Loss of Commitment? An Evolutionary Analysis of Bagwell's Example[1]

Jörg Oechssler[2]
Department of Economics
Humboldt University, Berlin

Karl H. Schlag[3]
Economic Theory III
University of Bonn

July, 1997

[1]We wish to thank Wolfgang Leininger for starting this project by asking us about the relevance of noise for evolutionary predictions.

[2]Institut für Wirtschaftstheorie, Spandauer Str. 1, 10178 Berlin, Germany, FAX: +49–30–2093-5619, E–mail: oechsler@wiwi.hu-berlin.de

[3]Adenauerallee 24-26, 53113 Bonn, Germany, E-mail: schlag@econ3.uni-bonn.de

Abstract

In a recent paper Bagwell (1995) pointed out that only the Cournot outcome, but not the Stackelberg outcome, can be supported by a *pure* Nash equilibrium when actions of the Stackelberg leader are observed with the slightest error. The Stackelberg outcome, however, remains close to the outcome of a *mixed* equilibrium.

We compare the predictions in various classes of evolutionary and learning processes in this game. Only the continuous best response dynamic uniquely selects the Stackelberg outcome under noise. All other dynamics analyzed allow for the Cournot equilibrium to be selected. In typical cases Cournot is the unique long run outcome even for vanishing noise in the signal.

JEL– classification numbers: C72, C73.

Key words: imperfectly observable commitment, evolution, imitation, learning, equilibrium selection.

1 Introduction

In a recent paper Bagwell (1995) pointed out that a "first mover advantage" in games depends crucially on the fact that the action taken by the first mover is perfectly observable. In fact, he showed that if the action is observed with the slightest bit of error, no commitment is achieved.[1] Bagwell used the example of Stackelberg competition in which the leader can either choose the quantity (L) of a Stackelberg leader or the Cournot quantity (C). He shows that if the quantity choice is observed with some error (i.e. if there is a small probability that the follower observes C when the leader, in fact, chose L), then the only equilibrium in *pure* strategies is the Cournot equilibrium.[2]

As noted by Bagwell (1995) there are – additionally to the Cournot equilibrium – two mixed equilibria, one of which is "close" to the Stackelberg outcome in the sense that it converges to the Stackelberg outcome as the noise vanishes.[3] By using a modification of Harsanyi and Selten's (1988) equilibrium selection theory Van Damme and Hurkens (1994) argue that this "noisy Stackelberg equilibrium" should be selected. However, the Cournot equilibrium is a strict equilibrium and therefore has many desirable properties.

Given the controversy over which equilibrium should be selected the purpose of this paper is to compare the predictions made by three classes of evolutionary dynamics for this game. First we consider a general class of smooth continuous time dynamics that include payoff monotone and payoff positive but not best response dynamics. Through the introduction of noise Cournot equilibrium becomes asymptotically stable. On first sight this might not be surprising as the Cournot equilibrium is the unique strict equilibrium in the game with noise. However, the payoff difference to the second best strategy vanishes as the noise goes to zero. Nevertheless, the underlying basin of attraction stays large when noise becomes small. Whether or not the

[1]Adolph (1996) shows that commitment is restored if additionally to the noise in signal transition players make mistakes in the *execution* of their strategies.

[2]This results has been generalized in several directions, see Van Damme and Hurkens (1994) and Güth, Kirchsteiger and Ritzberger (1995).

[3]For generalizations of this result to n player games see Güth et al. (1995).

Stackelberg equilibrium has similar properties depends on the specifications of the dynamic, under the standard replicator dynamic it does not.

Next we consider general finite population adjustment dynamics. Here, Cournot is selected if the population is large enough. For small populations the method is inconclusive in general. However, for a more limited class of dynamics based on imitation, Cournot is selected regardless of the population size.

Finally, we consider the continuous best response dynamic. Here, with or without noise, Stackelberg is the unique long run outcome.

Thus, we find that the Cournot equilibrium can no longer be ignored as a prediction under noise, often it is even the unique prediction. Only the strong, and somewhat unrealistic informational assumptions underlying the best response dynamic in infinite populations preserves the Stackelberg prediction.

2 Bagwell's example

Consider the following game in extensive form.[4]

Now suppose as in Bagwell's (1995) paper that player II can observe player I's choice only with some error. To be precise, we assume that with probability $1-\varepsilon$ player II observes the action of player I correctly. With probability $\varepsilon < 1/2$ he receives the wrong signal. This game of imperfect information yields the following normal form $\Gamma(\varepsilon)$.

	FF	FC	CF	CC
L	$2,1$	$2-2\varepsilon, 1-\varepsilon$	$2\varepsilon, \varepsilon$	$0,0$
C	$3,0$	$1+2\varepsilon, 1-\varepsilon$	$3-2\varepsilon, \varepsilon$	$1,1$

[4] The payoffs of this game do not match exactly those of Bagwell (1995). In particular, in the usual duopoly setting player II would receive a higher payoff from F when I plays C rather than L. We can simplify this without loss of generality since all that matter is that II prefers to play F following L and C following C. The best reply structures of both games are equivalent.

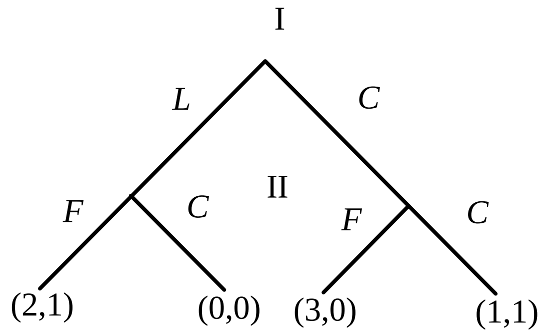

Figure 1: A simple extensive form game with perfect information.

