Learning in Games

Drew Fudenberg

Interactive Causal Learning Conference

May 2022

Review of Non-Cooperative Game Theory

- A *game* is an interactive decision problem, where each player's optimal choice—their best "strategy"—can depend on what the other players do.
- For now think of strategies as icons on a computer screen.
- So players need to predict each other's play.
- In a Nash equilibrium ("NE"), each player's strategy maximizes their payoff given the strategies being used by the other players.
- Equivalently, each player's strategy
  1. Is a best response to their beliefs about the strategies of the others, and
  2. Each player's beliefs are correct.
- Nash equilibrium generalizes the equilibrium notions of Cournot [1938] and Bertrand [1883], which were defined for particular oligopoly games.

Nash Equilibrium

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

- Two "pure-strategy" NE: (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 isn't a NE.
- When there are multiple NE, the outcome will only correspond to any equilibrium if players coordinate on the same one.
- Most applications of game theory focus on the Nash equilibria, or a subset of them selected by an "equilibrium refinement."

Equilibrium Analysis

- How should we expect play to change when e.g. some player's payoff function changes?
- *Naive approach*: Analyze change in that player's behavior assuming no other player changes their behavior.
- *Equilibrium analysis*: Solve for the new equilibrium points.
- In a decision problem an agent can't gain by lowering their payoff to some action w/o raising the payoff to any of them.

	L	R
U	$1, 3$	$4, 1$
D	$0, 2$	$3, 4$

- Unqiue NE is (U, L), player 1's payoff is 3.

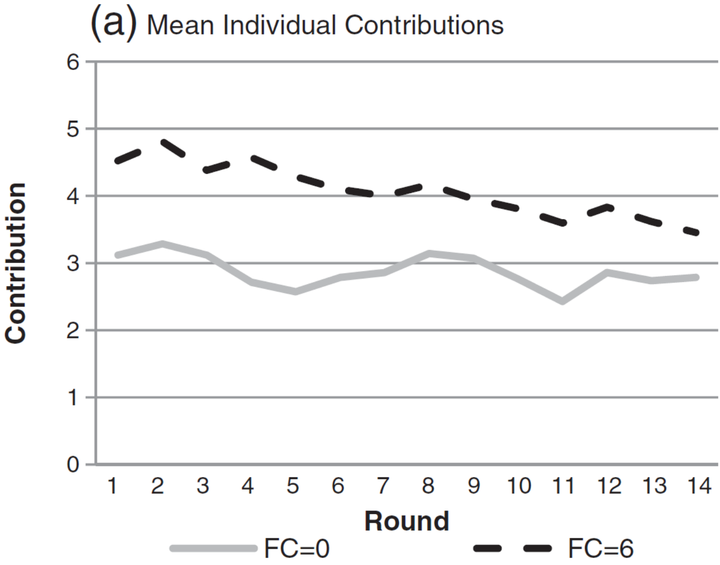
- Now lower player 1's payoff to U, resulting in the payoff matrix:

	L	R
U	$-1, 3$	$2, 1$
D	$0, 2$	$3, 4$

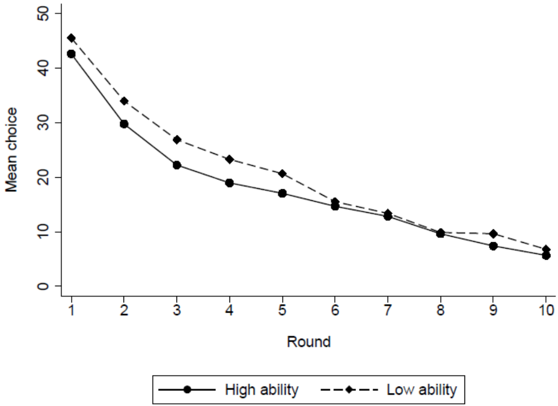
- Unique NE is now (D, R), so player 1 gets 3. The same would have happened if we simply changed the action set and made U impossible.
- Restricting player 1's choice set, or lowering her payoff to some strategy profiles can raise player 1's payoff because it can change the strategies chosen by her opponents.
- A similar observation applies to the amount/quality of information a player has: In a decision problem more information is always better, but in a game it can sometimes worse.

Learning and Equilibrium

- Rationality, even common knowledge of rationality, is neither necessary nor sufficient for Nash equilibrium
- Only sufficient in games with unique rationalizable outcome.
- With multiple NE, no reason for play to look like any of the equilibria without some explanation for why players all expect the *same* equilibrium.
- Not necessary either in theory (replicator dynamic can converge to NE) or empirically (convergence to approximate NE seen in colonies of bacteria).
- Yet equilibrium seems a decent approximation of the outcomes of some (not all!) experiments, and has been useful in empirical analyses of field data.



