UCLA IPAM July 2015

- Learning in (infinitely) repeated games with n players.
- Prediction and stability in one-shot large (many players) games.
- Prediction and stability in large repeated games (big games).
- Prediction and stability cycles in big changing games.

Small Sample of Related Literature

Large Cooperative: Aumann and Shapley (1974); Mailath and Postlewaite (1998)
Stochastic games: Shapley (1953) ; Mertens and Neyman, (1981)
One shot, Continuum: Schmeidler (1973); Rashid (1983); Mas-Colell (1984); Khan and Sun (1999, 2013); Al-Najjar (2004);
One shot, Unknown number of players: Myerson (1998),
One shot, asymptotically large, stability: Kalai (2004); Cartwright and Wooders (2009); Gradwohl, Reingold, Yadin, Yehudayoff (2009); Gradwohl and Reingold (2010); Carmona and Podzeck (2012); Azrieli and Shmaya (2013)
Large dynamic: E. Green (1978); Sabourian (1990); Fudenberg, Levine and Pesendorfer (1996); Al-Najjar and Smorodinsky (2000)
Large markets: Dubey, Mas-Colell, Shubik (1980), Rustichini, Satterthwaite and Williams (1994)
Large mechanisms: Azevedo and Budish (2012), Bodoh-Creed (2012)
Mean Field: Lions (2012), Johari (2010) ….
Big Bayesian: Kalai and Shmaya (2014)

Learning in repeated games: Kalai and Lehrer (1993), Sorin (1999) (also Fudenberg and Levine (1992) in a different context) Neyman (2013).

UCLA IPAM, July 2015

Lecture 1

Rational Learning in Repeated Games

Ehud Kalai and Eran Shmaya
Northwestern University

Definition:
An n-person strategic game is a function:

$u: S \to R^n$ with $S = \times_i S^i$.

S^i is the set of *strategies* of player I,

S is the set of *strategy profiles* or *strategy configurations*, and

u^i is the *payoff function* of player i.

Nash equilibrium

Extends the idea of equilibrium from supply and demand to general behavior between interacting players

Has taken over as a major analytical tool in economics

Operations management, political science and computer science are going through similar transformations

Coincides with behavior predicted by survival of the fittest.

Definition: A **Nash equilibrium s*** is a configuration of individual strategies, each optimal (best response) relative to the others, i.e., no player has an incentive to unilaterally deviate from the configuration.

$$u^i(s^{*1},\ldots,s^{*i-1},s^i, s^{*i+1},\ldots,s^{*n}) \leq u^i(s^*)$$

Simple familiar examples:

Everybody driving on the right side of the road.

Markets, real and on the web.

Complementarities in production:

a. Simultaneous production of software and of hardware

also

b. No production of software with no production of hardware.

But

production of software without production of hardware is not an equilibrium.

Common language, common system of measurements…

Example: be generous or selfish

Aka Prisoners' dilemma

when a $1 donation yields your opponent $3.

She

He	generous	selfish
generous	2, 2	-1, 3
selfish	3,-1	0, 0

The only Nash equilibrium is non-cooperative:
Both players choose the selfish action.

Example: shy woman / bold man,

Aka match pennies

he wants to be with her she wants to be alone.

		She: .5 in	She: .5 out
He	.5 in	1, 0	0, 1
He	.5 out	0, 1	1, 0

This game has no "pure strategy" Nash equilibrium, but it has a "mixed strategy" equilibrium: each player chooses one of the two options with equal probability.

Nash PhD thesis: if we allow for mixed strategies, every finite game has a Nash equilibrium.

Computer choice game with known types

Each has to choose $\mathcal{PC}$ or $\mathcal{M}$.
He likes $\mathcal{PC}$ she likes $\mathcal{M}$, but they also like to make the same choice.

His payoff: 1 if they choose the same computer (0 otherwise)
+ .2 if he chooses $\mathcal{PC}$ (0 otherwise).

Her payoff: 1 if they choose the same computer (0 otherwise)
+ .2 if she chooses $\mathcal{M}$ (0 otherwise).

This game has
two pure strategy equilibria: (1) both $\mathcal{PC}$ and (2) both $\mathcal{M}$;

and one mixes strategy equilibrium:
he randomizes .6 to .4 between $\mathcal{PC}$ and $\mathcal{M}$, and
she randomizes .4 to .6 between $\mathcal{PC}$ and $\mathcal{M}$.

Aka Prisoners' dilemma

Example: be generous or selfish, **repeated play**

- The same two players play the "stage game" in periods 1,2,…, with "perfect monitoring."

He \ She	generous	selfish
generous	2, 2	-1, 3
selfish	3,-1	0, 0

- At the end of every period, each is told the choice of his opponent and receives a payoff according to the table.
- A **strategy** is a function form histories of play to period choices ex. ***f^1(g\g,g\s,g\g) = (gen w.p. .9, selfish w.p. .1)***
- Present value is computed with a discount parameter d.

(f^1,f^2) is an **equilibrium** if each f^i maximizes the expectated present value of the total future payoffs.

Example: both play tit for tat ; average payoff = 2,2 (if d>1/3) .

Example of a Bayesian game

<u>Computer choice game with **unknown types**</u>

- Each of two players have to choose $\mathcal{PC}$ or $\mathcal{M}$.
- Each is of one random type: likes $\mathcal{PC}$ or likes $\mathcal{M}$, with prob .50 - .50
- Each player knows his own type, but only the probabilities of the opponent's type.
- Identical individual payoff functions:

1 if you choose the same computer as your opponent (0 otherwise)
\+ .2 if you chooses the computer you like (0 otherwise).

In Bayesian Nash equilibrium strategies are type dependent.

For example: choose your favorite computer, i.e.,
Choose $\mathcal{PC}$, if you like $\mathcal{PC}$; and choose $\mathcal{M}$ if you like $\mathcal{M}$.

Notice that the equilibrium is efficient if they happen to be of the same type (with prob. .5); and inefficient otherwise.

The mechanism design literature deals with fixing such inefficiencies.

Example of a **Bayesian repeated game**

Computer choice game with unknown types played repeatedly

- First, players are assigned types as above, to remain fixed throughout the repeated play.
- Then, the computer choice game is played in periods 1,2,… with perfect monitoring and with discounted sum of payoffs.

Strategy profile $\boldsymbol{f} = (\boldsymbol{f}^1, \boldsymbol{f}^2)$:
a vector of vectors of repeated game strategies.

For Player 1, for example: $\boldsymbol{f}^1 = (\boldsymbol{f}^{1,\mathcal{PC}}, \boldsymbol{f}^{1,\mathcal{M}})$

- $\boldsymbol{f}^{1,\mathcal{PC}}(h)$ describes the computer choice probabilities of his $\mathcal{PC}$- type, after the history of observed past choices h.
- $\boldsymbol{f}^{1,\mathcal{M}}(h)$ is the same, but for his $\mathcal{M}$- type.

Example of a **Baye** **n repeated game**

Computer choic known types played repeatedly

- First, pla above, to remain f
- Then, th periods 1,2,… with p discounted sum of payoffs.

For example

$f^{1,\mathcal{PC}}(\emptyset) = (.80,.20)$ means that initially he chooses $\mathcal{PC}$ wp .80 and $\mathcal{M}$ wp .20.

$f^{1,\mathcal{PC}}\left(\begin{smallmatrix}\mathcal{MMM}\\ \mathcal{MMM}\end{smallmatrix}\right) = (0,1)$ means that if they both chose $\mathcal{M}$s in the first three periods, in the fourth period he would choose $\mathcal{M}$ for sure.

Strategy profile $\boldsymbol{f} = (\boldsymbol{f}^1, \boldsymbol{f}^2)$:
a vector of vectors of repeated game strategies.

<u>For Player 1, for example</u>: $\boldsymbol{f}^1 = (\boldsymbol{f}^{1,\mathcal{PC}}, \boldsymbol{f}^{1,\mathcal{M}})$

- $\boldsymbol{f}^{1,\mathcal{PC}}(h)$ describes the computer choice probabilities of his $\mathcal{PC}$- type, after the history of observed past choices h.
- $\boldsymbol{f}^{1,\mathcal{M}}(h)$ is the same, but for his $\mathcal{M}$- type.

