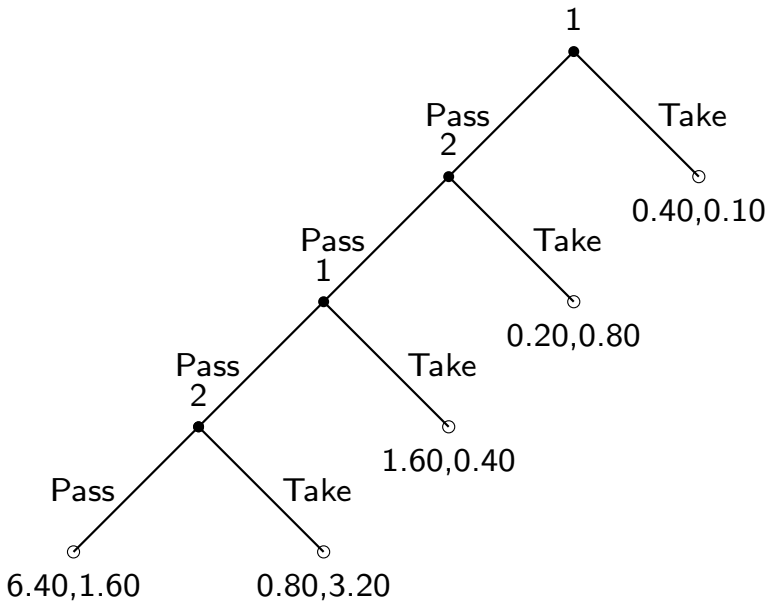
Centipede Game and Agent Quantal Response Equilibrium

Niels Boissonnet

May 2020

Extensive Form Representation



1. P1 chooses Take (and the payoffs are $0.40, 0.10$) or Pass...
2. If P1 chose Pass, then P2 can choose Take (and the payoffs are $0.20, 0.80$) or Pass...
3. If P2 chose Pass, then P1 can choose Take (and the payoffs are $1.60, 0.40$) or Pass...
4. If P1 chose Pass, then P2 can choose Take (and the payoffs are $0.80, 3.20$) or Pass (and the payoffs are $6.40, 1.60$)

Backward induction implies that P1 should choose to Take and the play stops after the first node .

Paper 1:
An Experimental Study of the Centipede Game

McKelvey and Palfrey (1992)

The experiment

- They ask subjects to participate in ten centipede games of either two rounds (four moves) or of three rounds (six moves).
- Before each game subjects were matched with a random subjects (they do not have the same opponent from one game to the other).

Motivations

- Backward induction yields a solution that is intuitively unsatisfactory, so it is worth observing experimentally how subjects behave in this game,
- Explaining deviation with
  1. The existence of an altruistic player;
  2. The fact that some DMs make mistakes.

Results: descriptive statistics

Let

- f_i be the proportion of games ending at the ith node;
- p_i is the probability of stopping the game

$$p_i = \frac{f_i}{\sum_{j=i}^{n+1} f_j}$$

TABLE IIA

PROPORTION OF OBSERVATIONS AT EACH TERMINAL NODE

	Session		N	f_1	f_2	f_3	f_4	f_5
Four Move	1	(PCC)	100	.06	.26	.44	.20	.04
	2	(PCC)	81	.10	.38	.40	.11	.01
	3	(CIT)	100	.06	.43	.28	.14	.09
	Total	1–3	281	.071	.356	.370	.153	.049
High Payoff	4	(High-CIT)	100	.150	.370	.320	.110	.050
Six Move	5	(CIT)	100	.02	.09	.39	.28	.20
	6	(PCC)	81	.00	.02	.04	.46	.35
	7	(PCC)	100	.00	.07	.14	.43	.23
	Total	5–7	281	.007	.064	.199	.384	.253

TABLE IIB[a]

