The Impact of Experts and Error in Observation on Informational Cascades

Rohit Mittapalli

Abstract

Often models are used to study Bayesian agents where limited information is available. Models, in which these agents sequentially decide to accept or reject an option while observing previous agent's decisions, can produce informational cascades. An informational cascade is said to occur when an agent relinquishes their own private information in exchange for a pattern evident in the decisions of previous agents. The model in this paper has two types of agents each with differing signal accuracies: experts and non-experts. The model also incorporates an element of noise causing agents to flip their decision at a low probability. With this model, we studied the probability the crowd cascades correctly with respect to noise and expert concentration. The paper showed showed a non-monotonic relationship between both noise and expert concentration and the probability of a correct cascade suggesting that a lower noise and a higher expert concentration does not necessarily increase the probability of a correct cascade.

1 Introduction

When people are connected, they are capable of observing each other's decisions and behavior allowing people to influence each other's decision. This phenomenon has nearly limitless application: the services people use, fashion or fads, to buy or not are few of the endless scenarios[1]. This influence people have upon others can result in social conformity. Figure a scenario where agents line up to buy a product. Agents decide to buy or not by factoring in two components: their private knowledge and their observations of earlier agents. Bikhchandani, Hirshleifer, and Welch (BHW) uncovered the possibility that agents may find it optimal to follow the decision of others even though they may disagree [1]. This phenomenon, where agents relinquish their own private information in exchange for the decision of their predecessors, is coined an informational cascade. Not only do informational cascades cause a loss of information but it also creates the possibility of a wrong cascade.

In order to analyze the characteristics of informational cascades, researchers use computer models to simulate real world applications. Each of these model contain 3 major components: state of the world, signals, and payoffs. In the majority of models discussed, the state of the world is binary: either good or bad (each with probability $\frac{1}{2}$). Agents in the model try to decide by guessing the state of the model through their observations and their signal. Their signal tells them the state of the world correctly with a certain probability. An agent receives a payoff based on their decision. If an agent rejects, they receive a payoff of 0. If an agent accepts, the agent receives a payoff according to the state of the world, positive if good, negative if bad [2].

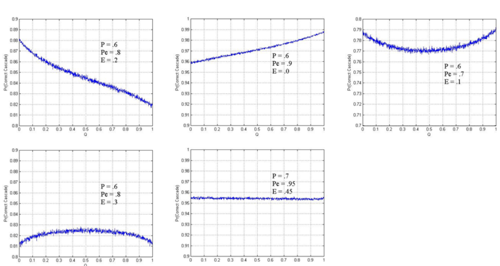
The BHW model was the first of its kind. The model studied the probability of an incorrect cascade occurring. BHW showed that with little information, cascades can occur quickly. BHW noting that cascades were founded on little to none information, began to study the fragility of cascades by introducing new information into the simulations [1]. The model essentially set the stage with a simple model providing room for many variations to follow.

After the BHW model in 1992, many variations stemmed out. Sasaki (2005) tested the effect of the order of the agents on the model. She organized them in two ways: seniority (highest to lowest signal accuracy) and anti-seniority (lowest to highest signal accuracy). She found that although seniority has more complete cascades, it is more likely to cause a wrong cascade [6]. Pastine and Pastine altered the BHW model such that the signal accuracies were no longer symmetric. A model is considered symmetric if the probability of a correct cascade is equal regardless of the state of the model. They showed that the slightest variation in a model's symmetry causes signal accuracy to have a non-monotonic effect on the probability of an correct cascade [5]. The model discussed in this paper is completely symmetrical.

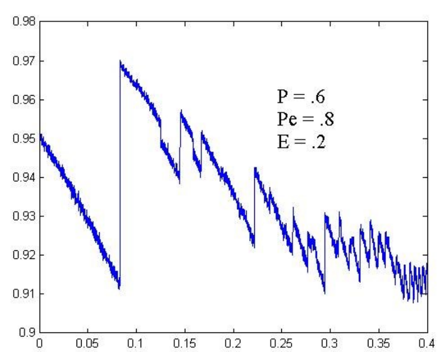
Wu [7] extended the BHW model to incorporate two types of agents: experts and laymen. Experts had higher-signal accuracy than laymen, but their type remained anonymous to the population. Her model shows that a mix between agents is strictly better than a homogeneous sample in all cases [7]. The model discussed in this paper also has two agent types as well: expert and non-experts. Similarly, experts have higher signal accuracy. The model differs from the Wu model because the identities of the experts are known and have a larger influence on an agent than a non-expert.

