

Social learning with subjective communication and self-selection*

Péter Kondor

London School of Economics

e-mail: p.kondor@lse.ac.uk

Gergely Ujhelyi

Harvard University

e-mail: ujhelyi@fas.harvard.edu

May 9, 2005

Abstract

We study social learning about a new product in a model with subjective communication (influenced by agents' beliefs) and self-selection of social contacts. Self-selection implies that consumers will be disappointed on average. Subjective communication implies that the messages sent will tend to be negative. In this model, unless the quality advantage of the new product is large enough, learning causes the set of consumers choosing the new product to shrink over time. Thus, learning may help efficiency by driving out inferior new products, but can hurt efficiency by reducing the consumption of new products that are superior. We show that learning in more "diverse" neighborhoods implies more consumption of the new product. However, these neighborhoods are more receptive to new products regardless of their quality. A number of extensions are studied, including more accurate communication, communication with subjective thresholds, and less accurate consumption experience. The results provide a learning-based explanation of why some communities seem resistant to new products or ideas.

*We thank Drew Fudenberg, Michael Kremer, Markus Möbius, József Molnár, Nathan Novemsky, and participants of the Theory and Development lunches at Harvard and the Whitebox Advisors conference on Behavioral Economics at Yale for comments and suggestions. Péter Kondor gratefully acknowledges the EU grant "Archimedes Prize" (HPAW-CT-2002-80054).

1 Introduction

We learn from the experience of others every day. We ask our friends about which movie to see and our colleagues about which seminar to attend; we observe strangers buying certain products rather than others, and we get exposed to publicly provided information. Various features of such *social* forms of learning have been studied by an extensive literature.¹ Two features that are perhaps among the most fundamental characteristics of social learning but that have not been studied are subjective communication and self-selection among one's social contacts.

Subjective communication. Explicit communication between agents is, in many settings, the most natural way of exchanging payoff-relevant information. Yet, part of the literature disregards communication altogether, and only considers learning from observing others' actions.² Some of the studies that do consider communication define it as the (possibly imperfect) observation of *payoffs*.³ Our view is that what distinguishes communication from direct observation is not their subject (payoffs or actions could both be observed directly or learned from word-of-mouth⁴) or their precision (direct observation and communication can both be imprecise). Rather, an important difference between the two is that communication is inherently subjective, in the sense of being influenced by individuals' beliefs and expectations. For example, a disappointed customer is unlikely to recommend a product, even if the disappointment was due to her excessively high expectations.⁵ Our model of social learning will incorporate this feature of communication.

Self-selection. When we want to learn from others' experience, we usually ask someone who has the relevant experience (or who we think might have it). Thus, when we want to know if a seminar is worth going to, we might ask someone who has already heard the speaker, has read the paper, or knows the catering company the organizers order from. But having or not having a particular experience is often the consequence of the choices one makes. For example, our colleague might have chosen to read the paper because he thought it would be good, because it was related to his work, or because the speaker is his friend. Thus, the set of agents with relevant experience that one might consult tends to be a self-selected sample from the population.⁶

¹See Chamley (2004) for a detailed survey.

²One reason that is sometimes given in favor of studying the observation of actions rather than "words" is that the latter is non-credible (cheap talk). While this can be true between competing producers or between a politician and the public, in many situations the parties lack strategic incentives, and the information they provide is credible. Friends, relatives, or even complete strangers usually have little incentive to misrepresent the information one might ask them to provide in every-day life.

³Ellison and Fudenberg (1993, 1995) assume observable payoffs. Fudenberg and Banerjee (2004) allow for direct observation of actions and a signal correlated with payoffs.

⁴In some situations, payoffs are observable without explicit communication (e.g., oil drillers may be able to see if oil was found on a neighboring track, and might have a pretty good idea about the associated payoffs, as in Hendricks and Kovenock (1989)). In other situations, actions are not observable unless communication occurs (e.g., if the relevant action was taken in the past).

⁵Januszewski (2004) provides evidence that in the airline industry "customer complaints are largely driven by disappointed consumers who received worse service than they expected" (23). The importance of having more realistic models of communication in social learning was also emphasized by Shiller (1995).

⁶When agents do not have access to complete histories (or their summary statistics), the literature usually assumes that contacts are a representative sample from the population. Banerjee and Fudenberg (2004) consider a biased rule which oversamples rare choices - this is essentially the opposite of the self-selection logic. Manski (2004) studies the selection problem that arises when only a subset of possible actions are chosen by the population.

As we show in this paper, the interaction of subjective communication and self-selection has important consequences for social learning, and can yield some surprising results. In particular, our model provides a learning-based explanation of why certain communities seem reluctant to embrace new products, technologies, or ideas.⁷

We study a common-value setting in which agents face a choice between an "Old" product with known quality and a "New" one whose quality they do not know. Agents base their decision on their opinions, which are private information, but which are, at least initially, an unbiased estimator of the true quality of the new product. In such a setting agents who decide to consume the new product overestimate its quality on average. Thus, they experience a negative "post-decision surprise" (Harrison and March, 1984).⁸ In the second round, a new set of agents has a chance to learn from this self-selected group of experienced agents before making their own consumption choices. Learning is Bayesian. We assume that communication is subjective in that it is influenced by the "surprise" of the experienced agents (the difference between their initial opinion and the consumption experience). Because surprise is on average negative, negative messages will be overrepresented in the communication. We study the resulting evolution of opinions and consumption choices in a benchmark model and several extensions.

In the benchmark model of Section 2, we make three strong assumptions. We assume that consumption completely reveals the product's quality, but that communication is limited to the sign of an experienced agent's surprise. Most importantly, we assume that an agent and his contacts share identical initial opinions. In this setting, social learning gradually eliminates the surprise in the population, so that on average consumers will eventually form correct opinions about the new product's quality. The main result of this section is that unless the new product is far superior to the old one, the negative messages generated in social learning will unambiguously reduce the measure of agents choosing the new product. Thus, social learning helps efficiency for new products that are inferior to the old one, as it speeds up their rejection by the population. When the new product is highly inferior to the old one, even limited subjective communication can transmit enough information to drive it out of the market immediately. However, learning also reduces the consumption of superior new products, and thus hurts efficiency in that case. If the new product has only a small quality advantage, its consumption might cease completely.

In Section 3, the benchmark model is compared to a setting where agents have contacts within a "neighborhood" i.e. with agents who have similar but not necessarily identical initial opinions to their own. In this model, when a receiver gets a message, she will not know the sender's initial opinion exactly, but she will know that the sender is in the neighborhood. Naturally, since messages are less informative in this setting, we find that opinions are slower to adjust towards the truth. As we show,

⁷In many social sciences, explanations of such resistance to novelty are centered on cultural constraints (see, e.g. Adongo et al. (1997) for "cultural factors constraining the introduction of family planning" in Ghana). Munshi and Myaux (2002) provide an economic explanation along these lines.