Strategy profile $f = (f^1, f^2)$ is **a Bayesian equilibrium** if for every player and type, his repeated game strategy maximized the type's expected present value of payoff, given the distribution of the opponent's types and their repeated game strategies.

Example of Kalai and Lehrer (1993) rational learning:

<u>assume that Pl 1 turns out to be a $\mathcal{PC}$ type and Pl 2 a $\mathcal{M}$ type:</u>

They will play $(f^{1,\mathcal{PC}}, f^{2,\mathcal{M}})$, with

$f^{1,\mathcal{PC}}$ being optimal against the .50 -.50 belief that he is facing $f^{2,\mathcal{PC}}$ or $f^{2,\mathcal{M}}$;

(similarly for Pl 2).

But maximizing the PV of future payoffs, implies that Pl 1 (and Pl 2) plays optimally relative to **Bayesian updated beliefs**:

Instead of the initial prior belief, *prob*$(t^2=\mathcal{PC})=.5$, after any history of past plays h, he uses the posterior belief, *prob*$(t^2=\mathcal{PC} \mid h)$, and optimize against it.

The *prob*$(t^2=\mathcal{PC} \mid h)$ must converge, but not necessarily to the true probability, which is zero since Pl 2 is a $\mathcal{M}$-type.

The convergence is a direct consequence of the Martingale convergence theorem.

But maximizing the PV of future payoffs, implies that Pl 1 (and Pl 2) plays optimally relative to **Bayesian updated beliefs**:

Instead of the initial prior belief, $prob(t^2 = \mathcal{PC}) = .5$, after any history of past plays h, he uses the posterior belief, $prob(t^2 = \mathcal{PC} \mid h)$, and optimize against it.

The $prob(t^2 = \mathcal{PC} \mid h)$ must converge, but not necessarily to the true probability, which is zero since Pl 2 is a $\mathcal{M}$-type.

Nevertheless, Pl 1's forecasts of the future play will become accurate, he will forecast Pl 2's choices under $f^{2,\mathcal{M}}$, as if he knew that she is the $\mathcal{M}$- type.

But maximizing the PV of future payoffs, implies that Pl 1 (and Pl 2) plays optimally relative to **Bayesian updated beliefs**:

Instead of the initial prior belief, $prob(t^{2}=\mathcal{PC})=.5$, after any history of past plays *h*, he uses the posterior belief, *prob(* … optimize against it.

This follows from the merging literature, see Kalai and Lehrer (1994) for sufficient conditions that hold in repeated games

The … to th … type.

Nevertheless, Pl 1's forecasts of the future play will become accurate, he will forecast Pl 2's choices under $f^{2,\mathcal{M}}$, as if he knew that she is the $\mathcal{M}$- type.

It follows that:

Theorem (Kalai and Lehrer 1993b) at any Bayesian equilibrium players converge to play a **subjective equilibrium** of the repeated game: At such an equilibrium they each play optimally relative to his beliefs, which may be false, but consistent with the observed data.

See von Hayek (1937) for the idea of subjective equilibrium, and

see also Battigalli (1987) Fudenberg and Levin (1993) on the idea of self-confirming subjective equilibria in different contexts

It follows that:

Theorem (Kalai and Lehrer 1993b) at any Bayesian equilibrium players converge to play a **subjective equilibrium** of the repeated game: At such an equilibrium they each play optimally relative to his beliefs, which may be false, but consistent with the observed data.

Theorem (Kalai and Lehrer 1993a) at any Bayesian equilibrium, the play will converge to an approximate Nash equilibrium of the repeated game, as if the types of both players are common knowledge.

The first theorem holds for any number of players.

For the second theorem with more than two players assume subjective independence: conditional on his realized type, every player believes that his opponents' types are independent of each other.

Example: Repeated Production

n players producing widgets at time periods 1,2,3,…

At the beginning of each period, a producer decides

1. the type of widgets he will produce, and
2. his selling price

At the end of the period he observes his competitors' choices and collects his period's profit.

Each producer knows only his own (constant) production capabilities and costs.

At a Bayesian equilibrium he maximizes the expected present value of all his future profits.

A producer may act strategically. For example, he may:

Learn: experiments with period choices to test the competitors responses.

Teach: sell widgets at low prices in selected periods to deceive his competitors about his cost and discourage their participation.

Nevertheless, with time:

The producers learn to predict the future choices of their competitors,

and play as if everybody's capabilities and costs are common knowledge.

The **rate of convergence in Theorem 1** may be arbitrarily bad. But when bad, it is because your forecasts are correct for many early periods. For example there may only be uncertainty about what your opponent will do in period 1 million.

However, from Sorin (1999) , there is only a finite number *K* of "learning periods" in which your forecast is significantly wrong. *K* depends on the accuracy of the initial beliefs of the players.

UCLA IPAM, July 2015

lecture 2

Learning, Predicting and Stability in Big Games

Ehud Kalai and Eran Shmaya
Northwestern University

Predictability and stability are critical for well functioning social systems.

Producers and consumers need reliable prices in order to plan their activities.

Predictable stable **driving patterns** are important for proper delivery of goods, planning of roads by traffic engineers, transportation of passengers …

Predictable stable demand and supply is important in **health delivery systems**, in the provision of **new technology**, etc. etc..

Predictability means that players can predict (with close to certainty) the outcome of a period before the play starts.

Hindsight **stability** means that the players have the incentives to follow the plans they each made prior to the start of a period, even after they observe the period's outcome (aka no regret, ex-post Nash)

Example, driving from the north suburbs to downtown Chicago in a morning rush hour; use one of two possible roads, E or W.

Predicting means that they know, with close to certainty, what the driving times on the two roads will be.

Stability means that once they started driving and hear the reported driving times on the radio, no driver has the incentive to deviate to a different road.

General principles:

- Periods that are not stable are (potentially) chaotic.
- Correct predictions are sufficient for stability, but not vice versa.
- Correct forecasts, i.e., assessing correct probabilities of driving times, are not sufficient for stability.
- With a small number of players we learn to forecast, but with a large number we learn to predict.
- Learning in period k if and only if period k's outcome is unpredictable.

Lecture Road Map

Part A: Motivation from Kalai, Econometrica (2004)

Hindsight stability in large **one-shot** games with **independent types,**

Part B: Learning, predicting and stability in big games, Kalai and Shmaya (DPs 2014a and 2014b)

- Markov perfect equilibrium in an **imagined-continuum model** of a repeated population game: A behaviorally simple ϵ-equilibrium, of a highly complex game.
- **Learning, predicting** and **hindsight Stability**.
- **Stability Cycles** in big games.

Lecture Road Map

Part A: Motivation from Kalai, Econometrica (2004)

Hindsight stability in large **one shot** games with **independent types,**

Part B: Learning, predicting and stability in big games, Kalai and Shmaya (DPs 2014a and 2014b)

- Markov perfect equilibrium in an **imagined-continuum model** of a repeated population game: A behaviorally simple ϵ-equilibrium, of a highly complex game.
- **Learning, predicting** and **hindsight Stability**.
- **Stability Cycles** in big games.

Example

Simultaneous-move Computer choice game: independent types

Players: i = 1,2,…,n;   each chooses an action: $a^i = \mathcal{PC}$ or $a^i = \mathcal{M}$

Player's types: iid $\Pr(t^i = \mathcal{PC}) = \Pr(t^i = \mathcal{M}) = .50$

Individual's payoff: $u^i = \text{prop}_{j \neq i}\,(a^j = a^i)^{1/3} + 0.2\delta_{a^i = t^i}$,

i.e., (the proportion of opponents he matches)$^{1/3}$
+ 0.2 if he chooses his computer type (0 otherwise).

Choose your favorite computer $(a^i = t^i)$ is a *Nash* equilibrium.

