Helpful Laymen in Informational Cascades

Jiemai Wu*

June 5, 2015

Abstract

This paper extends Bikhchandani, Hirshleifer and Welch's informational cascade model by introducing two types of players, experts with high signal accuracy and laymen with low signal accuracy. Assuming players randomize when indifferent, if a small enough group of laymen are present in the population, the probability of having a correct cascade is strictly higher than if no laymen are present. The same result is robust to almost every tie-breaking rule.

Keywords: Informational cascades, Bayesian learning, Heterogeneous signal accuracy

JEL classification: C72, D83

1 Introduction

It is conventional wisdom that smarter, more experienced people make better decisions. However, even a group of experts can cluster on the wrong choice. As shown by Bikhchandani, Hirshleifer, Welch (1992), and Banerjee (1992), when an agent observes both a private signal and the sequence of actions by previous agents, he can decide it is optimal to follow the choice of previous agents, even if his private signal indicates the opposite. This phenomenon, in which agents ignore their

*Department of Economics, Washington University in St. Louis, One Brookings Drive, Campus Box 1208, St. Louis, MO 63130, USA. Email: jiemaiwu@wustl.edu

private information, is called an "informational cascade." In particular, if agents early in the sequence picked the wrong choice because of faulty private signals, the entire sequence of agents will follow the wrong choice, making it a "wrong informational cascade." Wrong informational cascades have been successfully replicated in experiments (Anderson and Holt 1997, Celen and Kariv 2004).

If even a group of experts aren't exempt from wrong cascades, how much worse will people do if the population also includes laymen, who have poorer signals than the experts? The answer is: they may end up doing better, with a lower frequency of wrong cascades.

Specifically, this paper analyzes games in a Bikhchandani-Hirshleifer-Welch setting with two types of players that differ in private signal accuracy. Suppose a group of consumers face the same choice, for example, whether to buy an iPhone or an Android phone. Assume one platform works better than the other, but it's hard to tell which is better with absolute certainty. Some consumers, whom I call experts, receive accurate private signals that indicate the correct choice most of the time. Others, whom I call laymen, are less familiar with the smartphone industry, and have poorer private signals that are incorrect more often. This paper shows that one can always find a small enough fraction of laymen, such that the probability of wrong cascades strictly decreases if those laymen are present among a group of experts. The reason comes from the discreteness of individuals: a player rationally follows suit and starts a cascade only after observing a fixed number of people making identical choices in a row. Let's call this number the "cascade triggering number". While adding a small fraction of laymen only decreases the overall information quality by a little, the cascade triggering number can discretely jump up by 1. That is, a player requires one more count of evidence before rationally ignoring his private signal in decision-making, and this increased hesitation delays the start of a cascade. As a result, the delay allows more information to be revealed to the public, which enables later players to make better choices.

This paper is closely related to Bikhchandani, Hirshleifer and Welch's paper on informational cascades (1992). Their paper includes a scenario in which each player's (potentially different) signal accuracy is public information, and discusses how sensitive cascades are with respect to the order of players, e.g. whether an expert decides first. My paper instead assumes anonymity of heterogeneous players, and shows how cascades are sensitive to a change in the entire player distribution, e.g. how many experts there are in the population. Other papers that address the topic of signal accuracy in cascades includes Sasaki's experimental study (Sasaki 2005), which shows that if the order of the players is linked with their ranking of signal accuracy, then there is a higher frequency of cascades when experts choose first. A paper by Pastine I. and Pastine T. (2006) provides examples to show that when homogeneous players' conditional signal probabilities are asymmetric for good and bad states, it is possible that the probability of correct cascades is not monotonic in players' common signal accuracy. My paper, on the other hand, studies heterogeneous players with symmetric conditional signal probabilities, and I show that the non-monotonicity of Pr(correct cascade) with respect to the fraction of experts not only exists, but also persists for any parameter specification.

Other studies have extended the cascade literature in different dimensions of heterogeneity. Smith and Sorensen (2000) discussed the possibility of confounded learning when players have opposite preferences with respect to the true state of the world. Goeree, Palfrey, and Rogers (2006) assume that a player's payoff is partly determined by a private preference shock that is independent of the true state. They conclude that as long as the support of such shock is rich enough, players will always learn and converge to the true state asymptotically. Other papers explored the possibility of players being exposed only to a (potentially different) subset of past action history (Banerjee and Fudenberg 2004, Acemoglu, Daleh, Lobel, and Ozdaglar 2011). All these papers assume a homogeneous signal distribution.

My paper is organized as follows: Section 2 lays out the the model. Section 3 summarizes the learning dynamics of the game. Section 4 derives the necessary and sufficient conditions of a cascade. Section 5 introduces the main theorem which states that it's always possible to have a higher probability of correct cascades by having some laymen among experts. A discussion of the robustness of the results is included at the end.

2 Model set up

There are two states of nature, $V \in \{v^H, v^L\}$, with equal prior probabilities $P(v^H) = P(v^L) = \frac{1}{2}$. An infinite sequence of i.i.d. players, $i = 1, 2, 3, ...$, enter with an exogenous order. Players differ in types $t \in \{expert, layman\}$, and $\pi \in [0, 1]$ is the probability that player i is an expert $\forall i$.