Player II, the column player, has four pure strategies. E.g. FC stands for II's strategy of playing F in response to signal L and C in response to signal C. Let S_i denote player i's set of pure strategies and $\Delta(S_i)$ its mixed extension.[5] We will frequently write $A = A(\varepsilon) \in \mathbb{R}^{2\times 4}$ ($B = B(\varepsilon) \in \mathbb{R}^{2\times 4}$) for the payoff matrix of player I (II).

It is immediate that the Stackelberg strategies (L, FC), which are the unique subgame perfect equilibrium in the game of perfect information, are not an equilibrium in the game with noisy signals. The unique equilibrium in pure strategies is the Cournot equilibrium (C, CC). Note, that this equilibrium is strict. In addition, there are two mixed equilibria,

$$(\tilde{p}, \tilde{q}) := \left\{ (1-\varepsilon, \varepsilon), \left(\frac{1-4\varepsilon}{2-4\varepsilon}, \frac{1}{2-4\varepsilon}, 0, 0 \right) \right\}$$

and

$$(\hat{p}, \hat{q}) := \left\{ (\varepsilon, 1-\varepsilon), \left(0, \frac{1}{2-4\varepsilon}, 0, \frac{1-4\varepsilon}{2-4\varepsilon} \right) \right\}.$$

We call $(\tilde{p}, \tilde{q})$ the "noisy Stackelberg equilibrium" since it converges to the Stackelberg outcome as $\varepsilon \to 0$.

[5] If it leads to no confusion, we will continue to write (L, FF) or (e_L, e_{FF}) instead of $\{(1,0), (1,0,0,0)\}$.

3 Evolutionary dynamics

Evolutionary dynamics are a useful technique for testing the stability of a given Nash equilibrium. In the following this analysis will be undertaken using three different approaches. Generally it will be assumed that noise is realized independently across individuals.

3.1 Weakly payoff monotone dynamics

In this section we consider two infinite populations, one associated to each player in the game, in which individuals are continuously updating their actions. We search for states that are robust against rare mutations. Formally, we characterize asymptotically stable states. However, even if a state is asymptotically stable, we would consider it less plausible if its basin of attraction vanishes for $\varepsilon \to 0$.

Implicitly it is assumed that individuals obtain a payoff through random matching with one or several opponents from the opposite population and aim to maximize their expected payoffs. Individual experience and possible information about others leads to change in behavior. Instead of explicitly defining individual behavior we postulate weak conditions on aggregate behavior of the entire populations.

Changes in the population proportions are assumed to follow a *selection dynamic* (as defined by Samuelson and Zhang, 1992). This is a continuous time dynamic on $\Delta(S_1) \times \Delta(S_2)$

$$\begin{aligned} \dot{p}_i &= f_i(p,q), i \in S_1 \\ \dot{q}_j &= g_j(p,q), j \in S_2 \end{aligned}$$

with

1. $f_i, g_j : \Delta(S_1) \times \Delta(S_2) \to \mathbb{R}$ Lipschitz continuous,[6]

2. $\sum_{e_i \in S_1} f_i(p,q) = \sum_{e_j \in S_2} g_j(p,q) = 0$, and

[6] F_i is Lipschitz continuous if there exists $m_{F_i} > 0$ such that $F_i(p,q) - F_i(p',q') \leq m_{F_i} \|(p,q) - (p',q')\|$ for all $(p,q), (p',q') \in \Delta(S_1) \times \Delta(S_2)$.

3. $p_i = 0$ implies $f_i(p,q) \geq 0$, $q_j = 0$ implies $g_j(p,q) \geq 0$ for any $i \in S_1$ and $j \in S_2$.

The first condition guarantees that there is a unique solution. Moreover, it puts a bound on how much the gradient may change when the state (p,q) changes slightly. The other two conditions ensure that the dynamic stays in $\Delta(S_1) \times \Delta(S_2)$. Notice that the best response dynamic (Section 3.3) does not fit in this class since the gradient may change abruptly when there are small changes in the state (Lipschitz continuity fails). However, any dynamic that is based on individuals reacting to finite samples will belong to this class.

Definition 1 *We call a selection dynamic* weakly payoff monotone *in a given game if the following four conditions hold:*

1. $\left.\frac{f_k(p,q)}{p_k}\right|_{p_k=0} := \lim_{p_k \to 0} \frac{f_k(p,q)}{p_k}$ exists in $\mathbb{R} \cup \{-\infty, \infty\}$
2. $e_i A q \geq e_j A q$ for all j with strict inequality for some r such that $p_r > 0$ implies $\frac{f_i(p,q)}{p_i} > 0$.
3. $e_k A q \leq e_j A q$ for all j with strict inequality for some r such that $p_r > 0$ implies $\frac{f_k(p,q)}{p_k} < 0$.
4. The above properties also apply to $\dot{q}_j$ and $\frac{\dot{q}_j}{q_j}$ in their appropriate formulation.

We allow for infinite growth rates, which makes a scenario feasible where some individuals have enough knowledge of the game to stop playing some of their strategies (e.g., because they are strictly dominated). Instead of putting restrictions on growth rates of each strategy, we demand in conditions (2) and (3) that the growth rate of a best/worst response to the present state increases/decreases strictly if not all actions present achieve the same expected payoff.

The above definition generalizes several commonly used evolutionary dynamics. In particular, it covers the classes of payoff monotone (also known as compatible) and payoff positive dynamics (see Weibull, 1995, Chapt. 5, for definitions). Payoff monotonicity requires that growth rates of strategies are ranked according to their payoff. Payoff positivity requires that strategies earning above (below) average have positive (negative) growth rates.

