Journal of Psychology and Behavioral Science
June 2014, Vol. 2, No. 2, pp. 71-82
ISSN: 2374-2380 (Print) 2374-2399 (Online)

Published by American Research Institute for Policy Development

Is More Information Better? Social Learning with Confirmatory Bias

Matthew T. Clements[1]

Abstract

Confirmatory bias can lead to persistent divergence of opinion regarding matters of fact. Greater availability of information can foster greater agreement, but only if the degree of confirmatory bias is small; otherwise divergence is more likely when information is more easily obtained.

Keywords: imperfect information; cognitive bias; social learning

Is more information better? Social learning with confirmatory bias Public opinions clash for many reasons, including subjective issues. It is interesting when opinions diverge over matters of fact, when there is a concrete, objective sense in which one side is right and the other is wrong. One example is the controversy over the safety of childhood vaccinations. A 1998 study (Wakefield et al, 1998) claimed a link between the MMR vaccine and autism. The study met with skepticism from the medical establishment and was later thoroughly discredited. Nonetheless, opinions remain divided on the specific issue addressed in the study as well as other concerns about adverse effects of vaccines. The problem of global warming is another example in which public opinion often differs from expert opinion. Although these are complex issues, they tend to be seen as binary questions, where individual opinions take one of two sides.

One might think that ease of information transmission would alleviate this phenomenon, but this is clearly not the case in all instances: such differences of opinion can persist for long periods of time, even in the information age.

[1] Matthew T. Clements, Department of Finance and Economics, St. Edward's University, 3001 S. Congress Ave., Austin, TX 78704. E-mail: mattcl@stedwards.edu.

The lack of popular consensus on some issues can confound an unbiased observer, and a biased observer can often find agreement with whatever point of view is desired (although there may be a clear majority in favor of one opinion). This is true even for some distinctly binary, factual issues that would seem to be straightforward to verify, such as whether the Holocaust actually happened, or where Barack Obama was born.

There is a substantial literature on information aggregation among individuals, described in more detail below, that does not generally allow this kind of divergence. One possible explanation is that people are subject to confirmatory bias, meaning that they have some tendency either to disregard or misinterpret evidence that contradicts their opinions. The purpose of this paper is to examine the effects of confirmatory bias in a social learning setting. It is easy to imagine that this bias will potentially lead to divergence of opinion. However, it is not obvious how other factors interact with the bias, and how we can explain different outcomes. I present a model to address this. I find that confirmatory bias has some effects that are very intuitive, but that greater ease of information transmission can either encourage or inhibit divergence, depending on how severe the bias is.

In the following section, I review the related literature. I then present the model and its implications. In the next section I discuss the application of the model to examples, as well as potential complications to the model.

Literature

An extensive literature on social learning, beginning with and inspired by Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992), explores many aspects of the interplay between agents' private information and what inferences they draw from observing others. One focus of this literature has been the potential for an *information cascade*, or *herding*, in which each agent follows the actions of others regardless of the agent's own private information. The canonical example is the choice between two restaurants by imperfectly informed consumers: in choosing between restaurants A and B, a consumer may rationally infer that restaurant A is of higher quality because it has more customers, even if the consumer has some private information to the contrary.

A later piece by Bikhchandani, Hirshleifer, and Welch (1998) presents a thorough overview of the bulk of the literature, which includes many methodological variations on the original two papers. Notable issues are whether cascades aggregate information efficiently and under what circumstances a cascade can be dislodged. Even in the latter circumstance, it is not generally the case that different agents can simultaneously make different choices. Two exceptions to this are Callander and Hörner (2009) and Andreoni and Mylovanov (2012); in both cases, divergence exists because of heterogeneity of information available to agents. Further work concerning social learning and informational efficiency includes Ellison and Fudenberg (1993, 1995), Smith and Sorensen (2000), Banerjee and Fudenberg (2004), and Cao, Han and Hirshleifer (2011).

