14.771 Development Economics: Microeconomic Issues and Policy Models
Fall 2008

14.771: Recitation Handout #9
Herding and Technology Adoption

In the 1950s, economic development theory normally assumed that peasants in developing countries were overly conservative. The only way to achieve economic development was thus to IMPOSE new technology on them. The pendulum went the other way and in the mid-1970s, we started to rather emphasize that these farmers must make the optimal choice in their technology. Today, the emphasis is now on mechanisms that prevent the adoption of technology by rational farmers. We will emphasize today one paper (Banerjee, 1992) which has a suggestion why technologies may not be adopted optimally.

Herding behavior (Banerjee, 1992)

There are many real life examples where we see people choosing the same choice as the majority. It is known, for example, that voters tend to vote in the same direction as that predicted by poolsters. Similarly, researchers also seem to be focusing simultaneously on one "hot" topic of research. The purpose of this paper is to examine in what circumstances is this "herding" behavior rational. Could it be optimal to just "go with the flow"? And can this explain the lack of responsiveness of farmers to new technology?

A simple example: Choosing a restaurant

Let's start with a simple example. Let's say you're at a conference with many of your colleagues and in a new city. You have to decide between going to two restaurants, A and B. Your prior is that with 0.51 probability, A is better (let's assume that everyone shares this common prior). Each person receives a private signal about the quality of the restaurant (they heard from friends, tour guides, tour books, etc) but the signal could be wrong. All signals are assumed to be of the same quality.

Let us take the radical example where 99 out of 100 people trying to make this decision receive a signal that B is actually better but 1 receives a different signal. It is quite clear that B should be preferred if the signals are sufficiently good.

Now, let us assume that the one person who gets the A signal chooses the restaurant first. She picks to go to A because her prior and her signal tells her that this is the place to go. Now, the second person also sees that person 1 has gone to restaurant A. She then knows that person 1 received a signal for A. Given that all signals are assumed to be of the same value, her signal for B will cancel that of person 1 and she will simply use her priors and choose A. Now, person 2 would choose to go to restaurant A no matter what her signal is. Thus, her choice does not reveal any information sbout her signal to person 3. Person 3 then looks at his signal and then again, this signal is cancelled by the fact that person 1 got a signal for A. He thus uses his prior and chooses A. This will be the case for any subsequent person and everybody will go to restaurant A.

How would the outcome have been different if person 3 had been able to tell that person 2 got a signal for B? The fact that the second person ignored her signal and followed the herd is what the paper refers to as "herd externality".

A more general framework

Here are the basic assumptions of the model: There is a population of agents of size N trying to find the right option among a continuous set $i^* \in [0,1]$. Each agent may receive a signal with probability α. The signal says that the right option is $i\prime$. This signal is accurate with probability β - if the signal is wrong, i' is just a random draw with uniform weights on each i. *Is a signal with a low β (say, less than $\frac{1}{2}$) still useful? Why or why not?* Everyone is a rational Bayesian updater. *What type of equilibrium are we looking for?*

- *Payoffs*: $z > 0$ if the agent picks the right option i^* and zero otherwise.
- *Utility:* $u = z$ if they are right and $u = 0$ if they are wrong (risk neutrality)
- *Structure of the game:* Each agent must pick some $i \in [0.1]$ sequentially and people can observe the choices made and the order in which they have been made but agents cannot see whether or not agents who picked before them had a signal.
- At the end of the game, those (if any) who got the right answer get z and everyone else gets 0. If no one picked i^*, then everyone gets 0 and the right choice is not revealed.
- The structure of the game and the fact that everyone is a rational Bayesian is common knowledge

What should a strategy look like in this game?

The equilibrium depends on several tie-breaking assumptions, which are critical to the results. These rules are common knowledge - they are:

1. Whenever a decision maker has no signal and everyone else has chosen $i = 0$, she always chooses $i = 0$.
2. When decision makers are indifferent between following their own signal and following someone else's choice, they always follow their own signal
3. When a decision maker is indifferent between following more than one of the previous decision makers, she chooses to follow the one who has the highest value of i.

Do you think these decision rules will make herding more or less likely?

Now we can work out the equilibrium decision rule. First let's consider the first player:

- What will she do if she has a signal?
- What does she do if she does not?

