Adaptive Sampling Policies Imply Biased Beliefs: A Generalization of the Hot Stove Effect

Jerker Denrell (jerker.denrell@wbs.ac.uk)
Warwick Business School, University of Warwick,
Scarman Rd, Coventry CV4 7AL, West Midlands, UK

Abstract

The Hot Stove Effect is a negativity bias resulting from the adaptive character of learning. The mechanism is that learning algorithms that pursue alternatives with positive estimated values, but avoid alternatives with negative estimated values, will correct errors of overestimation but fail to correct errors of underestimation. Here we generalize the theory behind the Hot Stove Effect to settings in which negative estimates do not necessarily lead to avoidance but to a smaller sample size (i.e, a learner selects fewer of alternative B if B is believed to be inferior but does not entirely avoid B). We demonstrate formally that the negativity bias remains in this set-up. We also show that there is a negativity bias for Bayesian learners in the sense that most such learners underestimate the expected value of an alternative.

Keywords: learning; stopping; sampling; Bayesian models

Introduction

Learning from experience does not necessarily generate unbiased beliefs, partly due to psychological biases but also due to biases in the information learners sample and get exposed to. One important bias in sampling is the so called "hot stove effect" (Denrell and March, 2001) which refers to the asymmetry in error correction generated by adaptive learning processes. The key idea is that that the tendency to avoid alternatives with unfavourable past outcomes generates a biased set of experiences. Alternatives that are underestimated - believed to be worse than what they are – are unlikely to be tried and sampled again which implies that errors of underestimation are unlikely to be corrected. Alternatives that are overestimated - believed to be better than what they are – are likely to be tried and sampled again which implies that errors of overestimation are likely to be corrected. This asymmetry in error correction generates a biased set of experiences which in turn can give rise to biased judgments (Denrell, 2005), including ingroup bias and apparent risk averse behavior (Denrell, 2005, 2007).

There is good experimental support for the hot-stove effect at the individual level and researchers in psychology have relied on the hot stove effect to explain regularities in risk taking in experimental studies (Erev and Roth, 2014) and why people underestimate the trustworthiness of others (Fetchenhauer and Dunning, 2010). Researchers in finance (Dittmar and Duchin, 2016) have used field data to demonstrate that the hot stove effect can explain risk taking behavior by executives. The hot stove effect also has important implications for information aggregation and online reviews: if consumers avoid products with poor reviews, and consumers review products they buy, negative reviews will be more persistent than positive reviews, generating biased averages (Le Mens et al, 2018).

Past theoretical work on the hot stove effect have assumed that negative experiences may lead to avoidance, i.e., that the alternative is not tried at all (Denrell, 2005; 2007). Clearly, if no more information is available, a negative impression will persist. In many settings, however, a negative belief or impression may not lead to avoidance of the alternative but merely to a smaller sample size. An animal who has a more favourable impression of the energy content from plant of type A than from plant of type B, may search for plants of type A. During this search for plants of type A, some plants of type B may be incidentally be found. The result is that the animal samples more plants of type A (because the search is focused on such plants) than of type B, but the animal does not avoid plants of type B but simply samples fewer of them. Similarly, a firm may prefer to hire graduates from university A, but may nevertheless hire some, although fewer, graduates from university B if there are not enough graduates from A that accepts its offers.

In this paper I generalize the theory behind the hot stove effect and show that it holds even if a negative impression only leads to a reduction in the sample size, not necessarily to avoidance. Specifically, I show that for a broad class of learning algorithms in which the sample size is a function of the past belief, the final belief will be biased. If the sample size is higher if the past belief was more positive, there is a negativity bias: the final belief will be lower than the expected value of the random variable the learner is learning about. This result also applies to taking averages: the average of a sample will be biased if the total sample size is a function of the average based on an initial subset.

I also examine if the bias remains for a Bayesian learner. I show that there is no bias on average for a Bayesian learner: the average belief will be equal to the expected value of the random variable the learner is learning about. However, I also show that a majority of Bayesian learner will underestimate the variable they are learning about if they increase the sample size as a result of initial positive beliefs.

These results imply that a large class of sensible and adaptive learning processes can be expected to generate biased beliefs, even if decision-makers process the available information in a seemingly unbiased way (i.e. taking averages). Indeed, even rational Bayesian learners will tend to underes-

timate the expected value of an alternative (i.e., most will do so) if the total sample size is higher when the initially observed payoffs are high. Adaptive sampling processes are common in reality and are often necessary to reduce search costs. It is not sensible, for example, to continue to sample an alternative a fixed number of times if initial trials reveal that this alternative has much lower payoff than other available alternatives. The results in this paper show that even unbiased processing of information generated by such sampling policies can generate seemingly biased beliefs. Adaptive sampling policies thus offer an alternative explanation of biases in beliefs, such as a tendency to underestimate the extent to which others are trustworthy.

Illustration

To illustrate the basic ideas, we consider a simple two period set-up. In period one, a learner samples an alternative k times and observes the payoff generated. That is, the learner observes k payoffs, $x_{1,1}, x_{1,2}, ..., x_{1,k}$, each independently drawn from the payoff distribution $f(x)$, where $f(x)$ is assumed to be a normal distribution with mean zero and variance σ^2. Based on the observed payoff, the learner computes the average observed payoff after the first period: $\bar{x}_1 = (1/k)\sum_{j=1}^{k} x_{1,j}$.

