View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

MPRA
Munich Personal RePEc Archive

# Learning to be Biased

Ch'ng, Kean Siang and Zaharim, Norzarina
University Sains Malaysia

31. March 2009

Online at http://mpra.ub.uni-muenchen.de/14362/
MPRA Paper No. 14362, posted 31. March 2009 / 04:49

# Learning to be Biased

Ch'ng,K.S. [1] Mohd Zaharim, Norzarina [2]

This version: March 31, 2009

**Abstract**

We simulate societal opinion dynamics when there is confirmation bias in information gathering and spread. If decision making is influenced by confirmation bias, the agent puts more weight on positive information to confirm hypothesis or reservation in the learning process, which renders selectivity in information gathering. If the utility discovered post purchase is low, it is externalized rather than internalized (i.e., self blame) for the selectivity of information. This causes the agent to outweigh the negative information. These two mechanisms are simulated to investigate the societal opinion dynamics and explain behavioral patterns such as overconfidence, stickiness of response and "success breeds success" phenomenon.

***JEL Classification***: C15, C63, D81, D83.

***Keywords***: Confirmation bias; Opinion percolation and convergence; Selectivity in information search; Hypothesis testing

[1]Department of Economics, Universiti Sains Malaysia. Email: cks@usm.my
[2]Department of Psychology, Universiti Sains Malaysia, Email: norzarina@usm.my

# 1 Introduction

Opinion dynamics (OD) refers to the process of opinion adjustments and formations among interacting agents. We build on the *communication regime* by Hegselmann and Krause (2002) to describe opinion dynamics and convergence when agents exhibit confirmation bias in learning and spreading opinions. The bias results in average social characteristics such as over confidence, unequal distribution of revenue among movies and slow response to positive news. The bias reflects social psychological behavior in information selectivity and opinion dynamics.

It is common for humans to learn through observation and communication with other agents before making a purchase decision. This practice enables them to test the expectation or the hypothesis formed prior to the purchase. During the search process, hypothesis-consistent information is more easily adopted to confirm the hypothesis than hypothesis-inconsistent information to disconfirm the hypothesis. In psychology this learning mechanism is referred to as confirmation bias. Confirmation bias can be defined as the seeking or interpreting of evidence or information that conforms to the existing beliefs, expectations or hypothesis (Nickelson, 1998). In this article, we intend to describe this learning process and relate it to collective societal decisions.

Consumers begin to form hypotheses and beliefs on a particular product through observation and experience, which is then used as a benchmark for subsequent learning process. For example, through previous experience of a movie star or a director, consumers expect the same or better stunts and innovations from the newly released movie. If these prior expectations are successfully captured in the trailer or preview, the producer needs less compelling evidence to convince the consumers than when the movie is new. Similarly if an advertisement successfully erects a reference product in the mind of the consumers, the search process will look for information that conforms to the benchmark.

In the learning process with confirmation bias, people typically seek information that supports their hypotheses or beliefs and interpret information in ways that support these hypotheses or beliefs. However, people tend not to seek and perhaps even avoid information which are against the hypotheses or beliefs (Koriat *et al*, 1980). Several mechanisms are at play in this process. First, people tend to restrict their attention to a favored hypothesis. Second, they prefer information that increases their confidence in their favored hypothesis. Third, people look only or primarily for positive cases as examples to support their hypotheses, which then leads them away from discovering that a hypothesis is incorrect (Nickelson, 1998). Fourth, people tend to overweigh hypothesis-consistent evidence and underweigh hypothesis-inconsistent evidence. In other words, people tend to require less hypothesis-consistent evidence to accept a hypothesis than hypothesis-inconsistent information to reject a hypothesis (Pyszczynski and Greenberg,1987). Fifth, people see what they are looking for. If they are looking for evidence to support their existing hypothesis, they will see this supporting evidence.

Psychologists cite many sources of positive expectation, and among common ones are preferential treatment for supportive evidence for one's opinions or beliefs (Baron, 1995), overconfidence of one's belief or hypothesis (Pyszczynski & Greemberg, 1987), effects of one's expectations (Snyder and Campbell, 1984), the importance attached to drawing a correct conclusion (Pyszczynski & Greenberg,1987) and other determinants such as self-esteem and cognitive consistency. When people form hypotheses or beliefs, they tend to expect their beliefs to be confirmed rather than disconfirmed. However, this is not to suggest that agents deliberately ignore evidence, as once expectations are formed, the cognitive process of information search and evaluation take over, giving the sense that the search process is impartial and unbiased. Unaware of the consequences of bias, people engage in the search and evaluation process without intending to treat evidence in a biased way (Nickerson; 1998).

This paper intends to model confirmation bias in the micro level decision making process, and relate it to the macro level opinion formation. We relate the bias in information search to explain the societal decision making behavior, such as the skewness of revenue in the theatrical market, speed of acceptance between positive and negative news, overconfidence and loss aversion in information spread. We find that confirmation bias reinforces confidence and encourages opinion convergence when the hypotheses are confirmed. This is because positive and favorable information reinforces confidence level stronger than non-favorable and negative information. However, opinion divergence occurs when the hypotheses are disconfirmed. This is because agents externalize the blame on external factors rather than the selectivity in information search. This causes the spread of favorable information to be slower than the spread of negative news.

The paper is divided into 6 sections. Section 2 focuses on past studies on confirmation bias and reviews some of the applications of confirmation bias in different streams of literature. Section 3 shows how confirmation bias is formed and how it is applied in the learning process. Based on Section 3, simulations are conducted to test the opinion dynamics given the different levels of confirmation bias. This is done in Section 4. Section 5 discusses the results of the simulations, particularly the formation of opinion clusters. Section 6 explains the implications. Section 7 concludes.

## 2 Past Studies on Confirmation Bias

In the models by Weisbuch et al (2001) and Hegselmann and Keause (2002), opinion dynamics occur when opinions of the interacting agents are close and below a threshold value. Opinions which exceed the threshold level are ignored or opinions which are sufficiently different do not have influence on the agents concerned.

