Convergence in Finite Cournot Oligopoly with Social and Individual Learning

*Thomas VALLEE**

*Murat YILDIZOGLU***

* *LEN – IAE Nantes*

** *GREThA UMR CNRS 5113*

yildi@u-bordeaux4.fr

Cahiers du GREThA
n° 2007 – 07
Juin 2007

GRETHA UMR CNRS 5113
Université Montesquieu Bordeaux IV
Avenue Léon Duguit - 33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75 - Fax : +33 (0)5.56.84.86.47 - www.gretha.fr

Convergence dans l'oligopole de Cournot fini avec apprentissage social et individuel

Résumé

La convergence vers l'équilibre de Nash dans l'oligopole de Cournot est un problème qui apparaît de manière récurrente dans les études économiques. Le développement des jeux évolutionnaires a permis l'utilisation d'un concept d'équilibre en adéquation avec les dynamiques d'ajustement et la stabilité évolutionnaire de l'équilibre de Cournot a été étudiée par plusieurs articles. Ces articles montrent que l'équilibre Walrasien est la seule solution évolutionnairement stable du jeu de Cournot. Vriend(2000) propose l'utilisation des algorithmes génétiques pour l'étude des dynamiques d'apprentissage et il obtient la convergence vers la solution de Cournot avec l'apprentissage individuel. Nous montrons dans cet article pourquoi l'apprentissage social conduit à la solution de Walras et comment l'apprentissage individuel peut effectivement permettre la convergence vers la solution de Cournot. De plus, ces résultats sont obtenus dans un cadre plus général avec une application à l'aide d'expériences informatiques.

Mots-clés : Oligopole de Cournot ; Apprentissage ; Evolution ; Sélection ; Stabilité évolutionnaire ; Equilibre de Nash ; Algorithmes génétiques

Convergence in Finite Cournot Oligopoly with Social and Individual Learning

Abstract

Convergence to Nash equilibrium in Cournot oligopoly is a problem that recurrently arises as a subject of study 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 studied by several articles. Several articles show that the Walrasian equilibrium is the stable evolutionary solution of the Cournot game. Vriend (2000) proposes to use genetic algorithm for studying learning dynamics in this game and obtains convergence to Cournot equilibrium with individual learning. We show in this article how social learning gives rise to Walras equilibrium and why, in a general setup, individual learning can effectively yield convergence to Cournot instead of Walras equilibrium. We illustrate these general results by computational experiments.

Key words: Cournot oligopoly; Learning; Evolution; Selection; Evolutionary stability; Nash equilibrium; Genetic algorithms

JEL : L130; L200; D430; C630; C730

Convergence in Finite Cournot Oligopoly with Social and Individual Learning

Thomas Vallée[λ] and Murat Yıldızoğlu[θ]

[λ]LEN
IAE - Faculté des Sc. économiques et de gestion.
Chemin de la Censive du Tertre.
BP 52231. 44322 Nantes Cedex 3.
Thomas.Vallee@univ-nantes.fr

[θ]GREThA (UMR CNRS 5113)
Université Montesquieu Bordeaux IV
Avenue Léon Duguit.
F-33608 Bordeaux
yildi@u-bordeaux4.fr

July 9, 2007

Abstract

Convergence to Nash equilibrium in Cournot oligopoly is a problem that recurrently arises as a subject of study 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 studied by several articles. Several articles show that the Walrasian equilibrium is the stable ESS of the Cournot game. Vriend (2000) proposes to use genetic algorithm for studying learning dynamics in this game and obtains convergence to Cournot equilibrium with individual learning. We show in this article how social learning gives rise to Walras equilibrium and why, in a general setup, individual learning can effectively yield convergence to Cournot instead of Walras equilibrium. We illustrate these general results by computational experiments.

JEL codes: L130; L200; D430; C630; C730.
Keywords: Cournot oligopoly; Learning; Evolution; Selection; Evolutionary stability; Nash equilibrium; Genetic algorithms.

1 Introduction

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

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 studied by several articles. Vega-Redondo (1997) shows, for example, that the Walrasian equilibrium (**WE**) is the unique stable ESS situation in a quantity competition game with a homogenous product. The results of this literature strongly enhanced the doubts about the convergence of adjustment dynamics to Cournot equilibrium (**CE**). At the core of this literature lay the selection dynamics that are considered in the models. When one interprets these dynamics resulting from the learning process of boundedly rational firms under incomplete and imperfect information (about the demand function, the quantities of the competitors and their costs), the nature of this learning becomes the central element.

Evolutionary stability concept is more naturally connected to social learning of firms, through imitation of strategies and experimentation (random mutations) but it 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 tend to imply complex dynamics, and oligopoly is more naturally formulated as a *playing the field* game than a pairwise matching process. The analysis of these dynamics is very difficult under general conditions (*see* for example Stegeman and Rhode (2004) for a partial tentative) and recent studies have focused on the consequences of specific representations of firms' learning, using computational experiments.

Vriend (2000) advanced that genetic algorithms (**GA**) can be used to demonstrate the "essential difference between individual and social learning". This article shows that 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, the convergence to the CE. The main mechanism that is proposed as the source of this contrasted result is the *spite effect* that first appeared in evolutionary biology (*see* Hamilton (1970)). This effect is already underlined by Vega-Redondo (1997) in the explanation of the main result of this article. The spite effect corresponds to the fact that an individual can harm itself but if it even more harms others, it can gain an evolutionary advantage. These results have recently been questioned by Arifovic and Maaschek (forthcoming) who draw attention to the specific implementation adopted by Vriend in terms of the mechanisms of the GA–based learning and the cost structure of firms. As a consequence, the results on this problem are definitely puzzling and some general exploration of this question is welcome. We propose to inquire in this article a more general analysis of the convergence to equilibria in Cournot oligopoly.

In a first section, we study the properties of CE and WE in terms of evolutionary stability. The second section is dedicated to the analysis of the consequences of the spite effect on the evolutionary stability. The third section analyzes the general mechanisms that derive GA-based learning and their consequences on the outcomes of social and individual learning. We show why social learning can not yield convergence to CE and why individual learning can under some important conditions. The last section concludes the article.