In the second period, the learner takes an additional sample and observes m payoffs, $x_{2,1}, x_{2,2}, ..., x_{2,m}$, each independently drawn from the payoff distribution $f(x)$. We assume that the size of this sample, m, is a function of $\bar{x}_1$. For example, the learner may take a larger sample if the observed average first period payoff is high ($\bar{x}_1$ is high), than if the observed average first period payoff is low ($\bar{x}_1$ is low), because the alternative is believed to be more rewarding when the observed average first period payoff is high compared to when it is low. To illustrate the impact of such an adaptive sample size policy, suppose the sample size in period two is equal to $m = h(high)$ whenever $\bar{x}_1 > c$ and equal to $m = l(low)$ whenever $\bar{x}_1 \leq c$.

After the second period the learner computes the average of all the payoffs observed in period one and two: $\bar{x}_2 = (1/(k+m))[\sum_{j=1}^{k} x_{1,j} + \sum_{j=1}^{m} x_{2,j}]$. We are interested in whether this average is unbiased or not.

The answer is that this average will be biased. To illustrate this, suppose the learner samples two payoffs in the first period ($k = 2$), samples ten more if the first period average is positive ($h = 10$) but only samples one more if the first period average is negative ($l = 1$). The average of all observed payoffs after the second period will then be negative. When σ, the standard deviation of the payoff, equals 1, then $E[\bar{x}_2] = -0.141$ and the proportion of averages that are negative is 0.587. When $\sigma = 5$, $E[\bar{x}_2] = -0.705$ and the proportion of averages that are negative is 0.583. More generally,

$$E[\bar{x}_2] = -\frac{\sigma\sqrt{k}(h-l)}{\sqrt{2\pi}(k+l)(k+h)} e^{-c^2 k/2\sigma^2}, \qquad (1)$$

(see the appendix for a derivation of this equation). This equation shows that the bias depends on how the sample size varies with $\bar{x}_1$. If the learner takes a larger sample when $\bar{x}_1$ is high compared to when $\bar{x}_1$ is low ($h > l$), there is a negative bias: $E[\bar{x}_2] < 0$. If the learner instead takes a larger sample when $\bar{x}_1$ is low compared to when $\bar{x}_1$ is high ($h < l$), there is a positive bias: $E[\bar{x}_2] > 0$. Equation (1) also implies that the bias is larger when the payoffs are more variable, i.e., when σ^2 is larger. In summary, an average based on an adaptive sampling policy (adaptive in the sense that the sample size changes as a result of the initially observed average) will be biased and the bias depends on the type of the sampling policy (increasing or decreasing in the initial average). Because learners often do regulate sample sizes in response to feedback from initial samples, this implies that learning processes quite generally leads to biased beliefs, at least if the beliefs are based on the average payoffs observed. Note that this occurs even if the decision-maker does not process information in a biased way. The decision-maker simply take the average of the observed payoffs. This averaging process, which is unbiased when the sample size is fixed, becomes biased when the total sample size depends on the initially observed average.

Intuition

To understand the reason for the bias, note that the sum of all observations after the second period is the sum of two components: the sum of all observations after the first period ($\sum_{j=1}^{k} x_{1,j}$) and the sum of all observations in the second period ($\sum_{j=1}^{m} x_{2,j}$.). The sample size taken in the second period (m) will impact the relative weight of these two components. The reason for the negative bias, when the learner takes a larger sample when $\bar{x}_1$ is high compared to when $\bar{x}_1$ is low ($h > l$), is that the first component will be weighted relatively more when $\bar{x}_1$ is low which implies that the sample size taken in period two is low ($m = l$).

To illustrate this, suppose the average after two observations in the first period ($k = 2$) is $\bar{x}_1 = 1$. The sum in the first two periods was then equal to $k\bar{x}_1 = 2$. Suppose $c = 0$, implying that the learner will take a large sample ($m = h$), whenever $\bar{x}_1 > 0$. Suppose $h = 10$, i.e., the sample size is ten in period two if the first period average was positive. The average payoff observed in both periods is then

$$\bar{x}_2 = \frac{2 + \sum_{j=1}^{10} x_{2,j}}{2+10} = \frac{2}{2+10} + \frac{\sum_{j=1}^{10} x_{i,j}}{2+10}.$$

The expected value of an observation in the second period is zero, i.e., $E[x_{2,j} = 0]$. It follows that

$$E[\bar{x}_2 | \bar{x}_2 = 1] = \frac{2}{2+10} = \frac{1}{6},$$

which is close to zero. A positive first period average payoff thus tends to regress to the mean of the distribution, which is zero. The reason is that a large sample is taken in the second period which implies that the first period average will not be weighted much.

