens with a probability of 10%. If the cartel is discovered, you and your competitor will have to pay the full fine of 200 points.

Step 7: Summary

In this step you learn the choices made in the previous steps: your and your competitor's price choices and profits, your eventual fine, your eventual reward and your earnings.

If you paid a fine in this period, you will also know whether your competitor reported the cartel or the government discovered it.

In case a cartel was detected or reported in this period, you will not run any risk of being fined in future periods, unless you and your competitor discuss prices again.

Step 7 will last for 20 seconds.

Period ending and ending of the experimental session

After Step 7, a new period starts unless 20 or more periods have passed and the 15% probability of pair dismantling takes place. In that case, the experiment ends.

The following time line summarizes the seven steps of each round.

Throughout the experiment, a table will keep track for you of the history with your current competitor. For each previous period played with your current competitor, this table will show your price and profit, your competitor's price and profit as well as your eventual fine.

Payments

At the end of the experiment, your earnings in points will be exchanged in SEK. In addition you will be paid the show up fee of 50 SEK.

Before being paid in private, you will be asked to answer a short questionnaire about the experiment and you will have to handle back the instructions.

Please read now carefully the instructions on your own. If you have questions, raise your hand and you will be answered privately.

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

Bibliography

[1] AKERLOF, G. Procrastination and Obedience. *The American Economic Review 81*, 2 (1991), 1–19.

[2] APESTEGUIA, J., DUFWENBERG, M., AND SELTEN, R. Blowing the Whistle. *Economic Theory 31*, 1 (2007), 143–166.

[3] APESTEGUIA, J., HUCK, S., AND OECHSSLER, J. Imitation—theory and experimental evidence. *Journal of Economic Theory 136*, 1 (2007), 217–235.

[4] AUBERT, C., REY, P., AND KOVACIC, W. The impact of leniency and whistle-blowing programs on cartels. *International Journal of Industrial Organization 24*, 6 (2006), 1241–1266.

[5] AUMANN, R. J. Rationality and Bounded Rationality. *Games and Economic Behavior 21*, 1-2 (October 1997), 2–14.

[6] AUMANN, R. J., AND SORIN, S. Cooperation and bounded recall. *Games and Economic Behavior 1*, 1 (March 1989), 5–39.

[7] AXELROD, R. *The Evolution of Cooperation*. Basic Books, New York, 1984.

[8] BAKER, J. The Case for Antitrust Enforcement. *The Journal of Economic Perspectives 17*, 4 (2003), 27–50.

[9] BECKER, G. Crime and Punishment: An Economic Approach. *The Journal of Political Economy 76*, 2 (1968), 169–217.

[10] BLONSKI, M., OCKENFELS, P., AND SPAGNOLO, G. Co-operation in Infinitely Repeated Games: Extending Theory and Experimental Evidence, 2007.

[11] BLONSKI, M., AND SPAGNOLO, G. Prisoners' Other Dilemma, June 2003.
`http://www.wiwi.uni-frankfurt.de/Professoren/blonski/papers/PodRES.pdf`.

[12] BLUME, A., AND ORTMANN, A. The effects of costless pre-play communication: Experimental evidence from games with Pareto-ranked equilibria. *Journal of Economic Theory 132*, 1 (2007), 274–290.

[13] BROWN, G. Iterative solution of games by fictitious play. *Activity Analysis of Production and Allocation 13* (1951), 374–376.

[14] BUCCIROSSI, P., AND SPAGNOLO, G. The Effects of leniency on Illegal Transactions: How (Not) to Fight Corruption. *Stockholm School of Economics WP 456* (2001).

[15] BUCCIROSSI, P., AND SPAGNOLO, G. Leniency policies and illegal transactions. *Journal of Public Economics 90*, 6-7 (2006), 1281–1297.

[16] BUCCIROSSI, P., AND SPAGNOLO, G. Optimal Fines in the Era of Whistleblowers: Should Price Fixers Still Go To Prison? In *The Political Economy of Antitrust*, V. Goshal and J. Stennek, Eds. Elsevier, 2007.

