



Submission Number: ASSET2009-09-00011

Convergence in the finite Cournot oligopoly with social and individual learning

Murat Yildizoglu
GREQAM, Aix-Marseille University

Thomas Vallée
LEMNA-Université de Nantes

Abstract

Convergence to the Nash equilibrium in a Cournot oligopoly is a question that recurrently arises as a subject of controversy in economics. The development of evolutionary game theory has provided an equilibrium concept more directly connected with adjustment dynamics, and the evolutionary stability of the equilibria of the Cournot game has been extensively studied in the literature. Several articles show that the Walrasian equilibrium is the stable ESS of the Cournot game. But no general result has been established for the difficult case of simultaneous heterogeneous mutations. Authors propose specific selection dynamics to analyze this case. Vriend (2000) proposes using a genetic algorithm for studying learning dynamics in this game and obtains convergence to Cournot equilibrium with individual learning. The resulting convergence has been questioned by Arifovic and Maschek (2006). The aim of this article is to clarify this controversy: it analyzes the mechanisms that are behind these contradictory results and underlines the specific role of the spite effect. We show why social learning gives rise to the Walrasian equilibrium and why, in a general setup, individual learning can effectively yield convergence to the Cournot equilibrium. We also illustrate these general results by systematic computational experiments.

Convergence in the Finite Cournot Oligopoly with Social and Individual Learning

Thomas Vallée^λ and Murat Yıldızođlu^θ

^λLEMNA, IEMN - IAE
Université de Nantes
Chemin de la Censive du Tertre.
BP 52231. 44322 Nantes Cedex 3.
Thomas.Vallee@univ-nantes.fr

^θCorresponding author
Université Paul Cézanne (Aix-Marseille 3)
GREQAM (UMR CNRS 6579)
Centre de la Vieille Charité
2, rue de la Charité
13236 Marseille cedex 02
murat.yildizoglu@univ-cezanne.fr

January 29, 2009

Abstract

Convergence to the Nash equilibrium in a Cournot oligopoly is a question that recurrently arises as a subject of controversy in economics. The development of evolutionary game theory has provided an equilibrium concept more directly connected with adjustment dynamics, and the evolutionary stability of the equilibria of the Cournot game has been extensively studied in the literature. Several articles show that the Walrasian equilibrium is the stable ESS of the Cournot game. But no general result has been established for the difficult case of simultaneous heterogeneous mutations. Authors propose specific selection dynamics to analyze this case. Vriend (2000) proposes using a genetic algorithm for studying learning dynamics in this game and obtains convergence to Cournot equilibrium with individual learning. The resulting convergence has been questioned by Arifovic and Maschek (2006). The aim of this article is to clarify this controversy: it analyzes the mechanisms that are behind these contradictory results and underlines the specific role of the spite effect. We show why social learning gives rise to the Walrasian equilibrium and why, in a general setup, individual learning can effectively yield convergence to the Cournot equilibrium. We also illustrate these general results by systematic computational experiments.

JEL codes: L130; L200; D430; C630; C730.

Keywords: Cournot oligopoly; Learning; Evolution; Selection; Evolutionary stability; Nash equilibrium; Genetic algorithms.

1 Introduction

The starting point of this article is the recent contradiction that has arisen in the literature about the convergence to equilibrium in Cournot oligopoly games where boundedly rational agents' learning is represented using Genetic algorithms (**GA**). The interesting and contrasted results of Vriend (2000) have been questioned by more recent studies, in particular, by Arifovic and Maschek (2006). The main result of Vriend (2000) is convergence to Cournot equilibrium (**CE**) only when agents' learning is individual, while social learning yields convergence to the Walrasian equilibrium (**WE**). Arifovic and Maschek (2006) confirm the result for social learning, but question the convergence results under individual learning. The aim of this article is to clarify both the initial contrasted results, and the contradiction between these two computational articles. This contradiction adds to a rather old debate in economics.

Convergence to the Nash equilibrium in a Cournot oligopoly is indeed a problem that recurrently arises as a subject of controversy in economics. Already in the 60's, several articles in the Review of Economic Studies discussed the convergence under best reply dynamics (*see* for example Theocharis (1960)). This debate has led to the conclusion that best reply dynamics become unstable under standard assumptions about demand and cost conditions as soon as there are three or more firms in the oligopoly.

The development of evolutionary game theory (Maynard-Smith, 1982) has indirectly contributed to this debate by providing an equilibrium concept more directly connected with other types of adjustment dynamics (than best reply) which can arise when the learning of firms is taken into account. Several articles have studied the evolutionary stability of equilibria in the Cournot game. In a path-breaking article, Vega-Redondo (1997) shows that the Walrasian equilibrium (**WE**) is the unique evolutionary stable outcome in a quantity competition game with a homogenous product under certain specific assumptions about the interaction structure and strategy sets of the agents. The results obtained in the articles that followed Vega-Redondo (1997) cast more doubt on the potential convergence of adjustment dynamics to the Cournot equilibrium (**CE**). The nature of mutations and the selection dynamics are at the core of these results. These dimensions directly result from assumptions about the learning processes of boundedly rational firms with incomplete and imperfect information.

The evolutionary stability concept is very useful for analyzing the convergence to equilibrium with dynamics more naturally connected to social learning of firms through imitation of strategies and experimentation (random mutations). But this stability concept does not exclude other learning schemes, as long as they can be formulated as selection mechanisms operating at the level of firm population. Unfortunately, richer learning schemes generally imply more complex dynamics, and oligopoly is more naturally formulated as a *playing the field* game (each player playing against all the others in each period), than a pairwise matching process. The analysis of these dynamics is very difficult under general conditions and the existing literature proposes only partial results obtained under specific assumptions, as in Vega-Redondo (1997). Stegeman and Rhode (2004) or Bergin and Bernhardt (2004) consider relatively richer frameworks but they again obtain partial results that cannot be used to draw general conclusions about the consequences of different representations of firms' learning. This difficulty has motivated other recent studies to resort to agent based computational models to explore the question of convergence (*see* for example Kirman (1995) for a simulation-based analysis of three such specific learning rules).

Arifovic (1994) and Vriend (2000) develop models that adopt such an explorative strategy. They use a computational representation of learning that is quite commonly retained in the literature: genetic algorithms (*see* Dawid (1999), Vallée and Yıldızoğlu (2004)). GA represent the learning of agents as an exploratory combination process based on random experiments and combination of already discovered strategies. Arifovic (1994) initially obtained a convergence to the WE with this type of learning mechanisms. Vriend (2000) advances that genetic algorithms (**GA**) can be used to demonstrate the "essential difference between individual and social learning" of firms in a Cournot

oligopoly: a GA operating at the firm population level yields convergence, under social learning, to the WE, while a set of GAs each adjusting the strategies of an individual firm imply, under individual learning, convergence to the CE. These results apparently contrast with Arifovic (1994). They have also recently been questioned by Arifovic and Maschek (2006), who draw attention to the specific implementation adopted by Vriend in terms of the mechanisms of GA-based learning and the cost structure of firms. More precisely, Arifovic and Maschek confirm convergence to the WE under social learning, but they cast some doubt on the possibility of convergence to the CE under individual learning. Since all the results of these three articles are computational, it is difficult to understand why the convergence results under social learning are so robust and why the results for the individual learning are so controversial. It is important to verify if the difference between these two types of learning is really *essential*. If this is the case, then we must try to understand the fundamental mechanisms around which the debate about individual learning evolves.

The article is organized in two parts. The first part exposes the main results concerning imitation-based social learning. These results will be established by studying the evolutionary stability of equilibria in Cournot games (section 2), and by focusing on the main mechanism behind these results (section 3 on spite effect). These analytical convergence results will cover all general cases but one: the case with simultaneous and heterogenous mutations from equilibrium. Since this is the most interesting case for the Cournot game, it will be necessary to follow the analysis by combining partial analytical results with computational experiments. The second part (section 4) will consequently analyze the convergence conditions when simultaneous heterogenous mutations are allowed. In order to clarify the contradiction cited above, the analysis will be restricted to the special selection mechanism used by genetic algorithm based learning. When this is possible, we will rely on the general results of the first part in this analysis (in particular for the social learning case). Our results should also interest other domains where GA is used to represent individual or social learning (*see* for example Vallée and Basar (1999) or Yildizoglu (2002)). We will show why social learning cannot yield convergence to CE, and why individual learning can converge to CE if the interactions of the firms allow them to discover the decreasing relationship between the market price and their quantities. The last section concludes the article.

2 Evolutionary stability and Nash equilibrium in oligopoly

In this first section, we present the Cournot oligopoly game and its equilibria. We also define evolutionary stability in this game. We also study the evolutionary stability of these equilibria, before analyzing, in the section 3, the role of the spite effect.

2.1 A simple oligopoly model

We consider a standard symmetrical n -firms oligopoly model of quantity competition where all firms produce a homogenous product. The inverse demand function for this good is given by $p = p(Q)$, where $Q = \sum_{i=1}^n q_i$ and $dp/dQ < 0$. The common cost function of the firms is $C(q_i)$, with $C' > 0$ and $C'' > 0$. The profit function of a firm is: $\pi_i(q_i, \dots, q_n) = p(Q)q_i - C(q_i)$. Since the interaction between the strategies (quantities) of the firms only takes place through the common inverse demand function (and hence, through the sum of these quantities), a quantity profile (q_1, \dots, q_n) can be represented, from the point of view of a firm i , as (q_i, Q_{-i}) , where $Q_{-i} = \sum_{j \neq i} q_j$.

In this oligopoly, two different kinds of equilibrium can be defined: the Cournot-Nash equilibrium (**CE**) and the Walrasian equilibrium (**WE**).

Definition 1. A Cournot-Nash equilibrium (**CE**) is given by a quantity profile q^C and a market price p^C such as

1. Each firm maximizes its profit at this equilibrium

$$q_i^C = \operatorname{argmax}_{q_i} \pi_i(q_i, Q_{-i}^C) \Rightarrow p(Q^C) + q_i^C p'(Q^C) = C'(q_i^C), \forall i = 1, \dots, n \quad (1)$$

2. The market clears: $p^C = p(Q^C)$.

Definition 2. A Walrasian equilibrium (**WE**) is given by a quantity profile q^W and a market price p^W such as

1. Each firm uses marginal cost pricing (it is a price-taker)

$$q_i^W = \operatorname{argmax}_{q_i} \pi_i(q_i; p^W) \Rightarrow C'(q_i^W) = p^W, \forall i = 1, \dots, n \quad (2)$$

2. The market clears : $p^W = p(Q^W)$.

Since $dp/dQ < 0$ and $C'' > 0$, the conditions (1) and (2) imply the standard results on the comparison of this equilibrium:

$$q_i^C \leq q_i^W, \forall i \quad (3)$$

$$\Rightarrow Q^C \leq Q^W \text{ and } p^C \geq p^W \quad (4)$$

$$\pi_i^C = \pi_i(q_i^C, Q_{-i}^C) \geq \pi_i(q_i^W, Q_{-i}^W) = \pi_i^W, \forall i \quad (5)$$

Consequently, the Cournot equilibrium is preferred by the firms, while the consumer's surplus is maximal in the Walrasian equilibrium. Hence, agents are not indifferent to the equilibrium to which the market can converge. The analysis of the evolutionary stability of these equilibria sheds light on this convergence, under the assumption that successful strategies will have a tendency to diffuse through the population of firms.