IMPLIED TAKE PROBABILITIES FOR THE CENTIPEDE GAME

	Session	p_1	p_2	p_3	p_4	p_5	p_6
Four Move	1 (PCC)	.06 (100)	.28 (94)	.65 (68)	.83 (24)		
	2 (PCC)	.10 (81)	.42 (73)	.76 (42)	.90 (10)		
	3 (CIT)	.06 (100)	.46 (94)	.55 (51)	.61 (23)		
	Total 1–3	.07 (281)	.38 (261)	.65 (161)	.75 (57)		
High Payoff	4 (CIT)	.15 (100)	.44 (85)	.67 (48)	.69 (16)		
Six Move	5 (CIT)	.02 (100)	.09 (98)	.44 (89)	.56 (50)	.91 (22)	.50 (2)
	6 (PCC)	.00 (81)	.02 (81)	.04 (79)	.49 (76)	.72 (39)	.82 (11)
	7 (PCC)	.00 (100)	.07 (100)	.15 (93)	.54 (79)	.64 (36)	.92 (13)
	Total 5–7	.01 (281)	.06 (279)	.21 (261)	.53 (205)	.73 (97)	.85 (26)

Results: descriptive statistics

- *Backward induction is not satisfied.*
- The probability of stopping the Game is increasing.

TABLE IIA

PROPORTION OF OBSERVATIONS AT EACH TERMINAL NODE

	Session		N	f_1	f_2	f_3	f_4	f_5
Four Move	1	(PCC)	100	.06	.26	.44	.20	.04
	2	(PCC)	81	.10	.38	.40	.11	.01
	3	(CIT)	100	.06	.43	.28	.14	.09
	Total	1–3	281	.071	.356	.370	.153	.049
High Payoff	4	(High-CIT)	100	.150	.370	.320	.110	.050
Six Move	5	(CIT)	100	.02	.09	.39	.28	.20
	6	(PCC)	81	.00	.02	.04	.46	.35
	7	(PCC)	100	.00	.07	.14	.43	.23
	Total	5–7	281	.007	.064	.199	.384	.253

TABLE IIB[a]

IMPLIED TAKE PROBABILITIES FOR THE CENTIPEDE GAME

	Session	p_1	p_2	p_3	p_4	p_5	p_6
Four Move	1 (PCC)	.06 (100)	.28 (94)	.65 (68)	.83 (24)		
	2 (PCC)	.10 (81)	.42 (73)	.76 (42)	.90 (10)		
	3 (CIT)	.06 (100)	.46 (94)	.55 (51)	.61 (23)		
	Total 1–3	.07 (281)	.38 (261)	.65 (161)	.75 (57)		
High Payoff	4 (CIT)	.15 (100)	.44 (85)	.67 (48)	.69 (16)		
Six Move	5 (CIT)	.02 (100)	.09 (98)	.44 (89)	.56 (50)	.91 (22)	.50 (2)
	6 (PCC)	.00 (81)	.02 (81)	.04 (79)	.49 (76)	.72 (39)	.82 (11)
	7 (PCC)	.00 (100)	.07 (100)	.15 (93)	.54 (79)	.64 (36)	.92 (13)
	Total 5–7	.01 (281)	.06 (279)	.21 (261)	.53 (205)	.73 (97)	.85 (26)

Results: descriptive statistics

- In the five first games subjects are more altruist than in the five last: Is there some learning here?

The number of subjects that adopt the dominated strategy of passing on the last move drops from 25% to 15%.

TABLE IIIB

IMPLIED TAKE PROBABILITIES

COMPARISON OF EARLY VERSUS LATE PLAYS IN THE LOW PAYOFF CENTIPEDE GAMES

Treatment	Game	p_1	p_2	p_3	p_4	p_5	p_6
Four Move	1–5	.06 (145)	.32 (136)	.57 (92)	.75 (40)		
	6–10	.08 (136)	.49 (125)	.75 (69)	.82 (17)		
Four Move	1–5	.00 (145)	.06 (145)	.18 (137)	.43 (112)	.75 (64)	.81 (16)
	6–10	.01 (136)	.07 (134)	.25 (124)	.65 (93)	.70 (33)	.90 (10)

Results: descriptive statistics

- Some individuals (5%) passed at every opportunity they have;
- Some subjects' behavior does not seem to make sense indicating they make mistakes.