Each player i receives a private signal $S_i \in \{H, L\}$ conditional on the true state V, and player i's type t_i. The conditional probabilities are summarized in Table 1. An *expert* receives a more accurate private signal than a *layman*. Type t_i and signal S_i are both private information.

"Experts" (fraction π) have high signal precision:

Expert	$P(S_i = H\|V)$	$P(S_i = L\|V)$
V = v^H	p_E	$1 - p_E$
V = v^L	$1 - p_E$	p_E

"Laymen" (fraction $1 - \pi$) have low signal precision:

Layman	$P(S_i = H\|V)$	$P(S_i = L\|V)$
V = v^H	p_L	$1 - p_L$
V = v^L	$1 - p_L$	p_L

$\frac{1}{2} \leqslant p_L < p_E < 1$

Table 1: Conditional private signal distribution

A player faces two choices: *adopt* or *reject*. The payoff of rejection is 0. The payoff of adoption is 1 when $V = v^H$, and -1 when $V = v^L$. In other words, a player wishes to adopt when $V = v^H$, and reject when $V = v^L$.

Once a choice is made, it becomes public information for later players. Therefore, each rational player i observes the past action history $\{A_1, A_2, A_3 ... A_{i-1}\}$, his private type t_i, a private signal S_i, the fraction of experts π, and then chooses to adopt or reject.

Finally, player i's strategy when he's indifferent requires extra specification. Here I focus on the cases in which the player randomly chooses adoption with a fixed probability when indifferent. Under this tie-breaking rule, there is a unique perfect Bayesian equilibrium in which player i will

$$\textit{Adopt} \text{ if } P(v^H | A_1, ..., A_{i-1}, S_i, t_i) > P(v^L | A_1, ..., A_{i-1}, S_i, t_i);$$

$$\textit{Reject} \text{ if } P(v^H|A_1,...,A_{i-1},S_i,t_i) < P(v^L|A_1,...,A_{i-1},S_i,t_i);$$

$$\textit{Adopt} \text{ with probability } z \in [0,1] \text{ if } P(v^H|A_1,...,A_{i-1},S_i,t_i) = P(v^L|A_1,...,A_{i-1},S_i,t_i).$$

For example, if $z = 0.5$, the player flips a coin when indifferent. The main result of this paper is robust to almost every arbitrary tie-breaking rule. Section 5 and 6 include the detailed discussion.

As in the previous literature, an *informational cascade for type t* is said to occur when it is optimal for a type-t player to follow the choice of the preceding player regardless of his private signal. A *full informational cascade* occurs when it is optimal for a player to follow the choice of the preceding player regardless of his private signal *and* type.

It's intuitive that a cascade for experts occurs later than a cascade for laymen. Compared with laymen, experts are more confident in their private signals, so stronger evidence is needed to convince an expert to ignore his own signal and follow suit. Hence, a full cascade begins exactly when an expert enters a cascade.

3 Learning dynamics

To describe the dynamics of the game, let $\{l_n\}_{n=0}^{\infty}$ be a sequence of the *public likelihood ratio* where $l_0 \equiv 1$ and

$$l_n \equiv \frac{P(A_1,\ A_2,\ ...,\ A_n|v^H)}{P(A_1,\ A_2,\ ...,\ A_n|v^L)} \text{ for } n = 1,2,3,...$$

Hence the decision rule for player i can be rewritten as:

$$\textit{Adopt} \text{ if } l_{i-1} \cdot \frac{P(S_i|v^H,t_i)}{P(S_i|v^L,t_i)} > 1;$$

$$\textit{Reject} \text{ if } l_{i-1} \cdot \frac{P(S_i|v^H,t_i)}{P(S_i|v^L,t_i)} < 1;$$

$$Adopt \text{ with probability } z \in [0,1] \text{ if } l_{i-1} \cdot \frac{P(S_i|v^H, t_i)}{P(S_i|v^L, t_i)} = 1.$$

Conditional on the true state being 1, $\{l_n\}$ is a Markov chain:

Let

$$L_1 = \frac{1-p_E}{p_E},\ L_2 = \frac{1-p_L}{p_L},\ L_3 = \frac{p_L}{1-p_L},\ L_4 = \frac{p_E}{1-p_E}.$$

Since $\frac{1}{2} \leqslant p_L < p_E < 1$, $L_1 < L_2 < 1 < L_3 < L_4$. I use these four numbers as cutoff values to describe the evolution of l_n.

When $l_n < L_1$: No private signal can outweigh the strong public belief in favor of v^L. In this case a wrong full cascade of rejection occurs, and $l_{n+1} = l_n$ with probability 1.

When $l_n = L_1$: The current player is indifferent (and thus chooses adoption with probability z) only if he is an expert with signal H. Otherwise, he chooses rejection regardless of his signal. Therefore,

a. the current player *adopts* with probability $z\pi p_E$, and $l_{n+1} = l_n \cdot \dfrac{p_E}{1-p_E}$;

b. the current player *rejects* with probability $1 - z\pi p_E$, and $l_{n+1} = l_n \cdot \dfrac{1-z\pi p_E}{1-z\pi(1-p_E)}$.