2.2 Evolutionary stability of equilibrium

In standard evolutionary game theory, the selection dynamics are studied through symmetric pairwise interactions within a large population of players. A strategy is an evolutionary stable strategy (**ESS**) if, once adopted by the whole population, it cannot be invaded by a small *mass* of mutants (Maynard-Smith (1982), Weibull (1995)). Or, following the traditional definition (Weibull 2006):

Definition 3. A strategy profile q^* is said to be evolutionary stable (ESS) if, and only if, it meets the following conditions:

$$\pi_i(q_i, Q_{-i}^*) \leq \pi_i(q_i^*, Q_{-i}^*), \forall q_i, \forall i \quad (6)$$

$$\pi_i(q_i, Q_{-i}^*) = \pi_i(q_i^*, Q_{-i}^*) \Rightarrow \pi_i(q_i, Q_{-i}) < \pi_i(q_i^*, Q_{-i}) \quad (7)$$

It is clear that condition (6) implies that all ESS are Nash equilibrium (**NE**) strategies, but not all NE strategies are necessarily ESS (because of the condition (7)).

When we consider random, pairwise contests between individuals drawn from a large population, we arrive at this conclusion by assuming that both the predominant strategy and the mutant confront the same population profile. As already noticed by Riley (1979), a strategy which satisfies conditions (6 – 7), may, nevertheless, not be protected against invasion by a mutant strategy if we have a finite population game such as the Cournot game. In order to analyze an oligopoly situation we need, as shown by Alós-Ferrer and Ania (2005), a definition of an ESS adapted to a finite population of players that “play the field”, that is, all compete with each other simultaneously. Schaffer (1988) arrives at the same conclusion and proposes a concept of a *finite population ESS* defined as follows.

Definition 4. In a finite population game, a strategy profile q^* is said to be evolutionary stable if in the presence of a mutant firm j playing $q_j^m \neq q_j^*$

$$\pi_i(q_i^*, q_j^m, Q_{-i-j}^*) \geq \pi_j(q_j^m, Q_{-j}^*), \forall i \neq j, \forall q_j^m \neq q_j^* \quad (8)$$

where $Q_{-i-j}^* = \{q_1^*, \dots, q_{i-1}^*, q_{i+1}^*, \dots, q_{j-1}^*, q_{j+1}^*, \dots, q_n^*\}$.

In our oligopoly game, this definition means that the profit of firm i must be higher if it faces a mutant that plays q^m than if it is that mutant.

2.3 ESS and Nash equilibrium

As shown by Schaffer (1988), a finite population ESS is generally not a Nash equilibrium strategy, since condition (8) does not exclude the possibility that

$$\exists j, q_j^m, \pi_j(q_j^m, Q_{-j}^*) \geq \pi_j(q_j^*, Q_{-j}^*), \quad (9)$$

Indeed, the consequence of condition (8) together with condition (9) is as follows: even if an individual firm can obtain a net benefit by moving from the current equilibrium, the other firms will not imitate it and the population will not be invaded by this mutant strategy. Consequently, the current equilibrium strategy is ESS.

Proposition 1. (Schaffer (1989), Vega-Redondo (1997)) *In a finite population oligopoly game, the only ESS is the Walrasian equilibrium strategy.*

Consequently, it is quite paradoxical to have an oligopoly where, when each individual firm tries to adopt strategies with higher payoffs, the industry will, nevertheless, converge to WE where the profit of the firms is lower than in the CE.

The next section will analyze the mechanisms behind this result. We will pay special attention to the role of the spite effect in convergence to equilibrium.

3 Evolutionary stability of equilibrium and the spite effect

The mechanism that pushes the population towards the Walrasian equilibrium is well documented in evolutionary literature, and its relation with the spite effect has been established by Schaffer (1989). Hamilton (1970,1971) defines spiteful behavior as the capacity of an animal to harm both itself and another, and gain an evolutionary advantage if the harm on the other animal is greater than the self-harm. Schaffer establishes that the behavior of a firm that chooses to decrease its profit can diffuse in the population if it can decrease the profits of its competitors by a greater extent, and hence gain a comparative advantage.

Strictly speaking, the *spite effect* is the opposite of altruism. We should use the term "spite effect" only when the mutant firm j suffers from deviation ($\Delta\pi_j < 0$), but suffers less than the non-mutant firms ($\Delta\pi_i < \Delta\pi_j < 0, \forall i \neq j$). By *analogy*, in the positive direction, we may reinterpret this effect as the fact that the net gain of the mutant ($\Delta\pi_j > 0$) is greater than the net gain of the non-mutant firms ($\Delta\pi_j > \Delta\pi_i$) which is always true of course if the non-mutants suffer from the deviations ($\Delta\pi_i < 0$). Thus we have the following definition.

Definition 5. *In a finite population oligopoly game, given that a firm j will deviate from an equilibrium strategy, a spite effect exists when $\Delta\pi_i - \Delta\pi_j < 0, i \neq j$.*

The relationship between Nash equilibrium (**NE**), ESS and spite effect is straightforward if we consider strict Nash equilibrium. We can study the evolutionary stability of the NE under increasingly complex assumptions. We first allow one single strategy to mutate. Then we analyze evolutionary stability in the face of multiple identical mutations. Finally, we observe that obtaining a general result with multiple heterogenous mutations is very difficult. This is why this configuration is studied in the literature under more specific selection mechanisms, the second part of the article will consider such a mechanism. The last part of section 3 will also establish the relationship between ESS, spite effect, and maximization of relative payoffs, following Schaffer (1989), and Bergin & Bernhardt (2004).

3.1 Nash equilibrium and ESS in a finite population oligopoly game

We have already observed that any NE is not necessarily ESS. However we can show that in the absence of the spite effect, a strict Nash equilibrium is ESS. Consequently, the role of the spite effect is important.

Definition 6. A strategy $Q^N = (q_1^N, q_2^N, \dots, q_n^N)$ is a strict Nash equilibrium if

$$\pi_i(q_i^N, Q_{-i}^N) > \pi_i(q_i, Q_{-i}^N), \forall q_i \neq q_i^N \quad (10)$$

Proposition 2. In a symmetric finite population oligopoly game, a strict Nash equilibrium strategy profile is an ESS if a spite effect does not exist.

Proof.

Assume that we have a strict Nash equilibrium, q^N :

$$\pi_i(q_i^N, Q_{-i}^N) > \pi_i(q_i, Q_{-i}^N), \forall q_i \neq q_i^N \quad (11)$$

Recall that the spite effect does not hold if $\Delta\pi_i - \Delta\pi_j > 0$, that is if:

$$\left[\pi_i(q_j, Q_{-j}^N) - \pi_i(q_j^N, Q_{-j}^N) \right] - \underbrace{\left[\pi_j(q_j, Q_{-j}^N) - \pi_j(q_j^N, Q_{-j}^N) \right]}_{<0 \text{ by inequality (11)}} > 0$$

which means that a sufficient condition is:

$$\Delta\pi_i \equiv \pi_i(q_j, Q_{-j}^N) - \pi_i(q_j^N, Q_{-j}^N) > 0, \forall q_j \neq q_j^N$$

It is easy to check that, since by symmetry $\pi_i(q_j^N, Q_{-j}^N) = \pi_j(q_j^N, Q_{-j}^N)$, the last condition implies that $\pi_i(q_j, Q_{-j}^N) > \pi_j(q_j^N, Q_{-j}^N)$. And, by definition of the strict Nash equilibrium, we get the second part of the following condition

$$\pi_i(q_j, Q_{-j}^N) > \pi_j(q_j^N, Q_{-j}^N) > \pi_j(q_j, Q_{-j}^N)$$

which thus establishes the ESS condition (8). So, if the spite effect does not occur, the symmetric strict Nash equilibrium strategy is an ESS. \square

This proposition establishes the role of the spite effect in the evolutionary stability of a strict NE. We now establish more detailed results under increasingly complex assumptions about the allowed mutations: single mutation, multiple identical mutations, and multiple heterogenous mutations.

3.1.1 One single mutation case

Proposition 3. In a finite population symmetric oligopoly game, with one possible single mutation, a strategy profile $\hat{q} = (\hat{q}_1, \dots, \hat{q}_n)$ is an ESS, and no spite effect occurs, if

$$dq_i (p' dq_i + p - C') < 0, \forall i \quad (12)$$

Proof.

With \hat{q} the profits of the firms are

$$\pi_j(\hat{q}) = p(\hat{q})\hat{q}_j - C(\hat{q}_j) \quad (13)$$

$$\pi_i(\hat{q}) = p(\hat{q})\hat{q}_i - C(\hat{q}_i) \quad (14)$$

A deviation of the firm j from the equilibrium, $q_j \neq \hat{q}_j$, yields a new strategy profile $q^m = (\hat{q}_1, \hat{q}_2, \dots, \hat{q}_{j-1}, q_j, \hat{q}_{j+1}, \dots, \hat{q}_n)$ and the following profits:

$$\pi_j(q^m) = p(q^m)q_j - C(q_j) \quad (15)$$

$$\pi_i(q^m) = p(q^m)\hat{q}_i - C(\hat{q}_i), \forall i \quad (16)$$

The deviation will not be imitated if, with this new strategy profile q^m , the mutant firm j earns a lower profit than the non-mutant firms, that is if

$$\begin{aligned} & \pi_j(q^m) - \pi_i(q^m) < 0 \\ \Leftrightarrow & \pi_j(q^m) - \pi_j(\hat{q}) - \pi_i(q^m) + \pi_i(\hat{q}) < 0 \\ \Leftrightarrow & \Delta\pi_i - \Delta\pi_{-i} < 0 \end{aligned}$$

since, by the symmetry of the initial equilibrium, we have $\pi_i(\hat{q}) = \pi_j(\hat{q})$. The net gain of firm j , $\Delta\pi_j$ corresponds to (15) – (13):

$$\begin{aligned}\Delta\pi_j &= \pi_j(q^m) - \pi_j(\hat{q}) \\ &= [p(q^m)q_j - (p(\hat{q})\hat{q}_j)] - [C(q_j) - C(\hat{q}_j)]\end{aligned}$$

For a local deviation, $q_j = \hat{q}_j + dq_j$ we can compute the marginal net gain:

$$\Delta\pi_j = \left[p(\hat{q}_j + dq_j, \hat{Q}_{-j})(\hat{q}_j + dq_j) - p(\hat{q})\hat{q}_j \right] - [C_j(\hat{q}_j + dq_j) - C_j(\hat{q}_j)] \quad (17)$$

$$= \underbrace{\left[p(\hat{q}_j + dq_j, \hat{Q}_{-j}) - p(\hat{q}) \right] \hat{q}_j}_{\hat{q}_j dp} + \underbrace{\left[p(\hat{q}_j + dq_j, \hat{Q}_{-j}) - p(\hat{q}, \hat{Q}_{-j}) + p(\hat{q}) \right] dq_j}_{dp dq_j + \hat{p} dq_j} \quad (18)$$

$$- \underbrace{[C_j(\hat{q}_j + dq_j) - C_j(\hat{q}_j)]}_{dC_j} \quad (19)$$

with $\hat{p} = p(\hat{q})$. For any other firm $i \neq j$ we have:

$$\begin{aligned}\Delta\pi_i &= \pi_i(\hat{q}_j + dq_j, \hat{Q}_{-j}) - \pi_i(\hat{q}) \\ &= \underbrace{\left(p(\hat{q}_j + dq_j, \hat{Q}_{-j}) - p(\hat{q}) \right) \hat{q}_i}_{\hat{q}_i dp}\end{aligned} \quad (20)$$

And consequently:

$$\Delta\pi_j - \Delta\pi_i < 0 \quad (21)$$

$$\Rightarrow \hat{q}_j dp + dp dq_j + \hat{p} dq_j - dC_j - \hat{q}_i dp < 0$$

$$\Rightarrow dp(\hat{q}_j - \hat{q}_i) + dp dq_j + \hat{p} dq_j - dC_j < 0 \quad (22)$$

$$\Rightarrow dp dq_j + \hat{p} dq_j - dC_j < 0 \text{ since } \hat{q}_j = \hat{q}_i \quad (23)$$

Since $dp = p' dq_j$ and $dC_j = C' dq_j$,

$$p' dq_j dq_j + \hat{p} dq_j - C' dq_j < 0 \quad (24)$$

$$dq_j (p' dq_j + \hat{p} - C') < 0 \quad (25)$$

□

Before discussing the implications of social learning with genetic algorithms in more detail in the next section, we can already establish a first general result concerning social learning in the finite oligopoly game.