The model

Incomplete information: The opponent is

The model

Incomplete information: The opponent is

- selfish with probability q

The model

Incomplete information: The opponent is

- selfish with probability q
- altruist with probability $1 - q$

The model

Incomplete information: The opponent is

- selfish with probability q
- altruist with probability $1 - q$

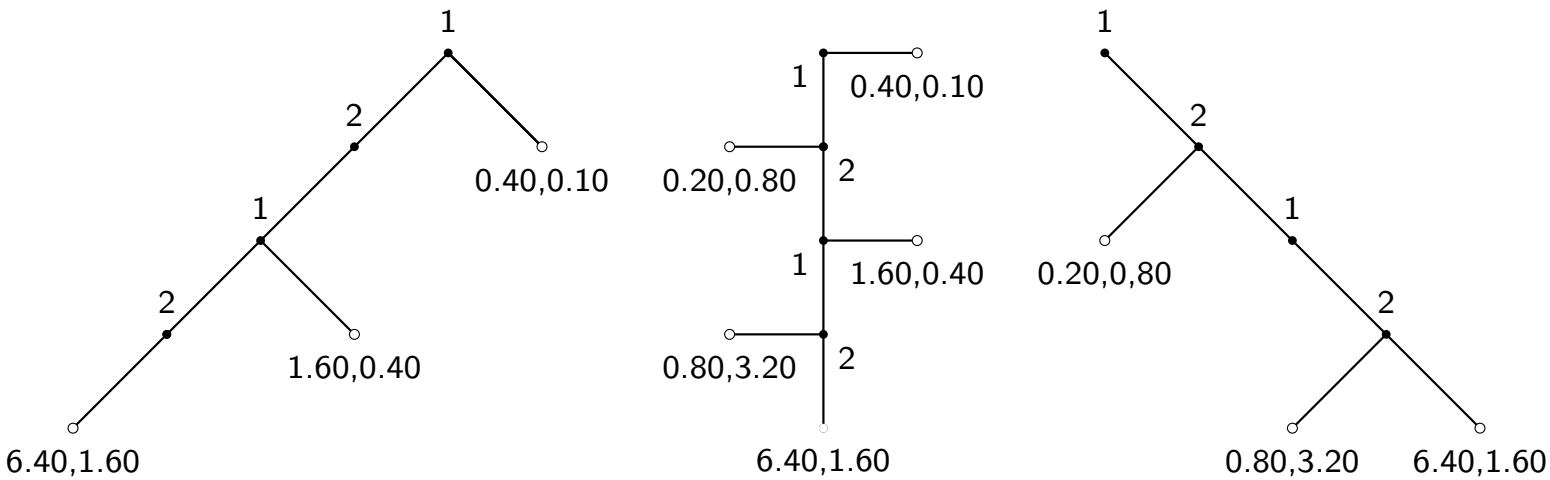


Figure: Extensive form representation

The model

Incomplete information: The opponent is