When $l_n \in (L_1, L_2)$: The current player adopts only if he is an expert with signal H. Otherwise, he chooses rejection regardless of his signal. Therefore,

a. the current player *adopts* with probability πp_E, and $l_{n+1} = l_n \cdot \dfrac{p_E}{1-p_E}$;

b. the current player *rejects* with probability $1 - \pi p_E$, and $l_{n+1} = l_n \cdot \dfrac{1-\pi p_E}{1-\pi(1-p_E)}$.

The transition of l_n when it falls in $\{L_2\}$, (L_2, L_3), $\{L_3\}$, or (L_3, L_4) can be deduced in a similar fashion. For each interval, pin down the type-signal combinations that lead to an adoption (respectively, rejection). Conditional on the action of the current player $A_i \in \{adopt, reject\}$, derive

$$l_{n+1} = l_n \cdot \frac{P(\text{type-signal combo that choose } A_i \,|\, v^H)}{P(\text{type-signal combo that choose } A_i \,|\, v^L)},$$

and the transitional probabilities accordingly. Finally, finish with the last possible scenario:

When $l_n > L_4$: No private signal can outweigh the strong public belief in favor of v^H. In this case a correct full cascade of adoption occurs, and $l_{n+1} = l_n$ with probability 1.

4 Conditions for a full cascade

The transition of the public likelihood ratio l_n describes how the game evolves; however, it requires much calculation to tell if a cascade has started in an arbitrary game. There's a quicker way to spot the rise of a cascade. Let's start with the following Lemma.

Lemma. $\forall \pi \in (0,1)$, *if a sequence of consecutive identical actions are observed from the beginning of the game, a cascade for laymen starts after exactly* 1 *player, and a cascade for experts (hence, a full cascade) starts after exactly N players, where N is the smallest integer larger than*

$$\frac{\ln\left[\frac{p_E}{1-p_E} \cdot \frac{\pi(1-p_E)+(1-\pi)(1-p_L)}{\pi p_E+(1-\pi)p_L}\right]}{\ln \frac{1-\pi(1-p_E)}{1-\pi p_E}} + 1.$$

Note that N is a decreasing function of π: the more experts there are in the population, the less evidence is needed to convince an expert to follow suit. For $\pi = 1$, $N = 2$. As $\pi \to 0$, $N \to \infty$ $\forall p_L, p_E$. Figure 1 in Section 5 also shows a plot of N for $p_L = 0.55$ and $p_E = 0.95$.

Proof. I here prove the lemma for cascades of adoption. The proof for cascades of rejection is symmetric. Let M be the smallest integer such that if player $M+1$ is a layman, and if all players $1, \ldots, M$ choose adoption, player $M+1$ also chooses adoption regardless of his private signal (i.e. he is in a cascade). Similarly, let N be the smallest integer such that if player $N+1$ is an expert, and if all players $1, \ldots, N$ choose adoption, player $N+1$ also chooses adoption regardless of his private signal. Then, M is simply the smallest integer s.t. $l_M > L_3$, and N is the smallest integer s.t. $l_N > L_4$, where $L_3 = \frac{p_L}{1-p_L}$, $L_4 = \frac{p_E}{1-p_E}$, as defined in Section 3. Since

$$l_M = \left[\frac{\pi p_E + (1-\pi)p_L}{\pi(1-p_E)+(1-\pi)(1-p_L)}\right]^M,$$

$l_M > L_3$ implies

$$M > \frac{\ln \frac{p_L}{1-p_L}}{\ln \frac{\pi p_E+(1-\pi)p_L}{\pi(1-p_E)+(1-\pi)(1-p_L)}} \equiv M^*.$$

Note that $M^* < 1\ \forall p_E, p_L$ s.t. $\frac{1}{2} \leqslant p_L < p_E < 1$. Therefore $M = 1$.

Similarly, since

$$l_N = \left[\frac{\pi p_E + (1-\pi)p_L}{\pi(1-p_E)+(1-\pi)(1-p_L)}\right]\left[\frac{1-\pi(1-p_E)}{1-\pi p_E}\right]^{N-1},$$

$l_N > L_4$ implies

$$N > \frac{\ln[\frac{p_E}{1-p_E}\cdot\frac{\pi(1-p_E)+(1-\pi)(1-p_L)}{\pi p_E+(1-\pi)p_L}]}{\ln \frac{1-\pi(1-p_E)}{1-\pi p_E}} + 1.$$

This completes the proof of the lemma. □

The above lemma identifies the start of a cascade when there is a sequence of identical actions from the very beginning of the game. The next proposition identifies the start of a cascade in a general case.

Proposition. *(Necessary and sufficient conditions for a full cascade) Following any history, if no cascade has yet started,*

1. at least 1, and at most 2 consecutive identical actions are needed to trigger a cascade for laymen;

2. at least N, and at most $N+1$ consecutive identical actions are needed to trigger a cascade for experts (and therefore, a full cascade).

Hence, a full cascade occurs with probability 1 as the number of players goes to ∞.

For example, let "A" denote "adopt" and let "R" denote "reject". If $N = 4$, then the proposition implies that no full cascade has started after action history "RRAARAAA". On the other hand, if the action history is "RRRAAAAA", a full cascade of adoption must be in action. In other words, "N consecutive identical actions" is the necessary condition for a full cascade; "$N+1$ consecutive identical actions" is the sufficient condition.