In all discussed models, agents perfectly observe the actions of those before them. Le, Subramanian, and Berry investigated the impact of errors in the observations of previous agents on total payoff. They concluded that a lower error level does not always mean a higher payoff [4]. The model discussed in this paper incorporates noise similarly to Le, Subramanian, and Berry.

This model studies the the probability of a correct cascade denoted as $Pr(CorrectCascade)$ in respect to Q (fraction of experts) and ε (error in observation). Graphs of $Pr(CorrectCascade)$ against Q exhibits a wide range of graphs some of which are non-monotonic, monotonic, U-Shaped, or almost flat (see Figure A). The graphs of $Pr(CorrectCascade)$ and ε is also non-monotonic characterized by "spikes" (see Figure B). By comparing the graphs to fluctuation in the model's parameters, some of these spikes can be explained.



(A) $Pr(CorrectCascade)$ in respect to Q



(B) $Pr(CorrectCascade)$ in respect to ε

2 Process

Initially the $Pr(CorrectCascade)$ was calculated through First-Step analysis discussed in detail later, but due to computational limitations the $Pr(CorrectCascade)$ had to be estimated through a Monte Carlo simulation. $Pr(CorrectCascade)$ was graphed in respect to the error in observation, ε, and expert concentration, Q. 3 phenomena were found in these graphs and their causes examined by overlaying graphs of model components.

3 Model Setup

The model assumes the state of the object, V, is either G (good) or B (bad) both with probability $\frac{1}{2}$. There are an infinite number of agents, listed $i = 1, 2, 3 \cdots$, in an exogenous order. An agent is an expert with probability $Q \in (0, 1]$. Non-experts have a signal accuracy of $P \in (0.5, 1)$, and experts have a signal accuracy of $P_{exp} \in (0.5, 1)$ given that $P_{exp} \geq P$. The variable, $\varepsilon \in [0, 0.5)$, is the probability that an agent incorrectly observes the previous agent's decision.

Agents' observations of their predecessors are documented through a Markov Chain. A non-expert's approval increases the states of the Markov Chain by $njump$ with probability a, and an expert's decision changes the current state by $ejump$ with probability b. Moving up the Markov Chain represents an agent observing their predecessor agreeing, correctly or incorrectly (See Figure C).

Thus the variable, a , is the probability that the $agent_{i+1}$ correctly observes $agent_i$ and $agent_i$ accepted plus the probability the $agent_{i+1}$ incorrectly observed $agent_i$ and $agent_i$ rejected.

$$a = P(1 - \varepsilon) + (1 - P)\varepsilon.$$

With the same logic:

$$a = P_{exp}(1 - \varepsilon) + (1 - P_{exp})\varepsilon.$$

The variables, $ejump$ and $njump$, must be set up such that when the Markov Chain is at 0, the agent should be completely indifferent of their predecessors. Thus:

$$Pr(V = G|Y_{exp_1}, Y_{exp_2} \cdots Y_{exp_{njump}}, N_{non-expert_1}, N_{non-expert_2}, \cdots N_{non-expert_{ejump}}) = \frac{1}{2}$$

By Bayes Rule:

$$Pr(Y_{expert}|S = G)^{njump} * Pr(N_{normal}|G)^{ejump} = Pr(Y_{expert}|S = B)^{njump} * Pr(N_{normal}|B)^{ejump}$$

By Substitution:

$$b^{njump} * (1 - a)^{ejump} = a^{ejump} * (1 - b)^{njump}$$

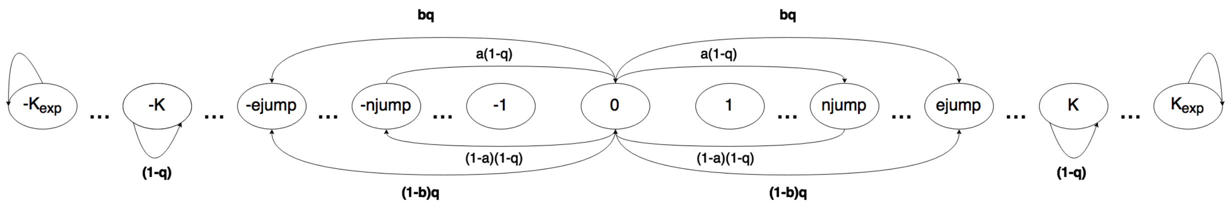
Resulting in:

$$\log_{\frac{a}{1-a}} \frac{b}{1-b} = \frac{ejump}{njump}$$

By inputting values of a and b computed by the inputted parameters, the model finds integer values for $ejump$ and $njump$ to fit the formula above. The formula above is also used to find the probability the chain will move up by 1. This is probability is denoted by ϑ and is later used.

$$\frac{b}{1-b}^{\frac{1}{ejump}} = \frac{a}{1-a}^{\frac{1}{njump}} = \frac{\vartheta}{1-\vartheta}$$

The Markov Chain has 4 trapping states each representing the start of a partial or full cascade: $-K_{exp}$,$-K$,K, K_{exp} (See Figure C). States $-K$ and K trap all non-experts creating a cascade of non-experts. However, this cascade can be broken by an expert. All agents, however, get stuck in the $-K_{exp}$ and K_{exp} trapping states; the model ends when either one is reached (See Figure C).



(C) The Markov Chain used for the model

An agent decides to buy based on his private information and his observations of previous agents. Thus, if (where F represents the current state of the Markov Chain and i_{signal} represents the signal of the agent)

$$P(V = G|S = F, i_{signal}) > \tfrac{1}{2}\ then\ the\ agent\ buys$$

$$P(V = G|S = F, i_{signal}) < \tfrac{1}{2}\ then\ the\ agent\ does\ not\ buy$$

$$P(V = G|S = F, i_{signal}) = \tfrac{1}{2}\ then\ the\ agent\ follows\ the\ crowd$$

An informational cascade occurs when an agent forgoes their own signal in exchange for their predecessor's choices. We also know that informational cascades occur when the current state is K. Thus:

$$P(V = G|S = K, Signal = B) = \tfrac{1}{2}$$

By Bayes theorem:

$$\frac{P(V=G)*P(Signal=B,S=K|V=G)}{P(V=G)*P(Signal=B,S=K|V=G)+P(V=B)*P(Signal=B,S=K|V=B)} = \frac{1}{2}$$

Therefore:

$$P(V = G) * P(Signal = B, S = K|V = G) = P(V = B) * P(Signal = B, S = K|V = B)$$

Through Substitution:

$$(1 - p) * \vartheta^k = p * (1 - \vartheta)^k$$

$$K = log_{\frac{\vartheta}{1-\vartheta}} \frac{p}{1-p}$$

We can substitute P_{exp} to get K_{exp}.

$$K_exp = log_{\frac{\vartheta}{1-\vartheta}} \frac{p_{exp}}{1-p_exp}$$

These values were used in both First-Step Analysis and the Monte Carlo simulation.

3.1 First Step Analysis

Initially we used First-Step Analysis to determine the exact $Pr(CorrectCascade)$. First-Step analysis in Markov Chains takes advantage of the discrete finite number of states to determine the probability of going from one state to another. In this case, First-Step analysis was used to determine the probability to go from state 0 to K_{exp} or $Pr(CorrectCascade)$.

Let $\pi(j)$ be the probability that an agent will go from state j to K_{exp}. $\pi(j)$ can be written in terms of possible states an agent can travel to from j multiplied by the probability an agent can get to K_{exp} from there. For example:

$$\pi(0) = a(1-q) * \pi(njump) + (1-a)(1-q) * \pi(-njump) + b * q * \pi(ejump) + (1-b) * q * \pi(ejump)$$

In the same way, all states can be written by the possible locations an agent can travel to. Let $[A]$ be a column vector containing $\pi(j)$ for all $j \in (-K_{exp}, K_{exp})$ with the first element being $-K_{exp}+1$ increasing by state by state to $K_{exp}-1$ such that the matrix has $2K_{exp}-1$ elements. Let $[X]$ be a $2K_{exp}-1$ by $2K_{exp}-1$ such that $X(l,m)$ is the probability that an agent travels from state l to state m. Let $[B]$ be a column vector of length $2K_{exp-1}$ to account for when K_{exp} can be reached with a single jump such that $B(l)$ is the probability of agent reaching K_{exp} from l. These definitions account for all terms in the system of equations such that $[A] = [X][A] + [B]$. Solving for $[A]$ and taking the middle term, $\pi(0)$, we get $Pr(CorrectCascade)$.