Proposition 4. *With imitation-based social learning, the strategies that imply a deviation from the WE will be eliminated while the strategies that imply a deviation from the CE in the direction of the WE will diffuse in the population.*

Proof.

Local (infinitesimal) deviation from the WE. At the WE, we have $\hat{p} = p^w = C'$ and the condition for stability (25) becomes $p'(dq_j)^2 < 0$. □

Proof. • $dq_j < 0$: Since $p' < 0$, the condition (25) holds. A single firm that deviates by decreasing its quantity will benefit, via a price increase, from a revenue increase lower than that offered to the firms who stay committed. So, the mutant strategy will not be replicated by imitation and the Walrasian strategy will persist. □

Proof.

- $dq_j > 0$: The firm that increases its quantity will decrease the market price to a level such that the profit of every firm will decrease. However the loss of the mutant firm will be higher than the loss of the firms continuing to play the Walrasian strategy. Again, this mutant strategy will not be replicated in the population.

Local deviation from CE. Remark that at the CE we have, by definition, an equality between marginal revenue and marginal cost. This means that $p'_j q_j^{CE} + p^{CE} - C' = 0$. If the deviation is local, we must have $|dq_j| < q_j^{CE}$, then $|p' dq_j| < |p' q_j|$ with $p' < 0$. Therefore the following results hold:

- $dq_j > 0$: Since $p^{CE} - C' > 0$, $p'_j q_j^{CE} + p^{CE} - C' = 0$ and $p' dq_j < p' q_j^{CE}$ (since $|dq_j| < q_j^{CE}$) we necessarily have $p' dq_j + p^{CE} - C' > 0$. As a consequence, condition (12) does not hold, and there is no spite effect. If one firm leaves the CE by increasing its quantity, it will decrease its profits, but this decrease will be less than the reduction of the other firms' profits. This strategy will be replicated and the deviation from CE will diffuse in the population.
- $dq_j < 0$: In this case, we have $p' dq_j > 0 > p' q_j^{CE}$ and $dq_j(p' dq_j + p^{CE} - C') < 0$. Thus, no one should imitate a strategy that moves away from the CE by decreasing the quantity again. That is, if one firm decreases its quantity from CE, it will earn less supplementary profits than the other firms. As a consequence, this strategy will not be replicated in the population.

□

This last proposition explains why imitation-based social learning drives the industry towards WE. This result can be generalized to the case of multiple mutants, when the mutations are identical.

3.1.2 Generalization to m identical mutations

Assume that m firms play a strategy $q_j = q$ while the other $n - m$ firms stay committed to the initial equilibrium strategy q^* .

Definition 7. In a finite population game with m mutations, an equilibrium strategy $(q_1^*, q_2^*, \dots, q_n^*)$, with $q_i^* = q^*, \forall i$, is a strong ESS if, $\forall q_j \neq q^*$,

$$\pi_{i_m}(q_j, Q_{-j_m}) > \pi_{j_m}(q_j, Q_{j_m})$$

with $Q_{j_m} \equiv (m - 1)q_j + (n - m)q^*$, j_m any mutant firm and i_m any non-mutant firm.

This means that if m firms play $q_j \neq q^*$, the profits of these m firms should be lower than the profits of the $n - m$ firms that stay committed to q^* . From this remark, Schaffer (1988) defines the concept of m -stable (finite population) ESS as follows.

Definition 8. An equilibrium strategy $(q_1^*, q_2^*, \dots, q_n^*)$ is called an m -stable (finite population) ESS if

$$\pi_i(Q_{\hat{m}}) > \pi_j(Q_{\hat{m}}) \text{ and } \pi_i(Q_{m+1}) < \pi_j(Q_{m+1})$$

with $1 \leq \hat{m} \leq m$, and with $Q_{\hat{m}} \equiv \hat{m}q_j + (n - \hat{m})q^*$ and $Q_{m+1} \equiv (m + 1)q_i + (n - (m + 1))q^*$.

The m -stable ESS means that as long as the number of the mutants is less than or equal to m the initial equilibrium strategy is an ESS one.

Proposition 5. In a finite population oligopoly game, a strategy profile $q^* = (q_1^*, q_2^*, \dots, q_n^*)$ is an ESS equilibrium whatever the number of mutants, if at least one non-mutant firm does not suffer from the spite effect implied by deviations from q^* .

Proof.

Note by I_m the set of mutant firms. With $q_j = \hat{q}, \forall j \in I_m$, $q_i = q_i^* = q^*, \forall i \notin I_m$ and $Q_m = m\hat{q} + (n - m)q^*$, the profits of the firms become

$$\text{For any mutant firm: } \pi_j(Q_m) = p(Q_m)\hat{q} - C_i(\hat{q}) \quad (26)$$

$$\text{For any non-mutant firm: } \pi_i(Q_m) = p(Q_m)q^* - C(q^*) \quad (27)$$

It is clear that the net gains $\Delta\pi_j$ are equivalent between all mutant firms and, $\Delta\pi_i$, between all non-mutant firms. If the spite effect does not work on at least one non-mutant firm, then it does not work on any non-mutant firm and these firms do not have any incentive to imitate the mutant strategy. As a consequence, the initial strategy profile is an ESS. \square

Corollary 1. *The Walrasian equilibrium strategy is an m -stable ESS, $\forall m < n - 1$.*

Proof.

Since the firms are identical, the oligopoly game with n firms and with m mutants is equivalent to a 2-player game with one single mutant. Thus, if the Walrasian equilibrium is an ESS with one single mutant for any finite population size n , it is an m stable ESS with m identical mutants. \square

3.1.3 The case of m heterogenous mutations

Let us assume that there exists m heterogenous mutants. In our oligopoly game, this means that some firms will increase their quantities while others will decrease their's. In such a framework, the firms that do not move will have an evolutionary stable strategy if they can at least get profits as high as the best of the mutants. In that case, the following definition holds.

Definition 9. *An equilibrium strategy $(q_1^*, q_2^*, \dots, q_n^*)$ is an ESS with m heterogenous mutants composing the set I_m under the following condition: if, for $\forall q_j \neq q^*, j \in I_m$, $\forall q_i = q^*, i \notin I_m$, and with \hat{q} , a strategy profile composed of m mutant strategies q_j and $n - m$ equilibrium strategies q^* , we have*

$$\pi_i(\hat{q}) > \max_{j \in I_m} \pi_j(\hat{q}), \forall i.$$

This definition means that whatever quantities the mutants choose, and whatever the new total quantity, the non-mutant firms must obtain a higher profit than any mutant firm. Otherwise, at least one mutation will be optimal to imitate, and it will diffuse.

As observed by Schaffer (1989), a strategy profile may be m -stable with homogenous mutants (e.g. mutations in the same direction) but unstable with heterogenous mutations.

It is very difficult to obtain general results with multiple heterogenous mutations since we can have mutations in opposite directions with an ambiguous and variable aggregate effect. Vega-Redondo (1997) and Bergin & Bernhardt (2004) limit this problem by adopting an approach that restricts the quantities to a discrete grid in order to structure possible deviations. Both obtain convergence to WE under imitation-based learning. In Bergin and Bernhardt (2004), convergence to CE only happens under specific cases of individual learning with memory. Unfortunately, given the specific constraints of these articles, they cannot provide a definitive conclusion to the controversy studied by our article, even if their results can guide us in our explorations (*see* the next section).

Following Schaffer (1989) and Bergin & Bernhardt (2004), we can reformulate the relation between spite effect and the ESS on the basis of relative payoffs of firms. This will also allow us to introduce the fitness concept that is central to the evolutionary learning mechanisms.

3.2 ESS and the importance of the relative payoffs

As observed by Schaffer (1989), there is a relationship between the finite population ESS concept and the “beat the average” game. Obviously, a strategy is an ESS if it is not possible to increase the

relative payoff by individual mutation, which means that it is a Nash equilibrium in the “beat the average” game reformulation of the initial oligopoly game.

Let us define the relative payoff function of a firm i by the difference of the firm’s profit and the population’s average profit:

$$\pi_i^R(q_i, Q_{-i}) \equiv \pi_i(q_i, Q_{-i}) - \frac{1}{n} \sum_{j=1}^n \pi_j(q_j, Q_{-j})$$

In an initial symmetrical configuration¹ ($\pi_j = \pi, \forall j$), we must have $\pi_i^R(q_i, Q_{-i}) = \frac{(n-1)}{n} \pi_i(q_i, Q_{-i}) - \bar{\pi}_{-i}$, with $\bar{\pi}_{-i} = \frac{(n-1)}{n} \pi$, with $\pi_j(q_j, Q_{-j}) = \pi \forall j \neq i$.

Proposition 6. *If q^* is not ESS, then there exists at least one mutant strategy \hat{q}_j such that the relative profit of the mutant firm j increases.*

Proof.

We know that q^* is not an (symmetrical) ESS if

$$\pi_j(\hat{q}_j, Q_{-j}^*) - \pi_i(\hat{q}_j, Q_{-j}^*) > 0$$

Since the non-mutant firms are symmetric, we can pose $\pi = \pi_i(\hat{q}_j, Q_{-j}^*), i \neq j$. Then we have from (8)

$$\begin{aligned} \frac{n-1}{n} (\pi_j(\hat{q}_j, Q_{-j}^*) - \pi_i(\hat{q}_j, Q_{-j}^*)) &> 0 \\ \pi_j(\hat{q}_j, Q_{-j}^*) \frac{(n-1)}{n} - \frac{(n-1)}{n} \pi &> 0 \\ \frac{(n-1)}{n} \pi_j(\hat{q}_j, Q_{-j}^*) - \bar{\pi}_{-j} &= \pi_j^R(\hat{q}_j, Q_{-j}^*) > 0 \end{aligned}$$

That is, the relative profit of the mutant is positive. Since we start from an initial symmetrical equilibrium where $\pi_j^R(q_j^*, Q_{-j}^*) = 0$, the relative profit of the mutant firm does indeed increase. \square

Another implication of this proposition can be formulated using the *relative fitness* (f_i) concept that we will use later in this article (see Definition 10, page 11):

$$f_i = \frac{\pi_i(q_i, Q_{-i})}{\sum_{j=1}^n (\pi_j(q_j, Q_{-j}))} = \frac{\pi_i(q_i, Q_{-i})}{\pi_i(q_i, Q_{-i}) + \Pi_{-i}}$$

where $\Pi_{-i} = \sum_{k \neq i} \pi_k(q_k, Q_{-k})$. The fitness measure is used in most biological games (and in nearly all evolutionary simulations) as the base for the selection process in a given population.

Corollary 2. *If q^* is not an ESS, there exists at least one strategy \hat{q}_j such that the fitness of the firm is increased.*

Proof.