In the following we present two examples of weakly payoff monotone dynamics (they are in fact aggregate monotone according to Samuelson and Zhang, 1992). The *standard continuous replicator dynamic* (Taylor, 1979) for a bimatrix game with payoff matrices (A, B) is defined as

$$\begin{aligned} \dot{p}_i &= p_i(e_i A q - pAq) \ , \\ \dot{q}_j &= q_j(pBe_j - pBq) \ , \end{aligned} \tag{1}$$

where e_i denotes the pure strategy i, $e_i = (0, 0, ..., 1, 0...0)$. Many individual learning models are approximated by this dynamic (e.g., Gale et al., 1995; Schlag, 1996).

A slightly modified version used mostly in biological applications, called the *adjusted continuous replicator dynamic* (Maynard Smith, 1982) is given by

$$\begin{aligned} \dot{p}_i &= p_i \frac{(e_i A q - pAq)}{pAq} \ , \\ \dot{q}_j &= q_j \frac{(pBe_j - pBq)}{pBq} \ . \end{aligned}$$

A state is called *(Lyapunov) stable* if trajectories starting sufficiently close stay arbitrarily close. A state is called *attracting* if there exists a neighborhood of this state such that trajectories starting in this neighborhood eventually converge to the state. *The basin of attraction* of an attracting state is the set of all states such that trajectories starting in such a state lead to the attracting state. *Asymptotic stability* means both stable and attracting. Sometimes the concept of asymptotic stability is too stringent and we need the following weaker concept. A closed set of rest points is called *interior asymptotically stable* if trajectories starting in the interior sufficiently close to the set stay arbitrarily close to the set and eventually converge to the set. This concept generalizes asymptotic stability to sets of rest points and additionally restricts attention to trajectories starting in the interior. An asymptotically stable state (or interior asymptotic stable set) can be viewed as a long run prediction of the underlying process given rare (interior) mutations.[7]

For the game $\Gamma(0)$ without noise and the standard continuous replicator dynamic Cressman and Schlag (1996) show that i) the Stackelberg equilibrium is contained in the unique interior asymptotically stable set $\left\{(e_L, (1-\lambda)\, e_{FC} + \lambda e_{FF}) : 0 \le \lambda \le \frac{1}{2}\right\}$ (This is the set of Nash equilibria that yield the Stackelberg outcome), and, ii) the Cournot equilibrium is (Lyapunov) stable but not asymptotically stable. The following proposition generalizes this result to our class of weakly payoff monotone dynamics.

Proposition 1 *For any weakly payoff monotone dynamic, when information is perfect (i.e., there is no noise) then both the Cournot and the Stackelberg equilibrium are Lyapunov stable. However, the Stackelberg equilibrium is contained in the unique interior asymptotically stable set.*

Proof. Let G be an interior asymptotically stable set. FC is a weakly dominant strategy for player II, hence, by Definition 1 (2), $\dot{q}_{FC} > 0$ in any interior state. Continuity of g_{FC} implies $\dot{q}_{FC}(p,q) \ge 0$ for any (p,q). Hence, G must contain a state in which $q_{FC} = 1$. Moreover, since $\dot{p}_L \ge 0$ holds when q_{FC} is sufficiently large, the Stackelberg equilibrium (L, FC) must be contained in G. Especially, (L, FC) is Lyapunov stable.

By definition, an interior asymptotically stable set is a closed set of rest points. This set must be connected by the stability requirement. Consequently, $G \subset \{(e_L, (1-\lambda)\, e_{FC} + \lambda e_{FF}) : 0 \le \lambda \le 1\}$.

If an element of G is not a Nash equilibrium then trajectories lead initially away from G (use the same trick as when proving "stability implies Nash", see e.g., Weibull, 1995, Proposition 4.8). Consequently G is a set of Nash equilibria. In the following we will show that G is equal to the Nash equilibrium component containing the Stackelberg equilibrium, i.e.,

$$G = \left\{(e_L, (1-\lambda)\, e_{FC} + \lambda e_{FF}) : 0 \le \lambda \le \frac{1}{2}\right\}.$$

[7] For formal definitions see Weibull (1995) and Cressman and Schlag (1996).

If $p_L = 1$ and $q_{FF} + q_{FC} > 0$ then Definition 1 (3) implies $\frac{g_{CF}(p,q)}{q_{CF}} < 0$ and $\frac{g_{CC}(p,q)}{q_{CC}} < 0$. Since g_j is continuous, with Definition 1 (1) it follows that $\frac{g_j(p,q)}{q_j}$ is a continuous function on the extended reals $\mathbb{R} \cup \{-\infty, \infty\}$. Hence, for $\tau > 0$ but sufficiently small, $\frac{g_{CF}(p,q)}{q_{CF}} < 0$, $\frac{g_{CC}(p,q)}{q_{CC}} < 0$ and hence $\dot{q}_{CC} + \dot{q}_{CF} \leq 0$ when $p_C < 3\tau$ and $q_{FF} + q_{FC} > \tau$. Since FC is a weakly dominant strategy for player II, continuity implies there exists $\mu > 0$ such that $1 - \tau > p_L > \tau$ and $q_{FC} > \frac{1}{4}$ implies $\dot{q}_{FC} > \mu$. If $q_{FF} = q_{FC} = \frac{1}{2}$ then $e_L A q = e_C A q$ and continuity of $f_L(p, \cdot)$ implies $\dot{p}_L = 0$. Consequently, there exists $0 < \nu < \tau$ such that $\dot{p}_L > -\mu$ and hence $\dot{p}_L + \dot{q}_{FC} > 0$ when $p_L > \tau$ and $\frac{1}{2} - \nu < q_{FC} < \frac{1}{2} + \nu$. Let $\alpha > 0$ be such that $e_L A q \geq e_C A q$ and hence $\dot{p}_L \geq 0$ when $q_{FC} > \frac{1}{2} + \nu$ and $q_{CC} + q_{CF} < \alpha$. Consequently, trajectories starting in $\left\{(p, q) : q_{FC} > \frac{1}{2} - \nu, p_C < \tau + 2\nu, q_{FC} - p_C > \frac{1}{2} - \nu - \tau, q_{CC} + q_{CF} < \alpha\right\}$ stay in this set. Moreover, $\dot{q}_{FC} \geq 0$ implies that trajectories converge to G. Since τ was arbitrary as long as it was sufficiently small it follows that G is an interior asymptotically stable set.