[17] BUCCIROSSI, P., AND SPAGNOLO, G. Corporate Governance and Cartel Deterrence. In *Issues in Competition Law and Policy*, W. Collins, Ed. American Bar Association, Antitrust Section, forthcoming.

[18] CAMERER, C. *Behavioral Game Theory: Experiments on Strategic Interaction*. Princeton University Press, Princeton, USA, 2002.

[19] CAMERER, C., AND HO, T. H. Experience-Weighted Attraction Learning in Coordination Games: Probability Rules, Heterogeneity, and Time-Variation. *Journal of Mathematical Psychology 42*, 2-3 (June 1998), 305–326.

[20] CAMERER, C., AND HO, T. H. Experienced-Weighted Attraction Learning in Normal Form Games. *Econometrica 67*, 4 (July 1999), 827–874.

[21] CAMERER, C., HO, T. H., AND CHONG, J. K. Sophisticated Experience-Weighted Attraction Learning and Strategic Teaching in Repeated Games. *Journal of Economic Theory 104* (2002), 137–188.

[22] CASON, T., AND GANGADHARAN, L. An Experimental Study of Compliance and Leverage in Auditing and Regulatory Enforcement. *Economic Inquiry 44*, 2 (2006), 352.

[23] CHEN, J., AND HARRINGTON JR, J. The Impact ofthe Corporate Leniency Program on Cartel Formation and the Cartel Price Path. In *The Political Economy of Antitrust*, V. Goshal and J. Stennek, Eds. Elsevier, 2007.

[24] CHO, I.-K. Introduction to learning and bounded rationality. *Journal of Economic Theory 124*, 2 (October 2005), 127–128.

[25] COSTA-GOMES, M., AND CRAWFORD, V. Cognition and Behavior in Two-Person Guessing Games: An Experimental Study. *The American Economic Review 96*, 5 (2006), 1737–1768.

[26] COSTA-GOMES, M., CRAWFORD, V. P., AND BROSETA, B. Cognition and Behavior in Normal-Form Games: An Experimental Study. *Econometrica 69*, 5 (September 2001), 1193–1235.

[27] CRANDALL, R., AND WINSTON, C. Does Antitrust Policy Improve Consumer Welfare? Assessing the Evidence. *The Journal of Economic Perspectives 17*, 4 (2003), 3–26.

[28] CRAWFORD, V. Look-ups as the Windows of the Strategic Soul: Studying Cognition via Information Search in Game Experiments. In *Perspectives on the Future of Economics: Positive and Normative Foundations*, A. Caplin and A. Schotter, Eds. Oxford University Press, in press.

[29] DAL BÓ, P. Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games. *American Economic Review 95*, 5 (2005), 1591–1604.

[30] DAL BÓ, P., AND FRÉCHETTE, G. The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence, October 2006.

[31] DIXON, H. D. Keeping up with the Joneses: competition and the evolution of collusion. *Journal of Economic Behavior & Organization 43*, 2 (October 2000), 223–238.

[32] ECKEL, C., AND GROSSMAN, P. Men, Women and Risk Aversion: Experimental Evidence. In *Handbook of Experimental Results*, C. Plott and V. Smith, Eds. Elsevier, New York, forthcoming.

[33] ELLISON, G. Bounded Rationality in Industrial OrganizationI. In *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, R. Blundell, W. Newey, , and T. Persson, Eds. Cambridge University Press, forthcoming.

[34] FEHR, E., AND GÄCHTER, S. Altruistic punishment in humans. *Nature 415* (2002), 137–140.

[35] FISCHBACHER, U. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*, 2 (2007), 171–178.

[36] FISHMAN, P., AND POPE, D. G. Punishment Induced Detterence: Evidence from the Video Rental Market, November 2006. `http://opimweb.wharton.upenn.edu/OPIM_Content_Management/OPIM_Files/Research_Files/2nd%20Paper%20-%20Pope.pdf`.

[37] FOLKES, V. The Availability Heuristic and Perceived Risk. *The Journal of Consumer Research 15*, 1 (1988), 13–23.

