Information Sampling, Conformity and Collective Mistaken Beliefs

Jerker Denrell (Jerker.Denrell@wbs.ac.uk)
University of Warwick, United Kingdom

Gaël Le Mens (Gael.Le-Mens@upf.edu)
Universitat Pompeu Fabra, Spain

Abstract

Societies sometimes stick to the status quo instead of switching to superior technologies and institutions. Existing explanations often attribute this to a coordination failure due to payoff externalities: people may know that another alternative is superior but nobody has an incentive to switch unless many others do so. We show that a simple learning argument can provide an alternative explanation. When people learn about the alternatives from their own experiences but tend to adopt the behaviors of others, they will mistakenly learn to believe that a popular alternative is superior to a better, but unpopular alternative. Our model neither assumes that agents engage in motivated cognition nor that they transmit mistaken information to others. Rather, it emphasizes the role of a fundamental asymmetry in access to information about popular versus unpopular alternatives. Our model thus provides a novel, sampling-based, explanation of how conformity in behavior can lead to private acceptance.

Keywords: Social Change, Learning, Conformity, Popularity

Explaining why collective mistakes emerge and persist is central to understanding social change, immobility, and differences in welfare (e.g. Elster, 1978; North, 1990). In developing countries, people keep on using poor domestic hygiene practices even though simple changes would save many lives (Curtis, Cairncrosss & Yonli, 2000), and farmers fail to use fertilizer despite their potential for large increases in productivity (Duflo, Kremer & Robinson, 2009). In western countries, firms persist in making abundant use of temporary workforce despite the inefficiency of this arrangement (Pfeffer, 1998). How come large groups of people could persist in using inefficient technologies, practices or institutions when better courses of actions are available?

Prior literature has proposed two classes of explanations. The first perspective has demonstrated that such mistakes can occur when payoffs increase with the number of people taking the same action (Arthur, 1989; Elster, 1978; North, 1990). This type of explanation, which relies on network externalities, generally assumes that people know the values of the alternatives. The problem is that nobody has an incentive to switch unless many others do so. The second perspective proposes that collective mistakes can emerge because agents sometimes believe that an alternative that is in fact suboptimal is the best. Rather than emphasizing a coordination failure, such explanations rely on the fact that people may not be aware of the qualities of the alternatives. This perspective assumes that agents use popularity as a signal of quality (Banerjee, 1992; Bikhchandani, Hirshleifer & Welch, 1992). Explanations that fall in this tradition are thus only valid when agents infer the qualities of the alternatives of the basis of the choices of others.

Here, we show that collective mistakes can still emerge and persist even if payoffs do not depend on the choices of others and people do not use popularity as a signal of quality. Instead of assuming that popularity directly affects an agent's evaluation of the available alternatives, we analyze situations where popularity only affects agents' sampling decisions: In our model, agents are more likely to try the most popular alternative, i.e., the alternative believed by most to be the superior. But agents' quality estimates are solely based on their own experiences with the alternatives. There are many situations where one might expect to see such conformity in behavior (people choose the popular alternative) but not in attitudes (their attitudes depend only on their own experience). For example, people may decide to go along with the majority and select the most popular alternative to avoid being seen as deviant (Cialdini & Goldstein, 2004), or because of adverse reputation effects to receiving a poor outcome with an unusual alternative (Keynes, 1936; Scharfstein & Stein, 1990). For example, it is difficult to avoid learning about the research tradition that is dominant in your department.

We show that when agents are more likely to be exposed to popular alternatives, a sampling bias emerges that leads most people to believe that the quality of popular alternatives is superior to that of unpopular alternatives. The intuition behind this result pertains to how people sample information about the available alternatives (Denrell & Le Mens, 2007; Le Mens & Denrell, 2011). Consider a medical doctor who has to select one of two possible treatments to cure a patient. Treatment P is popular among other doctors in her reference group and patients whereas treatment R is rarely chosen. Suppose the doctor selects R and the initial outcome is disappointing. While this might be a signal of the poor quality of R, it could also have resulted from other causes. But because patients may not want to continue a treatment with an unpopular drug with disappointing initial results, the doctor is likely to abandon R. In doing so, she will fail to discover that R might be an efficacious treatment. Compare this to what would have happened, had she selected P instead. She might have continued with P, even following disappointing initial results, because patients have heard that P has been efficacious in the past. In doing so, she would have acquired additional information about P and she might have discovered that this treatment was in fact efficacious.