Consider now the Cournot equilibrium (e_C, e_{CC}). $e_L A e_{CC} < e_C A e_{CC}$ together with Definition 1 (3) implies $\lim_{p_L \to 0} \frac{f_L(p, e_{CC})}{p_L} < 0$. Lipschitz continuity of f_L implies there exists $N > 0$ such that $\dot{p}_L \leq -N p_L$ in a neighborhood U of (e_C, e_{CC}). $e_C B e_{CC} \geq e_C B e_j$ for all j implies $\dot{q}_{CC}(e_C, e_{CC}) = 0$. Lipschitz continuity implies there exists $M > 0$ such that $\dot{q}_{CC} \geq -M p_L$ in a neighborhood $U' \subset U$ of (e_C, e_{CC}). W.l.o.g. let $U' = \{(p, q) : M p_C + N q_{CC} > \beta\}$ for some $0 < \beta < M + N$ chosen sufficiently large. Consequently, $M \dot{p}_C + N \dot{q}_{CC} = -M \dot{p}_L + N \dot{q}_{CC} \geq 0$ in U' which implies that $M p_C + N q_{CC}$ is a local Lyapunov function. Trajectories starting in U' stay in U' and hence (e_C, e_{CC}) is stable. ■

Notice that non decreasing best responses are sufficient to guarantee Lyapunov stability of the Stackelberg equilibrium.

Now we will investigate dynamic stability for constant noise and as noise varies. Noise effects individual payoffs, similar to uncertainty due to random matching. However, since populations are infinite, we will assume that these stochastic effects are washed out, and more importantly, that aggregate behavior is only effected by expected payoffs. Although a common assumption in the literature, it is not a priori clear that "plausible" individual behavior will aggregate in this fashion.[8] Hence we consider evolution in the game generated by the payoff matrices $A(\varepsilon)$ and $B(\varepsilon)$.

Comparing dynamic stability under different degrees of noise requires that we specify how the dynamic changes as the underlying payoffs in the game change. We will only need the following very weak condition. Let $\tilde{A}$ be an alternative payoff matrix for player I and let $\tilde{f}$ be the associated selection dynamic of the player I population. If $e_i\tilde{A}e_k \geq e_iAe_k$ and $e_r\tilde{A}e_k \leq e_rAe_k$ for all $r \neq i$ and all k then $\tilde{f}_i(p,q) \geq f_i(p,q)$ for all (p,q). Similar properties should also hold for changes in player II's payoffs. Notice that $\dot{p}_i$ need not be continuous in the underlying payoffs as this would be too strong a condition in many cases.[9]

When signals are received with noise the long run prediction of play changes drastically. The Cournot equilibrium can be selected by any generalized payoff monotone dynamic, whereas robustness of the Stackelberg equilibrium depends on the exact specification of the dynamic. For the most common representative in the class of payoff monotone dynamics, the standard replicator dynamic, the stability properties of the Stackelberg equilibrium are inferior to that of the Cournot equilibrium.

Proposition 2 *Consider now the game with noise. For any generalized payoff monotone dynamic the Cournot equilibrium (C, CC) is asymptotically stable. Furthermore, the basin of attraction of (C, CC) does not vanish as $\varepsilon \to 0$. The only other candidate for an asymptotically stable state is the Stackelberg equilibrium $(\tilde{p}, \tilde{q})$; under the standard replicator dynamic $(\tilde{p}, \tilde{q})$ is stable but not asymptotically stable; under the adjusted replicator dynamic $(\tilde{p}, \tilde{q})$ is asymptotically stable.*

[8] Of course this assumption is justified when individuals use the proportional imitation rule yielding population behavior according to the standard replicator dynamic (1) (see Schlag, 1996).

[9] E.g., assume that player I has two actions 1 and 2, action 1 (2) yielding 0 (x) with certainty. Here it does not seem reasonable that a learning dynamic must be continuous at $x = 0$.

Remark 1 *Notice that we do not make any claim about interior asymptotically stable sets when there is noise. In this game all Nash equilibria are singletons, hence an interior asymptotically stable set corresponds to an asymptotically stable state (see arguments used in the proof of Proposition 1).*

Proof. First we show that (C, CC) is asymptotically stable with non vanishing basin of attraction. Consider a slightly modified game Γ' with the same payoff matrix as $\Gamma(0)$ except that the payoffs of CF are evaluated at $\varepsilon = \frac{1}{4}$, i.e., $e_L A e_{CF} = \frac{1}{2}$, $e_C A e_{CF} = \frac{5}{2}$ and $pBe_{CF} = \frac{1}{4}$. Retracing the steps in the proof of Proposition 1 it follows that there exists a neighborhood U' of (C, CC) where $Mp_C + Nq_{CC}$ is a local Lyapunov function (i.e., $M\dot{p}_C + N\dot{q}_{CC} \geq 0$) in Γ' for appropriate constants $M, N > 0$. Comparing Γ' to $\Gamma(\varepsilon)$ with $\varepsilon < \frac{1}{4}$ we see that our monotonicity condition implies that both $\dot{p}_C$ and $\dot{q}_{CC}$ increase. Consequently, $Mp_C + Nq_{CC}$ is also a local Lyapunov function in U' for $\Gamma(\varepsilon)$ for any $\varepsilon < \frac{1}{4}$. Especially, U' was constructed such that $\dot{p}_C \geq Np_L$ which means that $p_C \to 1$ as $t \to \infty$. For given ε and sufficiently large p_C it follows that $\dot{q}_{CC} > 0$ when $0 < q_{CC} < 1$ and hence trajectories starting in U' (which is independent of ε) converge to (C, CC).

