Pathological Outcomes of Observational Learning

by Lones Smith and Peter Sorensen

Econometrica, March 2000

In the classic herding models of Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992), we see both:

1. Herds–starting at some point, all agents take the same action,

2. Information Cascades–starting at some point, behavior is not informative and beliefs do not change.

Smith and Sorensen offer a more general framework with heterogeneous preferences and continuous signal distributions, where herds and cascades are de-coupled.

Suppose that a herd has started, and someone chooses an action contrary to the herd action.

–Maybe the last agent had a strong enough signal. Model should allow for a continuous signal distribution.

–Maybe the last agent was irrational or the choice was an accident. Allow for noise.

–Maybe the last agent has different preferences. Allow for multiple preference types.

What are the possible "pathological" outcomes where learning ceases?

–type-specific herds.

–confounded learning, or a cascade in which history is uninformative and agents act based on their private signals.

Motivating Example: Texas highway drivers, with two states H and L. Some drivers are going to Houston and some to Dallas. Payoffs depend on the type (destination) and the state.

In state H, Houston types should take the high road and Dallas types should take the low road.

In state L, Houston types should take the low road and Dallas types should take the high road.

Suppose that 70% of drivers are Houston types and 30% are Dallas types.

What sorts of cascades are possible?

1. A type-specific herd. Eventually, all Houston types take the high road and all Dallas types take the low road. (Absent a strong signal to the contrary, a Dallas driver should take the less crowded road and a Houston driver should take the more crowded road.)

2. A type-specific herd in which all Houston types take the low road and all Dallas types take the high road.

3. Confounded learning. Beliefs about the probability of state H, denoted by q, are such that the fraction of drivers taking the high road is the same in both states.

$$\psi_H(q) = \psi_L(q).$$

Since the public history of choices is completely uninformative, drivers of both types base their decisions on private signals.

We know that we have

$$\begin{aligned} \psi_H(0) &= \psi_L(0) = 0.3 \\ \psi_H(1) &= \psi_L(1) = 0.7 \end{aligned}$$

Since ψ_H and ψ_L are continuous in q, we have confounded learning if there is an interior intersection.

Smith and Sorensen show that confounded learning exists robustly, and that the dynamics converging to confounded learning are locally stable.

The Model

Two states equally likely ex ante, $S = \{H, L\}$

An infinite sequence of agents, $1, 2, ..., n, ...$

Private belief of nth agent that state is H (belief based only on one's signal) is denoted by p_n

Assume that private beliefs conditional on the state are i.i.d., with distribution function F^s and density function f^s. Also, no signal *perfectly* reveals the state.

Action set has two elements, $\{1, 2\}$.

Game Timing: For each n, agent n observes her private signal and choices of previous agents, and chooses an action.

There are $t \geq 1$ rational types and 2 "crazy" types. A fraction of the population, κ_1 always choose action 1, and a fraction of the population, κ_2 always choose action 2. The fraction of rational types is $\kappa = 1 - \kappa_1 - \kappa_2$.

In equilibrium, agents know the strategies of the other agents, and can therefore compute the probability that the state is H after any history of choices of previous agents. We call this probability the public belief, denoted by q.

Given public belief q and private belief p, posterior belief is

$$r(p, q) = \frac{pq}{pq + (1-p)(1-q)}. \tag{1}$$

Define the public likelihood ratio (of state L vs. state H) before agent n observes her signal as

$$\ell_n = \frac{1 - q_n}{q_n}. \tag{2}$$

Fact: The stochastic process $< q_n >$ is a martingale, so we have $E(q_{n+1}|q_n) = q_n$. Because q_n is bounded, it almost surely converges to a random variable. That is, the sequence converges to some limit point, but the limit point is random.