First note that $\Pi_{-i} = (n-1)\pi_{-i} = (n-1)\pi$. From the previous proof we know that

$$\begin{aligned} \pi_j(\hat{q}_j, Q_{-j}) \frac{(n-1)}{n} - \frac{(n-1)}{n} \pi &> 0 \\ \pi_j(\hat{q}_j, Q_{-j})(n-1) - (n-1)\pi &> 0 \\ n\pi_j(\hat{q}_j, Q_{-j}) &> \pi_j(\hat{q}_j, Q_{-j}) + \Pi_{-i} \\ \frac{\pi_j(\hat{q}_j, Q_{-j})}{\pi_j(\hat{q}_j, Q_{-j}) + \Pi_{-i}} &> \frac{1}{n} \end{aligned}$$

¹For a non-homogenous initial situation, the proposition can be trivially proved by considering a deviation consisting in the imitation of the best strategy.

Since initially the fitness is

$$f_i = \frac{\pi_i}{\pi_i + \Pi_{-i}} = \frac{\pi}{n\pi} = \frac{1}{n}$$

then the *fitness* of the mutant firm increases. □

The study of general learning process of firms in a Cournot oligopoly implies the consideration of simultaneous heterogenous mutations and suffers from the difficulty we highlighted at the end of the preceding section. As a consequence, the literature has adopted an experimental strategy to tackle this problem: several articles introduce specific learning processes in order analyze, in more detail, the convergence of learning dynamics to equilibria, under heterogenous mutations. However, it appears that different learning processes can yield contrasting convergence results under genetic algorithm-based selection dynamics, as shown by the debate that followed the interesting, contested results obtained by Vriend (2000). These selection dynamics are driven by the relative fitness of firms' different strategies, as in Definition 10. As a consequence, this debate is around a particular case of the general framework studied in the preceding section, as shown in Corollary 2. We will be able to clarify, in the next section, the mechanisms that are behind the convergence results obtained with GA-based learning processes considered by the literature. We will observe that Corollary 2 will directly imply our main result concerning the social learning case. This analysis will also give us the opportunity to introduce individual learning, which should be contrasted with the imitation-based social learning. The consideration of this type of learning will directly be subject to the difficulty outlined in Definition 9. A specific theoretical analysis and computational experiments will help us to obtain partial results in this difficult case.

4 Learning, selection, and convergence to equilibrium

We now consider the possibility for firms to adapt their production levels as a consequence of their learning. Arifovic (1994) has introduced genetic algorithm (**GA**) as a tool for modeling adaptive learning of firms in a Cournot oligopoly. The main result of this article is the convergence to the WE under social and individual learning. More recently, Vriend (2000) questioned these results by showing that convergence to CE can be observed under individual learning. Arifovic and Maschek (2006) has cast some doubt on this result by showing the difficulty of converging to CE even with individual learning, except under the very specific assumptions adopted by Vriend (2000). Based on the general results established in the preceding section, we will analyze some general mechanisms that encompass all these contradictory results, and explain why they are obtained under the corresponding assumptions. After having discussed the selection mechanism implied by the GA, we will first analyze the convergence to equilibrium under social learning. The results of this first step are directly derived from the first part of our article; we just transpose them to the specific selection mechanism corresponding to GA. Then we will study individual learning and show why the expectations-based learning assumed by Arifovic (1994) and by Arifovic and Maschek (2006) cannot converge to CE, and why repetition-based learning of Vriend (2000) can. We will generalize the specific results of these articles under more standard assumptions for the Cournot oligopoly game.

4.1 Modeling learning with genetic algorithms (GA)

A GA is based on mechanisms inspired by biological evolution: selection, crossover, and mutation. The *canonical genetic algorithm* makes a population of chromosomes, $\{A_j\}$, $j = 1..m$ evolve. The size of the population (m) is given and is the source of one of the strengths of GA: implicit parallelism (the exploration of the solution space using several candidates in parallel). The population of chromosomes at step t (a generation) is denoted $P(t) = \{A_j\}_t$ with $\#P(t) = m$, and $\forall t = 1, 2..T$ with T , the given number of generations. Notice that the size of T is the other source of the strengths of the GA. The algorithm (randomly) generates an initial population $P(0)$ of candidate chromosomes that are evaluated at each period using the fitness (value) function. These chromosomes are used for

composing a new population at the next period $P(t+1)$. Each chromosome has a probability of being **selected** that increases with its relative fitness. The members included in the new population are recombined using a **crossover** mechanism. The crossover operation introduces controlled innovations in the population, combining the candidates already selected in order to create new candidates with a potentially better fitness. In addition, the **mutation** operator randomly modifies the candidates and introduces some random experimenting in order to explore more extensively the state space and escape local optima. Typically, the probability of mutation is relatively low in comparison with the probability of crossover because otherwise the disruption introduced by excessive mutations could destroy the hill-climbing capacity of the population. Finally, an **elitism** operator can be used to ensure that the best chromosome of a population will be carried to the next generation.

In the context of oligopoly, the GA is used to represent the learning capacity of the firms: it can make a population of production levels evolve based on increasing fitness, the profits resulting from these quantities. Vriend (2000) contrasts two different implementations of this approach and claims that this comparison can prove an essential difference between social and individual learning. In the case of *social learning*, the population of strategies that evolve through GA contains one strategy per firm and the operations of the GA correspond to imitation between firms (crossover operator) and to random experimenting by some firms (mutation operator). In the case of *individual learning*, the GA operates on the individual strategy population of each firm (we have as many GAs as firms in the economy) and the operations of the GA correspond to recombination of strategies already found (crossover operator) and to random experimenting in the strategy population of the firm (mutation operator).

By confronting these two different learning approaches, Vriend (2000) obtains a convergence to WE under social learning and convergence to CE, under individual learning. Arifovic and Maschek (2006) question these results by stressing the difficulty of converging to CE. Our preceding discussion would indicate that the results of Vriend (2000) are robust, even if they are obtained in this article under very specific demand and cost conditions, and with a very specific application of the GA methodology. We will now explore the general mechanisms behind these different results in the literature. Our analysis will be focused on the main operator in the evolution of the population: fitness-based selection. We will show that convergence occurs even though we ignore crossover and elitism operators in our analysis. This focalization will have the advantage of allowing us to connect the results on learning with the general results we obtained in the preceding section. In our computational experiments, we use a uniform rate of exactly one mutation in each run of the GA. Appendix A.2 specifies the structure of the GA used in our experiments.

In any population of strategies, the convergence to WE will occur in general if the selection mechanism favors strategies with higher quantities when the market outcome is below the WE and the strategies with lower quantities when the market is above the WE. Since the selection is based on a relative fitness of a strategy (whatever the particular definition of this fitness), this condition is equivalent to a necessarily increasing relationship between quantities and the relative fitness of strategies below the WE, and a decreasing relationship above the WE.

Definition 10. *If we denote the fitness of a strategy in the population of strategies by F_j , its relative fitness will be given by*

$$f_j = \frac{F_j}{\sum_j F_j} \quad (28)$$

In the context of the oligopoly, we will, in general, have a fitness function that will be positively related to the profit of the firm: $f_j = f\left(\pi_j\right)$. If the share of a strategy j in the population is noted by λ_j , the population of strategies will only be invaded by Walrasian strategies if we have the following correlation

$$\text{corr}(\lambda_j, q_j) \begin{cases} > 0 & \text{if } q_j \leq q^W \\ < 0 & \text{if } q_j > q^W \\ 0 & \text{if } q_j = q^W \end{cases}$$

In the case of the standard genetic algorithm, the selection operates through a *roulette wheel* process where the probability of reproduction of a strategy in the population is given by its relative fitness. In the case of the oligopoly, this fitness is directly related to the profits² of the firms and we will show that the way these profits are compared during reproduction and selection is *the essential difference* between social and individual learning. We first analyze the convergence to equilibrium under social learning.

4.2 Social learning

In the case of social learning, each strategy is played by a specific firm and the probability of the survival of a strategy results from the comparison of the profit of the firm i that uses this strategy (q_i) with the profits of other firms

$$f_i = \frac{\pi_i(q_i, Q_{-i})}{\sum_{i=1}^n \pi_i(q_i, Q_{-i})}$$

If we denote $\Pi_{-i} \equiv \sum_{k \neq i} \pi_k(q_i, Q_{-i})$, we can note this relative fitness as

$$f_i = \frac{\pi_i(q_i, Q_{-i})}{\Pi_{-i} + \pi_i(q_i, Q_{-i})}$$

Proposition 7. *The relative fitness function of the GA for social learning will be favorable to higher quantities which will push the population towards the WE iff*

$$\Leftrightarrow \varepsilon_{\pi_i, q_i} > \varepsilon_{\Pi_{-i}, q_i}$$

where $\varepsilon_{y,x}$ represents the elasticity of y with x .

Proof.

The relative fitness depends on the quantity corresponding to the strategy of the firm i in the following way

$$\begin{aligned} \frac{\partial f_i}{\partial q_i} &= \frac{\frac{\partial \pi_i}{\partial q_i} (\Pi_{-i} + \pi_i) - \pi_i \left(\frac{\partial \Pi_{-i}}{\partial q_i} + \frac{\partial \pi_i}{\partial q_i} \right)}{(\Pi_{-i} + \pi_i(q_i, q_{-i}))^2} \\ \frac{\partial f_i}{\partial q_i} &= \frac{\frac{\partial \pi_i}{\partial q_i} \Pi_{-i} - \pi_i \frac{\partial \Pi_{-i}}{\partial q_i}}{(\Pi_{-i} + \pi_i)^2} \end{aligned} \quad (29)$$

Hence, we observe that

$$\text{sgn} \left(\frac{\partial f_i}{\partial q_i} \right) = \text{sgn} \left(\frac{\partial \pi_i}{\partial q_i} \Pi_{-i} - \pi_i \frac{\partial \Pi_{-i}}{\partial q_i} \right). \quad (30)$$

This relative fitness will consequently be favorable to higher quantities in a way to push the population towards the WE iff

$$\begin{aligned} \text{sgn} \left(\frac{\partial f_i}{\partial q_i} \right) &> 0 \\ \frac{\partial \pi_i}{\partial q_i} \Pi_{-i} &> \pi_i \frac{\partial \Pi_{-i}}{\partial q_i} \end{aligned} \quad (31)$$

$$\begin{aligned} \Leftrightarrow \frac{q_i}{\pi_i} \frac{\partial \pi_i}{\partial q_i} &> \frac{q_i}{\Pi_{-i}} \frac{\partial \Pi_{-i}}{\partial q_i} \\ \Leftrightarrow \varepsilon_{\pi_i, q_i} &> \varepsilon_{\Pi_{-i}, q_i} \end{aligned} \quad (32)$$

²We will directly use the profits as fitness in the following.

□

Consequently, the tendency towards the WE equilibrium will depend, in this case, on the comparison of the impact of the quantity increase on a firm's profit with the same impact on the total profits of its competitors. If this impact is positive, the increase of the quantity of the firm i will increase the survival probability of this strategy if, and only if, the impact on its own profit is greater than the impact on the other firms' total profit. If this impact is negative, this is directly related to the *spite effect*: a firm that will move from a CE quantity may decrease its profit ($\varepsilon_{(\Pi_i, q_i)} < 0$), but it will decrease the competitors' profits even more ($0 > \varepsilon_{(\Pi_i, q_i)} > \varepsilon_{(\bar{\Pi}_{-i}, q_i)}$) and it will, therefore, increase its relative fitness. We must distinguish two cases in the evaluation of these impacts: infinite population and finite population market games. The latter is, of course, much more relevant to analyzing an oligopoly.

