Learning to Face Stochastic Demand

SJUR DIDRIK FLÅM*   MARIA SANDSMARK†

September 11, 2000

ABSTRACT. We consider repeated interaction among several producers of a homogeneous, divisible good, traded at a common market. Demand is uncertain, and its law is unknown. We explore an adaptive scheme leading such producers, over time, to face correct demand data. Extensions include noncooperative games in which strategic interaction is felt via exactly two real parameters.
Key words: repeated play, learning, Nash equilibrium, Cournot oligopoly, mean-value iterates, divergence, stochastic approximation.
JEL classification: C72, D83, L13.

1. INTRODUCTION

This paper considers a fixed, finite set I of producers of a homogenous good, traded at a common market. Each individual $i \in I$ is quite competent in regulating his output $q_i \geq 0$ and also in measuring his marginal cost $MC_i(q_i)$. *But* - by assumption - he is little informed about market demand or strategic interaction. A main issue naturally emerges in this setting: *How can such agents eventually come to identify salient features of the demand curve*?

In analyzing this question our interest coincides with a substantial literature on bounded rationality and learning, dating back to Simon (1955).[1] That literature, already rich, offers new perspectives on equilibrium. In fact, such perspectives are dominant in recent books on game theory, including Weibull (1995), Young (1998), Fudenberg and Levine (1998). They are also amply brought out in macro-economic studies of how learning may converge to rational expectations, see Evans and Honkapohja (1997) for a survey.

This paper has a similar objective, namely to tell a simple story about the learning of market demand. The leading question is: *how does individual i get to grasp the appropriate version of his marginal revenue* MR_i? That entity is, of course, indispensable for good decision making - and part of his optimality condition

$$MR_i = MC_i(q_i). \tag{1}$$

*Norwegian School of Economics and Business Administration; e-mail: sjur.flaam@econ.uib.no.
Thanks are due Meltzers høyskolefond and Ruhrgas.

†This author thanks for support from NFR project Quantec.

[1] The study of less than rational behavior is supported by evidence brought forth by psychologists and experimental economists - let alone introspection - which often discard the unbounded rationality hypothesis assumed to govern decision makers in traditional economic theory. See Conlisk (1996) for a discussion and references.

Presumably, the right hand side $MC_i(q_i)$ in (1) derives directly, and maybe easily, from a well known, convex, lower semicontinuous, cost curve $c_i(\cdot)$. The left hand side, MR_i, of (1) requires, however, insights that are less immediate.[2] Indeed, we shall consider instances where i earns random profit

$$\pi_i(\xi, Q, q_i) := P(\xi, Q)q_i - c_i(q_i)$$

and worships maximization of its expected value $E\pi_i(\xi, Q, q_i)$. A salient feature here is that q_i must be committed in face of demand uncertainty, before the underlying random variable ξ is unveiled or realized. We take each i to be risk neutral and the game to be repeated time and again. The expectation operator E thus reflects the main concern with long run average profit. Maximization of average profit is, however, rendered difficult since the mapping $(Q, q) \mapsto E\pi_i(\xi, Q, q_i)$ hardly is available. At least two sorts of hurdles explains its unavailability: First, the arguments - the exogenous variable ξ or the endogenous aggregate supply $Q := \sum_{i \in I} q_i$ - could be hard to observe; second, the price curve $(\xi, Q) \mapsto P(\xi, Q)$ or the distribution of ξ might remain unidentified or hidden. Such obstacles certainly make the task of i difficult. How can he optimize when no closed-form, explicitly known objective is within reach?

To make progress in these matters we posit that $\pi_i(\xi, q_i + \sum_{j \neq i} q_j, q_i)$ - or more generally, that $E\pi_i(\xi, q_i + \sum_{j \neq i} q_j, q_i)$ - be concave with respect to own decision q_i. Then, granted appropriate differentiability of the expected market curve $Q \mapsto EP(\xi, Q)$, condition (1), comes in the form

$$MR_i := \frac{\partial}{\partial q_i}\left[EP(\xi, Q)q_i\right] \in \partial c_i(q_i) =: MC_i(q_i), \tag{2}$$