⁸This result is related to the Winner's curse and similar phenomena. See Van den Steen (2004) for a discussion. Israel (2004) finds empirical evidence from the automobile insurance industry that "consumers who choose to join the firm tend to be excessively optimistic about its quality and thus disappointed by actual experiences", which, he notes, is "consistent with the idea that, while the average consumer has an accurate prior about the firm's quality, those consumers who choose to join a firm are the most optimistic about its quality" (p4).

this result has the striking implication that in larger neighborhoods, the measure of consumers of the new product is also larger, regardless of whether the new product is better or worse than the old one. Thus, more "diverse" communities seem more receptive to the New, be it good or bad. The intuition for this result comes from four ways in which learning from neighbors is different from learning from oneself (or from someone similar to oneself). The *experienced-neighbor* effect occurs when people learn from others who have tried a product which they themselves would not have tried. The *misleading-neighbor* effect means that a neighbor might send the opposite message than what the individual would have sent to herself. The *inexperienced-neighbor* effect operates when someone does not learn the bad news about an inferior new product because her neighbor did not try it. Finally, the *learning-from-neighbors* effect results from the fact that the disappointing messages are discounted when coming from people who are different from oneself. As we show, these four effects together imply that more diverse neighborhoods are more favorable to the diffusion of both inferior and superior new products.⁹

Sections 4, 5, and 6 provide some robustness checks on the model. In Section 4, we allow for more precise communication, and show that while the effects are smaller, the basic findings generalize. In Section 5, we relax the assumption that consumption fully reveals the quality of the product. We show that an imprecise signal of quality still results in messages that are negative in expectation, but the bias becomes less severe as we increase the variance of the signal. In the final extension of Section 6, we consider a slightly different communication technology. Specifically, we allow agents to communicate their actual experience, so that initial expectations do not influence the content of the messages, but we assume that messages will be sent only if there was a sufficiently large surprise. Thus we have selection effects both in participation and communication. We show that our finding of disproportionately negative messages survives in this environment as well.

2 The benchmark model

2.1 Self-selection and disappointment

In this section we describe a benchmark model of social learning by consumers who have *chosen* to consume a new product. There is a continuum of agents, uniformly distributed over the unit interval. Every agent can choose between an "old" product with fixed and known quality c and a "new" product with some fixed but unknown quality θ .¹⁰ We define the surplus that the new product represents relative to the old one as $\delta \equiv \theta - c$, which is unknown to the agents. Ideally, agents would like to choose the new product if and only if $\delta > 0$.

Before making a choice, agent i observes a private signal x_{0i} of the quality θ , where $x_{0i} = \theta + \varepsilon_{0i}$ and $\varepsilon_{0i} \sim U(-\frac{1}{2}, \frac{1}{2})$. We will refer to agents with $x_{0i} > \theta$ and $x_{0i} < \theta$ as "optimists" and "pessimists", respectively. After observing the private signal x_{0i} , the agent's belief about θ is uniform on $[x_{0i} -$

⁹See Bala and Goyal (1998) for different implications of learning from neighborhoods.

¹⁰A more consistent formulation would assume that agents have improper priors on θ , as in the global games literature (see Morris and Shin (2002) for an overview). The distinction will not matter here.

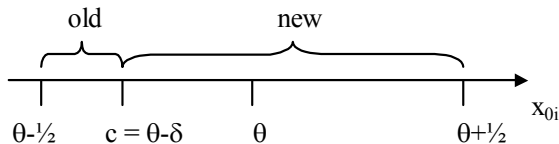


Figure 1: The benchmark model

$\frac{1}{2}, x_{0i} + \frac{1}{2}]$, with mean x_{0i} . In what follows, we will need to distinguish between expectations formed by an agent who does not know θ , and those formed by an outside observer who knows θ . Because the agent has less information, we will call expectations conditional on his information set "opinions", denoted o_i . Thus, we write $o_i(\theta|x_{0i}) = x_{0i}$, and the agent will choose the new product as long as $x_{0i} > c$. Figure 1 illustrates.

If the agent decides to consume the new product, she receives a second signal, x_{1i} . In this section, we make the assumption that experience fully reveals the quality of the product.

Assumption 1 (Experience is fully revealing) $x_{1i} = \theta$.

Under Assumption 1, after consumption, the agent updates her opinion to $o_i(\theta|x_{0i}, x_{1i}) = \theta$.¹¹ We define the "surprise" of an agent i as the difference between her opinion after and before consumption:

$$s_i \equiv o_i(\theta|x_{0i}, x_{1i}) - o_i(\theta|x_{0i}) = \theta - x_{0i}. \quad (1)$$

Definition (1) is consistent with the notion that (Bayesian) agents should not expect to be surprised, in other words, their opinion about their surprise is 0. Indeed, from (1), we get $o_i(s_i|x_{0i}) = o_i(\theta|x_{0i}) - x_{0i} = 0$. However, whether (from the outside observer's point of view) *expected* surprise equals 0 is a different question. If $\delta \geq \frac{1}{2}$, then it is true that $E(s_i|x_{0i} > c, \theta) = \theta - E(x_{0i}|x_{0i} > c, \theta) = 0$. However, if $\delta \in (-\frac{1}{2}, \frac{1}{2})$, then $E(x_{0i}|x_{0i} > c, \theta)$ is easily computed to be $\frac{1}{2}(c + \theta + \frac{1}{2})$. This expression is greater than θ , implying that the expected surprise of an agent will be negative:

$$E(s_i|x_{0i} > c, \theta) = \theta - E(x_{0i}|x_{0i} > c, \theta) = \frac{1}{2}(\delta - \frac{1}{2}) < 0. \quad (2)$$

If only those with sufficiently high initial opinions consume a good, on average consumers will be disappointed. This is simply because those with the lowest initial opinions, who could have a positive experience, choose not to consume. Note that a higher surplus δ means that more agents will consume, implying less self-selection. Average disappointment decreases with δ , and becomes 0 when $\delta = \frac{1}{2}$. Thus, the surplus δ determines the magnitude of the selection effect. We will be interested in the empirically relevant case when $\delta \in (-\frac{1}{2}, \frac{1}{2})$, so that some agents choose the new product, while others pick the old one.

¹¹The agent's initial opinion $o_i(\theta|x_{0i})$ is x_{0i} , even if she takes into account the additional signal she might get if she consumes the good, as $E[E(\theta|x_{0i}, x_{1i})|x_{0i}] = x_{0i}$ by the law of iterated expectations.

The rest of the paper examines the consequences of the simple observation captured in (2) for social learning.

2.2 Communication and learning in the second round

Now let us assume that there is a second group of agents, J , without any information. Each of them knows some agents from the first group, I , and can talk to them. Thus, before making his consumption decision, agent j observes messages m_{ij} , which may provide information on θ , as well as her initial opinion x_{0j} .

As discussed in the Introduction, this paper extends previous literature on social learning by assuming that the messages m_{ij} are subjective in the sense of being influenced by initial expectations or opinions, as well as the consumption experience itself. The notion that communication is influenced by initial expectations (disappointment and positive surprises) is both intuitive and supported by a large body of literature. First, disappointing consumption experiences clearly occur across a wide variety of products and settings. Israel (2004) shows that customers switch automobile insurance companies as their expectations become disappointed at their original insurer. Miguel and Kremer (2004) show that disappointing experiences with deworming drugs in Kenya reduced take-up in the population. Second, that consumers' *reported* satisfaction level with a product is a function of both the actual experience and prior expectations, is a well-documented fact in consumer psychology. See Oliver (1980) for a classic study and survey of the early literature on the "disconfirmation paradigm".¹² In a recent meta-analysis, Szymanski and Henard (2004) conclude that disconfirmation of expectations is one of the strongest determinants of consumer satisfaction. In the economics literature, a recent paper by Januszewski (2004) provides evidence that formal customer complaints in the airline industry are explained by the difference between expected and actual quality, rather than actual quality per se.¹³ Third, more generally, a number of studies in the anthropology literature show that communication about new products or technologies is subjective and limited, especially when the adopting community lacks the language to talk about specific attributes. On the diffusion of modern contraceptive practices in Northern Ghana, Adongo et al. (1997) write: "The popular term for family planning is *adog-maake*; a literal translation of this means "to have enough births," a notion that most men and women greet with antipathy or fear. Few women have a sophisticated understanding of what this term actually means, how methods work, or what contraceptives look like, apart from a vague belief that contraception is a powerful medicine that prevents birth." (p 1799).¹⁴

¹²Oliver (1980) considered a federal flu vaccination program, and found that consumers' reported satisfaction levels as well as their expressed intention to re-purchase the vaccine was a function of "disconfirmed expectations".

