Social Learning

Duarte Gonçalves

University College London

Topics in Economic Theory

Overview

1. Social Learning: Checking Neighbours' Crop Yields

2. Warm-Up: Gaussian-Quadratic Model

3. Information Cascades

4. Further Topics

Overview

1. Social Learning: Checking Neighbours' Crop Yields
2. Warm-Up: Gaussian-Quadratic Model
3. Information Cascades
4. Further Topics

Social Learning

Numerous situations in which we try to learn from others' actions

- Knowing a film is successful is informative about its quality and genre (unlikely to be awful, but also unlikely to be an indie film)
- Seeing price of a stock going up: infer more positive outlook on firm's fundamentals more likely.
- Product purchases, migration decisions, bachelor choices, technology adoption, take-up of microfinance, etc.

Can the crowd get it wrong? When?

Setup

Basic Ingredients

$\theta \in \Theta$ unknown parameter; $\theta \sim P_\Theta$.

Action: Each agent n takes action $a_n \in A \subseteq \mathbb{R}$.

Public History: $h_t := \{a_\ell, \ell < t\}$, $h_0 = \emptyset$.

Payoffs: $u(a_n, \theta, r_n)$; Type: Agent n's type r_n; R types.

Private Information: Each agent n has private information about θ.

Important how large A; coarser set restricts how much info conveyed.

Examples

Binary Actions: $\Theta = \{H, L\}$, $A = \{0, 1\}$, e.g. buy/invest/adopt/migrate/etc vs don't. WLOG $u(a, \theta, r) := a(\mathbf{1}_{\{\theta = H\}} - r)$.

Gaussian-Quadratic: $\Theta = A = \mathbb{R}$, e.g., predict variable: $u(a, \theta, r) := -(\theta - a)^2$ or investment problem $u(a, \theta, r) := 2\theta a - a^2$.

Setup

Private and Public Information:

Private Signal: $s_n \sim P_S(\cdot \mid \theta)$ iid (conditional on θ); $p_n^0(\theta) \propto p_0(\theta) P_S(s_n \mid \theta)$.
$p_n^0 = f_p(s_n)$, $s_n \sim P_S(\cdot \mid \theta)$; hence $p_n^0 \sim P_P(\cdot \mid \theta)$ for some P_P s.t. $\mathbb{E}[p_n^0] = P_\Theta$.

Private Belief: $p_n^0 \sim P_P$.

Public Belief: $q_t(\theta) = \mathbb{P}(\theta \mid h_t)$.
$\pi(h_t \mid \theta) := \mathbb{P}(h_t \mid \theta)$; $q_t(\theta) \propto p_0(\theta)\pi(h_t \mid \theta)$.

Private Posterior Belief: Private belief + updating based on public history.
$p_t(\theta) \propto p_0(\theta) P_S(s_t \mid \theta)\pi(h_t \mid \theta) = q_t(\theta) p_t^0(\theta)/p_0(\theta)$.
Updating mapping $BU : (p_t^0, q_t) \mapsto BU(p_t^0, q_t) = p_t$.

Timing: Within each period t:

1. Agent t observes private signal s_n and forms interim private belief p_t^0.
2. Agent observes public history $h_t = (a_\ell, \ell < t)$ and forms private posterior belief p_t.
3. Agent t takes action a_t.

wPBE (focusing on on-path)

Herd, Cascade

Definition

(i) Actions converge if $\sum_t a_t/t \to a_\infty$ for some $a_\infty \in A$.

(ii) A herd occurs at t if, given the public belief, agent t's optimal action is the same as agent $(t-1)$'s.

(iii) A herd starts at T if it occurs at any $t \geq T$.

(iv) A cascade occurs at t if, given the public belief, agent t's optimal action is a.s. independent of t private signal.

(v) A cascade starts at T if it occurs at any $t \geq T$.

(vi) A cascade that occurs at ∞ is a limit cascade.

Remark

Cascade in finite time $\implies$ Limit cascade, Herd in finite time $\implies$ Actions converge.
If A finite, Limit cascade $\implies$ Herd in finite time.

Overview

1. Social Learning: Checking Neighbours' Crop Yields
2. Warm-Up: Gaussian-Quadratic Model
3. Information Cascades
4. Further Topics

Warm-Up: Gaussian-Quadratic Model

$u(a,\theta) = -(\theta - a)^2$; $\theta \sim N(\theta_0, \rho_\theta^{-1})$; $s_t = \theta + \varepsilon_t$; $\varepsilon_t \sim N(0, \rho_\varepsilon^{-1})$ iid; indep. from θ.

