Non-Nash Probabilistic Equilibria: Correlated and Self-Confirming Equilibrium

P. Battigalli
Bocconi University
Game Theory: Analysis of Strategic Thinking

October 6, 2023

Abstract

Nash equilibrium (pure or mixed) has been justified as characterizing (1) self-enforcing non-binding agreements, or (2) steady states of learning dynamics in recurrent strategic interactions. Yet both justifications yield weaker equilibrium concepts: (1, *correlated equilibrium*) self-enforcing non-binding agreements can be stochastic and yield "spurious" correlation between the actions of different players; (2, *self-confirming equilibrium*) in steady states of learning dynamics players best respond to conjectures that may be incorrect, but are nonetheless consistent with evidence.

[These slides summarize and complement parts of Section 5.4 of Ch. 5 (on pure Nash equilibrium) and Sections 6.2 and 6.3 of Ch. 6 (on probabilistic equilibria) of GT-AST]

Introduction: Justifications of Nash Equilibrium

- The Nash equilibrium concept has been justified along different lines. None of these justifications holds water *in general*:
  - NE is the *solution of a multi-person decision problem obtained by strategic reasoning*. Yet, if strategic reasoning is modeled by Rationality and Common Belief in Rationality, this works only in games with a unique rationalizable outcome.
  - NE is a *necessary condition for a non-binding agreement to be self-enforcing*. Yet, we show below that such self-enforcing agreements may be stochastic and feature ("spurious") correlation between the choices of different players.
  - NE represents the limit steady states of learning processes. Given sufficient observability, within a population scenario, this yields the *mixed* NE concept. Yet, *observability may be imperfect*, and we show below that the learning story yields a weaker solution concept (pure or mixed) *self-confirming equilibrium*.

Example: Correlation in Battle of the Sexes

$\mathbf{1}\backslash 2$	B_2	S_2
B_1	$\mathbf{3}$, *1*	$\mathbf{0}$, *0*
S_1	$\mathbf{0}$, *0*	$\mathbf{1}$, *3*

weather	bad_2	sunny_2
bad_1	$\frac{1}{2}$	0
sunny_1	0	$\frac{1}{2}$

- **Rowena** (**pl. 1**) and *Colin* (*pl. 2*) want to agree in advance on how to play the BoS. Binding agreements are not feasible, the agreement has to be **self-enforcing**: *each player must want to comply assuming that the other complies.*
  - Each NE (B_1, B_2) and (S_1, S_2) works as a self-enforcing non-binding agreement, but each one is unfair: (B_1, B_2) favors Rowena, (S_1, S_2) favors Colin.
  - Smart idea to implement *fairness in expectation*: make the action depend on weather, an extraneous, payoff-irrelevant random variable, *go to* the (covered) **S***tadium if* **s***unny, to the* **B***allet if* **b***ad weather.*
  - Each player observes the same weather, bad and sunny are equally likely. *The weather works as a coordination device.* Each player has an *interim* (post observation) *incentive to comply.*

Example: Correlation in a Dove-Hawk game

$\mathbf{1}\backslash 2$	D_2	H_2
D_1	$\mathbf{6}$, 6	$\mathbf{2}$, 7
H_1	$\mathbf{7}$, 2	$\mathbf{0}$, 0

rand.var	d_2	h_2
d_1	$x : \frac{1}{3}$	$z : \frac{1}{3}$
h_1	$y : \frac{1}{3}$	

- **Rowena** (**pl. 1**) and *Colin* (*pl. 2*) can play in a "dovish" (non-aggressive) or "hawkish" (aggressive) way.
  - The agreement to play (D_1, D_2) would be fair with high payoffs, but it is not self-enforcing (not incentive compatible).
  - The asymmetric NEs (H_1, D_2) and (D_1, H_2) are—of course—self-enforcing, but do not attain a high total payoff.
  - *Smart idea*: Row and Col imperfectly and asymmetrically observe a random variable with equally likely realizations x, y, z (see table). They agree on choosing H_i if h_i and D_i if d_i.
  - This works! Given $\mathrm{h}_1 = \{y\}$ Row is certain of D_2 and best responds with H_1, given $\mathrm{d}_1 = \{x, z\}$ she deems D_2 and H_2 equally likely, thus $EU(D_1|\mathrm{d}_1) = \frac{1}{2}6 + \frac{1}{2}2 = 4 > 3.5 = \frac{1}{2}7 + \frac{1}{2}0 = EU(H_1|\mathrm{d}_1)$ (similarly for Col). **Note:** *Not a convex combinations of NEs!*