¹³The notion that satisfaction is affected by prior expectations is directly related to the theory of reference points, as in Kahnemann and Tversky's (1979) prospect theory. See Köszegi and Rabin (2004) for an economic model in which prior expectations form the reference point.

¹⁴Adongo et al. (1997) go on to write "The Kassem translation of the term family planning is *lorilao* meaning "to have a good birth." Although this term does not have the connotation of danger that is associated with *adog-maake*, it nevertheless emphasizes the birth process rather than the positive aspects of child spacing. Child spacing is a tradition that is respected, while contraception remains a subject that is embarrassing to divulge...". See also Kohler (1997) who cites several studies from the demography and anthropology literature showing that in developing countries, communication about contraceptives is limited and does not include several relevant details of women's experience with a given method. Bernardi (2003) discusses similar evidence from a developed country.

In this section, we model the notion of subjective communication in a simple manner, extensions are discussed in Sections 4 and 6. We assume that an experienced agent i can communicate one of three messages to an agent j from the second group. She either says that she had a good surprise ($m_{ij} = 1$ if $s_i > 0$), or that she had a bad surprise ($m_{ij} = -1$ if $s_i < 0$), or she says nothing ($m_{ij} = 0$ if $x_{0i} < c$).¹⁵

Assumption 2 (Limited subjective communication)

$$m_{ij} = \left\{ \begin{array}{l} 1 \text{ if } s_i > 0 \\ -1 \text{ if } s_i < 0 \\ 0 \text{ otherwise} \end{array} \right\}$$

In this model with limited subjective communication, the result in (2) implies that, on average, there will be more negative than positive messages sent. In particular, the probability that an experienced agent i asked at random would indicate that the product was a disappointment is given by

$$\Pr(m_{ij} = -1) = \Pr(s_i < 0 | x_{0i} > c, \theta) = \Pr(\theta < x_{0i} | x_{0i} > c, \theta) = \left\{ \begin{array}{ll} \frac{1}{1+2\delta} & \text{if } \delta > 0 \\ 1 & \text{if } \delta \leq 0 \end{array} \right\} \quad (3)$$

The probability of receiving a positive message is the complementary probability $1 - \Pr(m_{ij} = -1)$, implying that for every experienced contact, the expected value of the message going to an agent in the second group is negative:

$$E(m_{ij}) = \frac{2\delta - 1}{2\delta + 1} < 0 \quad (4)$$

if $\delta > 0$ (and $E(m_{ij}) = -1$ otherwise). The rest of the section deals with how these negative messages might affect consumers' opinions and consumption decisions in subsequent rounds.

Since unless agents in the second round have some extra information about first round agents' beliefs, learning about a surprise is totally uninformative (as it is simply the difference between two noise terms), we will have to specify what agents' know about their contacts. In the benchmark model, we assume that every agent has exactly one contact, with beliefs that are identical to his. Denoting $C(j) \subseteq I$ the set of contacts of an agent $j \in J$, we impose

Assumption 3 (Close contacts) $I = J$ and $C(i) = \{i\} \forall i$.

This assumption captures in an extreme manner the notion that one's social contacts tend to be similar to oneself (such as friends or village neighbors). See Section 3 for further discussion and extensions. Two cases in which this extreme assumption might be warranted are:

¹⁵Even this extreme formulation might not be that unrealistic. There is evidence in the psychology literature that limited messages or memories about the frequency with which a product has performed well guide consumer decisions. Alba and Marmorstein (1987) provide experimental evidence that consumers use the number of positive/negative attributes of a product as a decision heuristic, ignoring the magnitude or the relative importance of these attributes.

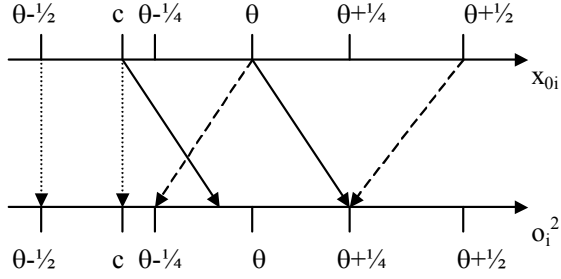


Figure 2: Adjusting opinions between rounds 1 and 2

(i) Agents are persons facing the same consumption decision over and over again, they would therefore like to recall the experience they previously had with the good. However, people’s ability to recall past memories can have several limitations. When “memories” m_{ii} only take on values from $\{-1, 1\}$, the agent only remembers her previous opinion (x_{0i}), and if she had a positive or negative surprise when she first consumed the good. This might be realistic if priors x_{0i} are determined by the environment (such as culture or family) and are therefore easier to recall than the signals x_{1i} related to infrequent consumption experiences.¹⁶

(ii) In the spirit of Manski (2004), an agent might represent a dynasty, defined by a prior x_{0i} which “runs in the family”. This interpretation assumes that children learn their parent’s surprise, but not their actual experience.

We now analyze the formation of opinions and the resulting consumption patterns under Assumptions 1-3. Upon receiving a message m_{ii} , a Bayesian agent who knows x_{0i} will believe that θ is uniformly distributed over the interval $[x_{0i} - \frac{1}{2}, x_{0i}]$ if $m_{ii} = -1$, or the interval $[x_{0i}, x_{0i} + \frac{1}{2}]$ if $m_{ii} = 1$. When $m_{ii} = 0$, no new information is learned. Thus we can write the agent’s initial opinion in round two as

$$o_i^2(\theta|x_{0i}, m_{ii}) = x_{0i} + m_{ii}\frac{1}{4} \quad (5)$$

where we use the superscript “2” to denote opinions formed in the second round. To simplify the notation, we will use o_i^2 without arguments to denote the opinion of the agent about θ conditional on whatever information (initial belief and messages) she has. Figure 2 illustrates the adjustment in opinions from round 1 to 2.

Note that while accurate conditional on the information of the agent, the opinions in (5) will generally be incorrect in expectation (from the outside observer’s point of view). In particular, for $\delta > 0$, there is no selection among the optimists, so (5) is accurate conditional on $m_{ii} = -1$: $E((o_i^2|_{m_{ii}=-1})|x_{0i} > c, \theta) = E(x_{0i}|x_{0i} > \theta, \theta) - \frac{1}{4} = \theta$ (see the dotted arrows in Figure 2), but for pessimists, $E((o_i^2|_{m_{ii}=1})|x_{0i} > c, \theta) = E(x_{0i}|\theta > x_{0i} > c, \theta) + \frac{1}{4} = \theta - \frac{1}{2}(\delta - \frac{1}{2}) > \theta$. For $\delta < 0$, no pessimist consumed the new product in the first round, but there is selection among optimists, so that

¹⁶See Lynch et al. (1988) for studies in the psychology literature which find that people remember (and use in their decisions) overall evaluations, rather than information about the specific attributes of a consumption experience.

$$E((o_i^2|_{m_{ii}=-1})|x_{0i} > c, \theta) = E(x_{0i}|x_{0i} > c, \theta) + \frac{1}{4} = \theta - \frac{1}{2}(\delta - \frac{1}{2}) > \theta.$$

In the second round, an agent will choose the new product iff $o_i^2 > c$. We are interested in how consumption choices and opinions change relative to the first round as a result of the messages transmitted. It is best to discuss separately the case of a superior ($\delta > 0$) and that of an inferior ($\delta < 0$) new product.

