Learning and Equilibrium in Games

Drew Fudenberg
Marshall Lectures
Cambridge University
October 2009

Game Theory: The study of interdependent decisions where an agent's payoff can depend on the decisions of other agents.

Key: best course of action can depend on what others do.

Most applications of game theory use equilibrium analysis: They identify the outcome of the game either with the set of Nash equilibria or with some subset of the Nash equilibria.

But when and why should we expect observed play to approximate Nash or any other sort of equilibrium?

Learning in Games: Equilibria are the long-run outcome of non-equilibrium learning or other forms of adaptation.

This theory helps explain how and when Nash equilibrium arise in static simultaneous-move games.

It can also suggest which Nash equilibria are more likely.

Learning theory suggests that other equilibrium concepts are more appropriate for dynamic games.

OUTLINE

I. Review of Game Theory and Nash Equilibrium

II. Learning in Static Games, with Digression on Evolutionary Game Theory

III. Strategies and Nash Equilibrium in Extensive-Form Games

IV. Learning in Extensive-Form Games and Self-Confirming Equilibrium

V. Experimentation and Subgame-confirmed Equilibrium

I. Introduction/Review of Game Theory

Informal game theory Hume [1739], Rousseau [1775], etc.

First formal analysis Cournot [1838] duopoly model:

- Only two firms, so they can't neglect effect of own output on price.
- Firms know how price varies with total output
- But they need to forecast the output of the other firm

Cournot's model:
Firms 1 and 2 simultaneously set output levels.
Price adjusts to clear the market.

A Cournot equilibrium is a pair of outputs such that each firm's output maximizes its profit given the output of the other firm.

This is an "equilibrium" in the sense that neither firm has an incentive to change.

Cournot also proposed a dynamic adjustment process whose steady states are the equilibria he defined.

-Important precursor of the learning-theoretic interpretation of equilibrium.

-But perhaps not so realistic.

Following Cournot a number of other games were analyzed as special cases: Bertrand, Edgeworth, Hotelling etc.

Their equilibrium concepts were generalized by Nash [1950].

Before Nash: von Neumann [1928] on the minmax theorem.

von Neumann and Morgenstern [1944]

The Theory of Games and Economic Behavior:

> "*We hope to establish…that the typical problems of economic behavior become strictly identical with the mathematical notions of suitable games of strategy*."

This book gave the first general definition of a "game."

It also introduced the general idea of a "strategy" as a contingent plan, as opposed to the uncontingent choice of output in the Cournot duopoly models- we will come back to this when we talk about extensive-form games.

Hugely important and useful book-played a key role in the development of game theory and mathematical social science. But it has a very different interpretation of game theory, and thus of equilibrium, than in Cournot or Nash.

Von Neumann and Morgenstern saw game theory's task as providing

> "*mathematically complete principles which define 'rational behavior' in a social economy*."

They also wanted their principles to have a unique, unambiguous prediction.

In two-player constant-sum games the minmax solution does what they were looking for.

But most of the games economists care about are not constant sum.

Problems with the Von Neumann-Morgenstern program:

- In general, a unique prediction is too much to ask for.
- The assumption of rational behavior does very little to restrict the possible outcomes.
- Even “common knowledge of rationality” isn’t enough to do what vN-M wanted.

So we need some other way of predicting what will happen in games.

Nash [1950] defined what we now call "Nash equilibrium," which generalizes Cournot equilibrium (and the subsequent equilibrium definitions of Bertrand, etc.)

Think of “strategies” for now as buttons on a computer terminal- players push buttons simultaneously, not seeing their opponent’s choice.

A *strategy profile* is simply a specification of a strategy for each player or player role- e.g. player 1 plays “U” and player 2 plays “L”. (there may be many players in a given player role.)

A *game in strategic form* is a collection of sets of feasible strategies- one feasible set per player- and a *payoff function* that specifies the payoffs of each player for each strategy profile.

Definition: A strategy profile is a *Nash equilibrium* if each player's strategy maximizes that player's payoff given the strategies being used by the other players. That is, no player can get a higher payoff by changing his strategy, holding fixed the strategies of the other players.

Nash Equilibria are "consistent predictions of play":
It is not consistent for all players to predict the <u>same</u> non-Nash outcome, for they should all know that someone has an incentive to change.

This consistency explains why Nash equilibria might persist *if* for some reason players started out predicting that it would, but it doesn't say why they should predict it in the first place.