A related literature considers departures from full rationality in learning and aggregation of information. Rabin and Schrag (1999) present a model of confirmatory bias from the perspective of the individual agent as well as reviewing the psychology literature on the existence and nature of the bias. More recent evidence that confirmatory bias exists includes Jones and Sugden (2001) and Kataria (2012). DeMarzo, Vayanos, and Zwiebel (2003) and Eyster and Rabin (2010) specifically consider the effects of naiveté in agents' process of inference on the aggregation of information. As in the earlier literature, a primary issue is the possibility of inefficient herding, as well as the potential for agents to be overconfident in their beliefs.

The present model starts with a similar conceptualization of confirmatory bias to that of Rabin and Schrag, but goes further in considering the interaction of a population of agents, all of whom may be subject to the bias. This model is particularly useful in exploring the effects of easier transmission of information. Unlike Callander and Hörner (2009) and Andreoni and Mylovanov (2012), the present model leads to potential divergence among agents who all have access to the same information.

Model and Results

Define the state of the world as $\Phi \in \{0, 1\}$. This is a factual statement that may be either true or false: for example, "Barack Obama was born in the United States." There are N agents that each receive a private signal about the state of the world, $S \in \{0, 1\}$, and otherwise have no prior belief about Φ.

Let *S0* and *S1* be the respective numbers of agents that receive signals $S = 0$ and $S = 1$. Without loss of generality, I assume $S1 \geq S0$. Agents form beliefs about Φ through the following process, with further description of each stage below:

1. Initial beliefs are equal to private signals.
2. Agents publicly signal their beliefs.
3. Public signals are observed but may be misperceived.
4. Agents update beliefs.
5. Steps 2 through 4 are repeated until a steady state is achieved.

In stage 2, let *m* be the number of agents with belief $\Phi = 0$, and *n* with $\Phi = 1$, where $m + n = N$. Agents are subject to confirmatory bias in the observation of public signals: if a given signal contradicts an agent's private signal, the agent misperceives that signal as confirmatory with probability α (and agents correctly perceive all signals that actually do confirm the agent's prior belief).1 The agent adopts the belief of the (perceived) majority, relying on the private signal to break ties. Once an agent perceives a signal, whether or not this perception is correct, it does not change except possibly when the signal itself changes; i.e., agents only re-evaluate signals that have changed.

It is difficult to make sense of the idea of persistent divergence of opinions if agents make irreversible choices, which is typically the case in the literature cited above.2 Allowing agents to change beliefs also removes much of the fragility and path dependence found in the cascade literature. Different modeling assumptions can lead to complex dynamics, but the focus of this paper is the relative likelihoods of different steady states. Another contrast to the cascadeliterature is that the primary import of private signals is their effect on the direction of the agent's bias.

A steady state, defined by (*m*, *n*), is a set of beliefs and signals that form a perfect Bayesian equilibrium: in a steady state, each agent's beliefs are consistent with the signals that agent perceives, and there is no further updating. *Convergence* is the mathematical operationalization of consensus: a steady state in which either $m = 0$ or $n = 0$, i.e. agents all adopt the same belief, as opposed to *divergence*.

A complication to this framework, discussed below, is that transmission of information is somehow costly. Proofs of the following propositions are in the appendix.

The following case serves as a baseline and illustrates some of the contrast with the cascade literature.

Proposition 1

If $\alpha = 0$ and $S0 \neq S1$, convergence on the belief that $\Phi = 1$ is certain.

In the absence of confirmatory bias, except for the specific case in which initial beliefs are evenly divided, there is always agreement on the majority belief. The proof is trivial: all agents perceive that there is a majority and adopt the majority belief. If $S0 = S1$, agents all maintain their prior beliefs. In either case, steady state is achieved immediately.

When confirmatory bias is present, some of the effects are fairly intuitive, as illustrated in the following proposition.

Proposition 2

The probability of convergence is decreasing in α. When $\alpha = 1$, divergence is certain. The probability of convergence is decreasing in *S0*, and the expected size of the minority coalition is increasing in *S0*. The results are more subtle with respect to the effect of population size, which interacts with the degree of bias.

Proposition 3

The probability of convergence is decreasing in N if α is large, and increasing in N if α is small.

Stronger confirmatory bias magnifies the effects of population size.

If the degree of confirmatory bias is severe, divergence is more likely in the presence of more signals because agents are likely to interpret these signals in favor of their own beliefs. When only mild bias is present, divergence is unlikely but is more likely when there are fewer signals. There is a greater chance that any one agent will misperceive enough of the (few) available signals to misperceive which is the majority view. As the population increases, it is more likely that the majority view dominates.