The new product is superior. When the surplus from the new product is relatively large, so that $\delta \geq \frac{1}{4}$, (5) implies that we have $o_i^2|_{m_{ii} \neq 0} > c$, and all those who consumed the new product in the first round will do so in the second round as well (this is the case shown in Figure 2). In this case, learning only affects the opinions of a fixed set of consumers.

How do these opinions compare to those in the first round? When $\delta \geq \frac{1}{4}$, expected surprise after second round consumption, $E(s_i^2|o_i^2 > c, \theta)$, can be computed as follows:

$$\begin{aligned} \theta - [E((o_i^2|_{m_{ii}=1})|x_{0i} > c, \theta) \Pr(m_{ii} = 1|x_{0i} > c, \theta) + E((o_i^2|_{m_{ii}=-1})|x_{0i} > c, \theta) \Pr(m_{ii} = -1|x_{0i} > c, \theta)] \\ = \left(\frac{\delta}{2} - \frac{1}{4}\right) \Pr(m_{ii} = 1|x_{0i} > c, \theta) = \frac{\delta}{2} \frac{2\delta - 1}{2\delta + 1} \end{aligned} \quad (6)$$

where we used the probabilities derived in (3). This expression is negative: on average there is still disappointment in the population. However, it is smaller in absolute value than (2) expected surprise in the first round. The information learned from the first round helped reduce the average distortion in opinions in the second round.

When δ is smaller than $\frac{1}{4}$, some consumers will consume the new product in the first round but not in the second. Note in particular that if a pessimistic agent consumed the new product in the first round, she will always consume it in the second round (since her opinion increases according to (5)). However, some of the optimistic agents will consume the new product in the first round but not in the second: those with x_{0i} between θ and $c + \frac{1}{4}$ will form opinions $o_i^2 < c$ and will switch to the old product.¹⁷ The measure of agents consuming in round two is given by

$$\Pr(o_i^2 > c) = \delta + \frac{1}{2} - (c + \frac{1}{4} - \theta) = 2\delta + \frac{1}{4}, \quad (7)$$

which is smaller than the measure of agents consuming in the first round, $\delta + \frac{1}{2}$. For new products that are "slightly" superior ($0 < \delta < \frac{1}{4}$), the negative messages arriving from experienced agents lead to a decrease in agents' willingness to consume in the second round.

To calculate the expected surprise of those who do consume, use¹⁸

$$\begin{aligned} \Pr(m_{ii} = 1|o_i^2 > c) &= \frac{\delta}{2\delta + \frac{1}{4}} \\ \Pr(m_{ii} = -1|o_i^2 > c) &= \frac{\delta + \frac{1}{4}}{2\delta + \frac{1}{4}} \end{aligned}$$

¹⁷Note that agents with x_{0i} above $c + \frac{1}{4}$ will again take the new product, implying that, in contrast to the first round, second round willingness to consume the new product is monotonic in the initial belief x_{0i} .

¹⁸To derive these expressions, note that $x_{0i}|(o_i^2 > c, m_{ii} = 1, \theta) \sim U[\theta - \delta, \theta]$ and $x_{0i}|(o_i^2 > c, m_{ii} = -1, \theta) \sim U[\theta - \delta + \frac{1}{4}, \theta + \frac{1}{2}]$.

to get

$$\begin{aligned} E(s_i^2 | o_i^2 > c) &= \theta - [E((o_i^2 |_{m_{ii}=1}) | o_i^2 > c) \Pr(m_{ii} = 1 | o_i^2 > c) + E((o_i^2 |_{m_{ii}=-1}) | o_i^2 > c) \Pr(m_{ii} = -1 | o_i^2 > c)] \\ &= \left(\frac{\delta}{2} - \frac{1}{4}\right) \frac{\delta}{2\delta + \frac{1}{4}} + \left(\frac{\delta}{2} - \frac{1}{8}\right) \frac{\delta + \frac{1}{4}}{2\delta + \frac{1}{4}} = \frac{1}{2} \left(\delta - \frac{1}{4} \frac{3\delta + \frac{1}{4}}{2\delta + \frac{1}{4}}\right). \end{aligned}$$

Again, one can check that $E(s_i^2 | o_i^2 > c) < 0$ so that there is disappointment on average, but it is smaller than in the first round.

The new product is inferior. When the surplus is negative (the new product is worse than the old one), so that $\delta \leq 0$, the set of consumers shrinks, as only negative messages are transmitted. In particular, when $\delta < -\frac{1}{4}$, no one will consume the new product in the second round, as the highest o_i ($= \theta + \frac{1}{2}$) is reduced to $o_i^2 = \theta + \frac{1}{4} < c$. When $\delta > -\frac{1}{4}$, the measure of consumers is given by $\Pr(o_i^2 > c) = \delta + \frac{1}{4}$. In the latter case, opinions in the second round will be distributed uniformly between c and $\theta + \frac{1}{4}$, implying that

$$E(s_i^2 | o_i^2 > c, \theta) = \theta - E((o_i^2 |_{m_{ii}=-1}) | o_i^2 > c, \theta) = \theta - \left(\theta - \frac{\delta}{2} + \frac{1}{8}\right) = \frac{1}{2}(\delta - \frac{1}{4}).$$

Again, disappointment decreased compared to the first round (see (2)).

The preceding discussion is summarized in the following proposition.

Proposition 1 (i) *The disappointment among consumers of the new product decreases from round 1 to round 2: $E(s_i^1 | o_i^1 > c, \theta) < E(s_i^2 | o_i^2 > c, \theta) < 0$.*

(ii) *The measure of agents choosing the new product decreases: $\Pr(o_i^1 > c) \geq \Pr(o_i^2 > c)$.*

The results summarized in Proposition 1 may be contrasted with the evolution of consumption choices and opinions in the absence of communication. With no communication, the measure of agents consuming the new product would be $\delta + \frac{1}{2}$ in round two, since exactly those with $x_{0i} > c$ would consume it. This would keep expected disappointment and the measure of consumers constant across periods. Unsurprisingly, Proposition 1(i) implies that, by providing information, social learning reduces expected disappointment. More interestingly, Proposition 1(ii) implies that, by reducing the measure of consumers of the new product, social communication helps efficiency when the new product is inferior to the old one, but hurts efficiency when it is superior. This result captures the natural consequence of learning from consumers who were disappointed on average. The negative messages they send reduce the measure of consumers irrespective of the quality of the product.

Combining the effects in parts (i) and (ii) suggests that if messages are sent over subsequent rounds, the efficiency consequences of social learning might be mitigated. Because the result in (i) states that disappointment decreases, we expect the negative effect of social learning on the measure of consumers to become smaller. This might help efficiency for a superior new product, but only if the negative effect is attenuated fast enough, before too many consumers stop consuming it. At the same time, this attenuation might slow down the rejection of an inferior new product by the population of consumers. We look at these issues in more detail below.