2 Evolutionary stability and Nash equilibria in oligopoly

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 $p = p\left(Q\right)$, where $Q = \sum_{i=1}^{n} q_i$ and $dp/dQ < 0$. The common cost function of the firms is $C\left(q_i\right)$, with $C' > 0$ and $C'' > 0$. The profit function of a firm is hence standard: $\pi_i(q_i, \ldots, q_n) = p(Q)q_i - C(q_i)$. Since the interaction between the strategies of the firms only takes place through the common inverse demand function (and hence, through the sum of these quantities), a quantity profile $(q_1, \ldots, q_n)$ can be represented, from the point of view of a firm i, as $\left(q_i, Q_{-i}\right)$, where $Q_{-i} = \sum_{j \neq i} q_j$.

In this oligopoly two different kinds of equilibria 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*

$$q_i^C = argmax_{q_i} \pi_i\left(q_{\,i}, Q_{-i}^C\right) \Rightarrow p\left(Q^C\right) + q_i^C p'\left(Q^C\right) = C'\left(q_i^C\right), \ \forall i = 1, \ldots n \tag{1}$$

2. *The market clears: $p^C = p\left(Q^C\right)$.*

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 a marginal cost pricing (it is a price-taker)*

$$q_i^W = argmax_{q_i} \pi_i\left(q_{\,i}; p^W\right) \Rightarrow C'\left(q_i^W\right) = p^W, \ \forall i = 1, \ldots n \tag{2}$$

2. *The market clears : $p^W = p\left(Q^W\right)$.*

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

$$q_i^C \leq q_i^W, \forall i \tag{3}$$
$$\Rightarrow Q^C \leq Q^W \text{ and } p^C \geq p^W \tag{4}$$
$$\pi_i^C = \pi_i\left(q_i^C, Q_{-i}^C\right) \geq \pi_i\left(q_i^W, Q_{-i}^W\right) = \pi_i^W, \forall i \tag{5}$$

As a consequence, the Cournot equilibrium is preferred by each firm and the consumer's surplus is maximal in the Walrasian equilibrium. The possibility that, under some selection dynamics, an oligopoly where each individual firm is profit maximizer can nevertheless converge to WE instead of CE is hence quite a paradoxical result that explains why this debate has been continuing for such a long time in economics. This paradoxical result is established through the analysis of the evolutionary stability of these equilibria.

2.2 Evolutionary stability of equilibria

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:*

$$\pi_i(q_i, Q^*_{-i}) \leq \pi_i(q^*_i, Q^*_{-i}), \forall q_i, \forall i \tag{6}$$
$$\pi_i(q_i, Q^*_{-i}) = \pi_i(q^*_i, Q^*_{-i}) \Rightarrow \pi_i(q_i, Q_{-i}) < \pi_i(q^*_i, Q_{-i}) \tag{7}$$

It is clear that the conditions $(6, 7)$ implies that all ESS is a Nash equilibrium strategy but all NE strategy is not necessarily ESS..
This definition, while considering random, pairwise contests between individuals drawn from a large population, is obtained by assuming that both the predominant player (e.g. strategy) and the mutant confront the same population profile. In a finite population game, as already noticed by Riley (1978), a strategy which satisfies the previous conditions $(6-7)$, may not be protected against invasion by a mutant strategy. In order to analyze an oligopoly situation, we need, as showed by Alós-Ferrer and Ania (2005), a definition of an ESS for a finite population of players which "play the field", that is all compete with each other simultaneously. Schaffer (1988) arrives to the same conclusion and proposes a concept of *finite population ESS* which is defined as follows[1]

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

$$\pi_i(q^*_i, q^m_j, Q^*_{-i-j}) \geq \pi_j(q^m_j, Q^*_{-j}), \forall i \neq j, \forall q^m_j \neq q^*_j \tag{8}$$

where $Q^*_{-i-j} = \left\{ q^*_1, \ldots, q^*_{i-1}, q^*_{i+1}, \ldots, q^*_{j-1}, q^*_{j+1}, \ldots, q^*_n \right\}$.

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

2.3 ESS and Nash equilibrium

As shown by Schaffer (1988), a finite population ESS is not generally a Nash equilibrium strategy. To see this, one should notice that the condition (8) does not exclude the possibility that

$$\exists j, \pi_j(q^m_j, Q^*_{-j}) \geq \pi_j(q^*_j, Q^*_{-j}), \forall q^m_j \tag{9}$$

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

A direct consequence of this result implies the paradox we have underlined in the preceding paragraph:

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

The next section will analyze the mechanisms which are behind this result.

3 Evolutionary stability of equilibria and the spite effect

The mechanism which pushes the population towards the Walrasian equilibrium is well known in the evolutionary literature:

[1]Sloth and Witta Jacobson (2006) define in a similar way the concept of *Darwinian equilibrium.*

"[...] a firm maximising its own profit may help its non-maximising competitors to do even better. Put another way, a firm which does not maximise its profit may still earn profits which are larger than those of its profit-maximising competitors, if the costs to itself of its deviation from maximisation are smaller than the costs it imposes on the maximising competitors. [...]This result is a consequence of the Darwinian definition of economic natural selection, whereby it is the fittest firms which survive. The above result is essentially an application to economics of Hamilton's theory of spite in evolutionary biology [Hamilton (1970.1971)]. An act by an animal is spiteful if the animal harms both itself and another. Hamilton demonstrated that such a trait could be selected for if the population was not very large. The condition for the selection of a spiteful trait is that the decrease in an animal's own Darwinian fitness is smaller than the decrease in the fitness of the average member of the rest of the population; since the holder of the spiteful trait thus has a higher fitness than that of his intraspecies competitors, the trait will be selected for. [...]This result is directly analogous to the spitefulness of the evolutionary biology ESS in finite populations [...]"
Mark E. Schaffer (1989)

Strictly speaking, the *spite effect* is the opposite of altruism. We should used the terms of spite effect only when the mutant firm j suffers from deviation ($\Delta\pi_j < 0$) but it 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 strategy, ESS and spite effect is straightforward.