Correlated Equilibrium (Hints)

- In the previous examples, each player i observes the realization t_i (type) of an extraneous, payoff-irrelevant random variable $\tau_i : \Omega \to T_i$ defined on a probability space (Ω, p). The probability of observing any t_i is

$$p\left(\tau_i^{-1}(t_i)\right) = p\left(\{\omega : \tau_i(\omega) = t_i\}\right).$$

  - A **probabilistic self-enforcing agreement**, or **correlated equilibrium (CE)** specifies a strategy (decision function) $\sigma_i : T_i \to A_i$ for each i so that the following incentive constraints hold: for all t_i and a_i,

$$p\left(\tau_i^{-1}(t_i)\right) > 0 \Rightarrow \mathbb{E}_{p,\sigma_{-i}}\left(u_i\left(\sigma_i(t_i), \cdot\right) | t_i\right) \geq \mathbb{E}_{p,\sigma_{-i}}\left(u_i\left(a_i, \cdot\right) | t_i\right),$$

w/ $\mathbb{E}_{p,\sigma_{-i}}\left(u_i\left(a_i, \cdot\right) | t_i\right) = \sum_{\omega \in \tau_i^{-1}(t_i)} u_i\left(a_i, \sigma_{-i}\left(\tau_{-i}(\omega)\right)\right) \frac{p(\omega)}{p\left(\tau_i^{-1}(t_i)\right)}$.

  - **Note:** The correlation among actions in a CE is **"spurious"**.

Steady States: Wrong Choice of a Safe Action

$\mathbf{1}\backslash 2$	ℓ	r
t	**2**, *0*	**2**, *1*
b	**0**, *0*	**3**, *1*

- Row *knows* her payoff function u_1, but *ignores* that of Col, u_2 (alternatively, she is not sure that Col is rational). They play (infinitely) many times. After each play they get *feedback*: Row just *observes her realized payoff*.
  - If Row is confident that Col plays r, she plays b, observes **3**, infers that Col played r, becomes even more confident of r, plays b again, and so on. Pair (b, r), the unique NE, obtains also in the limit, and Row assigns probability 1 to r in the long run.
  - If Row is sufficiently afraid that Col plays ℓ, she plays safe action t, observes $u_1 = \mathbf{2}$ independently of a_2; thus, she cannot infer anything and keeps playing t *keeping the same incorrect belief*.

Self-Confirming Equilibrium

- It has been argued that NE (maybe mixed NE, within a population-game scenario) is the necessary result of learning when the same game is played (infinitely) many times, or at least that steady states of learning dynamics must be NEs (if learning does not necessarily converge). The previous example suggests that this is not the case.
- To *characterize the steady states of learning dynamics*, we must first represent *information feedback*:
    - each i observes *ex post* a "message" $m_i = f_i(a_i, a_{-i})$, where $f_i : A_i \times A_{-i} \to M_i$ is i's **feedback function** (e.g., $f_i = u_i$);
    - let $f_{i,a_i} = f_i(a_i, \cdot) : A_{-i} \to M_i$ denote the **section** of f_i at a_i; if i plays a_i and observes m_i then i infers that the coplayers' action profile must be in $f_{i,a_i}^{-1}(m_i) := \{a_{-i} : f_i(a_i, a_{-i}) = m_i\}$;
    - in a (pure) *steady state* $\left(a_i^*, \mu^i\right)_{i \in I}$, called **self-confirming equilibrium (SCE)**, for each i, (1, B.R.) $a_i^* \in r_i\left(\mu^i\right)$ and (2, CONF) μ^i is *confirmed*, that is, $\mu^i\left(f_{i,a_i^*}^{-1}\left(f_i\left(a^*\right)\right)\right) = 1$; hence, each i keeps the same conjecture and plays the same action time and again.