Now consider the second player. Her strategy will be conditional on the first player's choice:

- What does she do if she has no signal?
  - $i_1 = 0$?
  - $i_1 \neq 0$?
- What does she do if she has a signal?
  - $i_1 = 0$?
  - $i_1 \neq 0$?

Now, consider the third player. What possible histories can she see? What will she do at each history?

Note - what is the probability that a player's signal matches the a location selected by a previous player *and* that the current player's signal is wrong? What does this imply about the decision rule? Suppose our third player sees both players on some $i > 0$ that does not match her signal - what should she do?

We can generalize this to note that once one option has been chosen by two people, the next person should always follow that option unless her signal matches one of the other options that have already been chosen. This is because even if the option with more than one player on it is the one with the highest i, everyone's signal is just as good as everyone else's, but if two people agree, there is some chance that this means that they both had the same signal, so the player should always pick this location rather than her own. We can generalize these arguments to get to teh first proposition in the paper: Under Assumptions 1, 2, and 3, the unique (Nash) equilibrium decision rule that everyone will adopt is decision rule D, which is characterized by the following principles:

1. The first decision makers follows her signal if she has one and chooses $i = 0$ otherwise
2. For $k > 1$, if the k^{th} decision maker has a signal, she will choose to follow her own signal either if and only if (a) holds or if (a) does not hold, (b) holds, where (a) and (b) are given by:
   (a) Her signal matches some option that has already been chosen.
   (b) No option other than $i = 0$ has been chosen by more than one person
3. Assume that the k^{th} decision maker has a signal. If any option (among those already chosen) other than the one with the highest i has been chosen by more than one person, the k^{th} decision maker will choose this option, unless her signal matches one of the other options that has already been chosen. In this case she follows her signal.
4. Assume that the k^{th} decision maker has a signal. If the option with the highest i (among those already chosen) has been chosen by more than one person, she will choose this option unless her signal matches one of the other options already chosen - in this case she follows her signal.
5. Assume that the k^{th} decision maker does not have a signal. Then she will choose $i = 0$ if and only if that is what everyone else has chosen. Otherwise, she chooses the option with the highest value of i that has already been chosen unless one of the other options (excluding $i = 0$) has been chosen by more than one person. In this case, she chooses the latter option.

The following diagram from the paper makes all of this a bit clearer:

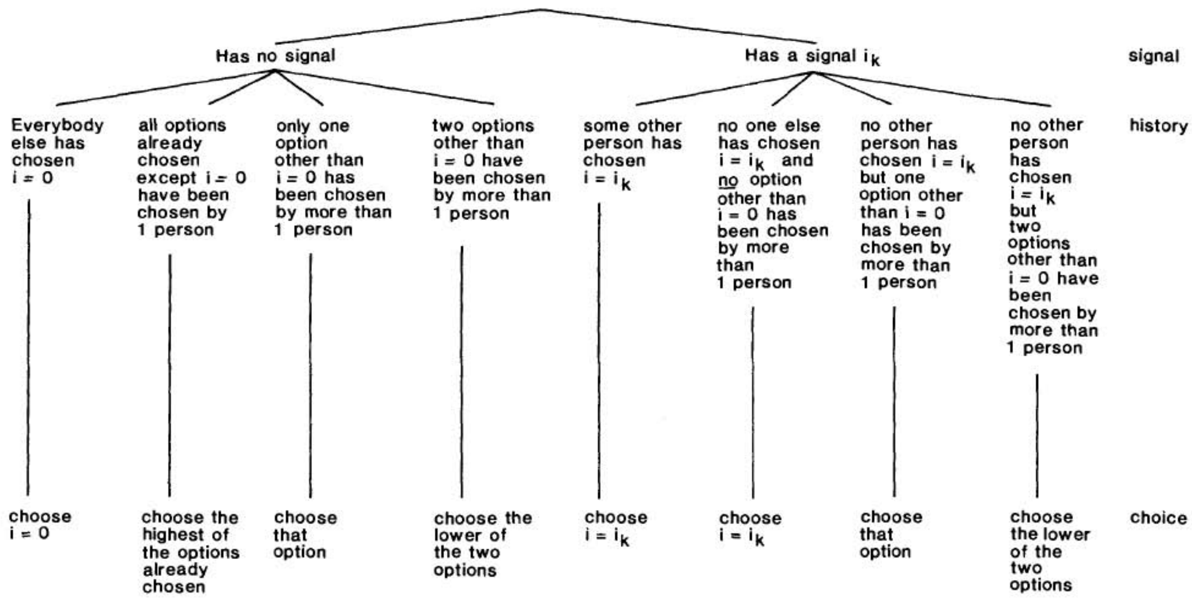

FIGURE I
The kth decision maker's choice problem ($k > 2$)