It is "asymptotically **hindsight stable,**" as the number of players increases.
due to predictability obtained through the laws of large numbers

Even more, it is "asymptotically **structurally robust**": it remains an equilibrium in all extensive-game alterations that (1) start with the same initial information, (2) preserve the players' strategic possibilities and (3) do not alter the players' payoffs.

For example, it survives under **sequential play** (no herding),
and, more generally, under **general dynamic play**: **revision of choices, information leakage, cheap talk, delegation possibilities**, and more..

Modelling Partially-specified games

Students choosing computers on the web

Instructions: "Go to web site xyz before Friday and click in your choice $\mathcal{PC}$ or $\mathcal{M}$."   Types, choices, and payoffs as before.

Need to know: who are the players? the order of play? monitoring? communications? commitments? delegations? revisions?... **Impossible**

But under structural robustness: any equilibrium of the one-shot simultaneous-move game, (e.g., choose your favorite computer) remains equilibrium no matter how you answer the above.

Price formation in Shapley Shubik market games

Hindsight stability → price stability

Kalai Econometrica (2004): In n-player one-shot simultaneous-move Bayesian games with independent types:

many players,
semi-anonymity,
continuity &
~~Independent types~~

All Nash equilibria
are asymptotically
hindsight stable

So hindsight stability is important

But it fails when player types are correlated,
common in economic interaction.

~~**Price stability** in market games~~

~~All Nash equilibria are asymptotically **structurally robust**~~

Hindsight stability fails with correlated types

will be used repeatedly in the following slides

Computer choice game with correlated types.

Players: i = 1,2,…,n, each chooses $\mathcal{PC}$ or $\mathcal{M}$.

Unknown **state of nature**: the computer with better overall features is:

s = $\mathcal{PC}$ or s = $\mathcal{M}$ with prob .50 , .50 .

Player types: iid conditional on s: $\Pr(t^i = s) = 0.7$, $\Pr(t^i = s^c) = 0.3$.

Payoffs: as before.

Equilibrium: everybody chooses her favorite computer.

It is **not hindsight stable** when n is large.

But notice: after the one-shot play they all know the state of nature and now their types are (conditionally) independent.
This suggests the study taken next

What happens with hindsight stability in large **repeated games** with **correlated types?**

Hindsight stability fails with correlated types

will be used repeatedly in the following slides

Computer choice game with correlated types.

Players: i = 1,2,…,n, each chooses $\mathcal{PC}$ or $\mathcal{M}$.

Unknown **state of nature**: the computer with better overall features is:

s = $\mathcal{PC}$ or s = $\mathcal{M}$ with prob .50 , .50

Player types: iid conditional on s: $\Pr(t^i = s) = 0.7$…

Payoffs: as before.

E…

Answer to follow:
If the number of players is large, with the exception of a finite number of chaotic learning periods, all periods are asymptotically hindsight stable.

…play they all know the state of …types are (conditionally) independent.

…ggests the study taken next

What happens with hindsight stability in large **repeated games** with **correlated types?**

Lecture Road Plan

Part A: Motivation from Kalai, Econometrica (2004)

Hindsight stability in **one shot** games with **independent types,**

Part B: Learning, predicting and stability in big games, Kalai and Shmaya (DPs 2014a and 2014b)

- Markov perfect equilibrium in an **imagined-continuum model** of a repeated population game: A behaviorally simple equilibrium, of a highly complex game.
- **Learning, predicting** and **hindsight Stability**.
- **Stability Cycles** in big games.

The Repeated Game with fixed unknown fundamentals

A symmetric anonymous repeated game of proportions with:

1. A large but unknown number of players n.
2. Fixed types, correlated through an unknown state of nature (game fundamentals).
3. Imperfect monitoring.

An imagined-continuum equilibrium:

- every player computes her best response based on expected values, as if she is negligible in a continuum of players.
- But (as game theorists) we compute probabilities of events in the actual n-person process, in which the n players follows the imagined-continuum reasoning above.

Will illustrate the concepts through a

Repeated computer choice game with correlated types.

Prior probabilities:

One (unknown) **state of nature** s,
$\theta_0(s)$ is the known prior prob that the state is s.

$\theta_0(S = \mathcal{PC}) = \theta_0(S = \mathcal{M})$=1/2,
each computer is equally likely to have the better overall features

Players' privately known **types** are distributed by a conditional iid,
τ_s(t) is the prob of a player being of type t when the state is s.

$\tau_{\mathcal{PC}}(t = \mathcal{PC})$ = .7 and $\tau_{\mathcal{PC}}(t = \mathcal{M})$ = .3:
when the state is $\mathcal{PC}$, independently of the others every player has prob of 0.7 to be a $\mathcal{PC}$ type and 0.3 to be a $\mathcal{M}$ type.

Symmetrically, $\tau_{\mathcal{M}}(t = \mathcal{M})$ = .7 and $\tau_{\mathcal{M}}(t = \mathcal{PC})$ = .3

The Stage Game, played in periods k = 0,1,2,…:

Actions available to every player are denoted by $a \epsilon A$

a = $\mathcal{PC}$ or a = $\mathcal{M}$, chooses $\mathcal{PC}$ or chooses $\mathcal{M}$.

$e_k(t, a)$ is the **empirical proportion** of players *who* are of type t and choose the action a.

$e_k(\mathcal{PC}, \mathcal{PC})$ the proportion of players who like $\mathcal{PC}$ and chose $\mathcal{PC}$,
$e_k(\mathcal{PC}, \mathcal{M})$ the proportion of players who like $\mathcal{PC}$ but chose $\mathcal{M}$, etc.

A random **outcome** x, is chosen with prob $\chi_{s,e}(x)$ and made public at the end of the period.

A sample with replacement of J computer users is taken $\boldsymbol{x} = x(\mathcal{PC})$ is the sample proportion of $\mathcal{PC}$ users. $x(\mathcal{M}) \equiv 1 - x(\mathcal{PC})$.

$u(t, a, x)$ is the **period payoff** of a player of type t who took the action a in a period with the outcome x.

u = (the proportion of users in the sample that his choice matches)$^{1/3}$
+ 0.2 if his choice is the same as his type, $a = t$.

The Repeated game

Can be:

Infinitely repeated with discounting.

Finitely repeated with the average of the periods payoffs.

Any function that is continuous and strictly monotonic in the periods payoffs. Will elaborate.

The repeated computer choice game is infinitely repeated with individual discount parameters.

Strategies and equilibrium terminology

We restrict the players to simplified symmetric equilibrium that does not depend on n, but we do a full Bayesian asymptotic analysis of such n players, as $n \to \infty$.

A **common strategy** F is a symmetric profile in which all the players play F.

F is **Markov**, if it depends only on the player's type and the "public-belief" over the unknown state, will elaborate.

In the **imagined play path**, periods' random empirical distributions of types & actions are replaced by their deterministic conditional expectations. The only uncertainty is about the unknown state s.

A **Markov strategy** is a function $\mathrm{F}: \Delta(S) \times \mathrm{T} \rightarrow \Delta(A)$

$F_{\theta,t}(a)$ is the probability of choosing the action a by a player of type t in periods in which the *Markov state* is θ, where θ is the prob distribution that describes the public belief about the unknown state s, to be described.

An $\boldsymbol{\alpha}$ threshold strategy : With prob 1:

Choose your type of computer t in periods with

$\alpha < \theta(t)$, i.e., $F_{\theta, t}(t) = 1$

choose the other computer t^c in periods with

$\theta(\mathrm{t}) \leq \alpha$ i.e., $F_{\theta, t}(t^c) = 1$

The **public beliefs** (in the imagined game) about s under a common strategy *F*

The **initial** public belief is θ_0 .

Continuing inductively, in a period that starts with the imagined public belief θ:

1. to every state s associate the (imagined) deterministic **empirical distribution**

the expected values from the continuum game.

$$\mathrm{d}_\theta(t,a) = \tau_s(t) \cdot F_{\theta,t}(a),$$ and

2. after observing the period outcome x, compute the **posterior** public belief by Bayes rule

$$\hat{\theta}_{\theta,x}(s) \equiv \frac{\theta(s)\cdot\chi_{s,d_\theta}(x)}{\sum_{s'} \theta(s')\chi_{s',d_\theta}(x)}$$

Suppose a period's **prior public belief** is $\theta(s = \mathcal{PC}) = .6,\ \theta(s{=}\mathcal{M}) = .4$ and that under $F_{\theta,t}$ each player chooses his computer type with probability 1.