which is both necessary and sufficient for optimality. The left hand side of (2) is the partial derivative $\frac{\partial}{\partial q_i}$ of expected earnings $EP(\xi, Q)q_i$, and the right hand side $\partial c_i(q_i)$ denotes the generalized (sub-)differential of convex analysis (Rockafellar, 1970). Under fairly innocuous regularity hypotheses on P, we can recast the preceding inclusion (2) in a more familiar mold

$$EP(\xi, Q) + E\frac{\partial}{\partial Q}P(\xi, Q)q_i \in \partial c_i(q_i), \tag{3}$$

which allows us to reiterate the problem more precisely: Suppose each players knows only his own production and lacks the experience or knowledge needed to execute E. In particular, suppose no player can observe data pertaining to his rivals, be it their actions or cost functions. *Then, how can such uninformed, unexperienced, or*

[2] To estimate how the price responds to increased supply, producers could let marketing research provide data from surveys or panel studies, or from benefit-cost studies geared at estimating customers' willingness-to-pay. We assume here that such analyses require too much effort or excessive outlays.

uneducated agents eventually succeed in having (3) *satisfied for all* i? *And how can such satisfaction finally result from non-coordinated, decentralized actions?*

A simple observation brings us somewhat forward. Note that the prediction problem, of each and every firm, reduces to the learning of equilibrium values of merely *two* parameters, namely:

$$\text{the expected price } p_1 := EP(\xi, Q) \text{ and the expected slope } p_2 := E\frac{\partial}{\partial Q}P(\xi, Q). \quad (4)$$

So, we ask: Can the concerned parties - via an explicit process - finally get to form a correct, common prevision $p = (p_1, p_2)$? If so, then, in the long run, each $i \in I$ brings forth a quantity $q_i = q_i(p)$ that satisfies

$$p_1 + p_2 q_i \in \partial c_i(q_i), \quad (5)$$

and the associated aggregate $Q = Q(p) := \sum q_i(p)$ confirms p as required in (4). Such a distinguished pair p, embodying rational expectation in the mean, is here called a *steady state*. Each steady state p generates an associated *Cournot-Nash equilibrium* $q(p) := (q_i(p))_{i \in I}$ of the underlying oligopoly. We shall reasonably assume that there exists a nonempty, finite set of steady states (and whence of equilibria).

Our concern is whether and how convergence to that discrete set may obtain in the long run. Under fairly weak hypotheses Section 2 provides an affirmative and constructive answer. We invent and offer there a rather appealing, simple story about the learning of market demand as characterized at any stage by the prevailing pair $p = (p_1, p_2)$. Moreover, we produce a tractable and novel algorithm much akin to so-called fictitious play, but one which requires little memory, no statistical skills, and only the capacity to solve (5) repeatedly. Our approach is therefore different from the adaptive learning processes found in Bray (1982), Thorlund-Petersen (1990), and Kalai and Lehrer (1995).[3] The main vehicle will be stochastic approximation theory (Benaim, 1996) coupled with the Bendixon-Poincaré results on two-dimensional flows (see e.g. Nemytskii and Stepanov, 1960). The use of stochastic approximation methods in analyzing convergence of learning heuristics in economic contexts was initiated by Marcet and Sargent (1989), see also Sargent (1993) and Evans and Honkapohja (1994).

The rest of the paper is organized as follows. Section 2 presents the adaptive learning scheme and establishes convergence of repeated updating. Section 3 looks briefly at three classical market instances: monopoly, Cournot oligopoly, and competitive behavior. Section 4 concludes.

2. REPEATED MARKET INTERACTION

A few words on notation and technicalities are in order. Any pair $p = (p_1, p_2)$ must belong to a nonempty, compact, convex set $\mathbb{P} \subset \mathbb{R}_+ \times \mathbb{R}_-$ prescribed a priori. The form and nature of that set shall not occupy us any further.

[3] Other related studies include Foster and Vohra (1997), and Flåm (1998a).