Public Belief: $N(\mu_t, \rho_t^{-1})$.

Posterior Belief = Private Signal + Public Belief: $N(\tilde{\mu}_t, \tilde{\rho}_t^{-1})$

$\tilde{\mu}_t = (1 - w_t)\mu_t + w_t s_t$; $\tilde{\rho}_t = \rho_t + \rho_\varepsilon$; $w_t = \rho_\varepsilon / \tilde{\rho}_t$.

Conditional on posterior belief, $a_t = \arg\max_a \mathbb{E}[u(a,\theta) \mid h_t, s_t] = \mathbb{E}[\theta \mid h_t, s_t] = \tilde{\mu}_t$.

Social Learning: As μ_t, a_t, w_t are known, agent $t+1$ can infer $s_t = (a_t - (1 - w_t)\mu_t)/w_t$

$\Longrightarrow$ $\mu_{t+1} = \tilde{\mu}_t$ and $\rho_{t+1} = \tilde{\rho}_t$.

As $a_t = \mu_t = \frac{\rho_\theta}{\rho_t}\theta_0 + \frac{t\rho_\varepsilon}{\rho_t}\sum_{\ell \leq t} s_\ell \to \theta$ a.s. $\Longrightarrow$ learning is complete.

No herd, no cascade in finite time.

Imitation: $\mathbb{E}[|a_t - a_{t-1}|] = w_t \mathbb{E}[|s_t - a_{t-1}|] \to 0$: actions more similar over time.

Heterogeneity: $u(a,\theta,r) = -(\theta + r - a)^2$, $r \sim N(0, \rho_r^{-1})$; results hold, but learning slower.

Warm-Up: Gaussian-Quadratic Model

Costly signal: $u(a,\theta) = -(\theta - a)^2 - c(\rho_\varepsilon)$; $c \geq 0, c' > 0, c'' > 0$.

Claim: Social learning stops in finite time, i.e.,

$$\exists T < \infty : \forall t > T,\ \rho_t = \rho_T \iff \rho_{\varepsilon,t} = 0.$$

- $\mathbb{E}[u(a_t,\theta) \mid h_t, s_t] = -\mathbb{E}[(\theta - \mathbb{E}[\theta \mid h_t, s_t])^2 \mid h_t, s_t] - c(\rho_{\varepsilon,t}) = -(\rho_t + \rho_{\varepsilon,t})^{-1} - c(\rho_{\varepsilon,t})$.

  FOC: $0 = (\rho_t + \rho_{\varepsilon,t})^{-2} - c'(\rho_{\varepsilon,t})$.
- Then: $\rho_t^{-2} \leq c'(0) \implies (\rho_t + \rho_{\varepsilon,t})^{-2} - c'(\rho_{\varepsilon,t}) \leq \rho_t^{-2} - c'(\rho_{\varepsilon,t}) \leq 0 \implies \rho_{\varepsilon,t} = 0$.
- $c'(\rho_{\varepsilon,t-1}) = (\rho_{t-1} + \rho_{\varepsilon,t-1})^{-2} = \rho_t^{-2} \geq (\rho_t + \rho_{\varepsilon,t})^{-2} = c'(\rho_{\varepsilon,t})$

  $\implies \rho_{\varepsilon,t}$ decreasing in t.
- Suppose that $\rho_{\varepsilon,t} \geq \delta\ \forall t$. Then, $\exists T : \forall t \geq T$,

  $\rho_t = \rho_0 + \sum_{\ell=1}^{t-1} \rho_{\varepsilon,\ell} \geq \rho_0 + (t-1) \geq c'(0)^{-1/2} \implies \rho_{\varepsilon,t} = 0$, contradiction.

Warm-Up: Gaussian-Quadratic Model

Costly signal: $u(a,\theta) = -(\theta - a)^2 - c(\rho_\varepsilon)$; $c \geq 0, c' > 0, c'' > 0$.

Claim: Social learning stops in finite time, i.e.,

$\exists T < \infty : \forall t > T,\ \rho_t = \rho_T \iff \rho_{\varepsilon,t} = 0$.

- $\mathbb{E}[u(a_t,\theta) \mid h_t, s_t] = -\mathbb{E}[(\theta - \mathbb{E}[\theta \mid h_t, s_t])^2 \mid h_t, s_t] - c(\rho_{\varepsilon,t}) = -(\rho_t + \rho_{\varepsilon,t})^{-1} - c(\rho_{\varepsilon,t})$.