This selectivity is referred to as confirmation bias in psychological literature,(for comprehensive review see Nickerson; 1998). People tend to give greater weight to information that supports their view rather than information that runs counter to it. This preferential treatment for the supportive information is tested based on the tendency of the participants to recall or produce reasons supporting the view they have. In the experiment by Perkins, Farady and Bushley (1991), participants were more able to generate reasons which support their views than reasons which run counter to it. Indeed, psychologists have perceived this tendency as the desire to believe.

The possible explanations for confirmation bias are as follows. First, people have the desire to believe. They find it easier to believe propositions that they would like to be true than propositions that they would prefer to be false. Second, people value consistency and rationality. They would like beliefs and evidence to be consistent, and consistency is an important requirement for rationality. Third, people have cognitive limitations. People tend to acquire information about only one hypothesis at a time, and this hypothesis is considered to be either true or false, but never both simultaneously. People may also fail to consider an alternative hypothesis because they simply do not think to do so. Fourth, people tend to have positivity bias. When there is no compelling evidence, people tend to assume a statement to be true than false (Nickerson, 1998). Fifth, people are pragmatic and tend to avoid making errors. People are more likely to be concerned about desirable outcomes than about the truth or falsity of hypotheses (Schwartz, 1982). Another reason for the presence of bias is deduced from the prior information dependence behavior. Studies by Snyder and Gangested (1981) and Snyder and Swann(1978) involved personality tests highlighting the degree to which people see or remember a particular occurrence corresponding to what they look for in a similar situation in future settings.

The bias is present in many learning processes. In a study by Baron (1995), many people judged the quality of arguments to be higher for one-sided arguments than for two-sided arguments. This finding suggests that the bias for one-sided arguments may be partially due to common beliefs about what makes an argument strong. In another study (Kuhn, 1989), children and young adults did not acknowledge evidence that was inconsistent with a favored theory or saw the evidence selectively or distortingly. The same evidence was also interpreted differently in relation to a favored theory versus an unfavored theory. Matlin and Stang (1978) argued that people find it easier to believe propositions they would like to be true than propositions they would prefer to be false. This argument is furthered by Weinstein (1989) who proposed that people demand very little compelling information for a conclusion that they would like to accept.

# 3 Information Search and Adjustment of Opinions

We start with a population of N agents with continuous beliefs or hypotheses (R) which are randomly distributed among agents on a lattice. We adopt the communication regime by Hegselmann and Krause (2002) and assume all agents can perceive all *local* agents' opinions. We limit local interaction to eight agents. In each time step, an agent can interact with one agent locally, and that agent interacts with his neighboring agents locally. There are a total of 50 x 50 agents on the lattice.

Each individual is represented by a vector on a site, which consists of four elements. The first element is the opinion, which is represented by $X \in \{0,1\}$, $R \in \{0,1\}$ is the reservation value, $U \in \{0,1\}$ is the utility and $I \in \{R,1\}$ is the information spread in the lattice. Non-adoption occurs when opinion is less than one or $X \neq 1$. Reservation or hypothesis or belief is the benchmark value against information spread in the lattice. During interactions, the reservation value is compared to information relayed in the lattice to evaluate whether the information is positive or negative. Information which exceeds this value is considered positive, otherwise negative.

At each time step any two agents meet; they seek information about a particular product from the agent they interact with. Adjustment of opinion takes place between these two agents. Suppose two agents $x_i$ and $x_j$ meet and the adjustment of opinions takes place;

$$X_i = x_i + \mu\left(\{I_j\}, R_i\right), \tag{1}$$

where $i \neq j$ and $\mu : [P(I_j), R_i] \rightarrow \Re$, where $\Re$ is positive real number. And the information (I) spread is defined as:

$$I_i = \beta\left(U_i, R_i\right), \tag{2}$$

The $\mu$ and $\beta$ are convergence parameters which are determined by information (I) and reservation or hypothesis (R), and utility (U) and reservation (R) respectively. In equation [1], if $I_j > R_i$, the adjustment of $\mu$ is in the range of 0.1 to 0.3, and in the range of 0 to 0.1 otherwise. If $U_i > R_i$, the value of $\beta$ will be in the range of $(R, 1)$. This means that if the utility is higher than reservation or hypothesis, agent will spread information (I) at the level which is higher than the reservation point (R), and if it is lower than reservation value agent will spread information (I) lower than his reservation value. Therefore, $\beta$ will be in the range of $(0, R)$. In the model, we assume the threshold or reservation value R to be constant throughout the interaction.

The agent practices confirmation bias in three different ways. *First*, we assume that the agent will increase his preference for a product by 0.3, i.e., $x + 0.3$ each interaction if he receives positive information from either one of his eight neighboring agents; that is, the quality information received from one of the agents is $I_j > R_i$, provided that the neighboring agent has adopted the product, $X_j = 1$ and the agent himself has not adopted the product, $X_i = 0$. However, it takes *three* agents to persuade that the product is not good. We denote á as a subset of the agent's eight neighbors, Ä={N,NE,E,SE,S,SW,W,NW} and á $\in$ Ä where á=3 can be any three agents from set Ä who have had a bad experience with a particular product. If á display negative feedback, i.e., $U_{á} < R_{á}$ and for this reason the agents relay $I_{á} < R_c$, where agent c is the agent we want to investigate. If in the first interaction, á={E, N, W}and second interaction, á={S,NW,SW} the adjustment of opinion with confirmation bias can be illustrated as;

$$X_i(t) = x_i$$

$$X_i(t+1) = X_i(t) + \tilde{\mu}(I_E(t), I_N(t), I_W(t), R_{E,N,W}(t))$$

$$X_i(t+2) = X_i(t+1) + \tilde{\mu}(I_S(t+1), I_{NW}(t+1), I_{SW}(t+1), R_{S,NW,SW})$$

we set $\tilde{\mu}(\bullet)$ = -0.1 for both cases. Thus, the opinion dynamic is biased for *reservation or hypothesis consistent* information.