- selfish with probability q
- altruist with probability $1 - q$

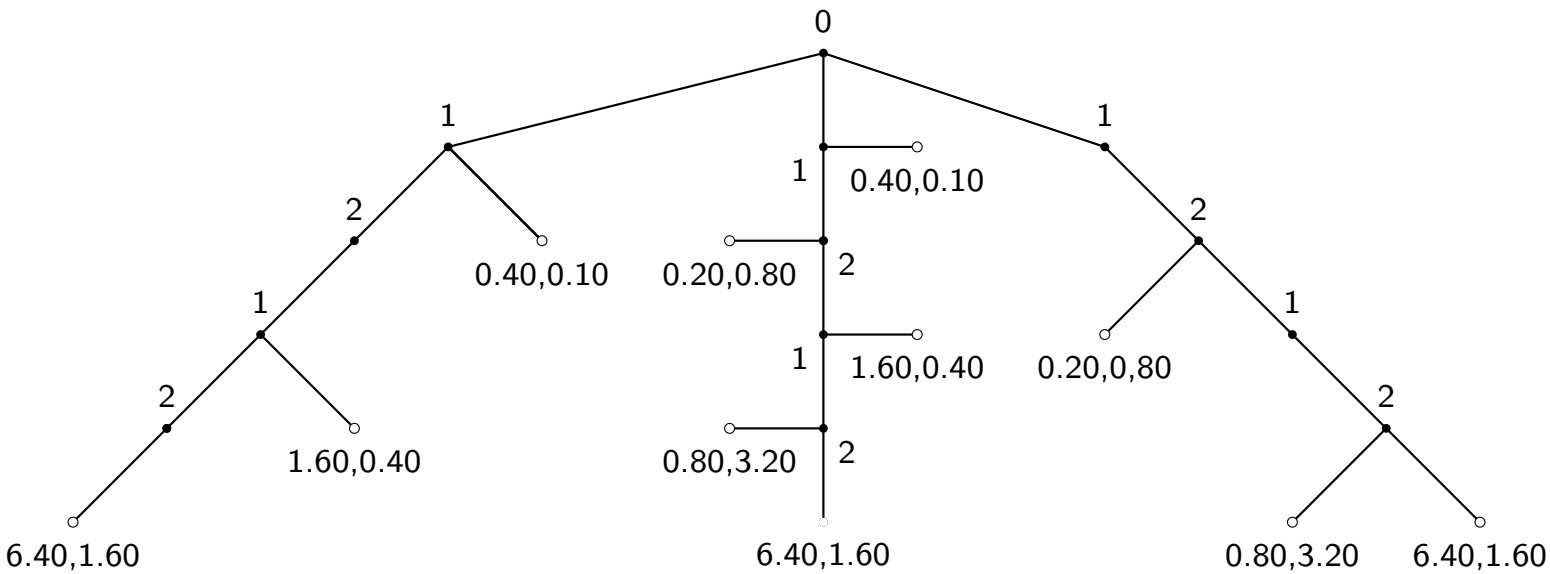


Figure: Extensive form representation

The model

Incomplete information: The opponent is

- selfish with probability q
- altruist with probability $1 - q$

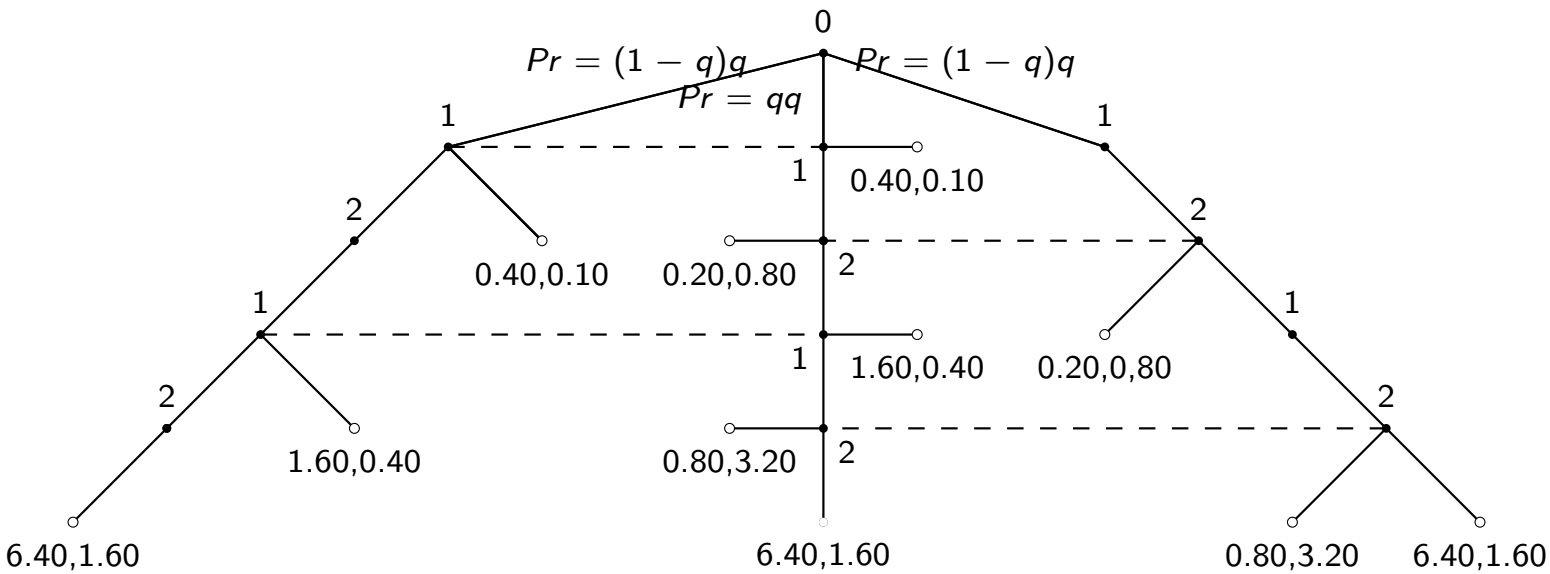


Figure: Extensive form representation

Proposition

Let q the common belief that a selected player is selfish. For a generic q there exists a sequential equilibrium denoted $p(q)$ with the following properties.

