Social learning, search and heterogeneity of payoffs

Edward Cartwright*
Department of Economics,
Keynes College,
University of Kent,
Canterbury, Kent.
CT2 7NP. UK.
E.J.Cartwright@kent.ac.uk

February 2007

.

Keywords: Social learning, search, herding
JEL Classification: C7, D8

Abstract

We consider a simple model that combines elements of search and social learning. Acting in sequence, and observing the action adopted by a previous agent, agents must search for an action. We explore why agent heterogeneity may increase expected payoffs and demonstrate that social learning may be most effective if agents are heterogenous.

*The research reported here was supported by ESRC Grant RES-061-23-001, 'Social learning and the Theory of Search'.

1 Introduction

Economic agents must often search for, say, goods that are of high quality or cheap price. Modelling such search has been the subject of a large literature (including Stigler 1961, Kohn and Shavell 1974, Rosenfield and Shapiro 1981). Often, however, agents begin search having observed the choice of one or more other agent (McFadden and Train 1996, Bikhchandani, Hirshleifer and Welch 1998, Banerjee and Fudenberg 2004). This means that costly search could potentially be avoided by social learning and imitation. In this paper we consider a simple model in which elements of search and social learning combine.

To briefly outline the model: Agents must adopt some action where action could be interpreted as, for example, choice of restaurant or grocery store. Given uncertainty over, say, the quality of a restaurant or cheapness of a grocery store agents can, as in a standard model of search, sample any number of actions before committing to an adopted action. Agents are assumed, however, to act in sequence with each agent being able to observe the action adopted by the agent previous to him in the sequence. An agent can imitate the action that he observes. Note that agents will be assumed rational and so our approach is distinguished from much of the literature on social learning where agents exhibit bounded rationality (e.g. Ellison and Fudenberg 1995).

The sequential nature of choice means that our model is similar to models of herding (Bikhchandani, Hirshleifer and Welch 1998). One crucial difference is that in herding models an agent only chooses once. Thus, an agent who adopts an action based on social learning cannot 'go back on his choice with hindsight'. In our model an agent can imitate a previous agent, find that imitating this agent did not produce the desired outcome and sample for himself. For example, an agent can try the restaurant that some other agent adopted, find that he does not like it, next time try some other restaurant and continue to search until he finds a restaurant that he does like.

The obvious benefit from observing another agent's choice is that an agent can imitate and avoid 'costly search'. For this to work requires that the agent should like what he imitates. For example, if one agent's 'ideal restaurant' is an others 'worst choice' then an agent who imitates may still be required to pursue costly search if he is 'not happy' with the choice he imitates. This would suggest that agents do best if preferences are homogenous. We shall demonstrate, however, that preference heterogeneity may be optimal.

Agent heterogeneity can prove beneficial in terms of 'accumulated learning'. To illustrate contrast two cases. *Scenario 1:* agent 1 finds and adopts an action that yields a 'very high utility'. Agent 2 will imitate the action, and provided preferences are not too heterogenous, will find that it yields him a 'very high utility' and so adopt it. Thus, if agent 1 finds an action that yields a 'very high utility' this will 'get passed on to' agent 3 and so on. *Scenario 2:* agent 1 finds and adopts an action that yields a 'just sufficient utility'. Agent 2 will imitate the action. If agents 1 and 2 are homogenous then agent 2 will also find the action yields a 'just sufficient utility' and so adopt it. If, however, agent 2 has slightly different preferences to agent 1 then he may find that the imitated action yields a 'not quite sufficient utility'. In this case agent 2 will decide to search and may potentially find and adopt an action that yields a 'very high utility'. This need not be good news for agent 2, who has the cost of search, but is good news for agent 3. In this way heterogeneity can prove beneficial. In short, 'high utility actions' will always get 'passed down the sequence' but 'low utility actions' get 'weeded out' by preference differences.

In general, the consequences of heterogeneity are more search and this has benefits and costs for a particular agent. The benefits are that heterogeneity results, as seen above, in more accumulated sampling so that an agent well into the sequence can expect to imitate a 'high utility action'. Heterogeneity means, however, that each agent has a greater chance of having to perform costly search himself. Characterizing these trade-offs through a heterogeneity function we provide a limit case. 'Going below' this limit case means that payoffs are maximized if agents are identical: the costs of heterogeneity outweigh the benefits. 'Going above' the limit case means that agents who come later in the sequence do best if there is heterogeneity: the benefits of heterogeneity eventually outweigh the costs.

Our results highlight the potential advantages of social learning. That social learning can prove advantageous in a model of search would seem obvious because it allows an agent to avoid the costs of search. We find, however, more surprisingly, that social learning can do even better. It not only allows agents to avoid the costs of search but also results in, on average, higher utility actions being adopted. Thus, social learning results in both less search and the adoption of more preferable actions. Interpreting our results another way, agents would be willing to pay to observe the choice of some other agent. The literature on social learning is diverse but often suggests that social learning can lead to 'inefficiencies' if agents ignore private signals to 'follow the herd' (e.g. Bikhchandani, Hirshleifer and Welch 1992). This can not be the case in our model where social learning proves unambiguously beneficial to agents.

Preference heterogeneity has been considered in models of search and social learning (e.g. McFadden and Train 1996, Smith and Sorenson 2000 and Goeree et. al. 2006). Our understanding, however, of the consequences of preference heterogeneity for social learning still remain limited (Keller, Rady and Cripps 2006). In this paper, we find that the benefits of social learning may be greater if there is preference heterogeneity. Equivalently, agents may pay more to observe others when there is agent heterogeneity than when all agents are identical. This is a surprising result that illustrates the importance of investigating further the relationship between preference heterogeneity and social learning.

A more detailed interpretation of the results and discussion of related literature is contained in Section 5. We shall illustrate the arguments using a simple 'uniformly distributed' model. This is done for transparency and generalizations are discussed in Section 4. Section 2 outlines the model, Section 3 characterizes optimal behavior and provides the main results. All proofs are contained in Section 6.

2 Model

There exists a set of *agents* $N = \{1, 2, 3, ...\}$, a set of *actions* $X = [0, 1]$ and discrete time *periods* $t = 1, 2, ...$ In periods $t = 1, 2, 3, ..$ agent 1 must choose an action. He can choose an action by *sampling* or using *experience* (both to be explained below). If agent 1 uses experience in period t_1 then agent 2 must choose an action in period $t_1 + 1, t_1 + 2, ...$ He can sample, use experience or *imitate agent 1*. If agent 2 uses experience in period t_2 then agent 3 must choose an action in periods $t_2 + 1, t_2 + 2, ...$ He can sample, use experience or *imitate agent 2*. And so on. Let x_t^i denote the action chosen by agent i in period t.

For each agent i there exists a *utility function* $u_i(x) : X \to [0, 1]$. Agent i receives utility $u_i(x_t^i)$ in period t. If agent i uses **experience** in period t_i then he selects action $b = \max_{t<t_i} u_i(x_t^i)$ and we say that agent i *has adopted* action $b_i := x_{t_i}^i$. Note (the formal argument will follow) that once an agent uses experience he uses experience in all subsequent periods and thus b_i will not change. If an agent **samples** then his action chosen is a random draw from set X determined by distribution $F[0, 1]$ where we use $F[l, h]$ to denote

the uniform distribution over interval $[l,h]$.[1] If agent $i+1$ **imitates** then he chooses action b_i. That is he chooses the action adopted by the agent previous to him in the sequence.