2.3 Many rounds

After the second round, each agent who consumed the good again sends a message, denoted m_{ii}^2 to a contact facing a consumption decision in round 3. This message reveals if the agent was positively or negatively surprised, leading the third round agent to form an opinion o_i^3 on θ . Consumption decisions are made and messages are sent in this fashion in every subsequent round n . Thus, the message sent in round $n - 1$ and received in round n is given by

$$m_{ii}^{n-1} = \left\{ \begin{array}{l} 1 \text{ if } s_i^{n-1} > 0 \\ -1 \text{ if } s_i^{n-1} < 0 \\ 0 \text{ if } o_i^{n-1} < c \end{array} \right\}$$

In round $n > 1$, a Bayesian agent who knows $(x_{0i}, o_i^{n-1}, m_{ii}^{n-1})$ forms the opinion $o_i^n(\theta|x_{0i}, o_i^{n-1}, m_{ii}^{n-1}) = o_i^{n-1} + m_{ii}^{n-1} \left(\frac{1}{2}\right)^n$, which generalizes the formula given in (5). Because of these recursive opinions, knowing $(x_{0i}, o_i^{n-1}, m_{ii}^{n-1})$ is equivalent to knowing (x_{0i}, h^{n-1}) , where $h^{n-1} \equiv (m_{ii}^1, m_{ii}^2, \dots, m_{ii}^{n-1})$ is the complete history of messages. We may therefore write initial opinions in round n as

$$o_i^n(\theta|x_{0i}, h^{n-1}) = x_{0i} + \sum_{k=1}^{n-1} m_{ii}^k \left(\frac{1}{2}\right)^{k+1}. \quad (8)$$

Clearly, $m_{ii}^{n-1} = 0$ implies $o_i^{n-1} = o_i^n$ for all n , so that once an agent consumes the old product, he learns nothing about the new one, and will thus keep consuming the old product forever.

As before, we investigate the evolution of consumption patterns and opinions among those who consume the new product. We have the following generalization of Proposition 1.

Proposition 2 (i) For all n , $E(s_i^n | o_i^n > c, \theta) \leq E(s_i^{n+1} | o_i^{n+1} > c, \theta) \leq 0$.

(ii) $\Pr(o_i^n > c)$ decreases monotonically in n up to round n_0 , where $n_0 = \max \{n : |\delta| < \left(\frac{1}{2}\right)^n\}$. If $\delta > 0$, $\Pr(o_i^n > c) = \left(\frac{1}{2}\right)^{n_0} + n_0\delta$ for all $n \geq n_0$, and if $\delta < 0$, $\Pr(o_i^n > c) = 0$ for all $n \geq n_0 + 1$.

Proof. (i) Over time, learning causes opinions to converge towards θ , reducing average disappointment.

(ii) It is best to consider three cases in turn: $\delta \geq \frac{1}{4}$, $\frac{1}{4} > \theta > 0$, and $0 \geq \delta$.

Case 1: $\delta \geq \frac{1}{4}$. Then $n_0 = 1$, and $\Pr(o_i^n > c) = \delta + \frac{1}{2}$ for all n .

Case 2: $\frac{1}{4} > \delta > 0$. As we saw in the previous section, at the start of round 2, agents with x_{0i} between θ and $c + \frac{1}{4}$ will form opinions $o_i^2 < c$ and will not consume in this round. More generally, opinions in (8) imply that at the start of round $n > 2$, a measure of agents equal to

$$\left[\theta + \sum_{k=3}^n \left(\frac{1}{2}\right)^{k-1} \right] - \left[c + \sum_{k=2}^n \left(\frac{1}{2}\right)^k \right] = \left(\frac{1}{2}\right)^n - \delta$$

stop consuming. The set of consumers of the new product will keep shrinking until round n_0 when this becomes zero. The measure of agents consuming from round n_0 onwards is $1 - \sum_{k=1}^{n_0} \left[\left(\frac{1}{2}\right)^k - \delta \right] = \left(\frac{1}{2}\right)^{n_0} + n_0\delta$.

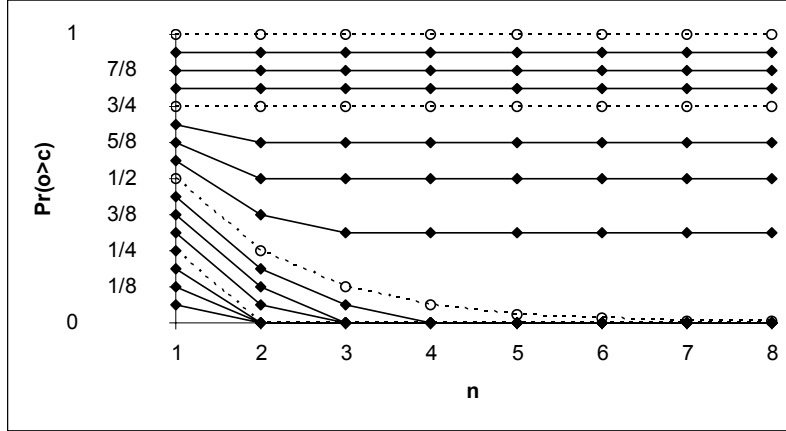


Figure 3: The evolution of $\Pr(o_i^n > c)$. (From top to bottom, paths correspond to $\delta = \frac{8}{16}, \frac{7}{16}, \dots, -\frac{7}{16}$, respectively.)

Case 3: $\delta \leq 0$. In this case, $m_{ii}^n = -1$ for all n . From (8), the measure of agents consuming in round n evolves according to

$$\Pr(o_i^n > c) = \frac{1}{2} + \delta - \sum_{k=2}^n \left(\frac{1}{2}\right)^k = \delta + \left(\frac{1}{2}\right)^n.$$

This expression declines monotonically in n , and reaches 0 at n_0 . ■

Proposition 2 generalizes the second round results of Proposition 1 in an intuitive manner. As shown in the previous section, disappointment results in negative messages, which generally lead to a decrease in the measure of consumers of the new product in the second round. However, the messages sent convey useful information, and help reduce average disappointment in the population. As Proposition 2 shows, if messages are sent over several periods, the reduction in disappointment attenuates the negative effect on consumption. As n increases, the measure of consumers of the new product eventually stabilizes.

This pattern is illustrated in Figure 3, which shows the paths of $\Pr(o_i^n > c)$ for different values of δ . The paths shown correspond to multiples of $\frac{1}{16}$, with the highest path corresponding to $\delta = \frac{8}{16}$, the next one to $\delta = \frac{7}{16}$ and so on down to $\delta = -\frac{7}{16}$. The paths corresponding to $\delta = \frac{1}{2}, \frac{1}{4}, 0$ and $-\frac{1}{4}$ are marked with dotted lines. As Proposition 2(ii) shows, social learning does not affect the measure of consumers if the new product is highly superior to the old one ($\delta \geq \frac{1}{4}$). However, if the new product is only slightly better ($\frac{1}{4} > \delta > 0$), the negative messages transmitted cause the measure of consumers to decline over several rounds. If the surplus is small, this might continue for a long time, so that only very few consumers end up choosing the new product. To see this formally, note that according to Proposition 2(ii), $\lim_{\delta \rightarrow 0} [\Pr(o_i^n > c) = \left(\frac{1}{2}\right)^{n_0} + n_0 \delta] \leq \lim_{\delta \rightarrow 0} \delta(1 + n_0) = 0$. Thus, the combination of self-selection and limited communication studied here does not affect efficiency when the new product is highly superior to the old one, but hurts efficiency for new products that are only slightly superior.