Bracha, Menietti, and Vesterlund [2011] on a voluntary contribution game where the Nash equilibrium is for both players to contribute 3.



Gill and Prowse [2016] on the “beauty contest.”

- The idea that equilibrium is the result of learning goes back to Cournot.

- It underlies Nash's *mass action* interpretation of equilibrium: "*participants are supposed to accumulate empirical information on the relative advantages of the various pure strategies at their disposal.*"

- To understand Nash equilibrium and related solution concepts, study the long-run behavior of non-equilibrium dynamic processes, and ask when they converge to equilibrium.

- Many sorts of adjustment processes, including biological evolution, have been said to involve "learning" in a broad sense.

- Won't have time for a broad survey.

Common themes in the learning-in-games literature

Non-equilibrium adjustment:

- Pointless to explain equilibrium in a game by assuming equilibrium in some larger adjustment game.
- So to explain how equilibrium arises, we *must allow for* players who "are making a mistake" in the sense that their behavior isn't a best response to the behavior of the others.
- The issue is *not* whether a model can generate suboptimal play, but rather whether players would notice that some other adjustment rule would be better.
- Focused on LR Play:
  - initial play determined by initial beliefs
  - easiest to obtain results for long run
  - but would like to know more about speed of learning, and implications of learning for short and medium run
- Convergence not guaranteed in general games.
- Learning can suggest equilibrium refinements.

- *Most but not all papers*: Play repeatedly without playing a "repeated game."

- Explained by reference to "large population model" with many "agents" in each "player role.

- *Large population*: unlikely to play their current opponent again for a long time, even unlikely to play anyone who played anyone who played him. So not worth sacrificing current payoff to influence this opponent's future play.

- Leading case anonymous random matching: Each period all agents are matched to play the game; they only see their own matches.

- This treatment is used in most experiments.

Fictitious Play

- Introduced by Brown [1951] as a way to compute equilibrium in two-player games. (hence “fictitious” play)
- Fudenberg and Kreps [1993] showed it also describe behavior of Bayesin agents learning from real data.
- Easy to analyze. Too simple to match experimental data, but has been used as a foundation for more complex models that fit the data better.
- *Motivation*: Agents know the strategy spaces and payoff functions but not how other people play.
- All they observe is the outcome of play in their own matches: don’t observe what happens in other matches, or opponents’ past play.
- Agents believe they are playing against a randomly drawn opponent from a large population, don’t try to influence current opponent’s future play.

- Agents act as if they are Bayesian expected utility maximizers, facing an unknown stationary distribution of opponents' strategies.

- Stationarity a reasonable first hypothesis in many situations, might expect players to stick with it when it is approximately right but to reject stationarity given sufficient evidence to the contrary—as when there is a strong time trend.

- Because what the agent observes is independent of own action, no incentive for "experimentation."

- Fictitious play has a specific functional form, which corresponds to a specific class of priors, but the functional form isn't needed for most results.

Important Facts about (2-player) FP

- If actions converge they converge to pure strategy NE.

- If time averages of empirical marginals converge the joint distribution is a NE. (*proof sketch: if player 2's marginal converges to σ_2, 1's beliefs converge to σ_2. If the mixed action σ_1 corresponding to the limit of 1's empirical marginal is not a best response to σ_2, some s_1 is strictly better than any other strategy in the support of σ_1 for all large enough times. At such times 1 must play only s_1.*)

- The above holds in any belief-based learning model that is "asymptotically empirical" (beliefs converge to empirical frequencies) and "asymptotically myopic" (eventually players choose actions that are myopic best responses to their beliefs).

Stochastic (or "Smooth") Fictitious Play

Like FP but with a smooth (continuous) "stochastic best response function" that assigns a mixed strategy response to each belief.

Advantages

- If beliefs converge behavior does too; not the case with standard fictitious play
- Allows convergence to mixed-strategy equilibria in fictitious play-like models: Actual play in FP can't converge to a mixed equilibrium.
- Avoids the discontinuity in standard fictitious play, where a small change in the data can lead to an abrupt change in behavior.

Stochastic Approximation

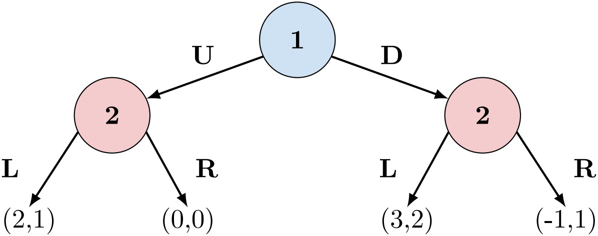
- Bayesian updating slows down at rate $1/t$, so can characterize the long run behavior of fictitious play using stochastic approximation.
- This allows us to characterize the long-run behavior of the stochastic discrete-time system by analyzing the recurrent sets of an associated deterministic continuous-time system.
- Fudenberg and Kreps [1993], Benaïm [1999], and Hofbauer and Sandholm [2002] do this for smooth fictitious play.
- Benaım, Hofbauer, and Sorin [2006] does this for exact fictitious play using results from Benaïm, Hofbauer, and Sorin [2005] on stochastic approximation with differential inclusions.