Recall that each cost function $c_i : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous convex. Let the interval

$$domc_i := \{q_i \in \mathbb{R} : c_i(q_i) < +\infty\}$$

be its so-called effective domain. Then, for every parameter configuration $p_1 \geq 0, p_2 < 0$, Eq. (5) has a unique solution $q_i = q_i(p) \in domc_i$ depending continuously on p. In the exceptional case when $p_2 = 0$, we assume that (5) still has a unique, continuously dependent solution $q_i = q_i(p) = q_i(p_1)$, $p_1 \geq 0$. So, in all circumstances, we write $Q(p) := \sum_{i \in I} q_i(p)$ to denote the aggregate supply generated by p. If individual i solves (5), and $c_i(\cdot)$ is twice differentiable convex on its effective domain $domc_i$, then

$$\frac{\partial q_i(p)}{\partial p_1} \geq 0 \text{ and } \frac{\partial q_i(p)}{\partial p_2} \geq 0.$$

Consequently,

$$\frac{\partial Q(p)}{\partial p_1} \geq 0 \text{ and } \frac{\partial Q(p)}{\partial p_2} \geq 0. \tag{6}$$

The inequalities (6) are easy to grasp: They simply say that aggregate supply increases as the expected price $p_1 \geq 0$ or the expected slope $p_2 \leq 0$, figuring in (5), become more favorable.

We are now ready to look at market interaction as a dynamic process. For our purposes it begins at time 0 and happens thereafter at discrete time periods $t = 1, 2, \ldots$. The agents start with a common, initial belief $p^0 := (p_1^0, p_2^0)$, determined by accident or historical factors not discussed here. Generally, the point $p^t := (p_1^t, p_2^t)$ has the status of a common belief, or a public prevision, held by everyone just before decentralized production is undertaken at time t. Then, based on p^t, each individual $i \in I$ produces a quantity q_i^t which solves

$$p_1^t + p_2^t q_i^t \in \partial c_i(q_i^t).$$

In other words, everyone invariably provides a best response to his actual beliefs. Thereafter the market clears. That is, aggregate demand equals total supply $Q^t = \sum_{i \in I} q_i^t$, and all producers see the realized pair

$$\hat{p}^t := \left[P(\xi^t, Q^t), \frac{\partial}{\partial Q} P(\xi^t, Q^t) \right], \tag{7}$$

belonging by assumption to $\mathbb{P}$. (Note that the underlying pair (ξ^t, Q^t) itself may well be unobservable or hidden.) Typically, the realized $\hat{p}^t$defeats the predicted p^t. Therefore, the latter must be compromised with the first. So, after market closure, at the end of period t, all individuals update their belief like a moving average as follows:

$$p^{t+1} := (1 - \lambda_t) p^t + \lambda_t \hat{p}^t. \tag{8}$$

Here the weight $\lambda_t \in [0, 1]$ strikes a balance between the most recently held opinion p^t and the fresh observation $\hat{p}^t$. To see this learning mechanism more clearly, recall that $Q(p) = \sum_{i \in I} q_i(p)$, with $q_i(p)$ being the solution of (5). Now define a function

$$F(p, \xi) := \left[P(\xi, Q(p)), \frac{\partial}{\partial Q} P(\xi, Q(p)) \right],$$

which records the realized price and its slope. Inserting $\hat{p}^t = F(p^t, \xi^t)$ in (8), we get a first-order, stochastic, typically non-linear, difference equation

$$p^{t+1} = (1 - \lambda_t) p^t + \lambda_t F(p^t, \xi^t). \tag{9}$$

We stress that the initial belief - and the subsequent learning scheme as well - is common to all players. Clearly, these assumptions are very restrictive. They are justified here partly by tractability. Relaxation of either assumption would make a $2|I|$-dimensional dynamical system replace our simple one having planar habitat $\mathbb{R}^2$. High dimensions make convergence analysis more difficult or less transparent. Additional justification for the commonality of (9) could be statistical in nature or refer to available expertise, pointing to p^0 as the most likely initial guess - and to λ_t the most reasonable appropriate weight at stage t.