So what will someone do is she sees 2 people on $i = .45$ and 200 people on $i = .7$? What mechanism do you foresee causing the herding phenomenon?

In this set-up, there is a possibility that everybody chooses the wrong choice. If the first 6 people have signals but the 7th does not, she will pick the highest i. Everybody will then herd around that point unless someone has a signal that is identical to one already selected by the first 6.

For β small enough, the probability that no one chooses the correct option tends to 1. In particular, given α and β, the probability that the correction option is not chosen by anyone is given by:

$$\frac{(1-\alpha)(1-\beta)}{1-\alpha(1-\beta)}$$

If people were not using the information from peers, the probability of choosing the right option would be $\alpha\beta$ which can be bigger than the one where choices are observed. There is thus an information externality in seeing other people's choices

A social planner would actually implement the following rule:

- If no signal and no option has been selected twice, choose 0
- If you have a signal and no option has been selected twice, choose i_k
- If one option has been selected twice, choose that option

Can someone explain why this is a better option? Would agents in our model want to play along according to the first rule? Can you think of other mechanisms that might lead to a better outcome vs. the herding equilibrium?

Robustness

Let us look at a couple of the assumptions made in the paper and discuss whether they matter.

- what if the choices were discrete rather than continuous?
- what if agents could trade in signals?
- what if the pay-offs were not the same for everybody?

What is the key assumption that generates the information externality?

Some exercises

Exercise 1:

Nature randomly determines whether the best decision for people is to adopt a new technology (V=1) or to reject it (V=0). The probability of each outcome is 1/2.

Subsequently, a sequence of individuals has to decide whether or not to adopt the new technology. Each player takes one action: adopt or reject.

Each individual who makes the correct decision (the same decision as nature) gets a payoff of 1, regardless of the payoff to others; an individual who makes the wrong decision gets a payoff of 0.

Before making her decision, each individual i observes a signal X_i that is independent of the signals of all the other players and can take on the value of A or R with the following probabilities:

	$\Pr(X_i = A\|V)$	$\Pr(X_i = R\|V)$
$V = 1$	p	$1 - p$
$V = 0$	$1 - p$	p

and p>1/2.

All players have a flat prior about V-that is they understand that V=1 and V=0 can happen with equal probabilities. Each player forms a posterior belief about V, taking into account her own signal and the actions of all those that moved before her. As a tie-breaking convention, assume that an individual indifferent between adoption and rejection adopts or rejects with an equal probability of 1/2.

(a) How does the move of the first player depend on her signal?

(b) How does the second player decide to move? Assume that adopt is the high payoff action (V=1). Conditional on V=1, what is the probability that the first two players choose (adopt, adopt)? That they choose (reject, reject)? That the first two moves are different?

(c) Given taht the first two moves were (adopt, adopt), what does player 3 choose if her signal is $X_3 = A$? If her signal is $X_3 = R$? Explicitly compute the posterior likelihood of (adopt, adopt, $X_3 = A$) given V=1 and V=0. Given that the first two moves were (reject, reject), what does player 3 choose if her signal is $X_3 = A$? If her signal is $X_3 = R$? How likely is inefficient herding after the first two moves as a function of p? Give the intuition.

(d) What is the probability that no herd has formed after 4 players have moved? After n players have moved (n even)?

Exercise 2:

Suppose that a new type of medecine has been introduced. Assume the following:

Medicines can either be low effectiveness or high effectiveness. Low effectiveness medicines work in a given patient with probability α, and high effectiveness medicines work in a given patient with probability β. The probability of getting well as a function of the effectiveness of the medicine is given in the following table, where W=1 if a given medicine works on a specific patient and W=0 otherwise; H=1 is if the medicine is highly effective, and H=0 is when it is not.