Extensive-Form Games

- Extensive-form games are used to model games with sequential and/or multiple moves.
- Here strategies are "complete contingent plans" that specify the action to be taken in every situation (at every "information set") that could arise in the course of play.
- The actions that a player uses can depend on the actions of others, but we can think of players simultaneously choosing strategies before the game is played.
- Can associate a unique strategic form with given extensive form.

Notation

- $I + 1$ players in the game, player $i = I + 1$ is Nature.
- Finite game tree with nodes $x \in X$, information sets h that partition the non-terminal nodes.
- Each h is associated with a single player, and the player who moves at h know which $x \in h$ prevails.
- A pure strategy for player i is now a *complete contingent plan*, i.e. a rule that specifies the action to play at each information set.
- S_i is the set of pure strategies for player i, $s \in S$ denotes a strategy profile for all players.
- A mixed strategy σ_i for player i is a probability distribution over S_i.
- Terminal nodes $z \in Z$; player i's payoff function u_i is a function of z.



a) extensive form

	(L,L)	(L,R)	(R,L)	(R,R)
U	$2,1$	$2,1$	$0,0$	$0,0$
D	$-1,1$	$3,2$	$-1,1$	$3,2$

b) strategic form

Nash equilibrium in Extensive-Form Games

- The **definition** of Nash equilibrium applies without change: A strategy profile such that no player can increase their payoff by changing their strategy, holding fixed the strategies of the other players.
- The learning foundation of equilibrium **does** change, as agents playing a game repeatedly need not learn the consequences of deviating from the equilibrium path.
- Exploration/exploitation trade-off: people may choose to engage in "active learning" or "experimentation."
- W/o experimentation, incorrect beliefs about off-path play could persist.
- For this reason, myopic learning only justifies the weaker concept of self-confirming equilibrium (Fudenberg and Levine [1993a], Fudenberg and Levine [1993b]).
- Self-confirming equilibrium has proved useful in many areas of economics, including economic history, experimental economics, industrial organization, and macroeconomics

More Notation

- Behavior strategies π_i specify probability distributions over feasible actions at each information set.
- Each s determines a probability distribution p over terminal nodes—what all players will see when the game is played.
- Player i's beliefs about opponents' play given by a probability measure μ_i over other players' behavior strategies.
- Note that a player 1's current belief about expected play can correspond to a correlated distribution even if the player is certain that the true distribution of his opponents' play in fact corresponds to independent randomizations.
- For example there is a difference between
  a) believing that your opponents are playing the correlated strategy $(\frac{1}{2}(A, A), \frac{1}{2}(B, B))$ and
  b) thinking they are either playing (A, A) or (B, B).
- Same marginal over first period observations, get updated differently because they imply different correlations between period 1 and period 2: either nothing to learn, or one observation is fully informative.

- Ruling out correlation and requiring that all players in the same role have the same beliefs corresponds to "unitary independent self-confirming equilibrium."

Definition

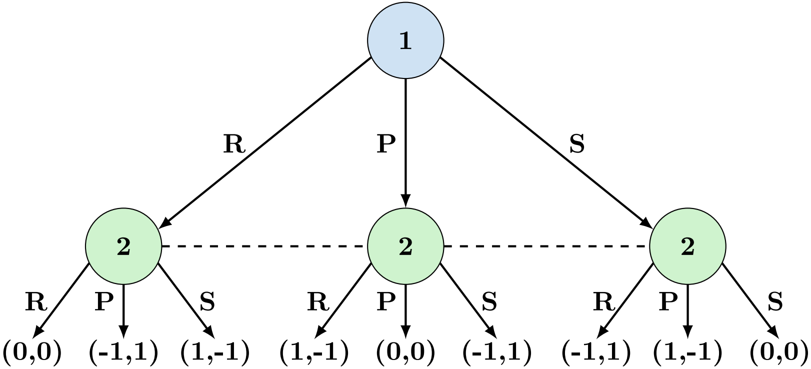
Strategy profile σ is a unitary independent self-confirming equilibrium (SCE) if for each player i there is a conjecture $\hat{\sigma}_{-i}$ about the play of the other

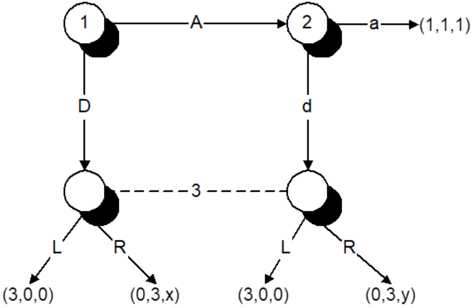
(a) s_i is a best response to $\hat{\sigma}_{-i}$, and

(b) The outcome when i plays σ_i is exactly what they expected.