Anyway, as more experience accumulates, i.e., when many observations have been incorporated into the prevailing belief, new observations should be assigned less weight. To account for this feature, and also for the importance of continued learning, we posit that

$$\sum_{t=0}^{\infty} \lambda_t = +\infty, \text{ and } \sum_{t=0}^{\infty} \lambda_t^2 < +\infty. \tag{10}$$

In particular, the choice $\lambda_t = \frac{1}{1+t}$ would be applicable, most natural, and yield an empirical average $p^{t+1} = \frac{p^0 + \hat{p}^1 + \cdots + \hat{p}^t}{1+t}$ which complies with Bayesian updating. Note that $\lambda_t \to 0$. Note also that (9) is accompanied - and driven - by the exogenous stochastic process $\xi^0, \xi^1, \ldots$. *We assume that these variables are independent and identically distributed.*[4] We now state our main result forthwith:

Theorem 1 (Convergence to beliefs that are confirmed in the mean) *Suppose P is compact convex. Also suppose that the averaged function*

$$f(p) := EF(p, \xi) := \left[EP(\xi, Q(p)), E\frac{\partial}{\partial Q} P(\xi, Q(p)) \right]$$

[4]This assumption may be relaxed in terms of so-called mixing conditions (Benaim, 1996). Also, to capture mutual interdependence between several markets, these possibly being charachterizered by ξ, there might exist a Markovian feedback mechanism from p to ξ. If so, one would presume that each fixed p yields an ergodic Markov chain in ξ. We shall not elaborate on such generalities; see Benveniste et al. (1990).

only has isolated fixed points in P and is continuously differentiable there with divergence

$$div f(p) := \frac{\partial f_1(p)}{\partial p_1} + \frac{\partial f_2(p)}{\partial p_2} < 2.$$

Then, for any initial $p^0 \in P$, the sequence $\{p^t\}$ converges almost surely to a fixed point $p = f(p)$.

Proof. For the sake of the argument, let $\tau_0 := 0$ and $\tau_k := \lambda_0 + \cdots + \lambda_{k-1}$, $k = 1, 2, \ldots$ define a new intrinsic time scale. Observe via (10) that $\tau_k \to +\infty$ and $\tau_{k+1} - \tau_k = \lambda_k \to 0$. Also observe that (9) can be rewritten in the form

$$p^{t+1} = p^t + \lambda_t \left[F(p^t, \xi^t) - p^t\right],$$

or equivalently, using the notation $p(\tau_k) = p^k$, it comes as a difference quotient

$$\frac{p(\tau_{k+1}) - p(\tau_k)}{\tau_{k+1} - \tau_k} = F(p(\tau_k), \xi^k) - p(\tau_k) = f(p(\tau_k)) - p(\tau_k) + e_k$$

where $f(p) = EF(p, \xi)$ is the averaged function, and where

$$e_k := F(p(\tau_k), \xi^k) - f(p(\tau_k))$$

denotes an "*error*" or a deviation from the mean value. Thus, in heuristic terms, when ignoring the error e_k, process (9) can be seen as a numerical integration scheme à la Euler of the ordinary differential equation

$$\frac{d}{d\tau} p(\tau) = f(p) - p. \tag{11}$$

We next discuss the asymptotic behavior of (11). Note first that any solution trajectory of (11), starting in $\mathbb{P}$, will remain inside that set forever. The simple reason is that since $f(p) \in \mathbb{P}$, the velocity $f(p) - p$ points inwards at every boundary point $p \in \mathbb{P}$. In particular, since $\mathbb{P}$ is compact, every orbit must be bounded. The Bendixon-Poincaré theory of bounded planar flows says that any solution trajectory of (11) must converge either to a stationary point or to a periodic solution (or a to limit cycle), cf. Nemytskii and Stepanov (1960) and Katok and Hasselblatt (1995). We see that

$$div\,[f(p) - p] = div f(p) - 2 < 0,$$

and hence, by Bendixon's criterion, (11) has no periodic solution or limit cycle in the bounded domain $\mathbb{P}$. The upshot is that any solution $p(\tau)$ of (11) converges to a steady state as $\tau \to +\infty$, that is, the system is globally asymptotically stable with steady states as the only possible limits.