Proof. I here prove the proposition for cascades of adoption. The proof for cascades of rejection is symmetric.

If a sequence of adoption starts from the beginning of the game, see Lemma. Otherwise, suppose a sequence of adoption starts after some history that ends with a rejection:

... (some history)..., R, A, A, A, A, A...

Denote the player associated with the rejection in the above example as player k.

Lemma implies that when $l = l_0 = 1$ (as at the beginning of the game), exactly N adoptions are needed to trigger a cascade for experts (and therefore, a full cascade).

The fact that player k chooses to reject implies that after receiving his private signal, player k's private posterior probability for v^L is at least 0.5. The public learning process captures this information by having $l_{k+1} \leqslant 1 \; \forall l_k$ whenever player k rejects. Therefore, since exactly N adoptions are needed to start a cascade when $l = 1$, in this less favorable scenario with $l_{k+1} \leqslant 1$, *at least* N adoptions are needed to trigger a cascade for experts after player k .

Next let's focus on player $k+1$, the first player who chooses adoption in the sequence. Similarly, $l_{k+2} \geqslant 1 \; \forall l_{k+1}$ whenever player $k+1$ adopts, so we need *at most* N more adoptions (which means $N+1$ adoptions in total) to trigger a cascade for experts.

The proof of cascades for laymen is obtained by simply replacing N with 1 in the argument above.

Finally, since N is finite, the probability of having a sequence of identical actions with the required length converges to 1, so a full cascade occurs with probability 1 in the limit. To formally show this, let M denote the number of players. Note that if starting from somewhere in the sequence, there are $N+1$ consecutive experts all receiving signal H, they will all choose adoption and this starts a cascade. Therefore,

$$\begin{aligned} P(cascade) &> P(\text{there exists } N+1 \text{ consecutive experts with signal } H) \\ &\geqslant 1-\left[1-(\pi p_E)^{N+1}\right]^{\frac{M}{N+1}} \\ &\to 1 \ \text{ as } M \to \infty. \end{aligned}$$

□

5 Probability of correct cascades

5.1 Example

If a few consecutive players happen to receive wrong signals, a cascade starts where everyone later in the sequence chooses the wrong option even when they receive correct signals. Such unfortunate events always occur with a positive probability. From a welfare point of view, therefore, it is meaningful to study the probability of landing on a correct cascade, and in particular, how this probability changes with the demographics of the population.

Suppose $p_E = 0.95$, $p_L = 0.55$. That is, experts receive correct signals 95% of the time, and laymen only receive correct signals 55% of the time. Also assume $z = 0.5$, which means players flip a coin when indifferent. Plot A in Figure 1 describes how N, the length of the sequence of identical actions needed to trigger a cascade, changes with π. Plot B, which is the result of 100,000 Monte Carlo simulation trials, shows the frequency of correct cascades for each π.

Observe that in plot A, N decreases with π. The more experts there are, the less it takes to trigger a full cascade. Let $f(\pi) \equiv N$ evaluated at π. Since N only takes integer values, f is a