- In one-shot simultaneous-move games like rock-scissors-paper, there are no off-path information sets, and the terminal node reveals the strategy profile.

- So in these games SCE is the same as NE.





- (A, a, L) is a SCE.
- But Nash equilibrium requires players 1 and 2 to make the same (correct) forecast of each player 3's play, and if both make the same forecast, at least one of the players must choose D.
- It wouldn't matter that 1 and 2 have different beliefs about 3's play at information set h if only player 1 can cause that information set to be reached.

- A game has observed deviators if whenever a deviation by some player i leads to an information set off the equilibrium path, there is no deviation by i's opponents that leads to the same information set.

- Games of perfect information satisfy this condition, as do all multistage games with observed actions.

- Also always satisfied in two-player games of perfect recall: With two players, both players must know who deviated. Not satisfied in the "horse" game.

Theorem (Fudenberg and Levine [1993a] and Kamada [2010])

In games with observed deviators, the outcome of any independent unitary self-confirming equilibrium is the outcome of a Nash equilibrium.

Now drop the restriction to unitary beliefs:

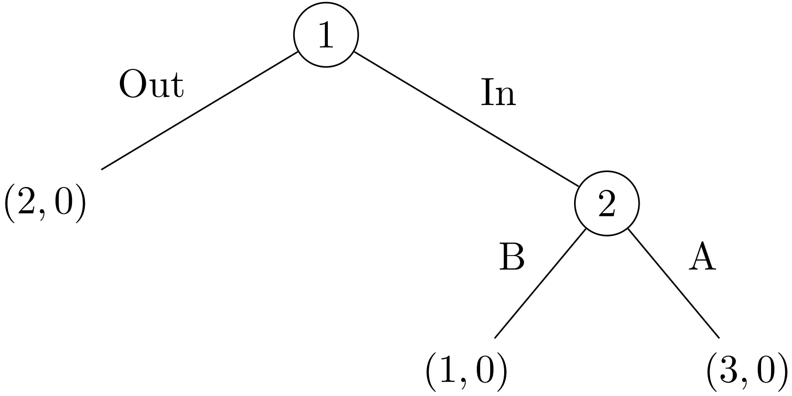
Definition

σ is an independent heterogeneous self-confirming equilibrium (SCE) if for each player i and each s_i with $\sigma_i(s_i) > 0$ there is a conjecture $\hat{\sigma}_{-i}$ such that

(a) s_i is a best response to $\hat{\sigma}_{-i}$, and
(b) The outcome when i plays σ_i is exactly what they expected..

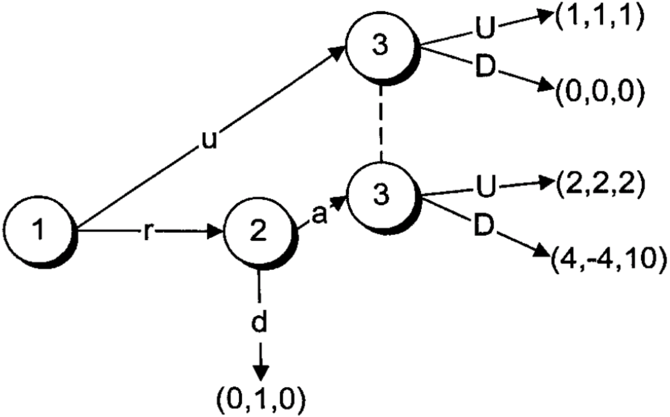
- Like unitary SCE, this reduces to Nash in one-shot simultaneous-move games.
- But allows player i to rationalize each strategy in the support of σ_i with a different belief.
- *Interpretation*: Many agents in the role of each player, and different agents in the role of player i may have observed play at different nodes. (*heterogeneity is a non-issue for Nash equilibrium*)
- Heterogeneous beliefs are important in many game theory experiments (see Fudenberg and Levine [1997]).

In this game there is no Nash equilibrium with outcome distribution $\left(\frac{1}{2}Out, \frac{1}{2}(In, A)\right)$, but it's a heterogeneous SCE.



Adding Prior Information about Payoffs

- In SCE the only constraint on beliefs is what players observe about others' play—players aren't required to use information about opponents' payoff functions.
- In some cases, players do have some prior information about their opponents' payoffs.
- In experiments, giving subjects information about other players' payoff functions can make a difference.
- This difference corresponds to the distinction between SCE and Rationalizable Self Confirming Equilibrium or RSCE (Dekel, Fudenberg, and Levine [1999]).
- RSCE is "unitary"—a single belief for all players, and all players see the same distribution on terminal nodes.
- RSCE has implications beyond the intersection of SCE and rationalizability, that come from the assumption that the outcome path is public information.