There are two ways in which the model can incorporate the ease with which information is transmitted among agents. Here I consider one of them, that agents only observe some fraction of public signals. The other is that agents incur a cost for each public signal that they observe. There would have to be some benefit of obtaining signals in order for agents to have incentive to obtain them. Depending on the nature of the benefit, the two formulations could lead to the same implications. A direct implication of Proposition 3 is that a restriction on the signals that agents receive that is uncorrelated with other parameters has effects indistinguishable from those of population size.

Corollary

If each agent observes a random fraction θ of public signals, $0 \leq \theta \leq 1$, then the probability of convergence is decreasing in θ if α is large and increasing in θ if α is small. As is clear from the proof of Proposition 3, perceiving a random fraction of signals has the same effect on the probability of convergence as the size of the population. Other effects would arise if agents acquire information systematically: for example, if agents are more likely to notice signals that agree with their own, this would exacerbate confirmatory bias as it is defined here. We could interpret θ as the distance from which signals can be perceived, where agents always perceive signals from those closest to them. This interpretation is consistent with the results as long as prior beliefs are uncorrelated with the spatial distribution of agents. We could then imagine disjoint pockets of agents within the population as a whole, with agreement within pockets but potential disagreement across them. For the results above, there is no need for divergent coalitions to be cohesive in any sense other than agents within each coalition hold the same opinion.

Similarly, if agents incur a cost of acquiring more signals, the results would be consistent with the above as long as the benefit to the agent is uncorrelated with other model parameters.

Discussion and Conclusion

In addition to experimental evidence that confirmatory bias exists, there are many popular indications that it plays a role in the formation of opinions. For example, the Economist (2012) reports that lack of scientific literacy does not explain the widespread disagreement with experts about global warming; rather, "individuals try to fit their interpretations of scientific evidence into pre-formed cultural philosophies." Similarly, Enos (2012) concludes from survey data that political advertisements are more likely to tap into voters' pre-existing opinions than they are to change opinions.

Whatever the source of confirmatory bias—upbringing, early life experience, wishful thinking, etc.—its existence is well established. Taking the presence of the bias as given, this paper demonstrates a fundamental problem with greater availability of information: it becomes easier to obtain confirmation of pre-existing beliefs. The Internet in particular facilitates finding others with similar beliefs, and this effect can be most dramatic for those with minority beliefs.

Consider "birthers," who subscribe to the belief that Barack Obama is not a natural-born citizen of the United States. This is the kind of belief that might not be expected to gain any traction if not for the Internet: if the believers are isolated from each other, the belief can easily die out. In the context of the model above, there would simply have to be a minority of agents with the belief, sufficiently strong confirmatory bias, and sufficient information transmission. The minority belief could then coexist indefinitely with the majority. In this and other cases, when there is divergence in a setting in which information is easy to transmit, we could take the divergence to be an indication of a high degree of confirmatory bias.

Several complications can be introduced into the model for the sake of greater realism. Some of these would clearly not change the qualitative nature of the results. For example, the reception of signals may be noisy, so that with some probability an agent misperceives signals that agree with the agent's prior. In an overlapping generations model, agents would receive their initial beliefs from the preceding generation, creating some momentum in the system.

The degree of bias could be endogenous, as in Rabin and Schrag (1999): the agent starts with no prior, or places equal weight on both states of the world, and is more likely to perceive any given signal as agreeing with what has been perceived in the past. Agents could be heterogeneous with respect to the degree of bias, with perhaps some agents that are not biased at all.

It would also be realistic to assume that not all agents send a signal, i.e. that some form an opinion but do not broadcast it. Another possibility is that some agents remain undecided, or more generally that agents have varying degrees of certainty about their belief. Depending on how certainty is parameterized, it could be related to the degree of bias (less biased agents require a greater majority of signals to be convinced), or it could have effects similar to difficulty in information transmission (agents perceive a benefit of certainty, which they weigh against the cost of obtaining more information).

