Learning in Games

Nash equilibrium describes a situation in which players have identical and exactly correct beliefs about the strategies each player will choose. How and when might the players come to have correct beliefs, or at least beliefs that are close enough to being correct that the outcome corresponds to a Nash equilibrium? Fudenberg and Kreps 1993

- If critics understood economics they wouldn't criticize rationality and rational expectations, they would criticize the assumption of common knowledge
- I know what you are going to do and you know what I am going to do and I know that you know what I am going to do and so forth and so on
- a non-economic example: husband and a wife love each other, and both know that the other loves them, and each knows that the other knows that they are loved and so forth and so on

Common Knowledge

- Perhaps true in marriage
- true of stock traders?
- Rational expectations: we share the same beliefs
- Common knowledge: we have a mutual deep understanding of each others beliefs
- Neither one assumed by economists: a conclusion not an assumption

History of Game Theory in Economics

- Game theory's big impact in economics: late 1970s and early 1980s
- assumptions of rationality and common knowledge taken for granted
- late 1980s: inadequacy of the strong rationality assumptions underlying game theory creating discontent among game theorists
- 1988 Drew Fudenberg and David Kreps examined how equilibrium in games might arise from a process of boundedly rational learning rather than hyper-rational introspection
- two decades of research on learning in games since then

The Mystery in Human Learning

- not why people learn so badly – why they learn so well.
- behavioral economists, psychologists, economists and computer scientists model human learning using naïve and primitive models.
- models designed by computer scientists to make the best possible decisions cannot come close to the learning ability of the average human child, chimpanzee or even rat.
- equilibrium models and rational expectations: if we have to choose between best models of learning and perfect learning – for most situations of interest to economists perfect learning fits the facts better

What Learning Theory Tells Us

- you can only learn to the extent that you have experience or other data to learn from
- absent information no reason to imagine that people are unbiased, or do not exhibit “irrational” or “behavioral” modes of decision making
- incorrect beliefs are fundamental to learning theory – beliefs were always correct there would be nothing to learn about

Self Confirming Equilibrium

- beliefs are correct about things that players see
- beliefs may be incorrect about things they do not see
- in the context of an extensive form game: you only see the terminal node

Notation and Definitions

$s_i \in S_i$ pure strategies for i; $\sigma_i \in \Sigma_i$ mixed

H_i information sets for i

$\bar{H}(\sigma)$ reached with positive probability under σ

$\pi_i \in \Pi_i$ behavior strategies

$\hat{\pi}(h_i | \sigma_i)$ map from mixed to behavior strategies

μ_i a probability measure on Π_{-i}

$u_i(s_i | \mu_i)$ preferences

$$\Pi_{-i}(\sigma_{-i} \big| J) \equiv \{\pi_{-i} \big| \pi_i(h_i) = \hat{\pi}(h_i | \sigma_i), \forall h_i \in H_{-i} \cap J\}$$

Notions of Equilibrium

Nash equilibrium

a mixed profile σ such that for each $s_i \in \mathrm{supp}(\sigma_i)$ there exist beliefs μ_i such that

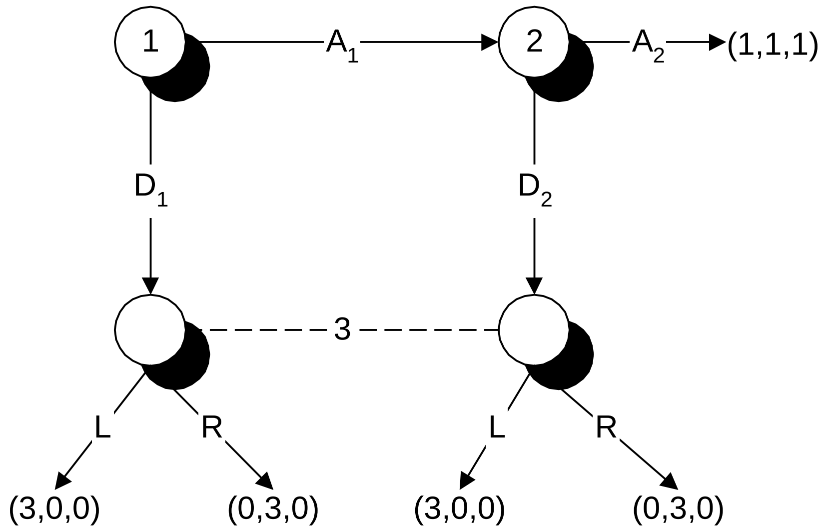
- s_i maximizes $u_i(\cdot|\mu_i)$
- $\mu_i(\Pi_{-i}(\sigma_{-i}|H)) = 1$