However as K_{exp} increases, the number of equations increase and it becomes increasingly difficult to solve the systems of equations. As an alternative, we use Monte Carlo simulations to provide a precise estimate.

3.2 Monte Carlo Simulation

The model stores all four inputted parameters $(P, P_{exp}, Q, \varepsilon)$ and stores the state at which the Markov Chain is at with variable, *currentstate*. The simulation continues to run until *currentstate* is greater than $K_{exp}-1$ or less than $-K_{exp}+1$. An agent is made an expert by comparing a randomly generated variable against Q. If made an expert, another randomly generated number is compared to b. If greater than b, the *currentstate* increases by *ejump* otherwise it decreases by *ejump*. If a non-expert, a randomly generated number is compared to a, by the same manner *currentstate* moves by *njump*. The program divides the number of time K_{exp} is reached by the number of times the program runs to estimate a probability.

The Monte Carlo estimation and the exact probability calculated by First-Step analysis shared a minimum of 2 significant digits, but in most cases shared up to 4 significant digits. Monte Carlo simulations run much quicker providing for easier graphing in exchange for some static.

4 Conclusions and Future Work

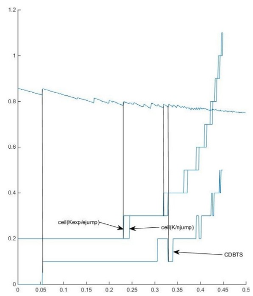
In order to analyze the model, we graphed $Pr(correctcascade)$ against ε for various Q, P, and P_{exp}. We also graphed $Pr(CorrectCascade)$ against Q also for various ε, P, and P_{exp}. In these graphs, three major phenomena arose.

4.1 $Pr(CorrectCascade)$ against Q Graph Shapes

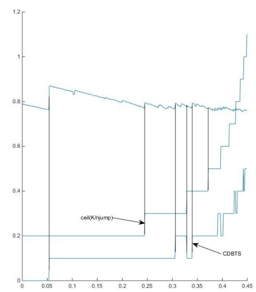
Without analysis, one would think with additional experts, $Pr(CorrectCascade)$ should continually rise. However, simulations runs show various graph shapes (See Figure A). No clear patterns were found. We can then conclude that the varied shapes are due to a combination of factors and cannot be predicted.

4.2 Spikes in $Pr(CorrectCascade)$ against ε

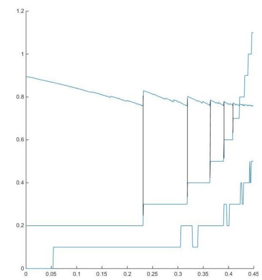
As the noise in the simulation increases, there are spikes in the graphs (See Figure B). In order to explain these spikes we overlaid key components of the graphs as they change with ε in order to see any correlation. We compared the graphs with $\lceil \frac{K}{njump} \rceil$, $\lceil \frac{K_{exp}}{ejump} \rceil$ and the Corrected Distance between Trapping States (CDBTS) (See Figure D). By doing so some of the spikes could be explained, but the cause of others are unknown. As $|(\varepsilon - .5)|$ gets closer to zero, there are more spikes that cannot be explained (See Figure D). $\lceil \frac{K}{njump} \rceil$ is the amount of normal agents required to get past the first trapping state thus when a majority of agents are non-experts (when Q is small) most spikes are explained by $\lceil \frac{K}{njump} \rceil$ (See Figure E). Similarly, when majority of agents are experts (Q is larger), $\lceil \frac{K_{exp}}{ejump} \rceil$ explains the majority of the spikes (See Figure F).