1. *For any q Player 2 chooses TAKE with probability 1 on its last move.*
2. *if $q = 1$ they choose TAKE*
3. *If $1 - q > 1/7$, both Players always choose PASS, except on the last move*
4. *If $1 - q < 1/7$ the equilibrium involves mixed strategies.*

The form of the imitation in our setting is obvious: selfish players sometimes pass, to mimic an altruist.

Limitations

For any value of q, there is at least one outcome with a 0 or close to 0 probability of occurrence.

- So, the model will fit poorly data in which all of the possible outcomes occur.
- Nor can it account for any consistent patterns of learning in the data, for substantial variations across individuals, or for some of the irregularities described earlier.

Limitations

For any value of q, there is at least one outcome with a 0 or close to 0 probability of occurrence.

- So, the model will fit poorly data in which all of the possible outcomes occur.
- Nor can it account for any consistent patterns of learning in the data, for substantial variations across individuals, or for some of the irregularities described earlier.

To account for these features of the data, they introduce two additional elements:

- *The subject explore the experiment to see what happens;*
- *Subjects have heterogeneous beliefs about their opponents' type.*

The subject explore the experiment to see what happens

- **Noisy play:** In game t, if p^* is the equilibrium probability of TAKE that the player at that node attempts to implement, we assume that the player actually chooses TAKE with probability $(1-\epsilon_t)p^*$, and makes a random move (i.e. TAKE or PASS with probability 0.5) with probability ϵ_t.
- **Learning:** Individuals follow an exponential learning curve with δ as the initial erro rate.

$$\epsilon_t = \epsilon e^{-\delta(t-1)}$$

Subjects have heterogeneous beliefs about the type.

- Q the true probability that individuals are selfish but that each player i has a belief, q_i that her opponent is a selfish players, where $q_i \neq Q$ is possible;
- Beliefs of the individuals are independently drawn from a Beta distribution with parameters (α, β), where the mean of the distribution is simply equal to $q = \frac{\alpha}{\alpha+\beta}$;
- **Egocentric model hypothesis:** Each player plays the game as if it were common knowledge that the opponent had the same belief, i.e., while the econometrician assumes there is heterogeneity in beliefs, players do have heterogeneous beliefs, but believe that everyone's beliefs are alike;
- **Rational expectations hypothesis:** The average belief $\frac{\alpha}{\alpha+\beta}$ equals the true proportion of selfish players.

Estimation

If Q is the true population parameter for the fraction of selfish players, then the likelihood of observing i choosing D_i, conditionally on q is

$$Pr_i(D_i|Q,\epsilon,\delta,q) = QPr_i^{selfish}(D_i|\epsilon,\delta,q) + (1-Q)Pr_i^{altruist}(D_i|\epsilon,\delta,q)$$

Since q is drawn from the Beta distribution with parameters (α,β), and density $B(q|\alpha,\beta)$, then the likelihood of observing D_i without conditioning on q is given by:

$$s_i(D_i|Q,\epsilon,\delta,\alpha,\beta) = \int_0^1 \pi_i(D_i|Q,\epsilon,\delta,q).B(q|\alpha,\beta).dq$$

Thus, to estimate your parameters you maximize

$$L(D|Q,\epsilon,\delta,\alpha,\beta) = \sum_{i=1}^{I} log\left(s_i(D_i|Q,\epsilon,\delta,\alpha,\beta)\right)$$

Tested hypotheses

- **Errors in action:** Is ϵ different from 0?
- **Heterogeneity of beliefs:** Is the variance of the estimated distribution of priors significantly different from 0 ?
- **Rational Expectations:** Is the estimated value of Q equal to the mean of the estimated distribution of priors, $\frac{\alpha}{\alpha+\beta}$?
- **Learning:** Is the estimated value of 8 positive and significantly different from 0?