  FOC: $0 = (\rho_t + \rho_{\varepsilon,t})^{-2} - c'(\rho_{\varepsilon,t})$; $\rho_{\varepsilon,t}$ decreasing in t; $\rho_t^{-2} \leq c'(0) \implies \rho_{\varepsilon,t} = 0$.
- Suppose that $\rho_{\varepsilon,t} > 0$ and $\rho_{\varepsilon,t} \downarrow 0$.

  From FOC:

$$\begin{aligned}\rho_{\varepsilon,t} &= c'(\rho_{\varepsilon,t})^{-1/2} - \rho_t = c'(\rho_{\varepsilon,t})^{-1/2} - \rho_{t-1} - \rho_{\varepsilon,t-1} \\ &= c'(\rho_{\varepsilon,t})^{-1/2} - \rho_{t-1} - c'(\rho_{\varepsilon,t-1})^{-1/2} + \rho_{t-1} = c'(\rho_{\varepsilon,t})^{-1/2} - c'(\rho_{\varepsilon,t-1})^{-1/2}.\end{aligned}$$

  Moreover

$$\begin{aligned}\rho_t &= \rho_0 + \rho_{\varepsilon,0} + \sum_{\ell=1}^{t-1} \rho_{\varepsilon,\ell} = \rho_0 + \rho_{\varepsilon,0} + \sum_{\ell=1}^{t-1} \left[c'(\rho_{\varepsilon,\ell})^{-1/2} - c'(\rho_{\varepsilon,\ell-1})^{-1/2}\right] \\ &= \rho_0 + \rho_{\varepsilon,0} + c'(\rho_{\varepsilon,t-1})^{-1/2} - c'(\rho_{\varepsilon,0})^{-1/2} \\ \implies & \lim_{t\to\infty} \rho_t = \rho_{\varepsilon,1} + c'(0)^{-1/2}.\end{aligned}$$

Warm-Up: Gaussian-Quadratic Model

Costly signal: $u(a, \theta) = -(\theta - a)^2 - c(\rho_\varepsilon)$; $c \geq 0, c' > 0, c'' > 0$.

Claim: Social learning stops in finite time, i.e.,

$\exists T < \infty : \forall t > T,\ \rho_t = \rho_T \iff \rho_{\varepsilon,t} = 0$.

- $\mathbb{E}[u(a_t, \theta) \mid h_t, s_t] = -\mathbb{E}[(\theta - \mathbb{E}[\theta \mid h_t, s_t])^2 \mid h_t, s_t] - c(\rho_{\varepsilon,t}) = -(\rho_t + \rho_{\varepsilon,t})^{-1} - c(\rho_{\varepsilon,t})$.

  FOC: $0 = (\rho_t + \rho_{\varepsilon,t})^{-2} - c'(\rho_{\varepsilon,t})$; $\rho_{\varepsilon,t}$ decreasing in t; $\rho_t^{-2} \leq c'(0) \implies \rho_{\varepsilon,t} = 0$.

- Suppose that $\rho_{\varepsilon,t} > 0$ and $\rho_{\varepsilon,t} \downarrow 0$.

  Since $\lim_{t\to\infty} \rho_t = \rho_{\varepsilon,1} + c'(0)^{-1/2}$, $\exists T : \forall t \geq T$,

  $\rho_t > c'(0)^{-1/2} \iff \rho_t^{-2} < c'(0) \implies \rho_{\varepsilon,t} = 0$.

Crashes and Booms

Adjustment to economic shock tends to be first slow and then abrupt (e.g., a market crash). Why?

Model sketch: population every period takes irreversible investment decisions; start with high public confidence in 'good times' but state is bad.

- Sigmoid-shaped adjustment of beliefs from social learning to change in state features:
  (1) slow evolution (somewhat business as usual),
  (2) fast adjustment to new reality (negative shock), and
  (3) stabilising around 'new normal'.
- Caplin and Leahy (1994 AER); Chamley (2010 Ch. 4.6).

Fast changing state would temper these abrupt changes.

Overview

Setup

Version of Bikchandani, Hirschleifer, and Welch (1992 JPE).
(but using martingale tools employed by Smith and Sørensen (2000 Ecta).)

Setup

State: $\theta \in \Theta = \{H, L\}$, $p_0 = \mathbb{P}(\theta = H) \in (0, 1)$.

Signal: $s_n \in S = \{1, 2, ..., N_S\}$; $s_n \sim P_S(\cdot \mid \theta)$. Assume $P_S(s \mid \theta) > 0 \quad \forall s, \theta$.
(No fully revealing signals with positive probability.)