Unitary Self-Confirming Equilibrium

- $\mu_i(\Pi_{-i}(\sigma_{-i} \mid \bar{H}(\sigma))) = 1$

(=Nash with two players)

Fudenberg-Kreps Example



A_1, A_2 is self-confirming, but not Nash

any strategy for 3 makes it optimal for either 1 or 2 to play down

but in self-confirming, 1 can believe 3 plays R; 2 that he plays L

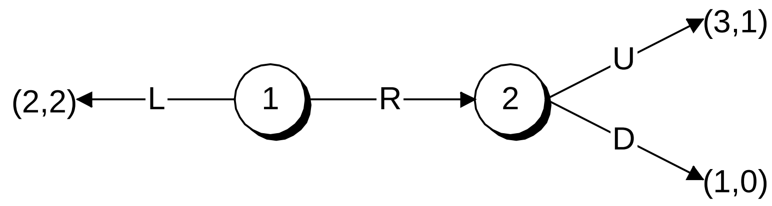
Heterogeneous Self-Confirming equilibrium

- $\mu_i(\Pi_{-i}(\sigma_{-i} \mid \bar{H}(s_i, \sigma))) = 1$

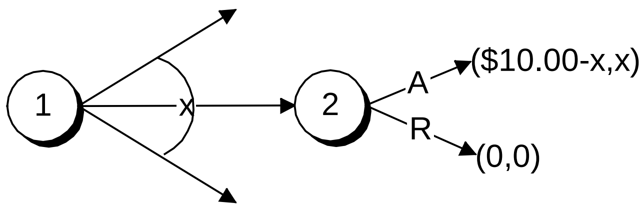
Can summarize by means of “observation function”

$$J(s_i, \sigma) = H, \bar{H}(\sigma), \bar{H}(s_i, \sigma)$$

Public Randomization



Ultimatum Bargaining Results



Raw US Data for Ultimatum

x	*Offers*	*Rejection Probability*
$2.00	1	100%
$3.25	2	50%
$4.00	7	14%
$4.25	1	0%
$4.50	2	100%
$4.75	1	0%
$5.00	13	0%
	27	

US $10.00 stake games, round 10

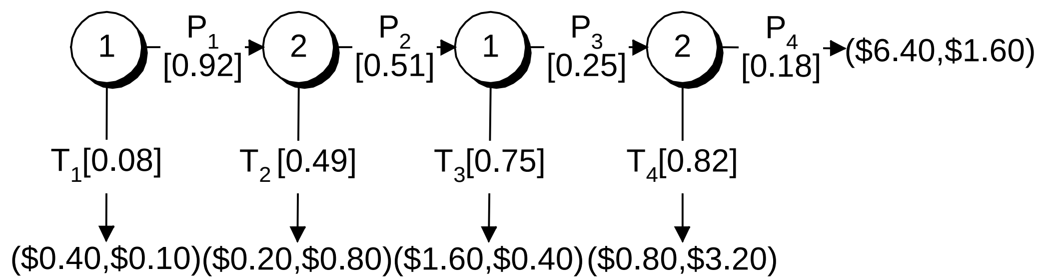
Trials	Rnd	Cntry Stake	Case	Expected Loss			Max Gain	Ratio
				Pl 1	Pl 2	Both		
27	10	US	H	$0.00	$0.67	$0.34	$10.00	3.4%
27	10	US	U	$1.30	$0.67	$0.99	$10.00	9.9%
10	10	USx3	H	$0.00	$1.28	$0.64	$30.00	2.1%
10	10	USx3	U	$6.45	$1.28	$3.86	$30.00	12.9%
30	10	Yugo	H	$0.00	$0.99	$0.50	$10?	5.0%
30	10	Yugo	U	$1.57	$0.99	$1.28	$10?	12.8%
29	10	Jpn	H	$0.00	$0.53	$0.27	$10?	2.7%
29	10	Jpn	U	$1.85	$0.53	$1.19	$10?	11.9%
30	10	Isrl	H	$0.00	$0.38	$0.19	$10?	1.9%
30	10	Isrl	U	$3.16	$0.38	$1.77	$10?	17.7%
	WC		H			$5.00	$10.00	50.0%