Each asymptotically stable state is a Nash equilibrium (a trivial generalization of Friedman, 1991 to our class of generalized payoff monotone dynamics). The best reply structure close to $(\hat{p}, \hat{q})$ resembles that of a coordination game where $(\hat{p}, \hat{q})$ is the unstable interior mixed equilibrium. This will make $(\hat{p}, \hat{q})$ unstable. Consider $G = \left\{(p, q) : p_C > 1 - \varepsilon, q_{CC} > \frac{1-4\varepsilon}{2-4\varepsilon}\right\}$, then $(\hat{p}, \hat{q})$ is an accumulation point of G. Starting in G, C is the unique best response for player I and CC is the unique best response for player II which implies that $\dot{p}_C > 0$ and $\dot{q}_{CC} > 0$. Especially trajectories starting in G converge to (C, CC) which means that $(\hat{p}, \hat{q})$ is not stable, especially it is not asymptotically stable.

Consider now the Stackelberg equilibrium $(\tilde{p}, \tilde{q})$. The support of $(\tilde{p}, \tilde{q})$ is contained in $\Delta\{L, C\} \times \Delta\{FF, FC\}$. On this face, $\Gamma(\varepsilon)$ resembles matching pennies. Consider the standard continuous replicator dynamic. Trajectories cycle on this face (see, e.g. Weibull, 1995). Especially, this means that $(\tilde{p}, \tilde{q})$ is not asymptotically stable. However, restricting the dynamic to this face $(\tilde{p}, \tilde{q})$ is stable. Moreover, since $BR(\tilde{p}, \tilde{q}) = \Delta\{L, C\} \times \Delta\{FF, FC\}$ it follows that $(\tilde{p}, \tilde{q})$ is also stable in the entire space (this follows from centre manifold theory, Wiggins, 1990, see Cressman and Schlag, 1996, for an explanation of its application and for some examples). In the adjusted continuous replicator dynamic, $(\tilde{p}, \tilde{q})$ is asymptotically stable on the face $\Delta\{L, C\} \times \Delta\{FF, FC\}$ (see again Weibull, 1995). Now the fact that $BR(\tilde{p}, \tilde{q}) = \Delta\{L, C\} \times \Delta\{FF, FC\}$ makes $(\tilde{p}, \tilde{q})$ asymptotically stable. ■

3.2 Discrete selection dynamics

A large part of the recent literature on evolution and learning assumes a setting with discrete time and a finite number N of individuals in the role of each of the players. The dynamics considered result from the composition $\mathcal{M}(\mathcal{S})$ of a selection process $\mathcal{S}$ and a mutation process $\mathcal{M}$. Most selection processes either are a version of a myopic best reply process (see e.g. Kandori, Mailath and Rob, 1993, and Young, 1993) or some sort of imitation process (Schlag, 1996). Below we consider two alternative approaches to encompass both kinds of dynamics.

The *discrete selection process* $\mathcal{S}$ (which should not be confused with the continuous selection dynamics defined in the previous section) is represented by a finite Markov chain on the state space $\{(p, q) \in \Delta(S_1) \times \Delta(S_2)$ s.t. $p_i \cdot N \in \mathbb{N}$, $q_j \cdot N \in \mathbb{N}$ for all $i, j\}$ with the following property. Most evolutionary processes are characterized by an element of inertia. We model this by assuming that each period with a fixed and independent probability $\theta > 0$ an individual is not able to adopt a new strategy.

3.2.1 Payoff sensitive dynamics

In our first approach we consider dynamics based on the game $\Gamma(\varepsilon)$ in which the expected effects of noise are incorporated. The class of discrete selection processes we consider are called payoff sensitive, a property which is defined next. Let p and q denote the frequency distribution of strategies in population one and two, respectively.

Definition 2 *A discrete selection dynamic* $\mathcal{S}$ *is called* payoff sensitive *if*

(a) $prob(p_i^{t+1} > p_i^t) > 0 \Rightarrow \exists k \neq i$ *with* $p_k > 0$ *and* $e_i Aq \geq e_k Aq$.

(b) *If* $\exists i$ *with* $p_i > 0$ *and* $e_i Aq \geq e_k Aq$, $\forall k$ *with* $p_k > 0$ *and strict inequality for some* k, *then* $\exists j$ *with* $e_j Aq \geq e_i Aq$ *s.t.* $prob(p_j^{t+1} > p_j^t) > 0$.

(c) *Equivalent conditions hold for* q^t *and* q^{t+1}.

Condition (a) states that the frequency of a strategy can only be increased if there is another strategy present which performs weakly worse. Condition (b) states that unless all current strategies perform equally, either a currently best strategy or some other strategy, which does at least as well, increases in frequency with positive probability. Condition (b) demands in particular that the process does not come to a halt unless all present strategies perform equally.

The definition allows for dynamics in which new superior strategies enter the system (e.g. best responses) and for dynamics in which only strategies can be chosen that are already represented in the population (as in imitation processes). It covers weakly monotone dynamics (Samuelson, 1994), and therefore best response and "Darwinian" dynamics (Kandori, Mailath and Rob, 1993). But it also covers some imitation dynamics, in which strategies which currently perform better in round–robin matchings are imitated, e.g., the proportional imitation rule and "imitate if better" (Schlag, 1996, see Corollary 1 below).

The mutation process $\mathcal{M}$ results from assuming that in each round, with an independent probability $\phi > 0$, an agent randomizes uniformly over all of his strategies. The process is therefore ergodic.[10]

In the game without error ($\varepsilon = 0$) the results with respect to a discrete payoff sensitive dynamic are inconclusive. Since neither of the pure equilibria is strict, the class of payoff sensitive dynamics is too general to make a unique prediction.[11]

[10] For a good introduction to the graph–theoretic methods used in this section see Vega–Redondo (1996). They were originally introduced by Freidlin and Wentzell (1984).