Order signals: $\frac{P_S(1|H)}{P_S(1|L)} < \cdots < \frac{P_S(N_S|H)}{P_S(N_S|L)}$.
MLRP WLOG when $|\Theta| = 2$; strict MLRP: bundle signals with same likelihood together.

Payoffs: $u(a, \theta, r) := a(\mathbf{1}_{\{\theta=H\}} - r)$. Tie-breaking: Invest iff belief $\theta = H > r$.

Setup

Cascade Sets and Other Definitions

Public History: $h_t = (a_\ell)_{\ell<t}$; $h_0 = \emptyset$.

Public Belief: $q_t = \frac{p_0 \pi(h_t|H)}{p_0 \pi(h_t|H) + (1-p_0)\pi(h_t|L)}$.

Private Interim Belief: $p_t^0(s) = \frac{p_0 P_S(s|H)}{p_0 P_S(s|H) + (1-p_0) P_S(s|L)}$. $p_t^0 \sim P_P$.

Let $\underline{b} :=$ inf supp$(P_P) > 0$ and $\bar{b} :=$ sup supp$(P_P) < 1$. (Recall: finite S.)

Private Posterior Belief:

$$p_t(s) = \frac{q_t P_S(s|H)}{q_t P_S(s|H) + (1-q_t) P_S(s|L)} = \frac{q_t\, p_t^0/p_0}{q_t\, p_t^0/p_0 + (1-q_t)\,(1-p_t^0)/(1-p_0)} =: BU(p_t^0, q_t).$$

Action a's Basin: $I_a^r := \{p \in [0,1] \mid a = \min(\arg\max_{a'} p\, u(a', H, r) + (1-p)\, u(a', L, r))\}$.

$\exists 0 = \hat{p}_0 < \hat{p}_1 < \hat{p}_2 = 1 : I_0^r = [\hat{p}_0, \hat{p}_1], I_1^r = (\hat{p}_1, \hat{p}_2]$.

(Argument extends beyond binary actions)

Cascade Set for a: set of public beliefs at which agent chooses a regardless of private information.

$J_a^r := \{q \mid BU(p, q) \in I_a^r, \forall p \in \text{supp}(P_P)\}$.

$\exists 0 = \hat{q}_0^{r,-} < \hat{q}_0^{r,+} < \hat{q}_1^{r,-} < \hat{q}_1^{r,+} = 1 : J_0^r = [\hat{q}_0^{r,-}, \hat{q}_0^{r,+}]$ and $J_1^r = (\hat{q}_1^{r,-}, \hat{q}_1^{r,+}]$,

with $BU(\underline{b}, \hat{q}_a^{r,-}) = \hat{p}_a^r$ and $BU(\bar{b}, \hat{q}_a^{r,+}) = \hat{p}_{a+1}^r$.

$q \in J_a^r \iff \hat{p}_a \leq BU(\underline{b}, q) < BU(p_0, q) = q < BU(\bar{b}, q) \leq \hat{p}_{a+1} \implies J_a^r \subseteq \text{int}(I_a^r)$.

Define $J^r := \cup_a J_a^r$.

Not Socially Learning the Wrong Thing

Proposition

Fix $\theta = H$. In any equilibrium, a fully wrong information cascade ($q_t \to 0$) a.s. never starts.

Proof

Need to show wp1 $q_t > 0$ $\forall t$ and $\lim_{t\to\infty} q_t > 0$.

Public Likelihood Ratio: $l(q_t) = (1 - q_t)/q_t$.

Claim: Conditional on $\theta = H$, $l(q_t)$ is a martingale.

$$\begin{aligned}\mathbb{E}[l(q_{t+1}) \mid h_t, \theta = H] &= \mathbb{E}\left[\frac{(1-p_0)\pi^L(h_{t+1})}{p_0\pi^H(h_{t+1})}\middle| h_t, \theta = H\right] \\ &= \mathbb{E}\left[\frac{\mathbb{P}(a_t \mid h_t, \theta = L)(1-p_0)\pi^L(h_t)}{\mathbb{P}(a_t \mid h_t, \theta = H)p_0\pi^H(h_t)}\middle| h_t, \theta = H\right] \\ &= l(q_t)\sum_a \mathbb{P}(a_t = a \mid h_t, \theta = H)\frac{\mathbb{P}(a_t = a \mid h_t, \theta = L)}{\mathbb{P}(a_t = a \mid h_t, \theta = H)} = l(q_t)\end{aligned}$$

Note: q_t is a martingale; conditional on $\theta = H$ it is a submartingale.