Rnds=Rounds, WC=Worst Case, H=Heterogeneous, U=Unitary

Comments on Ultimatum

- every offer by player 1 is a best response to beliefs that all other offers will be rejected so player 1's heterogeneous losses are always zero.
- big player 1 losses in the unitary case
- player 2 losses all knowing losses from rejected offers; magnitudes indicate that subgame perfection does quite badly
- tripling the stakes increases the size of losses a bit less than proportionally (losses roughly double).

Centipede Game: Palfrey and McKelvey



Numbers in square brackets correspond to the observed conditional probabilities of play corresponding to rounds 6-10, stakes 1x below.

This game has a unique self-confirming equilibrium; in it player 1 with probability 1 plays T_1

Trials/	Rnds	Stake	Case	Expected Loss			Max	Ratio
Rnd				Pl 1	Pl 2	Both	Gain	
29*	6-10	1x	H	$0.00	$0.03	$0.02	$4.00	0.4%
29*	6-10	1x	U	$0.26	$0.17	$0.22	$4.00	5.4%
	WC	1x	H			$0.80	$4.00	20.0%
29	1-10	1x	H	$0.00	$0.08	$0.04	$4.00	1.0%
10	1-10	4x	H	$0.00	$0.28	$0.14	$16.00	0.9%

Rnds=Rounds, WC=Worst Case, H=Heterogeneous, U=Unitary

Comments on Experimental Results

- heterogeneous loss per player is small; because payoffs are doubling in each stage, equilibrium is very sensitive to a small number of player 2's giving money away at the end of the game.
- unknowing losses far greater than knowing losses
- quadrupling the stakes very nearly causes $\bar{\varepsilon}$ to quadruple
- theory has substantial predictive power: see WC
- losses conditional on reaching the final stage are quite large--inconsistent with subgame perfection. McKelvey and Palfrey estimated an incomplete information model where some "types" of player 2 liked to pass in the final stage. This cannot explain many players dropping out early so their estimated model fits poorly

Summary: Social Preferences versus Learning

Many more losses due to incomplete learning than due to social preferences

Self-confirming Equilibrium and Economic Policy

- example adapted from Sargent, Williams and Zhao [2006a] by Fudenberg and Levine [2009]
- game between a government and typical or representative consumer
- first: government chooses high or low inflation
- second: consumers choose high or low unemployment
- consumers always prefer low unemployment
- government prefers low inflation to high inflation, but cares more about unemployment being low than about inflation
- “full” rationality (subgame perfection): the consumer will always choose low unemployment; government recognizing this will always choose low inflation

Self-confirming Equilibrium

- government believes incorrectly that low inflation leads to high unemployment
- widespread belief at one time
- care more about employment than inflation so keep inflation high
- never learn that beliefs about low inflation are false
- in practice information about consequences of low inflation generated by policy “mistakes” and random shocks
- Sargent, Williams and Zhao [2006a] use sophisticated dynamic model of learning to analyze how U.S. Federal Reserve policy evolved post World War II to ultimately result in the conquest of U.S. inflation
- In particular they explain why it took so long – a cautionary note for economic policy makers.