	L	R
U	0, 0	5, 5
D	5, 5	0, 0

• Two “pure-strategy” Nash equilibria: (U,R) and (D,L).
(Also a “mixed-strategy” equilibrium: if half the player 1’s play U and half play D then the player 2’s are indifferent so half can play L and half play R.)

• If player 1 expects (U,R)- so plays U- and player 2 expects (D,L)- and plays L- the result is (U,L) which is <u>not</u> a Nash equilibrium.

Can’t mix-and-match parts of an equilibrium profile.

• When there are multiple Nash equilibria, the outcome will only correspond to *any* of the equilibria is if all of the players coordinate on the *same* equilibrium.

- This coordination is not an implication of rationality, so rationality does not imply equilibrium play.

- Once the results are tabulated we will play this game again- with a different "protocol" : each agent matched against the entire distribution of other side's play, with payoffs 10p per game.

- Nash understood equilibrium in the sense of Cournot, as the long-run outcome of an adjustment process.

- Work on "learning in games" investigates this idea.

- Implications are simplest in *static Games*: one shot, simultaneous move, like the coordination example.

II. Learning and Equilibrium in Static Games

Self-interested players (as opposed to players who cooperate to find an equilibrium.)

No point in explaining equilibrium in a given game by assuming an equilibrium of some larger "adjustment game."

Thus non-equilibrium learning models need to allow for players whose adjustment rules aren't a best response to the adjustment rules of the others.

In this sense some player may be "making mistakes," so we can't use "no mistakes" as a criterion for selecting learning rules.

Instead ask that play be “plausible,” which includes “no really obvious mistakes.”

Various adjustment processes have been analyzed, with varying degrees of rationality on the part of the agents.

Focus here on “**Belief-based” models**:
Agents form expectations about the play of their opponents and use beliefs to choose actions, as in **fictitious play”** (Brown [1951]), and **“smooth fictitious play”** (Fudenberg-Kreps [1993]).

The details of these models are less important than three general properties:

***Passive learning*, *strategic myopia*, and *asymptotic empiricism*.**

Passive learning: Agent's actions have no impact on their observations.

- In particular, assume that each time the game is played, agents see the <u>strategy</u> played by their opponent.

- May see strategies of other players as well but this is independent of own play.

- So no reason to "experiment" to gain information.

Strategic myopia: Agents play repeatedly without trying to influence the future play of their opponents- they learn to play the "stage game" as opposed to playing a "repeated game."

Motivation: "Large populations" with many agents in each player role, and anonymous random matching.

Two versions of large populations

a) anonymous random matching:

Each period all agents are matched to play the game, and are told only play in their own match. Used in most experiments.

b) **aggregate statistic model:**

Each period all agents play. At the end of each period, agents learn the aggregate distribution of play, and each agent's payoff is determined by this distribution. Easier to implement in ……

Large-Population Interpretation of Mixed Equilibrium: each agent uses a pure action, but different agents play differently.

Asymptotic empiricism: Beliefs converge to the time average of opponents' play. (*so priors are eventually swamped by data.)*

Motivation: Agents believe they face a stationary, but unknown, distribution of opponents strategies. Then asymptotic empiricism is a consequence of Bayesian updating. (*if priors are "non-doctrinaire")*

Note: If all agents are strategically myopic and asymptotically empirical, the overall distribution of play <u>won't</u> be stationary unless initial play is a Nash equilibrium.

So acting as if the environment is stationary is a mistake.

But stationarity is a reasonable first hypothesis in many situations.

People do seem likely to reject the stationarity hypothesis given sufficient evidence to the contrary, e.g. if the system cycles rapidly.

For example if opponent/other side’ play has been
H,T,H,T,H,T,H,T,
you might expect the next plays is very likely to be H,
as opposed to 50-50.

(Such fast cycling can occur with fictitious play but not with smooth fictitious play)

Conversely if the system converges to a steady state then agents might maintain the assumption of stationarity- as near a stable steady state the system is “almost stationary.”

Result: With asymptotic empiricism and strategic myopia, *if* the population converges to repeatedly playing the same strategy profile, that strategy profile must be a Nash equilibrium.

Reason: If the aggregate play in the population is constant, beliefs come to look like the distribution of play.
So if play isn't a Nash equilibrium, some player(s) would change their strategy.

Notes:

- Asymptotic empiricism and strategic myopia don't require that the players know game theory.

- Play can converge to Nash equilibrium even if the agents have never heard of Nash equilibrium. This is analogous to the fact that consumers don't have to take economics classes for the market outcome to approximate a competitive equilibrium.

- People may never play the <u>exact</u> same game very often, but they may also extrapolate between games and learn from the experiences of others.