This stylized story illustrates an important asymmetry in opportunities for error correction: an error of underestimation of the efficacy of a popular treatment is less likely to persist than an error of underestimation the efficacy of an unpopular

treatment.

Prior research has identified a number of psychological mechanisms and biases that explain why people might come to prefer what is available or most popular even when it is not the best alternative (Bem, 1972; Festinger, 1957; Heider, 1958). These explanations assume that people adjust their evaluations of the alternatives based on what is seen as a norm. Here, we analyze a simple learning model that demonstrates that it is not necessary to invoke such intra-psychic adjustment. Our model also has novel policy implications, as it suggests that exposing initially skeptical adopters to a new practice may be sufficient to enhance its diffusion.

Model

In our model, agents make a sequence of choices among two uncertain alternatives and learn about the qualities of the alternatives from their own experiences. Our model is an extension of standard models of the evolution of coordination (Arthur, 1989). Consider a growing population. In each period, $t = 1, 2, \ldots$, one new agent enters the population. Each agent makes a choice, in every period, between two competing alternatives. Everything else equal, agents prefer to select the alternative that is chosen by the largest proportion of others in the previous period. People might care about popularity because of payoff externalities (i.e. when an alternative is more useful if widely spread) or because of adverse reputation effects to receiving a poor outcome with an unusual alternative (Keynes, 1936; Scharfstein & Stein, 1990). An agent may forego some of the benefits of coordination, however, and choose the less popular alternative, if she believes it is of a higher quality. These assumptions are implemented in the model by the use of the following logistic choice rule. The the probability that Agent i selects Alt. 1 in period $t+1$ is:

$$pC_1^i(\hat{q}_{1,t}^i, \hat{q}_{2,t}^i, p_{1,t}) = \frac{e^{ap_{1,t}+b\hat{q}_{1,t}^i}}{e^{ap_{1,t}+b\hat{q}_{1,t}^i} + e^{a(1-p_{1,t})+b\hat{q}_{2,t}^i}}, \quad (1)$$

where a and b are positive constants, $p_{k,t}$ is the share of the population choosing Alt. k in period t, and $\hat{q}_{k,t}^i$ is Agent i's estimate of the quality of Alt. k at the end of period t. The logistic choice rule is often used to model choice under uncertainty (Erev & Baron, 2005; Sutton & Barto, 2005) and provides good fit to experimental data on sequential choices (Denrell, 2005).

Whenever Agent i chooses Alt. k, she can observe some information about the quality of that alternative. The observation, $z_{k,t}^i$, is a random variable drawn from the underlying quality distribution, which is assumed to be normally distributed with expected value μ_k and common variance σ^2. We assume a simple updating rule: the revised estimate is a weighted average of the past estimate and the new observation (Busemeyer & Myung, 1992; Denrell, 2005),

$$\hat{q}_{k,t}^i = (1-\lambda)\hat{q}_{k,t-1}^i + \lambda z_{k,t}^i, \quad (2)$$

where λ is between 0 and 1. If i does not select Alt. k in period $t-1$, the estimate remains the same: $\hat{q}_{k,t}^i = \hat{q}_{k,t-1}^i$. Agent i's initial estimate, when she enters the population, is a random variable drawn from a normal distribution with mean zero and variance σ^2.

Note that this formulation assumes that agents only learn from personal experience: an agent only updates her estimate of the quality of the alternative she personally observes. Thus, in this model, agents do not infer the qualities of the alternatives based on the choices of others, nor do they learn from the observations of others.

Analysis

It is possible to do a formal analysis of estimates and choices when the number of period becomes large. In the Appendix, we derive an explicit expression for the joint distribution of the asymptotic quality estimates ($\hat{Q}_1$,$\hat{Q}_2$) of the two alternatives (eq. 10). We then solve, numerically, for the limiting values of p_1, the proportion of agents choosing Alt. 1. Finally, for any given value of p_1, we compute, by numerical integration, the probability that Alt. 1 is considered superior.