Agents act to maximize their (expected) discounted stream of utility using discount rate r. To avoid confusion we shall use utility to refer to the in-period payoff and expected payoff to refer to the discounted stream of utilities. A *strategy* for agent i dictates whether he should sample, use experience, or imitate in each period $t > t_{i-1}$ conditional on the history of choices made.

The assumption that agent i only chooses once agent $i-1$ has used experience is made for modelling convenience. Essentially we are modelling a setting where agent $i+1$ will begin search with 'additional information' from his observation of agent i. We could equivalently consider a model where there exists only one agent at any time but this agent is periodically replaced.[2] Note also that while, in principle, an agent may never use experience for realistic parameter values one could expect experience to be used within a 'reasonable number of periods'.

We are using what Kohn and Shavell (1974) call the experience case in that an agent gets utility $u_i(x_i^t)$ irrespective of whether he imitates, samples or uses experience. More generally, one may want to impose, say, a fixed cost to sampling and imitation. Our results, and the proofs as written, do extend, however, without qualification to a general formulation for sampling and imitation costs (such as that in Kohn and Shavell 1974).[3] Note that there are indirect costs to sampling (and imitation) if an agent expects to

[1] That is,

$$F[l,h](x) := \begin{cases} 0 \text{ if } x < l \\ \frac{x-l}{h-l} \text{ if } x \in [l,h) \\ 1 \text{ if } x \geq h \end{cases} .$$

[2] This does make the analysis less transparent. For example, if an agent expects to 'live' only a finite number of strategies it alters his optimal sampling strategy. Also, if agent i has not used experience but is replaced by agent $i+1$ then $i+1$ will be obliged to continue sampling. This is analytically irrelevant but does mean we cannot be so precise in saying, for example, what agent 3 would do as compared to agent 2.

[3] Sampling costs prove irrelevant because we taken as given a switchpoint w (see the next section). Sampling costs alter the value of w but nothing else. Thus, our results remain valid. Imitation costs prove irrelevant because there are only two possibilities: (i) imitation costs are low enough that all agents imitate or (ii) imitation costs are high enough that no-one imitates. In case (i) our results remain valid. In case (ii) our results change but in a trivial way.

receive a lower utility by sampling than could have been achieved through using experience.

We make a number of assumptions about the utility function u_i. Note that $u_i(x)$ itself is not known (otherwise agents have no reason to search) but agents will be assumed to know how utilities are distributed across actions.

Assumption 1: If agent i samples then his expected utility $u_i(x)$ is a random draw from $[0, 1]$ determined by distribution $F[0, 1]$.

Before detailing the remaining assumptions it is useful to briefly outline the model with the aid of figure 1. For each action x there exist real numbers $\underline{u}_x$ and $\overline{u}_x$ where $\underline{u}_x \leq x \leq \overline{u}_x$ is such that if agent i chooses action x then his payoff $u_i(x)$ is, ex-ante, a random draw on *uncertainty interval* $[\underline{u}_x, \overline{u}_x] \subset [0, 1]$. It must be emphasized that agent i's payoff is not stochastic but just unknown ex-ante. So if agent i chooses action x then he always receives utility $u_i(x)$ but $u_i(x)$ could be anything between $\underline{u}_x$ and $\overline{u}_x$. Equivalently we can interpret the model (as in Goeree et. al. 2006) that when agent i chooses action x he receives utility $u_i(x) = \pi_x + v_x^i$ where π_x is a common-value utility of action x and v_x^i is a private-value utility of action x drawn from some distribution f_x. In our model we set $\pi_x = x$ and $f_x = F[\underline{u}_x - x, \overline{u}_x - x]$.

There is a *degree of heterogeneity* parameter $\delta \in [0, 1]$ that determines the size of the uncertainty interval $\overline{u}_x - \underline{u}_x$. The smaller is δ then the larger is the interval and thus the greater the heterogeneity in preferences or, put another way, the relatively larger is the private-value utility. Specifically, if $\delta = 1$ then we set $u_i(x) = x$ for all x and so all agents are identical (see Figure 1a). In this case, when agent i imitates agent $i - 1$ he knows that he will receive the same payoff as agent $i - 1$. If $\delta = 0$ then there is maximal heterogeneity with, for example, the value $u_i(0.5)$ being a random draw from interval $[0, 1]$ (see Figure 1d). In this case, when agent i imitates agent $i - 1$ he can only form an expectation of how close his payoff will be to that of agent $i - 1$.

In stating the values of $\overline{u}_x$ and $\underline{u}_x$ we define an *upper bound* $h(x, \delta) := 1 - \delta(1 - x)$, a *lower bound* $l(x, \delta) := \delta x$ and introduce a *heterogeneity function* $\theta : X \times [0, 1] \to [0, 1]$. If $x \geq 0.5$ then we set $\underline{u}_x = \theta(x, \delta)$ and $\overline{u}_x = h(x, \delta)$ so the value $u_i(x)$ is, ex-ante, uniformly distributed on interval $[\theta(x, \delta), h(x, \delta)]$. If $x < 0.5$ then we set $\underline{u}_x = l(x, \delta)$ and $\overline{u}_x = \theta(x, \delta)$ so the value $u_i(x)$ is, ex-ante, uniformly distributed on interval $[l(x, \delta), \theta(x, \delta)]$. The upper and lower bounds $h(x, \delta)$ and $l(x, \delta)$ are fixed (essentially without loss of generality) but the bound $\theta(x,\delta)$ is left free to allow flexibility in the model. This is illustrated by comparing Figures 1b and 1c and will prove crucial in stating our results as we shall explain below.[4]

Before formally stating the remaining assumptions we note that, given Assumption 1, the bounds will prove irrelevant to an agent who samples and so are only relevant in looking at social learning.

Assumption 2: If $\delta = 1$ then $u_i(x) = x$ for all i and x. If $\delta < 1$ then utility $u_i(x)$ is a random draw for all i determined by distribution $F[\theta(x,\delta), h(x,\delta)]$ if $x \geq 0.5$ and $F[l(x,\delta), \theta(x,\delta)]$ if $x < 0.5$ where $h(x,\delta) := 1 - \delta(1-x)$ and $l(x,\delta) := \delta x$.

Assumption 2 formalizes the way that we shall model heterogeneity of payoffs. Note that if $E(x,\delta)$ denotes the expected utility of choosing action x then $E(x,\delta) = (\theta(x,\delta) + h(x,\delta))/2$ for $x \geq 0.5$ and $E(x,\delta) = (\theta(x,\delta) + l(x,\delta))/2$ for $x < 0.5$.

Assumption 3: The value of $E(x,\delta)$ is a continuous, non-decreasing function of x and a continuous monotonic function of δ. Also, $E(x,\delta) \leq x$ for all $x \geq 0.5$.

That the expected utility of action $x \geq 0.5$ be never more than x is not necessary for our results but desirable in interpretation. It means that payoff heterogeneity can never be of direct benefit to agents. In particular, if $\delta = 1$ then an agent who imitates someone using action x will get utility x. If $\delta < 1$ then they can expect utility at most x and possibly less (depending on $\theta(x,\delta)$).

We require one final, more technical assumption.

Assumption 4: If $x' > x \geq 0.5$ then $\theta(x',\delta) - \theta(x,\delta) \geq \delta(x' - x)$. If $x < x' < 0.5$ then $\theta(x,\delta) \leq \theta(x',\delta)$. Finally, $\theta(0.5,\delta) = \delta/2$ and $\lim_{x \longrightarrow 0.5, x<0.5} \theta(x,\delta) = 1 - \delta/2$.

Assumption 4 guarantees that the bounds $\overline{u}_x$ are $\underline{u}_x$ are continuous, increasing functions of x. This follows from continuity of $E(x,\delta)$ and the limits on $\theta(x,\delta)$ near $x = 0.5$. Assumption 4 simplifies the analysis but could be relaxed.