Another complication with straightforward implications is that false information can be disseminated in order to sway those predisposed to believe it. Oreskes and Conway (2010) argue that interested parties have spread disinformation regarding the dangers of smoking and second-hand smoke and anthropogenic threats to the environment including acid rain and global warming. Confirmatory bias is one contributor to the success of a disinformation campaign.

In a similar vein, agents could conceivably place more weight on sources of information that are perceived to be more credible: for example, news broadcasts vs. the opinions of friends. Regardless of whether such sources are better in any objective sense, prominent opinions can act as coordinating devices. According to Rochman (2011), 24% of Americans place some trust in celebrities. Model and actress Jenny McCarthy has had considerable impact on public opinion about the effects of vaccinations, in spite of a lack of any particular expertise.

This raises another issue: how to change public opinion in the face of confirmatory bias. In the context of the model, disrupting a steady state could result from either changing agents' bias, or introducing more information. The former would depend on where the bias originates and would perhaps not be a promising possibility. It is a more straightforward proposition to introduce enough new information to change agents' opinions, especially if some sources of information are thought to be more credible.

This would be a similarity to much of the cascade literature, in which a cascade can be dislodged when new information becomes available. Clearly the difficulty of accomplishing this is increasing in the proportion of agents to be convinced as well as the magnitude of confirmatory bias. This is arguably what has happened for the issues discussed in Oreskes and Conway (2010): public opinion—about whether smoking causes cancer, for example—is initially very divided but reaches consensus as enough information becomes available that it becomes too difficult to deny a common conclusion.

The primary purpose of this paper is not to offer a solution to this kind of divergence of opinion, but to make clear the nature of the problem. In particular, if confirmatory bias is a cause of disagreement, simply making information easier to obtain, without somehow increasing the credibility of the information, may not help. If there is no feasible way to reduce or eliminate the bias, then social and political institutions may have to operate in the face of persistent disagreement.

References

Andreoni, J., and T. Mylovanov (2012). Diverging opinions. American Economic Journal: Microeconomics, 4, 209-232.

Banerjee, A. (1992). A simple model of herd behavior. Quarterly Journal of Economics, 107, 797-817.

Banerjee, A., and D. Fudenberg (2004). Word-of-mouth learning. Games and Economic Behavior, 46, 1-22.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy, 100, 992-1026.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1998). Learning from the behavior of others: conformity, fads, and informational cascades. Journal of Economic Perspectives, 12, 151-170.

Callander, S., and J. Hörner (2009). The wisdom of the minority. Journal of Economic Theory, 144, 1421-1439.

Cao, H., B. Han, and D. Hirshleifer (2011). Taking the road less traveled by: does conversation eradicate pernicious cascades? Journal of Economic Theory, 146, 1418-1436.

DeMarzo, P., D. Vayanos, and J. Zwiebel (2003). Persuasion bias, social influence, and unidimensional opinions. Quarterly Journal of Economics, 118, 909–68. Economist (2012). Group dynamics. May 30.

Ellison, G., and D. Fudenberg (1993). Rules of thumb for social learning. Journal of Political Economy, 101, 612-643.

Ellison, G., and D. Fudenberg (1995). Word-of-mouth communication and social learning. Quarterly Journal of Economics, 110, 93-125.
Enos, R. (2012). Romney relies on race to attack Obama on welfare. Model Politics, September 12.
Eyster, E., and M. Rabin (2010). Naïve herding in rich-information settings. American Economic Journal: Microeconomics, 2, 221–243.
Jones, M., and R. Sugden (2001). Positive confirmation bias in the acquisition of information. Theory and Decision, 50, 59–99.
Kataria, M. (2012). The role of preferences in disagreements over scientific hypothesis: evidence on cognitive bias in formation of beliefs. Journal of Socio-Economics, 41, 364-369.
Oreskes, N., and E. Conway (2010). Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming. New York: Bloomsbury Press.
Rabin, M., and J. Schrag (1999). First impressions matter: a model of confirmatory bias.Quarterly Journal of Economics, 114, 37-82.
Rochman, B. (2011). Jenny McCarthy, vaccine expert? A quarter of parents trust celebrities. Time, April 26.
Smith, L., and P. Sorensen (2000). Pathological outcomes of observational learning. Econometrica, 68, 371–98.
Wakefield, A.J., S.H. Murch, A. Anthony, J. Linnell, D.M. Casson, M. Malik, M. Berelowitz, A.P. Dhillon, M.A. Thomson, P. Harvey, A. Valentine, S.E. Davies, J.A. Walker-Smith (1998). Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. The Lancet, 351, 637–41. (Retracted)