Self-Confirming Equilibrium and Economic Crises

- prior to current crisis Sargent, Williams and Zhao [2006b] examined series of Latin American crises
- assume that consumers have short-run beliefs that are correct, but have difficult correctly anticipating long run events (the collapse of a “bubble”)
- periodic crises arise as growth that is unsustainable in the long run takes place, but consumers cannot correctly foresee that far into the future

Who Got it Wrong?

- Economists anticipated this type of event, but understand the impossibility of predicting the timing
- Bankers cried all the way to the bank: if you can pay yourself bonuses during the upswing, and have the government cover your losses on the downswing – not much reason to worry about the business cycle

The Persistence of Superstition

What could be more irrational than superstition? If people are rational learners, won’t they learn that their superstitions are wrong? How can superstition persist in the face of evidence?

Example: code of Hammurabi

> *If any one bring an accusation against a man, and the accused go to the river and leap into the river, if he sink in the river his accuser shall take possession of his house. But if the river prove that the accused is not guilty, and he escape unhurt, then he who had brought the accusation shall be put to death, while he who leaped into the river shall take possession of the house that had belonged to his accuser.* 2nd law of Hammurabi

The Puzzle

- based on a superstition that we do not believe to be true – we do not believe that the guilty are any more likely to drown than the innocent
- if people can be easily persuaded to hold a superstitious belief, why such an elaborate mechanism? Why not simply assert that those who are guilty will be struck dead by lightning?

Three Games

First player (culprit) decides whether or not to commit a crime

Prefers to commit the crime if unpunished but not if he expects to be punished

Hammurabi game: if crime is commited a second player (witness), must decide whether or not to correctly identify the culprit

> Witness prefers to identify his worst enemy rather than the true culprit
>
> after testifying the criminal is tossed in the river – and most likely drowns

Game without the river: witness points the finger and whoever she accuses is punished

Lightning game: no witness, culprit – regardless of whether or not the crime was committed – has small chance of being struck dead by lightning

Superstitions that Inhibit Crime

Hammurabi game: witness believes accuser will survive if she lies and drown if she tells the truth (she is wrong, he will drown in both cases)

her beliefs lead her to tell truth; knowing this culprit does not commit the crime

Game without the river: culprit believes that witness will tell truth (he is wrong: without any chance of punishment, the witness will lie and identify her enemy)

His belief leads him not to commit the crime

Lightning game: culprit believes if he commits a crime he will be struck dead by lightning (he is wrong: he is very unlikely to be struck by lightning regardless of whether he commits a crime)

His belief leads him not to commit the crime

Which Superstitions will Survive? Lightning Game

not Nash: culprit should know probability of being struck dead by lightning doesn't depend on whether a crime is committed

self-confirming: culprit chooses not to commit the crime, belief that he will be struck dead by lightning is purely hypothetical and is never confronted with evidence

- self-confirming equilibrium is a reasonable description of short-run behavior
- an unlikely basis for a social norm
- over time people will commit crimes for one reason or another and will not be struck dead by lightning
- would not expect "being struck dead by lightning" to be basis of criminal justice systems, and historically it does not seem to be

Which Superstitions will Survive? Game Without the River

Nash equilibrium for culprit not to commit the crime and witness to tell the truth

witness is never called upon to testify so (by Nash logic) indifferent between telling the truth and lying

- no more plausible in the long-run than the lightning equilibrium
- if people do occasionally commit crimes, the witness – being in fact called upon to testify – will lie, and eventually the superstition that witnesses tell the truth should die in the face of evidence

Equilibrium is not subgame perfect

Subgame Confirmed Equilibrium

- when people are patient enough to try crimes to see if they “can get away with it” kind of equilibrium that results ia hybrid between subgame perfection and self-confirming equilibrium called subgame confirmed equilibrium
- deviations from a Nash equilibrium into a subgame should result in a self-confirming equilibrium

Which Superstitions will Survive? Hammurabi Game

- culprit commits crime rarely to “see if he can get away with it”
- can’t get away with it since the witness will tell the truth
- only commits crime to verify this fact
- witness tells truth because she superstitiously believes it gives her a better chance of accused drowing in river
- why does not witness lie occasionally to see if “she can get away with it?”
- experimenting is an investment in the future; if you verify you are right you go back to what you are doing and suffer a one period loss; if you find you were wrong, then you make a gain forever after – a future benefit

Why Invest in the Future?