Also, *given state* H, $< \ell_n >$ is a convergent martingale. To see this,

$$\begin{aligned}\ell_{n+1} &= \frac{pr(L|m_n, \ell_n)}{pr(H|m_n, \ell_n)} = \frac{pr(\ell_n)pr(L|\ell_n)pr(m_n|\ell_n, L)}{pr(\ell_n)pr(H|\ell_n)pr(m_n|\ell_n, H)} \\ &= \ell_n \frac{pr(m_n|\ell_n, L)}{pr(m_n|\ell_n, H)}.\end{aligned}$$

Therefore, we have

$$\begin{aligned}E(\ell_{n+1}|\ell_n, H) &= \sum_{m_n} pr(m_n|\ell_n, H)\left[\ell_n \frac{pr(m_n|\ell_n, L)}{pr(m_n|\ell_n, H)}\right] \\ &= \ell_n.\end{aligned}$$

To finish the description of the game, we need payoffs. For $m = 1, 2$, the payoff of an agent of type t with posterior beliefs r from action m is

$$ru_t^H(m) + (1 - r)u_t^L(m). \tag{3}$$

Here $u_t^s(m)$ is a parameter of the game representing the payoff from action m when the state is known to be s.

Denote a type-t agent's preferred action in state H as a_2^t and her unpreferred action in state H as a_1^t.

Note: one interpretation of "crazy" types is that they have the same preferred action in both states.

Then in equilibrium agents use a threshold strategy that depends on the type and the public likelihood ratio. There is a threshold belief $p^t(\ell)$ such that a_2^t is chosen if and only if $p^t \geq p^t(\ell)$.

For some likelihood ratios, it is possible that all type-t agents choose the same action, independent of their signal. This defines two cascade sets, with boundaries $\underline{\ell}$ and $\overline{\ell}$.

For $\ell \leq \underline{\ell}$, all type-t agents choose action a_2^t, and for $\ell \geq \overline{\ell}$, all type-t agents choose action a_1^t.

If $0 < \underline{\ell} < \overline{\ell} < \infty$ holds, then private beliefs are said to be *bounded*, and if $\underline{\ell} = 0$ and $\overline{\ell} = \infty$ holds, then private beliefs are said to be *unbounded*.

Some more notation:

$\rho^t(m|s,\ell)$ denotes the probability that a type-t agent chooses action m given state s and public likelihood ratio ℓ.

$\psi(m|s,\ell)$ denotes the probability that an agent of unknown type chooses action m.

λ^t denotes the proportion of rational agents who are type t. Thus, $\lambda^1 + \lambda^2 = 1$.

$\varphi(m,\ell)$ denotes next period's public likelihood ratio when the current period's ratio is ℓ and the agent chooses action m.

Equilibrium Transitions

$$\ell_0 = 1, q_0 = \frac{1}{2}$$

$$\rho^t(a_2^t|s,\ell) = 1 - F^s(p^t(\ell)) \tag{4}$$

$$\rho^t(a_1^t|s,\ell) = F^s(p^t(\ell))$$

$$\psi(m|s,\ell) = \kappa_m + \kappa \sum_{t=1}^{2} \lambda^t \rho^t(m|s,\ell) \tag{5}$$

$$\varphi(m,\ell) = \ell \frac{\psi(m|L,\ell)}{\psi(m|H,\ell)} \tag{6}$$

Example 1

One rational type, so $\lambda^1 = 1$

$$\begin{aligned} u^H(1) &= u^L(1) = 0 \\ u^H(2) &= u \\ u^L(2) &= -1 \end{aligned}$$

(Think of action 2 as "investing," which pays off in state H.)

From (3), an agent is indifferent between actions when posteriors are given by

$$r = \frac{1}{1+u}.$$

From (1) and (2), the belief threshold is therefore

$$p(\ell) = \frac{\ell}{u+\ell}. \tag{7}$$

Example 1A: Unbounded beliefs with no crazy types.