Anonymous (Mixed) Self-Confirming Equilibrium

- Consider a population-game scenario, as we did to motivate the mixed NE concept.
- In a steady state $(\alpha_i)_{i\in I}$, for each population $i \in I$, each action a_i played by a positive fraction $\alpha_i(a_i) > 0$ of agents must be B.R. to some conjecture $\mu^i_{a_i}$ that agrees with the long-run frequency of messages induced by α_{-i} given a_i:
  - For any a_i, μ^i, α_{-i}, m_i the **predicted long-run frequency** of m_i is

$$\mathbb{P}^f_{a_i,\mu^i}(m_i) := \mu^i\left(f^{-1}_{i,a_i}(m_i)\right) = \sum_{a_{-i}:f_i(a_i,a_{-i})=m_i} \mu^i(a_{-i}),$$

    the **actual long-run frequency** of m_i is

$$\mathbb{P}^f_{a_i,\alpha_{-i}}(m_i) := \alpha_{-i}\left(f^{-1}_{i,a_i}(m_i)\right) = \sum_{a_{-i}:f_i(a_i,a_{-i})=m_i} \prod_{j\neq i} \alpha_j(a_j).$$

  - An **anonymous SCE** is a profile $\left(\alpha_i, \left(\mu^i_{a_i}\right)_{a_i\in\mathrm{supp}\alpha_i}\right)_{i\in I}$ s.t. $\forall i, \forall a_i \in \mathrm{supp}\alpha_i$, (1, B.R.) $a_i \in r_i\left(\mu^i_{a_i}\right)$, (2, CONF) $\mathbb{P}^f_{a_i,\mu^i_{a_i}} = \mathbb{P}^f_{a_i,\alpha_{-i}}$.

Properties of Feedback: Observable Payoffs

- The feedback function f_i satisfies "observable payoffs" if, given each action a_i, each possible message reveals the realized payoff.
- Formally, f_i satisfies **observable payoffs** if, for all $a_i \in A_i$, $a'_{-i}, a''_{-i} \in A_{-i}$,

$$f_i\left(a_i, a'_{-i}\right) = f_i\left(a_i, a''_{-i}\right) \Rightarrow u_i\left(a_i, a'_{-i}\right) = u_i\left(a_i, a''_{-i}\right),$$

  that is, for each a_i, section $u_{i,a_i} : A_{-i} \to \mathbb{R}$ is *constant* on each subset $f_{i,a_i}^{-1}(m_i)$ $(m_i \in M_i)$.
- Special cases in which the property holds trivially:
  - the feedback is the realized payoff: $M_i \subseteq \mathbb{R}$, $f_i = u_i$;
  - the feedback is the action profile just played (**perfect feedback**): $M_i = A_i \times A_{-i}$, $f_i = \mathrm{Id}_{A_i \times A_{-i}}$, $f_i(a_i, a_{-i}) = (a_i, a_{-i})$.

Properties of Feedback: Own-Action Independence

- For each a_i, define the **ex post information partition** of A_{-i} given a_i: $f_{i,a_i}(A_{-i}) \subseteq M_i$ is the set of messages that i can get, each $m_i \in f_{i,a_i}(A_{-i})$ reveals that the coplayers' action profile belongs to $f_{i,a_i}^{-1}(m_i)$; the collection of such subsets is the partition $\mathcal{F}_{-i}(a_i)$ induced by a_i, that is,

$$\mathcal{F}_{-i}(a_i) := \left\{ C_{-i} \in 2^{A_{-i}} : \exists m_i \in f_{i,a_i}(A_{-i}), C_{-i} = f_{i,a_i}^{-1}(m_i) \right\}.$$