[4] If we think of the amount of heterogeneity as, say, $h(x,\delta) - \theta(x,\delta)$ then heterogeneity depends on both δ and $\theta(x,\delta)$. For example, comparing Figures 1b and 1c there is greater heterogeneity in Figure 1b but δ could be the same in both.

In interpretation there are broadly two ways we could proceed. First, we could assume that agent i becomes aware of $\theta(x, \delta)$ and $h(x, \delta)$ when he observes someone adopting action x. Intuitively, this would correspond to agent i observing both action b_{i-1} and utility $u_{i-1}(b_{i-1})$. Alternatively, we could assume that agent i only observes action b_{i-1} and remains ignorant of $\theta(b_{i-1}, \delta)$ and $h(b_{i-1}, \delta)$. The general correlation between utility functions is, however, known. Intuitively, this would correspond to agent i forming an expected utility from choosing action b_{i-1} on the basis that if agent $i-1$ has adopted action b_{i-1} then it must have yielded him a relatively high utility. We proceed using the later interpretation but mathematically both are identical. Either way, the function θ ties together the utility functions of agents in the sense that if agent i observes someone adopting action x then function θ allows him to form an expectation over the utility that he would receive from choosing x.

In a more general model the upper bound $h(x, \delta)$ and lower bound $l(x, \delta)$ would also be undetermined. To a large extent, however, given the freedom of $\theta(x, \delta)$ we can fix $h(x, \delta)$ and $l(x, \delta)$ without loss of generality. Basically, by manipulating $\theta(x, \delta)$ we can obtain the same degree of flexibility as if all bounds were left free. The present formulation will allow sharper results. Note also, that what happens for $x \geq 0.5$ is much more important for our purposes given that this is where utilities are highest and thus actions are likely to be adopted.

3 The optimal amount of heterogeneity

We begin by defining an optimal strategy (making use of Kohn and Shavell 1974). Let w be the real number called the *switchpoint* that solves

$$\int_{x>w} \frac{x-w}{r} dx = w - \frac{1}{2}. \tag{1}$$

We define the optimal strategy:

Optimal strategy: Agent 1: Sample in period $t+1$ if $u_1(x_t^1) < w$ and use experience otherwise. Agent $i \geq 2$: Imitate agent $i-1$ in period $t_{i-1}+1$. In period $t+1 > t_{i-1}+1$, sample if $u_i(x_t^i) < w$ and use experience otherwise.

Fixing a δ and assuming that agents behave according to the optimal strategy an expected payoff for agent i can be calculated. Let δ^i denote the value of

δ that maximizes the expected payoff of agent i (assuming that all agents use the optimal strategy). Our first result shows that the optimal strategy is indeed the optimal strategy. Formally, we find a Nash equilibrium in the sense that agent i's optimal strategy maximizes his payoff if the other agents use the optimal strategy.[5]

Proposition 1: If $\delta = 1$ or $\delta = \delta^j$ for some $j \in N$ and agents $1, 2, ..., i-1$ use the optimal strategy for some $i \in N$, then agent i maximizes his payoff by using the optimal strategy.[6]

If agent 1 uses the optimal strategy then his expected payoff is

$$U_1 := w\frac{(1+r)}{r}.$$

To explain, when $u_i(x_t^i) = w$ agent i is indifferent between using experience and sampling and if he was to use experience his payoff would be U_1.[7] If $u_i(x_t^i) < w$ and agent i uses the optimal strategy then his expected payoff is the same as if $u_i(x_t^i) = w$. Thus, U_1 is as above. For agents $i > 1$ their 'default payoff' is also U_1. More precisely, if agent i imitates agent $i-1$ but decides not to adopt action b_{i-1} then his expected payoff becomes U_1.

Before exploring the implications of heterogeneity consider the case where $\delta = 1$ and agents are homogenous. Applying Proposition 1 and the definition of optimal strategy we observe that Agent 1 will sample until he observes an action x such that $u_1(x) \geq w$. Agents $2, 3, 4, ...$ will then imitate and adopt action x. Thus, agent 1 'determines everything'. This is an example of what we shall call lock in where some action x, adopted by agent 1 in this example, will be adopted by all subsequent agents in the sequence. This is an example

[5] It is the unique Nash equilibrium (discounting tie-breaking rules) because agent i's optimal strategy is unique given the strategy of agent $i-1$ and agent 1 has a unique optimal strategy.

[6] There is no circularity in this argument: If $\delta = \delta^j$ then agents will want to behave according to the optimal strategy detailed in Lemma 1. If agents behave according to optimal strategy detailed in Lemma 1 then agent j wants $\delta = \delta^j$. More generally, agents may not want to imitate if δ is, say, low and so the 'optimal strategy' need not be consistent with Nash equilibrium.

[7] Using that his payoff is

$$w + \frac{w}{1+r} + \frac{w}{(1+r)^2} + ... = w + \frac{w}{r}$$

of herding (Bikhchandani, Hirshleifer and Welch 1998). Given that the action adopted by agent 1 is drawn from interval $[w, 1]$ according to the uniform distribution, the expected payoff of agents $2, 3, 4, ...$ is

$$U_i^1 := \frac{(1+w)}{2}\frac{(1+r)}{r} > U_1.$$

Agent heterogeneity may increase or decrease expected payoffs relative to U_i^1 as we now explain:

The *benefit of heterogeneity*: For δ sufficiently large there exists action $\mu^\delta > w$ such that $\theta(\mu^\delta, \delta) = w$. If some agent i adopts action $x \geq \mu^\delta$ then, given Proposition 1 and Assumptions 2 and 3, every agent $j > i$ will also adopt action x because $u_i(x) \geq w$. When this occurs we say that there is **lock in** to action $x \geq \mu^\delta$. We refer to μ^δ as the *lock in threshold.* If $\delta = 1$ then $\mu^\delta = w$ and so lock in occurs with agent 1 and on any action $x \geq w$. As δ increases then μ^δ increases. This implies that lock in will occur on some action $x \geq \mu^\delta > w$. Consequently more heterogeneity means that 'low actions get weeded out' and agents can expect to do better when lock in occurs. Note that for δ small we may have $\theta(1, \delta) < w$ and so no lock in threshold exists and lock in is not possible.

The *cost of heterogeneity*: If an agent imitates action x then his expected utility is $(\theta(x, \delta) + h(x, \delta))/2$. At best his expected payoff is x. More generally, his expected utility is less than x and, most importantly, a decreasing function of δ. Thus, the greater is heterogeneity the lower, ceteris paribus, is his expected utility from the imitated action.

Once the cost and benefit of heterogeneity are combined there is a trade-off that can mean heterogeneity is or is not desirable depending on θ. This is illustrated in Figure 2. There is no gain in locking in to a high action if the expected utility from imitating this action is still very low as seen in Figure 2b. Conversely, there is much to be gained from locking in on a high action if the expected utility from imitating this action is likely to be high as seen in Figure 2c. All will depend on the function θ as the following result shows.

Proposition 2: If $\theta(x, \delta) \leq \delta x$ for all $x \geq 0.5$ and δ then $\delta^i = 1$ for all i. If $\theta(x, \delta) > \delta x$ for all $x \geq 0.5$ and $\delta < 1$ then there exists finite $i^* > 2$ such that $\delta^{i^*} = 1$ for all $i < i^*$ and $\delta^{i^*} < 1$ for all $i > i^*$.