Not Socially Learning the Wrong Thing

Proposition

In any equilibrium, a fully wrong information cascade a.s. never starts.

Proof

Need to show wp1 $q_t > 0$ $\forall t$ and $\lim_{t\to\infty} q_t > 0$.

Public Likelihood Ratio: $l(q_t) = (1 - q_t)/q_t$.

Claim: Conditional on $\theta = H$, $l(q_t)$ is a martingale.

$$\mathbb{E}[l(q_{t+1}) \mid h_t, \theta = H] = l(q_t)$$

$l(q_t) \geq 0$. By Doob's martingale convergence theorem, conditional on $\theta = H$, $l(q_t) \to l(q_\infty) < \infty$ a.s.

Information Cascades Happen

Theorem

In any equilibrium, an information cascade starts in finite time a.s.

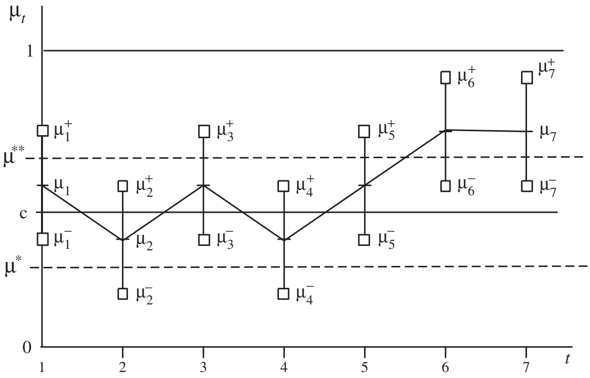

Figure 4.2 Representation of a cascade. In each period, the middle of the vertical segment is the public belief; the top and the bottom of the segment are the beliefs of an optimist (with a private signal $s = 1$) and of a pessimist (with signal $s = 0$). The private signals are $s_1 = 0,\ s_2 = 1,\ s_3 = 0,\ s_4 = 1,\ s_5 = 1$.

Information Cascades Happen

Theorem

In any equilibrium, an information cascade starts in finite time a.s.

Proof

Claim: $\text{supp}(q_\infty) \subseteq [0, \hat{q}_0^+] \cup [\hat{q}_1^-, 1]$.

Suppose $q^* \in (\hat{q}_0^+, \hat{q}_1^-) \cap \text{supp}(q_t) \implies \forall \varepsilon > 0, \mathbb{P}(q_\infty \in (q^* - \varepsilon, q^* + \varepsilon)) > 0$.

Hence wp>0 $\exists T < \infty : \forall t \geq T, q_t \in (q^* - \varepsilon, q^* + \varepsilon)$.

Take ε arbitrarily small. $(q^* - \varepsilon, q^* + \varepsilon) \subset (\hat{q}_0^+, \hat{q}_1^-)$, an active learning region.

In the active learning region, prob agent takes 0 or 1 is different depending on the true state, and so $\mathbb{P}(a_t = 1 \mid q_t, \theta = H) \neq \mathbb{P}(a_t = 1 \mid q_t, \theta = H)$.

$a_t = 1 \implies q_{t+1} > q_t$ and $a_t = 0 \implies q_{t+1} < q_t$.

Changes are discrete, so can find small $\varepsilon > 0$ s.t. exit $(q^* - \varepsilon, q^* + \varepsilon)$ wp>0 starting from $q_t \in (q^* - \varepsilon, q^* + \varepsilon)$. Contradiction.

Information Cascades Happen

Theorem

In any equilibrium, an information cascade starts in finite time a.s.

Proof

Claim: $\mathbb{P}(q_\infty = \hat{q}_1^-) = 0$.

Suppose $\mathbb{P}(q_\infty = \hat{q}_1^-) > 0$.

Hence for positive prob sample of paths $q_t \uparrow \hat{q}_1^-$.

$\implies$ for any $\varepsilon > 0$, wp>0 $\exists T < \infty : \forall t \geq T, q_t \in (\hat{q}_1^- - \varepsilon, \hat{q}_1^-]$.

Note $a_t = f(s_t, q_t)$, hence $\text{supp}(q_{t+1} \mid q_t = q)$ is independent from t.

$\implies$ q_{t+1} is time-homogeneous Markov process.

For small enough ε, one has $\forall q_t \in (\hat{q}_1^- - \varepsilon, \hat{q}_1^-] \implies (a_t = 0 \iff s_t = 1)$.

Information Cascades Happen

Theorem

In any equilibrium, an information cascade starts in finite time a.s.

Proof

Claim: $\mathbb{P}(q_\infty = \hat{q}_1^-) = 0$.