(u, U) is a Nash outcome (so self-confirming) but not RSCE: If player 1 knows 2 knows 3 is playing U, can use this knowledge and knowledge of player 2's payoffs to deduce that 2 will play a.

- RPCE (Fudenberg and Kamada [2015]) extends RSCE to cases where agents don't necessarily all see the terminal node at the end of each play: For example if player 1 stays home and player 2 and 3 go to an industry meet-up, player 1 doesn't observe what the meet-up was like.

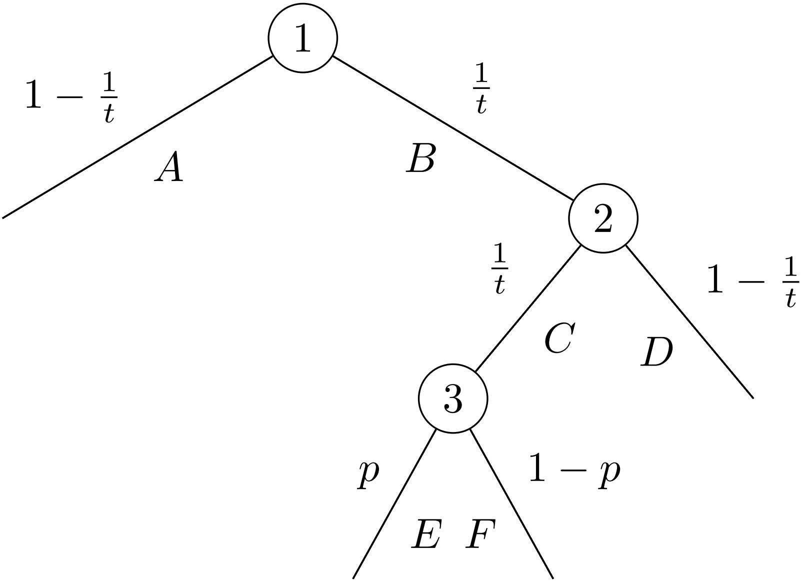
- Fudenberg and Kamada [2018] further extends this to heterogeneous beliefs: some agents in the role of player 1 go to the meetup and others stay home.

- These concepts have only been defined for special classes of games.

- Moreover, while RSCE and RPCE are motivated by learning, they don't yet have explicit learning foundation.

Experimentation and Learning

- SCE corresponds to learning by myopic agents.
- Non-myopic agents may perceive an option value to "experimentation," i.e. playing actions that *probably* don't give good payoffs but might.
- Suppose agents mostly play according to a fixed strategy profile, but "experiment" at rate $1/t$; i.e. there is a lower bound on the probability of each action, and the bound goes to 0 over time at rate $1/t$.
- Then the information sets that can be reached with a single deviation are reached infinitely often, because $\sum_{t=1}^{\infty} 1/t = \infty$.
- So if (as with correctly specified Bayesians) the agents' beliefs converge to their data, then from the law of large numbers beliefs at relevant information sets become correct.
- But the $1/t$ rule needn't lead to correct beliefs at nodes that take 2 or more deviations to reach, because $\sum_{t=1}^{\infty} 1/t^2 \neq \infty$.


1
$1 - \frac{1}{t}$
$\frac{1}{t}$
A
B
2
$\frac{1}{t}$
$1 - \frac{1}{t}$
C
D
3
p
$1 - p$
E
F

- Will player 2 experiment enough to learn 3's play? If they experiment at rate $1/t$ they won't, but $1/t^{1/2}$ would do it.

- More generally, how much experimentation will agents choose to do, and which actions will they experiment with?

- We could directly specify agent heuristics or behaviors.

- Or as in Fudenberg and Levine [1993b] and subsequent work, *derive* experimentation from dynamic programming.

- Find that when players are patient they experiment enough to rule out non-Nash equilibria.

- And by considering which experiments are used more often can derive *equilibrium refinements*, i.e. restrictions on which Nash equilibria can arise.

Steady State Learning in Extensive-Form Games

- Assume agents are Bayesian who believe that they are facing a fixed steady-state distribution of opponents' play, but don't know what the steady state is.
- Each time the game is played, the agents observe the resulting terminal node (but not what opponents would have done at information sets that weren't reached) and update their beliefs about the prevailing distribution of play.
- A large (continuum) population of agents in each player role are randomly matched to play a fixed stage game, so agents don't try to influence how their current partners will play in future matches.
- Assume priors are *non-doctrinaire*, meaning they assign positive probability to all opponent strategies so Bayes rule well defined. (*This requires positive probability even on opponent strategies that are strictly dominated.*).