1. The **imagined** empirical distribution $d_\theta(t, a)$ is

For s=$\mathcal{PC}$
$d_\theta(\mathcal{PC}, \mathcal{PC}) = .7 \cdot 1 = .7$
$d_\theta(\mathcal{PC}, \mathcal{M}) = .7 \cdot 0 = .0$
$d_\theta(\mathcal{M}, \mathcal{PC}) = .3 \cdot 0 = 0$
$d_\theta(\mathcal{M}, \mathcal{M}) = .3 \cdot 1 = .3$

For s=$\mathcal{M}$
$d_\theta(\mathcal{PC}, \mathcal{PC}) = .3 \cdot 1 = .3$
$d_\theta(\mathcal{PC}, \mathcal{M}) = .3 \cdot 0 = .0$
$d_\theta(\mathcal{M}, \mathcal{PC}) = .7 \cdot 0 = 0$
$d_\theta(\mathcal{M}, \mathcal{M}) = .7 \cdot 1 = .7$

Recall, the period outcome x($\mathcal{PC}$) is the proportion of $\mathcal{PC}$ users in a sample with replacement of J computer users from the population.

2. Suppose that in J=20 observations there were 11 $\mathcal{PC}$ users, then the **posterior public belief** is :

$$\hat{\theta}_{\theta,11/20}(s = \mathcal{PC}) = \frac{0.6 \cdot B_{20,0.7}(11)}{0.6 \cdot B_{20,0.7}(11) + 0.4 \cdot B_{20,0.3}(11)} = 0.89$$

$$\hat{\theta}_{\theta,11/20}(s = \mathcal{M}) = \frac{0.4 \cdot B_{20,0.3}(11)}{0.6 \cdot B_{20,0.7}(11) + 0.4 \cdot B_{20,0.3}(11)} = 0.11$$

$B_{20,0.7}(11)$ is binomial prob of 11 successes in 20 trials with success prob 0.7

Private beliefs in a period with public belief θ of a player of type t is

$$\theta^{(t)}(s) \equiv \frac{\theta(s) \cdot \tau_s(\mathrm{t})}{\sum_{s`} \theta(s') \cdot \tau_{s'}(\mathrm{t})}$$

Period expected payoff (in the imagined game) of a type t who chooses the action a when the common strategy is F and the public belief is θ

$$u_F(\theta, t, a) \equiv \sum_x \left(\sum_s \theta^{(t)}(s) \cdot \chi_{s,d_\theta}(x) \right) u(t, a, x)$$

The probability that the player assigns to the outcome x

The probability that the player assigns to the outcome x, for a given s,

Computed with the public belief θ, not with $\theta^{(t)}$, because the strategies of the players are conditioned on the public belief, θ.

Recall: $\mathrm{d}_\theta(t, a) = \tau_s(t) \cdot F_{\theta,t}(a)$,

Definition: A common strategy F is a **Markov equilibrium** (in the imagined game) if for every public belief $\theta \epsilon \Delta(S)$ and every type $t \epsilon T$, if $F_{\theta,t}(a) > 0$, then a maximizes $u_F\ (\theta, t, a)$.

The players needs no information about the size of the population.

Kalai and Shmaya (2014a) define an equilibrium in the imagined game without the Markov and Myopicity properties, and show that:

- Myopicity is a result, not an assumption.
- When the number of players is large:
  1 Period probabilities in the imagined game approximate the real probabilities.
  2 Best response strategies in the imagined game are uniform e - best response uniformly, for all n> n_0 (same for Nash equi.)

Due to myopicity

- Markov equilibrium and the predictability / stability results are applicable to many repetition-payoff structures, e.g., finitely repeated games with average payoff, short and long lived players, overlapping generations, etc.
- The equilibrium may be used in segments within the big repeated games with changing fundamentals.
- Existence and equilibrium computation are simple matters.

For the common strategy G in which every player chooses her favorite computer, define $\boldsymbol{\alpha} = \theta(\mathcal{PC})$ at which $u_G\,(\theta, \mathcal{PC}, \mathcal{PC}) = u_G\,(\theta, \mathcal{PC}, \mathcal{M})$,

α is the **tipping value**: when everybody chooses her favorite computer, how low must the prior on $S = \text{PC}$ be to make PC types choose M.

$\boldsymbol{\alpha} < ½$; it is the same for $\mathcal{PC}$ and for $\mathcal{M}$.

The common α threshold strategy is a Markov equilibrium.

- It is trivial to check: direct from the definition.
- It is easy to play: start by choosing your computer type, continue by updating the public belief and following the threshold rule.
- Coordination: from some time on they will all use the same computer.
- It is a real ε Nash equilibrium if the number of players is large.
- If in addition the sample size is large, they will all be using the better computer from the second period on.

Uniform Learning to Predict

Definition: consider a common Markov strategy F, period k is (uniformly) <u>asymptotically predictable</u> up to $[\, r, \delta, \rho \,]$, if with sufficiently many players

$$\Pr_F \left(\begin{array}{l} \text{Every player assigns probability} \geq 1-\delta \text{ to the ball} \\ \text{of radius } r \text{ around the true outcome of period } k \end{array} \right) \geq 1-\rho$$

Theorem 1: For every positive δ and ρ there is a finite integer K s.t. under any Markov strategy F and any positive r, all but at most K periods are asymptotically predictable up to $[\, r, Q(r)+\delta, Q(r)+\rho \,]$.

The lack of concentration of the outcome function, i.e., the measure of the set of outcomes that cannot fit into a ball of diameter r in the worst case (over all s and e).

Corollary: Suppose the outcome function has a variance σ^2. For every positive δ and ρ there is a finite integer K s.t. under any Markov strategy F and any positive r, all but at most K periods are asymptotically predictable up to $[\, r, 4(\sigma/r)^2+\delta, 4(\sigma/r)^2+\rho \,]$.

In the computer choice game,

for arbitrarily small δ and ρ there is a finite *K* s.t. under any Markov strategy *F* and for any positive *r*, all but at most *K* periods are asymptotically predictable up to

$$[\, r,\ 1/Jr^2 + \delta\ ,\ 1/Jr^2 + \rho\ ],$$

i.e., with sufficiently many players

$$\Pr_F \left(\begin{array}{l} \text{every player assigns probability} > 1- (1/Jr^2 + \delta) \\ \text{to the ball of radius } r \text{ around the true} \\ \text{outcome of period } k \end{array} \right) > 1- (1/Jr^2 + \rho)$$

With a large sample size *J*, there is a high probability of approximate uniform predictability.

Hindsight Stability

Definition: A common Markov strategy *F* is asymptotically hindsight stable in period *k* up to $[\varepsilon, \rho]$, if with sufficiently many players

$$\Pr_F \left(\begin{array}{l} \text{after observing the period's outcome, by a unilateral change of} \\ \text{her action } \textbf{some} \text{ player can improve her payoff by more than } \varepsilon \end{array} \right) \leq \rho$$

Theorem 2. For every positive ε, ρ there is an integer *K* s.t. in every Markov equilibrium F and every d > 0 all but at most K periods are hindsight stable up to $[\, 2d + 2Q(w^{-1}(d)) + 2\varepsilon,\ Q(w^{-1}(d)) + \rho \,]$

$W^{-1}(d)$ is the modulus of continuity of u, a generalized Lipschitz value for points that are d units apart.