- History and culture can help coordinate expectations.

- Schelling's "label salience" (see *The Strategy of Conflict,* 1960) is an illustration of this.

Consider a single population playing a symmetric coordination game- everyone choose between the same two strategies

$$\begin{array}{cc} 1,1 & 0,0 \\ 0,0 & 1,1 \end{array}$$

Note that the names of the strategies aren't displayed- they don't matter for the set of Nash equilibria…

Holmes and Watson on a train line
"Heads" or "Tails"
Meeting for lunch.

Conclusion: Shared history of play and/or shared culture can coordinate expectations and lead to Nash equilibrium play in static games.

The idea of culture, and playing based on what one knows about the other agent, seem to rely on agents not only thinking but having a theory of what others are thinking, a "theory of mind."

This isn't necessary for belief-based learning.
But asymptotic empiricism does require that agents have a memory and keep track of opponents' past play.

However, players don’t have to be rational or have even have a memory for play to end up at a Nash equilibrium, as shown by work in evolutionary game theory.

Digression in honor of Charles Darwin…

Evolutionary Game Theory

Payoffs correspond to reproductive fitness

Players are genetically programmed to play various actions (or express various phenotypes).

Fitter strategies/phenotypes have more offspring so their share of the population increases.

Fitness is not absolute, but is “frequency-dependent”:
It can depend on the actions or phenotypes of others.
In this case “survival of the fittest” is inherently game-theoretic-
In the coordination game we played neither U nor D is “fitter” in an absolute sense.

Replicator Dynamic

- Standard evolutionary adjustment process.
- Single continuum population
- Deterministic adjustment in continuous time.

(variants include adjustment in a stochastic environment as in Fudenberg-Harris [1992], discrete time as in Dekel-Scotchmer [1992], finite-population models, multiple populations...)

- State of the system = fractions of the population using each action.

- Reproduction rate of each individual= payoff in the game.

- So total number of offspring of A-players at time *t* is
(Mass of agents playing A) *(Payoff to A at time t).

- Can show that the fraction of the population playing A grows only if its current payoff is higher than the average.

- At a steady state every action that has positive population share must be equally fit- as otherwise the share of the fitter ones would grow.

- So a steady state where all actions are played must be a Nash equilibrium.

- There can be other sorts of steady states that are not Nash equilibria, as non-existent strategies have no offspring and there is no mechanism (in this stark model) to reintroduce strategies that are "extinct" and.

- Non-Nash states are not locally stable: if mutation introduced a small number of agents playing strategy that has a higher payoff (given the population distribution) then the share of these mutants would grow so the state would move away from this non-Nash point.

This motivates the idea of an “evolutionarily stable strategy,” or ***ESS***:

Suppose that the population is originally at some profile σ, and then a small share of "mutants" start playing some other strategy σ'.

ESS asks that the existing population gets a higher payoff against the resulting mixture than the mutants do.

More formally, σ is an ESS if

σ is a NE: there is no other strategy σ' with
$u(\sigma',\sigma) > u(\sigma,\sigma)$

<u>and</u>

If $u(\sigma',\sigma) = u(\sigma,\sigma)$
(so σ' is an *alternate best reply* to σ)
then $u(\sigma',\sigma') < u(\sigma,\sigma')$.

Any <u>strict</u> Nash equilibrium (where the equilibrium strategy yields a strictly higher strategy than any alternative) is an ESS, and any strict Nash equilibrium is locally stable under the replicator dynamic.

And for typical ("*generic*") static games, all pure-strategy equilibria are strict. For these games the pure-strategy ESS are the same as the pure-strategy Nash equilibria.

ESS has proved useful in explaining a wide range of biological phenomena. Examples include mutualisms and parasites in games between species, animal behaviors such as territorialism, and *genomic imprinting*.

Humans and most mammals inherit two copies of most genes, one from the mother and one from the father.

For most genes both copies are functional, but for “imprinted” genes production of the corresponding protein sometimes occurs from only one of the two alleles, depending on which parent passed on the gene. Examples of this in insects, mammals, and plants; in mammals it seems most often associated with placental growth and development.

Explanation: (Haig [1997]) Genomic imprinting arise from the conflicting interests of the father and mother, as represented in a game where the players are the genes.

An individual has a maternal and paternal allele that disagree over the amount of some factor (e.g. growth of current embryo).

Suppose the amount X of the factor is the sum of the independent production by the maternal allele (m) and the paternal allele (p); $X = m + p$.