*Second*, in the learning process, agent will decrease 0.1 of his preference if he finds there are three agents in the neighborhood who have not adopted the product. Let $(x_i)_{i=1}^{\infty}$ denote the hypotheses formed in a lattice, and the local interaction of $x_i$ is $(x_i)_{i=1}^{Ä}$. If $\sum x_{i=1}^{3}$=0, $\mu_i \rightarrow -0.1$.

*Thirdly,* after adoption, the agent will relay two types of information: (1) bias up and (2) bias down. Bias up occurs when agent outweighs the positive information. If the agent is neutral, and no bias is applied, the positive information relayed ranges from R to 1 when $U > R$. But if it is outweighed, the information relayed becomes $I_{biasup} = (R + \epsilon, 1)$, where $\epsilon$ is some positive number. And if $U < R$, agent will apply bias on the negative information relayed, $I_{biasdown} = (0, R - \epsilon)$. In a situation when agent outweighs negative information than positive information, then $I_{biasdown} \succ I_{biasup}$ where " $\succ$ " is denoted as preference over. But if agent is biased for positive information, then $I_{biasup} \succ I_{biasdown}$. This third assumption is based on the behavior of consumers who like to externalize the hypothesis inconsistent information. Instead of blaming himself for his selectivity of information as in the previous two learning processes explained above, the agent blames the low quality of the product and the producer. [1]

# 4 Simulation

We simulate the opinion-dynamics when agents practice confirmation bias in information gathering and spreading. We hypothesize that;

- The convergence of opinions occurs if information spread condorms to reservation or hypothesis.
- If the information is confirming, the adoption rate is higher than when the information is disconfirming; this renders different speed of adoption in the two situations.
- Disconfirming information is over-weighted more than confirming information after purchase which causes opinion divergence.

Confirmation bias renders higher sensitivity of consumers to positive news than negative news. Figure 1 illustrates the opinion dynamics of different levels of information spread and reservation value. If we imagine R to be different levels of hypothesis of a consumer, any level of information ( information spread depends on the level of utility spread in the lattice) which is disconfirming (low U), will not be adopted. However, when the information conforms to the hypothesis, as when the level of U is high, the adoption of the information is smooth. This different response renders various slopes of adoption curves in the figure for different values of U. As shown in the figure, the adoption rate is higher when there is hypothesis consistent information (high level of U) than when there is hypothesis inconsistent information.

[1] The program of confirmation bias in Mathematica is available upon request from reader.

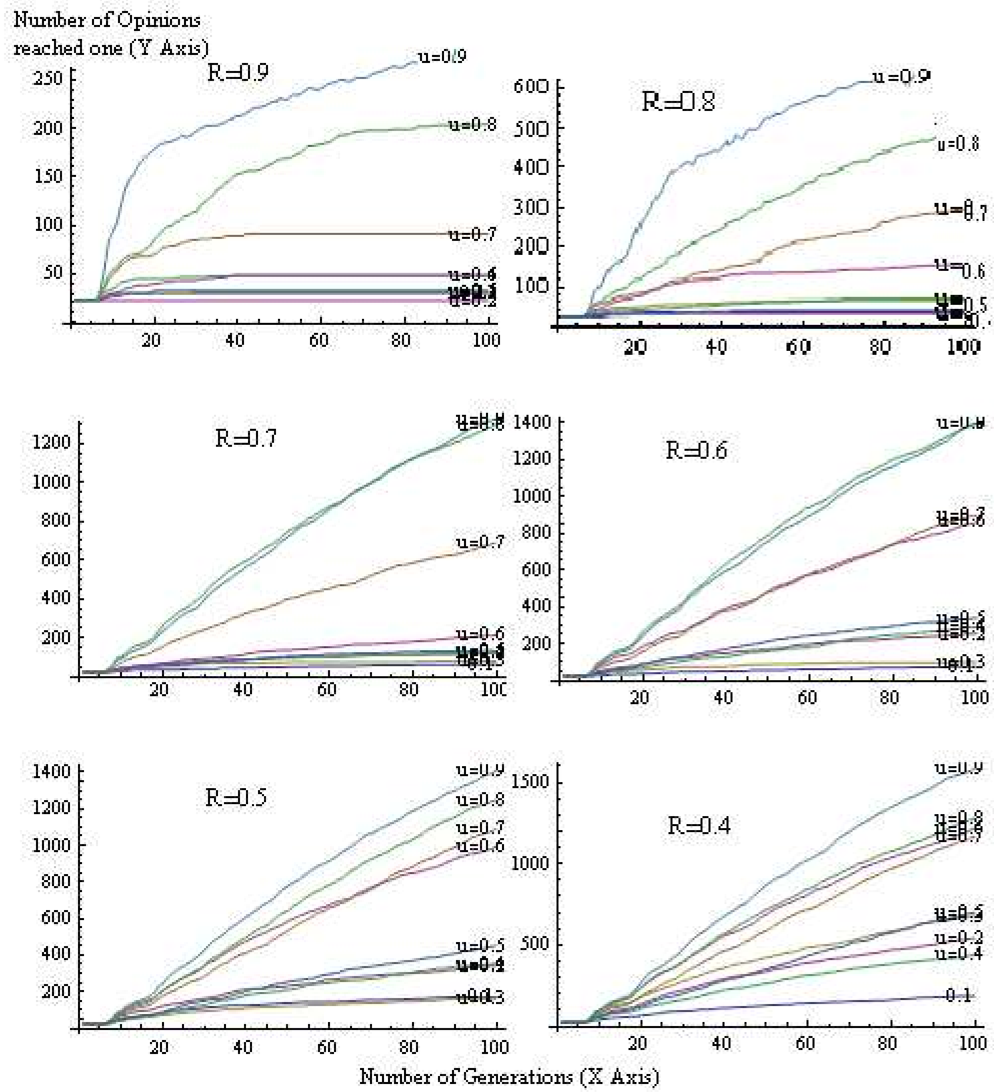


Figure 1: The opinion dynamics for different levels of information spread (I) and different levels of utility (U) . Axis y denotes the number of opinions reaching 1, and axis x denotes number of interactions when R varies.