The predictions of the asymptotic analysis are depicted by the solid lines on the graphs of Fig. 1; the diamonds represent simulated estimates after 500 periods. The left panel shows that when the weight of popularity in the sampling rule (parameter a, see eq. 1, is small, most agents select the first alternative, which has a higher average quality. If a is sufficiently large, however, it is possible that most agents select the second, inferior, alternative instead. The right panel shows that, in this case, most agents will also come to believe that the second alternative has a higher quality, i.e., $P(\hat{Q}_1 < \hat{Q}_2) > 0.5$. In summary, our model implies that, in all cases, most agents will come to believe that the alternative chosen by most has the higher quality, even if such belief is mistaken.

More generally, the expected asymptotic quality estimate of an alternative, for a randomly chosen agent, is an *increasing* function of the limiting proportion of agents who choose it:

$$E[\hat{Q}_k] = \mu_k - (1-p_k)\frac{b\lambda}{(2-\lambda)}\sigma^2. \quad (3)$$

This equation demonstrates that the choices of other agents create a systematic externality on the quality estimates of an agent, in spite of the fact that she learns only from her own experience. More precisely, the quality of a rarely chosen alternative is systematically underestimated, and the lower the number of agents who choose it, the more severe the underestimation. The negativity bias also occurs for the popular alternative but is of much lower magnitude. In particular, when the limiting proportion of agents who select the popular alternative is close to 1, there is almost no bias in the quality estimate of that alternative.

The probability that the process will converge to the inferior alternative, and that this alternative will be believed to have a higher quality, depends crucially on the difference in average qualities and on the variability of the observations. If the mean returns differ substantially, agents will quickly identify the best alternative. For example, suppose $\mu_1 = 1, \mu_2 = 0$

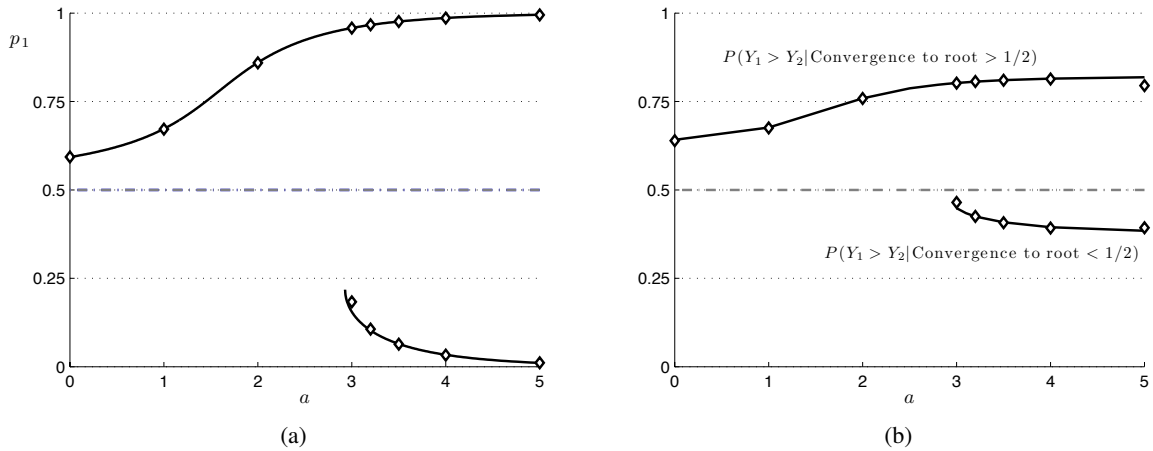

Figure 1: Simulated estimates (open diamonds) match the predictions of the asymptotic theory (solid line). The left panel displays the possible limiting values for the proportion of agents who select Alt. 1. These are the roots of the self-consistency equation (eq. 12). The right panel displays the limiting proportion of agents who believe that Alt. 1 is the best. When most agents choose Alt. 1 ($p_1 > 1/2$), most of them also believe that it is the best. And inversely, when most agents choose Alt. 2 ($p_1 < 1/2$), most agents believe that it is the best. Simulation results are averages based on 500 runs of 500 periods each (with $b = 1.5, \mu_1 = 0.25, \mu_2 = 0$ and $\sigma^2 = 1$).

and $\sigma^2 = 1$. It is then highly unlikely that agents will converge to Alt. 2 and believe it has higher quality. The presence of the variance of the observations in eq. 3 also shows that convergence on the inferior alternative is much more likely if alternatives are difficult to distinguish because observations are noisy. Noisy observations often happen when there is some delay between choices and observations of the corresponding outcomes which makes the association between actions and outcomes difficult. In those settings, the dynamics of experiential learning could be of particular importance.

