Lecture 23: Observational Learning and Herd Behavior

Alexander Wolitzky

MIT

6.207/14.15: Networks, Spring 2022

Bayesian Social Learning in Networks

For our last topic, we return to social learning—how do individuals form beliefs/opinions from others in their social network?—but now from a Bayesian perspective.

Lots of empirical evidence that individuals form beliefs from talking to or observing others, about a wide range of important issues:

- Information about jobs and investment opportunities
- Financial planning (how/how much to save)
- Product choice (which brand to buy, what book/movie/restaurant to consume)
- Technology choice (which production technology to use, e.g. which medical treatment to use, which crop to plant)
- Voting (who to vote for)

These Issues Could Not be More Timely

- Disinformation and “fake news”
- Political persuasion and polarization
- Groupthink and herd behavior
- Privacy and spamming

How should government/society regulate traditional and social media, and more generally our shared informational environment?

Models of Social Learning

We previously studied social learning through the DeGroot model: repeated averaging of neighbors' opinions.

Now we take the Bayesian perspective that agents rationally update their beliefs based on what they see others do.

Main question for today: when does rational observational learning eventually lead to correct learning, and when does it lead to "herd behavior"?

- And how does this depend on the network?

Observational Learning and Herd Behavior

We'll cover the **herding model** (also known as the **sequential social learning model**), due to Banerjee (1992) and Bikchandani-Hirshleifer-Welch (1992).

The model explains why and when rational agents "herd" by copying others' (possibly incorrect decisions) rather than relying on their own information.

Many theoretical extensions and empirical applications.
Among the latter:

- Do stock analysts rely on their own information or just copy others' investments?
- What determines the course of asset "bubbles" or "panic selling"?
- What determines "fads" for certain goods?

Model

Underlying informational environment similar to the static voting model from last class:

- Payoff-relevant state $\theta \in \{0, 1\}$.
- Prior belief is $\Pr(\theta = 1) = p$.
- N individuals. Each individual i gets conditionally iid signal $s_i \in \{0, 1\}$, with $\Pr(s_i = \theta) = q > \frac{1}{2}$.

Key differences from last class:

- Instead of the group making a collective decision $x \in \{0, 1\}$ (where the group wants to take $x = \theta$), each individual takes her own decision $x_i \in \{0, 1\}$ (and each wants to take $x_i = \theta$).
- Instead of all voting simultaneously, the agents now move in **sequence**, and **each observes the earlier agents' actions.**
- Can think of this as social learning on the line network, where everyone observes their predecessors' action. Later on consider more general networks.

Story/Example

- There are two new restaurants in town: French (restaurant 0) and Italian (restaurant 1).
- Prior belief of everyone in town is that Italian is better w/ prob p.
- Everyone in town gets an independent piece of information about which restaurant is better (the signals s_i).
- At 6pm the first night the restaurants open, the first customer chooses where to go based on her signal.
- The second customer sees where the first customer goes (and can thus infer her signal), and chooses where to go based on this and his own signal.
- And so on.

Story/Example (cntd.)

- Suppose that in fact French is better, but the first 5 customers get signals that Italian is better (unlikely but possible).
- The first 5 customers go to Italian.
- Suppose customer 6 gets signal that French is better.

Story/Example (cntd.)

- Suppose that in fact French is better, but the first 5 customers get signals that Italian is better (unlikely but possible).
- The first 5 customers go to Italian.
- Suppose customer 6 gets signal that French is better.
- Since the first 5 customers went to Italian, this may mean they all got signals that Italian was better (whether or not it means this depends on their equilibrium strategies, as we'll see).

Story/Example (cntd.)

- Suppose that in fact French is better, but the first 5 customers get signals that Italian is better (unlikely but possible).
- The first 5 customers go to Italian.
- Suppose customer 6 gets signal that French is better.
- Since the first 5 customers went to Italian, this may mean they all got signals that Italian was better (whether or not it means this depends on their equilibrium strategies, as we'll see).
- If it does, customer 6 has to weigh 1 pro-French signal against 5 pro-Italian signals. She may well decide to go to the Italian restaurant despite her own signal.

Story/Example (cntd.)

- Suppose that in fact French is better, but the first 5 customers get signals that Italian is better (unlikely but possible).
- The first 5 customers go to Italian.
- Suppose customer 6 gets signal that French is better.
- Since the first 5 customers went to Italian, this may mean they all got signals that Italian was better (whether or not it means this depends on their equilibrium strategies, as we'll see).
- If it does, customer 6 has to weigh 1 pro-French signal against 5 pro-Italian signals. She may well decide to go to the Italian restaurant despite her own signal.
- If she does, then her signal is "lost to society": no future customers know about this pro-French signal, since all they see is everyone going to the Italian restaurant.

Story/Example (cntd.)