3.1 Nash equilibrium and ESS in a finite population oligopoly game

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

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

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

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

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

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

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

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 Nash equilibrium strategy is an ESS. □

3.1.1 One single mutation case

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

$$dq_i\left(p'dq_i + p - C'\right) < 0, \forall i \tag{12}$$

Proof.
At $\hat{q}$ the profits of the firms are

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

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

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

$$\pi_j(q^m) = p(q^m)q_j - C(q_j) \tag{15}$$

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

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 $\pi_i(\hat{q}) = \pi_j(\hat{q}_i)$.
The net gain of firm j, $\Delta\pi_j$ corresponds to (15) − (13):

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

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

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

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

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

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

$$\Delta\pi_i = \pi_i(\hat{q}_j + dq_j, \hat{Q}_{-j}) - \pi_i(\hat{q})$$

$$= \underbrace{\left(p(\hat{q}_j + dq_j, \hat{Q}_{-j}) - p(\hat{q})\right)\hat{q}_i}_{\hat{q}_i dp} \tag{20}$$

And consequently:

$$\Delta\pi_j - \Delta\pi_i < 0 \tag{21}$$
$$\Longrightarrow \hat{q}_j dp + dpdq_j + \hat{p}dq_j - dC_j - \hat{q}_i dp < 0$$
$$\Longrightarrow dp(\hat{q}_j - \hat{q}_i) + dpdq_j + \hat{p}dq_j - dC_j < 0 \tag{22}$$
$$\Longrightarrow dpdq_j + \hat{p}dq_j - dC_j < 0 \text{ since } \hat{q}_j = \hat{q}_i \tag{23}$$

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

$$p'dq_j dq_j + \hat{p}dq_j - C'dq_j < 0 \tag{24}$$
$$dq_j \left(p'dq_j + \hat{p} - C'\right) < 0 \tag{25}$$

□

Before discussing more in detail in the next section the implications of learning with genetic algorithms, 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 can 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 $p'(dq_j)^2 < 0$.

- $dq_j < 0$: Since $p' < 0$, the condition (25) holds. A firm that deviates alone by decreasing its quantity will benefit, via a price increase, from a revenue increase lower that the one it will offer to the firms who stay committed. So, the mutant strategy will not be replicated by imitation and the Walrasian strategy will persist.
- $dq_j > 0$: The firm who increases its quantity will decrease the market price to a level such that the profit of every firm will decrease. But the loss of the mutant firm will be higher than the loss of the firms continuing to stick to the Walrasian strategy. Thus, this mutant strategy will again 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 last condition 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, the condition 12 does not hold and a spite effect is not observed. 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 move away from the CE by decreasing again the quantity. 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.

□

3.1.2 Generalization to m identical mutants

Identical mutants

Assume now 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* $(q_1^*, q_2^*, .., q_n^*)$, *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 plays $q_j \neq q^*$, the profits of these m firms should be lower than the profits of the $n - m$ firms that stayed committed to q^*. From this remark, Schaffer (1988) defines the concept of $m - stable$ (finite population) ESS as follows.

Definition 8. *An equilibrium strategy* $(q_1^*, q_2^*, .., q_n^*)$ *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. But if the number of mutants increases above m, this is no more the case.

Proposition 5. *In a finite population oligopoly game, if a strategy profile* $q^* = (q_1^*, q_2^*, ..., q_n^*)$, *does not yield a spite effect for at least one firm, it is an ESS equilibrium whatever is the number of mutants.*

Proof.
With the m simultaneous mutants belonging to I_m, the set of mutant firms, the profits are, with $Q_m = (m-1)\hat{q} + (n-m)q^*$, and with $q_j = \hat{q}, \forall j \in I_m$ and $q_i = q_i^* = q^*, \forall i \notin I_m$

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

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

It is clear that the $\Delta\pi_j$ are equivalent between all mutant firms and $\Delta\pi_i$, between all non mutant firms. If the spite effect holds for at least one non mutant firm then it holds for all non mutant firms. As a consequence, the initial strategy profile is an ESS. □

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 a n firms and with m mutants is equivalent to a 2–players game with one single mutant. And 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. □

Non-invadable and invading strategies

As noticed by Alós-Ferrer and Ania (2006), one can reinterpret the ESS (i.e. 1–stable ESS) and the $(n-1)$–stable ESS as, respectively, an non-invadable strategy and invading strategy.
First, if a strategy q^* is an ESS, no matter what the mutation strategy is, the mutant will always find itself earning lower payoffs than the rest of the population. As a consequence, this mutation should disappear. So, the (1–stable) ESS is a non-invadable strategy.
Second, if a strategy is $(n-1)$–stable ESS, then, whenever the whole population but one agent mutate to another strategy, the non-mutant player will earn larger payoffs than the mutants. It is possible to reinterpret the previous result by saying that the non-mutant player was the mutant one. Since its strategy is the best one, it will invade the rest of the population: the $(n-1)$–stable ESS becomes an invading strategy.

Heterogeneous mutants with heterogeneous deviations

Let assume now that there exists m heterogeneous mutants. In our oligopoly game, this means that some of the firms will increase their quantities while some others will decrease it. In such a framework, the firms that do not move will have an evolutionary stable strategy if, at least, they can get higher profits than the best of the mutants. So, the following definition holds.

Definition 9. *An equilibrium strategy* $(q_1^*, q_2^*, .., q_n^*)$ *is an* ESS with m heterogeneous mutants *belonging to* i_m *if, for* $\forall q_j \neq q^*, j \in i_m, \forall q_i = q^*$*,and* $\hat{q}$ *a strategy profile composed of* m *mutant strategies* q_j *and* $n-m$ *equilibrium strategies* q^*

$$\pi_i(\hat{q}) > \max_{j \in i_m} \pi_j(\hat{q})$$

This definition means that whatever are the quantities that the mutants will choose, and whatever is the new total quantity, the non mutant firms must obtain a higher profit than any mutant firms. 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 homogeneous mutants (e.g. mutations in the same direction) but unstable with heterogeneous mutations.