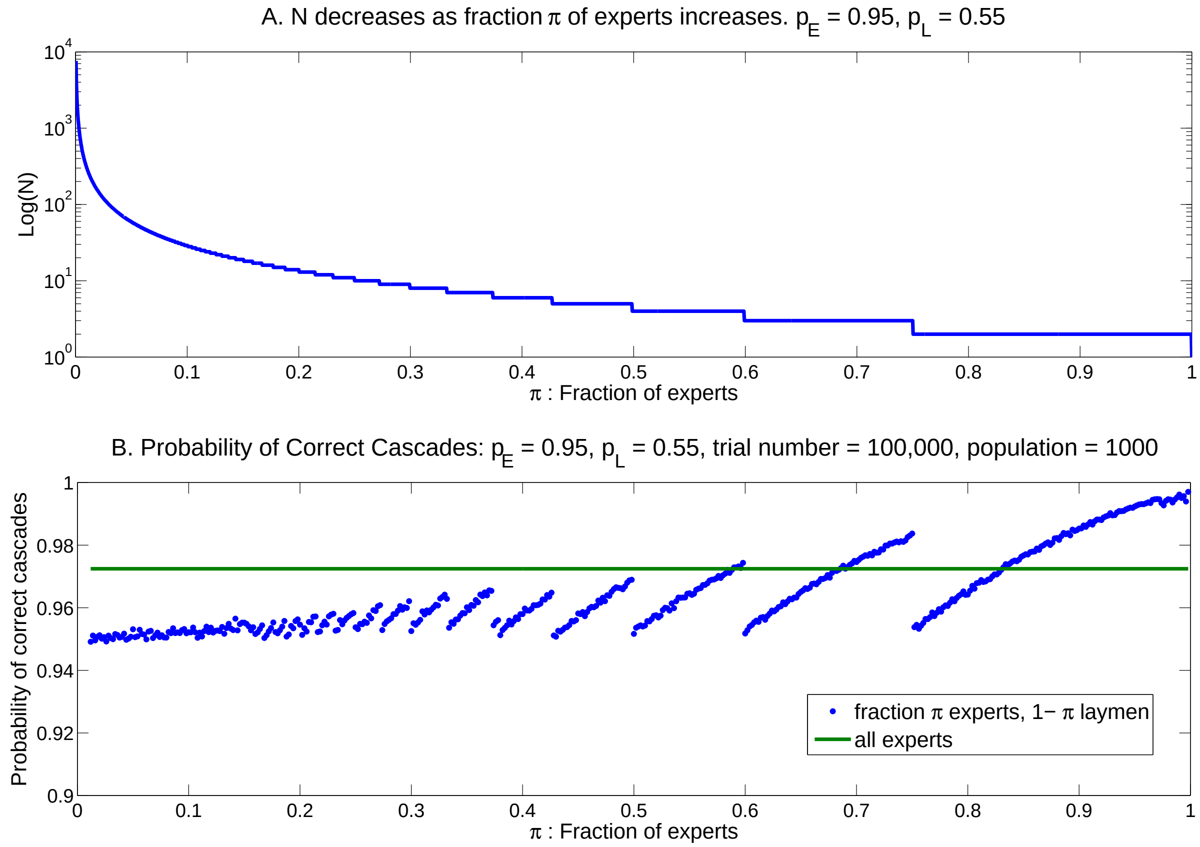

Figure 1: Probability of correct cascades exhibits discontinuous drop when integer N decreases

discontinuous function of π, and so I define a set of "turning points" $\{\pi_N\}_{N=2}^{\infty}$ s.t. $f(\pi_N) = N$ and $\lim_{\varepsilon\to 0} f(\pi_N - \varepsilon) = N+1$. In the example, $\pi_2 = 1$, $\pi_3 \approx 0.75$, $\pi_4 \approx 0.6$, $\pi_5 \approx 0.5$.

In plot B, observe that the probability of correct cascades $p_{correct}(\pi)$ (blue dots in the picture) increases in π until it drops when π reaches π_N for some N. A particularly interesting fact is that, for a range of $\pi \in (0.59, 0.6) \cup (0.69, 0.75) \cup (0.83, 1)$, $p_{correct}(\pi) > p_{correct}(1)$ (blue dots above the green line). When about 40% of the players are laymen, the probability of landing on the correct cascade is even higher than the case in which all players are experts.

What explains the drops in $p_{correct}$? And why is $p_{correct}$ higher when laymen are present?

To answer the first, note that the likelihood function of each sequence of type-signal realization (i.e. player 1 is an expert with H, player 2 is a layman with L, etc.) is continuous at each π_N, and therefore its left-sided limit at π_N is equal to its value at π_N. However, $f(\pi_N) = \lim_{\varepsilon\to 0} f(\pi_N - \varepsilon) - 1$, which implies that although the overall population composition and signal quality at π_N and $\pi_N - \varepsilon$ are almost the same, players wait one less period to start a cascade at π_N. With an earlier start of the cascade, all later players' decisions are based on a smaller set of information, which, inevitably, lead to more mistakes and a lower $p_{correct}$.

The reason for a higher $p_{correct}$ when laymen are present follows the same logic. When there are laymen around, experts wait longer before following suit, and everyone benefits from this little hesitation. To be more specific, when all players are experts, suppose the first player adopts. Then even if the second player receives signal L, he's indifferent between adoption and rejection, and will therefore choose randomly. In contrast, when there is a small group of laymen, the expert second player with signal L no longer trusts the first player as much as himself, and instead strictly follows his own signal to reject. This added bit of conservativeness sends out a clearer message to later players; they know the second player adopts if and only if the signal is H. Because later players now have better information to work with, they end up making the correct choice more often. The theorem in the next section generalizes this idea.

5.2 Theorem on probability of correct cascades

In the example, when 40% of the players are laymen, the probability of correct cascades is higher than the case in which no players are laymen. The following theorem generalizes this result by showing that for any values of p_E and p_L, there always exists a small enough fraction, such that if this fraction of laymen are present, the probability of correct cascades is higher than if no laymen are present.

Theorem. *Let $p_{correct}(\pi)$ be the probability of correct cascades when fraction π of the population are experts. Then $\forall p_E, p_L$ with $\frac{1}{2} \leqslant p_L < p_E < 1$, $\exists\, \underline{\pi} \in (0,1)$ s.t. $\forall \pi \in (\underline{\pi}, 1)$,*

$$p_{correct}(\pi) > p_{correct}(1).$$

Proof. It suffices to prove that $\lim\limits_{\pi \to 1} p_{correct}(\pi) > p_{correct}(1)$.

Let $l_\pi(A_1,A_2)$ denote the public likelihood ratio after two actions $A_1,A_2 \in \{\text{A, R}\}$ when fraction π are experts. Similarly define $l_{expert}(A_1,A_2)$ when $\pi = 1$: all players are experts.