- Invest when patient and the delay is not too long
- Culprit immediate rewards: you discover you can get away with murder, you can start on a life of crime straight away
- Witness not called to testify very often
- benefit of “trying it out to see” is that at some far distant date when called upon to testify again can again lie
- not a good investment

superstition about something that lies “off the equilibrium path”

in other words: superstition about something that happens very infrequently when the social norm is adhered to

according to learning theory: “Hammurabi had it exactly right: (our simplified interpretation of) his law uses the greatest amount of superstition consistent with patient rational learning”

Self-Confirming, Nash Equilibrium and Agreeing to Disagree

no-trade theorem: with underlying common knowledge and in the absence of other reasons to trade, people should not trade merely based on informational differences

when you offer to bet that a particular horse will win at the race-track I should refuse the bet on the grounds that the only reason you are willing to make the offer is because you know something I don't know

- not very realistic it seems, so reject common knowledge?
- learning theory alone gives rise to the no-trade theorem
- if on average lose by trading some of us must be losers – eventually the losers should quite
- self-confirming equilibrium also says no-trade

Which Assumption is Violated?

Rationality?

No reason to trade?

- At the race-track some people bet because they enjoy betting and happily lose money
- That provides an incentive for everyone else to bet based on information differences
- Or: there is a sucker born every minute and new ones show up to take the slack

Keynes Beauty Contest

His explanation of how stock markets work

- players must choose the most beautiful woman from six photographs
- players who pick the most popular face win

a boring coordination game: every face is a Nash equilibrium

Keynes made a fortune for King’s College Cambridge through his stock market investments; also lost it periodically, so he was lucky to quit while ahead

Nagel Experiment

choose a number between 0 and 100.