Learning process of the firms in Cournot oligopoly will generally imply simultaneous heterogenous mutations. We now consider the consequences of these mutations in the context of the debate on individual *v.s.* social learning with genetic algorithms.

3.2 Spite effect and importance of 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 its relative payoff by individual mutation, which means that it is a Nash equilibrium of the "beat the average" game reformulation of the initial oligopoly game.

Let define the function of a firm i by the difference of the firm's profit and the population 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 configuration[2] ($\pi_j = \pi, \forall j$), we must have $\pi_i^R(q_i, Q_{-i}) = \frac{(n-1)}{n}\pi_i(q_i, Q_{-i}) - \overline{\pi}_{-i}$, with $\overline{\pi}_{-i} = \frac{(n-1)}{n}\pi$, with $\pi_j(q_j, Q_{-j}) = \pi \ \forall j \neq i$.

Proposition 6. *If* q^* *is not ESS, than there exists at least one mutant strategy* $\hat{q}_j$ *such that the relative profits of the mutant firms is increasing.*

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 have from (8)

$$\frac{n-1}{n}\left(\pi_j(\hat{q}_j, Q_{-j}^*) - \pi_i(\hat{q}_j, Q_{-j}^*)\right) > 0$$
$$\pi_j(\hat{q}_j, Q_{-j})\frac{(n-1)}{n} - \frac{(n-1)}{n}\pi > 0$$
$$\pi_j^R(\hat{q}_j, Q_{-j}) > 0$$

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

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

Another implication of this proposition can be formulated using the *fitness* concept that we will use later in this article:

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

where $\Pi_{-i} = \sum_{k \neq i} \pi_j(q_i, Q_{-i})$. The fitness measure is used in most biology evolution game as the base of the selection process in a given population.

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

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

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

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. □

3.2.1 ESS, variation of profits and spite effect

The ESS is hence related to the variation of the relative profits, but it can also formulated in terms of variation of profits. Let $\Delta\pi_j = \pi_j(\hat{q}_j, Q_{-j}^*) - \pi_j(q_j^*, Q_{-j}^*)$ be the net evolution of the profit of the mutant firm j who decides to produce a quantity $\hat{q}_j$ rather than an equilibrium quantity q_j^*. For the each of the non-mutant firms ($i \neq j$), the evolution of profit is given by $\Delta\pi_i = \pi_i(\hat{q}_j, Q_{-j}^*) - \pi_i(q_j^*, Q_{-j}^*)$. Since, $\pi_j(q_j^*, Q_{-j}^*) = \pi_i(q_j^*, Q_{-j}^*)$ in a symmetrical equilibrium, the following definition is directly related to definition 4.

Definition 10. *In a finite population game, a strategy $q^* = \{q_1^*, q_2^*, ...q_j^*, ...q_n^*\}$ is said to be evolutionary stable if, for any strategy $\hat{q}_j \neq q_j^*$,*

$$\Delta\pi_i - \Delta\pi_j \geq 0$$

This means that the net gain (or loss if $\Delta\pi_j < 0$) of the mutant firm must be lower (greater) than the net gain (or loss if $\Delta\pi_i < 0$) of the non mutant firms. The situation where the variation of profits is negative is directly related to the spite effect.

```
procedure evolution program
begin
    t ← 0
    (1) initialize P (t)
    (2) evaluate P (t)
    while (not termination–condition) do
    begin
        t ← t + 1
        (3) select P (t)from P (t − 1)
        (4) alter P (t)
        (5) evaluate P (t)
    end
end
```

Figure 1: The structure of an evolutionary program (Michalewicz, 1996)

4 Learning, selection and convergence to equilibria

If we allow firms to adapt their production levels as a consequence of learning, this adaptation will imply a specific selection mechanism in the evolution of their strategies. The recent debate on the convergence to CE with learning firms (*see* Vriend (2000), Arifovic and Maaschek (forthcoming)) has focused on learning with genetic algorithms (GA).

A GA is based on mechanisms inspired by biological evolution: selection, crossover and mutation (*see Figure* 1). The *canonical genetic algorithm* makes evolve a population of chromosomes. The size of the population m is given. It is the source of one of the strengths of the 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 total number of generations. Notice that T is the other source of the strengths of the GA. The algorithm (randomly) generates an initial population $P(0)$ of candidate chromosomes which are evaluated at each period using the fitness (value) function. They are used for composing a new population at the next period $P(t + 1)$. Each chromosome has a probability of being **selected** that is increasing in its fitness. The members included in the new population are recombined using a **crossover** mechanism (see Figure **??**). The crossover operation introduces controlled innovations in the population since it combines the candidates already selected in order to invent new candidates with a potentially better fitness. Moreover, the **mutation** operator randomly modifies the candidates and introduces some random experimenting in order to more extensively explore the state space and escape local optima. Typically, the probability of mutation is rather low in comparison with the probability of crossover because otherwise the disruption introduced by excessive mutations can destruct the hill-climbing capacity of the population. Finally, an **elitism** operator can be used which ensures that the best individual of a population will be carried to the next generation.

The GA is used to represent the learning capacity of the firms, since it can make evolve a population of production levels on the base of a fitness based on the profits resulting from these quantities. Vriend (2000) confronts two different implementations of this approach and claims that this confrontation 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 by firm and the operations of the GA correspond to the imitation between firms (crossover) and to random experimenting by some firms (mutations). 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) and random experimenting in the strategy population of the firm (mutation).

By confronting these two different approaches to learning by GA, Vriend (2000) obtains a convergence to WE under social learning and to CE, under individual learning. Arifovic and Maaschek (forthcoming) question these results. 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 that lay behind these different results in the literature. Our analysis will be focused on the main operator in the evolution of the population: the fitness based selection. This focalization will also allow us to connect the results on learning with the general results we have obtained in the preceding sections.

In any population of strategies, the convergence to WE will occur in general if the selection mechanism favours 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 the relative fitness of the strategies (whatever is the particular definition of this fitness), this condition is equivalent to a necessarily increasing relationship between the relative fitness of the strategies and quantities below the WE and a decreasing relationship above the WE.