Thus, if function θ is 'sufficiently high' the expected payoff of agents later in the sequence is maximized when there is agent heterogeneity. If $\theta(x, \delta) = \delta x$ then the expected utility of imitating action x is.

$$E(x, \delta) := \frac{\theta(x, \delta) + h(x, \delta)}{2} = \frac{1 - \delta}{2} + \delta x.$$

As $\theta(x, \delta)$ is less than or greater than δx the expected utility of imitating action x is less than or greater than $E(x, \delta)$. Proposition 2 demonstrates that when the expected utility of imitating is $E(x, \delta)$ or lower then agents do best if they are identical (and $\delta = 1$). The costs of heterogeneity outweigh the benefits. If there is agent heterogeneity then an agent who imitates someone who has adopted a high utility action expects to choose an action that yields him a relatively low payoff. If the expected utility of imitating is strictly greater than $E(x, \delta)$ then some agents do best if there is heterogeneity (and $\delta < 1$). The potential benefits of heterogeneity now outweigh the costs. Indeed, the method of proof for Proposition 2 provides another interpretation. If we fix a $\delta < 1$ so that some $\mu^{\delta} \in (w, 1)$ exists and thus lock in is possible then there will exist some i^* such that agents $i \geq i^*$ do strictly better than if $\delta = 1$. Thus, it is not that we need 'pick δ carefully', basically, as long as lock in is possible, all agents sufficiently far into the sequence do better with heterogeneity. Note, however, that Agent 2 always does best if $\delta = 1$ and agents $2, 3, .., i^* - 1$ also do best if $\delta = 1$.

Proposition 2 says nothing about the value of i^* or δ^i for $i > i^*$. Especially given that our proof will be one of limiting arguments this may lead to questions whether i^* can be small and or δ^i small. That is, will agents early in the sequence desire significant heterogeneity. We can address this point by looking at a limiting example. If $\theta(x, \delta) + h(x, \delta) = 2x$ then the expected utility of imitating action x is x. In terms of the earlier discussion the costs to heterogeneity are minimized.

Proposition 3: If $\theta(x, \delta) = x - (1 - \delta)(1 - x)$ for all $x \geq 0.5$ and δ then $\delta^i < 1$ for all $i \geq 3$ and $\lim_{i \to \infty} \delta^i = 0$.

The proof of Proposition 3 proceeds by showing that agent 3 would prefer $\delta = 0$ to $\delta = 1$. While this does not preclude that δ^3 is still close to 1 it does illustrate how significant heterogeneity, even if not necessarily optimal, need not be undesirable. Ultimately, we observe that agents prefer maximal heterogeneity.

Finally, we can comment on expected payoffs. Let U_i^δ denote the expected payoff of agent i. We have already defined U_1 (which is independent of δ) and U_i^1. Comparing U_1 and U_i^1 we see that social learning leads to a significant rise in expected payoffs.

Proposition 4: If $\delta < 1$ then $U_{i+1}^\delta > U_i^\delta$ for all $i \geq 2$. If $\delta^i < 1$ then $U_i^{\delta_i} > U_i^1$.

The first part of the Proposition demonstrates that agents do best the further they are into the sequence. This is an illustration of how there can be accumulated learning even though each agent only observes the agent previous to him in the sequence. This, in itself, does not tie down the relative size of payoffs but the second part of the Proposition does so. If $\delta = \delta^i < 1$ then all agents $j \geq i$ do strictly better given agent heterogeneity than they would do given $\delta = 1$. This does not rule out some agents $j < i$ also doing better.

To summarize our results. We find that social learning can significantly increase expected payoffs. This, in itself, is not surprising as social learning allows agents to 'free-ride' and avoid 'costly search'. We find, however, that when there is agent heterogeneity imitation may actually do better. It not only means that agents avoid the cost of search it also means that they can expect to adopt a higher utility action than if they had searched themselves. Specifically, we know that agent 1 (who cannot use social learning) has expected payoff wR where $R := (1+r)/r$. If $\delta = 1$ then agents $2, 3, ...$ have expected payoff $(1+w)R/2$. The benefit in being 'able to free ride' and avoid search could thus be measured by $(1-w)R/2$. Agent heterogeneity means that social learning can result in even higher benefits. If $\theta(x,\delta) + h(x,\delta) = 2x$ for example, then, applying Proposition 3, we see that $\lim_{i\to\infty} U_i^{\delta_i} = (w+3)R/4$ and so agent heterogeneity can result in an additional, significant, gain of $(1-w)R/4$. The consequence of this is that an agent may prefer a setting where he will observe someone with different preferences to himself. Interpreting our results another way an agent would be willing to pay to observe some other agent. If $\delta = 1$ then we see that an agent would be willing to pay $(1-w)R/2$. If $\delta < 1$ and there is agent heterogeneity an agent may be willing to pay up to $(1-w)3R/4$.[8] This illustrates

[8] Note that, given a choice, an agent does best to observe the agent who most recently adopted an action. Also, if observing the action of an agent is costly, then it may be optimal for an agent to observe only one previous agent (even if he does not adopt the imitated action). This provides some motivation for the model considered in this paper.

the potential gains from social learning and agent heterogeneity.

4 Generalizations and extensions

Our analysis essentially relies on four factors: (1) Lock in can occur despite heterogeneity of preferences. (2) Lock in occurs on the actions that yield highest utilities. (3) An agent only observes the action adopted by the agent previous to him in the sequence. (4) Agents are homogenous in their sampling distribution. With these four factors the analysis would easily generalize because heterogeneity would act to weed out low utility actions and enable lock in on high utility actions. We should, however, question what happens in the absence of these four factors.

Lock in would not occur if with positive probability $u_i(x) < w$ for all i and all x. This could result from an enlargement in the support of the distribution over $u_i(x)$. For example, it may be that most agents like a restaurant of a very high quality (even if preferences do vary) but occasionally there is some 'extreme' agent who would dislike the restaurant. This should make little difference provided that the probability $u_i(x) < w$ is small. Lock in would no longer be possible but the expected loss from 'extreme type agents' would be small (because extreme type agents are rare) and compensated for by the gains to agent heterogeneity. If, however, the probability that $u_i(x) < w$ is large then preference heterogeneity may be undesirable. For example, suppose that there are two 'types of agent', say, a Chinese food type and an Indian food type where a high quality Chinese restaurant is not liked by those who prefer Indian food and vice versa. If proportion δ of agents are of one type and proportion $1-\delta$ of the other type then preference heterogeneity will be undesirable. This is because heterogeneity would stop lock in from occurring but not lead to any significant gains elsewhere.

Relaxing factor (2) would likely change our conclusions because heterogeneity in this case would act to weed out the high utility actions with lock in on actions yielding 'just sufficient utility'. Intuitively, factor (2) does not seem unreasonable so this qualification seems a minor one. This issue can also be related to our existing results and analysis. We find, Proposition 2, that if heterogeneity decreases the expected utility of imitation by enough then heterogeneity may no longer be beneficial. This could be interpreted as saying that lock in does not occur on high utility actions.

In relaxing factor (3) suppose that an agent can observe two or more other agents and possibly the whole sequence. This would matter if there exists some agent $i-1$ where agents $1,...,i-2$ adopt action b_1 but agent $i-1$ 'rejects' action b_1, searches and adopts action b_2. In the model we have analyzed agent i only observes action b_2. Observing b_1 as well as b_2 can clearly be of no disadvantage to agent i but presumably is to his benefit. In particular, he can imitate both b_1 and b_2 and can adopt the action yielding highest utility.[9] This reduces the probability that he would have to sample and means that the 'accumulated learning' by agents $1,..,i-2$ is not lost. Here heterogeneity has a further benefit in creating multiple imitation opportunities. Less immediate in this setting is the optimal strategy that agents should use. Note that it is also not critical for our results that each agent observes the agent immediately previous to him in the sequence. We could, for instance, assume that an agent observes the action adopted by one agent randomly selected from those previous in the sequence.