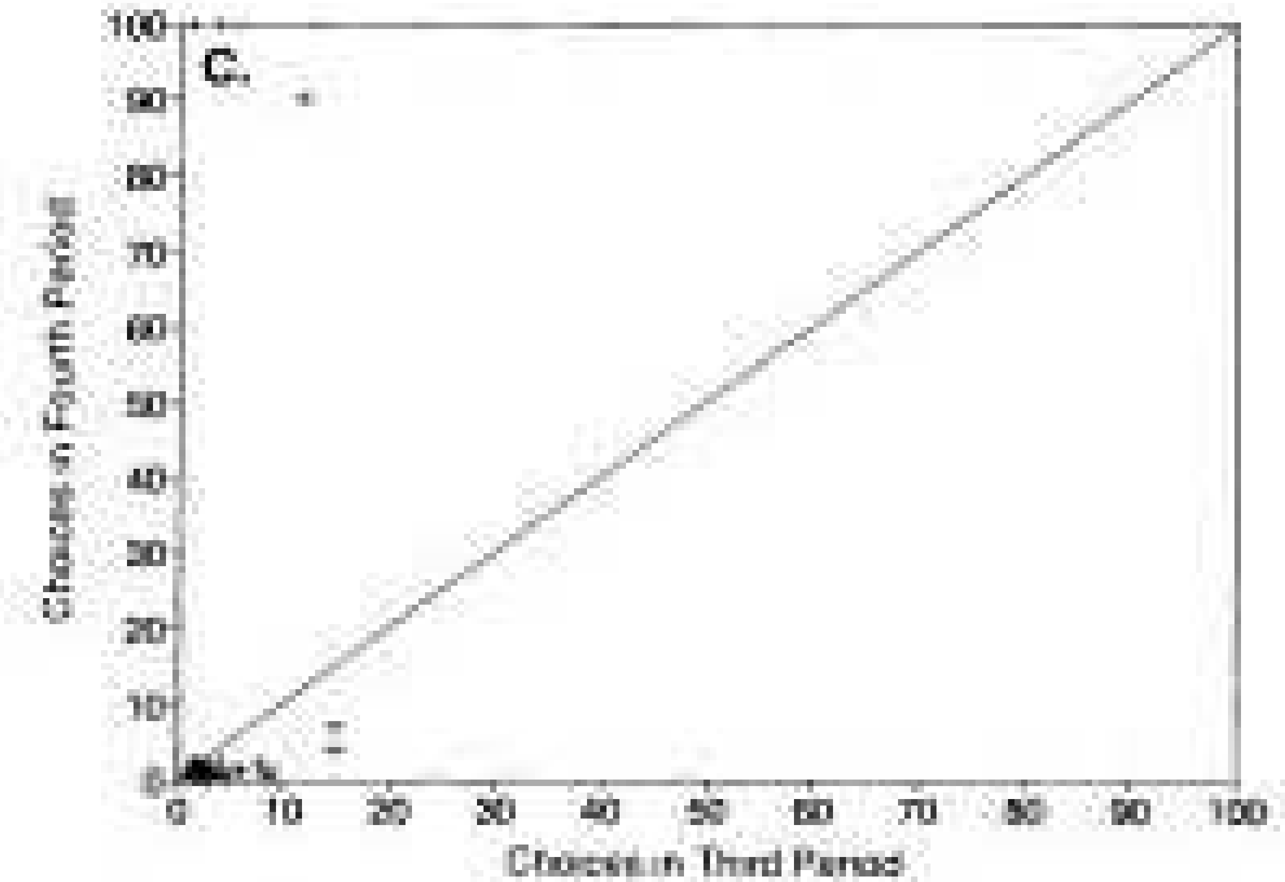
players closest to half the average value win