Definition 11. *If we note by F_j the fitness of a strategy in the population of strategies, its* relative fitness *will be given by*

$$f_j = \frac{F_j}{\sum_j F_j} \tag{28}$$

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: $f_j = f\left(\underset{+}{\pi_j}\right)$. In the simplest case, and the articles we cite use this formulation, the profit will directly correspond to the fitness of the strategy. As a consequence, 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

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

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 related to the profits of the firms and the way these profits are compared is an essential difference between social and individual learning.

4.1 Social learning

In the case of the 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 that uses this strategy with the profits of the other firms

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

when firm i is the user of the strategy j. If we note $\Pi_{-i} \equiv \sum_{k \neq i} \pi_k(q_i, Q_{-i})$, we can note this relative fitness as

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

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

$$\frac{\partial f_i}{\partial q_i} = \frac{\frac{\partial \pi_i}{\partial q_i}(\Pi_{-i} + \pi_i) - \pi_i(\frac{\partial \Pi_{-i}}{\partial q_i} + \frac{\partial \pi_i}{\partial q_i})}{(\Pi_{-i} + \pi_i(q_i, q_{-i}))^2}$$

$$\frac{\partial f_i}{\partial q_i} = \frac{\frac{\partial \pi_i}{\partial q_i}\Pi_{-i} - \pi_i \frac{\partial \Pi_{-i}}{\partial q_i}}{(\Pi_{-i} + \pi_i)^2} \tag{29}$$

As a consequence, we observe

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

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

$$\begin{aligned} \frac{\partial \pi_i}{\partial q_i}\Pi_{-i} &> \pi_i \frac{\partial \Pi_{-i}}{\partial q_i} \\ \Leftrightarrow \frac{q_i}{\pi_i}\frac{\partial \pi_i}{\partial q_i} &> \frac{q_i}{\Pi_{-i}}\frac{\partial \Pi_{-i}}{\partial q_i} \\ \Leftrightarrow \varepsilon_{\pi_i, q_i} &> \varepsilon_{\Pi_{-i}, q_i} \end{aligned} \tag{31}$$

where $\varepsilon_{y,x}$ represents the elasticity of y with x. As a consequence, the tendency towards the WE equilibrium will depend in this case on the comparison of the impact of the quantity increase of the firm on its profit with the same impact on the total profits of its competitors. If the these impacts are positive, the increase of the quantity of the firm i will increase its probability to survive, if and only if, the impact on its own profit is greater than the impact of the other firms total profit. If these impacts are negative on the contrary, this condition is of course 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 even more decrease the competitors profits ($0 > \varepsilon_{(\Pi_i, q_i)} > \varepsilon_{(\overline{\Pi}_{-i}, q_i)}$) and it will increase its relative fitness. We must distinguish two cases in the evaluation of these impacts.

4.1.1 Infinite population of firms

When the population of the firms is large, we have, by the atomistic property, that the variation of the quantity of one particular firm i does not significantly modify neither the market price nor 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 (31):

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

and

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

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

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

since q^W is the Walrasian strategy and it verifies, by definition, $p - C' = 0$. We hence 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 of this equilibrium. This is the main mechanism that pushes the industry towards the WE instead of the CE.

4.1.2 Finite population of firms

In a finite population, the atomistic property is not assured. As a consequence, we can have $\partial p(q_i, Q_{-i})/\partial q_i < 0$ and $\partial \Pi_{-i}/\partial q_i \neq 0$. If we consider the last derivative

$$\begin{aligned}
\frac{\partial \Pi_{-i}}{\partial q_i} &= \frac{\partial (\sum_{j \neq i}(pq_j - C(q_j))}{\partial q_i} \\
&= \sum_{j \neq i}(\frac{\partial p}{\partial q_i} q_j) = \frac{\partial p}{\partial q_i} \sum_{j \neq i} q_j = \frac{\partial p}{\partial q_i} Q_{-i} \\
&= p^{'} Q_{-i}
\end{aligned} \tag{32}$$

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 compute

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

As a consequence, the equation (30) becomes

$$sgn\left(\frac{\partial f_i}{\partial q_i}\right) = sgn\left((p^{'} q_i + p - C')\Pi_{-i} - p^{'} Q_{-i}\pi_i\right) \tag{33}$$

$$= sgn\left((p - C')\Pi_{-i} + p^{'}(q_i \Pi_{-i} - Q_{-i}\pi_i)\right) \tag{34}$$

The first term shows that the impact of an increase of its quantity depends on the position of the firm in terms of marginal cost and price (in a similar way to the infinite population case). But it also depends, in a more complicated way, on the relative quantities and profits of the firms, in a way not disconnected from the spite effect. We can note that the last effect in this equation can be reformulated:

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

since $p' < 0$. Hence, when $\Pi > 0$, this effect will favour increasing quantities if the market share of the firm is lower than its profit share. As a consequence, if we have $\Pi_{-i} > 0$ and $\Pi > 0$,

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

since $p^{'} q_i + p - C' < 0$ in equation (33) if $q_i^C < q_i$, and $p = C'$, $q_i/Q = \pi_i/\Pi$ if $q_i = q^W, \forall i$.

As a consequence, the strategies will advance towards the WE if the firms can obtain positive profits at the neighborhood of WE.

4.2 Individual learning

When the learning of the firms is individual, the firms can not count on their interactions on the market for directly learning the strategies from the competitors (imitation). Each firm must extract information on the relevance of its actual strategy from the global market variables that result from the interactions of the firms (mainly, the market price) that it can observe, and from the profit it obtains with its strategy under these market conditions.

As a consequence, when the learning dynamics are not based on imitation, the spite effect cannot directly play. The main question again is the consequences of these dynamics in terms of convergence to equilibrium. Since convergence to 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 on which the firm can count during its learning process.

As Alós-Ferrer (2005) summarizes it: "The intuition is that if firms remember past profits, destabilizing Cournot will not be such an easy task. After a single mutation [...], the mutant may earn more than the nonmutants, but the largest profit remembered will still be that of the Cournot equilibrium, and hence, the mutant will 'correct the mistake', even in absence of any strategic considerations."

Riechmann(2006), on the other hand, suggests 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".