- Suppose that in fact French is better, but the first 5 customers get signals that Italian is better (unlikely but possible).
- The first 5 customers go to Italian.
- Suppose customer 6 gets signal that French is better.
- Since the first 5 customers went to Italian, this may mean they all got signals that Italian was better (whether or not it means this depends on their equilibrium strategies, as we'll see).
- If it does, customer 6 has to weigh 1 pro-French signal against 5 pro-Italian signals. She may well decide to go to the Italian restaurant despite her own signal.
- If she does, then her signal is "lost to society": no future customers know about this pro-French signal, since all they see is everyone going to the Italian restaurant.
- The same logic applies to customer 7, who will also go Italian regardless of her own signal.

Story/Example (cntd.)

- Suppose that in fact French is better, but the first 5 customers get signals that Italian is better (unlikely but possible).
- The first 5 customers go to Italian.
- Suppose customer 6 gets signal that French is better.
- Since the first 5 customers went to Italian, this may mean they all got signals that Italian was better (whether or not it means this depends on their equilibrium strategies, as we'll see).
- If it does, customer 6 has to weigh 1 pro-French signal against 5 pro-Italian signals. She may well decide to go to the Italian restaurant despite her own signal.
- If she does, then her signal is "lost to society": no future customers know about this pro-French signal, since all they see is everyone going to the Italian restaurant.
- The same logic applies to customer 7, who will also go Italian regardless of her own signal.
- From this point on, everyone goes to the Italian restaurant: an incorrect herd has formed!

Remarks

- Everyone in the story is behaving rationally—we've described an equilibrium.
- If people saw earlier players' **signal**, by LLN eventually everyone would take the right action. The problem is that you only see earlier players' **actions**, and these actions are uninformative if the earlier players themselves are herding.
- Since everyone in the story is rational, people don't become more and more convinced that the herd is right as it continues to persist. Everyone understands that actions are uninformative once the herd forms, so no further belief updating at this point. Still, in the positive-probability event that the herd formed on the wrong action, everyone is taking the wrong action.

Analysis

Back to the model: say a **herd** forms at period T if, starting from period T, everyone takes the same action.

Theorem

If N is sufficiently large, then there exists a period T at which a herd forms with positive probability (there can be many such periods). Moreover, the herd forms on the wrong action with positive probability.

The proof fleshes out the logic of the example.

Proof Sketch

Since $q > \frac{1}{2}$, the signals convey some information.

Hence, given what she has observed about earlier players, a given player will either follow her signal (play $x = s$), or she will always take the same action regardless of her signal.

- She does the former if the information content of observing the earlier actions is weaker than her signal; does the latter otherwise.

Proof Sketch

Since $q > \frac{1}{2}$, the signals convey some information.

Hence, given what she has observed about earlier players, a given player will either follow her signal (play $x = s$), or she will always take the same action regardless of her signal.

- She does the former if the information content of observing the earlier actions is weaker than her signal; does the latter otherwise.

Since all agents face the same problem and have equally informative signals, once one player takes the same action regardless of her signal, all subsequent players will do the same.

- Once one player "follows the crowd," all later players will do so, too, because they have exactly the same "social information" as the first player to follow the crowd.

Proof Sketch

Since $q > \frac{1}{2}$, the signals convey some information.

Hence, given what she has observed about earlier players, a given player will either follow her signal (play $x = s$), or she will always take the same action regardless of her signal.

- She does the former if the information content of observing the earlier actions is weaker than her signal; does the latter otherwise.

Since all agents face the same problem and have equally informative signals, once one player takes the same action regardless of her signal, all subsequent players will do the same.

- Once one player "follows the crowd," all later players will do so, too, because they have exactly the same "social information" as the first player to follow the crowd.

Note: If in equilibrium the play of players 1 through 5 induces player 6 to follow the crowd, then player 7 understands this and hence does not draw any inference from player 6's action.

Proof Sketch (cntd.)

Therefore, every equilibrium must take the following form:

- Up to some time T (random, and possibly $= \infty$), everyone follows her signal.
- Starting at time T, everyone takes the same action regardless of their signals.

It remains to prove that with positive probability, T is finite and the herd forms on the wrong action.

Proof Sketch (cntd.)

Suppose $\theta = 0$, but up to T everyone has gotten signal $s = 1$.

- By hypothesis, up to T everyone follows her signal, so up to T everyone plays $x = 1$.

Then if player T gets signal $s = 0$, her posterior belief that $\theta = 1$ is

$$\frac{\Pr\left(\theta = 1 \cap \left(T-1 \text{ right signals}\right) \cap \left(1 \text{ wrong signal}\right)\right)}{\left(\begin{array}{c} \Pr\left(\theta = 1 \cap \left(T-1 \text{ right signals}\right) \cap \left(1 \text{ wrong signal}\right)\right) \\ +\Pr\left(\theta = 0 \cap \left(T-1 \text{ wrong signals}\right) \cap \left(1 \text{ right signal}\right)\right) \end{array}\right)}$$

$$= \frac{pq^{T-1}\left(1-q\right)}{pq^{T-1}q + \left(1-p\right)\left(1-q\right)^{T-1}q}$$

$$= \frac{1}{1 + \frac{1-p}{p}\left(\frac{1-q}{q}\right)^{T-2}}$$

Since $q > \frac{1}{2}$, this converges to 1 as $T \to \infty$.