- Time is discrete and doubly infinite, i.e. $t \in \{..., -2, -1, 0, 1, 2, ...\}$.
- Overlapping generations: Each period some agents leave the population and are replaced. (some papers assume fixed finite lifetimes T, others assume geometrically distributed lifetimes).
- Agents try to maximize their expected discounted lifetime payoffs versus the unknown distribution of opponents' play—so they face dynamic programming problems.
- The resulting system has steady states even though individual agents learn, because departing agents take their information with them.
- Steady states are fixed points of this map.

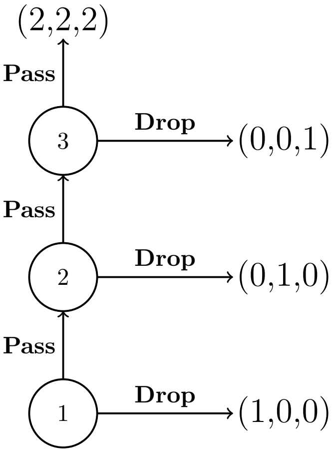
Patiently Stable Strategy Profiles

- Patiently stable strategy profiles are the limits of steady states when agents are long-lived (so they can acquire sufficient data) and also patient (so have an incentive to experiment).
- Theorem: Every patiently stable strategy profile is a Nash equilibrium.
- *High-level proof sketch*:
  - Long lived agents learn the path of play.
  - They eventually stop experimenting, which is why the stable steady states must be self-confirming equilibrium.
  - But if the steady state isn't a Nash equilibrium, some player has a profitable deviation.
  - Use assumption of non-doctrinaire prior to show that when agents in this player role are patient, they will perceive a non-zero option value to experimenting with the deviation, a contradiction.

Refinements of Nash Equilibrium

- Nash equilibrium is not sufficient for patient stability.
- But even when players are long lived and patient they may choose not to experiment if their decision node is only reached very rarely.
- Fudenberg and Levine [2006] show that patient stability doesn't yield backward induction.
- But off-path play isn't completely arbitrary: nodes one step off the path are reached infinitely often, and so play there looks like a SCE.
- Paper argues that incomplete off-path learning can explain the persistence of some "superstitions."
- The model says superstitions can persist forever; a better take-away is that superstitions that are subgame-confirmed can persist a lot longer than those that aren't.

- Fudenberg and Levine [2006] study “simple games” such as this:



- Unique backward induction solution is for all to Pass.
- But (Drop, Drop, Pass) is patiently stable.

Signaling games

- A large literature starting with Spence [1973] studies "signaling games," where an informed sender sends a signal to a receiver.
- Key issue here is what the receiver should believe about the sender's "type" (e.g. productivity in Spence) when the sender chooses a signal that the equilibrium says should never be sent.
- Beliefs here aren't determined by Bayes rule, and the freedom to specify different beliefs lets many outcomes be supported by Nash equilibrium.
- Many equilibrium refinements have been proposed for these games, based on what the receivers "should" believe.
- In the learning framework, these off-path signals are experiments that are chosen by young players.
- And receivers will learn this, which constrains their beliefs and replies.

- Fudenberg and He [2018] uses the Gittins index from multi-armed bandit theory to characterize the implications of patient learning in signaling games.
- Fudenberg and He [2020] adds the restriction that agents know each others' payoff functions to the Fudenberg and He [2018] model: that is, it assumes players assign probability 0 to opponent strategies that are strictly dominated.
- Can do this because receivers have no experimentation incentives, so never play strictly dominated strategies, and senders with independent beliefs won't either.
- Clark and Fudenberg [2021] studies learning in a model where senders can send cheap-talk messages in addition to signals.
- Assumes senders play the game much more often than the receivers do, which fits settings where senders are institutions who interact with a very large number of individual receivers.

Taking Stock

- In "simple" games, sequential equilibrium predicts the backward induction solution, but every subgame-confirmed equilibrium is patiently stable: much weaker predictions than from classic equilibrium refinements

- In signaling games, patient learning can rule out some sequential equilibria, even those that pass the Intuitive Criterion.

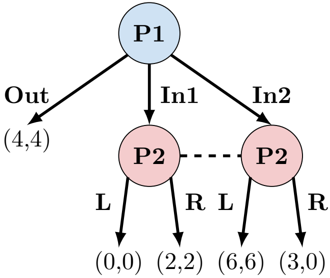
- Signaling games are different in two ways:
  - The relative probabilities of experiments matter, as there are non-singleton information sets.
  - But receivers don't need to experiment; they see sender's type at the end of each round.
  - Patient stability and sequential equilibrium *do* coincide in simple games of length 2.