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

Models of individual learning necessarily introduce dynamics for the evolution of the individual strategies. In the discussion above we have deliberately focused on a population based approach, in line with the recent literature on this debate. Effectively, Vriend (2000), Arifovic and Maaschek (forthcoming), and other articles, discuss the convergence problem with a representation of strategy evolution based on different formulations with genetic algorithms. 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, and learning without expectations, evaluating one strategy in each period. Other intermediate cases have also been considered, like in Vriend (2000) where the GA does not intervene in each period. When this is relevant, we will also point to more traditional learning schemes based on best replies.

4.2.1 Individual learning with hypothetical profits

In this case, 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 pseudocode if the oligopoly game has a duration of T periods (see Figure 2).

Notice that each strategy is evaluated, in each period t, using a same given price p^{t-1}. Thus, with the roulette wheel selection, the fitness of a given strategy j of the player i would result from:

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

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

- 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:
  $$\pi_{ij}(q_{ij}, p_1) = p^{t-1}q_{ij} - C(q_{ij}), \text{ for all } i.$$
  - Using GA procedures (selection, crossover and mutation) in each strategy population, a new population of strategies q_i^t is defined for each firm on the base of these profits.
  - Again the hypothetical profits are calculated $\{\pi_{i1}(q_{i1}, p^{t-1}), ..., \pi_{ik}(q_{ik}, p^{t-1})\}$.
  - 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 2: Pseudocode of the learning model with hypothetical profits

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

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

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

As a consequence, the best fitness will be attained under the same conditions as Walrasian pricing rule $\left(C^{'} = p\right)$ since:

- During the roulette wheel procedure of the GA, given a price p:
  - higher quantities will spread in the population, as long as $p > C^{'}$;
  - lower quantities will diffuse if $p < C^{'}$.
- 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 $j \in \max_{q_{il}} f_{il}, \forall l$.

As we already shown above there is a strong relationship between evolution of the fitness and ESS.

Proposition 7. *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 never move, $Q^C_{-i} = \{q^C_1, ..q^C_{i-1}, q^C_{i+1}, ..q^C_n\}$. In this case, the learning of player i is identical to the learning of its monopolistic optimal quantity, given that it doesn't know the exact price relationship $p(q_i, Q^C_{-i})$. 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 straightforward to see that the fitness evolution, as a maximization of the hypothetical profit, will lead to the Walrasian condition: $p_{t-1} = C^{'}$. As a consequence, the CE equilibrium is not ESS, while the WE is. □

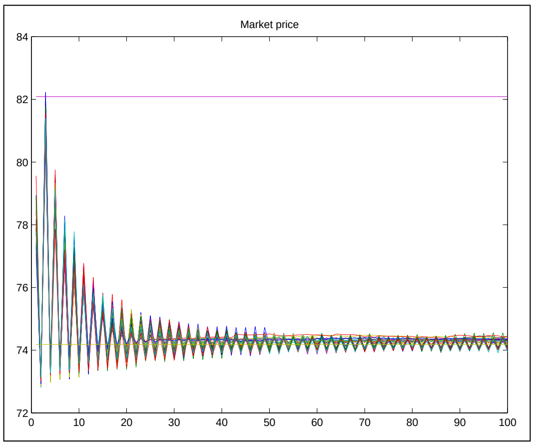


Figure 3: Convergence to WE price with hypothetical learning

We use computational experiments to illustrate our results in this section and the following one. The details of the simulation protocol are given in Appendix A. The Figure 3 clearly shows the market price quickly converges to WE price in all simulations of a standard oligopoly with hypothetical profits.

- 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 ($Period \pmod{\tau} = 0$): Using GA procedures, as roulette wheel selection, a new set of strategies, q_i^{ϕ}, is defined for each firm.
  - Each player plays a strategy $q_{ij} \in q_i^{\phi}$. The corresponding price market, p_t, is calculated. The realized profit is calculated ($\pi_{i,t}(Q_t, p_t)$).

Figure 4: Pseudocode of learning model without expectations

4.2.2 Individual learning without expectations

In this approach, the firm must play strategies in order to evaluate their relative fitness. Each population of strategies is played without modification during τ periods (an *epoch* corresponding to the $GArate$ in Vriend (2000)). A special case would be an epoch of k periods where each strategy is played once, following a sequential order, but one can consider more sophisticated configurations. During an epoch, strategies from given populations meet following the strategy-selection scheme used (random, roulette wheel, play the best, etc.). Profits of strategies are computed as a consequence of this matching process and, at the end of the epoch, the GA modifies populations on the base of these profits. Figure 4 gives the pseudocode of this configuration.

This learning algorithm differs from the hypothetical profit case by one important aspect. Each strategy is evaluated based on the base of the real observed profit that it yields by repeatedly meeting the strategies of other firms. This observed profit is necessarily based on the effective price corresponding to these strategies and it can now include information on the demand curve if the same population of quantities can meet different, but given, quantities of the competitors. This obviously will possible if the $GArate$ is long enough in comparison with the strategy population size.

Given a $GArate$, each individual strategy of the firm q_{ij} will play $GArate$ times against the strategy populations of other firms $\{q_{lj}\}_{l=1\ldots n, l\neq g}$. On average, the profit of this strategy will be given by

$$E\left[\pi_i\left(q_{ij}, Q_{-i}\right)\right] = E\left[p\left(q_{ij}, Q_{-i}\right)\right] q_{ij} - C\left(q_{ij}\right) = \overline{p} q_{ij} - C\left(q_{ij}\right)$$

with

$$\overline{p} = \sum_{l\neq i} p\left(q_{ij}, \sum_{\underline{\ } \ l} q_{lj}\right) f\left(q_{lj}|GArate\right)$$

where $f\left(q_{lj}|GArate\right)$ is the frequency of each individual strategy of other firms, conditioned by the strategy selection process (the selection of the strategies to be played on the market) and $GArate$. We can observe that

$$GArate \to \infty \Rightarrow f\left(q_{lj}|GArate\right) \to f\left(q_{lj}\right)$$