Suppose agents receive private signals, $\sigma \in (0,1)$, with densities

$$\begin{aligned} g^H(\sigma) &= 2\sigma \\ g^L(\sigma) &= 2(1-\sigma). \end{aligned}$$

Then we can calculate $p = \sigma$, so the density and distribution functions for private beliefs are given by

$$\begin{aligned} f^H(p) &= 2p \text{ and } F^H(p) = p^2 \\ f^L(p) &= 2(1-p) \text{ and } F^L(p) = 2p - p^2. \end{aligned}$$

From (7) and the fact that p can be arbitrarily close to 0 and to 1, it is clear that we have unbounded beliefs.

Substitution into (4)-(6) yields

$$
\begin{aligned}
\rho(1|H,\ell) &= \frac{\ell^2}{(u+\ell)^2} \\
\rho(1|L,\ell) &= \frac{\ell(\ell+2u)}{(u+\ell)^2} \\
\psi(m|s,\ell) &= \rho(m|s,\ell) \\
\varphi(1,\ell) &= \ell\frac{\psi(1|L,\ell)}{\psi(1|H,\ell)} = \ell + 2u \\
\varphi(2,\ell) &= \ell\frac{\psi(2|L,\ell)}{\psi(2|H,\ell)} = \frac{u\ell}{u+2\ell}.
\end{aligned}
$$

Given state H (w.l.o.g.), the likelihood ratio is a convergent martingale. Convergence requires either $\varphi(1, \ell) = \ell$ or $\varphi(2, \ell) = \ell$, from which we conclude:

$\ell_\infty = 0$ with probability one. The probability of an infinite subsequence of action 1 is zero, so there is a herd on action 2, the correct choice in state H.

The situation is very different from discrete examples, like BHW, where there is a positive probability of a herd on the wrong choice.

An interesting feature of the equilibrium is that, when ℓ is close to zero and there has been a string of action 2 choices, there is always a positive probability that an agent gets a signal below $p(\ell) = \frac{\ell}{u+\ell}$. In that case, the agent will choose action 1 and beliefs change dramatically to $\ell + 2u$.

Example 1B: Bounded beliefs with no crazy types.

Replace the previous signal densities with

$$\begin{aligned} g^H(\sigma) &= 1 \\ g^L(\sigma) &= \frac{3}{2} - \sigma. \end{aligned}$$

Then we can calculate

$$p(\sigma) = \frac{2}{5 - 2\sigma}$$

which implies

$$\begin{aligned} F^H(p) &= \frac{5p - 2}{2p} \\ F^L(p) &= \frac{(5p - 2)(p + 2)}{8p^2}. \end{aligned}$$

The range of possible private beliefs, as σ ranges from 0 to 1, are: $\frac{2}{5} < p < \frac{2}{3}$.

Since $p(\ell) = \frac{\ell}{u+\ell}$, the range of likelihood ratios is given by $\underline{\ell} = \frac{2u}{3}$ and $\overline{\ell} = 2u$.

Dynamics

If $u \geq \frac{3}{2}$, we have $\ell_0 \leq \frac{2u}{3}$, and we herd on action 2 from the beginning.

If $u \leq \frac{1}{2}$, we have $\ell_0 \geq 2u$, and we herd on action 1 from the beginning.

For the interesting case, $\frac{2u}{3} < \ell_0 < 2u$, equations (4)-(6) imply

$$\begin{aligned}
\rho(1|H,\ell) &= \frac{3\ell - 2u}{2\ell} \\
\rho(1|L,\ell) &= \frac{(3\ell - 2u)(3\ell + 2u)}{8\ell^2} \\
\psi(m|s,\ell) &= \rho(m|s,\ell) \\
\varphi(1,\ell) &= \ell\frac{\psi(1|L,\ell)}{\psi(1|H,\ell)} = \frac{3\ell}{4} + \frac{u}{2} \\
\varphi(2,\ell) &= \ell\frac{\psi(2|L,\ell)}{\psi(2|H,\ell)} = \frac{\ell}{4} + \frac{u}{2}.
\end{aligned}$$

Given state H, the likelihood ratio converges, which requires either $\varphi(1, \ell) = \ell$ or $\varphi(2, \ell) = \ell$.