Interestingly, in the $\frac{1}{4} > \delta > 0$ case, in any period n , the derivative of the measure of consumers

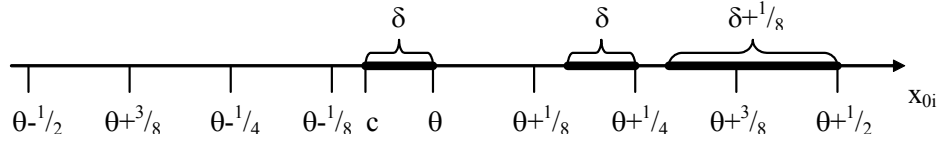


Figure 4: Set of consumers (in bold) for $n \geq n_0 = 3$.

with respect to initial beliefs, $\partial \Pr(o_i^n > c) / \partial x_{0i}$, can be highly non-monotonic. As n increases, disjunct sets of optimistic agents switch to the old product in each round. The outcome of the process is illustrated in Figure 4 for the case of $n_0 = 3$. The first set of missing optimists, with x_{0i} just above θ , switch to the old product in round 2. The second set, with x_{0i} just above $\theta + \frac{1}{4}$, switch in round 3.

Finally, when the new product is worse than the old one, social learning clearly helps efficiency by transmitting the negative messages. Eventually, the consumption of an inferior new product stops completely, although the limited nature of communication means that this might take very long if the quality difference is small: $\lim_{\delta \rightarrow 0} n_0 \rightarrow \infty$. Learning does especially well when the new product is highly inferior to the old one. As we already saw in the previous section, when $\delta < -\frac{1}{4}$, consumption of the new product ceases completely after the first round.

To summarize, this model predicts that highly superior new products will be consumed by a large set of consumers, whose disappointment will decrease over time. New products that are only slightly superior, on the other hand, might end up being consumed by only a small fraction of the population. At the same time, inferior new products will eventually be driven out of the market. In particular, the limited communication studied here is sufficient to drive out inferior new products immediately when the quality difference is large enough. Since without social learning, the set of consumers would remain unchanged in this model, the combination of social learning and self selection helps efficiency when the new product is inferior to the old one, but hurts efficiency when it is superior.

The analysis above has yielded some stark results, but relied on three strict assumptions: (i) that the agent only learns from someone with identical prior to his; (ii) that the experience fully reveals the new product's quality; and (iii) that communication (or memory) is imperfect, and only surprises are communicated. Below, we study the effect of modifying each of these assumptions in turn.

3 Neighbors

In this section we relax Assumption 3. We assume that agents in the second round can get messages from their “neighbors”, who have similar (but not identical) priors to theirs. There is extensive evidence from several literatures that the people one consults before making consumption decisions about experience goods are socially close “neighbors”. When choosing among contraceptive methods, women rely on the advice of other women, who are their village neighbors, friends or parents (see, e.g., Entwisle et al. (1996) on Thailand, and Bernardi (2004) on Italy). Munshi and Myaux (2002) present evidence on the role of ethnic origin and religious affiliations in determining learning patterns

about contraceptives in Bangladesh. Romani (2003) presents similar evidence on the diffusion of agricultural innovations in Cote d'Ivoire. If people who are socially close share similar opinions, and the agent knows this, then learning about the surprise of a neighbor is informative. We investigate the implications of using such information below.

We assume that agents do not know the initial beliefs of their neighbors exactly. They only know that it is close to theirs. In particular, let us assume that each agent j in the second round observes messages from her "neighborhood" $0 < \omega < \frac{1}{2}$, i.e. $C(j) = \{i | x_{0i} \in [x_{0j} - \omega, x_{0j} + \omega]\}$.

As in the previous section, we start by computing the expected value of messages that are sent after the first round. If we randomly pick an agent from the second round who received a message from her ω neighborhood, the probability that her message was negative is given by

$$\Pr(m_{ij} = -1 | \theta, m_{ij} \neq 0) = \Pr(x_{0i} > \theta | \theta, m_{ij} \neq 0) = \frac{\Pr(x_{0i} > c, x_{0j} - \omega \leq x_{0i} \leq x_{0j} + \omega, \theta < x_{0i})}{\Pr(x_{0i} > c, x_{0j} - \omega \leq x_{0i} \leq x_{0j} + \omega)}.$$

We find

$$\Pr(m_{ij} = -1 | \theta, m_{ij} \neq 0) = \Pr(x_{0i} > \theta | \theta, m_{ij} \neq 0) = \frac{1 - \frac{\omega}{2}}{1 - \frac{\omega}{2} + 2\delta}.$$

Hence, the expected value of the message received by the agent is given by

$$E(m_{ij} | m_{ij} \neq 0, \theta) = \left\{ \begin{array}{ll} -1 & \text{if } \delta \leq 0 \\ \frac{2\delta - 1 + \frac{\omega}{2}}{2\delta + 1 - \frac{\omega}{2}} & \text{if } \frac{1}{2} - \omega > \delta > 0 \end{array} \right\},$$

which is always negative. If the new product is worse than the old one, $\delta < 0$, all those who try it will be disappointed, just as in the $\omega = 0$ case, so each j agent who receives a message receives $m_{ij} = -1$. When the new product is better than the old one and we keep the neighborhood relatively small, the larger the neighborhood, the higher the expected message. Although the expected message remains negative, a larger neighborhood mitigates this distortion. The intuition is that a larger neighborhood increases the range of potential j agents to whom a given i agent can send her message. However, the neighborhood of a very optimistic i agent (who sends negative messages) increases relatively less, so that on average fewer negative messages are sent.¹⁹

Next, consider the opinion of an agent in the second round, given that there is a message. In the Appendix, we prove the following Lemma.

¹⁹For example, the neighborhood of the most optimistic i agent, $x_{0i} = \theta + \frac{1}{2}$, is only ω large, while less optimistic agents have a neighborhood of size 2ω . Hence, as the neighborhood grows, the negative messages of the optimistic agents reach relatively less people than the positive message of pessimistic agents. One might think that this effect is the artifact of the bounded support assumption on x_{0i} and x_{0j} given θ . However, all is needed is for the distribution of x_{0i} and x_{0j} to be symmetric around θ . This is so because self-selection ensures that the most optimistic agents will tend to be further away from θ than the most pessimistic agents. Consequently, a larger neighborhood will increase the range of potential contacts of very optimistic i agents less than that of very pessimistic ones.

Lemma 1 (i) If $c + \omega < x_{0j}$, then

$$o_j^2(\theta|x_{0j}, m_{ij}) = x_{0j} + m_{ij} \left(\frac{1}{4} - \frac{1}{8}\omega \right).$$

(ii) If $c < x_{0j} < c + \omega$, then

$$\begin{aligned} o_j^2(\theta|x_{0j}, 1) &= \frac{\omega}{\omega + x_{0j} - c} \left(x_{0j} + \frac{1}{4}\omega + \frac{1}{4} \right) + \frac{x_{0j} - c}{\omega + x_{0j} - c} \left(\frac{1}{2}x_{0j} + \frac{1}{2}c + \frac{1}{4} \right) \\ o_j^2(\theta|x_{0j}, 0) &= \frac{1}{4}(3x_{0j} - \omega + c) \\ o_j^2(\theta|x_{0j}, -1) &= \frac{\omega}{\omega + x_{0j} - c} \left(x_{0j} + \frac{1}{2}\omega - \frac{1}{4} \right) + \frac{x_{0j} - c}{\omega + x_{0j} - c} \left(\frac{3}{4}x_{0j} + \frac{1}{4}c - \frac{1}{4} \right) \end{aligned}$$