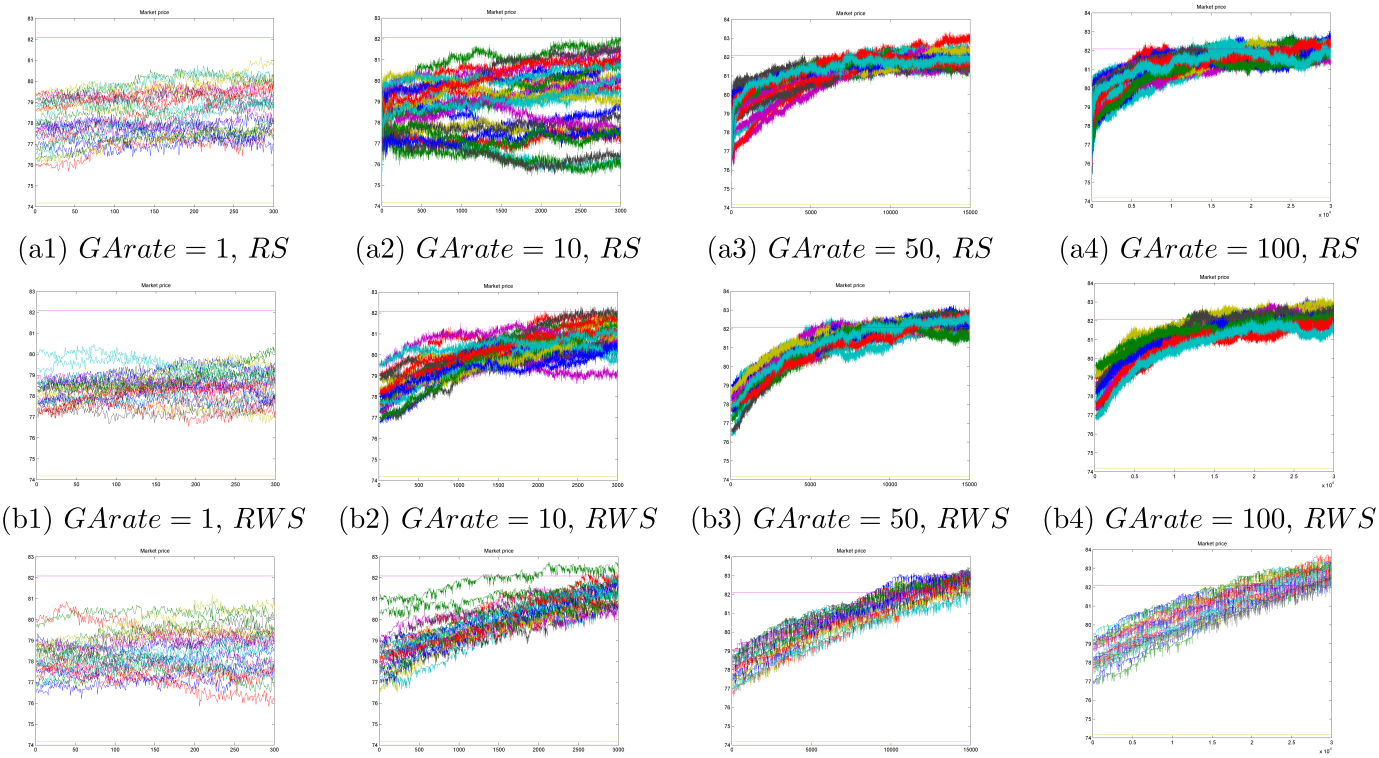
which is the real empirical frequency of each strategy in the strategy population of the firm l. Moreover, given the populations of other firms, the average profit of q_{ij} that results from the matching process now implies $\partial\overline{p}/\partial q_{ij} < 0$ and the selection process will be able to use this information on the demand curve. As a consequence, the strategies that correspond to $C' = p$ will be dominated by strategies that are closer to the condition $C' = \overline{p} + q_{ij}\partial\overline{p}/\partial q_{ij} < \overline{p}$. This mechanism will push the strategies towards CE instead of WE.

Again we use computational experiments to illustrate these results. We gives results on convergence conditioned by two main dimensions that arose in the preceding discussion: the $GArate$ and the mechanism that is used to select in the population the strategy to be played on the market (*see* Appendix A). The results of our simulations (a batch of 25 simulations with each configuration) are given in Figure 5. In each simulation we run exactly 300 generations for the GA.

Graphics $(a1 - c1)$ and $(a2 - c2)$ show that if the strategies do not meet enough frequently, the information on demand is quite difficult to extract and firms have some difficulty to converge to CE price, whatever the selection mechanism used for choosing the played quantities. This effect is stronger for the random selection (RS) process since the matching of the strategies is not oriented at all in this case. We observe that other mechanisms already fair quite well for $GArate = 10$.

Graphics $(a3 - c3)$ and $(a4 - c4)$ show that matching mechanisms integrating some intrinsic randomness (in complement of the randomness implied by the GA – mutations) give rise to convergence to CE price when $GArate$ is higher. The PTB mechanism that systematically plays for each firm its supposedly best strategy of the period has some difficulty to converge and 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. We also observe in graphics $(a3 - b3)$ and $(a4 - b4)$ that random selection secures an earlier convergence to CE by more systematically extracting information on demand, in accordance with the general result we have established above.

These results show that one does not need very specific Cournot oligopoly game and GA setup in order to show the convergence to CE in the case of individual learning. A simple quadratic cost oligopoly game and an elementary, real value based, GA with only selection and low mutation is sufficient to observe this convergence.