If X increases resources devoted to this embryo (as opposed to saved for future ones) then the maternal allele will 'favor' a smaller amount of this factor, as the mother cares (either somewhat or a lot) more than the father about her future offspring. Suppose the maternal allele favors $X = M$ and paternal allele favors $X = P > M$.

Then the ESS strategy for alleles to express P when paternally-derived and 0 when maternally-derived.

This explains why imprinting can be an all-or-none phenomenon in which one allele is silenced

And in some cases there is an “arms race” with the maternal allele triggering the release of a buffer that soaks up the factor. (Wilkins and Haig [2001]).

…back to belief-based learning

Just as with the replicator dynamic, we can study the stability properties of Nash equilibria with belief-based learning.

$1,1$	$0,0$			$1,-1$	$-1,1$
$0,0$	$1,1$			$-1,1$	$1,-1$

Intuitively, the mixed equilibrium in the coordination game on the left looks unstable.

In the game on the right ("matching pennies") stability/instability of the equilibrium is less obvious.

- Formal analysis requires more explicit dynamic model, e.g. smooth fictitious play.

- Analysis simplest with a single agent on each side, or if all agents in a population have the same beliefs. Here the mixed equilibrium is unstable in the coordination game and stable in matching pennies. (Fudenberg-Kreps [1993], Benaïm-Hirsch [1999].)

- Single agent per side isn't consistent with strategic myopia, and in large populations different agents probably receive different information – e.g. they might only observe play in own matches.

- But stability similar results in large population with heterogeneous beliefs provided the matching process is close enough to uniform. (Fudenberg-Takahashi [2009]).

- There are games where the unique Nash equilibrium is unstable under standard belief-based dynamics.

- Here it is not clear that the dynamics make sense, nor is it clear what actually happens, though there is some evidence that experimental play can indeed cycle (Benaïm, Hofbauer, Hopkins [2009]).

- There are also classes of games where play always converges to Nash equilibrium, for example "potential games" (Hofbauer-Sandholm [2002])

Consider the following game:

Each player picks an integer j between 1 and 100. The players whose choice is closest to 2/3 of the median are "winners;" the winner's payoff is \$10/(number of winners) . Losers all get 0.

Here we have a form of the aggregate-statistic model.

(survey in Nagel [1998])

- Unique Nash equilibrium: everyone chooses 1.

- This also the implication of (some forms of) “common knowledge of rationality.

Claim: If we repeated this enough times most people would choose 1. Once again the Nash equilibrium arises from learning but doesn’t describe initial play.

Conclusions so far:

- No reason to expect Nash equilibrium the first time people play a game, especially if it is unfamiliar to them.
- If play in a static game converges the long-run outcome should look like a Nash equilibrium.
- No presumption that non-equilibrium play always converges.

So far: strategies as computer keys or static choices.

Next time:
Learning and Equilibrium in Extensive Form Games.

- Extensive-form games are used to study issues of commitment, signalling, reputation, etc. that are inherently *dynamic*.

- Here strategies are complete contingent plans" that specify the action to be taken in every situation (*i.e. at every "information set")* that could arise in the course of play.

The <u>definition</u> of Nash equilibrium applies without change:

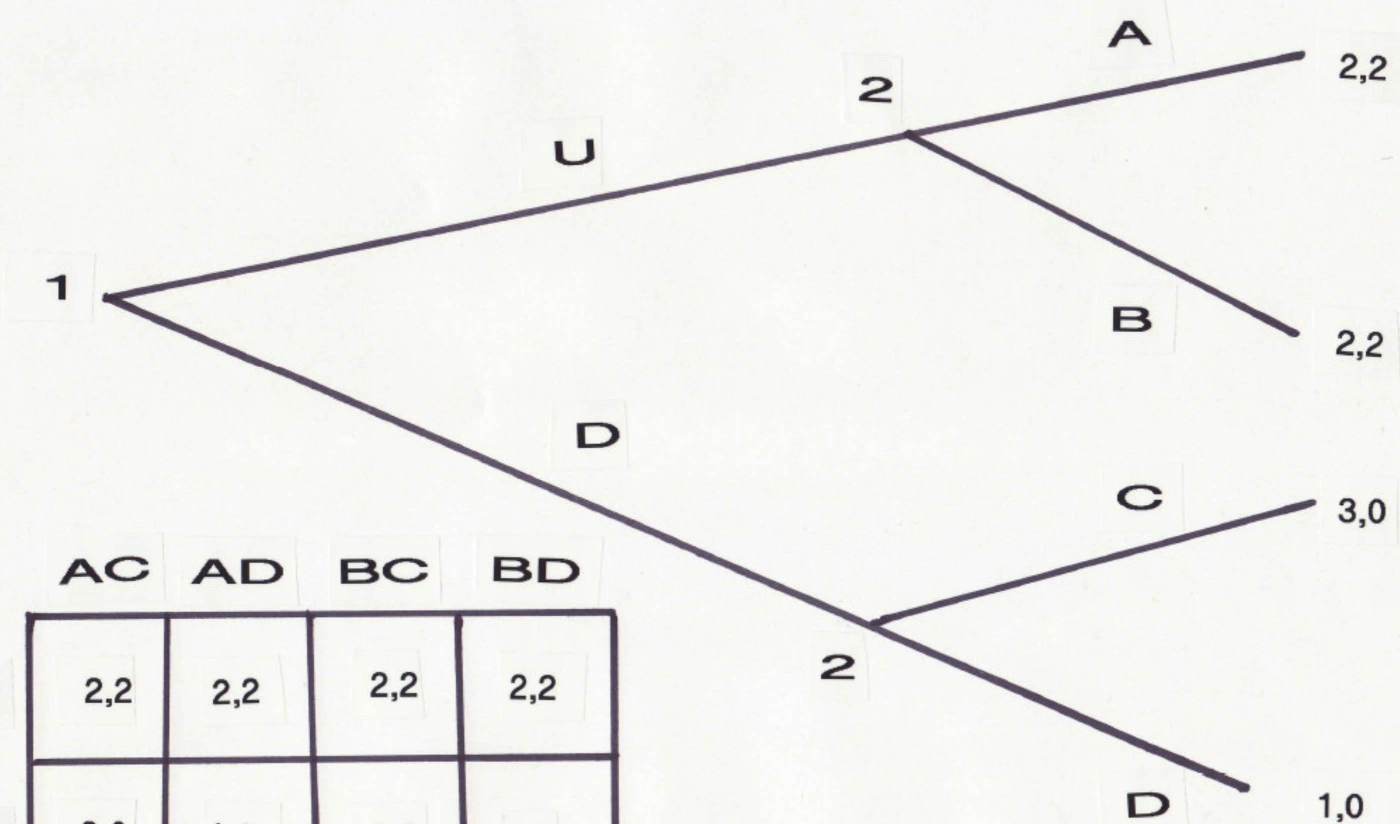
A Nash equilibrium is (still) a strategy profile such that no player can increase their payoff by changing their strategy, holding fixed the strategies (*not* the actions) of the other players.

But the implications of learning theory are different in the extensive form game than if players are pushing buttons corresponding to the strategies in the strategic form.

In particular, play can converge to non-Nash states where the players have incorrect beliefs but these beliefs are consistent with their observations.

So other equilibrium concepts are needed to capture the effect of learning…

Consider the following extensive form game:



	AC	AD	BC	BD
U	2,2	2,2	2,2	2,2
D	3,0	1,0	3,0	1,0

The long-run outcome of learning depends on what the players observe when the game is played.

So the possible long-run outcomes can be different if player 1 observes player 2's **strategy** than if player 1 only observes player 2's **action.**

In the first case, learning leads to Nash equilibrium as before.

In the second it only leads to the larger set of **self-confirming equilibrium.** (Fudenberg and Levine [1993]).

To define this concept, for a given strategy profile, say that an information set is “reached” if it has positive probability and say that it is “unreached” or “off-path” if it has probability zero.

If players don’t observe opponents’ play at unreached information sets, they may not learn it, so their actions may not be optimal given the way that opponents respond to deviations.

Self-confirming equilibrium requires that each player’s beliefs are consistent with their observations, but not necessarily correct; this is in the spirit of Hahn’s [1977] “conjectural equilibrium.”

Definition: σ is a *self-confirming equilibrium* (SCE) if for each player i and each strategy s_i with $\sigma_i(s_i) > 0$ there are beliefs $\mu_i(s_i)$ such that

(a) s_i maximizes player *i's* payoff given his beliefs $\mu_i(s_i)$,

and

(b) $\mu_i(s_i)$ is consistent with what player i see when he plays s_i. (*More formally,* $\mu_i(s_i)$ *is correct at every information set that is reached under* (s_i, σ_{-i}).)

Convergence to self-confirming equilibrium has been shown in a variety of learning models, which differ in the rationality of the agents, interaction structure, observation structure, etc.
And experimental play of extensive-form games often results in outcomes that are consistent with self-confirming equilibrium but not Nash equilibrium.

Tomorrow: Self-Confirming Equilibrium, Learning in Extensive Form Games, and the Code of Hammurabi