### 4.1 The Basic Case: Small Threshold

We start with the simulation when the threshold value between reservation and *ex post* utility is at the minimum, i.e., $U - R \geq 0$. For a small threshold, opinion dynamics is thwarted by a high reservation value; although the *ex post* utility is high, the reservation value in the neighborhood is also high.

- For a small threshold value between reservation and *ex post* utility, many clusters of opinion will be formed.
- The non-consensus of opinions causes heterogenous information to be relayed. Convergence occurs only at the local level; global convergence is difficult to achieve.

### 4.2 High Threshold When $U > R$

When the *ex post* utility is higher than the reservation value of an agent, he will relay $I_{biasup} = (R + \epsilon, 1)$. We observe a convergence of opinions into a single cluster on the lattice. This is due to the overconfidence and favorable treatment of positive information. Overconfidence occurs when agents spread bias up information, and the contacted agents increase their opinions faster when the information is confirming rather than disconfirming information.

### 4.3 High Threshold When $U < R$

When the utility is lower than the reservation, disconfirming information will be outweighed and causes global non-consensus. Agents start to blame the low quality of a product rather than the selectivity of information during the search process. The bias down information, i.e., $I_{biasdown} = (0, R - \epsilon)$ will be further enhanced by lower reserved agents. If two contacting agents are related through a chain of interactions, the bias down information will spread and convince other agents from different reservation clusters.

## 5 Results and Discussion

When comparing the opinion dynamics for different thresholds, decreasing difference in reservation and utility results in a larger variety of final opinions (opinion clusters). Observing the initial and final opinions in Figure 2, one sees the dynamics converge to many opinion clusters. Initial opinions gathered are later segregated due to low difference between utility and reservation values. Many outliers are apparent in the plot. A similar pattern is observed in Figure 3 with a higher value of reservation and utility level.

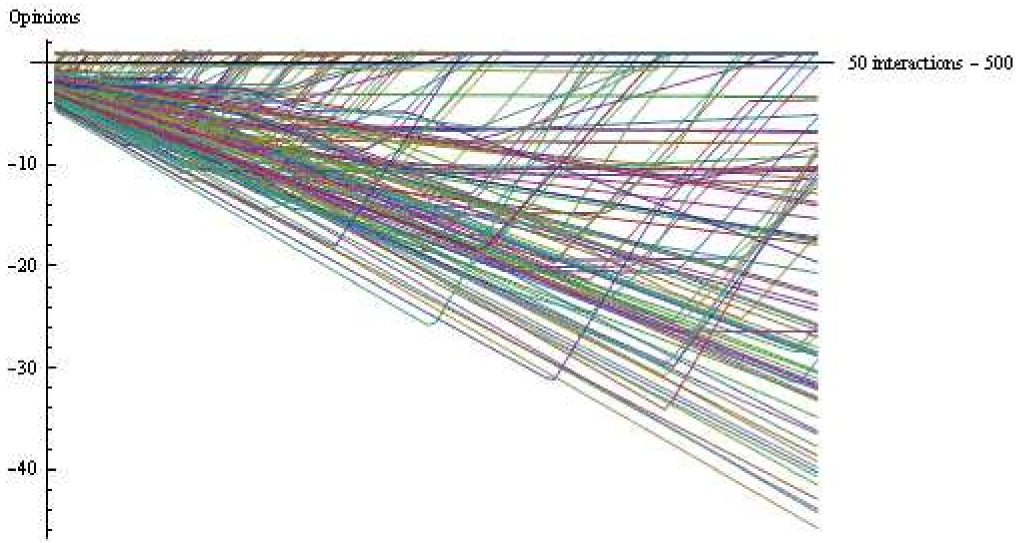


Figure 2: Slow Opinion convergence when R=0.3,1 and U=0.3,1, 200 individuals and 50 to 500 interactions. Each line represents individual opinion.

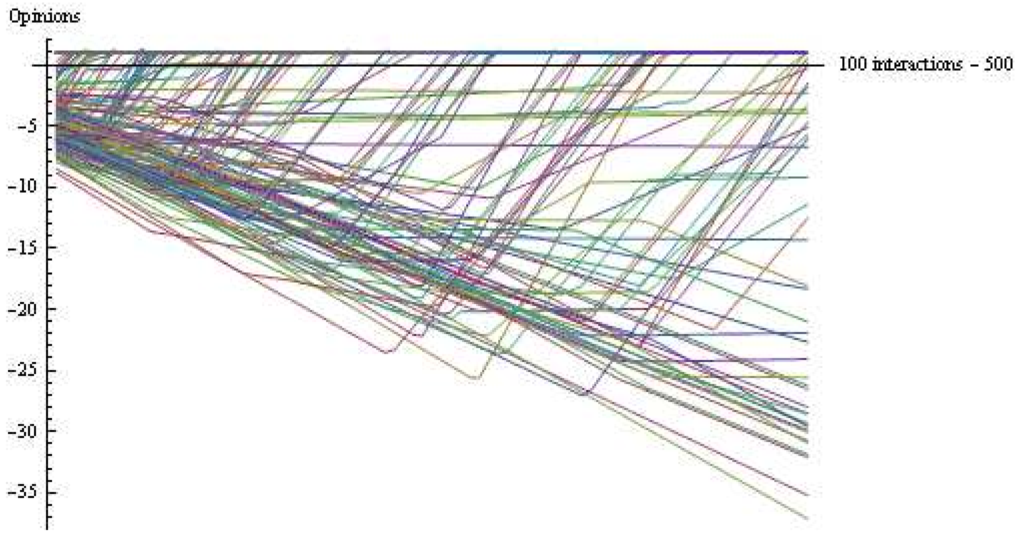


Figure 3: Opinions clusters when R=0.8, 1 and U=0.8,1, 200 individuals and 100 to 500 interactions.