Suppose, in contrast, that the average after two observations in the first period ($k = 2$) was $\bar{x}_1 = -1$. The sum in the

first two periods was then equal to $k\bar{x}_1 = -2$. When $c = 0$, the learner will take a small sample ($m = l$) in the second period because $\bar{x}_1 = -1 < c$. Suppose $l = 1$: only one sample is taken if the first period belief was negative. The average payoff observed in both periods is then

$$\bar{x}_2 = \frac{-2 + x_{2,1}}{2+1} = \frac{-2}{2+1} + \frac{x_{2,1}}{2+1}.$$

Again, the expected value of an observation in the second period is zero, i.e., $E[x_{2,1} = 0]$. It follows that

$$E[\bar{x}_2 | \bar{x}_2 = -1] = \frac{-2}{3},$$

which is further away from zero than $E[\bar{x}_2 | \bar{x}_2 = 1] = \frac{1}{6}$ is. A negative first period average payoff thus tends to regress less to the mean of the distribution than a positive first period average payoff does. The reason is that a small sample is taken in the second period if the first period average is negative. The negative first period average will thus be weighted more in the final average.

As this example shows, negative first period averages are more 'persistent' in the sense that they are given a larger weight than positive first period averages are when the learner samples more after a positive compared to a negative first period average. Because negative and positive first period averages of the same magnitude are equally likely, when the payoff distribution is normal with mean zero (this distribution is symmetric around zero), the greater persistence of negative first period averages explains the overall negative bias.

The impact of variance can also be understood in a similar way. If the payoff distribution is more variable, first period averages will tend to differ more from zero, both in the positive and negative directions. Positive first period averages will tend to regress more to the mean (which is zero) than negative averages do. When the negative first period averages are more extreme, because the variance is larger, the result is a stronger bias.

General Theorem about Biased Averages

The bias generated by an adaptive sampling size policy does not only hold for the normal distribution and the specific binary sampling policy considered above (high above zero, low below) but holds for any distribution and a large class of adaptive sampling policies.

Theorem 1: In period one, a learner observes k payoffs, $x_{1,1}, x_{1,2}, ..., x_{1,k}$, each independently drawn from the payoff distribution $f(x)$, with expected value $E[x] = u$. In period two the learner samples $n(\bar{x}_1)$ payoffs, each independently drawn from the payoff distribution $f(x)$. Here $n(\bar{x}_1) \geq 1$, and $n(\bar{x}_1)$ is a function of the first period average payoff: $\bar{x}_1 = (1/k)\sum_{j=1}^{k} x_{1,j}$. Let $\bar{x}_2$ be the average observed payoff during both periods one and two. Then i) $E[\bar{x}_2] < u$ whenever $n(\bar{x}_1)$ is a strictly increasing function of $\bar{x}_1$, and ii) $E[\bar{x}_2] > u$ whenever $n(\bar{x}_1)$ is a strictly decreasing function of $\bar{x}_1$.

Proof: See the Appendix.

Alternative Learning Models

So far we have assumed that the learner computes the average of the observed payoffs, but a similar bias holds for several other learning models. Suppose, for example, that the encounters objects sequentially and gives more weight to the most recently observed payoff. This results in a similar bias:

Theorem 2: In period one, a learner observes k payoffs, $x_{1,1}, x_{1,2}, ..., x_{1,k}$, each independently drawn from the payoff distribution $f(x)$, with expected value $E[x] = u$. The learner updates his or her belief, z after each observed payoff, giving more weight to the most recently observed payoff. Specifically, if the prior belief was z_t, the new belief is $z_{t+1} = (1-b)z_t + bx_t$, where b is the weight on the most recently observed payoff. The initial belief is assumed to be unbiased: $z_0 = u$. In period two the learner samples $n(z_{1,k})$ payoffs, each independently drawn from the payoff distribution $f(x)$. Here $n(z_{1,k}) \geq 1$, and $n(z_{1,k})$ is a function of belief after the k observed payoffs in the first period, $z_{1,k}$. Let z_2 be the belief after both periods one and two. Then i) $E[z_2] < u$ whenever $n(z_{1,k})$ is a strictly increasing function of $\bar{x}_1$, and ii) $E[z_2] > u$ whenever $n(z_{1,k})$ is a strictly decreasing function of $\bar{x}_1$.

Proof: See the Appendix.

Bayesian updating

One might suspect that the bias occurs because the learner does not take sample size into account when updating. A rational learner should after all update less when the sample size is small and update more if the sample size is large, which would seem to work against the above bias. This leads to the question whether the bias persists if the learner is a Bayesian updater, who takes sample size into account when updating, and whose belief is the conditional expectation given the observed data, i.e., $b_2 = E[u|X]$? The answer is that there is no bias on average in the following sense: the expected value of the belief after the two periods, $E[b_2]$, is equal to the expected value of the prior. For example, if Bayesian learners are learning about the mean of a random variable, u_i, and the means are drawn from a prior distribution, $f(u_i)$, with expected value $E[u_i] = m$, then $E[b_2] = m$. Averaging over different learners, who draw different values of u_i, thus generates no bias. Nevertheless, the distribution of beliefs may still be 'biased' (or skewed) in the following sense: a majority of Bayesian updaters will have a belief after two periods below m.