The consequences of relaxing factor (4) are more complex. If agents have different sampling distributions (known to them) then each agent has a different switchpoint. Agents with a 'poor' sampling distribution would have a low switchpoint and be more inclined to adopt the action imitated. Agents with a 'good' sampling distribution would have a high switchpoint and thus be inclined to sample even if the imitated action yields a high utility. Lock in need not occur in this setting because agents with a good sampling distribution need 'never be satisfied' but this is not necessarily a bad thing. In particular, lock in will occur for those agents with a poor sampling distribution in the sense that they always adopt the imitated action rather than sample. If there are agents with a good sampling distribution who push the expected utility of any adopted action higher then this should be beneficial. We only have to check whether those agents with a good sampling distribution benefit, given their increased chance of having to search. Agents only search, however, because it increases their payoff and an agent with a good sampling distribution could behave, if she wished, the same as an agent with a poor sampling distribution.

[9] If imitation is costly he may not opt to do this on the basis that $b_1 = b_2$ with positive probability.

5 Discussion and related literature

In the setting of this paper imitation proves advantageous for agents. That social learning could allow agents to avoid the costs associated with search is trivial. We find, however, that when there is agent heterogeneity social learning actually does better. It not only means that agents avoid the cost of search it also means that they can expect to adopt a higher utility action than if they had searched themselves. This can be contrasted with much of the literature on social learning, particularly that on herding and information cascades, where it is typically shown that social learning can lead to inefficiencies. In herding models social learning can prove inefficient if agents ignore private information and follow the herd (Banerjee 1992, Bikhchandani, Hirshleifer and Welch 1992, 1998). The reason that social learning proves efficient in our model is the possibility of 'rejecting advice'. In our model an agent can imitate, gain full information about an action and then subsequently search. In models of herding an agent has only one opportunity to use an action. So in herding models, unlike the model of this paper, an agent cannot learn from experience. This leads to a contrast in the desirability of lock in. In our model lock in is 'good' because it indicates that a high utility action has been found. In herding models lock in can be 'bad' because it encourages agents to disregard informative signals. The crucial difference is whether agents have the possibility to 'act on advice' but subsequently 'reject the advice' if they find it was not successful for them. It does seem reasonable that in many contexts social learning need not lead to a binding commitment, like that assumed in herding models.

The importance of considering social learning in conjunction with agent heterogeneity is also illustrated by our results. For instance, (unobserved) agent heterogeneity creates a divergence between private signal and observed action that results in the type of issues familiar from the herding literature. Also, many models of social learning lead to 'lock in type stories' where lock in may or may not prove advantageous. Agent heterogeneity is one source through which a 'story of lock in' may change. In the model of this paper agent heterogeneity implies lock in on more preferable actions and thus proves advantageous. Related results are obtained by Goeree et. al. (2006) and Smith and Sorenson (2000). Both papers consider models of social learning with multiple agent types where agents with 'extreme types' guarantee that lock in cannot occur. This perturbs the dynamic and can result in 'complete learning' and an 'efficient outcome'. Note, however, that this result relies on 'extreme heterogeneity' where the set of agent types has full support. Our result is more easily interpreted in terms of 'slight heterogeneity' in agent preferences. Also related is the work of Banerjee and Fudenberg (2004). They consider a model of social learning where behavior converges to the optimum. Again, lock in on an inefficient equilibrium is avoided by the presence of agents who 'break the chain', but in this case it results from each agent only observing a sample of previous agents actions. This is similar to our model where the fact that only one previous action is observed implies that any amount of 'accumulated learning' can 'be lost'.

What can we learn with respect to agent behavior? Clearly, agents should imitate and let as many agents as possible act before they do. More interesting is to question whether or not the degree of heterogeneity is endogenous and if so what the implications are. The degree of heterogeneity may be endogenous if agents have some choice over who they observe and imitate. An agent may, for example, choose to observe only the agent he believes has the most similar preferences to himself. If so, then a tension may arise between the interests of the population and individual. As we have seen, an agent can do best if the population is heterogeneous. An agent also does best, however, if he imitates someone with identical preferences to himself. That is, if agent i had the choice then he would want agents $1, 2, .., i-1$ to be heterogenous but $i-1$ to be identical to himself. Intuitively, the only thing that i could control is his similarity with agent $i-1$. Hence, a tension between what is best for the population and best for the individual.

If search is conducted simultaneously amongst agents then free-riding or encouragement effects may emerge reflecting strategic concerns (McFadden and Train 1996, Bolton and Harris 1999, Bergemann and Valimaki 2000 and Keller, Rady and Cripps 2005). The framework of this paper was deliberately chosen to abstract away from such strategic issues in order to focus on the consequences of agent heterogeneity in a model of social learning. Our results do, however, point towards an incentive to free ride or delay 'commitment' as long as possible because the further back is an agent in the sequence the higher his utility. Our analysis would also point towards a more surprising result, the greater is agent heterogeneity then the more incentive to free ride because the greater are the gains from being later in the sequence. In particular with heterogeneity there are two motivations to free ride: to avoid search and to ultimately imitate a better action.

A related issue is that of endogenous sequencing. As is standard in models of herding we have assumed an exogenous sequencing of agents. Clearly, however, agents would care where in the sequence they are and this leads to questions of whether agents should be able to choose when to act (Gale 1996). This, as already highlighted in the previous paragraph, would lead to strategic incentives to delay acting as long as possible. Within our framework agents are ex-ante identical and so we do not gain much insight into which agents would have the least incentive to delay acting. Agent heterogeneity does, however, point towards potential consequences for the 'dynamics of sequencing' particularly if heterogeneity is observable in some way. In particular, we know that agents do best if they follow agents most identical to them but we also know that agents benefit from being later in the sequence. So, even if we fix a sequence in which all except one agent will act it is an interesting question to ask where this agent would choose to enter the sequence.

6 Proofs of the Propositions

Let $\beta(x)$ denote the probability that an agent i would adopt action x, i.e. $\beta(x) = \Pr\{u_i(x) \geq w\}$. Note that this is not agent specific. Also, note that $\beta(x) = 1$ if and only if $x \geq \mu^\delta$ for some lock in threshold $\mu^\delta \in [w, 1]$.

Lemma 1: If $\delta < 1$ then β is a continuous, non-decreasing function of x.