For small enough ε, one has $\forall q_t \in (\hat{q}_1^- - \varepsilon, \hat{q}_1^-] \implies (a_t = 0 \iff s_t = 1)$.

Recall: s_1 is lowest signal, induces $p_t^0 = \underline{b}$, and $BU(\underline{b}, \hat{q}_1^-)$ leaves DM indifferent between 0 and 1; tie-breaking in favour of action 0.

Note that, since q_t is a martingale, $\mathbb{E}[q_{t+1} \mid q_t] = q_t$ and then

$$\frac{q_t P_S(1|H)}{q_t P_S(1|H)+(1-q_t)P_S(1|L)} < q_t < \frac{q_t(1-P_S(1|H))}{q_t(1-P_S(1|H))+(1-q_t)(1-P_S(1|L))}.$$

Pick $\varepsilon > 0$ s.t. $\frac{\hat{q}_1^- P_S(1|H)}{\hat{q}_1^- P_S(1|H)+(1-\hat{q}_1^-)P_S(1|L)} < \hat{q}_1^- - \varepsilon < \hat{q}_1^- < \frac{(\hat{q}_1^- - \varepsilon)(1-P_S(1|H))}{(\hat{q}_1^- - \varepsilon)(1-P_S(1|H))+(1-(\hat{q}_1^- - \varepsilon))(1-P_S(1|L))}$.

i.e., (i) $\hat{q}_1^- - \varepsilon >$ posterior ($\hat{q}_1^-$ & s_1) and (ii) $\hat{q}_1^- <$ posterior ($\hat{q}_1^- - \varepsilon$ & s_ℓ, $\ell > 1$);
(yes, there is such $\varepsilon > 0$.)

Then, $q_t \in (\hat{q}_1^- - \varepsilon, \hat{q}_1^-] \implies q_{t+1} \notin (\hat{q}_1^- - \varepsilon, \hat{q}_1^-]$. Contradiction.

Information Cascades Happen

Theorem

In any equilibrium, an information cascade starts in finite time a.s.

Proof

We have $\operatorname{supp}(q_\infty) \subseteq J^r$. It remains to be shown that a.s. $\exists T < \infty : \forall t \geq T, q_t \in J^r$.

Note $J^r = \operatorname{int} J^r \cup \{\hat{q}_0^+\}$. Let $B = [0, 1] \setminus \operatorname{int} J^r$.

Then, since $q_\infty \in J^r$ and B closed $\implies$ either (a) $\exists t < \infty : q_t \in \operatorname{int} J^r$ or (b) $\forall t, q_t \in B \implies q_\infty \in B \implies q_\infty = \hat{q}_0^+$.

If (a), done. Suppose (b).

WTS that, on event $\{q_\infty = \hat{q}_0^+\}$, a.s. $\exists T < \infty : \forall t \geq T$, $q_t = \hat{q}_0^+$.

Information Cascades Happen

Theorem

In any equilibrium, an information cascade starts in finite time a.s.

Proof

WTS that, on event $\{q_\infty = \hat{q}_0^+\}$, a.s. $\exists T < \infty : \forall t \geq T$, $q_t = \hat{q}_0^+$.

Suppose not. Then wp>0, $\{q_\infty = \hat{q}_0^+\} \cap_{t\in\mathbb{N}} \{q_t > \hat{q}_0^+\}$.

On $\{q_\infty = \hat{q}_0^+\} \cap_{t\in\mathbb{N}} \{q_t > \hat{q}_0^+\}$, a.s. $\forall \varepsilon > 0, \exists T < \infty : \forall t \geq T$, $q_t \in (\hat{q}_0^+, \hat{q}_0^+ + \varepsilon)$.

Take ε small such that $\forall q_t \in (\hat{q}_0^+, \hat{q}_0^+ + \varepsilon)$, $a_t = 1 \iff s_t = N_S$.

$a_t = 1 \implies q_{t+1} = \frac{q_t P_S(N_S|H)}{q_t P_S(N_S|H) + (1-q_t) P_S(N_S|L)} > q_t$ and, since q_t is martingale,

$a_t = 0 \implies q_{t+1} = \frac{q_t(1-P_S(N_S|H))}{q_t(1-P_S(N_S|H)) + (1-q_t)(1-P_S(N_S|L))} < q_t$.

Can take $\varepsilon > 0$ small enough s.t. also $\frac{(\hat{q}_0^+ + \varepsilon)(1-P_S(N_S|H))}{(\hat{q}_0^+ + \varepsilon)(1-P_S(N_S|H)) + (1-(\hat{q}_0^+ + \varepsilon))(1-P_S(N_S|L))} < \hat{q}_0^+$.