Now there are two possibilities, $\ell_\infty = 2u$, in which case we herd on action 1, or $\ell_\infty = \frac{2u}{3}$, in which case we herd on action 2.

Since we have a martingale, $E(\ell_\infty | H) = \ell_0$, so we can compute the probability of an action 2 herd, π, solving

$$\ell_0 = \pi(\frac{2u}{3}) + (1 - \pi)2u.$$

Note: for $\frac{2u}{3} < \ell_0 < 2u$, beliefs never enter the cascade set, even though a herd starts with probability one. There is always a (vanishing) probability that an agent will go against the herd. If so, beliefs change drastically.

Example 1C: Bounded beliefs with "crazy" types.

Introducing crazy types to the previous example will affect the dynamics, but not the cascade sets.

We have $\underline{\ell} = \frac{2u}{3}$ and $\overline{\ell} = 2u$.

But here, unlike the previous example, when ℓ is near $\frac{2u}{3}$ or $2u$, beliefs are continuous in actions. That is,

$$\varphi(1, \frac{2u}{3}) = \ell \quad \text{and} \quad \varphi(2, 2u) = \ell.$$

Extremely unlikely actions are attributed to noise.

Example 2: Two rational types with opposing preferences. No noise, bounded beliefs.

type "U" has preferences

$$u^H(1) = 0, \quad u^H(2) = u, \quad u^L(1) = 1, \quad u^L(2) = 0,$$

type "V" has preferences

$$u^H(1) = 1, \quad u^H(2) = 0, \quad u^L(1) = 0, \quad u^L(2) = v.$$

w.l.o.g., assume $v \geq u$.

Thresholds for the two types are:

$$p^U(\ell) = \frac{\ell}{u + \ell} \text{ and } p^V(\ell) = \frac{\ell}{v + \ell}.$$

For $p > p^U(\ell)$, type U chooses action 2.

For $p > p^V(\ell)$, type V chooses action 1.

Assume the same bounded belief structure as in Example 1B:

$$
\begin{aligned}
g^L(\sigma) &= \frac{3}{2} - \sigma \\
g^L(\sigma) &= 1 \\
p(\sigma) &= \frac{2}{5 - 2\sigma} \\
F^H(p) &= \frac{5p - 2}{2p} \\
F^L(p) &= \frac{(5p - 2)(p + 2)}{8p^2},
\end{aligned}
$$

and suppose that $2u > \frac{2v}{3}$ and $\ell_0 \in (\frac{2v}{3}, 2u)$.

Transitions for type U are the same as in Example 1B:

$$\rho^U(1|H,\ell) = \frac{3\ell - 2u}{2\ell}$$

$$\rho^U(1|L,\ell) = \frac{(3\ell - 2u)(3\ell + 2u)}{8\ell^2}.$$

Transitions for type V can be computed as

$$\rho^V(1|H,\ell) = \frac{2v - \ell}{2\ell}$$

$$\rho^V(1|L,\ell) = \frac{(2v + \ell)(2v - \ell)}{8\ell^2}.$$

The cascade sets are:

$\ell \in [0, \frac{2u}{3}]$ (all type U choose action 2)

$\ell \in [2v, \infty]$ (all type V choose action 2)

$\ell \in [2u, \infty]$ (all type U choose action 1)

$\ell \in [0, \frac{2v}{3}]$ (all type V choose action 1)

From (5), we have

$$\psi(1|H,\ell) = \lambda^U \left[\frac{3\ell - 2u}{2\ell}\right] + \lambda^V \left[\frac{2v - \ell}{2\ell}\right] \quad (8)$$

$$\psi(1|L,\ell) = \lambda^U \left[\frac{(3\ell - 2u)(3\ell + 2u)}{8\ell^2}\right] + \lambda^V \left[\frac{(2v + \ell)(2v - \ell)}{8\ell^2}\right]. \quad (9)$$

If there is an ℓ^* such that $\psi(1|H,\ell^*) = \psi(1|L,\ell^*)$ holds, then actions are uninformative when $\ell = \ell^*$.