Proof: If $x \geq \mu^\delta$ then $\beta(x) = 1$. If $x \in [0.5, \mu^\delta)$ then

$$\beta(x) = \frac{h(x, \delta) - w}{h(x, \delta) - \theta(x, \delta)}.$$

Comparing, $\mu^\delta > x' > x \geq 0.5$ and using Assumption 4 we get that $h(x', \delta) - \theta(x', \delta) \leq h(x, \delta) - \theta(x, \delta)$. Clearly, $h(x', \delta) \geq h(x, \delta)$. Thus, $\beta(x') \geq \beta(x)$. If $x < 0.5$ then

$$\beta(x) = \frac{\theta(x, \delta) - w}{\theta(x, \delta) - l(x, \delta)}.$$

Comparing $0.5 > x' > x \geq 0$ and using Assumption 4 we get $\theta(x', \delta) \geq \theta(x, \delta)$. If $\theta(x, \delta) - l(x, \delta) \geq \theta(x', \delta) - l(x', \delta)$ then $\beta(x') \geq \beta(x)$. Otherwise $\theta(x', \delta) - \theta(x, \delta) > l(x', \delta) - l(x, \delta) = \delta(x' - x) > 0$. But given that $w > 0.5$

this also implies $\beta(x') \geq \beta(x)$.[10] Finally, we can compare $x' = 0.5 > x$. Using Assumption 4 we get that $\theta(x', \delta) = \delta/2 = \lim_{x \to 0.5, <0.5} l(x, \delta)$ and $h(x') = 1 - \delta/2 = \lim_{x \to 0.5, <0.5} \theta(x, \delta)$. Thus, $\beta(x') = \lim_{x \to 0.5, <0.5} \beta(x)$. This shows that β is an non-decreasing function of x. Continuity is apparent from the continuity of both upper and lower limits and that $\delta < 1$ (and so there exists x where $\beta(x) \in (0, 1)$).■

Let π_i and Π_i denote the ex-ante probability distribution over the action that agent i will adopt (as viewed from period $t = 0$). That is, $\Pi_i(x) = \int_0^b \pi_i(x)$ is the probability that agent i adopts action $x \in [0, b]$. [Existence of function π_i is demonstrated in Lemma 2.] Let

$$D_i := \int_0^1 \pi_{i-1}(x)\beta(x)dx$$

denote the ex-ante probability that agent $i \geq 2$ will adopt action b_{i-1}. That is, D_i denotes the probability that agent i will adopt the action of agent $i-1$. Again, this is ex-ante as viewed from period $t = 0$ and thus before b_{i-1} is known. Note that if $\delta < 1$ then $D_i < 1$ for all $i \geq 2$. For example, with positive probability agents $1, 2, .., i-1$ will adopt an action $x \in [w, w + \varepsilon]$ (for some small positive number ε). Using assumptions 2 and 3, with positive probability agent i will not adopt action x.

Lemma 2: If agents use the optimal strategy, $\delta < 1$ and there exists a lock in threshold $\mu^\delta \in [w, 1]$ then

1: π_i is a continuous, non-decreasing function of x for all i;

2: $\pi_i(x) = \pi_i(x')$ for all $x, x' \geq \mu^\delta$ and $i \in N$;

[10] The point is intuitively clear. To give a more formal argument. The 'loss in probability' is at least

$$Loss := \frac{\theta(x', \delta) - \theta(x, \delta)}{\theta(x', \delta) - \delta x'}.$$

The 'gain in probability' is at most

$$\begin{aligned} Gain \quad : \quad &= \left[\frac{1}{\theta(x, \delta) - \delta x} - \frac{1}{\theta(x', \delta) - \delta x'}\right]\left[\theta(x, \delta) - \frac{1}{2}\right] \\ &= \frac{[\theta(x', \delta) - \theta(x, \delta) - \delta(x' - x)]}{(\theta(x, \delta) - \delta x)(\theta(x', \delta) - \delta x')}\left[\theta(x, \delta) - \frac{1}{2}\right]. \end{aligned}$$

Using that $\theta(x, \delta) - \delta x > \theta(x, \delta) - \frac{1}{2}$ and $\delta(x' - x) > 0$ we obtain $Gain < Loss$.

3: $\Pi_i(x) \geq \Pi_{i+1}(x)$ for all x and $i \geq 2$;

4. $\Pi_i(\mu^\delta) > \Pi_{i+1}(\mu^\delta)$ for all $i \geq 2$

5: $D_i < D_{i+1}$ for all $i \geq 2$.

Proof: If $x \geq \mu^\delta$ then $\pi_1(x) = 1/(1-w)$. If $x < \mu^\delta$ then $\pi_1(x) = \Lambda\beta(x)$ where constant Λ satisfies

$$\Lambda \int_0^{\mu^\delta} \beta(x)dx = 1 - \frac{1-\mu^\delta}{1-w}.$$

By Lemma 1 we see that π_1 is a continuous, non-decreasing function of x. For any agent $i \geq 2$: if agent i adopts action b_{i-1} then his distribution over actions adopted is given by π_{i-1}. If agent i samples then his distribution over actions adopted is given by π_1. So,

$$\pi_i(x) = \pi_{i-1}(x)\beta(x) + (1-D_i)\pi_1(x). \tag{2}$$

From equation (2) and Lemma 1 it is simple to derive the five claims of the lemma by iteratively deriving π_i. That both π_{i-1} and β are continuous non-decreasing functions of x gives claims 1 and 3. The additional property that β must be strictly increasing over some range of x (because $\beta(x) = 1$ for $x > \mu^\delta$ and $\beta(x) < 1$ for $x < \mu^\delta$) gives claim 5. If claim 2 holds for π_{i-1} (and we have already demonstrated that it holds for π_1) then using $\beta(x) = 1$ for all $x \geq \mu^\delta$ we see that π_i satisfies claim 2. Using that $D_i < 1$ for all i, $\beta(x) = 1$ for all $x \in [\mu^\delta, 1]$ and $\Pi_1(\mu^\delta) > 0$ gives claim 4.■

Lemma 3: Agent 2's expected payoff is maximized when $\delta = 1$ and he imitates agent 1.

Proof: If agent 2 does not imitate agent 1 then his expected payoff is U_1 irrespective of δ. If $\delta = 1$ then agent 2 has an expected payoff of $U_i^1 > U_1$. So, agent 2 must imitate agent 1 to maximize his expected payoff. Recall that $\pi_1(x) = 1/(1-w)$ for all $x \geq w$ and $\pi_1(x) = 0$ for all $x < w$ if $\delta = 1$. If $\delta < 1$ then by Assumptions 1 and 2 $\pi_1(x) \leq 1/(1-w)$ for any $x \geq w$ and $\pi_1(x) > 0$ for some $x < w$. Further $E(x, \delta) \leq x$ for all x. This implies that $U_2^\delta < U_i^1$ for any $\delta < 1$. Thus, agent 2's utility is maximized when $\delta = 1$.■

Lemma 4: If $\delta = \delta^j$ for some j then any agent $i \geq 2$ maximizes his expected payoff by imitating agent $i - 1$.

Proof: If agent i does not imitate agent $i-1$ then his expected payoff is U_1. If $\delta = 1$ and agent i imitates then his expected payoff is $U_i^1 > U_1$. Now consider $\delta^j < 1$. We have immediately that $U_j^{\delta^j} \geq U_i^1$ and so agent j must imitate agent $j-1$. Given Claim 3 of Lemma 2 if it is optimal for agent i to imitate then it is optimal for every agent $l \geq i$ to imitate. We know that j will imitate and so it remains to consider agents $i < j$. Suppose that agent $j-1$ does not imitate. That agent j wishes to imitate agent $j-1$ implies, by symmetry, that agent 2 would want to imitate agent 1. But, if agent 2 imitates then all agents $i > 2$ would want to imitate. Suppose that agent $j-2$ does not imitate. That agent $j-1$ wishes to imitate agent $j-2$ implies that agent 2 would want to imitate agent 1. Again, we have a contradiction. Iterating the argument gives the desired result.■

Proof of Proposition 1: The optimal strategy for agent 1 is detailed by Kohn and Shavell (1974) but easily explained in this context. For the moment, without loss of generality set $u_1(x) = x$ for all x. To sample is the optimal strategy for agent 1 in period t if and only if the expected payoff from sampling in period t and then using experience thereafter is greater than the payoff of using experience in periods t onwards. That is, if b is the current best available action, when

$$\frac{1}{2} + \int_{x \leq b} \frac{b}{r} dx + \int_{x > b} \frac{x}{r} dx \geq \left(1 + \frac{1}{r}\right) b. \tag{3}$$

Rearranging, this becomes

$$\int_{x > b} \frac{x - b}{r} dx \geq b - \frac{1}{2} \tag{4}$$

The larger is b then the smaller is the LHS of (4) and the larger is the RHS. Thus, there exists a unique switchpoint w where agent 1 should sample if and only if $b \leq w$.[11] We have framed the discussion in terms of agent 1 but can clearly generalize to all agents. Consider agent $i \geq 2$. Agent i has an opportunity to imitate agent $i-1$. Clearly, there is no gain to agent i in delaying imitation and so he should either imitate or not in period $t_{i-1}+1$. As seen in Lemma 4, agent i should imitate. Imitation is a one shot opportunity and so agent i should sample or use experience in all subsequent periods.