And can take $\varepsilon > 0$ small enough s.t. also $\frac{\hat{q}_0^+ P_S(N_S|H)}{\hat{q}_0^+ P_S(N_S|H) + (1-\hat{q}_0^+) P_S(N_S|L)} > \hat{q}_0^+ + \varepsilon$.

Hence, for small enough ε, $q_t \in (\hat{q}_0^+, \hat{q}_0^+ + \varepsilon) \implies q_{t+1} \notin (\hat{q}_0^+, \hat{q}_0^+ + \varepsilon)$. Contradiction.

Done!

Wrong Cascades Happen

Theorem

In any equilibrium, for $p_0 \notin J^r$, a wrong information cascade starts wp> $\mathbf{0}$.

It is possible everyone unwittingly just takes the suboptimal action!

This is despite the fact every period new information arives.

Markets can get it very wrong due to informational externalities.

Wrong Cascades Happen

Theorem

In any equilibrium, for $p_0 \notin J^r$, a wrong information cascade starts wp> 0.

Proof

Consider $\theta = H$; symmetric for $\theta = L$.

J^r is absorbing/steady state set: $q_t \in J^r \implies q_{t+1} = q_t$.

$\tau := \inf\{t \geq 0 \mid q_t \in J^r\} = \inf\{t \geq 0 \mid l(q_t) \notin (l(\hat{q}_0^+), l(\hat{q}_1^+)]\}$.

Let $w^H := \mathbb{P}(q_\infty \in J_0^r \mid \theta = H) = \mathbb{P}(l(q_\infty) \in< l(\hat{q}_0^+) \mid \theta = H)$;

$l_{0,\infty} = \mathbb{E}[l(q_\infty) \mid \theta = H, q_\infty \in J_0^r]$; and $l_{1,\infty} = \mathbb{E}[l(q_\infty) \mid \theta = H, q_\infty \in J_1^r]$.

$l(q_t)$ is martingale $\implies$ by Doob's optional stopping theorem,

$l(p_0) = \mathbb{E}[l(q_\tau)|\theta = H] = w^H l_{\infty,0} + (1 - w_H)l_{1,\infty}$.

Finally: $p_0 \notin J^r \implies l_{0,\infty} < l(q_0) < l_{1,\infty} \implies w_H > 0$.

Robustness of Cascades

Smith and Sørensen (2000 Ecta) ask how robust are the results

- ‘Crazy’ action/person; need to ‘ignore deviation’.
- Agent with different preferences; others may still learn from their actions.
- Agent has more precise information; contrarian action.

Robustness of Cascades

'Crazy type' $m \in A$:

Always plays action m. (e.g., preference s.t. strictly dominant to play m)

Arrives with probability $K\kappa_m \geq 0$, where $\kappa_m, K \in [0,1]$ and $\sum_{m \in \mathcal{A}} \kappa_m = 1$.

'Rational type' $r \in R = \{1, ..., N_R\}$:

Type r determines preference $u(a, \theta, r)$.

Arrives with probability $(1 - K)\rho_r > 0$.

N_A^r non-weakly dominated actions.

Order actions $a_1^r, ..., a_{N_A^r}$ s.t. $u(a_1^r, H, r) < ... < u(a_{N_A^r}, H, r)$.

I_a^r, J_a^r, etc. all defined the same.

Cascade from t if $q_t \in J := \cap_r J^r$. Common knowledge we're in a cascade.

Limit cascade $q_\infty \in J$.

Robustness of Cascades

Private Interim Beliefs:

Private beliefs bounded if $0 < \underline{b} < \bar{b} < 1$; unbounded if $0 = \underline{b} < \bar{b} = 1$.
(Recall $\bar{b} = \sup \operatorname{supp} P_P$ and $\underline{b} = \inf \operatorname{supp} P_P$.)

$P_P^{\theta}(\cdot) = P_P(\cdot \mid \theta)$. Assume P_P^H and P_P^L mutually abs. continuous (no fully-revealing signals wp> 0).

Immediate that with unbounded private beliefs, $J^r_{a^r_1} = \{0\}$ and $J^r_{a^r_{N^r_A}} = \{1\}$, and for other actions $J^r_a = \emptyset$.

Robust and Less Robust Results

Theorem

A fully wrong information cascade a.s. never starts.

Condition on $\theta = H$.

(a) With a single 'rational type', a not-fully-wrong limit cascade occurs a.s.