## 5.1 Large Difference Between Reservation and Utility

Large differences of reservation and utility, particularly when $U > R$, renders fast adoption due to selectivity of information to confirm hypothesis during the learning process. An explanation offered is that adopting a confirming information is faster than accepting a disconfirming information. When agents become overconfident, clusters of opinions break down. Figure 4 shows the opinion convergence.

However, in the case when $R > U$, initial opinions are segregated further by agents over-weighting the negative quality information. When the *ex post* utility is low, an agent will tell a very bad story about the product to other interacting agents. This bias in information spread will kill the product. Figure 5 shows the divergence of opinions when agents overweigh negative information.

## 5.2 Opinion Clusters for Different Values of R and U

For a large threshold value, when $U - R > 0.2$, opinion convergence or fewer clusters are observed. Many opinion clusters are formed if $U - R < -0.1$. The opinion dynamics become static or stagnant when $U - R$ is in the range

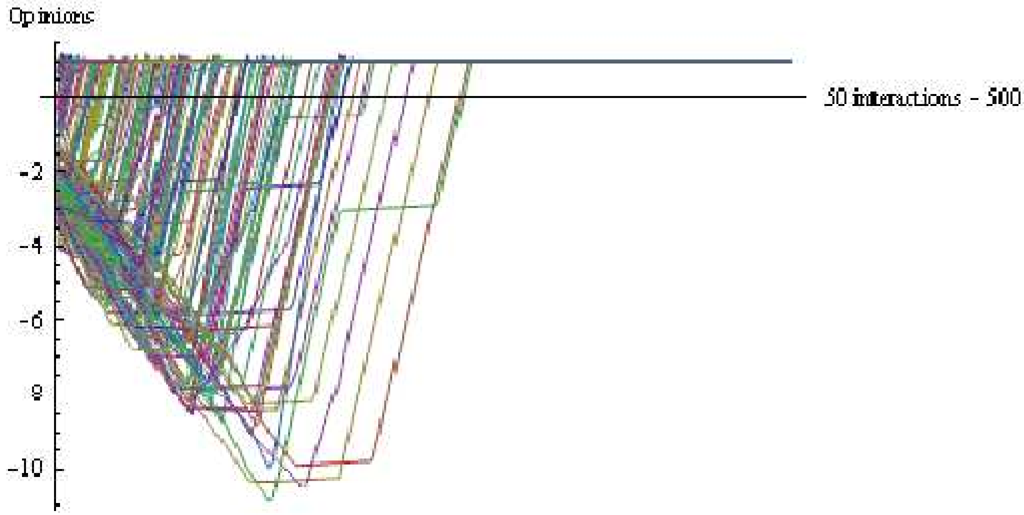


Figure 4: Opinion convergence when R=0.3,1 and U=0.8,1, 200 individuals and 50 to 500 interactions and the convergence occurs around 260 interactions

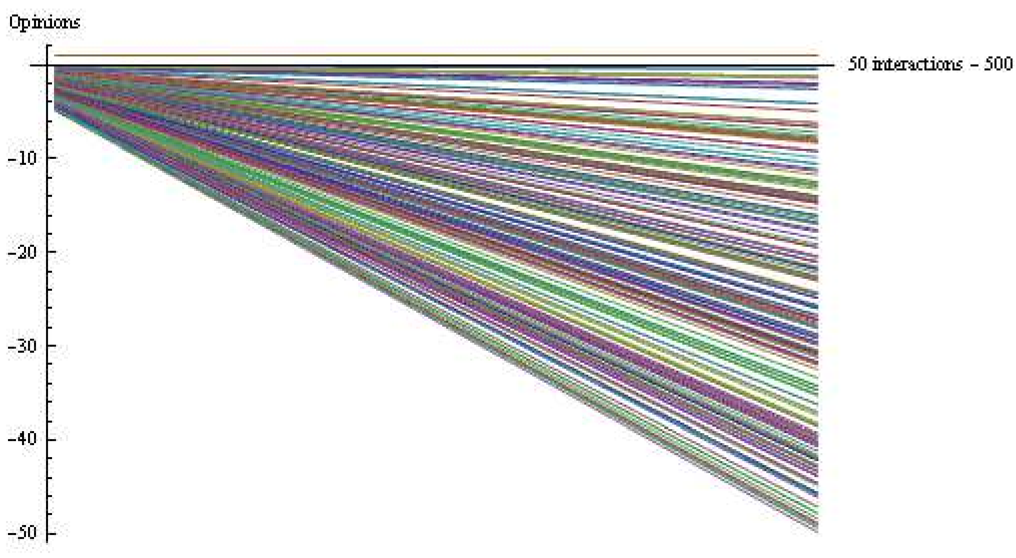


Figure 5: Opinion divergence when R=0.8,1 and U=0.3,1, 200 individuals and 50 to 500 interactions

of $\{0, 0.2\}$. Figure 1 shows the opinion clusters for different values of utility and reservation.

On the squared lattice, any agent can interact with his eight connected neighbors. At each time step, a pair is randomly selected among the connected agents and their opinions are updated based on equations [1] and [2]. Figures 6 and 7 show the opinion percolation when the reservation value is high. From the figures, the percolation (which is denoted by black dots on the lattice) is local, i.e., islands of opinion do not connect to each other. Consensus is reached when the threshold value, i.e., $U - R$ is more than 0.2. Figure 8 shows the global percolation of opinion when the threshold value is high, $R = 0.3$ and $U = 0.8$. This is due to incrementation of confidence when agents select hypothesis-related information.