Corollary for payoff with Lipschitz constant *L* and outcome with variances $\leq \sigma^2$. For every positive ε, ρ there is an integer *K* s.t. in every Markov equilibrium *F* all but at most *K* periods are asymptotically hindsight stable up to

$$[8(\sigma L / \varepsilon)^2 + 2\varepsilon\,,\ 8(\sigma L / \varepsilon)^2 + \rho]\ .$$

In the computer choice game,

for arbitrarily small ε and ρ there is a finite K s.t. under any Markov equilibrium F all but at most K periods are uniformly asymptotically stable up to

$$[\, 4\varepsilon + 2 / J\varepsilon^6 ,\ \rho + 1 / J\varepsilon^6 \,],$$

i.e., with sufficiently many players

$$\Pr_F \left(\begin{array}{l} \text{With hindsight, by a unilateral change of his} \\ \text{action some player can improve his payoff by} \\ \text{more than } 4\varepsilon + 2 / J\varepsilon^6 \end{array} \right) < \rho + 1 / J\varepsilon^6$$

- Reducing the noise in the publicly observed outcomes (bigger sample size J in our example) improves hindsight stability.
- But with substantial noise in the observed outcomes, hindsight **in**stability is unavoidable, regardless of the number of players.

Rough intuition about the proof of learning to predict

(That hindsight stability follows from predictability is intuitively clear)

Consider first the |T | imagined processes, in which for every s the t-types hold deterministic beliefs about the probabilities of the period empirical distributions of type and actions, $d_\theta(t,a)$.

- Merging, under the automatic grain of truth, implies that with high probability, except for a finite number of learning periods, the forecasted probabilities over the outcome of the periods are appx accurate. (Fudenberg-Levin, Sorin, Kalai-Lehrer), i.e., the same as would be forecasted with knowledge of the unknown state.
- High concentration (small variance in our example) of the outcome distribution, combined with the fact that the empirical distributions in the imagined processes are deterministic conditional on the states, implies that with high probability at the non-learning periods they predict the realized period outcomes (not just their probabilities).

So in the imagined processes, in all non-learning periods players will have approximately correct predictions.

So in the imagined processes, in all non-learning periods players will have approximately correct predictions.

But what about in the real process, in which the players observe **the randomly realized** real outcomes?

Building on Kalai (2005), Kalai and Shmaya (2013) show that when the number of players is large and outcome probabilities are continuous, real probabilities of period events are approximated well by the probabilities in the imagined process. Thus appx correct predictions holds with (real) high probability in the non-learning periods.

Remarks:

1. **Predictability is a result of "no further learning" from some time on.** Similar to multi arm bandit problems, the players do not necessarily learn the real state of nature, or even learn to play "as if" they know it.

2. **On the rate of getting to predictability**: We know from Sorin (1999) that the number of chaotic periods is monotone in the size of the grain of truth, which is bounded below in our population game. Thus the number of unpredictable periods is bounded above.

Stability Cycles in Big Games

by

Ehud Kalai and Eran Shmaya

Abstract. In a **big game a** large anonymous population plays an infinitely repeated (stochastic) game in which:

(1) game fundamentals (stochastic state) and the set of players change over time,

(2) players private types are correlated through the fundamentals, and

(3) information about fundamentals and play is incomplete and imperfect.

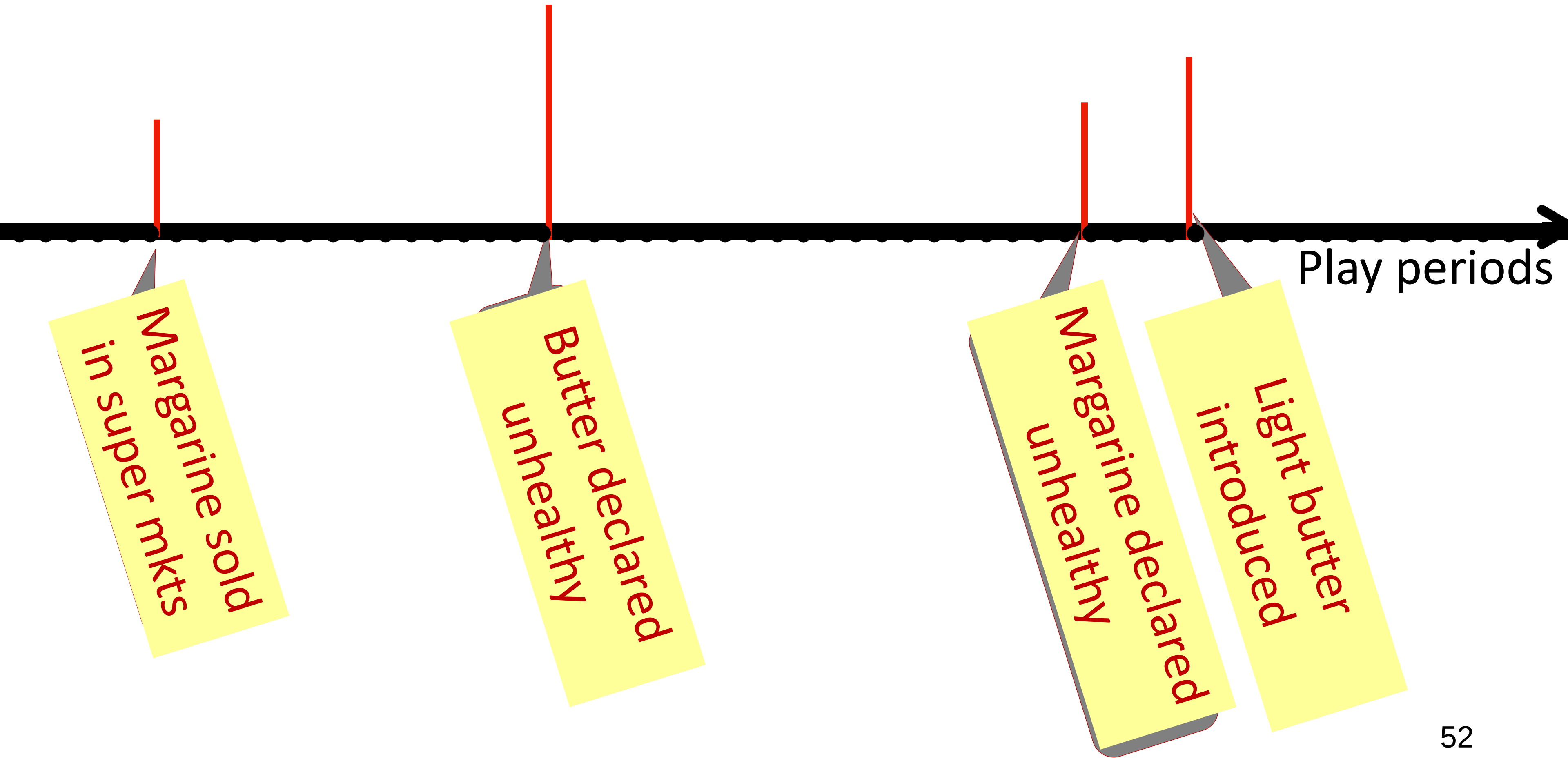
Important games, but difficult to analyze.