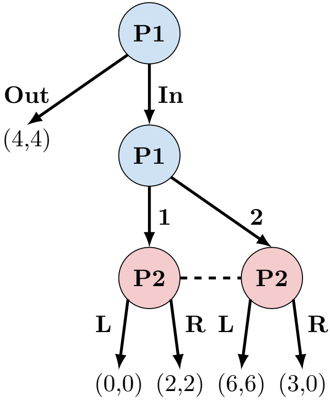
Normal-Form Invariance

- When should two extensive forms have the same solution? E.g. what transformations of the extensive form should leave our predictions unchanged?
- Kohlberg and Mertens [1986] say that the reduced normal form "captures all the relevant information for decision purposes..."
- To this Clark, Fudenberg, and He [2022] add "for fixed beliefs about the play of the opponents."
- We should only expect the set of learning outcomes to be invariant to transformations that are
  - decision invariant: lead to the same best responses as a function of opponent strategies, and
  - information invariant: provide the same feedback to the agents in their learning problems.

Splitting Decision Nodes

- Learning makes the same predictions when a node is split:
  - Decision invariant: same decision problems for fixed beliefs.
  - Information invariant: same feedback to other players.





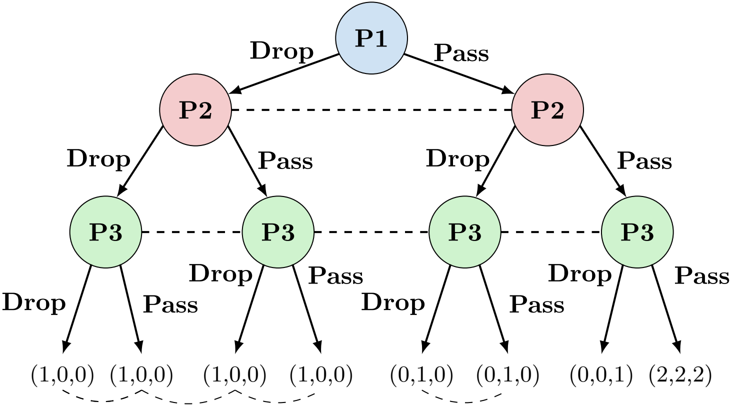
- The outcome **Out** is patiently stable in either extensive form.
- It's a sequential equilibrium outcome in the left-hand game but not when 1's node is split.

Normal Form Versions of Perfect-Information Games

- The normal form is a one-shot simultaneous-move game where players choose strategies simultaneously, and the terminal nodes correspond to the strategies played.
- In the normal form of the Drop/Pass game, the only stable outcome is (**Pass**, **Pass**, **Pass**):
  - Non-doctrinaire priors imply that P3s never play the weakly dominated strategy **Pass**.
  - And unlike in the extensive form, P2s do not need to experiment to learn this.
  - Once P2s learn that P3s **Pass**, they **Pass** as well.
  - And because P1s always observe how P2s and P3s play, they find it better to **Pass**.
- More generally, if agents play the normal form corresponding to a simple game, then stable outcomes must satisfy backward induction for the extensive form.
- The normal form doesn't determine the outcome of learning.

Terminal Node Partitions

- The normal form has the decision-relevant information.
- But for invariance of learning outcomes, we need to replicate the feedback players get in the perfect information extensive form.
- So augment the normal form with a terminal node partition.
- Here the terminal node partition is that if P1 Drops, players do not observe the choices of P2 and P3, and that if P1 Passes and P2 Drops they do not observe the choice of P3.



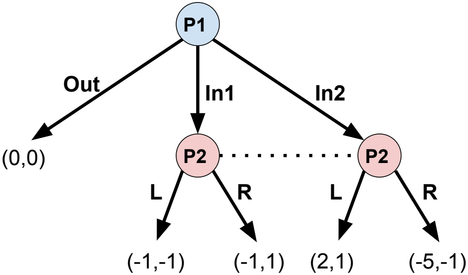
- Terminal node partitions have already been used in defining learning-inspired equilibrium concepts.
- Our work incorporates them into explicit learning models.

Forward Induction

- Kohlberg and Mertens [1986] and subsequent papers argue that solution concepts should imply the iterated deletion of weakly dominated strategies or the related concept of "forward induction."

- Many definitons of forward induction in the literature, but none of them has been accompanied by a theory of how players come to have equilibrium beliefs, or why they should maintain their beliefs in the equilibrium play after observing a deviation.

Violation of Forward Induction



	L	**R**
Out	0,0	0,0
In1	-1,-1	-1,1
In2	2,1	-5,-1

- **In1** is strictly dominated by **Out** for P1.
- When **In1** is deleted, $(\mathbf{In2}, \mathbf{L})$ is the only sequential equilibrium so it is the only strategically stable equilibrium.
- **In1** and **In2** provide same info about P2's action, and **In1** is better if P1's current belief puts high probability on **R**.
- So for some priors a patient P1 agent may choose to experiment with **In1**, and **Out** can be patiently stable.