# 6 Implications

### The Stickiness of Response When Information is Positive

One of the causes of confirmation bias is the loss aversion behavior; where the utility gained from the confirmed preference is lower than the utility loss if the preference is disconfirmed. (Pyszczynski and Greenberg, 1987; Tversky and Kahneman, 1991) . Thus after the purchase, if the quality is lower than expected, instead of blaming himself for making the choice, the agent blames the producer or product, which causes him to spread outweighed negative news in the market. This makes the product less welcome in the subsequent interactions among the agents and the percolation becomes localized. The presence of bias also highlights that negative news are taken more seriously than positive news and the speed of adoption of bad information is higher than the speed of adoption of positive information.

### Overconfidence

It is usually not difficult to find evidence to substantiate a belief when a preference or hypothesis is formed. If we imagine the lattice sites as confidence and belief, Figure 8 displays the opinion-dynamic when agents exhibit overconfidence. In the figure, information search involves positive information that confirms the opinion. For example, the opinion is incremented by 0.3 when the positive information is received, i.e., $I > R$, which is higher than 0.1 if the agent received negative information. Although he may encounter negative information in the subsequent interactions, the increase in confidence is higher than the decrease in confidence level. Over time, this reinforces the positive opinion of the agent. Therefore, in Figure 8, although the agent starts with low confidence (i.e., low value of x), information search in the interaction process can help reinforce the confidence. Overconfidence leads to societal convergence.

**Divergence of Willingness to Pay for Similar Products**

Confirmation bias occurs when people want to believe a particular belief is true. If two different belief reference points are formed, say by advertisements, subsequent information received will be benchmarked against the points. Different reference points are tantamount to different expectations, e.g., expectation on quality, comfort, performance, etc. This leads to different levels of willingness to pay (maximum amount willing to pay).

**Success Breeds Success**

The presence of bias in information search and relay can in particular be found among movie goers. Due to prior experience of a director or star, consumers expect the same or higher degree of innovation in the newly released movie. If these expectations are captured in the preview or trailer before release, the subsequent information spread will be in favor of this movie. The spread of information also becomes more in favor of the movie as consumers become more confident.

If the information is negative and disconfirms the preference or expectation, consumers will spread and outweigh the negative information. These two phenomena, overconfidence when the choice is optimal and outweighing the negative information when the quality is not as expected, cause the unequal distribution of revenue among the movies.

# 7 Conclusion

In this paper, we simulate decision making when agents practice confirmation bias. When there is confirmation bias, it takes less hypothesis-consistent information to confirm reservation than hypothesis-inconsistent information to disconfirm reservation. This causes faster upward opinion adjustment when agents receive positive information than negative information. We simulate three different sources of confirmation bias as in past literature and relate them to societal decisions. They are preferential treatment on supportive information by agents, overconfidence of one's belief, and externalization of dissatisfaction when the quality after purchase is low. The first two learning processes lead agents to adopt positive information more easily than negative information, which leads to convergence of opinions. However, overconfidence with one's hypothesis leads agents to treat hypothesis-consistent information more seriously than hypothesis-inconsistent information. This leads agents to spread very negative information when the product is less than satisfactory and that causes divergence of opinions.

# References

Barron,J (1995), 'My side bias in thinking about abortion.', *Thinking and reasoning* **7**, 221–235.

Farady, M, P. and Bushley, B (1991), *Everyday Reasoning and the roots of intelligence.*, Hillsdale, NJ: Erlbaum.

Hegselmann,R and Krause, U (2002), 'Opinion dynamics and bounded confidence models: Analysis and simulation', *Journal of Artificial Societies and Social Simulation* **5**(3).

Koriat,A.,, Lichtenstein,S and B, Fischhoff (1980), 'Reasons for confidence', *Journal of Experimental Psychology: Human Learning and Memory* **6**, 107–118.

Kuhn, D. (1989), 'Children and adults as intuitive scientists', *Psyhological Review* **96**, 674–689.

Matlin M.W and Stang, D.J (1978), *The Pollyanna principle: Selectivity in languages, memory and thought.*, Cambridge, MA: Shenkman.

Nickerson, R. S. (1998), 'Confirmaion bias: A ubiquitous phenomenon in many guises.', *Review of General Psychology* **2**(2), 175–220.

Pyszczynski,T and Greenberg, J. (1987), *Toward and integration of cognitive and motivational perspectives on social inference: A biased hypothesis testing model. Advances in experimental social psychology.*, New York: Academic Press.

Snyder, M , and Campbell, B. (1980), 'Testing hypothesis about other people: The role of the hypothesis.', *Personality and Social Psychology Bulletin* **6**, 421–426.

Snyder,M and Gangestad.S (1981), *Hypothesis testing process. New Directions in attribution research.*, Vol. 3, Hillsdale, NJ: Erlbaum.

Snyder,M and SwannW.B (1978), 'Behavioral confirmation in social interaction: From social perception to social reality.', *Journal of Experimental Social Psychology,* **14**, 148–162.

Tversky, Amos and Daniel Kahneman (1991), 'Loss aversion in riskless choice: A reference dependent model.', *Quarterly Journal of Economics* **107**(4), 1039–1061.

Weinstein,N.D (1989), 'Optimistic bias about personal risks.', *Science*, **246**, 1232–1233.

Weisbuch, Gerard., Deffuant, Guillaume., A. and Jean Pierre, Nadal (2001), Interacting agents and continuous opinions dynamics, Technical report, Laboratoire associe au CNRS (URA 1306) , a I'ENS et aux Universities Paris 6 et Paris 7.