Good news

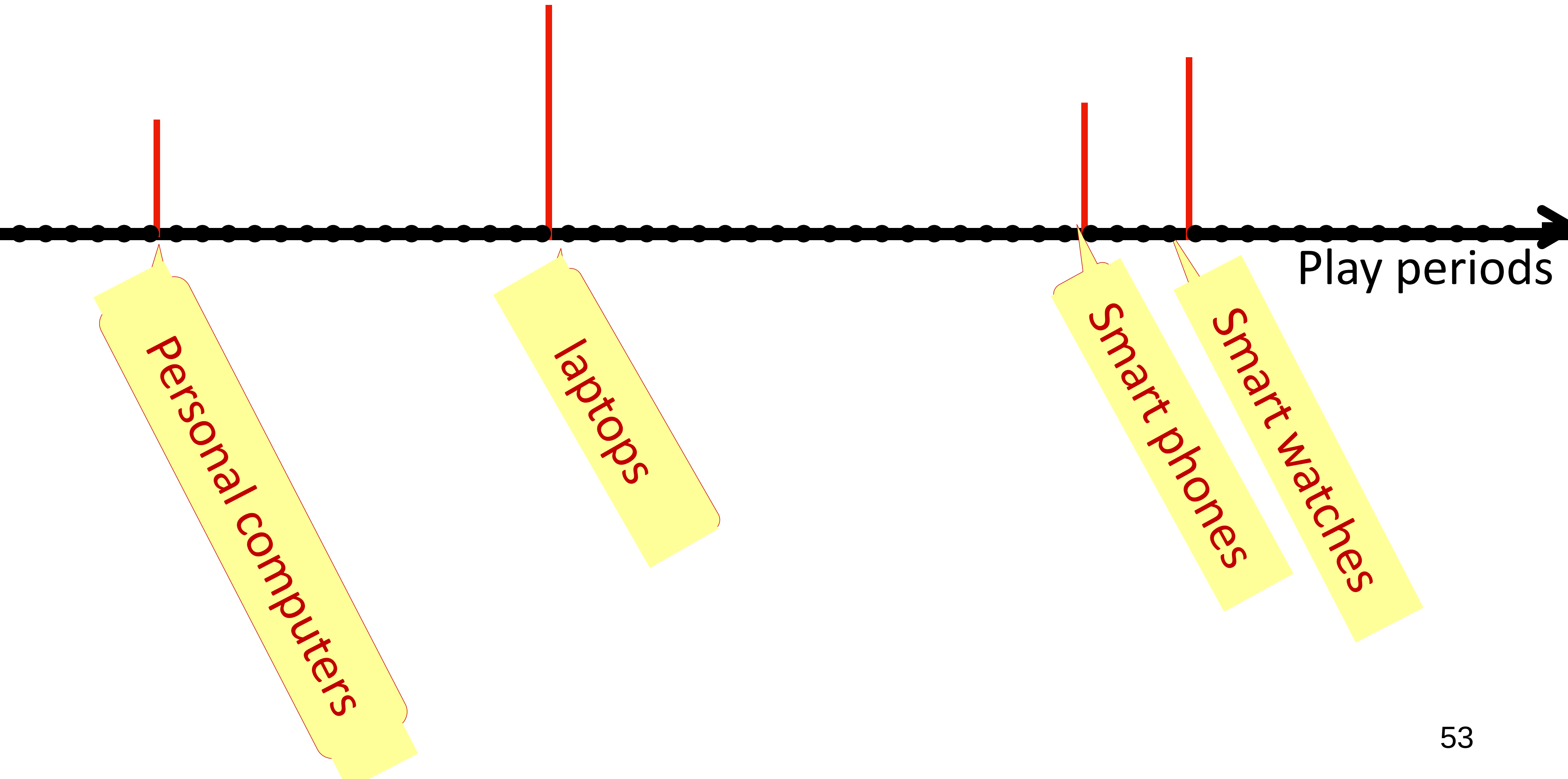
When fundamental changes are guided by aggregate population data:

- The play admits a simple behavioral myopic Markov perfect equilibrium, and
- the period outcomes are highly predictable and the play is hindsight stable, provided that fundamental changes are infrequent and external uncertainty is low.

Example: Market for Butter



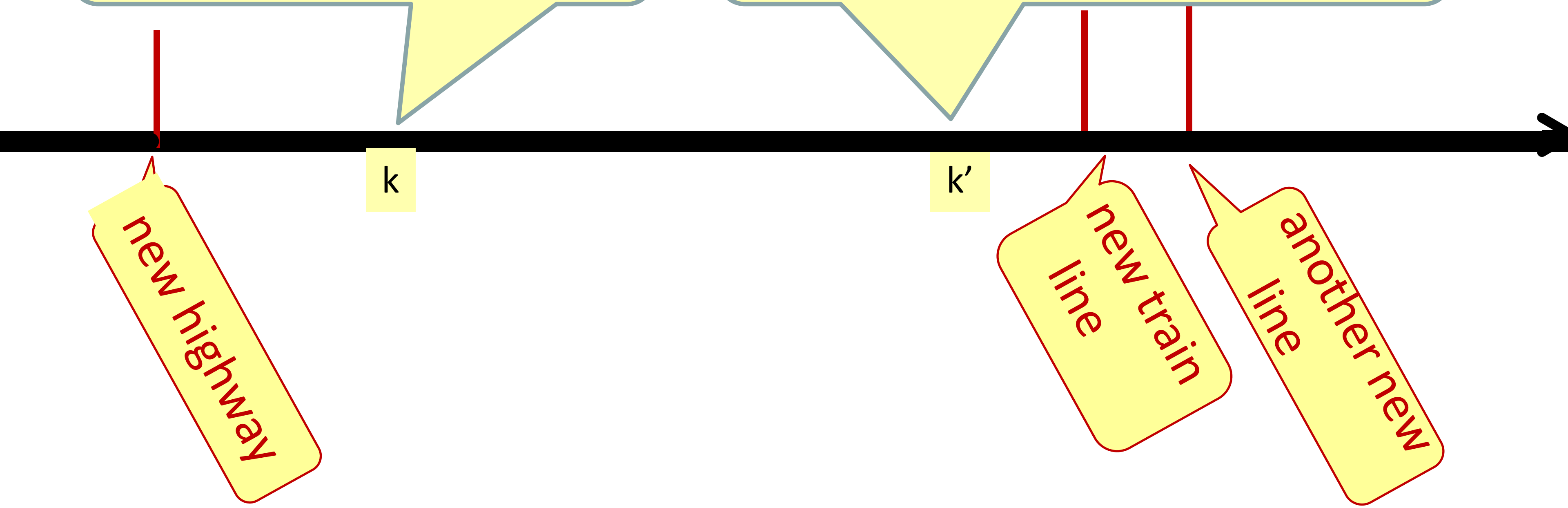
Example: Use of computing devices



Example: Repeated Rush-Hour Commute

Ex. predictability on day k: before observing the driving times, every driver assigns 99% to the 5 minutes ball around the driving times to be realized.

Ex. hindsight stability of chosen routes on day k': after observing the driving times, no player can gain more than 4 minutes by deviating from her chosen route.



Repeated Rush-Hour Commute

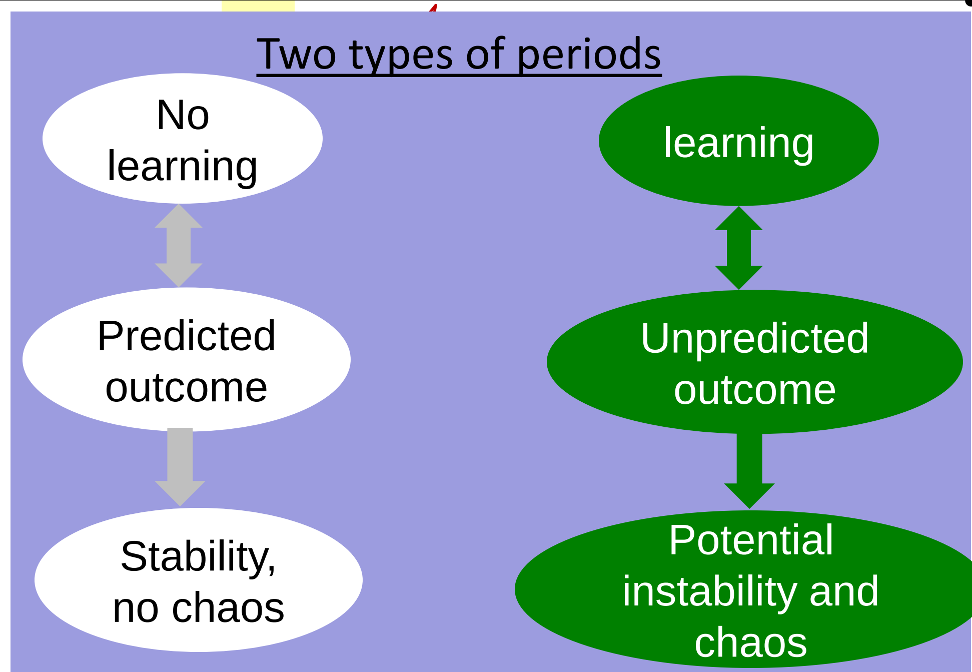
Ex. predictability on day k: before observing the driving times, every driver assigns 99% to the 5 minutes ball around the driving times to be realized.