[38] FREIDLIN, M., AND WENTZELL, A. *Random Perturbations of Dynamical Systems.* Springer Verlag, New York, 1984.

[39] FRIEDMAN, J. W., AND MEZZETTI, C. Bounded rationality, dynamic oligopoly, and conjectural variations. *Journal of Economic Behavior & Organization 49*, 3 (November 2002), 287–306.

[40] GABAIX, X., LAIBSON, D., MOLOCHE, G., AND WEINBERG, S. Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model. *American Economic Review 96*, 4 (Sept. 2006), 1043–1068.

[41] GAROUPA, N. Behavioral Economic Analysis of Crime: A Critical Review. *European Journal of Law and Economics 15*, 1 (2003), 5–15.

[42] GILBOA, I., AND SCHMEIDLER, D. Case-Based Decision Theory. *The Quarterly Journal of Economics 110*, 3 (August 1995), 605–39.

[43] HAMAGUCHI, Y., AND KAWAGOE, T. An Experimental Study of Leniency Programs, 2005.

[44] HARRINGTON, J. J. Modelling the Birth and Death of Cartels with an Application to Evaluating Antitrust Policy, 2007. `http://www.econ.jhu.edu/People/Harrington/CartelBirthDeath-6.07.pdf`.

[45] HARRINGTON, J. J. Optimal Corporate Leniency Programs. *Journal of Industrial Economics* (forthcoming). `http://www.econ.jhu.edu/People/Harrington/amnesty11-05.pdf`.

[46] HARRINGTON, W. Enforcement leverage when penalties are restricted. *Journal of Public Economics 37*, 1 (October 1988), 29–53. `http://www.sciencedirect.com/science/article/B6V76-458WNXV-3H/2/d3c35f75197b19adf3fbcd66920bae23`.

[47] HINLOOPEN, J., AND SOETEVENT, A. Trust and Recidivism; the Partial Success of Corporate Leniency Program in the Laboratory, 2006.

http://www.tinbergen.nl/discussionpapers/06067.pdf.

[48] HO, T., CAMERER, C., AND CHONG, J. Self-tuning experience weighted attraction learning in games. *Journal of Economic Theory 133*, 1 (2007), 177–198.

[49] HO, T. H., WANG, X., AND CAMERER, C. Individual Differences in EWA Learning with Partial Payoff Information, October 2006.

[50] HOLT, C. Industrial Organization: A Survey of Laboratory Research. In *The Handbook of Experimental Economics*, J. Kagel and A. Roth, Eds. Princeton University Press, 1995, pp. 349–443.

[51] HOLT, C., AND LAURY, S. Risk Aversion and Incentive Effects. *The American Economic Review 92*, 5 (2002), 1644–1655.

[52] HUCK, S., NORMANN, H.-T., AND OECHSSLER, J. Learning in Cournot Oligopoly–An Experiment. *Economic Journal 109*, 454 (1999), C80–95.

[53] HUCK, S., NORMANN, H.-T., AND OECHSSLER, J. Trial & Error to Collusion - The Discrete Case, May 2000. Bonn Econ. Discussion Papers.

[54] HUCK, S., NORMANN, H.-T., AND OECHSSLER, J. Through Trial and Error to Collusion. *International Economic Review 45*, 1 (02 2004), 205–224.

[55] IVALDI, M., JULLIEN, B., REY, P., SEABRIGHT, P., AND TIROLE, J. The Economics of Tacit Collusion, 2003. IDEI Working Papers.

[56] JACQUEMET, N., RULLIÈRE, J., AND VIALLE, I. Monitoring optimistic agents. *Journal of Economic Psychology* (2007).

[57] JANSSEN, M. Auctions as coordination devices. *European Economic Review 50*, 3 (2006), 517–532.

[58] JEHIEL, P. Learning to Play Limited Forecast Equilibria. *Games and Economic Behavior 22*, 2 (February 1998), 274–298.

[59] JEHIEL, P. Repeated games and limited forecasting. *European Economic Review 42*, 3-5 (1998), 543–551.