4.2.1 Infinite population of firms

When the population of firms is large, the influence of each firm is negligible, and the quantity variation of one particular firm i does not significantly modify the market price or the profits of the competitors: $\partial p(q_i, Q_{-i})/\partial q_i = 0$ and $\partial \Pi_{-i}/\partial q_i = 0$. This of course considerably simplifies the evaluation of the terms of the equation (32):

$$\varepsilon_{\pi_i, q_i} > \varepsilon_{\Pi_{-i}, q_i} = 0$$

and

$$\begin{aligned} \varepsilon_{\pi_i, q_i} &= \frac{q_i}{\pi_i} \left[p(q_i, Q_{-i}) - \frac{\partial p(q_i, Q_{-i})}{\partial q_i} q_i - \frac{dC}{dq_i} \right] \\ &= \frac{q_i}{\pi_i} (p - C') \end{aligned}$$

which implies, for $\pi_i > 0, q_i > 0$

$$\varepsilon_{\pi_i, q_i} \begin{cases} > 0 & \text{if } q_i \leq q^W \\ < 0 & \text{if } q_i > q^W \\ = 0 & \text{if } q_i = q^W \end{cases}$$

since q^W is the Walrasian strategy and it verifies, by definition, $p^W - C'(q^W) = 0$. Hence, we observe that increasing quantities will diffuse in the population when the actual outcome is below the WE, and their fitness will play against their reproduction when the industry is above this equilibrium. This is the main mechanism that pushes the industry towards the WE instead of the CE when the population of firms is very large.

4.2.2 Finite population of firms

In a finite population, the atomistic property is no longer verified. Consequently, it becomes possible to have $\partial p(q_i, Q_{-i})/\partial q_i < 0$ and $\partial \Pi_{-i}/\partial q_i \neq 0$. If we consider the latter derivative

$$\begin{aligned} \frac{\partial \Pi_{-i}}{\partial q_i} &= \frac{\partial(\sum_{j \neq i} (p q_j - C(q_j)))}{\partial q_i} \\ &= \sum_{j \neq i} \left(\frac{\partial p}{\partial q_i} q_j \right) = \frac{\partial p}{\partial q_i} \sum_{j \neq i} q_j = \frac{\partial p}{\partial q_i} Q_{-i} \\ &= p' Q_{-i} \end{aligned} \tag{33}$$

since $dC_j/dq_i = 0$ and $\partial p/\partial q_i = \partial p/\partial Q \times \partial Q/\partial q_i = \partial p/\partial Q = p'$. We can also easily observe that

$$\frac{\partial \pi_i}{\partial q_i} = p' q_i + p - C'$$

Consequently, the equation (30) becomes

$$\text{sgn} \left(\frac{\partial f_i}{\partial q_i} \right) = \text{sgn} \left((p' q_i + p - C') \Pi_{-i} - p' Q_{-i} \pi_i \right) \quad (34)$$

$$= \text{sgn} \left((p - C') \Pi_{-i} + p' (q_i \Pi_{-i} - Q_{-i} \pi_i) \right) \quad (35)$$

The first term shows that the impact of an increase of q_i depends on the position of the firm i in terms of marginal cost and price (similar to the infinite population case). However it also depends on the relative quantities and profits of the firms, in a more complicated way than can be easily linked to the spite effect. We can note that the second term in equation (35) can be reformulated as the following:

$$\begin{aligned} q_i \Pi_{-i} - Q_{-i} \pi_i &= q_i \Pi_{-i} + q_i \pi_i - q_i \pi_i - Q_{-i} \pi_i \\ &= q_i \Pi - Q \pi_i \\ &\Rightarrow \text{sgn} \left(p' (q_i \Pi_{-i} - Q_{-i} \pi_i) \right) = -\text{sgn} \left(\frac{q_i}{Q} - \frac{\pi_i}{\Pi} \right) \end{aligned}$$

since $p' < 0$. When $\Pi > 0$, this effect will favor increasing quantities if the market share of the firm is lower than its profit share. Consequently, if we have $\Pi_{-i} > 0$ and $\Pi > 0$,

$$\frac{\partial f_i}{\partial q_i} \begin{cases} > 0 & \text{iff } \pi_i > 0, q_i < q^W, \frac{q_i}{Q} < \frac{\pi_i}{\Pi} \\ \leq 0 & \text{iff } \pi_i < 0, q^C < q_i < q^W \\ = 0 & \text{iff } q_i = q^W \end{cases}$$

since $p' q_i + p - C' < 0$ in equation (34) if $q^C < q_i$, and $p = C'$, $q_i/Q = \pi_i/\Pi$ if $q_i = q^W, \forall i$. Hence, the strategies will move the quantities towards the WE if the firms can obtain positive profits in the neighborhood of WE. This result complies with Corollary 2.

Hence, we conclude that social learning will drive the market towards the WE if these basic conditions (such as the feasibility of WE), easily compatible with a standard Cournot oligopoly model, are verified. We now turn our attention to the more controversial case of GA-based individual learning.

4.3 Individual learning

When the learning of the firms is individual, the firms cannot count on directly learning strategies from competitors (imitation) through their interactions in the market. Each firm must extract the necessary information on the relevance of its current strategy from the global market variables observable by that firm, resulting from the interactions of the firms (mainly, the market price), and from the profit it obtains with its strategy under these market conditions.

Consequently, when the firm cannot observe the actions and the profits of other firms, learning dynamics are not based on imitation and the spite effect cannot directly play a part. The main question again is the consequence of these dynamics in terms of convergence to equilibrium. Since convergence to the WE with social learning is general, the capability of such a learning to allow the emergence of a CE becomes the main controversial point in this approach. This capability will strongly depend on the kind of information which the firm can obtain during its learning process.

Riechmann (2006b) says that "if agents do not know or simply neglect the state dependent nature of the problem, i.e. the fact that they do have an influence on the market price, the outcome of individual learning will be the Walras equilibrium. The reason for this is straightforward: if agents do not think or do not know they can influence the market price, the best thing they can do is to compute a best response to last period's equilibrium price". We will show that this intuition is relevant for understanding the conditions of convergence to the CE.

In this context, the learning of agents is based on the comparison of performances of different strategies inside each firm, while competition will imply an interaction at the market level between the selected strategies in each period. Consequently, the concepts of convergence and stability become more intricate. A quantity profile q^* will be stable at the market level if, and only if, every firm plays (selects) the corresponding equilibrium strategy, and every firm will play this strategy only if the learning process of the firm has effectively *converged* to it.

The coevolution of strategy populations of firms and convergence of the GA towards equilibrium, will strongly depend on the approaches used by the models for the selection process and the evaluation of the fitness. Vriend (2000), Arifovic and Maschek (2006), and other articles, discuss the convergence problem with a representation of strategy evolution based on different formulations for the genetic algorithm. Two contrasted approaches have been used in this formulation: learning with profit expectations based on the actual market price used for evaluating (through hypothetical profits) the population of strategies in each period (approach used by Arifovic (1994) and partly by Arifovic and Maschek (2006)), and learning without expectations, evaluating one strategy in each period. Other intermediate cases have also been considered, as in Vriend (2000) where the GA does not intervene in each period.

We now contrast the convergence properties of these two different learning schemes: with and without expectations.

4.3.1 Individual learning with expectation-based on hypothetical profits

In this case, the firms use expectations based on the last period price for evaluating the fitness of each strategy in their population. This hypothetical profit determines the relative fitness of each strategy and guides the selection process. The main algorithm of this game can be summarized as a simple pseudo-code if the oligopoly game has a duration of T periods (see Figure 1).

-
- Period 0: a population of k strategies, $q_{ij} \in q_i^0$, is drawn for each player i .
 - Period 1: each player plays (random choice) a particular strategy $q_{ij} \in q_i^0$. The corresponding market price, p^1 , is calculated.
 - While *Period* $\leq T$:
 - The hypothetical profit of each strategy $q_{ij} \in q_i^{t-1}$ is calculated with this price:

$$\pi_{ij}(q_{ij}, p_t) = p^{t-1}q_{ij} - C(q_{ij}), \text{ for all } i.$$
 - Using GA procedures (selection, crossover, and mutation) in each strategy population, a new population of strategies q_i^t is defined for each firm based on these profits.
 - Again the hypothetical profits are calculated $\{\pi_{i1}(q_{i1}, p^{t-1}), \dots, \pi_{ik}(q_{ik}, p^{t-1})\}$.
 - From this set, one (possibly the best) strategy is drawn in order to be played: q_{ij}^t
 - The new market price: p^t is calculated.
-

Figure 1: Pseudocode of the learning model with hypothetical profits

It is important to note that in each period t , each strategy is evaluated using the same given price p^{t-1} . Thus, with the roulette wheel selection, the relative fitness of a given strategy j of the player

i would correspond to:

$$\begin{aligned}
f_{ij} &= \frac{\pi_i(q_{ij}, p^{t-1})}{\sum_{j=1}^k \pi_i(q_{ij}, p^{t-1})} \\
&= \frac{\pi_i(q_{ij}, p^{t-1})}{\pi_i(q_{ij}, p^{t-1}) + \sum_{l=1, l \neq j}^k \pi_i(q_{i,l}, p^{t-1})} \\
&= \frac{\pi_i(q_{ij}, p^{t-1})}{\pi_i(q_{ij}, p^{t-1}) + \Pi_{i,-j}}
\end{aligned}$$

where $\Pi_{i,-j} = \sum_{l=1, l \neq j}^k \pi_i(q_{i,l}, p^{t-1})$.

Since now q_{ij} cannot modify the given price p^{t-1} , the total impact of q_{ij} on f_{ij} is given by

$$\begin{aligned}
\frac{\partial f_{ij}}{\partial q_{ij}} &= \frac{\left(\frac{\partial \pi_i(q_{ij}, p^{t-1})}{\partial q_{ij}}\right) \Pi_i - \frac{\partial \pi_i(q_{ij}, p^{t-1})}{\partial q_{ij}} \pi_i(q_{ij}, p^{t-1})}{\Pi_i^2} \\
&= \frac{\left(\frac{\partial \pi_i(q_{ij}, p^{t-1})}{\partial q_{ij}}\right) \Pi_{i,-j}}{\Pi_i^2} \\
&= \frac{(p^{t-1} - C') \Pi_{i,-j}}{\Pi_i^2} \begin{cases} > 0 & \text{if } C' < p^{t-1} \\ < & \text{if } C' > p^{t-1} \\ = 0 & \text{if } C' = p^{t-1} \end{cases}
\end{aligned}$$

with $\Pi_i = \sum_{l=1}^k \pi_i(q_{i,l}, p^{t-1})$ and if $\Pi_{i,-j} > 0$.

Consequently, the best relative fitness will be attained under the same conditions as the Walrasian pricing rule ($C' = p$) since:

- during the roulette wheel procedure of the GA, given a price p :
 - higher quantities will spread in the population, as long as $p > C' \Leftrightarrow q_j < q^W$
 - lower quantities will diffuse if $p < C' \Leftrightarrow q_j > q^W$
- the strategy that will be played will be the closest one to the competitive quantity if the selection of the strategy to play is based on maximal possible fitness: q_j such that $j \in \max_{q_{il}} f_{il}, \forall l$

We have already shown above that there is a strong relationship between the evolution of fitness and ESS.

Proposition 8. *Under hypothetical individual learning, the CE is not evolutionarily stable, while the WE is.*

Proof.