Ex. hindsight stability of chosen routes on day k’: after observing the driving times, no player can gain more than 4 minutes by deviating from her chosen route.

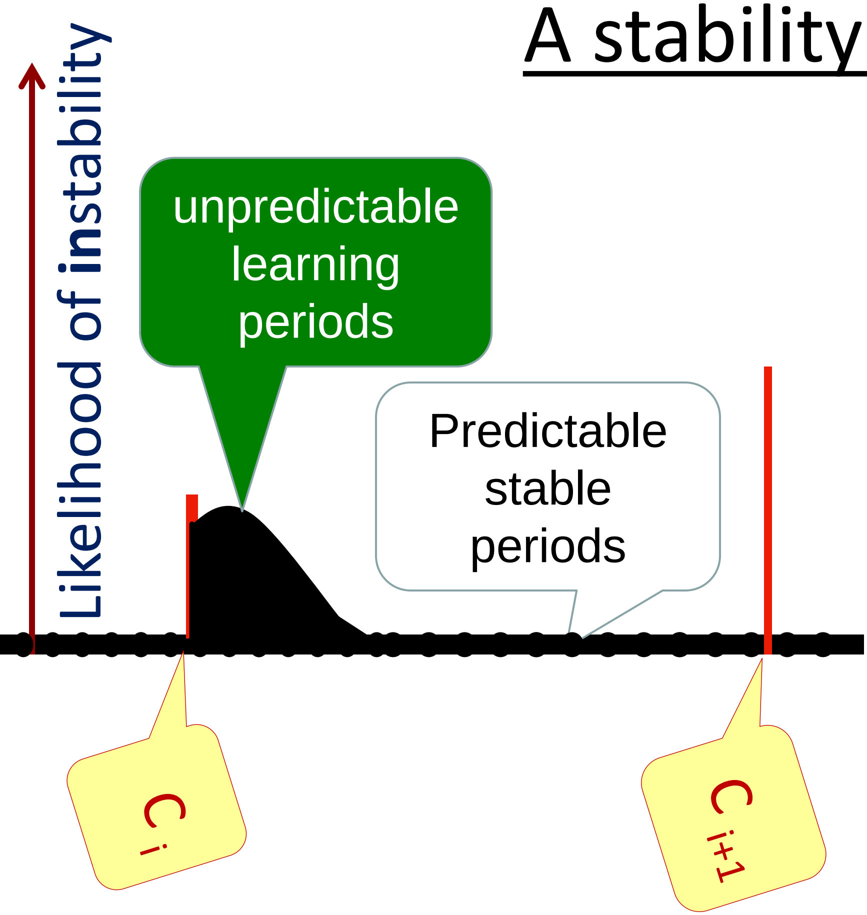
At equilibrium:
Predictability on day k implies hindsight stability on day k, but not the converse.

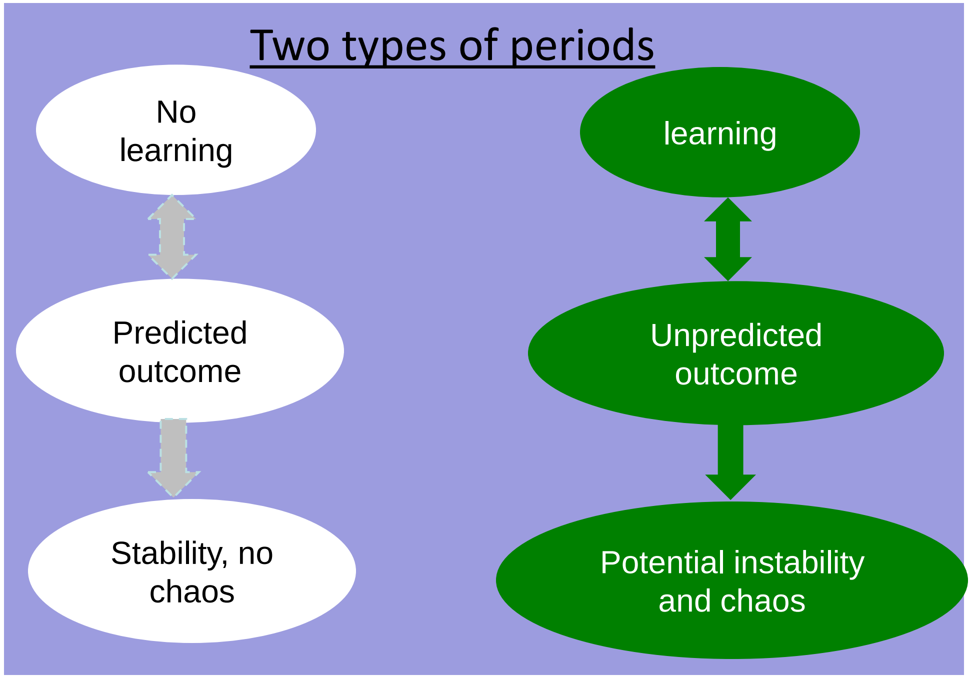
Thus, driving patterns on a day with unpredicted driving times is (potentially) unstable and chaotic.

Learning happens at the end of day k if and only if the observed driving times on this day were unpredicted.

