Evolution & Learning in Games
Econ 243B

Jean-Paul Carvalho

Lecture 3:
Social Learning

Bayesian Social Learning

Bikhchandani, Hirshleifer & Welch (1992), Banerjee (1992)

- n agents sequentially decide between two actions, $s_i \in \{a, b\}$.
- Two states of world: $\theta \in \{A, B\}$.
- Action a is the preferred action in state A, and b is the preferred action in state B. Payoffs:

$$\pi_i = \begin{cases} 1 & \text{if } s_i = \theta, \\ -1 & \text{if } s_i \neq \theta. \end{cases}$$

Private Information

At the outset, each agent i receives a private signal z_i, where signals are i.i.d. conditional on θ.

For all i:

- If $\theta = A$, $z_i = A$ with prob. $q > \frac{1}{2}$ and B with prob. $1 - q$.
- If $\theta = B$, $z_i = B$ with prob. $q > \frac{1}{2}$ and A with prob. $1 - q$.

Social learning: i's decision is based not only on z_i but also on information learned from other players.

— *Observable signals*. For n large, learn state and make correct choice with high prob. by LLN.

Learning by Observing Actions

Actions can provide information on private signals, but less than you might think due to **information cascades**.

- Suppose for the first-mover, $z_1 = A$. By Bayes' theorem, 1's posterior belief is

$$\begin{aligned} P(\theta = A | z_1 = A) &= \frac{P(z_1 = A | \theta = A)\, P(\theta = A)}{P(z_1 = A)} \\ &= \frac{q\frac{1}{2}}{q\frac{1}{2} + (1-q)\frac{1}{2}} \\ &= q \;>\; \tfrac{1}{2}\,. \end{aligned}$$

- Hence 1's expected payoffs from $s_1 = a$ is $q - (1-q) = 2q - 1 > 0$ and from $s_1 = B$ is $1 - 2q < 0$.

- Agent 1 chooses $s_1 = a$.

Information Cascade

Now suppose the second-mover's signal is $z_2 = A$.

- 2 infers from observing $s_1 = a$, that $z_1 = A$.
- By Bayes' theorem, 2's posterior belief is

$$\begin{aligned} P(\theta = A | z_1 = A, z_2 = A) &= \frac{P(z_1 = A, z_2 = A | \theta = A)\, P(\theta = A)}{P(z_1 = A, z_2 = A)} \\ &= \frac{q^2 \frac{1}{2}}{q^2 \frac{1}{2} + (1-q)^2 \frac{1}{2}} \ > \ q \ > \frac{1}{2}\,. \end{aligned}$$

- Hence $s_2 = a$.

Thenceforth, all subsequent agents choose $s_i = A$ regardless of z_i.

Decision Rule

General rule:

- Let d be the number of prior choices of A minus number of prior choices of B.
  - $d > 1$, choose A regardless of private signal.
  - $d = 1$, choose A if $z_i = A$ and anything if $z_i = B$.
  - $d = 0$, follow private signal.
  - Cases are symmetric for $d < 0$.

Precisely which cascade emerges is *path dependent*.

Likelihood of Cascades & Information Transmission

Example. $q = 0.51$.

- If choose action uniformly at random when indifferent, probability of cascade after first two individuals is approx. 0.75.
  - Cascade after AA and BB with prob. one, and after AB and BA with prob. $1/2$.
- Probability of correct cascade is 0.5133 (calulation in Bikhchandani, Hirshleifer and Welch 1992).
- Probability of correct belief without social learning is $q = 0.51$.

See Smith and Sorensen (2000) for generalization in which continuous action space improves information transmission and can lead to full revelation of beliefs.

Culture

‘*Culture is information that people acquire from others by teaching, imitation and other forms of social learning*. On a scale unknown in any other species, people acquire skills, beliefs, and values from the people around them, and these strongly affect behaviour. People living in human populations are heirs to a pool of socially transmitted information that affects how they make a living, how they communicate, and what they think is right and wrong.” (Boyd & Richerson 2005, p. 3).

> *Especially important is the ability to transmit information across generations.*

- Produced by vertical (parent to child), oblique (adult non-parent to child) and horizontal (peer-to-peer) transmission.
- In complex societies, organizations (media, state, schools, clerics) play an important role in cultural transmission, and
- External information storage devices (libraries, www) and big data allow for more complex forms of learning.