Assume that all players except i are already at the CE solution and that they will stick to it, $Q_{-i}^C = \{q_1^C, \dots, q_{i-1}^C, q_{i+1}^C, \dots, q_n^C\}$. In this case, the learning of player i is identical to the learning of a monopoly that ignores the exact price relationship $p(q_i, Q_{-i}^C)$. The hypothetical learning procedure leads to an evaluation of the profit at period t that is based on the previous price p_{t-1} : $\pi_{i,t}(q_{ij}, p_{t-1}) = p_{t-1} q_{ij} - C(q_{ij})$. It is evident that the fitness evolution, as a maximization of the hypothetical profit, will lead to the Walrasian condition: $C' = p_{t-1}$. Consequently, the CE equilibrium is not ESS, while the WE is. \square

We use computational experiments to illustrate our results for this section and for the following section. The details of the simulation protocol are given in Appendix A. Figure 2 represents the

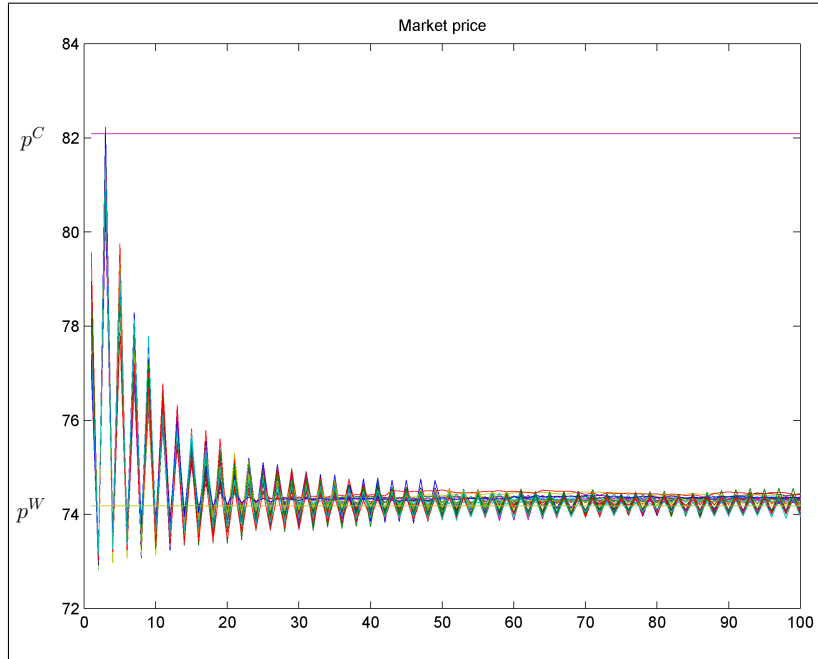


Figure 2: Convergence to WE price with hypothetical learning

evolution of market price under expectations-based individual learning. It clearly shows that the market price quickly converges to WE price in all simulations. In accordance with the theoretical results, this kind of learning cannot yield convergence to CE. Consequently, the numerical results obtained by Arifovic (1994) and by Arifovic and Maschek (2006) for this type of learning are perfectly normal.

We will now consider a type of learning that is based on the direct use of strategies, as in Vriend (2000) (or in Yildizoglu (2003) under different circumstances).

4.3.2 Individual learning without expectations

In this approach, the firm does not use the previous period's price, and it must experiment with strategies in order to evaluate their relative fitness. Such a learning process precludes any capacity of firms to generalize from their experience, in order to form expectations. The cognitive capacity of the firm and its forward looking ability are quite limited in this case. In such a framework, the only source of learning for the firm is the effective utilization of each strategy in its population.

If the environment of the firm (the structure of the market and the strategies used by the other firms) is completely stable, using once each strategy in the population would be sufficient to discover the relative fitness of this strategy. In a multi-player game, such as the Cournot oligopoly, the environment will not be stable as long as other firms continue to learn and change their quantities as a consequence of this learning. In such dynamic setups, each strategy must be used for a sufficiently long time, in order for the firm to discover its fitness against the strategies of other firms. This is the experimentation period or, as it is called in training based learning mechanisms, an *epoch*. Vriend (2000) introduces into the GA such a training period, named the *GARate*, of length 100 (the GA modifies the population of strategies once every *GARate* period). We will show below why this is an important condition in the convergence to the Cournot equilibrium. We will then establish, using computational experiments, the length of the epoch, in our special setup, that would assure the convergence to equilibrium under different selection mechanisms. We will show that the rather special learning setup used in Vriend (2000) is necessary in order to obtain the convergence to the CE with individual learning, and that the experimentation period is a very important dimension of

the convergence results when agents' learning is not forward looking.

In this new experimental setup, each population of strategies is played without modification during τ periods (an epoch). A special case would be an epoch of k periods ($k = \text{number of chromosomes}$) where each strategy is played only once, following a sequential order, but one can consider more sophisticated configurations. During an epoch, strategies from given populations meet each other following the actual strategy-selection scheme used by the firms (random, roulette wheel, play the best, etc.). Profits attributed to different strategies are updated following this matching process and, at the end of the epoch, the GA modifies populations based on these profits. Figure 3 gives the pseudo-code of this configuration. Profits are randomly initialized and they are only updated when they are used in the market or when the corresponding quantities are changed by a mutation.

-
- Period 0: a set of k strategies, $q_{ij} \in q_i^0$, is (randomly) drawn for each player i .
 - While *Period* $\leq T$:
 - if (*Period* (mod τ) = 0): using GA procedures, as the roulette wheel selection, a new set of strategies, q_i^ϕ , is defined for each firm. Profits of new strategies resulting from mutations are randomly drawn around the average profit of the firm's strategy population. Other strategies carry on their fitness from epoch to epoch.
 - each player plays a strategy $q_{ij} \in q_i^\phi$. The corresponding market price, p_t , is calculated. The realized profit is calculated $(\pi_{i,t}(Q_t, p_t))$.
-

Figure 3: Pseudocode of learning model without expectations

Consequently, this learning model differs from the hypothetical profit case in one important aspect. Each strategy is evaluated using the effectively observed profit that it yields by repeatedly meeting the strategies of other firms. This observed profit necessarily results from the market price corresponding to these strategies, and it can now include some information on the demand curve if the same population of quantities can meet different, but given, quantities of the competitors. This obviously will only be possible if τ is long enough in comparison with the strategy population size.

Given a τ , the strategies of the firm i will meet τ times the strategy populations of other firms $\{q_{lj}\}_{l=1\dots n, l \neq i}$. At the end of an epoch, and before the intervention of the GA, the average profit of an individual strategy q_{ij} will be given by

$$E[\pi_i(q_{ij}, Q_{-i})] = E[p(q_{ij}, Q_{-i})]q_{ij} - C(q_{ij}) = \bar{p}q_{ij} - C(q_{ij}) \quad (36)$$

with

$$\bar{p} = \sum_{l \neq i} p\left(q_{ij}, \sum_l q_{lj}\right) \cdot f(q_{lj}|\tau)$$

where $f(q_{lj}|\tau)$ is the frequency for q_{ij} of meeting each individual strategy of the other firms, conditioned by the strategy selection process (the selection of the strategies to be played on the market), and the length of the epoch during which the strategy populations remain constant: τ . When τ is very short in comparison with the strategy population size (k), then q_{ij} will only be able to meet a very limited subset of strategy profiles of other firms and i will have only partial information about the price mechanism. On the contrary, if τ is long enough, q_{ij} is able to meet all strategies of other firms and the firm i will have a precise idea about the price it can expect with different quantities in its population and the corresponding expected profits. We indeed observe that

$$\lim_{\tau \rightarrow \infty} f(q_{lj}|\tau) = f(q_{lj})$$

which is the real empirical frequency of each strategy in the strategy population of the firm l . The average profit of equation (36) will correspond in this case to the expected profit with the strategy $q_{ij}, \forall j = 1..k$. Consequently, given the populations of other firms, the average profit of q_{ij} that results from the matching process will imply in this case $\partial\bar{p}/\partial q_{ij} < 0$, and the selection process will be able to use this information about the demand curve. The strategies that correspond to $C' = p$ will then be dominated by strategies that are closer to the condition $C' = \bar{p} + q_{ij}\partial\bar{p}/\partial q_{ij} < \bar{p}$. This mechanism will push the strategies towards the CE instead of the WE.

Again we use computational experiments to illustrate these mechanisms. From the preceding discussion, the results on convergence will establish the role of two main dimensions of the learning process: the length of the epoch, τ and the mechanism that is used to select, in the strategy population of each firm, the quantity that is played on the market (*see* Appendix A.2). We study in three steps the convergence to equilibria in this setup.

First, we give some illustrative results in Figure 4: the results of simulations (a batch of 20 simulations with each configuration) for the baseline configuration of the oligopoly (*see* Appendix A.1) and the roulette wheel selection. In each configuration we run the GA exactly 300 generations (the duration of the simulations is hence equal to $T = 300\tau$). These first results clearly show that the convergence of the market price to the CE price is favored by longer τ : none of the 20 runs converge for $\tau = 1$, and only a weak convergence is observable for $\tau = 10$. Longer epochs are necessary for a clear convergence (especially, when $\tau = 100$ —the case of Vriend (2000)—the convergence is observed in all 20 runs).

The results of Figure 4 are aggregate and they do not indicate the dispersion that could hide behind these results: are the quantities of different firms close to the CE? For a more rigorous proof of convergence, we give, in the second step, more systematic results summarized in Table 1. To obtain these results, we simulate 20 runs with a GA of 200 generations in each run. We study the convergence of the best quantity in the strategy population of firms (the strategy with the highest relative fitness in the strategy population of each firm) to the Cournot equilibrium. We check this convergence by testing the null hypothesis that the observed average of these quantities in the last period of each run is equal to the CE quantity ($H_0 : q = q^C$). We use Student's t-tests for checking if we can reject this null hypothesis. If for a run we can not reject the hypothesis, we consider that convergence has occurred for this run. We also give the p -values corresponding to the probability of accepting by error this null hypothesis. In order to show the importance of the strategy selection mechanisms, we compare the results obtained using three different methods: Random Selection (**RS**), Roulette Wheel Selection (**RWS**) and Play the Best (**PTB**) (*see* Appendix A.2 for more details).

The first two lines of each part of the table show that if the strategies do not meet frequently enough ($\tau \leq 10$), the information on demand is quite difficult to extract and firms have some difficulty in converging to CE, no matter what the selection mechanism used for choosing the played quantities. This effect is stronger for the random selection (RS) mechanism since the matching of the strategies is not oriented at all in this case. We observe that other mechanisms already fare quite well, even for $\tau = 10$.

Matching mechanisms integrating some intrinsic randomness (in complement to the randomness implied by the mutations) give rise to convergence to the CE price when τ is higher. The PTB mechanism that systematically plays for each firm its supposedly best strategy of the period, has some difficulty in converging: it has a tendency to over-shoot by over estimating the strength of dp/dq_{ij} and hence by over reducing the production of firms. The convergence is not systematic with this mechanism, even for quite long epochs ($\tau = 50; 75$). RWS assures systematic convergence as soon as $\tau \geq 25$. We observe in the last column of the table that the likelihood of the convergence to the CE clearly increases with τ and highest values are observed for the RWS mechanism.