$\lim_{\pi\to 1} l_\pi(\text{A, A}) = \left(\frac{p_E}{1-p_E}\right)^2 > l_{expert}(\text{A, A}) = \frac{p_E\,[p_E + z\,(1-p_E)]}{(1-p_E)\,(1-p_E+zp_E)} > L_4 \Rightarrow$ cascade of adoption

$\lim_{\pi\to 1} l_\pi(\text{R, R}) = \left(\frac{1-p_E}{p_E}\right)^2 < l_{expert}(\text{R, R}) = \frac{(1-p_E)\,[1-p_E+(1-z)\,p_E]}{p_E\,[p_E+(1-z)(1-p_E)]} < L_1 \Rightarrow$ cascade of rejection

$\lim_{\pi\to 1} l_\pi(\text{A, R}) = \lim_{\pi\to 1} l_\pi(\text{R, A}) = l_{expert}(\text{A, R}) = l_{expert}(\text{R, A}) = 1 \Rightarrow$ back to the origin

For both scenarios, the only possible action history that can trigger a correct cascade is a pair of correct actions following several pairs of opposite actions (A, R) or (R, A). Therefore,

$$
\begin{aligned}
p_{correct}(1) &= \frac{1}{2}\sum_{k=0}^{\infty}[\text{P(opposite action pair} \mid v^H)]^k \cdot \text{P(A, A} \mid v^H) \\
&\quad + \frac{1}{2}\sum_{k=0}^{\infty}[\text{P(opposite action pair} \mid v^L)]^k \cdot \text{P(R, R} \mid v^L) \\
&= \frac{1}{2}\sum_{k=0}^{\infty}[p_E(1-p_E)]^k \cdot p_E \cdot [p_E + z\cdot(1-p_E)] \\
&\quad + \frac{1}{2}\sum_{k=0}^{\infty}[p_E(1-p_E)]^k \cdot p_E \cdot [p_E + (1-z)\cdot(1-p_E)] \\
&= \frac{1}{2}\sum_{k=0}^{\infty}[p_E(1-p_E)]^k \cdot p_E \cdot (p_E+1) \\
&= \frac{p_E(p_E+1)}{2(1-p_E+p_E^2)}.
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
\lim_{\pi\to 1} p_{correct}(\pi) &= \frac{1}{2}\sum_{k=0}^{\infty}\lim_{\pi\to 1}[\text{P(opposite pair | } v^H)]^k \cdot \text{P(A, A | } v^H) \\
&\quad + \frac{1}{2}\sum_{k=0}^{\infty}\lim_{\pi\to 1}[\text{P(opposite pair | } v^L)]^k \cdot \text{P(R, R | } v^L) \\
&= \frac{1}{2}\sum_{k=0}^{\infty}\lim_{\pi\to 1}[\text{P("H, L" or "L, H" | } v^H)]^k \cdot \text{P(H, H | } v^H) \\
&\quad + \frac{1}{2}\sum_{k=0}^{\infty}\lim_{\pi\to 1}[\text{P("H, L" or "L, H" | } v^L)]^k \cdot \text{P(L, L | } v^L) \\
&= \sum_{k=0}^{\infty}[2p_E(1-p_E)]^k \cdot p_E^2 \\
&= \frac{p_E^2}{1-2p_E+2p_E^2}.
\end{aligned}
$$

When $p_E \in (\frac{1}{2}, 1)$,

$$\frac{p_E(p_E+1)}{2(1-p_E+p_E^2)} < \frac{p_E^2}{1-2p_E+2p_E^2},$$

therefore $\lim_{\pi\to 1} p_{correct}(\pi) > p_{correct}(1)$.

□

From a slightly different point of view, as $\varepsilon \to 0$, at each $\pi = \pi_N - \varepsilon$ the game resembles one in which all players choose according to their private signals when they are indifferent, i.e. choosing adoption if and only if the private signal is H when the posterior probabilities are equal. Let's call such player a "*non-conformist*" for future reference. This tie-breaking rule helps to reveal more signals, and leads to a higher $p_{correct}$. At each $\pi = \pi_N + \varepsilon$, the game resembles one where all players copy the previous player when indifferent, i.e. when the posterior probabilities are equal, a player chooses adoption if and only if the previous player chooses adoption. Fewer signals are revealed in this case, resulting in a lower $p_{correct}$. When $\pi = \pi_N$, players randomize when indifferent, resulting in a middle case with a medium $p_{correct}$.

In particular, note that the derivation of $\lim_{\pi\to 1} p_{correct}(\pi)$ is independent of the player's strategy when indifferent. Moreover, $\lim_{\pi\to 1} p_{correct}(\pi) = p_{correct}^{non-conformist}(1)$, the probability of correct cascades when all players are experts *and* non-conformists, and this is true for all tie-breaking strategies. Therefore, $\lim_{\pi \to 1} p_{correct}(\pi) > p_{correct}(1)$ because $p_{correct}^{non-conformist}(1) > p_{correct}(1)$, i.e. games with non-conforming experts are more likely to have correct cascades than games with experts who randomize when indifferent. In this sense, the randomization tie-breaking rule is "inferior" to the non-conformist tie-breaking rule. The following corollary shows that the result of the last theorem holds for all ties-breaking rules that are different from the non-conformist rule.

Corollary. *For any tie-breaking strategy $\tau \neq \tau^{non-conformist}$, let $p_{correct}^{\tau}(\pi)$ be the probability of correct cascades when fraction π of the population are experts, then*

$\forall p_E, p_L$ *with* $\frac{1}{2} \leqslant p_L < p_E < 1$, $\exists\, \underline{\pi} \in (0,1)$ *s.t.* $\forall \pi \in (\underline{\pi}, 1)$,

$$p_{correct}^{\tau}(\pi) > p_{correct}^{\tau}(1).$$

Proof. $\forall \tau \neq \tau^{non-conformist}$, $\lim_{\pi \to 1} p_{correct}^{\tau}(\pi) = p_{correct}^{non-conformist}(1) > p_{correct}^{\tau}(1)$, and the result follows.