To explain this in more detail, consider a learner who observes independent draws from a random variable with distribution $f_i(X)$ with expected value u_i. The learner knows that u_i is drawn from a prior distribution with expected value m ($E[u_i] = m$). Let b_1 be the belief after having observed k independent draws $x_{1,1}, x_{1,2}, ..., x_{1,k}$ from $f_i(X)$ in the first period, $b_1 = E[u_i | x_{1,1}, x_{1,2}, ..., x_{1,k}]$. Let b_2 be the belief after all observations, in both period one and two: $b_2 = E[u_i | x_{1,1}, ..., x_{2,n(b_1)}]$.

It is now important to distinguish between two ways in which the belief of a learner could be 'biased'. Consider, first, a given value of $u_i = k$. We may ask if the expected belief is unbiased in the sense that $E[b_n|u_i = k] = k$, where b_n is the belief after a fixed sample of n observations (fixed in the sense that the number of observations is independent of the value of the observations). The average is unbiased in this sense but it is well-known that Bayesian updating is not. Suppose for example that u_i is drawn from a normal distribution with mean zero and variance one. The learner observes $x_i = u_i + \varepsilon_i$ where ε_i is drawn from a normal distribution with mean zero and variance one. For this set up it is well-known that expected value of the posterior after one observation is $x_i/2$ (e.g., DeGroot, 1970). Suppose now $u_i = 5$. The expected value of the belief after one observation, given that $u_i = 5$, equals $E[x_1/2|u_i = 5] = 2.5$ which is lower than 5. Bayesian updating can thus be biased in the sense that the expected belief, given a set of observations and $u_i = k$, is not equal to k. The reason is that the belief only gradually increases from the prior (zero) towards the expected value (five).

Bayesian updating is unbiased, however, if we average over all learners, in the following sense. Let b_i denote the belief of a Bayesian learner, i.e., $b_i = E[u_i|X]$, where X is the data observed. Then $E[b_i] = m$ where $m = E[u_i]$ is the expected value of the prior. For example, suppose u_i is drawn from a normal distribution with mean zero and variance one and the learner observes $x_i = u_i + \varepsilon_i$, where ε_i is drawn from a normal distribution with mean zero and variance one. We then have $E[b_i] = E[x_i/2] = E[(u_i + \varepsilon_i)/2] = 0$. Thus, $E[b_i] = E[u_i] = 0$.

We now show that Bayesian updating remains unbiased in this latter sense, even if the sample size in period two depends on the belief after period one. This is in fact true for any distribution:

Theorem 3: In period one the learner observes k independent draws $x_{1,1}, x_{1,2}, ..., x_{1,k}$, from distribution $f_i(X)$ with expected value u_i. The learner knows that u_i is drawn from a prior distribution, $f(u_i)$, with expected value $E[u_i] = m$. In period two, the learner observes $n(b_1)$ independent draws from distribution $f_i(X)$: $x_{2,1}, x_{2,2}, ..., x_{2,n(b_1)}$. The sample size in the second period, $n(b_1)$, is a function of the belief after the first period, $b_1 = E[u_i|x_{1,1}, ..., x_{1,k}]$. Let b_2 be the belief after all observations, in both period one and two: $b_2 = E[u_i|x_{1,1}, ..., x_{2,n(b_1)}]$. Then $E[b_2] = m$.

Proof of Theorem 3: The proof is very simple. We condition on the belief after the first period: $E[b_2|b_1]$. Because conditional expectations are martingales (e.g., Williams, 1991, p. 96), we have $E[b_2|b_1] = b_1$. That is, there is no change, in expectation, from the belief after period one to the belief after the information in the second period. This is true since if information that would lead to such a change could be anticipated in period one it should be already have been incorporated into the belief, of a rational agent, in period one. From $E[b_2|b_1] = b_1$ it follows that $E[b_2] = E_{b_1}[E[b_2|b_1]] = E[b_1]$. By the tower property of martingales (e.g., Williams, 1991, p. 88) we have $E[b_1] = E[E[u_i|x_{1,1}, x_{1,2}, ..., x_{1,k}]] = E[u_i] = m$. Thus, $E[b_2] = m$.

While there is no bias when averaging over the beliefs by all learners (i.e., $E[b_2] = m$), most Bayesian learners may nevertheless end up with a belief below m if positive beliefs lead to larger sample sizes. To illustrate this, suppose $f_i(X)$ is a normal distribution with mean u_i and variance σ_e^2. Moreover, suppose u_i is drawn from a normal distribution with mean m and variance σ_u^2. The learner observes k independent draws $x_{1,1}, x_{1,2}, ..., x_{1,k}$ from $f_i(X)$ in the first period. In period two, the learner observes $n(b_1)$ independent draws from distribution $f_i(X)$: $x_{2,1}, x_{2,2}, ..., x_{2,n(b_1)}$.

Theorem 4: If the payoff and the prior distributions are both normal, then i) whenever $n(b_1)$ is a strictly increasing function of b_1, most learners will, after the second period, underestimate the random variable they are learning about: $Pr(b_2 < m) > 0.5$. ii) Whenever $n(b_1)$ is a strictly decreasing function of b_1, most learners will, after the second period, overestimate the random variable they are learning about: $Pr(b_2 > m) > 0.5$.

Proof: See the Appendix.