In the last step, we consider the dispersion of quantities in the strategy populations of the firms (and not only for the best strategy in each population). Figure 5 exhibits the distribution of quantities observed in the last epochs of 20 simulations using the RWS and $\tau = 100$. Figure 5-(a) gives the distributions observed in the period $t = 20000$ (corresponding to the last epoch of 20000

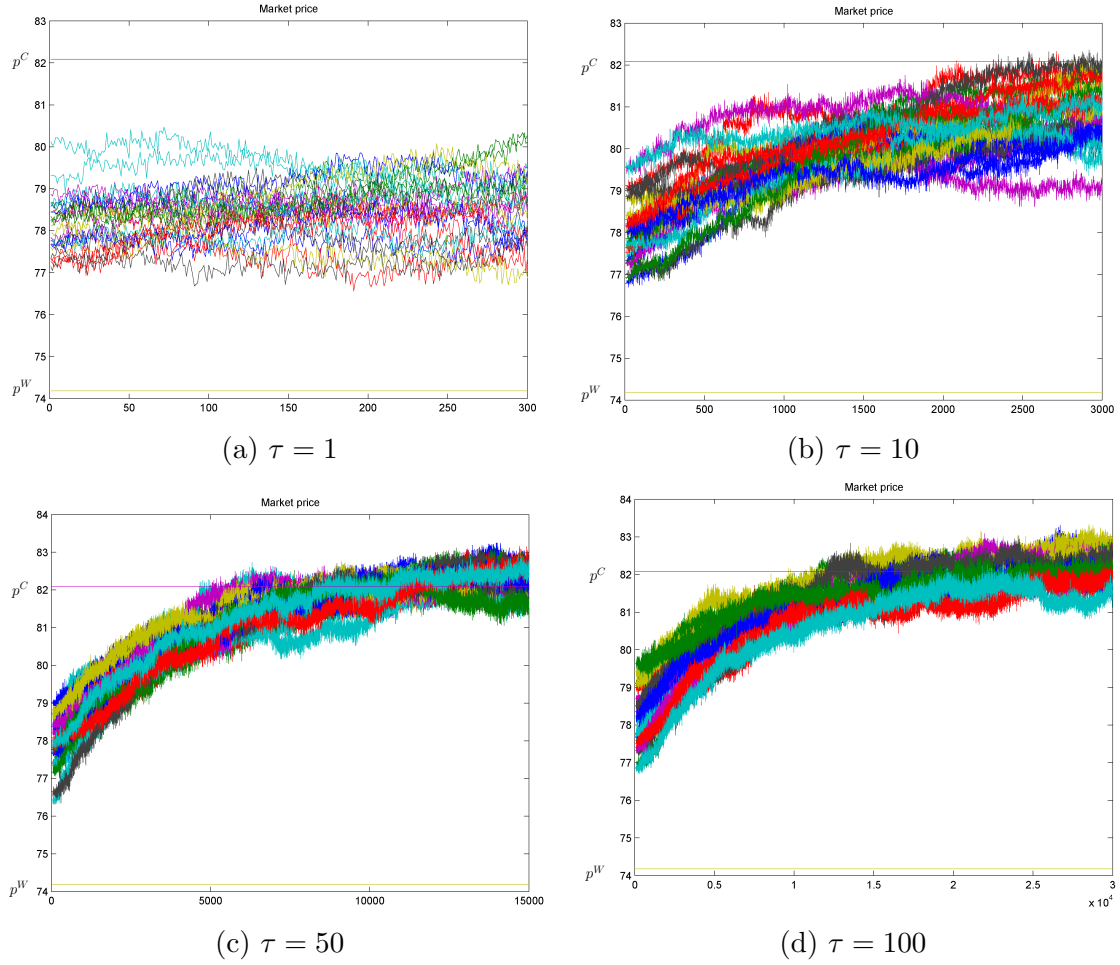


Figure 4: Evolution of the market price: Convergence to the Cournot equilibrium price: Individual learning without expectations and roulette wheel selection

Selection	τ	Convergence to CE	Average $p - value$
RS	1	1/20	0
	10	3/20	0.04
	25	12/20	0.19
	50	20/20	0.58
	75	20/20	0.67
	100	20/20	0.64
PTB	1	3/20	0.019
	10	15/20	0.172
	25	20/20	0.45
	50	19/20	0.56
	75	19/20	0.67
	100	20/20	0.56
RWS	1	0/20	0
	10	10/20	0.09
	25	20/20	0.6
	50	20/20	0.74
	75	20/20	0.76
	100	20/20	0.79

Table 1: Testing convergence to equilibrium (20 runs with each configuration, Student's t-test with $\alpha = 5\%$)

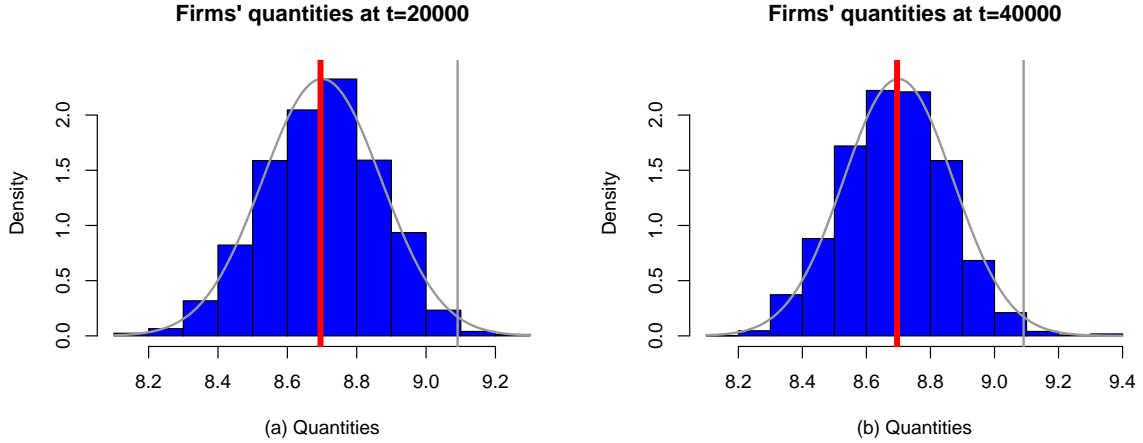


Figure 5: Distribution of quantities in all strategy populations. Observations from the last epochs of 20 runs for two different simulation lengths

	a	b	c	d	n	τ
$p - value$	+				+	+

Table 2: Determinants of convergence to CE

periods simulations, with 200 generations for the GA) and Figure 5-(b), in $t = 40000$ (last epoch of 40000 periods simulations, with 400 generations for the GA). The thick vertical line represents q^C , while the thin one corresponds to q^W . The gray curves give estimated normal densities for each distribution. These distributions show that our results are not subject to generic drift, and that we do not observe a divergence from the CE in longer simulations. On the contrary, the distribution is even more tightly centred on $q^C = 8.6957$ in the last observation set. The observed variance is mainly due to the constant and strictly positive mutation rate; this is shown by the strong symmetry of the distribution in this case.

These results show that the convergence to CE in the case of individual learning can be obtained even under less specific assumptions about the Cournot game than in Vriend (2000). A simple quadratic cost oligopoly game and an elementary, real value-based GA with only selection and low mutation is sufficient to observe convergence in this case.

We have also analyzed in a more systematic way, factors that influence the convergence to the CE with RWS. These factors are the demand and cost parameters (a, b, c, d), τ and the number of firms, n . We have run 1300 simulations with randomly drawn values for these parameters (see Appendix A.3). The results of Student's t-test, comparing the mean of best production levels of the last period to q^C , indicate that in 97.3% of these simulations we cannot reject the null hypothesis of convergence to CE ($H_0 : \bar{q} = q^C$). This is quite a strong result given that 51.1% of the simulations correspond to $\tau \leq 25$. We also determine the influence of these factors on the $p - value$ for the convergence test to CE. We remind you that the higher this value, the stronger the likelihood of convergence (non-rejection of H_0). Table 2 gives the factors that are significant at 5% and the direction of their influence on $p - value$. At a very global level, and under the assumption of linearity, these results show that factors favoring convergence are: a higher demand slope, a higher number of firms and a higher epoch length. The result on τ is perfectly in line with preceding theoretical results. The result on the number of firms can seem paradoxical (a higher number of firms is supposed to favor WE), but one should not forget that a higher number of firms also corresponds here to a higher capacity of exploration for the learning of the industry. A higher demand slope is favorable since it discriminates more clearly between quantities from the point of view of their impact on the market price.

5 Conclusion

The main objective of this article is to gain some general understanding of the convergence to equilibrium in an oligopoly game with learning dynamics. We have established the properties, in terms of evolutionary stability, of two types of potential equilibrium in this game: the Cournot equilibrium (CE) and the Walrasian equilibrium (WE). The first part of the article shows that the WE is quite robustly stable under general conditions when learning is based on imitation and random experimenting (mutations). This result stems from the spite effect that appears when learning possesses a social dimension (as in the case of imitation-based learning) and when dynamics are based on selection, and hence on the relative performance of firms. With the spite effect, learning through imitation and mutation diverts the attention of firms from their own profits and prevents the emergence of dynamics based on best replies that could assure the evolutionary stability of the CE. The WE becomes the only ESS equilibrium in this case. We must also consider the possibility of multiple mutations, since they better correspond to the *playing the field* aspect of the Cournot game (each firm meets in each period all other firms, and not a single, randomly drawn competitor). We show in this case that dynamics and stability are very difficult to analyze under general conditions. The second part of the article consequently focuses on a special setup that emerged in the literature for analyzing learning dynamics: learning through Genetic Algorithms (GA). This focus is interesting because this setup has yielded some puzzling results and controversies about the convergence to the CE. We first study the general conditions under which the selection in the GA can promote convergence to the CE. We show that when the GA represents social learning, the convergence can only occur to the WE. In the case of individual learning, the convergence to the CE is only possible if the interactions of the firms allow them to discover the decreasing relationship that prevails between the market price and their quantities. We show that this is not possible with learning based on hypothetical profits as considered by Arifovic in several articles and, more recently, in Arifovic and Maschek (2006). Some direct confrontation of quantities is necessary to assure the discovery of the impact of the quantities of firms on the market price, and to secure the convergence to the CE. We show, in a computational setup close to Vriend (2000), but under quite general cost and GA conditions, that the convergence to the CE arises when firms' strategies have enough opportunity to meet each other before the intervention of the GA. We also carry systematic Monte Carlo explorations that shed light on the influence of different parameters on this convergence. Our results clarify the results obtained in Vriend (2000) and Arifovic and Maschek (2006).

References

- Alós-Ferrer, C. and A. B. Ania, (2005), The Evolutionary Stability of Perfectly Competitive Behavior, *Economic Theory*, vol. 26:3, pp. 497–516.
- Alós-Ferrer, C. and A. B. Ania, (2006), Evolutionary Stability in Finite Populations : A survey of the recent literature and some additional results, Working Paper.
- Arifovic, J., (1994), Genetic Algorithm Learning and the Cobweb Model, *Journal of Economic Dynamics and Control*, Special Issue on Computer Science and Economics, vol. 18:1, pp. 3–28.
- Arifovic, J. and M. Maschek, (2006), Revisiting Individual Evolutionary Learning in the Cobweb Model - An Illustration of the Virtual Spite Effect, *Computational Economics*, vol. 28, pp. 333–354.
- Bergin, J. and D. Bernhardt, (2004), Comparative Learning Dynamics, *International Economic Review*, vol. 45:2, pp. 431–465.
- Dawid, H., (1999), *Adaptive Learning by Genetic Algorithms*, Springer.
- Goldberg, D. (1989) *Genetic Algorithms in Search, Optimization and Machine Learning*. New York: Addison-Wesley Publishing Company.
- Hamilton, W.D., (1970), Selfish and Spiteful Behavior in an Evolutionary Model, *Nature*, vol. 228, pp. 1218–1220.
- Hamilton, W.D., (1971), Selection of selfish and altruistic behavior in some extreme models. In : *Man and Beast: Comparative Social Behaviour*, J. F. Eisenberg and W. S. Dillon, eds, pp. 57–91.