[11] E.g. it depends on whether players who already play a best reply may switch to other strategies. If they are not allowed to do so, then generically (C, CC) is the unique

With noise the picture changes. (C, CC) is now a strict equilibrium and the remaining equilibria are mixed. Stochastic dynamics do not in general converge to mixed strategy equilibria in asymmetric games (see Oechssler, 1994, for some of the problems involved). Depending on the exact specification of the process a mixed equilibrium may even fail to be a restpoint of the selection dynamics. Hence, it is not surprising that the discrete payoff responsive dynamics select the strict equilibrium (C, CC) in $\Gamma(\varepsilon)$ if the population size is large enough.[12]

Proposition 3 *Let $\varepsilon > 0$ be given and consider an imitation dynamic. If the population size N is larger than $1/\varepsilon$, then the limit distribution of the dynamic $\mathcal{M}(\mathcal{S})$ for $\phi \to 0$ puts probability one on the equilibrium (C, CC).*

Proof. Note first that the support of the limit distribution of $\mathcal{M}(\mathcal{S})$ for $\phi \to 0$ is a union of absorbing sets of $\mathcal{S}$ (see e.g. Samuelson, 1994, Theorem 1). A set of states Q is absorbing with respect to $\mathcal{S}$ if $\mathcal{S}$ cannot cause the process to leave Q and any state in Q is reached within finite time from any other state.

Due to condition (b) of Definition 2 a singleton set can be absorbing only if all strategies present in a population earn the same payoff. Candidates for absorbing states are therefore all equilibria and all monomorphic states, that is, states in which all players of a population use the same strategy.

Given the best reply structure of $\Gamma(\varepsilon)$, inertia and condition (b) implies that from any non–absorbing state there exists a sequence of transitions, each occurring with positive probability, leading to some monomorphic state. Hence, each absorbing set contains either an equilibrium or a monomorphic state.

For $N > 1/\varepsilon$ it takes at least two mutations to leave the basin of attraction of (C, CC), that is, the set of states from which $\mathcal{S}$ returns to (C, CC) with

absorbing state. If they are allowed then it is possible to make (L, FC) the unique locally stable mutation connected component (see proof of Proposition 3).

[12] Note, however, that the result is more ambiguous than that of the last section as for a *given* population size there always exists an ε small enough such that no result of this kind can be obtained which holds for the entire class of payoff sensitive dynamics.

probability one. This follows because with only one mutation we have that $\forall j \in S_2$

$$e_C A \left[\frac{N-1}{N} e_{CC} + \frac{1}{N} e_j \right] > e_L A \left[\frac{N-1}{N} e_{CC} + \frac{1}{N} e_j \right]$$

and $\forall j \in S_2, j \neq CC$

$$\left[\frac{N-1}{N} e_C + \frac{1}{N} e_L \right] B e_{CC} > \left[\frac{N-1}{N} e_C + \frac{1}{N} e_L \right] B e_j \, .$$

Hence, by condition (a) the process must return to (C, CC) after one mutation.

Using the terminology of Nöldeke and Samuelson (1993) the collection of absorbing sets can be partitioned into (mutation connected) components. A component is called *locally stable* if it takes more than one mutation to reach any other component. Given that it takes at least 2 mutations to leave the basin of attraction of $\{(C, CC)\}$, this component is locally stable. We claim that the remaining components are not locally stable as one mutation is sufficient to reach some other component.

Consider first the monomorphic states. If monomorphic states belong to absorbing sets, then (C, FF), (C, FC), (L, FF), (L, FC) and (C, CF) belong to the same component as they form a cycle in the sense of Nöldeke and Samuelson (1993). Starting in (C, CF) , suppose there is one mutation to CC. By condition (a) the process can move only to states in which p_{CC} is increased. Therefore, the process converges to (C, CC) and the component is not locally stable. Likewise, (C, CC) can be reached from (L, CC) and (C, CF) can be reached from (L, CF). Consequently, (C, CC) is the unique monomorphic state contained in a locally stable component.

Next, consider $(\hat{p}, \hat{q})$. One mutation to CC puts the process in the basin of attraction of (C, CC).

Finally, the best reply structure on the face $\Delta\{FF, FC\} \times \Delta\{C, L\}$ are the same as in a Matching Pennies game. Due to the inertia assumption, with positive probability the dynamics spiral outwards and reach the set $\{(p, q) : p_L < \varepsilon \text{ and } p_{FF} + p_{FC} = 1\}$. From there by condition (a) CC will increase with positive probability. Hence, (C, CC) can be reached from $(\tilde{p}, \tilde{q})$

with one mutation. Consequently, $\{(C, CC)\}$ is the unique locally stable component.

By Proposition 1 of Nöldeke and Samuelson (1993) a state can appear in the support of the limit distribution only if it belongs to a locally stable component. Since $\{(C, CC)\}$ is the unique locally stable component and a limit distribution exists, (C, CC) has probability one in the limit distribution. ■

3.2.2 Imitation dynamics

In our second approach we explicitly incorporate noise as a stochastic effect. In each round, individuals from opposite populations are randomly matched. In a given match, $A(0)$ $(B(0))$ is the payoff matrix for player I (player II) with probability $1-\varepsilon$ and $A(1)$ $(B(1))$ is the payoff matrix for player I (player II) with probability ε. We focus on behavior that relies on information gathered through seeing the performance of others. More specifically, we consider discrete selection processes that result when individuals update their behavior by imitating others, e.g., using rules like "imitate if better" or the proportional imitation rule (Schlag, 1996).