(D) $Q = 0.5$



(E) $Q = 0.05$

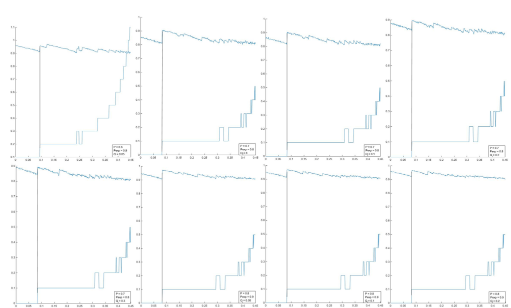


(F) $Q = 0.95$

Figure 3

4.3 $Pr(CorrectCascade)$ against $\varepsilon = 0$

For some graphs, $\varepsilon = 0$ is not the highest $Pr(CorrectCascade)$. In other words, in some cases having perfect observation is less effective than having some error in observation. Again we overlaid $\lceil \frac{K}{njump} \rceil$, $\lceil \frac{K_{exp}}{ejump} \rceil$, and CDBTS to determine the factors which caused the first spike. In all cases when $Pr(CorrectCascade)$ was not the highest point, the first spike was caused by an increase in CDBTS (See Figure G). Essentially, if the distance between trapping states increases before the noise becomes too large, $Pr(CorrectCascade)$ will be higher for some $\varepsilon > 0$. This shows that an additional state between the trapping states outweighs the effect noise has on the simulation.



(G) $Pr(CorrectCascade)$ in respect to ε. Examples where $Pr(CorrectCascde)$ at $\varepsilon > 0$ is greater than $Pr(CorrectCascade)$ at $\varepsilon = 0$

5 Discussion

These results imply that that a heterogeneous set of agents may increase the probability of the correct cascade. The simulation also showed that a lower noise level does not always correlate with a higher $Pr(CorrectCascade)$.

The results experiment closely paralleled previous works with some exception. Jieman Wu's model in Helpful Laymen in Informational Cascades [7] also contained two types of agents: experts and laymen; however, unlike the model discussed in this paper, the type each agent is not publicly known. The paper showed that a medley of experts and non-experts will always have a more efficient cascade. The model discussed in this paper highlights this as a likely possibility but not necessarily an ultimatum. This model also incorporated noise in an identical fashions as Le, Subramanian, and Berry [1].The model discussed in this paper, however, includes a second type of agent. Both models showed that $Pr(CorrectCascade)$ has a non-monotonic relationship with ε. Similar to this model, the model in the Value of Noise for Informational Cascades showed that all spikes correlated when K was increased [1]. However, the model discussed in this paper presents some cases where $\varepsilon = 0$ was not the global maximum, but the Value of Noise for Informational Cascades showed that in every condition no noise always the highest $Pr(CorrectCascade)$. This discrepancy is likely caused by the introduction of a new agent type.

This model provides another variation of the classic BHW model to examine Informational Cascades. Unlike any other model, it incorporated both noise as well as two types of agents. Further expansion on the work may include exploring under what conditions certain shapes of the $Pr(CorrectCascade)$ versus Q graph occur. By categorizing graphs in to different types, certain patterns between them may emerge. For example: certain graph types may transition into others as Q or ε increase. The model could also be expanded to compare the $Pr(CorrectCascade)$ when agents know who the experts are and when the agents do not. Another alteration worthy of pursuing would be making the experts more confident in their own private signal and thus less likely to be swayed by the crowd.

References

[1] 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.

[2] Easley, D., Kleinberg, J. (2010). Informational Cascades. In Networks, Crowds and Markets (1st ed., pp. 425-444). New York, New York: Cambridge University Press.

[3] Gale, D. (1996). What have we learned from social learning? European Economic Review, 40, 617-628.

[4] Le, T., Berry, R., and Subramanian, V. (2014). The Value of Noise for Informational Cascades. Retrieved June 6, 2015, from http://users.eecs.northwestern.edu/ rberry/Pubs/ISIT2014.pdf

[5] Pastine, I., and Pastine T. (2006), "Signal Accuracy and Informational Cascades," working paper.

[6] Sasaki, S. (2005), "Signal Qualities, Order of Decisions, and Informational Cascades: Experimental Evidence," Economics Bulletin, Vol. 3, No. 34 pp. 1-11.

[7] Wu, J. (2015). Helpful Laymen in Informational Cascades. Journal of Economic Behavior and Organization, 116, 407-415.

6 Appendix

Corrected Distance between Trapping States (CDBTS) – The number of agents required to cross between the trapping states.

$$\left\lceil \frac{K_{exp} - \lceil \frac{K}{njump} \rceil * njump}{ejump} \right\rceil$$