[60] JEHIEL, P. Limited Foresight May Force Cooperation. *Review of Economic Studies 68*, 2 (April 2001), 369–91.

[61] JEHIEL, P. Analogy-based expectation equilibrium. *Journal of Economic Theory 123*, 2 (August 2005), 81–104.

[62] JOHNSON, E. J., CAMERER, C., SEN, S., AND RYMON, T. Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining. *Journal of Economic Theory 104*, 1 (May 2002), 16–47.

[63] JOHNSON, E. J., SCHKADE, D. A., AND BETTMAN., J. R. Monitoring Information Processing and Decision: The Mouselab System, 1988. working paper, Graduate School of Industrial Administration, Carnegie-Mellon University.

[64] JOLLS, C. Behavioral Law and Economics, 2007. NBER Working Paper.

[65] JOLLS, C., SUNSTEIN, C., AND THALER, R. A Behavioral Approach to Law and Economics. *Stanford Law Review 50*, 5 (1998), 1471–1550.

[66] KAHNEMAN, D., SLOVIC, P., AND TVERSKY, A. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press, 1982.

[67] KAHNEMAN, D., AND TVERSKY, A. Prospect Theory: An Analysis of Decision under Risk. *Econometrica 47*, 2 (1979), 263–292.

[68] KARANDIKAR, R., MOOKHERJEE, D., RAY, D., AND VEGA-REDONDO, F. Evolving Aspirations and Cooperation. *Journal of Economic Theory 80*, 2 (June 1998), 292–331.

[69] KELLER, C., SIEGRIST, M., AND GUTSCHER, H. The Role of the Affect and Availability Heuristics in Risk Communication. *Risk Analysis 26*, 3 (2006), 631–639.

[70] KWOKA, J. The Attack on Antitrust Policy and Consumer Welfare: A Response to Crandall and Winston, 2003. http://www.economics.neu.edu/papers/documents/03-008.pdf.

[71] LESLIE, C. Antitrust Amnesty, Game Theory, and Cartel Stability. *Journal of Corporation Law 31*, 2 (2006), 453.

[72] LEVITT, S. Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error? *Economic Inquiry 36*, 3 (1998), 353–72.

[73] MILGROM, P., AND ROBERTS, J. Adaptive and sophisticated learning in normal form games. *Games and Economic Behavior 3*, 1 (1991), 82–100.

[74] MONDERER, D., AND SHAPLEY, L. Potential Games. *Games and Economic Behavior 14* (May 1996), 124–143.

[75] MOTTA, M. *Competition Policy: theory and practice*. Cambridge University Press, Cambridge,UK, 2004.

[76] MOTTA, M., AND POLO, M. Leniency Programs and Cartel Prosecution. *International Journal of Industrial Organization 21*, 3 (2003), 347–379.

[77] NEYMAN, A. Bounded Complexity Justifies Cooperation In The Finitely Repeated Prisoners' Dilemma. *Economics Letters 19*, 3 (1985), 227–229.

[78] NORMANN, H.-T., AND WALLACE, B. The Impact of the Termination Rule on Cooperation in a Prisoner's Dilemma Experiment, may 2005. http://personal.rhul.ac.uk/ulte/003/termination.pdf.

[79] OECHSSLER, J. Cooperation as a result of learning with aspiration levels. *Journal of Economic Behavior & Organization 49*, 3 (November 2002), 405–409.

[80] OFFERMAN, T., AND POTTERS, J. Does Auctioning of Entry Licenses Induce Collusion? An Experimental Study. *Review of Economic Studies 73*, 3 (2006), 769–791.

[81] OFFERMAN, T., POTTERS, J., AND SONNEMANS, J. Imitation and Belief Learning in an Oligopoly Experiment. *Review of Economic Studies 69*, 4 (October 2002), 973–97.

[82] PEYTON YOUNG, H. The Evolution of Conventions. *Econometrica 61*, 1 (January 1993), 57–84.