- Feedback function f_i satisfies **own-action independence of feedback about others (OAI)** if $\mathcal{F}_{-i}(\cdot)$ is *"essentially constant"*, that is, for all *justifiable* actions a_i', a_i'',

$$\mathcal{F}_{-i}(a_i') = \mathcal{F}_{-i}(a_i'').$$

Example: Cournot, observed price, independence

- Set of outputs $A_i = \{b, \ell, h\} = \{0, 1, 2\}$,
  $P(q_1 + q_2 + q_3) = (6 - \sum_i q_i)$, $MC = AC = 2$,
  $f_i(q_1, q_2, q_3) = P(q_1 + q_2 + q_3)$.
- Then, $\mathcal{F}_{-i}(q_i)$ is *constant* (we show the profit matrix for each output q_1, red ellipses form the partition):

h	0	0	0
ℓ	0	0	0
b	0	0	0
$q_1{=}b$	b	ℓ	h

h	1	0	-1
ℓ	2	1	0
b	3	2	1
$q_1{=}\ell$	b	ℓ	h

h	0	-2	-4
ℓ	2	0	-2
b	4	2	0
$q_1{=}h$	b	ℓ	h

Example: Cournot, observed profit, lack of independence

- $A_i = \{b, \ell, h\} = \{0, 1, 2\}$, $P(q_1 + q_2 + q_3) = (6 - \sum_i q_i)$, $MC = AC = 2$, $f_i(q_1, q_2, q_3) = \pi_i(q_1, q_2, q_3) = \left(4 - \sum_j q_j\right) q_i$.
- Then, $\mathcal{F}_{-i}(q_i)$ is *not constant* (we show the profit matrix for each output q_1, red ellipses form the partition):

h	0	0	0
ℓ	0	0	0
b	0	0	0
$q_1 = b$	b	ℓ	h

h	1	0	-1
ℓ	2	1	0
b	3	2	1
$q_1 = \ell$	b	ℓ	h

h	0	-2	-4
ℓ	2	0	-2
b	4	2	0
$q_1 = h$	b	ℓ	h

Sufficient Conditions for the Equivalence of SCE and NE

Lemma

Suppose that f_i satisfies observable payoffs and own-action independence of feedback about others. Then best replies to confirmed conjectures are also best replies to correct conjectures: for all a_i and α_{-i}, if there is μ^i such that (1) $a_i \in r_i\left(\mu^i\right)$ and (2) $\mathbb{P}^f_{a_i,\mu^i} = \mathbb{P}^f_{a_i,\alpha_{-i}}$, then $a_i \in r_i\left(\alpha_{-i}\right)$.

Theorem

If, for each $i \in I$, f_i satisfies observable payoffs and own-action independence of feedback about others, then every pure or mixed action profile $\left(\alpha_i^*\right)_{i\in I}$ is part of some (anonymous) SCE if and only if $\left(\alpha_i^*\right)_{i\in I}$ is a (mixed) NE.

SCE in Cournot Oligopoly

- Each firm knows its cost function $C_i(\cdot)$ and the inverse demand function $P(\cdot)$. Let $Q_{-i} = \sum_{j \neq i} q_j$ (competitors' total output).
  - If *feedback is realized market price*, viz. $f_i\left((q_j)_{j\in I}\right) = P\left(\sum_{j\in I} q_j\right)$, then observable payoffs and own-action independence (OAI) hold: given q_i and $p = P(q_i + Q_{-i})$ firm i can find
  $$Q_{-i} = P^{-1}\left(P\left(q_i + Q_{-i}\right)\right) - q_i.$$
  Thus, an output profile (q_i^*) is part of an SCE if and only if it is a (the) Cournot-Nash equilibrium.
  - If *feedback is realized profit*, viz. $f_i\left((q_j)_{j\in I}\right) = \pi_i\left((q_j)_{j\in I}\right)$, then OAI does not hold: $f_i(0, q_{-i}) = C_i(0)$, and
  $$q_i > 0 \Rightarrow Q_{-i} = P^{-1}\left(\frac{\pi_i\left((q_j)_{j\in I}\right) + C_i(q_i)}{q_i}\right) - q_i.$$
  The following are SCEs: $\left((q_j^*)_{j\in J}, \mathbf{0}_{I\setminus J}\right)$ with $(q_j^*)_{j\in J}$ NE of restricted game *with only* firms in J.

References

- Battigalli, P., E. Catonini, and N. De Vito (2023): *Game Theory: Analysis of Strategic Thinking*. Typescript, Bocconi University.
- Battigalli, P. (2023): *Mathematical Language and Game Theory*. Typescript, Bocconi University.
- Battigalli, P., and G. Lanzani (2023): *A Note on Self-Confirming Equilibrium and Stochastic Control*. Typescript, Bocconi University. (**Optional**)