To understand the intuition underlying these results, note that individuals accumulate biased samples of information about the qualities of the two alternatives. Negative estimates reduce the probability of further sampling, which implies that no further information is available and the quality estimate will not be updated (Denrell, 2005; Fazio, Eiser, & Shook, 2004). This implies, in turn, that the qualities of the alternatives will tend to be underestimated, as illustrated by eq. 3.

The magnitude of the information bias, however, is moderated by popularity. To see how, consider extreme cases. Suppose, for example, that Alt. 1 is much more popular than Alt. 2. In this case, decision makers will sample Alt. 1 almost no matter what their quality estimates are. There is thus almost no sampling bias for this alternative, the quality estimate is close to the true quality and there is almost no underestimation tendency for this alternative. When Alt. 1 is much less popular than Alt. 2, sampling will depend more strongly on quality estimates. The quality estimate will thus be subject to the systematic underestimation tendency described above.

It is important to note that our results do *not* require that the estimates of each individual converge toward a stable value. In fact, the magnitude of the sampling bias is strongest when $\lambda = 1$. In this case, agents' quality estimates correspond to their last observation of the alternative, and thus the estimates are subject to potentially large changes after each observation. But the population still converges to one of the two alternatives, and the estimates for this alternative become more positive than for the other alternative. More generally, simulations show that people will still tend to evaluate popular alternatives more positively under different assumptions regarding the estimate updating rule, such as when λ declines with the number of observations. For example, when the quality estimate of Alt. k is the average of all prior observations of that alternative, it can still happen that most people select the inferior alternative and mistakenly believe it to be the alternative of higher quality. But the synchronization of estimates and behavior occurs for values of σ that tend to be larger than when λ is constant. This is not surprising, because in that case, estimates integrate information better (when the environment is stable) than when λ is constant.[1]

Coordination and Synchronization of Estimates

We motivated our assumption that people are more likely to select popular alternatives by referring to settings where people want to conform to the majority. But another reason for wanting to select a popular alternative is the desire to coordinate one's behavior with others (e.g. Hardin, 1968). For example, people might want to go to the same venues as those in the same social group, or they might want to use the same computer platform as others so as to be able to exchange files

[1]This case is very close to Bayesian updating with a prior on μ_k that has a $N(\mu_k, \sigma)$ distribution.

more easily. It is possible to model such settings as follows: Suppose that there are N players who choose, in each period, one of M alternatives. The payoff of Alt. k follows a Gaussian distribution with mean μ_k and variance σ^2. That is, if $o^i_{k,t}$ denotes the payoff Player i receives from selecting Alt. k in period t, we have $o^i_{k,t} \sim N(\mu_k, \sigma)$. In addition, there is a coordination bonus: if i selects Alt. k in period t and $\eta^{-i}_{k,t}$ other players select Alt. k in that period, Player i receives a bonus of $c\eta^{-i}_{k,t}$ where c is a positive parameter. Suppose there are T periods and that the goal of each player is to maximize the total payoff she receives over the T periods.

These assumptions regarding the payoff structure and the goals of the decision makers specify a coordination game (e.g. Gibbons, 1992). In this setup, people have an incentive to select a popular alternative. This formulation defines a setting where people are likely to develop quality estimates consistent with the predictions of our theory without making assumptions about the estimate updating and the choice rules. This suggests a simple way to test our theory in the laboratory: make people play a coordination game such as the one just described, measure if a majority of players believe the alternative chosen by most people to be the superior one (even in cases where it is in fact the inferior alternative) and evaluate if the pattern of estimates can be well explained by an information bias in favor of the popular alternative.

Our assumptions regarding sampling and estimate updating define a heuristic that people can use to play this coordination game. But people can also adopt other heuristics. A widely studied strategy for playing coordination games is the best-reply strategy (e.g. Young, 1998). A decision maker uses a 'best-reply' strategy when she selects the alternative that has the highest subjective expected payoff, assuming that other players will choose the same alternative as they did in the prior period (choice is randomized if more than one alternative has maximal subjective expected payoff). We ran 10,000 simulations of the game, assuming that players use the best-reply strategy, with $N = 10$, $\mu_1 = 1$, $\mu_2 = 2$, $\sigma = 3$, $c = 0.1$ and $\lambda = 0.5$. Simulations show that the quality estimates of the players will tend to synchronize with each others, in a fashion similar to what happens in the model we have analyzed earlier in the paper. For example, after 20 periods, the correlation between the quality estimates of Player 1 for Alt. 1 and the sum of the estimates of Players 2 to 10 is 0.64. Furthermore, the players' choices sometimes coordinate on the inferior alternative (with the above parameters, the likelihood that 6 or more players prefer Alt. 1 is about 16%). This is not surprising because, if c is high enough, coordination on the inferior alternative is a Nash equilibrium.