It would follow that $\psi(2|H,\ell^*) = \psi(2|L,\ell^*)$ holds as well, so we have

$$\varphi(1,\ell^*) = \ell^* \frac{\psi(1|L,\ell^*)}{\psi(1|H,\ell^*)} = \ell^* \quad \text{and}$$

$$\varphi(2,\ell^*) = \ell^* \frac{\psi(2|L,\ell^*)}{\psi(2|H,\ell^*)} = \ell^*.$$

When $\ell = \ell^*$, beliefs stay at ℓ^* no matter what action is chosen. Confounded learning.

Equating the right sides of (8) and (9), we have confounded learning at ℓ^* iff

$$\frac{\lambda^U}{\lambda^V} = \frac{(2v - \ell)(3\ell - 2v)}{(2u - \ell)(3\ell - 2u)} \equiv h(\ell)$$

for $\ell = \ell^*$.

Since we have $\ell < 2u$ and $\ell > \frac{2v}{3} > \frac{2u}{3}$, $h(\ell)$ continuously maps $(\frac{2v}{3}, 2u)$ onto $(0, \infty)$.

Therefore, for any $\frac{\lambda^U}{\lambda^V}$, there exists a likelihood ratio with confounded learning.

One can show that $\varphi(1, \ell)$ and $\varphi(2, \ell)$ are both increasing in ℓ.

Dynamics away from ℓ^*:

For $\ell \in (\frac{2u}{3}, \ell^*)$ we have:

$\varphi(2, \ell) < \ell$

[after action 2, ℓ goes down but not below $\frac{2u}{3}$]

$\varphi(1, \ell) > \ell$

[after action 1, ℓ goes up but not above ℓ^*]

For $\ell \in (\ell^*, 2v)$ we have:

$\varphi(2, \ell) > \ell$

[after action 2, ℓ goes up but not above $2v$]

$\varphi(1, \ell) < \ell$

[after action 1, ℓ goes down but not below ℓ^*]

In either interval, ℓ is a bounded martingale that must converge to one of the endpoints.

If $\ell_0 \in (\frac{2u}{3}, \ell^*)$, it follows that ℓ_∞ must put positive probability on ℓ^* (confounded learning) and positive probability on the type-specific herd where type U chooses action 2 and type V chooses action 1.

If $\ell_0 \in (\ell^*, 2v)$, it follows that ℓ_∞ must put positive probability on ℓ^* (confounded learning) and positive probability on the type-specific herd where type U chooses action 1 and type V chooses action 2.

For this example, ℓ^* (confounded learning) must be locally stable.

General Results

Theorem 1: *Suppose w.l.o.g. that the state is H.*

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

(b) With a single rational type and unbounded private beliefs, $\ell_n \to 0$ almost surely.

(c) With $T \geq 2$ rational types with different preferences, only a limit cascade that is not fully wrong or a confounded learning outcome may arise.

Theorem 2: *Assume there are $T \geq 2$ rational types.*

(d) If belief distributions are discrete, confounded learning is nongeneric.

(e) With $M > 2$ actions and unbounded beliefs, confounded learning is nongeneric.

(f) At any confounding outcome, some pair of types has opposed preferences.

(g) Assume $M = 2$ and some types with opposing preferences. With atomless bounded beliefs and $T = 2$, a confounded learning point exists generically, provided both types are active over some public belief range. With atomless unbounded beliefs and $f^H(1), f^L(0) > 0$, a confounding point exists if the opposed types have sufficiently different preferences.

Theorem 3: *Assume a single rational type and no noise.*

(a) A herd on some action will almost surely arise in finite time.

(b) With unbounded private beliefs, individuals almost surely settle on the optimal action.

(c) With bounded private beliefs, absent a cascade on the most profitable action from the outset, a herd arises on another action with positive probability.

Theorem 5: *(c) For nondegenerate parameters, a confounded learning point ℓ^* is locally stable.*