The Evolution of Cultural Evolution

(Henrich and McElreath 2003)

- Humans can survive in a far wider range of environments than other primates.
- No evolved hardwired cognitive architecture for doing so, e.g. Burke and Wills.
- The main difference between human beings and other animals is the human capacity for social learning and the accumulation of knowledge over generations (e.g. hunting technologies, food processing methods).
- Call this **cultural learning**.
- Necessary to understand the *psychological mechanisms* that make cultural learning possible and the *population dynamics* produced by cultural learning.

A Model of Cultural Learning

(Giuliano and Nunn 2017, based on Rogers 1988)

- A continuum of agents.
- Discrete time: $t = 1, 2, \ldots$
- Two states of world: $\theta \in \{A, B\}$.
- Two actions: $s_i \in \{a, b\}$. (Action a is the preferred action in state A, and b is the preferred action in state B.)
- Payoffs:

$$\pi_i = \begin{cases} \beta & \text{if } s_i = \theta, \\ -\beta & \text{if } s_i \neq \theta. \end{cases}$$

Environmental Variability & Learning

State of world:

- With probability $1 - \Delta$, the state in period $t + 1$ is the same as in t.
- With probability Δ, the state is A or B, each with probability $1/2$.

Strategies:

- **Social Learner** (SL): copies action of an agent in the previous generation, chosen uniformly at random.
- **Individual Learner** (IL): pays cost $c > 0$ to learn state (for sure).

Population state: x is the proportion of social learners.

Social Learning

The share $1 - x$ of individual learners choose the correct action.

But so does a social learner if

(i) she copies an individual learner since the latest state change,

(ii) she copies from a social learner who copied from an individual learner since the latest state change,

(iii) copies from a social learner who copied from a social learner who copied from an individual learner since the latest state change,

(iv) and so on.

Up-to-date Social Learning

In equilibrium (i.e. $x_t = x$ for all t), the probability of

- (i) is $(1 - x)(1 - \Delta)$,
- (ii) is $x(1 - x)(1 - \Delta)^2$,
- (iii) is $x^2(1 - x)(1 - \Delta)^3$,

and so on.

Iterating and summing gives

$$\sum_{t=1}^{\infty} x^{t-1}(1 - x)(1 - \Delta)^t .$$

Out-of-date Social Learning

With complementary probability, a social learner chooses an action that was copied before the last state change:

$$1 - \sum_{t=1}^{\infty} x^{t-1} (1 - x) (1 - \Delta)^t .$$

After a state change, there is a $\frac{1}{2}$ chance of being in either state, so there is a 50% chance that a social learner is correct and a 50% chance she is incorrect.

Learning Payoffs

$$
\begin{aligned}
\pi_{SL} &= \left(\sum_{t=1}^{\infty} x^{t-1} (1-x) (1-\Delta)^{t} \right) \beta \\
&\quad + \frac{1}{2} \left(1 - \sum_{t=1}^{\infty} x^{t-1} (1-x) (1-\Delta)^{t} \right) \beta \\
&\quad + \frac{1}{2} \left(1 - \sum_{t=1}^{\infty} x^{t-1} (1-x) (1-\Delta)^{t} \right) (-\beta) \\
&= \frac{(1-x)(1-\Delta)}{1-x(1-\Delta)} \beta ,
\end{aligned}
$$

which declines monotonically from $(1-\Delta)\beta$ to 0 as x goes through $[0,1]$.

$$\pi_{IL} = \beta - c.$$

Equilibrium

Three regimes:

1. $c \leq \Delta\beta$. **IL Monomorphic Equilibrium**: $x = 0$.

2. $c \geq \beta$. **SL Monomorphic Equilibrium**: $x = 1$.

3. $\Delta\beta < c < \beta$. **Polymorphic Equilibrium**:

$$x^* = \frac{c - \Delta\beta}{(1 - \Delta)c}.$$

Environmental Variability & Social Learning

$$\frac{dx^*}{d\Delta} = -\frac{c(\beta - c)}{[(1-\Delta)c]^2} \; < 0$$

Therefore, environmental variability limits social learning by reducing the amount of information stored in the behavior of previous generations.

Giuliano and Nunn (2017) find that populations with ancestors who lived in regions with greater environmental variability (measured by temperature variation) are more ‘traditional’ today.

Welfare and Adaptation

- Notice that the mean payoff is the same as in a population of purely individual learners.
- Hence social learning may not be adaptive.
- However, Boyd and Richerson (1995) show that social learning leads to higher average payoff in the population if it allows the accumulation of behaviors that no individual learner could acquire in a lifetime.
- Cumulative cultural evolution is rare among animals because it only spreads when there is a critical mass of cultural learners (Boyd and Richerson 1996).