- Smithsonian Press, Washington, DC.
- Holland, J.H., 1992, *Adaptation in Natural and Artificial Systems. An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, 2nd ed. MIT Press, Cambridge, MA.
- Huck, S.; H. Th. Normann; J. Oechssler (1999) Learning in Cournot Oligopoly - An Experiment, *The Economic Journal*, vol. 109, No. 454, Conference Papers, pp. 80–95.
- Kirman A. P., (1995), Learning in Oligopoly: Theory, Simulations and Experimental Evidence, in Kirman, A.P. and Salmon, M. (eds), *Learning and Rationality in Economics*, Blackwell Publishers, Oxford.
- Maynard-Smith J., (1982), *Evolution and the theory of games*, MIT Press, Cambridge MA.
- Riechmann, Th., (2006a) Mixed motives in a Cournot game. *Economics Bulletin*, vol. 4:29), pp. 1–8.
- Riechmann, T., (2006b), Cournot or Walras? Long-Run Results in Oligopoly Games, *Journal of Institutional and Theoretical Economics*, vol. 127(4), pages 702–720.
- Riley, J. G., (1979) Evolutionary equilibrium strategies, *Journal of Theoretical Biology*, vol. 76:2, pp. 109–123.
- Schaffer, M., (1988), Evolutionarily Stable strategies for a Finite Population and a Variable Contest Size, *Journal of Theoretical Biology*, vol. 132, pp. 469–478.
- Schaffer, M., (1989), Are Profit-Maximisers the Best Survivors?, *Journal of Economic Behavior and Organization*, vol. 12, pp. 29–45.
- Stegeman M. and P. Rhode, (2004), Stochastic Darwinian equilibria in small and large populations, *Games and Economic Behavior*, vol. 49, 171–214.
- Theocharis, R. (1960). On the stability of the Cournot solution on the oligopoly problem, *Review Economic Studies*, vol. 73, pp. 133–4.
- Vallée T. and T. Basar, (1999), Off-line computation of Stackelberg solutions with the genetic algorithm, *Computational Economics*, vol. 13:3, p.201–209.
- Vallée, T. and M. Yildizoglu, (2004), Présentation Des Algorithmes Génétiques et de Leurs Applications En Économie, *Revue d'Économie Politique*, 114, 711–745.
- Vega-Redondo, F. (1997). The evolution of Walrasian behavior, *Econometrica*, vol. 65, pp. 375–84.
- Vincent Th. and J.S. Brown, (2005), *Evolutionary Game Theory, Natural Selection, and Darwinian Dynamics*, New York: Cambridge University Press.
- Vriend, N., 2000, An illustration of the essential difference between individual and social learning, and its consequences for computational analyses, *Journal of economic dynamics and control*, vol. 24, 1–19.
- Weibull J., (1995), *Evolutionary Game Theory*. Cambridge (USA): MIT Press.
- Weibull J., (2006), *Evolutionary stability and the replicator dynamic*, Lecture notes, Stockholm School of Economics.
- Yildizoglu, M., (2002), Competing R&D Strategies in an Evolutionary Industry Model, *Computational Economics*, vol. 19, pp. 52–65.

A Appendix

A.1 Baseline oligopoly model of the experiments

We define a simple general oligopoly game with

$$\begin{aligned}
 P(q_i, Q_{-i}) &= a - b(q_i + Q_{-i}), \quad a, b > 0 \\
 C_i(q_i) &= F + cq_i + dq_i^2 \\
 F, c &\geq 0, d \leq 0.
 \end{aligned}$$

For a n -firm oligopoly, the symmetrical equilibria are given as follows (with $Q_{-i} = (n-1)q_{-i}$):

$$\begin{aligned}
 \text{Cournot-Nash equilibrium (CE): } q_i^C &= \frac{a-c}{b+2d+bn} \\
 \text{Walrasian equilibrium (WE): } q_i^w &= \frac{a-c}{2d+bn} > q_i^C
 \end{aligned}$$

In our simulations we use the following numerical specification for the oligopoly:
 $a = 256, b = 1, F = 0, c = 56, d = 1, n = 20$.

$$\begin{aligned} \text{CE: } q_i^C &= 200/23 = 8.6957, p^C = 1888/23 = 82.087 \\ \text{WE: } q_i^W &= 100/11 = 9.0909, p^W = 816/11 = 74.1818182 \end{aligned}$$

A.2 Specifications of the GA

In all simulations we use the same genetic algorithm (GA): a real value-based GA with 20 chromosomes. We keep the number of generations constant over all simulations in order to let other parameters determine the learning capacity of firms. In order to keep a GA structure as close as possible to our theoretical results, we use a GA based only on selection (roulette wheel selection) and on mutations. We use a uniform rate of exactly one mutation in each run of the GA. In the case of social learning, this mutation takes place in the firm population. In the case of individual learning each mutation takes place in the strategy population of every firm (hence we have as many mutations as the number of firms in each intervention of the GA). The chromosomes of the first generation for each firm are randomly generated in the interval $[(1 - \gamma)q^C, (1 + \gamma)q^W]$, with $\gamma = 1\%$. In the case of the learning without expectations (**LWE**, epoch-based learning), we have randomly generated the profits corresponding to this first generation in the interval $[0, 10]$. Mutations introduce new strategies in the population and they are Gaussian, centered around the individual strategy population average

$$\hat{q}_{ij} = \bar{q}_i + \sigma \mathcal{N}(0, 1)$$

with initial profits drawn using the same structure for the LWE case

$$\hat{\pi}_{ij} = \bar{\pi}_i + \sigma \mathcal{N}(0, 1).$$

We use $\sigma = 5\%$ for the LWE case and $\sigma = 1\%$ in the learning with hypothetical profits case (**LHP**). In the LWE the profits associated to the quantities are carried from period to period (and from epoch to epoch) except when these quantities are used and the corresponding profits are updated or when they have mutated and inherited a new random profit level.

We use the following scaled fitness function for our experiments:

$$F_{ij} = \exp\left(100 \max\left\{0, \frac{\pi_{ij}}{\pi_i^*}\right\}\right)$$

where π_i^* is the maximal profit in the individual population of strategies.

In the LHP case, only the best strategy is played in accordance with Arifovic and Maschek (2006).

In the learning without expectations case, three strategy selection mechanisms are tested:

- random selection (**RS**): a strategy is uniformly selected in the population to be played on the market;
- roulette wheel selection (**RWS**): the probability of selection of a strategy is proportional to its relative fitness, $f_{ij} = F_{ij} / \sum F_{ij}$;
- play the best selection (**PTB**): the strategy with the highest relative fitness is played.

In this case, the populations of strategies are played without the intervention of the GA for a given number, τ , of periods (the *GArate* in Vriend (2000)). We test the following cases: $\tau \in \{1, 10, 25, 50, 75, 100\}$. The last case corresponds to the assumption used by Vriend (2000).

We run 20 simulations with each configuration and the graphics of the Figure 4 give the evolution of the market price in these 20 simulations. In Table 1, we give convergence results using Student's t-test in each configuration. We respectively test that the average of the firms' best strategies in the last period is equal to q^C . For each configuration, we count cases for which the convergence can be accepted with $\alpha = 5\%$, over the set of 20 simulations. We also give the average p -value in each configuration.

A.3 Monte Carlo simulations

We proceed with Monte Carlo simulations for a systematic analysis of determinants of convergence to equilibrium. We run 1,300 simulations where the main parameters of the selection mechanism and of the Cournot oligopoly are randomly drawn in given intervals:

$$a \in [55, 300], b \in [0.5, 2], d \in [0, 5], c \in [0, 50], n \in \{5, 7, 10, 12, 15, 17, 20\}, \tau \in \{10, 25, 50, 100\}.$$

The GA is run with 300 generations. We otherwise use the baseline values of the parameters and the RWS mechanism.

Results of the regression on p – value:

```
lm(formula = pvalueqc ~a + b + c + d + nbFirms + tau)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.72121	-0.21096	0.02297	0.21345	0.55502

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.871e-01	4.600e-02	6.241	6.44e-10 ***
a	-4.785e-06	1.200e-04	-0.040	0.96819
b	5.573e-02	1.935e-02	2.880	0.00406 **
c	3.200e-04	5.855e-04	0.547	0.58479
d	-2.598e-03	5.975e-03	-0.435	0.66382
nbFirms	1.527e-02	1.669e-03	9.148	< 2e-16 ***
tau	8.169e-04	2.481e-04	3.292	0.00103 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2679 on 993 degrees of freedom

Multiple R-squared: 0.09496, Adjusted R-squared: 0.0895

F-statistic: 17.37 on 6 and 993 DF, p-value: < 2.2e-16

A.4 Oligopoly model of Vriend (2000)

Vriend (2000) adopts a very specific configuration for the oligopoly model:

$$P(Q) = -1 \times 10^{-97} - 1.5 \times 10^{95} Q^{-39.99999997}$$

$$C_i(q_i) = -4.097 \times 10^{-94}$$

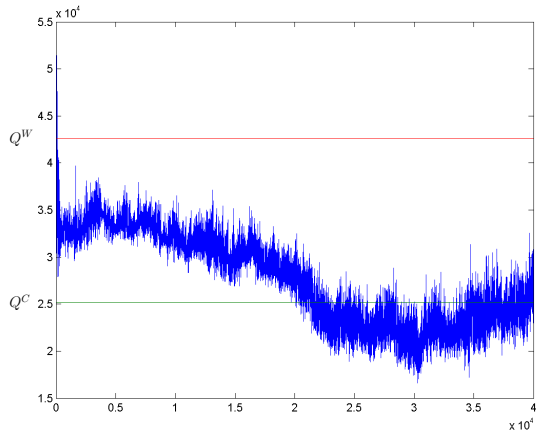
Under these specifications, we have an oligopoly model very strictly tailored for $N = 40$ firms. The equilibria are given by

$$\text{CE: } q_i^C = 942.4, p^C = 1.3333 \times 10^{-088}$$

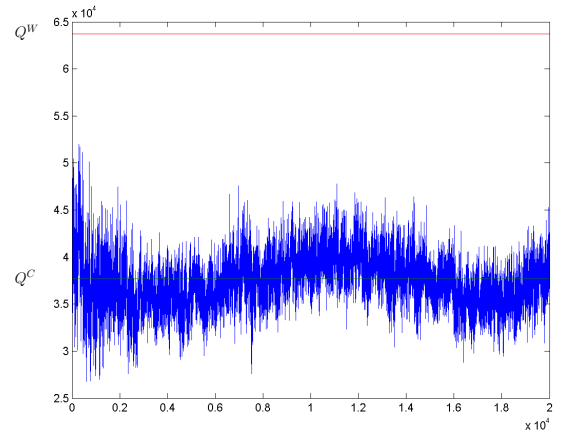
$$\text{WE: } q_i^W = 1593.5, p^W = 0$$

This is definitely not an ordinary oligopoly setup, and it cannot really be called *well behaved*. Under these very specific conditions, the convergence to CE is not easily obtained, even with some fine-tuning of the GA. More specifically, none of the selection mechanisms that we consider in the baseline model (RS, RWS or PTB) is able to converge to CE under reasonable conditions. One needs the exact tournament selection of Vriend (2000) to obtain a convergence such as in the Figure 6-(a) where the total quantity produced by the firms is plotted across time and one can observe that the oligopoly remains in the neighborhood of CE instead of converging to WE.

We also test the convergence under the alternative cost setup analyzed by Arifovic and Maschek (2006). Following our general conclusions, convergence to CE must also be observed under individual learning with this configuration. The cost function is the only difference with the baseline setup of Vriend (2000): $C_i(q_i) = -2.060 \times 10^{-84} + 1.0 \times 10^{-90} q_i$, $C' > 0$. Contrary to the conclusion of Arifovic and Maschek, we effectively observe convergence to CE, as it is shown in Figure 6-(b).



(a) - Baseline setup



(b) - Alternative setup ($C' > 0$)

Figure 6: Convergence in Vriend(2000) ($\tau = 100$)