See Appendix for the proof of $p_{correct}^{non-conformist}(1) > p_{correct}^{\tau}(1)$.

□

6 Conclusion

This paper extends the Bikhchandani-Hirshleifer-Welch informational cascades model by incorporating heterogeneity in private signal accuracy. I conclude that in the unique perfect Bayesian equilibrium associated with a tie-breaking rule under which players randomize when indifferent, the probability of correct cascades is higher when a small enough group of laymen are present among a population of experts. The corollary in Section 5 shows that this result is robust to all tie-breaking rules except what I have called the "non-conformist" rule (i.e. choose according to own signal when indifferent). But even in cases where the non-conformist tie-breaking rule is adopted, as long as players are discrete, the intuition in section 5.1 carries on, and the probability of correct cascades is still non-monotonic with discontinuous jumps resembling those in Figure 1. Therefore, for all tie-breaking rules including the non-conformist rule, there exist cutoff expert fractions from which adding a small fraction of laymen makes correct cascades more frequent.

Although much of the discussion involves cases in which players are indifferent, quantitatively they make a big difference. As seen in the example in section 5.1, a population with 40% laymen and 60% experts land on the correct cascade more frequently than a population of 100% experts, even though the signal accuracy differs dramatically for the two player types (0.55 vs. 0.95). From this perspective, it confirms how easily an informational advantage can be outweighed by insufficient learning when people follow suit.

Acknowledgments

I am grateful to John Nachbar and Paulo Natenzon for their helpful guidance and encouragement. I also thank Brian Rogers, David Levine, Maher Said, Xiaoyuan Huang, as well as participants in Econometric Society Australasian Meeting 2013 and Washington University theory bag lunch for their invaluable comments.

Appendix

Nonconformist tie-breaking rule yields the highest probability of correct cascade

Consider homogeneous players with signal accuracy $p = P(H \mid v^H) = P(L \mid v^L)$ who adopt an arbitrary tie-breaking rule. I here prove that the probability of having a correct cascade is the highest if all players adopt a nonconformist tie-breaking rule.

Definition. A tie-breaking rule is defined by $\tau \equiv \{l_h\}_{h \in \mathscr{H}}$, where $\mathscr{H}$ is the set of all possible history such that the next player following such history can be indifferent. l_h denotes the probability that if the player is indeed indifferent after history h, he picks an action in accordance to his private signal.

For example, if the first player chose A, the second player will be indifferent if he receives signal L. In this case l_A denotes the probability that the second player chooses R. Similarly, l_{ARR} denotes the probability that the 4th player with signal H chooses A when the action history $h = ARR$. For a counter example, $AA \notin \mathscr{H}$ because a cascade of adoption starts after AA, and the next player will not be indifferent regardless of his signal.

$\{l_h\}$ fully characterizes a tie-breaking rule because a player's strategy only depends on the previous action history and his private signal. Moreover, if $h \in \mathscr{H}$, only one of the two signals induces indifference, so it's sufficient to define the probabilities as functions of action history only.

Definition. Let $\tau^{nonconf}$ denote the nonconformist tie-breaking rule: always follow own signal when indifferent. I.e., $l_h = 1$ for all h.

Proposition. *Let $P^{correct}(\tau)$ denote the unconditional probability of correct cascades when the tie-breaking rule is τ. Then for all $\tau \neq \tau^{nonconf}$,*

$$P^{correct}(\tau) < P^{correct}(\tau^{nonconf}).$$

Proof. It suffices to only consider tie-breaking rules with $l_h > 0\, \forall h$. If after history h the next player is indifferent and he always copies the last player's action ($l_h = 0$), then his action conveys no information to the later players, and his presence has no effect on $P^{correct}$. Therefore, simply delete such player from the sequence, and only focus on the games played by players who follow own signal with positive probability when indifferent.

I prove the proposition in 3 steps. First, I show that $P^{correct}$ strictly increases in l_A and l_R. I then show that $P^{correct}$ strictly increases in l_h for all $h \in \mathscr{H}$. Finally, conclude that $P^{correct}$ is maximized only when $l_h = 1$ for all h, which corresponds to the nonconformist tie-breaking rule.

Claim 1: $\dfrac{\partial P^{correct}(\tau)}{\partial l_A} > 0$ and $\dfrac{\partial P^{correct}(\tau)}{\partial l_R} > 0$.