(iii) If $c - \omega < x_{0j} < c$, then

$$\begin{aligned} o_j^2(\theta|x_{0j}, 1) &= \frac{1}{4}(3x_{0j} + \omega + c + 1) \\ o_j^2(\theta|x_{0j}, 0) &= \frac{1}{4}(3x_{0j} - \omega + c) \\ o_j^2(\theta|x_{0j}, -1) &= \frac{1}{2}(x_{0j} + \omega + c - \frac{1}{2}). \end{aligned}$$

Proof. See Appendix A.1. ■

Part (i) of the Lemma generalizes the updating rule (5) of the benchmark case with $\omega = 0$. We see that when an agent has neighbors, she responds less to messages than when these messages come from herself (the adjustment is $m_{ij} \left(\frac{1}{4} - \frac{1}{8}\omega \right)$ compared to $m_{ij} \frac{1}{4}$ in the benchmark case). The reason is clear: whatever message (positive or negative) I get, I can never rule out the possibility that my neighbor and I have initial opinions on the opposite side of θ . For example, a positive message might still mean that $x_{0i} < \theta < x_{0j}$, in which case I would have sent myself a negative message instead. Thus agents attenuate their adjustment, and this attenuation is larger the larger the neighborhood. When $\omega = \frac{1}{2}$, agents do not respond to messages at all, and $o_j^2 = x_{0j}$ for all m_{ij} in part (i) of Lemma 1. This diminished responsiveness to messages will be the basis of the *"learning-from-neighbors"* effect we discuss below.

Parts (ii) and (iii) of the Lemma apply when an agent's initial belief is in the neighborhood of the cutoff c . In this case, from part of her neighborhood, an agent can only get the message $m_{ij} = 0$, since her neighbors with $x_{0i} < c$ will not have consumed the new product. Thus, messages convey more information, and the agent takes this into account when forming her opinions. Note in particular that no news is bad news: $m_{ij} = 0$ implies that someone in the neighborhood had a signal $x_{0i} < c$. This causes an agent receiving a 0 message to adjust her opinion downwards (naturally, the adjustment is smaller than after $m_{ij} = -1$). This will be the basis of the *"inexperienced-neighbor"* effect described below.

In the remainder of this section, we discuss how the results of Section 2 with respect to consumption patterns are modified if $\omega > 0$. To simplify the analysis, and to make the comparison with the

benchmark case more transparent, we assume that the neighborhood ω is small.²⁰ Our main result is the following.

Proposition 3 *The measure of agents consuming in the second round, $\Pr(o_j^2 > c)$, is increasing in ω for all $\delta \in (-\frac{1}{2}, \frac{1}{2})$.*

Proof. See Appendix A.2. ■

The proposition states that new products have a higher probability of being consumed if agents learn from larger neighborhoods. This result is striking because it applies both when the new product is better and when it is worse than the old one. We provide an intuitive explanation of the result below. We phrase the discussion in terms of four "effects" that larger neighborhoods have on second round consumption *relative* to the benchmark (or relative to smaller neighborhoods).

Consider first the case of a superior new product ($\delta > 0$), and take an agent with x_{0j} just below the cutoff c . This agent is quite pessimistic, so she would not receive a message or try the product in the $\omega = 0$ case. However, now she has neighbors who are less pessimistic than her, and therefore try the product in the first round. Precisely because the agent is quite pessimistic, her neighbors are pessimistic also, and will therefore have a positive experience with the new product if they try it. If our agent receives a message from such a neighbor, it will be a positive message, and will convince her to try the product. We call this the "*experienced-neighbor*" effect. Under this effect, an agent who would not get a message with $\omega = 0$ might get a message of the "right" sign. In effect, she gets the message she could not send to herself from someone else. This can only change consumption choices from old to new, therefore the *experienced-neighbor* effect always favors the new product.²¹

Second, take an agent with x_{0j} just *above* the cutoff c . Such an agent has neighbors who consumed the old product, she therefore has a positive probability of getting a 0 message. From Lemma 1(ii), we know that such a message is bad news, causing the agent to adjust her opinion downwards. Thus, while the agent always consumed the new product when $\omega = 0$, she might now choose the old one instead. We call this the "*inexperienced-neighbor*" effect, which reduces the measure of consumers for $\delta > 0$.

Third, take an agent with x_{0j} just below θ : $x_{0j} \in [\theta - \omega, \theta]$. When $c < \theta - \omega$, such an agent has a high enough x_{0j} to try the product without a message. However, precisely because her x_{0j} is relatively high, she might have an optimistic neighbor, who will get disappointed by the new product. If a pessimistic agent receives a negative message from such a neighbor, she will adjust her opinion downwards. When δ is not too large, the agent will not try the product in the second round, even though she would have tried it in the benchmark case of $\omega = 0$. We call this the "*misleading-neighbor*" effect, since in this case the message sent by the neighbor is the opposite of what the agent would

²⁰Specifically we assume that $\frac{1}{6} \geq \omega$ to guarantee that we have $o_j^2|_{m_{ij}=-1} < c$ in Lemma 1(ii), so that agents with x_{0j} just above c will stop consuming the new product if they get a negative message.

²¹Note that the *experienced-neighbor* effect implies that the measure of consumers might *increase* from round 1 to round 2, as agents between $c - \omega$ and c consume with positive probability in round 2. We conjecture that if we allow for learning to continue over several rounds, there is a positive probability that everyone will end up consuming the new product.

have sent to herself. We observe this same effect for the optimistic agents with $x_{0j} \in [\theta, \theta + \omega]$. In the benchmark case, these agents would get a negative message for sure and would therefore not consume in the second round for $\delta < \frac{1}{4}$. However, with $\omega > 0$, these slightly optimistic agents may have a pessimistic neighbor and might therefore get a positive message. The *misleading-neighbor* effect will cause them to consume the new product in the second round. As we see, the *misleading-neighbor* effect can favor both the old and the new product.

Finally, consider the case of an inferior new product ($\delta < 0$). In this case, we have two positive neighborhood effects. First, agents slightly above the cutoff c might have a neighbor who is below the cutoff, and therefore did not consume the product. These agents will not get a message, and some of them will consume the new product, even though they would have stopped consuming it after receiving a negative message in the benchmark case. This is the *inexperienced-neighbor* effect, which favors the inferior new product. With no information coming from a neighbor who did not try the product, the agent is forced to try the product and "see for herself".²²

To see the second neighborhood effect, recall that as Lemma 1(i) shows, even the agents who get a message with probability 1 adjust their opinions less than in the benchmark case (by $\frac{1}{4} - \frac{\omega}{8}$ instead of $\frac{1}{4}$). When $\delta < 0$, this implies that if δ is not too large in absolute value, some agents who would have formed opinions $o_j^2|_{m_{ij}=-1} < c$ in the benchmark case will now still have $o_j^2|_{m_{ij}=-1} > c$, and will thus consume the new product in round 2. We call this the "*learning-from-neighbors*" effect. Because my neighbor is different from me, I do not adjust my opinion as much in response to a message coming from her as I would if the message was coming from me. This effect always favors the new product.

To summarize, we can isolate two effects of having neighbors that unambiguously favor the new product (*experienced-neighbor*, and *learning-from-neighbors*), and two effects that can favor the old or the new product (*misleading-neighbor* and *inexperienced-neighbor*). As we show in Proposition 3, the neighborhood effects favoring the new product always dominate. Thus, having a larger neighborhood always has a positive impact on the probability that the new product will be consumed in the second round. This is true regardless of whether the new or the old product is better.