Knowing now the good limit properties of (11), the rest follows from stochastic approximation theory, Benaim (1996) Theorem 1.2 and Corollary 3.7. That theory says that, under our conditions, (9) and (11) have the same asymptotic limit sets. □

Corollary 1. (Convergence to Cournot-Nash oligopolistic market equilibrium) *Suppose* $\frac{\partial}{\partial Q}EP(Q,\xi) \leq 0$, *and* $\frac{\partial^2}{\partial Q^2}EP(Q,\xi) \leq 0$, *and that every cost function* c_i *is twice differentiable convex on* $dom c_i$. *If every individual updates his prevision*

$$p = \begin{cases} p_1 & \text{about the expected price } EP(Q,\xi), \\ p_2 & \text{about the expected slope } \frac{\partial}{\partial Q}EP(Q,\xi) \end{cases}$$

according to (9), then it follows that each producer's production converges almost surely to a long-run Cournot-Nash equilibrium level for the game in which i has objective $EP(\xi, q_i + \sum_{j \neq i} q_j)q_i - c_i(q_i)$.

Proof. Simply note via (6) that

$$div f(p) = \frac{\partial}{\partial Q}EP(Q,\xi)\frac{\partial Q}{\partial p_1} + \frac{\partial^2}{\partial Q^2}EP(Q,\xi)\frac{\partial Q}{\partial p_2} \leq 0$$

and invoke Theorem 1. □

Similar arguments are found in Kaniovski and Young (1995) and Flåm (1998b). See also Keenan and Rader (1985) and Corchon and Mas-Colell (1996).

3. EXAMPLES

We shall consider three market instances, ranked in terms of the number $|I|$ of oligopoly members.

3.1. Monopoly. An extreme setting has I as a singleton, i.e., $|I| = 1$, and $q_i = Q$. Then, most likely, the producer already knows, or quickly detects, that he is the only supplier. Such knowledge is not necessary though. Indeed, he may well believe, for quite a while or even persistently, that rival agents also supply the market, but he refrains from exploring that matter. It suffices instead for him to observe (7) at the end of each time period t, and to keep on responding as described above. In particular, he need not ascertain or record the underlying process (ξ^t, Q^t).

3.2. Cournot Oligopoly. Now, let $|I| \geq 2$. Our scenario then fits repeated play of a Cournot oligopoly, generalized here to comprise demand uncertainty. Instead of looking at iterated best responses to observed strategy profiles, we focus on recursive identification of merely *two* endogenous parameters, namely the average price and its slope in equilibrium. When $|I| > 2$ this is a useful simplification. In fact, we avoid the divergence that may plague fictitious play of nonzero-sum games. Besides, the informational requirements are here quite modest. Players may never get to see each other directly. They all act in front of uncertainty; they all respond to only two parameters.

It is instructive to consider briefly the degenerate instance featuring no uncertainty. Even then an individual i cannot easily consider or identify his optimality condition

$$P(Q) + P'(Q)q_i \in \frac{\partial}{\partial q_i} c_i(q_i),$$

in exact numerical form. Instead, not knowing $P(\cdot)$, and possibly not observing aggregate supply, all producers entertain a belief $p = (p_1, p_2)$ about $(P(Q), P'(Q))$ and solve their surrogate optimality conditions (5) for all i. Alternatively, but somewhat less appealing, one might say that each producer i believes he obtains profit

$$\pi_i(q_i, p) = (p_1 + \frac{p_2}{2} q_i)q_i - c_i(q_i).$$

Then his best response satisfies (5).

To solve (5) is often easy. It amounts though, when each c_i is smooth on $domc_i = [0, +\infty)$, to solve the system

$$\left.\begin{array}{l} p_1 + p_2 q_i = c_i'(q_i) - \mu_i, \\ q_i, \mu_i \geq 0 \text{ and } q_i \mu_i = 0 \end{array}\right\} \text{ for all } i \in I,$$

μ_i being the Lagrange multiplier associated with the constraint $q_i \geq 0$. To avoid these multipliers, and to accommodate nonsmooth cost functions, instead of addressing (5), we use the following alternative procedure:

Solve the stage game as follows: Let $t \longleftarrow t+1$ and $\lambda \longleftarrow \frac{1}{t}$. Given the actual common belief p, solve the auxiliary optimization problem

$$\begin{array}{ll} \text{maximize} & p_1 \sum_{i \in I} q_i + \frac{p_2}{2} \sum_{i \in I} q_i^2 - \sum_{i \in I} c_i(q_i) \\ \text{subject to} & q_i \geq 0 \text{ for all } i. \end{array}$$

Let $Q := \sum_{i \in I} q_i(p)$ be the ensuing total supply.