Estimation Results and Tests

Notations :

- $\hat{\mu} = \frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}}$
- $\sigma =$ variance of the distributed priors beliefs

TABLE IV

RESULTS FROM MAXIMUM LIKELIHOOD ESTIMATION[a]

Treatment	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{Q}$	$\hat{\varepsilon}$	$\hat{\delta}$	$-\ln L$
unconstrained	42	2.75	.939	.956	.18	.045	327.35
$\mu = Q$	44	2.75	.941	.941	.18	.045	327.41
$\delta = 0$	68	2.50	.965	.950	.21	.000	345.08
$\sigma = 0$	—	—	.972	.850	.23	.020	371.04

TABLE V

LIKELIHOOD RATIO TESTS

	Hypothesis	Treatment	d.f.	−2 log likelihood ratio
Theoretical Hypotheses	Heterogeneity ($\sigma = 0$)	4-move	1	87.38*
		6-move	1	181.78*
	Rational expectations ($\mu = Q$)	4-move	1	.12
		6-move	1	1.38
	Learning ($\delta = 0$)	4-move	1	35.46*
		6-move	1	37.88*

Paper 2:
Quantal Response Equilibria for Extensive Form Games

McKelvey and Palfrey (1998)

Agent quantal response equilibrium (AQRE)

Agent quantal response equilibrium (AQRE)

Generalization of AQRE to extensive forms McKelvey and Palfrey (1998).

- including games with incomplete information

Agent quantal response equilibrium (AQRE)

Generalization of AQRE to extensive forms McKelvey and Palfrey (1998).

- including games with incomplete information

Basic idea

- I_i denotes an information set of player i, σ denotes a completely mixed strategy profile, $\pi_i(s_i|\sigma)$ denotes the expected payoff of action $s_i \in S_i(I_i)$

Agent quantal response equilibrium (AQRE)

Generalization of AQRE to extensive forms McKelvey and Palfrey (1998).

- including games with incomplete information

Basic idea

- I_i denotes an information set of player i, σ denotes a completely mixed strategy profile, $\pi_i(s_i|\sigma)$ denotes the expected payoff of action $s_i \in S_i(I_i)$
- We assume that the DM observes a disturbed payoff $\hat{\pi}_i(s_i|\sigma) = \pi_i(s_i|\sigma) + \epsilon_{s_i}$ where ϵ_{s_i} is a random variable, privately observed by the DM, anbd independently distributed among agents.
- Then, a strategy is σ^* is an *agent quantal response equilibrium* if for all i, and all s_i

$$\sigma^*(s_i) = \mathbb{P}\left(\hat{\pi}_i(s_i|\sigma^*) \geq \hat{\pi}_i(s_i'|\sigma^*), \forall s_i' \in S_i(I_i)\right) =$$

$$\mathbb{P}\left(\pi_i(s_i|\sigma^*) - \pi_i(s_i'|\sigma^*) \geq \epsilon_{s_i'} - \epsilon_{s_i}, \forall s_i' \in S_i(I_i)\right)$$

Logit AQRE

- σ is an Agent logit equilibrium if ϵ_{s_i} is a Weibull Distribution with cumulative distribution function $F(\epsilon_{s_i}) = e^{-e^{\lambda s_i}}$
- This implies (see eg., McFadden, 1975) that for all actions in all information sets

$$\sigma(s_i) = \frac{\exp\{\lambda \cdot \pi_i(s_i|\sigma)\}}{\sum_{s_i' \in S_i(I_i)} \exp\{\lambda \cdot \pi_i(s_i'|\sigma)\}}$$

Logit AQRE

- σ is an Agent logit equilibrium if ϵ_{s_i} is a Weibull Distribution with cumulative distribution function $F(\epsilon_{s_i}) = e^{-e^{\lambda s_i}}$
- This implies (see eg., McFadden, 1975) that for all actions in all information sets

$$\sigma(s_i) = \frac{\exp\{\lambda \cdot \pi_i(s_i|\sigma)\}}{\sum_{s_i' \in S_i(I_i)} \exp\{\lambda \cdot \pi_i(s_i'|\sigma)\}}$$