Discussion and Conclusion

Our model illustrates a novel mechanism that explains why groups may be reluctant to switch to another practice, even when this other practice is superior. Our mechanism complements existing explanations that show how a concern for popularity can lead to lock-in. More precisely, our model demonstrates how a concern for popularity can also influence evaluations of the qualities of the alternatives even when people learn only from their own private observations. Our theory thus provides a simple mechanism for why public conformity in behavior (a tendency to choose popular alternatives) may lead to private acceptance (the belief that the popular alternative is the best) at the individual level and to collective illusions at the level of the group.

Existing psychological explanations of this synchronization of beliefs with behavior, such as cognitive dissonance theory, or self-perception theory, attribute it to motivated cognition (Bem, 1972; Festinger, 1957; Heider, 1958). Our explanation does not challenge the experimental evidence underlying those explanations. Rather, it suggests a complementary explanation that is likely to be important in realistic settings where popularity affects choices and access to payoff information. A distinctive feature of our analysis is that it assumes that people are good processors of information, but are naïve with respect to the nature of the sample they use to form their quality estimates. This is a standard assumption of the research program on sampling explanations of judgment biases that has received broad empirical support (Fiedler, 2012; Fiedler & Juslin, 2006; Juslin et al., 2007). We do not claim that people do not engage in motivated cognition. Instead, we point to a fundamental asymmetry regarding the information sample that people have about popular v.s. unpopular alternatives. Because cognition operates on the available sample of information, our sampling explanation operates at a different level of analysis and thus complements explanations that focus on cognitive processes.

How do our results relate to herding models (Banerjee, 1992; Bikhchandani et al., 1992), which also explain collective failures to identify the best alternative? These models also assume that people are good processors of information. But contrary to these models, we do not assume that agents use popularity as a signal of quality. Therefore, our model may be more suitable for contexts where the assumptions of herding models do not apply. In particular, our model fits contexts where people have different tastes or do not believe that others know best, but where their sampling behavior is still influenced by others (Sutton & Barto, 1998).

Our theory also differs from explanations of collective mistakes that attribute them to a coordination failure due to network externalities (e.g. Elster, 1978; North, 1990). In fact, our theory suggests that, if network externalities affect sampling behavior, the group may not only converge to the inferior alternative but, when this happens, most agents will also come to believe that the inferior alternative is of superior quality. Thus, in a vote about whether to switch to another alternative, most people would favor sticking to the status quo even if it is actually inferior. By contrast, explanations that rely on a coordination failure predict that people will switch if a vote could be organized and switching cost were low. Despite this difference in prediction, our theory complements explanations based on payoff externalities by suggesting that

payoff externalities can have a systematic effect on quality estimates that reinforces the possibility of a lock-in.

More generally, our results point to the systematic effect of unbiased experiences on beliefs. From a policy perspective, this illustrates the potential benefits of exposing initially skeptical adopters to an unpopular practice. Agents may appear to be resistant to unpopular practices not because they are risk averse or conservative, but because their own experiences with the unpopular practice are often skewed towards failures. In this case, inducing agents to try the unpopular practice again might help its acceptance, even when persuasive campaigns are not effective.

Appendix

Preliminaries

In this appendix, we analyze the asymptotic behavior of the model. Let $\hat{Q}_{1,t}$ (resp. $\hat{Q}_{2,t}$) be a random variable that refers to the quality estimate for Alt. 1 (resp. Alt. 2) of a randomly chosen agent, at the end of period t. Let $h_t(\hat{q}_{1,t},\hat{q}_{2,t})$ be the joint density of the quality estimates in period t. Let $P_{1,t}$ denote the proportion of the population choosing alternative 1 in period t. Capital letters denote random variables, and corresponding lower case letters denote realizations of the random variables.