Thus, even if Bayesian learners are learning about a symmetric distribution, and they have a prior which is symmetric around zero ($m = 0$), most learners will end up with a negative belief after sampling. This is because of the adaptive sampling policy. The intuition is similar to that for the bias for the learning policy that simply takes the average: negative initial beliefs are more persistent because the learner will then not take many additional samples and the initial few negative observations will be weighted heavily. Note that such a bias would not occur if the learner followed a fixed sampling policy and decided, at the outset, how many samples to take. Such a learner would, when learning about a normally distributed payoff with a mean taken from normally distributed prior with mean zero, be equally likely to have a negative or a positive belief. Note also that when sampling is adaptive all Bayesian learner know that only 50% of all means are negative. Still, most Bayesian learners end up believing that the mean they observe is negative.

That a majority of Bayesian learners end up with a negative belief may seem paradoxical since on average there is no bias: $E[b_2] = 0$ when $m = 0$, i.e., the average belief, averaging over all Bayesian learners, is equal to the mean of the prior. The paradox is resolved by noting that the learners who believe the mean is negative are less confident in this estimate than the learners who believe that the mean is positive and thus took a larger sample in the second period. This results in a skewed distribution of beliefs after the second period. Most beliefs are below $m = 0$ but those below $m = 0$ are less likely to be extreme beliefs than beliefs above $m = 0$ are since beliefs below $m = 0$ are based on smaller sample sizes than those above $m = 0$.

Because Bayesian updating with normal distributions relies on the observed average it may also seem puzzling that Bayesian updating is unbiased in the sense that $E[b_2] = m$

(Theorem 3) while averaging is biased in the sense that $E[\bar{x}|u_i = k] < k$ for all k (Theorem 1). Consider again the case when u_i is drawn from a normal distribution with mean m and variance σ_u^2. In period one the learner observes k independent draws $x_{1,1}, x_{1,2}, ..., x_{1,k}$, $x_{1,j} = u_i + \varepsilon_j$ where ε_j are independently drawn from a normal distribution with mean zero and variance σ_e^2. In period two, the learner observes $n(b_1)$ independent draws from the same distribution: $x_{2,1}, x_{2,2}, ..., x_{2,n(b_1)}$, where $n(b_1)$ is a function of $b_1 = E[u_i|x_{1,1}, ..., x_{1,k}]$. The Bayesian estimate of u_i given the observations in both periods equals (DeGroot, 1970): $E[u_i|X] = (\bar{x}\sigma_u^2 + (\sigma_e^2/n_2)m)/(\sigma_u^2 + (\sigma_e^2/n_2))$, where $n_2 = k + n(b_1)$, the total number of observations in both periods, and $\bar{x}$ is the average of all observations in both periods. As $\sigma_u^2 \to \infty$ $E[u_i|X] \to \bar{x}$, i.e., converges to the average. How can then the Bayesian estimate be unbiased while the average is biased? The puzzle is resolved by noting that when σ_u^2 grows large compared to σ_e^2, the signal to noise ratio (σ_u^2/σ_e^2) becomes large. When $\sigma_u^2 \to \infty$ the signal to noise ratio goes to infinity. The noise term is then vanishingly small compared to the values of u_i. Because the bias in the average depends on the possibility that the first period observations can far below (or above) u_i, i.e., depends on noise, there is only a vanishingly small bias in the average when $\sigma_u^2 \to \infty$, resolving the seeming paradox.

Implications for Understanding Learning

The bias resulting from adaptive sampling policies implies that even seemingly unbiased learning algorithms, such as averaging or Bayesian updating, can result in biased beliefs, at least in the sense that most learners underestimate (or overestimate) an alternative. This offers an alternative explanation of some judgment biases; an explanation that does not require a psychological bias that assumes biases in information processing. For example, suppose it is observed that a firm tend to underestimate the productivity of graduates from universities. Such a negativity bias. can be explained by an adaptive sampling policy (the firm tend to hire fewer people from universities it has had worse experience with) and does not require motivated reasoning or a cognitive bias.

Appendix: Proofs

Derivation of equation 1:

$$E[b_2] = P(\bar{x}_1 > c)E[b_2|\bar{x}_1 > c] + P(\bar{x}_1 < c)E[b_2|\bar{x}_1 < c].$$

We have

$$E[b_2|\bar{x}_1 > c] = E[\frac{S_2}{k+h}|\bar{x}_1 > c] + E[\frac{\bar{x}_1 k}{k+h}|\bar{x}_1 > c].$$

where S_2 is the sum of the observations in period two. Due to independence, $E[\frac{S_2}{k+h}|\bar{x}_1 > c] = \frac{1}{k+h}E[S_2] = 0$. Using the expression for the expectation of a truncated normal distribution (e.g. Greene 2000, p. 952) we have

$$E[\bar{x}_1|\bar{x}_1 > c] = \frac{\sigma_e}{\sqrt{k}}\frac{\phi(\alpha)}{1-\Phi(\alpha)},$$

where $\alpha = c/(\sigma_e/\sqrt{k}) = -c\sqrt{k}/\sigma_e$. Moreover, $P(\bar{x}_1 > c) = 1 - \Phi(\alpha)$. Hence,

$$P(\bar{x}_1 > c)E[b_2|\bar{x}_1 > c] = (1-\Phi(\alpha))\frac{k}{k+h}(\frac{\sigma_e}{\sqrt{k}}\frac{\phi(\alpha)}{1-\Phi(\alpha)}) = \frac{\sigma_e\sqrt{k}}{k+h}\phi(\alpha).$$