Proof Sketch (cntd.)

Hence, there is some first period T as which this posterior belief is above the cutoff belief for $x = 1$ to be optimal.

- At this point, a herd forms on action 1 (the wrong action).
- This sequence of events occurs with probability $(1 - p)(1 - q)^{T-1} q > 0$.

An Alternative Assumption

What if instead of assuming that $s_i \in \{0, 1\}$ with

$$\Pr(s_i = \theta) = q > \frac{1}{2},$$

we'd assumed $s_i \in \{-1, 0, 1, 2\}$ with

$$\begin{aligned} \Pr(s_i = \theta) &= q, \\ \Pr(s_i = \theta - 1) &= \frac{1-q}{2}, \\ \Pr(s_i = \theta + 1) &= \frac{1-q}{2}, \end{aligned}$$

so now signal is either equal to state or is off by 1.

In particular $s_i = -1$ can **only** happen if $\theta = 0$, while $s_i = 2$ can **only** happen if $\theta = 1$.

Can an incorrect herd form now?

An Alternative Assumption

What if instead of assuming that $s_i \in \{0, 1\}$ with

$$\Pr(s_i = \theta) = q > \frac{1}{2},$$

we'd assumed $s_i \in \{-1, 0, 1, 2\}$ with

$$\begin{aligned} \Pr(s_i = \theta) &= q, \\ \Pr(s_i = \theta - 1) &= \frac{1-q}{2}, \\ \Pr(s_i = \theta + 1) &= \frac{1-q}{2}, \end{aligned}$$

so now signal is either equal to state or is off by 1.

In particular $s_i = -1$ can **only** happen if $\theta = 0$, while $s_i = 2$ can **only** happen if $\theta = 1$.

Can an incorrect herd form now? **No**.

- A player with $s_i = -1$ or $s_i = 2$ learns the state with certainty and will therefore always take the correct action.
- This overturns any incorrect herd.

Implication

Incorrect herds form with positive probability whenever a sufficiently strong "social signal" can overturn any private signal.

Incorrect herds cannot form whenever a sufficiently strong private signal can overturn any social signal.

- Formally, the difference corresponds to whether the likelihood ratio of private signals $\Pr(s|\theta = 0) / \Pr(s|\theta = 1)$ is bounded or unbounded.

Implication

Incorrect herds form with positive probability whenever a sufficiently strong "social signal" can overturn any private signal.

Incorrect herds cannot form whenever a sufficiently strong private signal can overturn any social signal.

- Formally, the difference corresponds to whether the likelihood ratio of private signals $\Pr(s|\theta = 0) / \Pr(s|\theta = 1)$ is bounded or unbounded.

Implication: A small share of "experts" in the population can prevent herding, but they must have a high degree of expertise (i.e., very informative signals) and everyone must understand that they really are experts (i.e., an unexpected action against the herd must be viewed as a sign of expertise, not error).

Richer Observation Networks

Another special feature of the basic herding model is that the network of observations is extremely simple.

- Line network: everyone observes everyone before her in the line.

More generally, Acemoglu, Dahleh, Lobel, Ozdaglar (2008) show that social learning fails arise with positive probability whenever there exists a finite set of agents S who are **excessively influential**, meaning that there is an infinite set of agents T such that with positive probability every agent in T observes the actions of only agents in S.

Conversely, if

1. no set of agents is excessively influential, and
2. private signal informativeness is unbounded,

then social learning succeeds with probability 1.

- We omit the proof.

Sacrificial Lambs

Another case where incorrect herds cannot arise, even with bounded private signals:

- Consider the line example, but where every 100th agent sees only her private signal and **not** earlier agents' actions (and this information structure is common knowledge among all agents).
- These few agents must act based on their private signals.
- Eventually, everyone else will be able to learn the state based on their actions alone.

Sacrificial Lambs (cntd.)

Uninformed agents who end up facilitating social learning by following their private signals are sometimes called **sacrificial lambs**.

- Observing others is necessary for social learning to work, but having a few people in society who **don't** observe others can help social learning by injecting new information into the system.
- This is a simple example of the value of diversity in society for social learning: the presence of a few people with different information (even if it's very little information) prevents herding or "groupthink."

Lecture Summary

- Social learning on networks is an important and active topic, with many different approaches and applications to pressing social and political issues.
- Observational learning can lead to herd behavior, as subsequent information is lost to society once a herd forms.
- Whether or not a wrong herd can form depends on the strength of the most informative private signals (the presence of "experts") and whether there are excessively influential agents in society, or more generally whether the network structure ensures that new information can always enter the system.