(b) With a single 'rational type' and unbounded private beliefs there is complete learning: $q_t \to \mathbf{1}_{\{\theta=H\}}$ a.s.

(c) With bounded private beliefs and $q_t \notin J$, a wrong limit cascade happens with positive probability.

(d) Letting the support of private beliefs $[\underline{b}_n, \overline{b}_n] \to [0, 1]$, the chance of a wrong limit cascade vanishes (continuity at the limit).

Robust and Less Robust Results

Cascade a.s. in finite time artefact of finite S

For generic atomless P_P, prob contrarian agent becomes vanishingly small when public belief tends to cascade set.

Hence, from martingale property, variation in public belief converges to zero Public belief never actually reaches the cascade set. (see Chamley 2010 Ch. 4.2.)

Herds vs Cascades

Limit cascade implies herd (with finite A). Can have herd without a cascade being triggered.

Herd is triggered a.s. in finite time.

Fragile Herds

Cascades tend to be fragile wrt small shocks.

Overview

Further Topics

Social Non-Learning, See Kartik, Lee, Liu, and Rappoport (2024 Ecta).

- Binary state: Except when private beliefs are unbounded, learning is not complete and may herd on wrong action.
- Multiple states: unbounded beliefs okay; weaker condition on info + condition on preferences is sufficient
(intermediate preferences and subexponential location-shift information).

Heterogeneous Precision and Order of Moves

- Ottaviani and Sørensen (2001 JPubE): Social learning + reputation concerns. Study order of moves.
- Holding aggregate precision fixed, better to have it concentrated in single person.
- The anti-seniority rule not necessarily optimal.
- Increasing the precision of agent can harm.

Limited Observability

- Random sample of past actions (Smith and Sørensen, 2020). Smaller samples can improve efficiency (Chamley 2020, Ch. 5).
- Ability to observe others' actions/payoffs may cause insufficient exploration. E.g., favourable reviews deter exploration (Acemoglu, Makhdoumi, Malekian, and Ozdaglar (2022 Ecta)).

Further Topics

Endogeneous Timing in BHW model often just expands set equilibria.

See Chamley and Gale (1994 Ecta), Chamley (2010 Ch. 6).

See also Gul and Lundholm (1995 JPE), Caplin and Leahy (1994 AER).

Overconfidence can help (Bernardo and Welch (2001 JEMS); Arieli, Babichenko, Müller, Pourbabaee, and Tamuz (2024)).

Correlation neglect can harm (Eyster and Rabin (2010 AER)).

Non-EU Maximising

Uncertainty wrt others' signal precision and uncertainty aversion can deter breaking cascades even in unbounded private beliefs case (Chen (2025 AER)).

Further Topics

Costly Info

Ali (2018 JET): Flexible info.

Complete learning with finite A and costly signals iff $\mathbb{P}(\text{cost} \leq \varepsilon) > 0$ for any $\varepsilon > 0$.

Similar insight to unbounded beliefs. Also similar result if agents pay search costs for history (Mueller-Frank Pai (2016 AEJMicro)).

Changing State

Levy, Pęski, and Vieille (2024 Ecta).

Population each period. Agents sample actions from previous period and can acquire costly private signal.

At eqm, always acquire private signal and learning incomplete even with very precise signals.

(See also Dasaratha, Golub, and Hak (2023 REStud) for changing state + networks)

More: Social learning in markets, networks, contracting, ...

See Bikhchandani, Hirshleifer, Tamuz, and Welch (2024 JEL).

Further Topics

Testing for Social Learning

Technology adoption: wheat and rice varieties in India (Foster and Rosenzweig 1995 JPE, Munshi 2004 JDevE), new crop adoption in Mozambique (Bandiera and Rasul 2006 EJ), pineapple in Ghana (Conley and Udry 2010 AER), maize in Malawi (BenYishay and Mobarak 2019 REStud), microfinance in India (Banerjee, Breza, Chandrasekhar, Duflo, Jackson, and Kinnan 2024 REStud).

(Testing of diffusion and social learning was a hot topic in the field/dev RCT revolution.)

Social learning on networks: Mobius, Phan, and Szeidl (2015 NBER WP); Chandrasekhar, Larreguy, and Xandri (2020 Ecta).

Correlation neglect: Enke and Zimmermann (2019 REStud); Angrisani, Guarino, Jehiel, and Kitagawa (2018 AEJMicro).

Metastudy of experiments: Weizsäcker (2010 AER)

Social Learning

Duarte Gonçalves

University College London

Topics in Economic Theory