[11] Clearly $w < 1$ and $w > E$.

The problem facing agent i is now identical to that of agent 1 and thus he should use experience or sample as $u_i(x)$ is greater than or less than w.■

Proof of Proposition 2: Set $\theta(x,\delta)=\delta x$ for $x\geq 0.5$. Then

$$E(x,\delta)=\frac{1-\delta}{2}+\delta x \tag{5}$$

and there exists lock in threshold $\mu^\delta = w/\delta$ when $\delta > w$. Suppose that $\delta > w$ and there is lock in. Given Lemma 2 and (5) the expected payoff of subsequent agents is

$$R\frac{\delta}{\delta-w}\int_{\frac{w}{\delta}}^{1}\left[\frac{1-\delta}{2}+\delta x\right]dx = R\frac{1+w}{2}. \tag{6}$$

Note that this expected value does not depend on δ. Indeed, this is the expected payoff if $\delta = 1$. If $\delta = 1$ then lock in will occur for sure after agent 1 and the expected payoff of all agents $i \geq 2$ is U_i^1. By Assumption 3 the expected payoff of an agent that imitates before lock in has occurred is strictly less than U_i^1. So, if $\delta < 1$ the expected payoff of any agent $i \geq 2$ is strictly less than U_i^1. Thus, $\delta^i = 1$ for all i. The same reasoning can be used if $\delta \leq w$ and lock in can thus not occur. Also, if $\theta(x,\delta) \leq \delta x$ for $x \geq 0.5$ and $\delta < 1$ then the expected payoff after lock is at most U_i^1 and so we can again use the same reasoning.

If $\theta(x,\delta) > \delta x$ for all $x \geq 0.5$ then

$$E(x,\delta) > \frac{1-\delta}{2}+\delta x \tag{7}$$

for all $\delta < 1$. Fix a δ and corresponding $\mu^\delta \in (w,1)$. Find the $\delta' > \delta$ such that $\mu^\delta = w/\delta'$.[12] If lock in occurs then, using claim 2 of Lemma 2 and equations (6) and (7), the expected payoff of subsequent agents is

$$R\frac{\delta'}{\delta'-w}\int_{\frac{w}{\delta'}}^{1} E(x,\delta)dx = U_i^1 - \varepsilon.$$

for some positive real number $\varepsilon > 0$. Note that this expected value can be fixed independent of i. If L_i denotes the probability that lock in has

[12] We have that $w=\theta(\mu^\delta,\delta) > \delta\mu^\delta$ and $\mu^\delta > w$ so there exists $\delta' > \delta$ such that $\delta'\mu^\delta = w$.

occurred before agent i must first choose an action then clearly $\lim_{i\to\infty} L_i = 1$. Suppose, as a worst case scenario, that an agent i has expected payoff U_1 if lock in has not occurred before he must choose an action. For i sufficiently large $(1 - L_i)U_1 < \varepsilon$ and $U_i^\delta > U_i^1$. Noting that if $\delta = 1$ an agent's expected payoff is U_i^1 we find, as desired, that there exists some i^* such that $\delta^i < 1$ for all $i \geq i^*$.■

Proof of Proposition 3: Fixing a δ and setting $\theta(\mu^\delta, \delta) = w$ we have that $\mu^\delta = (w + 1 - \delta)/(2 - \delta)$ and $E(x, \delta) = x$. So, if lock in occurs then the expected utility of subsequent agents is

$$EU_\delta = \frac{2-\delta}{1-w}\int_{\frac{w+1-\delta}{2-\delta}}^{1} x dx = \frac{w+3-2\delta}{2(2-\delta)}$$

The value of EU_δ is maximized when $\delta = 0$ because

$$\frac{dEU_\delta}{d\delta} = \frac{w-1}{2(2-\delta)^2} < 0.$$

Note that EU_δ and its differential are independent of i. We know that the probability of lock in L_i tends to one as i tends to infinity provided that $\mu^\delta < 1$. Thus, δ^i tends to 0 as i tends to infinity.

Consider agent 3 and suppose that $\delta = 0$. We shall show that agent 3's expected payoff is greater than U_i^1. Setting $\delta = 0$ we have that $\mu^0 = (w+1)/2$ and $E(x, \delta) = x$. So, if lock in occurs then the expected utility of subsequent agents is

$$EU_{\delta=0} = \frac{2}{1-w}\int_{\frac{w+1}{2}}^{1} x dx = \frac{w+3}{4}.$$

With probability 0.5 agent 1 will sample an action $b_1 \geq \mu^0$. Let $p = 2(1 - D_2)$ denote the probability that agent 2 samples *conditional* on $b_1 < \mu^0$. By claim 5 of Lemma 2 the probability that agent 3 will sample conditional on $b_2 < \mu^0$ is is strictly less than p. If $b_2 < \mu^0$ but agent 3 adopts b_2 then, using Lemma 2 and that $E(x, \delta) = x$, we know that the expected payoff of agent 3 is strictly greater than

$$R\frac{2}{1-w}\int_{w}^{\frac{w+1}{2}} x dx = R\frac{3w+1}{4}.$$

The probability that lock in has occurred (i.e. that $b_2 \geq \mu^0$) is $1/2 + p/4$. If lock in has not occurred then with probability of at most p agent 3 will sample and with probability at least $1-p$ he will adopt action b_2. So, denoting agent 3's expected payoff by $U_3^{\delta=0}$ we get

$$\begin{aligned}\frac{U_3^{\delta=0}}{R} &\geq \left(\frac{1}{2}+\frac{p}{4}\right)\left(\frac{w+3}{4}\right)+\left(\frac{1}{2}-\frac{p}{4}\right)\left[pw+(1-p)\left(\frac{3w+1}{4}\right)\right] \\ &= \frac{8w+8+p^2(1-w)}{16} > \frac{w+1}{2}.\end{aligned}$$

Thus, agent 3 gets a higher expected payoff if $\delta = 0$ than $\delta = 1$. Consequently $\delta^3 < 1$. Applying Proposition 4 we see that agents $i > 3$ also prefer $\delta = 0$ to $\delta = 1$.■

Proof of Proposition 4: The first part of the Proposition is an immediate consequence of claims 3 and 4 of Lemma 2. The second part of the Proposition is an immediate consequence of δ^i being the i that maximizes i's expected payoff.■

References

[1] Banerjee, A., (1992) 'A simple model of herd behavior', *The Quarterly Journal of Economics* 107: 797-817.

[2] Banerjee, A., and D. Fudenberg, (2004) 'Word-of-mouth learning', *Games and Economic Behavior* 46: 1-22.

[3] Bergemann, D. and J. Valimaki, (2000) 'Experimentation in markets', *Review of Economic Studies* 67: 213-234.

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

[5] 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.

[6] Bolton, P. and C. Harris, (1999) 'Strategic Experimentation', *Econometrica* 67: 349-374.