Properties

Logit AQRE

- σ is an Agent logit equilibrium if ϵ_{s_i} is a Weibull Distribution with cumulative distribution function $F(\epsilon_{s_i}) = e^{-e^{\lambda s_i}}$
- This implies (see eg., McFadden, 1975) that for all actions in all information sets

$$\sigma(s_i) = \frac{\exp\{\lambda \cdot \pi_i(s_i|\sigma)\}}{\sum_{s_i' \in S_i(I_i)} \exp\{\lambda \cdot \pi_i(s_i'|\sigma)\}}$$

Properties

- For $\lambda < \infty$, all actions are chosen with positive probability, thus there are no information sets off the path, equilibrium refinement is implied

Logit AQRE

- σ is an Agent logit equilibrium if ϵ_{s_i} is a Weibull Distribution with cumulative distribution function $F(\epsilon_{s_i}) = e^{-e^{\lambda s_i}}$
- This implies (see eg., McFadden, 1975) that for all actions in all information sets

$$\sigma(s_i) = \frac{\exp\{\lambda \cdot \pi_i(s_i|\sigma)\}}{\sum_{s_i' \in S_i(I_i)} \exp\{\lambda \cdot \pi_i(s_i'|\sigma)\}}$$

Properties

- For $\lambda < \infty$, all actions are chosen with positive probability, thus there are no information sets off the path, equilibrium refinement is implied
- As λ approaches infinity, convergence to a sequential equilibrium

Logit AQRE

- σ is an Agent logit equilibrium if ϵ_{s_i} is a Weibull Distribution with cumulative distribution function $F(\epsilon_{s_i}) = e^{-e^{\lambda s_i}}$
- This implies (see eg., McFadden, 1975) that for all actions in all information sets

$$\sigma(s_i) = \frac{\exp\{\lambda \cdot \pi_i(s_i|\sigma)\}}{\sum_{s_i' \in S_i(I_i)} \exp\{\lambda \cdot \pi_i(s_i'|\sigma)\}}$$

Properties

- For $\lambda < \infty$, all actions are chosen with positive probability, thus there are no information sets off the path, equilibrium refinement is implied
- As λ approaches infinity, convergence to a sequential equilibrium

Can be combined with social preferences, too, e.g. in ultimatum games

Reinterpreting the data of the centipede game with AQRE

The Agent logit equilibrium when $S_i(I_i) = \{s_i, s_{i'}\}$ is

$$\sigma(s_i) = \frac{1}{1 - exp(\lambda(\pi_i(s_i'|\sigma) - \pi_i(s_i|\sigma))}$$

So,

$$p_4 = \frac{1}{1 - exp(\lambda(1.6 - 3.2))}$$

$$p_3 = \frac{1}{1 - exp(\lambda(0.8p_4 + 6.4(1 - p_4) - 1.6))}$$

$$p_2 = \frac{1}{1 - exp(\lambda(0.4p_3 + (1 - p_3)[3.2p_4 + 1.6(1 - p_4)) - 0.8))}$$

$$p_1 = \frac{1}{1 - exp(\lambda(2p_2 + (1 - p_2)(1.6p_3 + (1 - p_3)(0.8p_4 + 6.4(1 - p_4)) - 4))}$$

Paper 3:
An Experimental Study of Constant-sum Centipede Games

Fey, McKelvey, and Palfrey (1996)

Motivation:

In the previous paper there were two very different ideas:

1. With the existence of an altruist subjects, there is an incentive for the selfish player to choose PASS,
2. The subjects make errors thorough the game

To disentangle this idea they study a centipede game in which there are no Pareto improvements over the course of the game, so that no selfish player has interest to choose PASS.