Definition 3 *A discrete selection dynamic $\mathcal{S}$ is called an* imitation dynamic *if*

(a) $prob(p_i^{t+1} > p_i^t) > 0 \Leftrightarrow p_i > 0$ and $\exists j, r$ and $k \neq i$ with $p_k, q_j, q_r > 0$ and $\exists \varepsilon_1, \varepsilon_2 \in \{0, 1\}$ such that $e_i A(\varepsilon_1) e_j > e_k A(\varepsilon_2) e_r$.

(b) Equivalent conditions hold for q^t and q^{t+1}.

For imitation dynamics, in contrast to payoff sensitive dynamics, we find that the population size need not be sufficiently large in order to obtain Cournot as the unique long run outcome.

Corollary 1 *Let $\varepsilon > 0$ be given and consider an imitation dynamic. Then the limit distribution of the dynamic $\mathcal{M}(\mathcal{S})$ for $\phi \to 0$ puts probability one on the equilibrium (C, CC).*

Proof. It is easily shown that it takes at least one mutation in each population, and hence at least a total number of two mutations, to be able to leave the basin of attraction of (C, CC) under payoff imitative dynamics. The rest of the proof follows as in the proof of Proposition 3. ■

3.2.3 A note on the order of limits

The results above were obtained by considering – for fixed observational noise ε – the limit behavior of the dynamics when the probability of "mutations" ϕ vanishes. Especially, the dynamic conditions (Definitions 2 and 3) did not need to incorporate the fact that mutations exist since these were very small. It is also interesting to consider the reverse order of limits: What happens if we first let the probability of observational errors ε go to zero and then consider the dynamics as ϕ converges to zero?

Under payoff sensitive dynamics, noise could not be too small for given population size N in order to derive general results. Once $N < \frac{1}{\varepsilon} - 2$ one mutation suffices to leave (C, CC). Let us restrict attention to the best response dynamic. Once ε is sufficiently small for given N, (C, CC) becomes the unique absorbing state. Additionally, the dynamic does not depend on the exact size of ε which means that the order of limits does not effect the outcome for the best response dynamic. Hence, the limit distribution for vanishing mutation rate and vanishing noise puts probability one on the unique absorbing state (C, CC) (compare to Footnote 11).

Consider now payoff imitative dynamics. For fixed mutation rate ϕ and vanishing noise rate ε, unless CC is never imitated, we find that both Cournot and Stackelberg equilibrium belong to the same locally stable mutation connected component. Consequently, selection forces are too weak here to select between the two equilibria of interest.

Which order of limits is more plausible depends on whether one thinks that trembles in the *execution* of strategies or in their *perception* are more likely to occur.[13]

[13] See Adolph (1996) for a similar argument.

3.3 Continuous best response dynamic

The continuous best response dynamic (Matsui, 1992, Hofbauer, 1995) is defined as

$$\dot{p} = MBR(q) - p$$
$$\dot{q} = MBR(p) - q$$

where $MBR(x)$ is a (possibly discontinuous) selection from the (mixed) best response correspondence to the profile x. The interpretation is that at any instant of time a small fraction of each (infinite) population is allowed to adjust its strategy and chooses a best reply against the current profile. While each player chooses a pure strategy, mixtures are possible since different players may choose different pure strategy best responses.[14] Note that players are assumed to know the exact distribution of strategies in the other population when adjusting their strategies. This process does not result if each player samples only a finite number of players from the opposite population and plays a best response to his/her sample. In fact, Schlag (1997) shows that such best response play to finite samples does not even generate a weakly payoff monotone dynamic in Bagwell's example. Thus, the informational requirements underlying the continuous best response dynamic are quite strong.

When signals are observed without error, FC is the unique best reply for player II in any interior state. Given a sufficiently large proportion of player II individuals choosing FC, any player I individual will choose L . Consequently, we have the following result.

Remark 2 *Without noise, any trajectory starting in the interior will converge to the Stackelberg outcome.*

Below we will see that this unambiguous prediction of the Stackelberg outcome under the continuous best response dynamic will carry over when signals are noisy.

[14] Hofbauer (1995) shows that the continuous best reply process is in some sense equivalent to fictitious play.

Proposition 4 *For* $0 < \varepsilon < 1/2$ *there are two asymptotically stable states,* (C, CC) *and* $(\tilde{p}, \tilde{q})$. *While the basin of attraction of* (C, CC) *vanishes as* $\varepsilon \to 0$, *the basin of* $(\tilde{p}, \tilde{q})$ *converges to the set of all interior states.*

Proof. (C, CC) is a strict equilibrium and hence asymptotically stable.

We will show that trajectories starting in

$$M(\varepsilon) = \{(p, q) \in \Delta(S_1) \times \Delta(S_2) \mid p_2 < \min\{1 - \varepsilon, 1 - \sigma + \sigma q_2\}\},$$

where $\sigma := \frac{\varepsilon(2-4\varepsilon)}{1-4\varepsilon}$, converge to $(\tilde{p}, \tilde{q})$. This will complete the proof since for $\varepsilon \to 0$ $M(\varepsilon)$ converges to $\Delta(S_1) \times \Delta(S_2)$.

Consider a state $(p, q) \in M(\varepsilon)$. Since $p_2 < 1 - \varepsilon$, CC is not a best reply. Note that CF is strictly dominated for all $\varepsilon < 1/2$ and is therefore never a best reply. Thus, for all $(p, q) \in M(\varepsilon)$, FC or FF are best replies for player II.

We claim trajectories starting in $M(\varepsilon)$ stay in $M(\varepsilon)$. Initial states in which L is a best reply for player I are unproblematic since then $\dot{p}_1 > 0$ and $M(\varepsilon)$ cannot be left.