what you want to do depends on what you think average opinion is

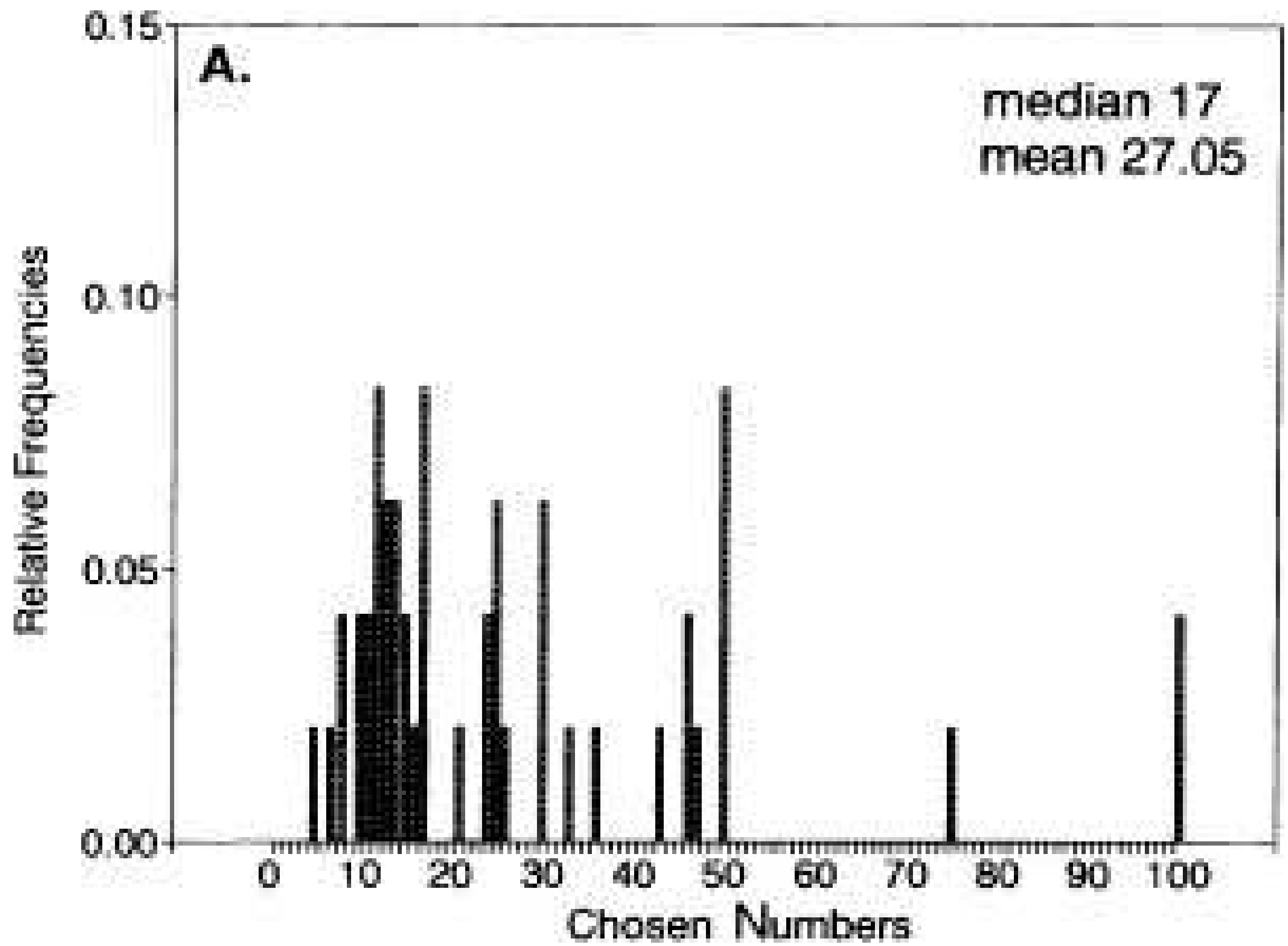
if you think that people choose randomly, average should be 50, so guess 25

(almost) a unique Nash equilibrium: zero and one

Third and Fourth Plays



First Time Play



Level-k Theory

One time play: not Nash equilibrium

Not expected to be Nash since one-time play

another theory does do a good job of explaining what is going on

originates with Nagel, developed by Stahl and Wilson [1994], Costa-Gomes., Crawford. and Broseta [2001] and Camerer, Ho and Chong [2004] and others

people differ in their sophistication

- naïve individuals – level-0 – play randomly
- less naïve individuals – level-1 – believe that their opponents are of level-0
- in general higher level and more sophisticated individuals – level-*k* – believe that they face a mixture of less sophisticated individuals – people with lower levels of *k*

Summary

- single common probability belief about relative likelihood of degree of sophistication can explain first time play in a variety of experimental games
- idea has yet to gain wide traction in economics
- types of games in which the theory has been shown to work are relatively simple and unlike the kinds of situations economists are interested in
- in real markets – stock markets for example – participants are generally relatively experienced, so more likely to exhibit equilibrium behavior than level-*k* behavior
- modern finance: “noise traders” (inexperienced,naïve, small in numbers) play important role both in transmission of information through prices and in price fluctuations; level-*k* theory may enable us to better model them

Learning Theory versus Behavioral Economics

Behavioral economics: we are most ignorant about the things we are the most familiar with

- With lifetime of evidence that we are procrastinators, we stubbornly stick to the belief that we are not
- Despite complete ignorance of the Becker-Degroot-Marschak elicitation procedure and absence of experience with second price auctions, we comprehend that procedure perfectly

Learning theory: we are most likely to be mistaken about are the things we know the least about

- common sense and supported by overwhelming evidence
- law of gravity is common knowledge based on lifetime of evidence
- broad disagreement about possibility and nature of life after death based on total absence of data