More on Forward Induction

- Forward induction arguments are not supported by learning, as dominated actions may be used to gain info about others' play.

- Govindan and Wilson [2009] on why forward induction is a desirable refinement: "*If an outcome does not satisfy forward induction... there is an equivalent game in which this outcome results only from Nash equilibria and not from any sequential equilibrium.*"

- But learning doesn't support the idea that equivalent normal forms will be played the same way.

Future Extensions?

- We expect agents to learn the consequences of actions at information sets closer to the path of play more quickly than at information sets further away. Can this intuition be formalized?
- In Clark and Fudenberg [2021], a natural difference in interaction rates between agents in two populations leads to interesting results. Are there other such settings?
- Can we extend learning in games to agents whose priors rule out the true distribution of play, as when agents have incorrect beliefs about the the extensive form of the game.
- Maximizing expected discounted utility is a natural benchmark and starting point, but it would be interesting to extend the learning foundations program to consider heuristics like upper confidence bound policies.

References I

[1] Michel Benaïm. "Dynamics of stochastic approximation algorithms." *Seminaire de probabilites XXXIII*. Springer, 1999.

[2] Michel Benaïm, Josef Hofbauer, and Sylvain Sorin. "Stochastic approximations and differential inclusions." *SIAM Journal on Control and Optimization* (2005).

[3] Michel Benaım, Josef Hofbauer, and Sylvain Sorin. "Stochastic approximations and differential inclusions, Part II: Applications." *Mathematics of Operations Research* (2006).

[4] Joseph Bertrand. "Book review of theorie mathematique de la richesse social and of recherches sur les principes mathematiques de la theorie des richesses." *Journal des savants* (1883).

References II

[5] Anat Bracha, Michael Menietti, and Lise Vesterlund. "Seeds to succeed?: Sequential giving to public projects." *Journal of Public Economics* (2011).

[6] George W Brown. "Iterative solution of games by fictitious play." *Activity analysis of production and allocation* (1951).

[7] Daniel Clark and Drew Fudenberg. "Justified communication equilibrium." *American Economic Review* (2021).

[8] Daniel Clark, Drew Fudenberg, and Kevin He. "Observability, Dominance, and Induction in Learning Models." *arXiv preprint arXiv:2201.00776* (2022).

[9] A Cournot. "Researches into the principles of wealth, 1963 (English Translation)." *Irwin Paperback Classics in Economics* (1938).

References III

[10] Eddie Dekel, Drew Fudenberg, and David K Levine. "Payoff information and self-confirming equilibrium." *Journal of Economic Theory* (1999).

[11] Drew Fudenberg and Kevin He. "Learning and Type Compatibility in Signaling Games." *Econometrica* (2018).

[12] Drew Fudenberg and Kevin He. "Payoff information and learning in signaling games." *Games and Economic Behavior* (2020).

[13] Drew Fudenberg and Yuichiro Kamada. "Rationalizable partition-confirmed equilibrium with heterogeneous beliefs." *Games and Economic Behavior* (2018).

[14] Drew Fudenberg and Yuichiro Kamada. "Rationalizable partition-confirmed equilibrium." *Theoretical Economics* (2015).

References IV

[15] Drew Fudenberg and David M Kreps. “Learning mixed equilibria.” *Games and Economic Behavior* (1993).

[16] Drew Fudenberg and David K Levine. “Self-confirming equilibrium.” *Econometrica* (1993).

[17] Drew Fudenberg and David K Levine. “Superstition and rational learning.” *American Economic Review* (2006).

[18] Drew Fudenberg and David K Levine. “Measuring players' losses in experimental games.” *The Quarterly Journal of Economics* (1997).

[19] Drew Fudenberg and David K Levine. “Steady state learning and Nash equilibrium.” *Econometrica* (1993).

[20] David Gill and Victoria Prowse. “Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis.” *Journal of Political Economy* (2016).

References V

[21] Srihari Govindan and Robert Wilson. "On forward induction." *Econometrica* (2009).

[22] Josef Hofbauer and William H Sandholm. "On the global convergence of stochastic fictitious play." *Econometrica* (2002).

[23] Yuichiro Kamada. "Strongly Consistent Self-Confirming Equilibrium." *Econometrica* (2010).

[24] Elon Kohlberg and Jean-Francois Mertens. "On the strategic stability of equilibria." *Econometrica: Journal of the Econometric Society* (1986).

[25] Michael Spence. "Job Market Signaling." *The Quarterly Journal of Economics* (1973).