[7] Ellison, G. and D. Fudenberg (1995) 'Word-of-mouth communication and social learning', *Quarterly Journal of Economics* 110: 93-125.

[8] Gale, D., (1996) 'What have we learned from social learning', *European Economic Review* 40: 617-628.

[9] Goeree, J., T. Palfrey and B. Rogers, (2006) 'Social learning with private and common values', *Economic Theory* 28: 245-264.

[10] Jovanovic, B. and G.M. MacDonald, (1994) 'Competitive Diffusion' The Journal of Political Economy 102: 24-52.

[11] Keller, G., S. Rady and M. Cripps, (2005) 'Strategic experimentation with exponential bandits' *Econometrica* 73: 39-68.

[12] Kohn, M.G. and S. Shavell, (1974) 'The Theory of Search', *Journal of Economic Theory* 9: 93-123.

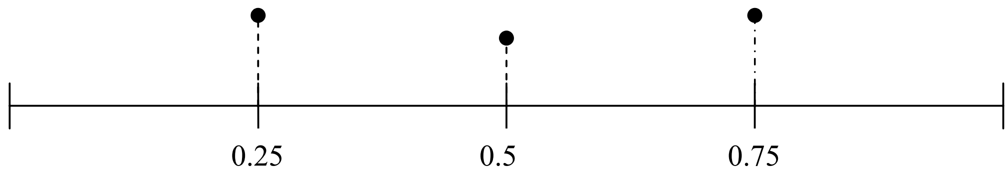
[13] McFadden, D. and K. Train, (1996) 'Consumers' evaluation of new products: learning from self and others', *Journal of Political Economy* 104: 683-703.

[14] Rosenfield, D.B. and R.D. Shapiro, (1981) 'Optimal adaptive price search', *Journal of Economic Theory* 25: 1-20.

[15] Smith, L. and P. Sorenson (2000) 'Pathological models of learning', *Econometrica* 68: 371-398.

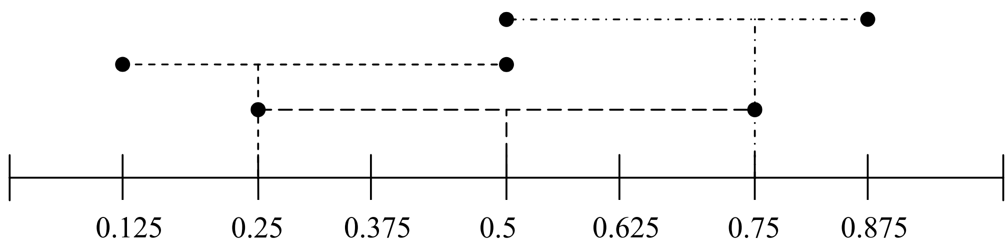
[16] Stigler, G.J., (1961) 'The economics of information', *Journal of Political Economy* 213-225.

Figure 1: The consequences of payoff heterogeneity:
Figure 1a: If δ = 1 then the utility of action x is x.

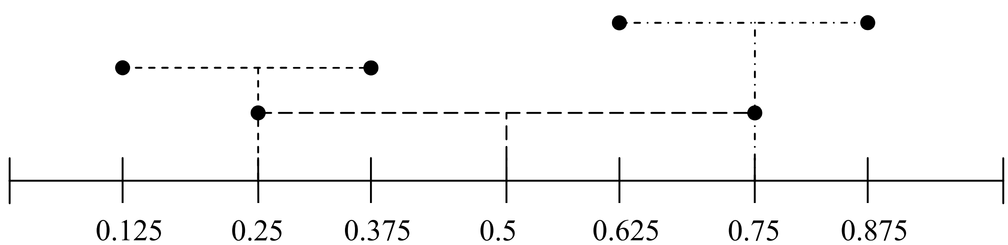


Figures 1b and 1c: If δ = 0.5 then the utility of action x < 0.5 could be anything in the interval [0.5x, θ(x, δ)] and the utility of action x > 0.5 could be anything in the interval [θ(x, δ), 1- 0.5x].

Figures 1b: Low expected utility; θ(x, δ) is relatively high for x < 0.5 and low for x > 0.5.



Figures 1c: High expected utility; θ(x, δ) is relatively high for x > 0.5 and low for x < 0.5.



Figures 1d: : If δ = 0 then the utility of action x can vary over the entire unit interval.

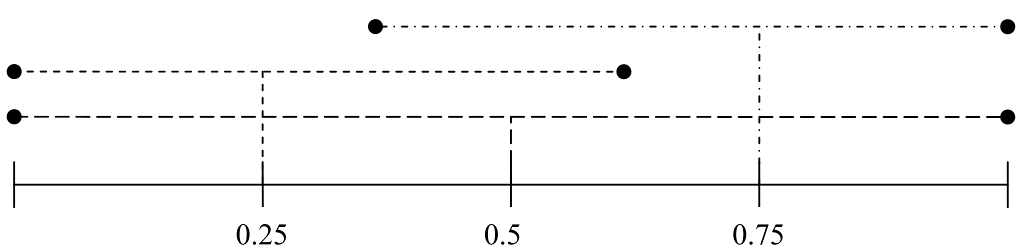
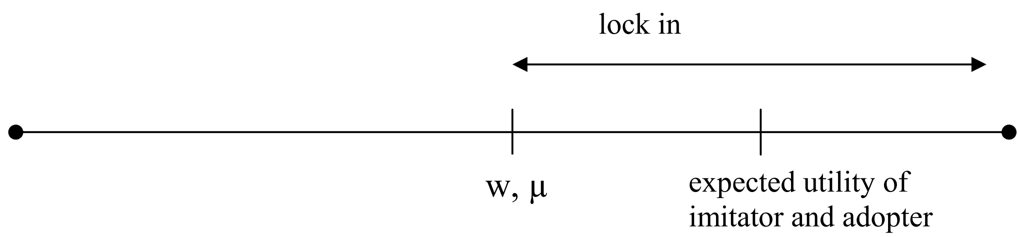


Figure 2: The consequences of heterogeneity

Figure 2a: If $\delta = 1$ then lock in occurs on any action greater than w and the expected utility of the adopted action is (w + 1)/2.



expected utility of adopter and imitator

Figure 2b: Heterogeneity is bad. For $\delta < 1$ lock in occurs on an action greater than w so the expected utility of an adopted action is $(\mu + 1)/2 > (w + 1)/2$ but the expected utility of an imitator is lower than (w + 1)/2.

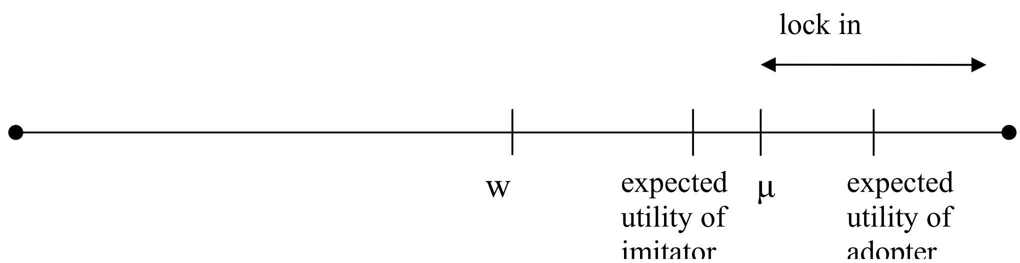


Figure 2c: Heterogeneity is good. For $\delta < 1$ lock in occurs on an action greater than w so the expected utility of an adopted action is $(\mu + 1)/2 > (w + 1)/2$. The expected utility of an imitator is lower than $(\mu + 1)/2$ but greater than (w + 1)/2.