The expected proportion of agents choosing Alt. 1 in period $t+1$ is:

$$E[P_{1,t+1}|p_{1,t}] = \iint\limits_{\hat{q}_{1,t},\hat{q}_{2,t}} pC_1(\hat{q}_{1,t},\hat{q}_{2,t},p_{1,t})h_t(\hat{q}_{1,t},\hat{q}_{2,t})d\hat{q}_{1,t}d\hat{q}_{2,t}, \quad (4)$$

One of the difficulties in analyzing the asymptotic behavior of equation 4 is that $P_{1,t}$ is a random variable, which varies from period to period. It is reasonable to suspect, however, that $P_{1,t}$ will converge to a constant, as $t \to \infty$ and the number of agents increases. Let p_1 denote this limiting proportion. The intuition is the same as for the law of large numbers: if the number of agents is very large, the proportion choosing Alt. 1 should converge to its expected value.

Another difficulty in solving the above equation is that $h_t(\hat{q}_{1,t},\hat{q}_{2,t})$ will change over time. Nevertheless, there is reason to believe that as $t \to \infty$, $h_t(\hat{q}_{1,t},\hat{q}_{2,t})$ will converge to a stationary distribution, denoted $h(\hat{q}_1,\hat{q}_2)$. As $t \to \infty$, more agents are added to the system and a larger proportion of agents will have had extensive opportunities to sample the two alternatives. While a new agent, with random estimates, is added to the system in each period, the influence of such agents should become vanishly small over time.

To calculate the equilibrium the model converges to, we adopt both of these simplifying assumptions. That is, we will assume, without a rigorous proof, that $P_{1,t}$ converges to a constant p_1 and that $h_t(\hat{q}_{1,t},\hat{q}_{2,t})$ converges to a stationary distribution $h(\hat{q}_1,\hat{q}_2)$. Under these assumptions, the following self-consistency equation must hold for p_1:

$$p_1 = \iint\limits_{\hat{q}_1,\hat{q}_2} pC_1(\hat{q}_1,\hat{q}_2,p_1)h(\hat{q}_1,\hat{q}_2)d\hat{q}_1d\hat{q}_2. \quad (5)$$

Moreover, for any given value of p_1, the stationary joint density of the quality estimates, $h(\hat{q}_1,\hat{q}_2)$, must satisfy

$$h(\hat{q}_1,\hat{q}_2) = \int_r h(r,\hat{q}_2)pC_1(r,\hat{q}_2,p_1)\tau_1(r,\hat{q}_1)dr + \int_k h(\hat{q}_1,k)pC_2(\hat{q}_1,k,p_1)\tau_2(k,\hat{q}_2)dk. \quad (6)$$

Here, $pC_2(\hat{q}_1,k,p_1)$ is the probability that an agent with quality estimates $\hat{q}_1$ and k will choose Alt. 2 and $\tau_k(r,\hat{q}_k)$ is the probability mass that the quality estimate for Alt. k transitions from r to $\hat{q}_k$ given that the agent samples Alt. k. Because the new estimate equals $\hat{q}_k = (1-\lambda)r + \lambda z_k$, where z_k is normally distributed with mean μ_k and variance σ^2, $\tau_k(r,\hat{q}_k)$ equals the probability mass that z_k is equal to $(\hat{q}_k - (1-\lambda)r)/\lambda$.

To explain the above equation (eq. 6), note that the terms to the right add up to the probability that the quality estimates for alternatives 1 and 2 are $\hat{q}_1$ and $\hat{q}_2$, after an agent has sampled one of the alternatives. The first term on the right hand side is the probability that the quality estimates for alternatives 1 and 2 are $\hat{q}_1$ and $\hat{q}_2$, and that the agent sampled Alt. 1 in the previous period. This set of estimates can only emerge, after the agent samples Alt. 1 , if this agent's estimate of the quality of Alt. 2 was equal to $\hat{q}_2$. Similarly, the second term on the right hand side is the probability that the quality estimates for alternatives 1 and 2 are $\hat{q}_1$ and $\hat{q}_2$ and that the agent sampled Alt. 2 in the previous period.

Below, we show how one can solve for $h(\hat{q}_1,\hat{q}_2)$, for any value of p_1. Using $h(\hat{q}_1,\hat{q}_2)$, we can then solve for the equilibrium value of p_1. Finally, we derive the expected quality estimates in the stationary state.