Proof of Claim 1:

$$
\begin{aligned}
P^{correct}(\tau) &= \frac{1}{2}P(AA \text{ or } ARAA \text{ or } RAAA \text{ or } ARARAA\ldots|v^H) \\
&\quad + \frac{1}{2}P(RR \text{ or } ARRR \text{ or } RARR \text{ or } ARARRR\ldots|v^H) \\
&= \frac{1}{2}\{p\,[p+(1-p)l_A] + p(1-p)l_A p\,[p+(1-p)l_{ARA}] + \ldots\} \\
&\quad + \frac{1}{2}\{p\,[p+(1-p)l_R] + (1-p)pl_A p\,[p+(1-p)l_{ARR}] + \ldots\}
\end{aligned}
$$

where $p = P(H|V=1) = P(L|V=0)$. Therefore,

$$
\begin{aligned}
\frac{\partial P^{correct}(\tau)}{\partial l_A} &= \frac{1}{2}\left\{p\,[p+(1-p)] + (1-p)p^2\,[p+(1-p)l_{ARA}] + \ldots\right\} \\
&\quad + \frac{1}{2}\left\{(1-p)p^2\,[p+(1-p)l_{ARR}] + \ldots\right\} \\
&= \frac{1}{2}(2p-1)\{p(1-p) + [p(1-p)]^2\,(l_{ARA}+l_{ARR}) \\
&\quad + [p(1-p)]^3\,[l_{ARA}(l_{ARARA}+l_{ARARR}) + l_{ARR}(l_{ARRAA}+l_{ARRAR})] \\
&\quad + \ldots\} \\
&> 0
\end{aligned}
$$

since $p > \dfrac{1}{2}$.

A similar argument proves $\dfrac{\partial P^{correct}(\tau)}{\partial l_R} > 0$. //

Claim 2: $\dfrac{\partial P^{correct}(\tau)}{\partial l_h} > 0$ for all $h \in \mathscr{H}$.

Proof of Claim 2: Proof by induction. First,

$$
\begin{aligned}
sgn\left(\frac{\partial P^{correct}}{\partial l_{ARA}}\right) &= sgn\left[P(AR)\cdot\frac{\partial P^{correct}}{\partial l_A}\right]=1\\
sgn\left(\frac{\partial P^{correct}}{\partial l_{RAA}}\right) &= sgn\left[P(RA)\cdot\frac{\partial P^{correct}}{\partial l_A}\right]=1\\
sgn\left(\frac{\partial P^{correct}}{\partial l_{ARR}}\right) &= sgn\left[P(AR)\cdot\frac{\partial P^{correct}}{\partial l_R}\right]=1\\
sgn\left(\frac{\partial P^{correct}}{\partial l_{RAR}}\right) &= sgn\left[P(RA)\cdot\frac{\partial P^{correct}}{\partial l_R}\right]=1
\end{aligned}
$$

where $sgn(\cdot)=1$ means a positive value.

Note that any history $h\in\mathscr{H}$ can be written as $h=ARh'$ or $h=RAh'$ for some history $h'\in\mathscr{H}$. So if $sgn\left(\frac{\partial P^{correct}}{\partial l_{h'}}\right)=1$, then $sgn\left(\frac{\partial P^{correct}}{\partial l_h}\right)=sgn\left[P(AR)\cdot\frac{\partial P^{correct}}{\partial l_{h'}}\right]$ or $sgn\left(\frac{\partial P^{correct}}{\partial l_h}\right)=sgn\left[P(RA)\cdot\frac{\partial P^{correct}}{\partial l_{h'}}\right]$. In both cases the sign is positive.

Therefore, by induction, conclude that $\frac{\partial P^{correct}}{\partial l_h}>0$ for all $h\in\mathscr{H}$. //

Finally, Claim 2 implies that $P^{correct}(\tau)$ is maximized when $l_h=1$ or all $h\in\mathscr{H}$, which is only true for $\tau=\tau^{nonconf}$. □

References

[1] Acemoglu, D., Dahleh, M., Lobel, I., and Ozdaglar, A. [2011], “Bayesian Learning in Social Networks,” *Review of Economic Studies*, 78, 1201-1236.

[2] Anderson, L., and Holt, C. [1997], "Information Cascades in the Laboratory," *The American Economic Review*, 87, 847-62.

[3] Banerjee, A. [1992], “A Simple Model of Herd Behavior,” *The Quarterly Journal of Economics*, Vol. 107, No. 3, 797-817.

[4] Banerjee, A., and Fudenberg, D. [2004], “Word-of-Mouth Learning,” *Games an Economic Behavior*, 46, 1-22.

[5] Bikhchandani, S., Hirshleifer D., and Welch, I. [1992], “A Theory of Fads, Fashion, Custom, and Cultural Exchange as Information Cascades,” *Journal of Political Economy*, 100, 992-1026.

[6] Bikhchandani, S., Hirshleifer D., and Welch, I. [1998], “Association Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades,” *The Journal of Economic Perspectives*, Vol. 12, No. 3, 151-170.

[7] Celen, B., and Kariv, S. [2004], “Distinguishing Informational Cascades From Herd Behavior in the Laboratory,” *The American Economic Review*, Vol. 94, No. 3 (Jun., 2004), 484-498.

[8] Goeree, J., Palfrey, T., and Rogers, B. [2006], “Social Learning With Private and Common Values,” *Economic Theory*, Vol. 28, 245-264.

[9] Pastine, I., and Pastine T. [2006], “Signal Accuracy and Informational Cascades,” working paper.

[10] Sasaki, S. [2005], “Signal Qualities, Order of Decisions, and Informational Cascades: Experimental Evidence,” *Economics Bulletin*, Vol. 3, No. 34 pp. 1-11

[11] Smith, L., and Sorenson, P. [2000], "Pathological Outcomes of Observational Learning," *Econometrica*, Vol. 68, No. 2, 371-398.