Cultural Learning with Coordination

Carvalho & McBride *in progress*

- It is hard to conceive of situations in which individuals learn socially but act in isolation (payoff not frequency dependent).
- Now let payoff depend on matching the state and coordinating with other agents. All else is the same.
- Define p as the proportion of agents choosing the "correct" action.
- Payoffs:

$$\pi_i = \begin{cases} \beta + \alpha p, & \text{if } s_i = \theta, \\ -\beta + \alpha\left(1 - p\right), & \text{if } s_i \neq \theta. \end{cases}$$

Learning Payoffs

Hence the expected payoff to a social learner is

$$\begin{aligned}
\pi_{SL} &= \left(\sum_{t=1}^{\infty} x^{t-1}(1-x)(1-\Delta)^t\right)(\beta+\alpha p) \\
&\quad +\frac{1}{2}\left(1-\sum_{t=1}^{\infty} x^{t-1}(1-x)(1-\Delta)^t\right)(\beta+\alpha p)+ \\
&\quad +\frac{1}{2}\left(1-\sum_{t=1}^{\infty} x^{t-1}(1-x)(1-\Delta)^t\right)(-\beta+\alpha(1-p)) \\
&= \frac{(1-x)(1-\Delta)}{1-x(1-\Delta)}\left(\beta+\alpha\left(p-\frac{1}{2}\right)\right)+\frac{1}{2}\alpha.
\end{aligned}$$

The payoff to an individual learner is

$$\pi_{IL} = \beta + \alpha p - c.$$

Solving for p

The proportion of individuals choosing the correct action, p, is

$$\begin{aligned}
p &= 1 - x + x\left(\sum_{t=1}^{\infty} x^{t-1}(1-x)(1-\Delta)^t\right) \\
&+ x\frac{1}{2}\left(1 - \sum_{t=1}^{\infty} x^{t-1}(1-x)(1-\Delta)^t\right) \\
&= 1 - x + x\left(\frac{(1-x)(1-\Delta)}{1-x(1-\Delta)} + \frac{1}{2}\left(1 - \frac{(1-x)(1-\Delta)}{1-x(1-\Delta)}\right)\right) \\
&= 1 - \frac{1}{2}\frac{x\Delta}{1-x(1-\Delta)}.
\end{aligned}$$

Clearly, $p > \frac{1}{2}$. Furthermore:

$$\frac{\partial p}{\partial x} = -\frac{1}{2}\frac{\Delta}{(1-x(1-\Delta))^2} < 0$$

with $p \to 1$ as $x \to 0$, and $p \to \frac{1}{2}$ as $x \to 1$.

Learning Payoffs with Coordination

Substituting for p in the payoff functions, the expected payoff to the individual learner is

$$\pi_{IL} = \beta + \alpha\left(1 - \frac{1}{2}\frac{x\Delta}{1 - x\left(1-\Delta\right)}\right) - c,$$

and the expected payoff to the social learner is

$$\begin{aligned}\pi_{SL} &= \frac{\left(1-x\right)\left(1-\Delta\right)}{1-x\left(1-\Delta\right)}\left(\beta + \alpha\left(1 - \frac{1}{2}\frac{x\Delta}{1-x\left(1-\Delta\right)} - \frac{1}{2}\right)\right) + \frac{1}{2}\alpha \\ &= \frac{\left(1-x\right)\left(1-\Delta\right)}{1-x\left(1-\Delta\right)}\left(\beta + \frac{1}{2}\alpha\frac{1-x}{1-x\left(1-\Delta\right)}\right) + \frac{1}{2}\alpha.\end{aligned}$$

Both payoffs are strictly decreasing in x.

Equilibrium

In a **polymorphic equilibrium**:

$$x^{**} = \frac{c - \Delta\beta - \left(p\left(x^{**}\right) - \frac{1}{2}\right)\Delta\alpha}{(1-\Delta)\,c}.$$

Recall that $p(x) > \frac{1}{2}$. Hence $x^{**} < x^{*} = \frac{c-\Delta\beta}{(1-\Delta)c}$.

- Social coordination (i.e., $\alpha > 0$) reduces the equilibrium proportion of social learners.
- *Direct effect* of $\alpha > 0$: Individual learners get a higher social coordination payoff because they always coordinate with the majority while not all social learners do so.
- *Indirect effect* of $\alpha > 0$: The direct effect reduces x and thus raises both π_{IL} and π_{SL}.