Update beliefs by setting $\left\{ \begin{array}{ll} p_1 \longleftarrow & (1-\lambda)p_1 + \lambda P(Q), \\ p_2 \longleftarrow & (1-\lambda)p_2 + \lambda P'(Q). \end{array} \right.$

Continue to solve the stage game until convergence.

A numerical illustration: To simplify suppose there is no uncertainty. Let I comprise firms $i = 1, \ldots, 5$, having cost functions $c_i(q_i) = \alpha_i q_i + 0.1 q_i^2$ with $\alpha_1 = 2, \alpha_2 = 4, \alpha_3 = 5, \alpha_4 = 7$, and $\alpha_5 = 8$, respectively. There is given an "unknown" demand curve $P(Q) = 1000 - 2Q$ and a common initial belief $p^0 = (50, -1)$. The game starts at time $t = 0$ and brings then out

$$P(Q^0) = 626.668 \text{ and } P'(Q^0) = -2.$$

The updated belief becomes $p^1 = (626.688, -2)$, which in its turn yields the realization

$$P(Q^1) = 0 \text{ and } P'(Q^1) = -2,$$

and the new belief $p^2 = (313.334, -2)$. The development in supply Q and individual output q_i is given in Table 1 below.

In this example, the correct slope is quickly learned whereas identification of the equilibrium price requires more time. Anyway, as the table shows, convergence comes fairly quickly, and we choose to stop after 13 iterations. The realized price and slope are then

$$P(Q^{13}) = 184.590 \text{ and } P'(Q^{13}) = -2,$$

which form a subsequent belief $p^{14} = (184, 590, -2)$.

Table 1

Development in Qand q_i

$t \backslash$	Q	q_1	q_2	q_3	q_4	q_5
0	186.666	40.000	38.333	37.500	35.833	35.000
1	1412.427	283.940	283.031	282.576	281.667	281.213
2	700.304	141.515	140.606	140.152	139.243	138.788
3	462.929	94.040	93.131	92.677	91.768	91.313
4	386.368	78.728	77.819	77.365	76.455	76.001
5	410.032	83.461	82.552	82.097	81.188	80.734
6	407.882	83.031	82.122	81.667	80.758	80.304
7	407.742	83.003	82.094	81.639	80.730	80.276
8	407.716	82.998	82.089	81.634	80.725	80.270
9	407.709	82.996	82.087	81.633	80.724	80.269
10	407.707	82.996	82.087	81.632	80.723	80.269
11	407.706	82.996	82.087	81.632	80.723	80.268
12	407.706	82.996	82.087	81.632	80.723	80.268
13	407.705	82.996	82.086	81.632	80.723	80.268

To check the above result, we can use the method in Murphy et al. (1982): If $Q^* = \sum_{i \in I} q_i^* = 407.705$ is such that solving the problem

$$\begin{array}{ll} \text{maximize} & P(Q^*) \sum_{i \in I} q_i + \frac{1}{2} P'(Q^*) \sum_{i \in I} q_i^2 - \sum_{i \in I} c_i(q_i) \\ \text{subject to} & \sum_{i \in I} q_i = Q^* \text{ and } q_i \geq 0 \text{ for all } i \end{array}$$

results in the multiplier $\mu(Q^*)$, associated with the constraint $\sum_{i \in I} q_i = Q^*$, being zero, then $(q_i^*)_{i \in I}$ must be a Cournot equilibrium. In our example $\mu(407.705) = -0.0002$, which we deem sufficiently close to zero.