The Stationary Distribution of the quality estimates

For a given value of p_1, the quality estimates for a representative agent follow a discrete time markov process, with a general state space $\mathbb{R}\times\mathbb{R}$. Because there is a positive probability, in any period, that the system could transition from any state to another, the markov process has a unique stationary distribution, which has to satisfy equation 6. The unique joint density that satisfies this equation is

$$h(\hat{q}_1,\hat{q}_2) = Kg_1(\hat{q}_1)g_2(\hat{q}_2)[e^{-ap_1-b\hat{q}_1} + e^{-a(1-p_1)-b\hat{q}_2}], \quad (7)$$

where K is a normalizing constant, i.e.,

$$1/K = \iint\limits_{\hat{q}_1,\hat{q}_2} g_1(\hat{q}_1)g_2(\hat{q}_2)[e^{-ap_1-b\hat{q}_1} + e^{-a(1-p_1)-b\hat{q}_2}]d\hat{q}_1d\hat{q}_2, \quad (8)$$

and $g_k(y) = \int_r g_k(r)\tau_k(r,y)dr$, i.e., $g_k(\cdot)$ is the distribution of the random variable the estimate of Alt. k would converge to if the probability that Alt. k is selected were equal to 1 in every period. When the quality distribution of Alt. k is normally distributed with mean μ_k and variance σ^2, it can be shown that $g_k(\cdot)$ is a normal density with mean μ_k and variance $\sigma^2\lambda/(2-\lambda)$.

Using appropriate algebraic manipulations, it can be easily verified that the joint density in eq. 7 satisfies the stability equation 6. The explicit formula for the normalizing constant is:

$$1/K = e^{\frac{b^2\sigma^2\lambda}{2(2-\lambda)}}\left(e^{-ap_1}e^{-b\mu_1} + e^{-a(1-p_1)}e^{-b\mu_2}\right). \quad (9)$$

The stationary joint density of the estimates is

$$h(\hat{q}_1,\hat{q}_2) = \frac{\frac{1}{2\pi\frac{\lambda}{2-\lambda}\sigma^2}e^{\frac{-(\hat{q}_1-\mu_1)^2}{2\frac{\lambda}{2-\lambda}\sigma^2}}e^{\frac{-(\hat{q}_2-\mu_2)^2}{2\frac{\lambda}{2-\lambda}\sigma^2}}[e^{-ap_1-b\hat{q}_1} + e^{-a(1-p_1)-b\hat{q}_2}]}{e^{\frac{b^2\sigma^2\lambda}{2(2-\lambda)}}\left(e^{-ap_1}e^{-b\mu_1} + e^{-a(1-p_1)}e^{-b\mu_2}\right)}. \quad (10)$$

The probability that an agent will consider Alt. 1 to have a quality higher than Alt. 2 is thus

$$P(\hat{Q}_1 > \hat{Q}_2) = \int\limits_{\hat{q}_1=-\infty}^{+\infty}\left(\int\limits_{\hat{q}_2=-\infty}^{\hat{q}_1} h(\hat{q}_1,\hat{q}_2)d\hat{q}_2\right)d\hat{q}_1, \quad (11)$$

which can be computed by numerical integration.

The equilibrium value of p_1

In the stationary state, the probability that Alt. 1 is selected is given by equation 5. The appropriate substitutions and algebraic manipulations imply

$$p_1 = \frac{1}{1+e^{-2a(p_1-0.5)}e^{-b(\mu_1-\mu_2)}}. \tag{12}$$

The values that p_1 could converge to are the stable roots of the self-consistency equation (eq. 5), which can be found numerically for any value of a by using equation 12. When the value of a is sufficiently low, such as when $a = 1$, there is only one root and this is above 0.5. However, if a is sufficiently large, such as when $a = 3.5$, the equation has three roots, one close to zero, one close to one, and another one, between the two other roots. The intermediary root is unstable, however. That is, the derivative of right-hand-side of the stationarity equation, evaluated at the intermediary root, is higher than 1. As a result, any small disturbance will tend to move the system away from the intermediary root.