Similarly,

$$P(\bar{x}_1 < c)E[b_2|\bar{x}_1 < c] = \Phi(\alpha)\frac{k}{k+l}(-\frac{\sigma_e}{\sqrt{k}}\frac{\phi(\alpha)}{\Phi(\alpha)}) = -\frac{\sigma_e\sqrt{k}}{k+l}\phi(\alpha).$$

Overall we get

$$E[b_2] = \frac{\sigma_e\sqrt{k}}{(k+l)(k+h)}[((k+l)\phi(\alpha) - (k+h)\phi(\alpha)] = -\frac{\sigma_e\sqrt{k}}{(k+l)(k+h)}[(h-l)\phi(\alpha)].$$

Now, $\phi(\alpha) = \frac{1}{\sqrt{2\pi}}e^{-\alpha^2/2}$. Hence,

$$E[b_2] = -\frac{\sigma_e\sqrt{k}(h-l)}{\sqrt{2\pi}(k+l)(k+h)}e^{-c^2k/2\sigma_e^2}.$$

Proof of Theorem 1: The average of all payoffs observed in periods one and two is, given the average observed in period one ($\bar{x}_1$), is

$$\bar{x}_2 = \frac{k\bar{x}_1 + \sum_{j=1}^{n(\bar{x}_1)} x_{2,j}}{k + n(\bar{x}_1)},$$

where $k\bar{x}_1$ is the sum of all observed payoffs in the first period. Thus,

$$E[\bar{x}_2|\bar{x}_1] = \frac{k\bar{x}_1 + E[\sum_{j=1}^{n(\bar{x}_1)} x_{2,j}|\bar{x}_1]}{k + n(\bar{x}_1)}.$$

Because $E[\sum_{j=1}^{n(\bar{x}_1)} x_{2,j}|\bar{x}_1] = un(\bar{x}_1)$, this can be written as

$$E[\bar{x}_2|\bar{x}_1] = \frac{k\bar{x}_1}{k + n(\bar{x}_1)} + \frac{n(\bar{x}_1)u}{k + n(\bar{x}_1)}.$$

Moreover, because $E(\bar{x}_2) = E_{\bar{x}_1}(E[\bar{x}_2|\bar{x}_1])$ we get

$$E(\bar{x}_2) = E_{\bar{x}_1}(\frac{k\bar{x}_1}{k + n(\bar{x}_1)}) + E_{\bar{x}_1}(\frac{n(\bar{x}_1)u}{k + n(\bar{x}_1)})$$

Let $g(\bar{x}_1) = \frac{1}{k+n(\bar{x}_1)}$. Because $E(\bar{x}_1 g(\bar{x}_1)) = E(\bar{x}_1)E(g(\bar{x}_1)) + cov(\bar{x}_1, g(\bar{x}_1))$ and $E(\bar{x}_1) = ku$ we get

$$E_{\bar{x}_1}(\frac{k\bar{x}_1}{k + n(\bar{x}_1)}) = kuE(g(\bar{x}_1)) + Cov(\bar{x}_1, g(\bar{x}_1)).$$

Note that

$$kE(g(\bar{x}_1)) = E(\frac{k}{k+n(\bar{x}_1)}) = 1 - E(\frac{n(\bar{x}_1)}{k+n(\bar{x}_1)})$$

Overall we get

$$\begin{aligned} E(\bar{x}_2) &= u[1 - E(\frac{n(\bar{x}_1)}{k+n(\bar{x}_1)}] \\ &+ Cov(\bar{x}_1, g(\bar{x}_1)) + uE(\frac{n(\bar{x}_1))}{k+n(\bar{x}_1)}) \\ &= u + Cov(\bar{x}_1, g(\bar{x}_1)). \end{aligned}$$

$Cov(\bar{x}_1, g(\bar{x}_1))$ is negative (positive) whenever $g(\bar{x}_1)$ is a strictly decreasing (increasing) function of $\bar{x}_1$. Because $g(\bar{x}_1) = 1/(k+n(\bar{x}_1))$, which is a strictly decreasing function of $n(\bar{x}_1)$, $Cov(\bar{x}_1, g(\bar{x}_1))$ is negative whenever $n(\bar{x}_1)$ is a strictly increasing function of $\bar{x}_1$ and $Cov(\bar{x}_1, g(\bar{x}_1))$ is positive whenever $n(\bar{x}_1)$ is a strictly decreasing function of $\bar{x}_1$.