Appendix

Proofs of Propositions

Proof of Proposition 2

In a steady state, there are potentially two coalitions: a minority that believes $\Phi = 0$, and a majority that believes $\Phi = 1$ (given the assumption that $S1 \geq S0$). Agents in the minority coalition must misperceive enough signals to believe that they are actually in the majority. The majority coalition consists of all *S1* agents with the private signal that $\Phi = 1$ as well as some of the remaining agents, each of whom does not misperceive enough signals to hold the belief that $\Phi = 0$. The probability that the population will arrive at steady state (*m*, *n*) is then

$$P(m,n) = \left\{ \sum_{i=\left\lfloor \frac{n-m}{2} \right\rfloor + 1}^{n} \binom{n}{i} \alpha^i (1-\alpha)^{n-i} \right\}^m \left\{ \sum_{j=0}^{\left\lfloor \frac{n-m}{2} \right\rfloor} \binom{\left\lfloor \frac{n-m}{2} \right\rfloor}{j} \alpha^j (1-\alpha)^{n-j} \right\}^{S_0 - m}$$

for $m \leq S0$. Let $P(C)$ be the probability of convergence. For a convergent steady state (in which $m = 0$), the expression above simplifies to the following:

$$P(C) = \left\{ \sum_{i=0}^{\lfloor n/2 \rfloor} \binom{\lfloor n/2 \rfloor}{i} \alpha^i (1-\alpha)^{n-i} \right\}^{S_0}$$

The effect of $S0$ on $P(C)$ is clear from the expression. The derivative of $P(C)$ with respect to α is

$$\frac{\partial P(C)}{\partial \alpha} = \sum_{i=0}^{\lfloor n/2 \rfloor} \binom{\lfloor n/2 \rfloor}{i} \left[\alpha^i (1-\alpha)^{n-i} \left\{ \frac{i}{\alpha} + \frac{i-n}{1-\alpha} \right\} \right]$$

Demonstrating that this is negative is purely an algebraic matter. To take an overview, first note that, when $\alpha > .5$, every term in the sum above is negative. For smaller α, some of these terms are positive, but the first few (or at least the first one) are negative. Note also that the smaller α is, the larger the magnitude of these negative terms. In every case, the negative values dominate, making the sum negative.

Proof of Proposition 3

Let $Xi = \alpha^i (1-\alpha)^{n-i}$ and consider its derivative with respect to n:

$$\frac{\partial X_i}{\partial n} = \left(\frac{\alpha}{1-\alpha} \right)^i (1-\alpha)^n \ln(1-\alpha) \leq 0$$

Note that this expression is equal to zero if and only if $\alpha = 0$ and decreases without bound as α approaches 1. I.e. the probability that an agent misperceives exactly n public signals is generally decreasing in n, and the magnitude of this rate of change is increasing in α. Taking $P(C)$ itself, it is not possible to differentiate because of its discrete nature. However, we may identify two effects on $P(C)$ of increases in n, in addition to the effect on Xi: there are more terms in the sum, and the combinatorial factor is larger. When α is sufficiently large, the derivative of Xi is large and negative, dominating the other two effects. When α is sufficiently small, the derivative of Xi is close to zero, and the other two effects dominate. The proposition is formally proven using the Gamma function, the continuous analog of the factorial: Given the continuity of $\Gamma(n)$, there is continuity of the derivative of $P(C)$ over α.

A complication in reasoning from the continuous to the discrete case is that Gamma has no mass at zero, whereas the combinatorial term for $i = 0$ is strictly greater than zero, and is relatively large if α is small, creating a potential contradiction to the proposition. Sufficiently large N eliminates this issue.

Footnotes

1 An alternative way to model confirmatory bias is to assume agents simply ignore contradictory signals with some probability. This would not affect the results qualitatively.

2 When a cascade can be dislodged, it is generally true that all agents would like to switch to the new cascade if they were able to do so.