	$\Pr(W=1\|H)$	$\Pr(W=0\|H)$
$H=1$	β	$1-\beta$
$H=0$	α	$1-\alpha$

Suppose that the population starts out with a prior that there is a chance p that newly introduced medicines are highly effective and a chance 1-p that newly introduced medicine are of low effectiveness (i.e. Pr(H=1)=p, Pr(H=0)=1-p).

(a) Suppose that the particular medicine is in fact highly effective (H=1). Suppose that someone meets a friend who had a good experience with the medicine. What will be his posterior belief about the medicine after being told about his or her friend's experience?

(b) Suppose that p=0. What is the posterior belief?

(c) Describe the set of prior and posterior beliefs in a model of information flow that would be necessary to explain that individuals in the Kremer and Miguel study who had friends who were treated were less likely to take the medecine? Why would this effect be more negative for individuals with more education?

(d) One finding from the earlier study is that the treatment reduces disease transmission. Could we interpret the result discussed in (c) without referring to social learning?

(e) The authors also find that this effect is less negative for students in school 2 (who were taking the pill for the second year in a row). Can we explain this through a social learning model?

Exercise 3:

Consider a village where someone is afflicted with an illness X. X could be of two types: Non-Self-Limiting (NS) or Self-Limiting (SL). The villager does not observe the type of illness, but has a prior on the illness type, given by: Prob (X=SL)=θ.

If an illness is SL, it has a probability of 1 of being cured next period, whether the villager undergoes any treatment or not. The cure probability for a NS illness, however, depends on the treatment. The following table describes the probability of being cured:

	no treatment	*untrained doctor*	*trained doctor*
$X=SL$	1	1	1
$X=NS$	0	0.5	1

Utility of a villager is given by: ln(C)-H where C is consumption exclusing health care costs and H is positive if the village is sick and 0 otherwise. The village has an income Y. Trained doctors are more expensive than untrained ones and no treatment costs nothing.

Patients are myopic and maximize per period utility, so they do not experiment to gain information for the future. Patients also do not communicate with each other.

After seeking treatment and observing whether the illness gets better, the villager updates beliefs on whether the illness is self-limiting. The illness may recur in later periods because the patient may be reinfected. After n periods in which the patient has been struck with the illness, denote a patient's belief is Pr(X=SL)=θ_n. The initial belief is θ_0.

(1) First, qualitatively characterize how the myopic treatment decisions of patients in the first period they have the illness will vary with income.

(2) Following the period in which they have the illness, what will be the beliefs about the illness of individuals who choose ϕ, depending on the actual nature of the illness.

(3) What will be the beliefs about the illness of individuals who choose the untrained doctor, depending on the actual nature of the illness?

(4) What will be the beliefs about the illness of individuals who choose the trained doctor, depending on the nature of the illness?

(5) Suppose the disease is self-limiting. Consider someone who is struck twice and sees practitioner the trained doctor both times. What is θ_2?

(6) Characterize the use of no treatment and each type of doctors by different income groups in the long-run after many periods in which the illness has struck. Remember the person is afflicted with the same illness in every round.

So far we have assumed that the cure probabilities for the doctor (untrained or trained) are known. Next, assume that the cure probabilities are unknown, but the cumulative distribution for the trained doctor first order stochastically dominates that for the untrained practitioner. In particular, assume that the Pr(untrained doctor cures a NS illness with probability 0.5)=0.5 and Pr(untrained doctor cures a NS illness with probability 1)=0.5 but Prob(trained doctor cures a NS illness with probability 0.5)=0.2 and the Pr(trained doctor cures a NS illness with probability 1)=0.8. Further, assume that $\theta = 1/4$, so that agents place a low probability on the illness being self-limited.

(7) Qualitatively discuss which treatment choice myopic patients will make in the first round, as a function of their income.

(8) Suppose a patient sees an untrained doctor and gets better. What is their updated probability that the illness is self-limiting?

(9) What is their updated probability that they have found an untrained practitioner who cures this illness with probability 1?

(10) What treatment, if any, will they choose the next time they experience the same symptoms?

(11) Could people continue to see an untrained doctor in the long-run after many periods in which the illness has struck?