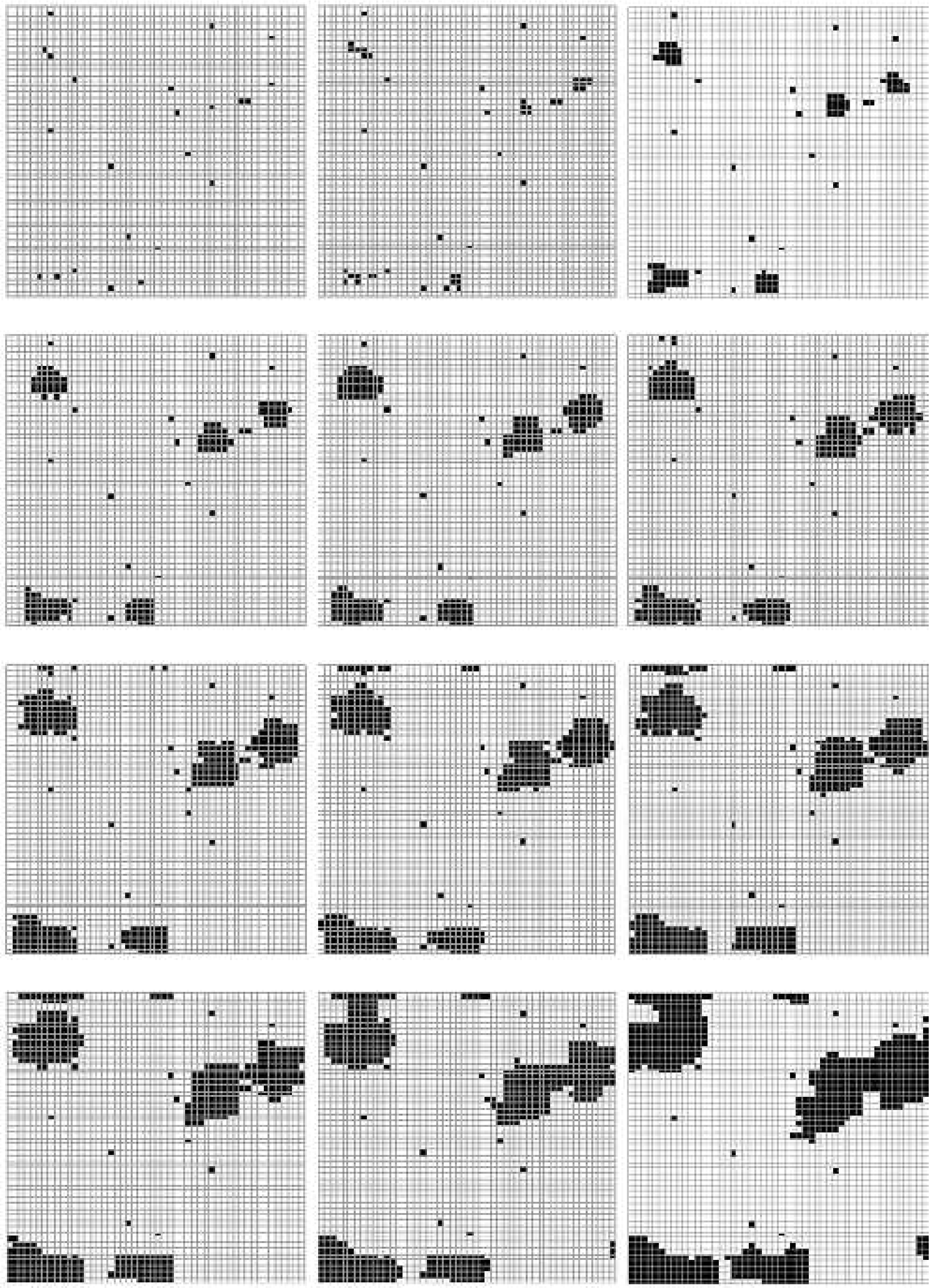

Figure 6: The percolation of opinions when U=(0.8,1), R=(0.8,1), I=0.5 for 1(23),10(48),20(97),30(147),40(197),50(246),60(286) ,70(327),80(370),90 (404), 100 (434) and 150 (567)interactions (number of opinions reached one).