3.3. Price-taking behavior. Suppose all $i \in I$ behave as though none of them affect the price. Accordingly, all firms remain convinced that $p_2 \equiv 0$ in any circumstance. If so, (5) reduces to

$$p_1 \in \partial c_i(q_i), \tag{12}$$

with solution $q_i = q_i(p_1)$. Learning, therefore, proceeds with respect to the expected price $p_1 = EP(\xi, Q)$, and the two-dimensional process (9) simplifies to

$$p_1^{t+1} = (1 - \lambda_t)p_1^t + \lambda_t F_1(p_1^t, \xi^t) \tag{13}$$

where

$$F_1(p_1, \xi) := P(\xi, \sum_{i \in I} q_i(p_1)).$$

Convergence now obtains under remarkably mild conditions:

Theorem 2. (Convergence to competitive equilibrium in one market) *Suppose the averaged function*

$$f_1(p_1) := EF_1(p_1, \xi)$$

is Lipschitz continuous and has only isolated equilibria. Then (13) converges almost surely for any initial p_1^0 to a fixed point $p_1 = f_1(p_1)$ at which all price-taking firms have maximized their expected profit.

Proof. The differential equation

$$\dot{p}_1 = f_1(p_1) - p_1 \tag{14}$$

is asymptotically stable. Indeed, consider the function

$$L(\tau) := -\int_{p_1(0)}^{p_1(\tau)} [f_1(p_1) - p_1] \, dp_1.$$

It is Lyapunov because $\frac{d}{d\tau}L(\tau) = -[f_1(p_1) - p_1]^2 \leq 0$ with strict inequality unless p_1 is an equilibrium. Stochastic approximation theory (Benaim, 1996) now shows that (13) and (14) have the same asymptotic limits. □

4. EXTENSIONS

We conclude by abstracting from the market context. To simplify we suppress uncertainty here. This may help to see other applications of the approach used above.

Suppose player $i \in I$ obtains payoff $\pi_i(p, x_i)$ - or records marginal payoff $m_i(p, x_i)$ - that depends on a *two-dimensional parameter* $p \in \mathbb{P}$ and own choice $x_i \in \mathbb{X}_i$. The vector p is endogenous, being the result of strategic interaction. Specifically, there exists an "aggregation mapping" A that associates to each strategy profile $x := (x_i)_{i \in I}$, in the ambient space $\mathbb{X} := \Pi_{i \in I} \mathbb{X}_i$, a two-dimensional parameter $p = A(x)$, belonging to some specified subset $\mathbb{P}$ of the Euclidean plane. (Absent uncertainty, in the above Cournot example we have $x_i := q_i$, $\mathbb{X}_i := domc_i$, $m_i(p, x_i) := p_1 + p_2 x_i - c_i'(x_i)$, and $Ax := \left[P(\sum_{i \in I} x_i), P'(\sum_{i \in i} x_i)\right]$.) In general let

$$B_i(p) := \arg\max_{x_i \in \mathbb{X}_i} \pi_i(p, x_i) \tag{15}$$

denote the *best response* of individual i to parameter pair p. Alternatively, suppose his optimality condition

$$0 \in m_i(p, x_i) := \frac{\partial}{\partial x_i} \pi_i(A(x), x_i) \,\big|_{A(x)=p} \tag{16}$$

admits a solution, also called best response $x_i = B_i(p) \in \mathbb{X}_i$. Here, as above, $\frac{\partial}{\partial x_i}$ denotes the generalized (super-)differential operator of convex analysis (Rockafellar, 1970). When using (16) it is tacitly assumed that $\pi_i(x) = \pi_i(A(x), x_i)$ be concave with respect to x_i. This ensures the optimality of the solution x_i to (16). Note that we say nothing about the nature of the decision spaces $\mathbb{X}_i, i \in I$. They can be quite general, except, of course, that format (16) forces us to let $\mathbb{X}_i$ be a convex subset of some topological vector space.

Whatever the specification and the origin of the choice $B_i(p)$, whether stemming from (15) or (16), we assume it fully characterizes optimality, is well defined, and unique for every $i \in I$ and $p \in \mathbb{P}$. Also, we tacitly posit, as a behavioral assumption, that each individual i, in producing his best response $B_i(p)$, regards p as exogenous and given. More precisely, he acts at any stage as though he has a firm and seemingly "rational" belief about p.