Consider next initial states (p, q) in which C is a best reply for player I, which implies that

$$q_2 \leq \hat{q}_2 = \frac{1}{2 - 4\varepsilon}.$$

Suppose first that FC is a best reply for player II, i.e. $p_2 \geq \varepsilon$. The best response dynamics are always pointed in the direction of the best replies, in this case, (C, FC). Thus, we have to show that

$$(1 - \lambda)(p, q) + \lambda(C, FC) \in M(\varepsilon),$$

$\forall \lambda < \frac{1}{2-4\varepsilon} \frac{1-q_2(2-4\varepsilon)}{(1-q_2)}$, i.e. for all λ such that the convex combination remains in the region where (C, FC) are best replies. In particular, it must hold that

$$(1 - \lambda)p_2 + \lambda \leq (1 - \lambda)(1 - \sigma + \sigma q_2) + \lambda,$$

which is satisfied by construction of $M(\varepsilon)$.[15]

[15] Note that for $q_2 < \frac{1}{2-4\varepsilon}$, $1 - \varepsilon > 1 - \sigma + \sigma q_2$.

Finally consider the case that FF is a best reply for player II at (p, q), which implies that $p_2 \leq \varepsilon$. In this case $\dot{p}_2 > 0$ but as long as $p_2 \leq \varepsilon$, $M(\varepsilon)$ cannot be left, which proves the claim.

Since only FC or FF can be best replies for player II in $M(\varepsilon)$, all trajectories starting in $M(\varepsilon)$ have limit points in the face $H := \Delta\{L, C\} \times \Delta\{FC, FF\}$. Trajectories starting in H stay in H. On H, $\Gamma(\varepsilon)$ is a rescaled version of 'matching pennies'. By Theorem 7 in Hofbauer (1995) the continuous best response dynamic on H converges to the unique Nash equilibrium $(\tilde{p}, \tilde{q})$ of this restricted game.

What remains to show is that trajectories approaching H behave like trajectories starting on H. This can be done by defining an appropriate distance function of the trajectory on H to $(\tilde{p}, \tilde{q})$ that decreases strictly over time for trajectories on H. Consequently, this distance also decreases strictly for trajectories that are sufficiently close to H. This can be used to show that trajectories starting in the interior converge to the noisy Stackelberg equilibrium $(\tilde{p}, \tilde{q})$.[16]

Finally, $(\hat{p}, \hat{q})$ is not stable since there are arbitrarily close points to it that belong to $M(\varepsilon)$, which means that there are trajectories that start close to it and converge to $(\tilde{p}, \tilde{q})$. ■

References

[1] Adolph, B. (1996), "Commitment, Trembling Hand Imperfection and Observability in Games", mimeo, Humboldt University, Berlin.

[2] Bagwell, K. (1995), "Commitment and Observability in Games", *Games and Economic Behavior*, 8, 271-280.

[3] Cressman, R., and Schlag, K. (1996), "The Dynamic (In)Stability of Backward Induction", SFB 303 Discussion Paper No. B-347, University of Bonn.

[16] Use a similar trick as in Hirsch and Smale (1974, Problem 2, p. 309).

[4] FRIEDMAN, D. (1991), "Evolutionary Games in Economics", *Econometrica*, 59, 637-666.

[5] GALE, J., BINMORE, K., and SAMUELSON, L. (1995), "Learning to be Imperfect: the Ultimatum Game", *Games and Economic Behavior,* 8, 56-90.

[6] GÜTH, W., KIRCHSTEIGER, G., and RITZBERGER, K. (1995), "Imperfectly Observable Commitments in n–Player Games", mimeo, Humboldt University, Berlin.

[7] HARSANYI, J., and SELTEN, R. (1988), *A General Theory of Equilibrium Selection in Games*, Cambridge, Mass.: MIT–Press.

[8] HIRSCH, M., and SMALE, S. (1974), *Differential Equations, Dynamical Systems and Linear Algebra*, Academic Press.

[9] HOFBAUER, J. (1995), "Stability for the Best Response Dynamics", mimeo, University of Vienna.

[10] KANDORI, M., MAILATH, G., and ROB, R. (1993), "Learning, Mutation, and Long Run Equilibria in Games", *Econometrica*, 61, 29-56.

[11] MATSUI, A. (1992), "Best Response Dynamics and Socially Stable Strategies", *Journal of Economic Theory*, 57, 343-362.

[12] MAYNARD SMITH, J. (1982), *Evolution and the Theory of Games*, Cambridge: Cambridge University Press.

[13] NÖLDEKE, G., and SAMUELSON, L. (1993), "An Evolutionary Analysis of Backward and Forward Induction", *Games and Economic Behavior,* 5, 425-454.

[14] OECHSSLER, J. (1994), "An Evolutionary Interpretation of Mixed Strategy Equilibria", mimeo Humboldt University, Berlin, forthcoming: *Games and Economic Behavior.*

[15] Samuelson, L. (1994), "Stochastic Stability with Alternative Best Replies" *Journal of Economic Theory*, 64, 35-65.

[16] Samuelson, L., and Zhang, J. (1992), "Evolutionary Stability in Asymmetric Games", *Journal of Economic Theory,* 57, 363–391.

[17] Schlag, K. (1996), "Why Imitate, and if so, How? A Bounded Rational Approach to Multi-Armed Bandits", Discussion Paper B-361, University of Bonn.

[18] Schlag, K. (1997), "Sophisticated Imitation", mimeo, University of Bonn.

[19] Taylor, P. (1979), "Evolutionary stable Strategies With Two Types of Players", *Journal of Applied Probability*, 16, 76-83.

[20] Van Damme, E., and Hurkens, S. (1994), "Games with Imperfectly Observable Commitments", CentER Discussion Paper, No. 9464.

[21] Vega–Redondo, F. (1996), *Evolution, Games, and Economic Behaviour,* Oxford: Oxford University Press.

[22] Weibull, J. (1995), *Evolutionary Game Theory*, Cambridge, Mass.: MIT–Press.

[23] Wiggins, S. (1990), *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Springer-Verlag: New York, Heidelberg, Berlin.

[24] Young, P. (1993), "Evolution of Conventions", *Econometrica*, 61, 57-84.