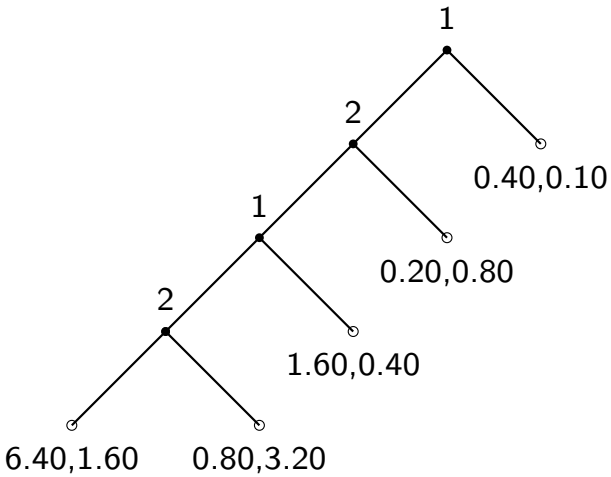


Figure: Classical Centipede Game

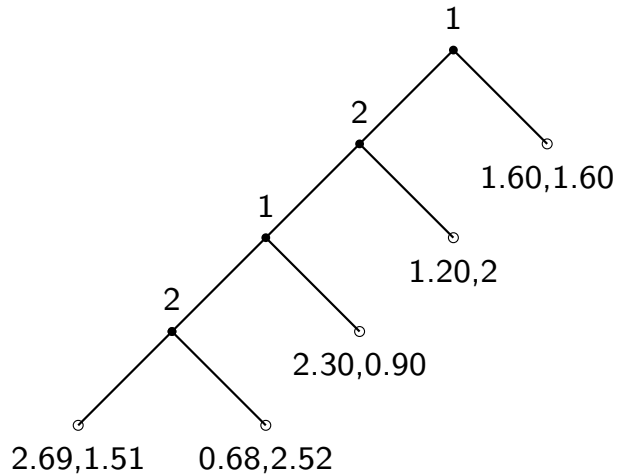


Figure: Zero-sum Centipede Game

The always *Take* Model

Compatible with backward induction, the Egalitarian model

This is a point prediction, in that the models rule out any other outcome as impossible. Econometrically, this makes it impossible to specify the likelihood of our data, which contain prohibited outcomes. They consider a statistical version of the model

- Every player, at every node reached in every match, chooses Take with a fixed probability p and chooses Pass with probability $1 - p$.
- Estimates of p close to 1 would indicate that the "Always Take" model does fairly well in explaining the data.

The always *Take* Model

- The DM chooses "Always Take" with probability

$$p_t = 1 - (1 - p_0)e^{-\alpha t}$$

with p_0 is the probability of Take used by subjects in the first match of the experiment and α represents the learning rate of the subjects.

- M be the total number of matches in the experiment, indexed by $i = 1....., M$.
- n_i the terminal node i
The log-likelihood is given by

$$Log\ L = \sum_{i=1}^{M} 1 - (1 - p_0)e^{-\alpha t} + \sum_{i=1}^{M} n_i(1 - p_0)e^{-\alpha t}$$

The AQRE with learning

- A player at node j in match number t faces a choice between Take, with payoff $\hat{u}_T = u_T + \epsilon_T$, and Pass, with expected payoff $\hat{u}_P + \epsilon_P$ where $\epsilon = \epsilon_T - \epsilon_P$ is drawn from a logistic distribution. So, the probability that a player at node jn in the match t chooses *Take* is

$$p_t^j = \mathbb{P}(\hat{u}_T > \hat{u}_P)$$
$$= \mathbb{P}(\epsilon > u_P - u_T) = \frac{1}{e^{\lambda(u_P - u_T)}}$$

- p_t^N is the probability that Player 2 chooses *Take* at N. Given this decision probability, we can compute the expected payoff a Pass by a Player 1 at the node $N-1$ ans determine: p_t^{N-1}.
- More generally we can obtain

$$\mathbf{p}_t = (p_t^1, p_t^2, ..., p_t^N)$$

The AQRE with learning

- let $\hat{f}_t^1 = p_t^1$, $\hat{f}_t^2 = (1 - p_t^1)p_t^2$, $\hat{f}_t^3 = (1 - p_t^1)(1 - p_t^2)p_t^3$, etc.
- Specifically, we assume that

$$\lambda_t = \lambda_0 + \beta t$$

- The log likelihood is

$$logL = \sum_{i=1}^{M} log(\hat{f}_t^{n_i}(\lambda, \beta))$$

AQRE vs "Always Take" model

- In eight of the nine individual experiments, the Agent Quantal Response Equilibrium model does better