The expected quality estimates

The marginal distribution of $\hat{q}_1$, given p_1, is given by integration of the joint distribution (eq. 10) over $\hat{q}_2$. After simplifications, we get:

$$E[\hat{Q}_1] = \mu_1 - (1-p_1)\frac{b\lambda}{(2-\lambda)}\sigma^2, \tag{13}$$

which is an increasing function of p_1. A similar calculation gives $E[\hat{Q}_2] = \mu_2 - p_1\frac{b\lambda}{(2-\lambda)}\sigma^2$, which is a decreasing function of p_1. Moreover,

$$E[\hat{Q}_1] - E[\hat{Q}_2] = \mu_1 - \mu_2 + (2p_1-1)\frac{b\lambda}{(2-\lambda)}\sigma^2, \tag{14}$$

which is an increasing function of p_1. Thus, $E[\hat{Q}_2] > E[\hat{Q}_1]$ even if $\mu_2 < \mu_1$ whenever

$$(1-2p_1)\frac{b\lambda}{(2-\lambda)}\sigma^2 > \mu_1 - \mu_2. \tag{15}$$

The maximum difference is obtained when $p_1 = 0$, and

$$\frac{\mu_1-\mu_2}{\sigma^2} < \frac{b\lambda}{(2-\lambda)}. \tag{16}$$

Acknowledgements

Gaël Le Mens was supported by Spanish Ministry of Science and Innovation Grant # ECO2010-17145, and the Barcelona School of Management.

References

Arthur, W. B. (1989). Competing Technologies, Increasing Returns, and Lock-In by Historical Events, *Economic Journal* 97:642-65.

Banerjee, A. V. (1992). A Simple Model of Herd Behavior. *Quarterly Journal of Economics* 107:797-817.

Bem, D. J. (1972). Self-perception theory. In *Advances in experimental social psychology* (Academic Press, New York, NY)., Vol 6 pp 1-62.

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

Busemeyer, J.R., Myung, I. J. (1992). An adaptive approach to human decision making: Learning theory decision theory and human performance. *Journal of Experimental Psychology: General* 121:177-194.

Cialdini, R. B., Goldstein, N.J., Social Influence: Conformity and Compliance. *Annual Review of Psychology* 55:591-621.

Curtis. V., Cairncross, S., Yonli, R. (2000). Domestic hygiene and diarrhoea – pinpointing the problem. *Tropical Medicine and International Health*, 5(1):22-32.

Denrell, J. (2005). Why most people disapprove of me: Experience sampling in impression formation. *Psychological Review* 112:951-978.

Denrell, J., Le Mens, G. (2007). Interdependent Sampling and Social Influence. *Psychological Review* 114 (2).: 398-422.

Duflo, E., Kremer, M., & Robinson, J. (2009). Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya. *NBER Working Paper No. 15131* .

Elster J (1978). *Logic and Society: Contradictions and Possible Worlds* (Wiley, New York, NY).

Erev, I., Barron, G. (2005). On adaptation maximization and reinforcement learning among cognitive strategies. *Psychological Review* 112:912-931.

Fazio, R. H., Eiser, J. R., & Shook, N. J. (2004). Attitude Formation Through Exploration: Valence Asymmetries. *Journal of Personality and Social Psychology*, 87(3), 293–311.

Festinger, L. (1957). *A theory of cognitive dissonance* (Stanford Univ Press, Palo Alto, CA).

Gibbons, R. (1992). *Game theory for applied economists.* Princeton University Press.

Hardin, G. (1968). The Tragedy of the Commons. *Science*, 162(5364), 1243-1248.

Heider, F. (1958). *The psychology of interpersonal relations* (Wiley, New York, NY).

Keynes, J. M. (1936). *General Theory of Employment Interest and Money* (Macmillan, London, UK).

Le Mens, G., Denrell, J. (2011). Rational learning and information sampling: On the naïvety assumption in sampling explanations of judgment biases. *Psychological Review*, 118:379–392.

Munshi, K. (2004). Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution. *Journal of Development Economics*, 73:185-213.

North, D. (1990). *Institutions, institutional change, and economic performance* (Cambridge Univ Press, Cambridge, UK).

Pfeffer, J. (1998). *The Human Equation* (Harvard Business School Press, Cambridge MA).

Scharfstein, D. S., Stein, J. C. (1990). Herd behavior and investment. *The American Economic Review* 80(3):465-479.

Sutton, R., Barto, A. G. (1998). *Reinforcement learning* (The MIT Press, Cambridge, MA).

Young, H. P. (1998). *Individual Strategy and Social Structure.* (Princeton University Press, Princeton).