Proof of Theorem 2: The belief after all payoffs observed in both periods one and two, $z_{2,n(z_{1,k})}$, given the belief at the end of period one ($z_{1,k}$), is equal to

$$z_{1,k}(1-b)^{n(z_{1,k})} + x_{2,1}b(1-b)^{n(z_{1,k})-1} + ... + bx_{2,n(z_{1,k})}.$$

The conditional expected value $E[z_{2,n(z_{1,k})}|z_{1,k}]$ equals

$$z_{1,k}(1-b)^{n(z_{1,k})} + E[x_{2,1}b(1-b)^{n(z_{1,k})-1} + ... + bx_{2,n(z_{1,k})}|z_{1,k}].$$

Because $E[x_{2,j}|z_{1,k}] = u$, this can be written as

$$E[z_{2,n(z_{1,k})}|z_{1,k}] = z_{1,k}(1-b)^{n(z_{1,k})} + (1-(1-b)^{n(z_{1,k})})u.$$

Moreover, because $E(Y) = E_X(E[Y|X])$, $E[z_{2,n(z_{1,k})}]$ equals

$$E_{z_{1,k}}[z_{1,k}(1-b)^{n(z_{1,k})}] + E_{z_{1,k}}[(1-(1-b)^{n(z_{1,k})})]u.$$

Let $g(z_{1,k}) = (1-b)^{n(z_{1,k})}$. Because $E(z_{1,k}g(z_{1,k})) = E(z_{1,k})E(g(z_{1,k})) + Cov(z_{1,k}, g(z_{1,k}))$, and $E(z_{1,k}) = u$, we get

$$\begin{aligned} E[z_{2,n(z_{1,k})}] &= uE(g(z_{1,k})) + Cov(z_{1,k}, g(z_{1,k})) + u(1 - E(g(z_{1,k})) \\ &= u + Cov(z_{1,k}, g(z_{1,k})). \end{aligned}$$

$Cov(z_{1,k}, g(z_{1,k}))$ is negative (positive) whenever $g(z_{1,k})$ is a strictly decreasing (increasing) function of $\bar{x}_1$. Because $g(z_{1,k}) = (1-b)^{n(z_{1,k})}$, which is a strictly decreasing function of $n(z_{1,k})$, $Cov(z_{1,k}, g(z_{1,k}))$ is negative whenever $n(z_{1,k})$ is a strictly increasing function of $z_{1,k}$ and $Cov(z_{1,k}, g(z_{1,k}))$ is positive whenever $n(z_{1,k})$ is a strictly decreasing function of $z_{1,k}$.

Proof of Theorem 4: Without loss of generality we focus on the case when $m = 0$. If m differs from zero, the distribution of payoffs is identical to a constant, m, plus a random variable with a normal distribution with mean zero. Let S_1 be the sum of the observed payoffs in period one. This sum is equally likely to be positive or negative. Suppose $S_1 = z > 0$. It follows that the belief after the first period, $E[u|S_1] = (1/k)S_1\sigma_u^2/(\sigma_u^2 + \sigma_e^2/k)$, is positive. Denote this belief by b_1, i.e., $b_1 = E[u|S_1]$. Note that b_1 is a strictly increasing function of S_1. This positive belief turns into a negative belief after period 2 whenever

$$b_2 = \frac{\frac{1}{k+n(b_1)}(S_1+S_2)\sigma_u^2}{\sigma_u^2 + \frac{\sigma_e^2}{k+n(b_1)}} < 0,$$

i.e., whenever $S_2 < -S_1$. We seek the probability of this event given the observed value of $S_1 = z$, i.e., we seek $Pr(S_2 < -S_1|S_1 = z)$.

Conditional on $S_1 = z$, the sum of the payoffs in the second period, S_2, is a normally distributed random variable with mean $n(z)E(u|S_1 = z) = n(z)z\sigma_u^2/(\sigma_u^2 + \sigma_e^2/k)$ and variance $n(z)^2Var(u|S_1 = z) + \sigma_e^2 n(z)$, where $Var(u_i|x_1 = z) = \sigma_u^2\sigma_e^2/(k\sigma_u^2 + \sigma_e^2)$ (see Lindgren (1993, p. 289). It follows that

$$Pr(S_2 < -S_1|S_1 = z) = \Phi(\frac{-z - n(z)(z/k)\sigma_u^2/(\sigma_u^2 + \sigma_e^2/k)}{\sqrt{n(z)^2\sigma_u^2\sigma_e^2/(k\sigma_u^2 + \sigma_e^2) + \sigma_e^2 n(z)}}).$$

Altogether, the probability that a positive belief after period one turns into a negative belief after period two equals

$$Pr(+ \to -) = \int_{z=0}^{z=+\infty} \Phi(\frac{-z - n(z)(z/k)\sigma_u^2/(\sigma_u^2 + \sigma_e^2/k)}{\sqrt{n(z)^2\sigma_u^2\sigma_e^2/(k\sigma_u^2 + \sigma_e^2) + \sigma_e^2 n(z)}}) f(z|z > 0)dz.$$

Because $f(z|z > 0) = f(z)/P(z > 0) = f(z)/0.5$, $Pr(+ \to -)$ equals

$$\int_{z=0}^{z=+\infty} \Phi(\frac{-z - n(z)(z/k)\sigma_u^2/(\sigma_u^2 + \sigma_e^2/k)}{\sqrt{n(z)^2\sigma_u^2\sigma_e^2/(k\sigma_u^2 + \sigma_e^2) + \sigma_e^2 n(z)}}) 2f(z)dz, \quad (2)$$

where $f(z)$ is the density of the sum of the payoffs in the first period which is a normal distribution with mean zero.