[83] POLINSKY, A., AND SHAVELL, S. The Economic Theory of Public Enforcement of Law. *Journal of Economic Literature 38*, 1 (2000), 45–76.

[84] RASSENTI, S., REYNOLDS, S. S., SMITH, V. L., AND SZIDAROVSZKY, F. Adaptation and convergence of behavior in repeated experimental Cournot games. *Journal of Economic Behavior & Organization 41*, 2 (February 2000), 117–146.

[85] REY, P. Toward a Theory of Competition Policy. In *Advances in Economics and Econometrics: Theory and Applications: Eighth World Congress*, M. Dewatripont, L. Hansen, and S. Turnovsky, Eds. Cambridge University Press, 2003, ch. 3.

[86] ROTH, A., AND EREV, I. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior 8*, 1 (1995), 164–212.

[87] RUBINSTEIN, A. *Modelling Bounded Rationality*. MIT press, Cambridge, USA, 1998.

[88] RUBINSTEIN, A. Instinctive and Cognitive Reasoning : A Study of Response Times, 2006. http://arielrubinstein.tau.ac.il/papers/Response.pdf.

[89] SAPIENZA, P., TOLDRA, A., AND ZINGALES, L. Understanding Trust, 2007.

[90] SELTEN, R., MITZKEWITZ, M., AND UHLICH, G. R. Duopoly Strategies Programmed by Experienced Players. *Econometrica 65*, 3 (1997), 517–556.

[91] SHAPLEY, L. S. Some topics in two-person games. In *Advances in Game Theory*, M. Dresher, L. S. Shapley, and A. W. Tucker, Eds. Princeton, NJ: Princeton Univ. Press., 1964, pp. 1–28.

[92] SIMON, H. A Behavioral Model of Rational Choice. *Quarterly Journal of Economics 69* (1955), 99–118.

[93] SIMON, H. Rational Choice and the Structure of the Environment. *Psychological Review 63* (1956), 129–138.

[94] SKRONDAL, A., AND RABE-HESKETH, S. *Generalized Latent Variable Modeling: multilevel, longitudinal, and structural equation models*. CRC Press, 2004.

[95] SLOVIC, P., FINUCANE, M., PETERS, E., AND MACGREGOR, D. Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason, Risk, and Rationality. *Risk Analysis 24*, 2 (2004), 311–322.

[96] SPAGNOLO, G. Optimal Leniency Programs, 2000. *Fondazione Eni Enrico Mattei* `http://www.feem.it/NR/rdonlyres/F1EEE777-C479-4827-A600-8AA93E9AAE47/868/4200.pdf`.

[97] SPAGNOLO, G. Self-defeating Antitrust Laws:: how Leniency Programs Solve Bertrand's Paradox and Enforce Collusion in Auctions, 2000. *Fondazione Eni Enrico Mattei*.

[98] SPAGNOLO, G. Divide et Impera: Optimal Leniency Programs. *SSRN eLibrary* (2004).

[99] SPAGNOLO, G. Leniency and Whistleblowers in Antitrust. In *Handbook of Antitrust Economics*, P. Buccirossi, Ed. MIT Press, Cambridge, MA., 2008.

[100] SPROUL, M. Antitrust and Prices. *The Journal of Political Economy 101*, 4 (1993), 741–754.

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

[102] VEGA-REDONDO, F. The Evolution of Walrasian Behavior. *Econometrica 65*, 2 (1997), 375–384.

[103] WALTMAN, L., AND KAYMAK, U. Q-learning agents in a Cournot oligopoly model. *Journal of Economic Dynamics and Control* (forthcoming).

[104] WERDEN, G. The Effect of Antitrust Policy on Consumer Welfare: What Crandall and Winston Overlook. *Antitrust Division Discussion Paper No. EAG 03-2* (2003).
`http://ssrn.com/abstract=384100`.

[105] WERDEN, G., AND SIMON, M. Why Price Fixers Should Go to Prison. *Antitrust Bulletin 32* (1987).

[106] WHINSTON, M. *Lectures on antitrust economics*. MIT Press, 2006.