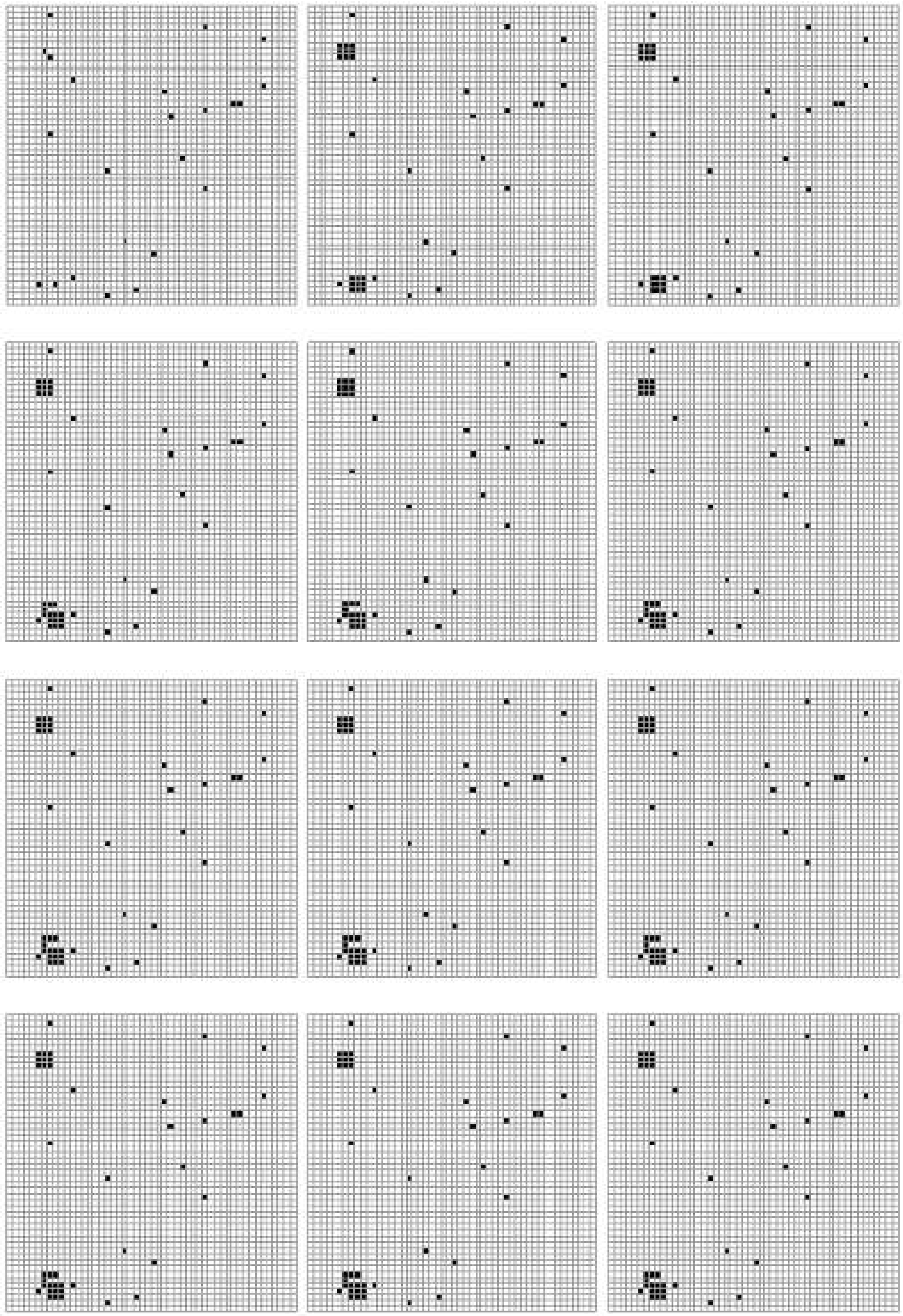

Figure 7: The percolation of opinions when U=(0.3,1), R=(0.8,1), I=0.5 for 1(23),10(38),20(38),30(43),40(43),50(43),60(43),70(43),80(43),90(43), 100(43) and 150(43) interactions

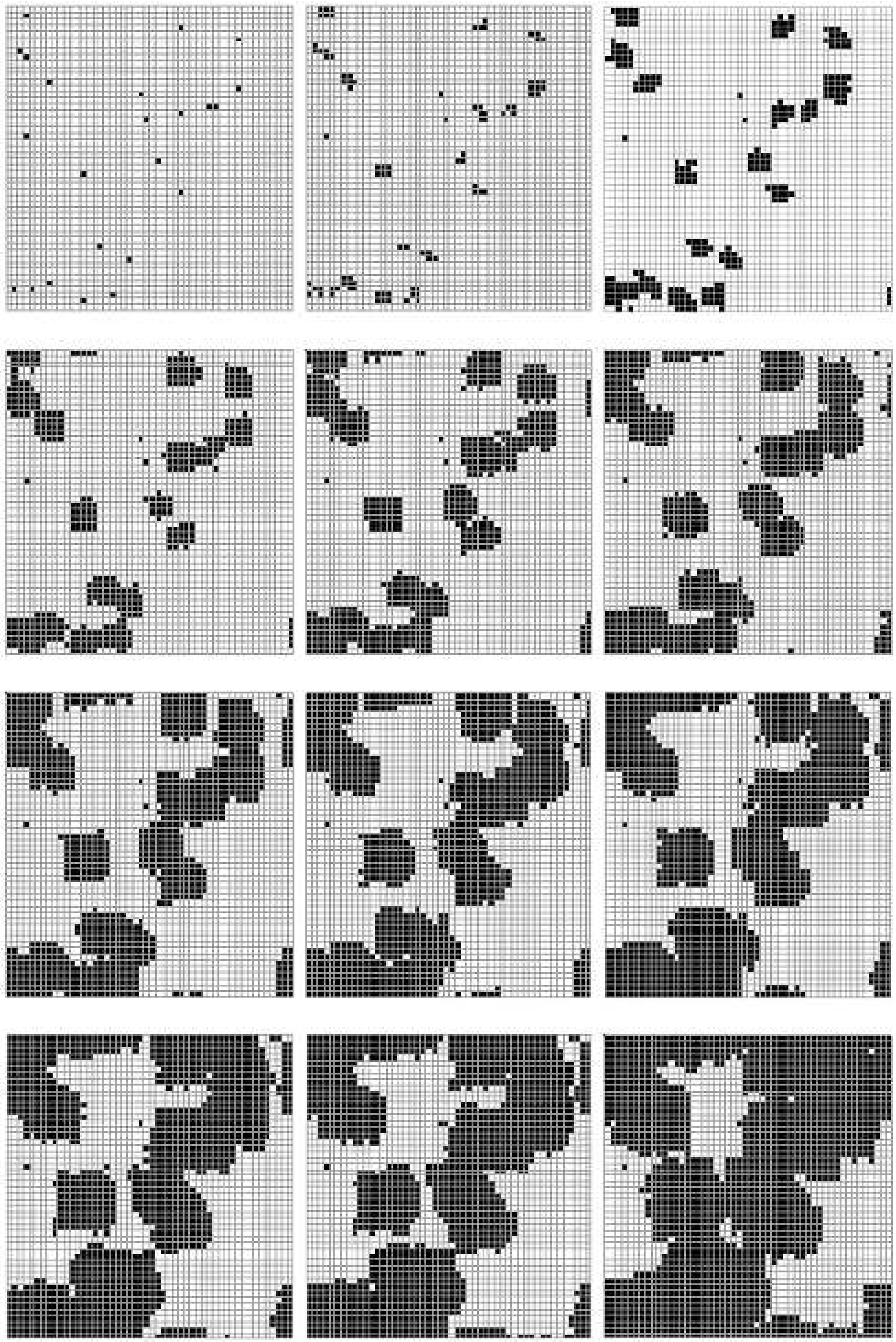

Figure 8: The percolation of opinions when U=(0.8,1), R=(0.3,1), I=(0,R-0.5) for 1(23),10(93),20(250),30(424),40(612),50(774),60(911),70(1045),80(1155),90(1265), 100(1354) and 150(1720) interactions.