Suppose next $S_1 = z < 0$ implying that $E[u|S_1] < 0$. This negative belief turns into a positive belief after period 2 whenever $S_2 > -z$. Following the same reasoning as above $Pr(S_2 > -S_1|S_1 = z)$ equals

$$1 - \Phi(\frac{-z - n(z)(z/k)\sigma_u^2/(\sigma_u^2 + \sigma_e^2/k)}{\sqrt{n(z)^2\sigma_u^2\sigma_e^2/(k\sigma_u^2 + \sigma_e^2) + \sigma_e^2 n(z)}}).$$

Because $1 - \Phi(y) = \Phi(-y)$, this can be written as

$$\Phi(\frac{z + n(z)(z/k)\sigma_u^2/(\sigma_u^2 + \sigma_e^2/k)}{\sqrt{n(z)^2\sigma_u^2\sigma_e^2/(k\sigma_u^2 + \sigma_e^2) + \sigma_e^2 n(z)}}).$$

Altogether, the probability that a negative belief after period one turns into a positive belief after period two, $Pr(- \to +)$, equals

$$\int_{z=-\infty}^{z=0} \Phi(\frac{z + n(z)(z/k)\sigma_u^2/(\sigma_u^2 + \sigma_e^2/k)}{\sqrt{n(z)^2\sigma_u^2\sigma_e^2/(k\sigma_u^2 + \sigma_e^2) + \sigma_e^2 n(z)}}) 2f(z)dz.$$

After the variable substitution, $z = -y$, this integral equals

$$\int_{y=0}^{y=\infty} \Phi(\frac{-y - n(-y)(y/k)\sigma_u^2/(\sigma_u^2 + \sigma_e^2/k)}{\sqrt{n(-y)^2\sigma_u^2\sigma_e^2/(k\sigma_u^2 + \sigma_e^2) + \sigma_e^2 n(-y)}})2f(y)dy.$$

where we have used the fact that $f(-y) = f(y)$.

We wish to show that $Pr(+ \rightarrow -) > Pr(- \rightarrow +)$, which requires us to show that

$$\Phi(\frac{-z - n(z)(z/k)\sigma_u^2/(\sigma_u^2 + \sigma_e^2/k)}{\sqrt{n(z)^2\sigma_u^2\sigma_e^2/(k\sigma_u^2 + \sigma_e^2) + \sigma_e^2 n(z)}}) >$$
$$\Phi(\frac{-z - n(-z)(z/k)\sigma_u^2/(\sigma_u^2 + \sigma_e^2/k)}{\sqrt{n(-z)^2\sigma_u^2\sigma_e^2/(k\sigma_u^2 + \sigma_e^2) + \sigma_e^2 n(z)}}),$$

over the positive domain of z. To do so, note first that the derivative of

$$g(z,w) = \frac{-z - wz\sigma_u^2/(\sigma_u^2 + \sigma_e^2/k)}{\sqrt{w^2\sigma_u^2\sigma_e^2/(k\sigma_u^2 + \sigma_e^2) + \sigma_e^2 w}},$$

with respect to w is

$$\frac{\delta g(z,w)}{\delta w} = \frac{z}{2w\sqrt{\frac{\sigma_e^2 w(\sigma_e^2 + \sigma_u^2 + \sigma_e^2 w)}{\sigma_u^2 + \sigma_e^2}}},$$

which is positive whenever $z > 0$. Hence $g(z,w)$ is an increasing function of $n(z)$. Because $\Phi(x)$ is a increasing function of x and $n(z)$ is an increasing function of z, implying that $n_2(z) > n_2(-z)$, it follows that $\Phi(g(z,n(z)) > \Phi(g(z,n(-z))$ implying $Pr(+ \rightarrow -) > Pr(- \rightarrow +)$.

References

DeGroot, M. H. (1970). *Optimal statistical decisions*. New York, NY: McGraw-Hill Book Company.

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

Denrell, J. (2007). Adaptive learning and risk taking. *Psychological Review*, *114*, 177–187.

Denrell, J., & March, J. (2001). Adaptation as information restriction: The hot stove effect. *Psychological Review*, *12*, 523–538.

Dittmar, A., & Duchin, R. (2016). Looking in the rear-view mirror: The effect of managers' professional experiences on corporate financial policy. *Review of Financial Studies*, *29*, 565–602.

Erev, I., & Roth, A. (2014). Maximization, learning, and economic behavior. *Proceedings of the National Academy of Sciences*, *111*, 10818–10825.

Fetchenhauer, D., & Dunning, D. (2014). Why so cynical?: Asymmetric feedback underlies misguided skepticism regarding the trustworthiness of others. *Psychological Science*, *21*, 189–193.

Le-Mens, G., Kovács, B., Avrahami, J., & Kareev, Y. (2018). How endogenous crowd formation undermines the wisdom of the crowd in online ratings. *Psychological Science*, *29*, 1475–1490.

Lindgren, B. W. (1993). *Statistical theory*. New York, NY: Chapman & Hall.

Williams, D. (1991). *Probability with martingales*. Cambridge, UK: Cambridge University Press.