A natural concern is then with the learning of justifiable beliefs during repeated play. Reflecting that concern, $p \in \mathbb{P}$ should be declared an *steady state* if and only if it satisfies the *fixed point condition*

$$p = A \circ B(p) =: f(p)$$

where $B := (B_i)_{i \in I}$ is the entire vector of best responses. The reason for naming the point p a steady state is that it generates a *Nash noncooperative equilibrium* $x \in B(p)$ in so far as the best response x to the belief p reproduces that same belief. In other words, a steady state embodies both confirmed expectations and optimal responses, these two properties being the defining features of Nash equilibrium.

When focus is on steady states and possible approaches to such distinguished outcomes, Theorem 1 may sometimes be directly applicable.

REFERENCES

[1] Benaim M., (1996). A dynamical approach to stochastic approximations, *SIAM Journal of Control and Optimization* 34(2) 437-472.

[2] Benveniste, A., M. Métivier, and P. Pirouret, (1990). *Stochastic Approximations and Adaptive Algorithms*, Springer-Verlag, Berlin.

[3] Bray, M. (1982). Learning, estimation, and the stability of rational expectations, *Journal of Economic Theory* 26, 318-339.

[4] Conlisk, J., (1996). Why bounded rationality? *Journal of Economic Literature*, 34, 669-700.

[5] Corchon, L. C. and A. Mas-Colell, (1996). A note on the stability of best reply and gradient dynamics with applications to imperfectly competitive models, *Economic Letters* 51, 59-65.

[6] Evans, G. W. and S. Honkapohja, (1994). Learning, convergence, and stability with multiple rational expectations equilibria, *European Economic Review*, 38, 1071-1098.

[7] Evans, G. W. and S. Honkapohja, (1997). Learning dynamics, in J.B. Taylor and M. Woodford (eds.) *Handbook of macroeconomics,* Elsevier, Amsterdam.

[8] Flåm, S. D., (1998a). Averaged predictions and the learning of equilibrium play, *Journal of Economic Dynamics and Control*, 22, 833-848.

[9] Flåm, S. D., (1998b). $2x2$ Games, Fictitious Play and Green's Theorem, *Proceedings of the IV Catalan Days of Applied Mathematics,* Universitat Rovira i Virgili, Tarragona, Spain 89-101.

[10] Foster, D. P. and R. V. Vohra, (1997). Calibrated learning and correlated equilibrium, *Games and Economic Behavior*, 21, 40-55.

[11] Fudenberg, D. and D. K. Levine, (1998). *The theory of learning in games*, The MIT Press, Cambridge.

[12] Kalai, E. and E. Lehrer, (1995). Subjective games and equilibria, *Games and Economic Behavior*, 8, 123-163.

[13] Kaniovski, Y. M. and H. P. Young, (1995). Learning dynamics in games with stochastic perturbations, *Games and Economic Behavior* 11, 330-363.

[14] Katok, A. and B. Hasselblatt, (1995). *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press.

[15] Keenan, D. and T. Rader, (1985). Market dynamics and the law of demand, *Econometrica* 53(2) 465-471.

[16] Marcet, A. and T. J. Sargent, (1989). Convergence of least squares learning mechanisms in self-referential linear stochastic models, *Journal of Economic Theory*, 48, 337-368.

[17] Murphy, F. H., H. D. Sherali, and A. L. Soyster, (1982). A mathematical programming approach for determining oligopolistic market equilibrium, *Mathematical Programming* 24, 92-106.

[18] Nemytskii, V. V. and V. V. Stepanov, (1960). *Qualitative Theory of Differential Equations*, Princeton University Press, Princeton.

[19] Rockafellar, R. T., (1970). *Convex analysis*, Princeton University Press, Princeton.

[20] Sargent, T. J.,(1993). *Bounded rationality in macroeconomics*, Oxford University Press, Oxford.

[21] Simon, H. (1955). A behavioral model of rational choice, *Quarterly Journal of Economics*, 279, 99-118.

[22] Thorlund-Petersen, L. (1990). Iterative computation of Cournot equilibrium, *Games and Economic Behavior* 2, 61-75.

[23] Weibull, J. W., (1995). *Evolutionary game theory*, The MIT Press, Cambridge.

[24] Young, H. P., (1998). *Individual strategy and social structure: An evolutionary theory of institutions*, Princeton University Press, Princeton.