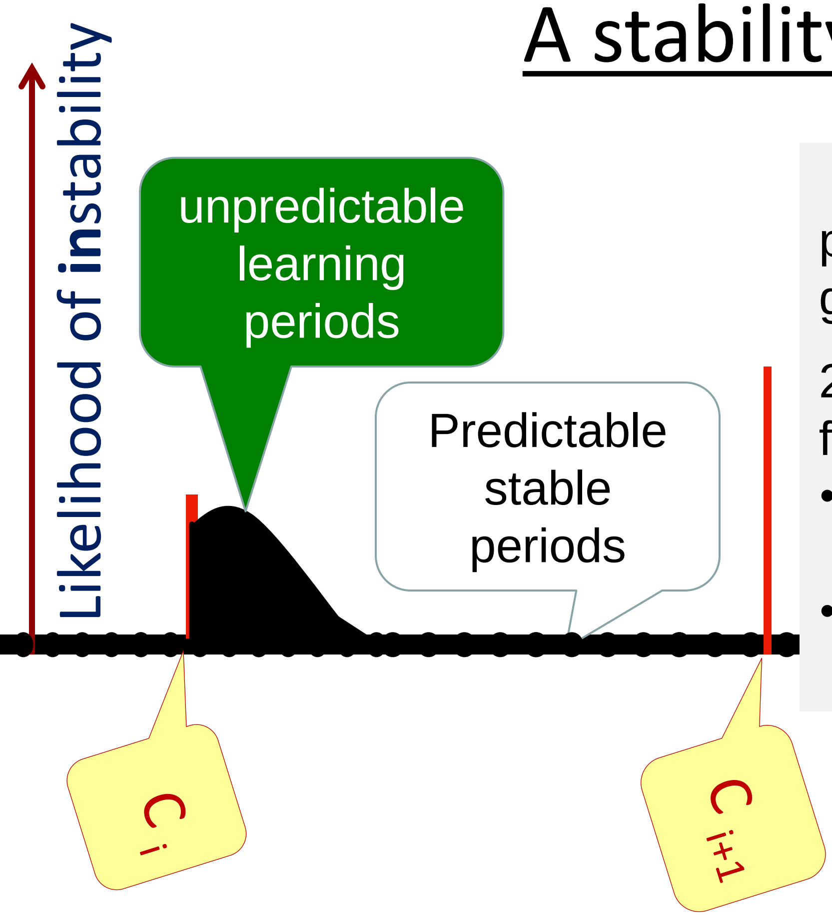


A stability cycle





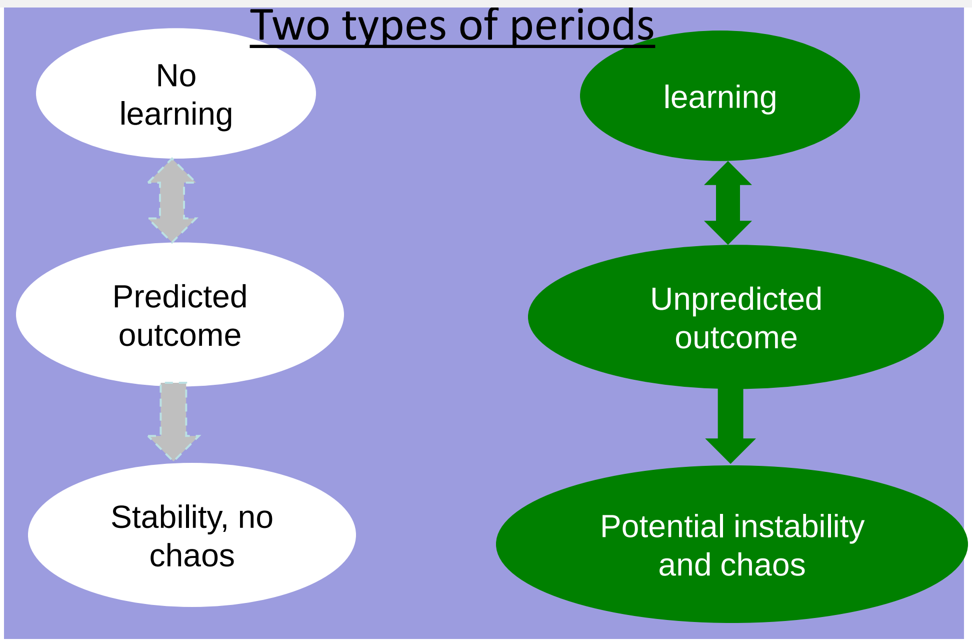
A stability cycle



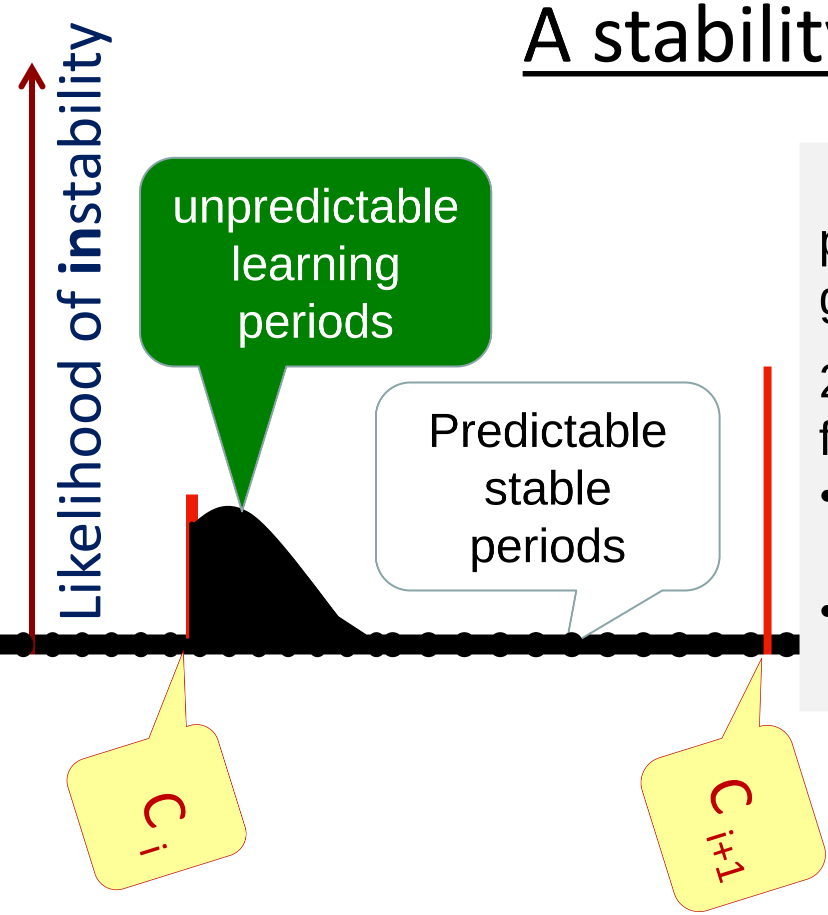
In any segment [C_i … C_{i+1})

1. the play admits a simple behavioral Markov-perfect myopic equilibrium of the infinitely repeated game.

2. The number of learning periods is bounded by a finite k that depends on:
- The accuracy of the players beliefs about the new parameters at C_i , **not at** C_{i+1} , and
- On the desired level of predictability and stability.



A stability cycle



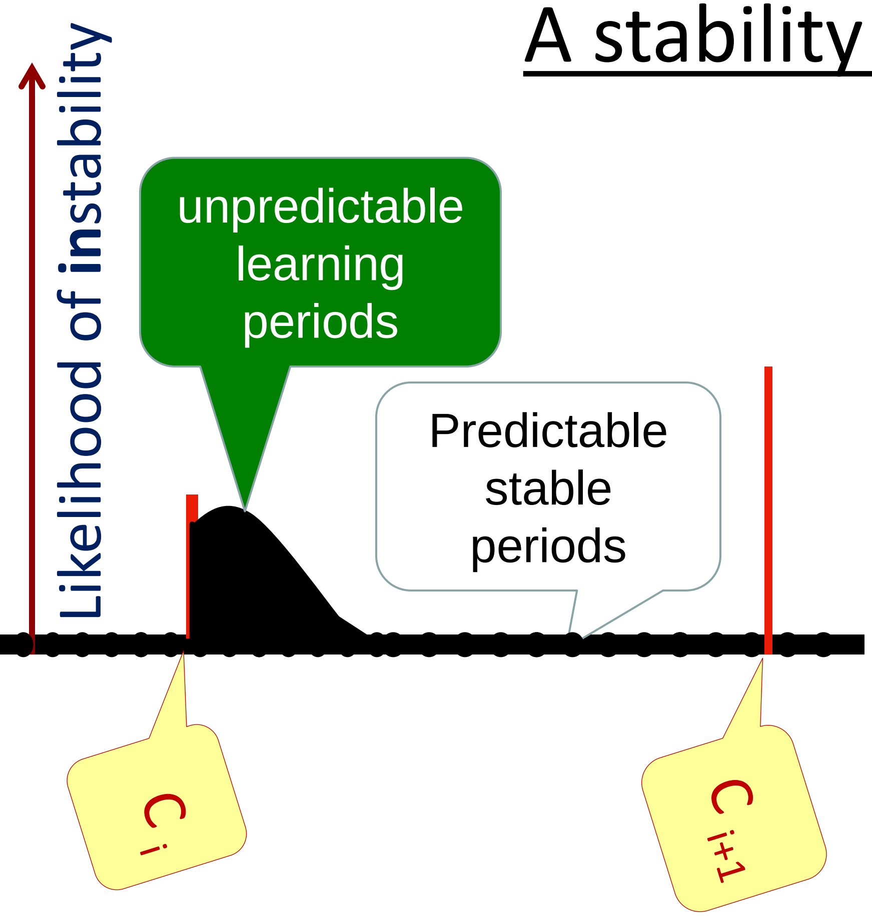
In any segment [C_i … C_{i+1})

1. the play admits a simple behavioral Markov-perfect myopic equilibrium of the infinitely repeated game.

2. The number of learning periods is bounded by a finite k that depends on:

- The accuracy of the players beliefs about the new parameters at C_i , **not at** C_{i+1} , and
- On the desired level of predictability and stability.

A stability cycle



Warning: not to be confused with the business cycle: the y-axis does not represent the quality of the period outcome, only the **inability to predict it.**

A stability cycle

The percentage of predictable stable periods increases with:

1. Lower external uncertainty in the outcome function,
2. Less frequent fundamental changes,
3. Players information about the new fundamentals.

Open questions

Many questions about big games with changing fundamentals, for example:

- What do players observe, if any, about the changing fundamentals?
- How to measure the level of changes that reflects on the number of learning periods?
- ….