(a1) $GArate = 1$, RS (a2) $GArate = 10$, RS (a3) $GArate = 50$, RS (a4) $GArate = 100$, RS

(b1) $GArate = 1$, RWS (b2) $GArate = 10$, RWS (b3) $GArate = 50$, RWS (b4) $GArate = 100$, RWS

(c1) $GArate = 1$, PTB (c2) $GArate = 10$, PTB (c3) $GArate = 50$, PTB (c4) $GArate = 100$, PTB

RS : Random selection; RWS : Roulette wheel selection; PTB : Play the best

Figure 5: Convergence to the Cournot equilibrium price: Individual learning without expectations

5 Conclusion

The main objective of this article is to gain some general understanding of the convergence to equilibria in an oligopoly game with learning dynamics. We have established the properties, in terms of evolutionary stability, of two types of potential equilibria in this game: Cournot equilibrium (CE) and Walrasian equilibrium (WE). The first part of the particle 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 is possess such a social dimension (like under imitation) and when dynamics are based on selection, hence on relative performance of firms. Under 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 reply in order to assure the evolutionary stability of the CE. The WE becomes the only ESS equilibrium in this case. As a consequence, we study the possibility of multiple mutations, since this is more in accordance with the *playing the field* aspect of the Cournot game (each firm meets in each period the totality of 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 hence focalizes a special setup for learning dynamics based on the properties of Genetic Algorithms (GA) because this setup has yielded some puzzling results and controversies about the convergence to CE. We first study the general conditions under which the selection in GA can promote convergence to CE. We show that when the GA represents social learning, the convergence can only occur to WE. In the case of individual learning, the convergence to CE is only possible if the interactions of the firms allow them to discover the decreasing relationship 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 in the forthcoming paper, in collaboration with Maaschek). Some memory of quantities associated to observed current profits is necessary in order to secure this convergence. We show, in a computational setup close to Vriend (2000), but under quite general cost and GA conditions, that the convergence to CE arises when firms' strategies have enough opportunity to meet each other before the intervention of the GA. Our results confirm and clarify the results obtained in Vriend (2000).

References

Alkemade F., J.A. La Poutré, and H.M. Amman, (2005), On Social Learning and Robust Evolutionary Algorithm Design in Economic Games, *Proceedings of the 2005 IEEE Congress on Evolutionary Computation (CEC 2005)*, IEEE Press, pages 2445 - 2452.

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. and M.Maschek, forthcoming, Revisiting Individual Evolutionary Learning in the Cobweb Model - An Illustration of the Virtual Spite Effect, *Computational Economics*.

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 behaviour 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.

Maynard-Smith J., (1982), *Evolution and the theory of games*, MIT Press, Cambridge MA.

Riechmann, Th., (2006) Mixed motives in a Cournot game. *Economics Bulletin*, Vol 4(29), pp. 1-8.

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*, 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.

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*, 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.

Appendix

A Numerical experiments

A.1 Oligopoly model of the experiments

Let define a simple general oligopoly game with

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

where $Q_{-i} = \sum_{j \neq i} q_j$.

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

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

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

$$\text{CE: } q_i^C = 200/23 = 8.695\,7, \ p^C = 1888/23 = 82.087$$
$$\text{WE: } q_i^W = 100/11 = 9.090\,9, \ p^W = 816/11 = 74.1818182$$

B GA and other specifications

In all simulations we have used the same genetic algorithm (GA): a real value based GA with 20 chromosomes. We have kept the number of generations constant over all simulations in order to let other parameters determine the learning capacity of the firms. In order to keep a GA structure as close as possible to our theoretical results, we used a GA only based on selection (roulette wheel selection) and mutations (with a uniform rate of one mutation in each run of the GA, for each firm).

The chromosomes of first generation for each firm are randomly generated in the interval $\left[(1-\gamma)\, q^C, (1+\gamma)\, q^W\right]$, 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

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

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

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

We use $\sigma = 5\%$ for the LWE case and $\sigma = 1\%$ in the learning with hypothetical profits case (**LHP**).
We have used the following scaled fitness functions for our experiments:

$$\text{Hypothetical profits: } F_{ij} = exp\left(100 \ max\left\{0, \frac{\pi_{ij}}{\pi_i^*}\right\}\right)$$
$$\text{Learning without expectations: } F_{ij} = exp\left(\frac{\pi_{ij}}{\pi_i^*}\right)$$

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

In the learning without expectations case, three strategy selection mechanisms have been 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 of periods (*GArate* in Vriend (2000). We test the following cases for the *GArate*: $\{1, 20, 50, 100\}$. The last case corresponds the assumption used by Vriend.

In the LWHP case, only the best strategy is played in accordance with Arifovic and Maaschek (forthcoming).

We run 25 simulations with each configuration and the graphics give the averages over firms for these 25 simulations, as in Vriend (2000) and Arifovic and Maaschek (forthcoming).

Cahiers du GREThA
Working papers of GREThA

GREThA UMR CNRS 5113

Université Montesquieu Bordeaux IV
Avenue Léon Duguit
33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75
Fax : +33 (0)5.56.84.86.47

www.gretha.fr

Cahiers du GREThA (derniers numéros)

2007-01 : GONDARD-DELCROIX Claire, *Entre faiblesse d'opportunités et persistance de la pauvreté : la pluriactivité en milieu rural malgache*

2007-02 : NICET-CHENAF Dalila, ROUGIER Eric, *Attractivité comparée des territoires marocains et tunisiens au regard des IDE*

2007-03 : FRIGANT Vincent, *Vers une régionalisation de la politique industrielle : l'exemple de l'industrie aérospatiale en Aquitaine*

2007-04 : MEUNIE André, POUYANNE Guillaume, *Existe-t-il une courbe environnementale de kuznets urbaine ? Emissions polluantes dues aux déplacements dans 37 villes*

2007-05 : TALBOT Damien, *EADS, une transition inachevée. Une lecture par les catégories de la proximité*

2007-06 : ALAYA Marouane, NICET-CHENA Dalila, ROUGIER Eric, *Politique d'attractivité des IDE et dynamique de croissance et de convergence dans les Pays du Sud Est de la Méditerranée*

2007-07 : VALLÉE Thomas, YILDIZOĞLU Murat, *Convergence in Finite Cournot Oligopoly with Social and Individual Learning*

2007-08 : CLEMENT Matthieu, *La relation entre les transferts privés et le revenu des ménages au regard des hypothèses d'altruisme, d'échange et de partage des risques*

2007-09 : BONIN Hubert, *French banks in Hong Kong (1860s-1950s): Challengers to British banks?*

La coordination scientifique des Cahiers du GREThA est assurée par Sylvie FERRARI et Vincent FRIGANT. La mise en page est assurée par Dominique REBOLLO.