These results suggest that when a community is more open or diverse in the sense that people exchange opinions with others who are different from them, the diffusion of a good product is further facilitated by social learning. This occurs because some people will learn from others who have tried the product which they themselves would have been too pessimistic to try (the *experienced-neighbor* effect); and because some optimistic people who would have been disappointed will instead get a positive message from their neighbor, and thus try the product (the *misleading-neighbor* effect). At the same time, such communities might be slow in driving out inferior products, because some people will not learn about an inferior product when talking to someone who has not tried it, and will thus have to try it themselves (the *inexperienced-neighbor* effect); and because the disappointing messages resulting from self-selection will be discounted when coming from people who are different from the agent, leading her to consume the inferior new product (the *learning-from-neighbors* effect). Thus,

²²The *inexperienced-neighbor* effect implies that, in contrast to the benchmark case in section 2, a positive measure of consumers consume the new product in round 2 even if it is highly inferior ($\delta < -\frac{1}{4}$).

more diverse communities have a tendency to be more receptive to the New, be it better or worse than the Old.

4 The accuracy of communication

In this section, we assume that the message m_{ii} can provide a more accurate description of the experience θ . We relax Assumption 2 to capture the fact that apart from communicating the sign of the surprise, people are often able to tell each-other how large their surprise was.

To model this formally, we assume that the message m_{ii} creates $2l$ partitions of equal size on the information set (conditional on x_{0i}) of each agent, $[x_{0i} - \frac{1}{2}, x_{0i} + \frac{1}{2}]$. The exogenous parameter l reflects the informativeness of communication. A message $m_{ii}(l)$ reveals which partition θ belongs to. The benchmark model corresponds to $l = 1$ or two partitions, where m_{ii} revealed whether θ was in the "upper" partition $[x_{0i}, x_{0i} + \frac{1}{2}]$, or in the "lower" partition $[x_{0i} - \frac{1}{2}, x_{0i}]$. As l becomes larger, $m_{ii}(l)$ locates θ more accurately.

Clearly, the exact value of the label m_{ii} is irrelevant, as long as each value identifies a partition unambiguously. To simplify the algebra, we assume that $m_{ii}(l)$ identifies each partition with an odd integer, with negative values corresponding to $\theta < x_{0i}$, and higher absolute values corresponding to larger distances from x_{0i} . For example, if $l = 2$, besides communicating the sign of the surprise in the first round, agent i can also indicate whether his surprise was "large and positive" ($m_{ii} = 3$), "small and positive" ($m_{ii} = 1$), "small and negative" ($m_{ii} = -1$), or "large and negative" ($m_{ii} = -3$). More generally,

$$m_{ii}(l) = \left\{ \begin{array}{l} -(2k - 1) \text{ if } \theta + \frac{k-1}{2l} < x_{0i} < \theta + \frac{k}{2l} \\ 2k - 1 \text{ if } \theta - \frac{k}{2l} < x_{0i} < \theta - \frac{k-1}{2l} \end{array} \right\} \quad (9)$$

where $k \in [1, l]$ is an integer such that $2k - 1$ identifies the partition. Under these messages, opinions in the second round are given by

$$o_i^2(\theta|x_{0i}, m_{ii}(l)) = x_{0i} + m_{ii}(l) \frac{1}{4l}, \quad (10)$$

which generalizes formula (5) in the benchmark case. Figure 5 illustrates the updating rule (10) for $l = 2$.

As before, an agent in the second round will choose the new product iff $o_i^2 > c$. When comparing this setup to the benchmark case, we obtain the following result.

Proposition 4 (i) *The measure of second-round consumers, $\Pr(o_i^2 > c)$, (weakly) increases in l when $\delta > 0$ and decreases in l when $\delta < 0$.*

(ii) *Expected disappointment $|E(s_i^2|o_i^2 > c, \theta)|$ decreases in l .*

Proof. Consider first the case of $\delta > 0$. Since (10) implies that an agent's opinion can fall below θ by at most $\frac{1}{4l}$, the set of consumers will remain unchanged after the first round as long as $\delta \geq \frac{1}{4l}$. To compute expected surprise in this case, recall that in the benchmark case, opinions conditional on

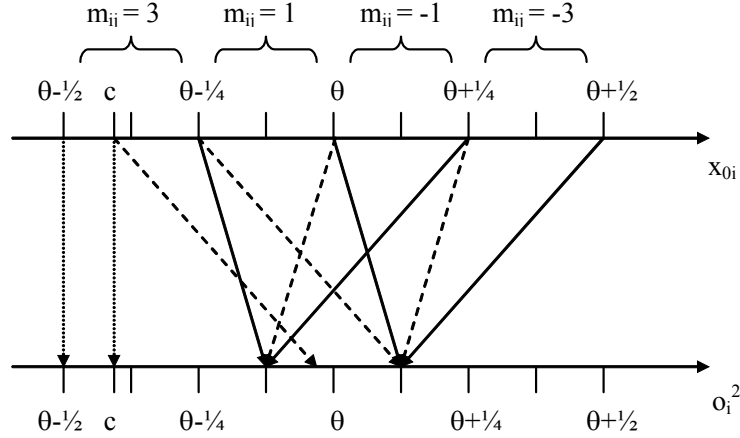


Figure 5: Adjustment of opinions when $l = 2$.

a message were accurate, as long as the message was not the highest one actually sent ($m_{ii} = 1$). The same result holds in the more general case. If $\theta - \frac{\bar{k}}{2l} < c < \theta - \frac{\bar{k}-1}{2l}$ for some \bar{k} , opinions in (10) imply that for all $k < \bar{k}$, a message $m_{ii}(l) = 2k - 1$ results in an opinion that is accurate conditional on the message. Opinions are only distorted conditional on the highest message actually sent, $m_{ii}(l) = 2\bar{k} - 1$. In particular,

$$o_i^2 |_{m_{ii}=2\bar{k}-1} = x_{0i} + (2\bar{k} - 1) \frac{1}{4l},$$

and

$$E((o_i^2 |_{m_{ii}=2\bar{k}-1}) | x_{0i} > c, \theta) = \frac{\theta - \frac{\bar{k}-1}{2l} + c}{2} + (2\bar{k} - 1) \frac{1}{4l} = \theta - \frac{\delta}{2} + \frac{\bar{k}}{4l} > \theta. \quad (11)$$

Because message $m_{ii}(l) = 2\bar{k} - 1$ is sent with probability $\Pr(m_{ii}(l) = 2\bar{k} - 1 | x_{0i} > c, \theta) = \frac{\theta - \frac{\bar{k}-1}{2l} - c}{\delta + \frac{1}{2}}$, expression (11) implies that the average surprise is given by

$$\begin{aligned} E(s_i^2 | x_{0i} > c, \theta) &= E(\theta - (o_i^2 |_{m_{ii}=2\bar{k}-1}) | x_{0i} > c, \theta) \Pr(m_{ii}(l) = 2\bar{k} - 1 | x_{0i} > c, \theta) \\ &= \left(\frac{\delta}{2} - \frac{\bar{k}}{4l} \right) \frac{\delta - \frac{\bar{k}-1}{2l}}{\delta + \frac{1}{2}} = \frac{\left(\delta - \frac{\bar{k}}{2l} \right) \left(\delta - \frac{\bar{k}}{2l} + \frac{1}{2l} \right)}{2\delta + 1} \end{aligned}$$

This expression generalizes the expected surprise (6) derived in the benchmark case. One can check that it is negative and increasing in l .

When δ is between 0 and $\frac{1}{4l}$, some consumers will consume the new product in the first round but not in the second. Optimistic agents with x_{0i} between θ and $c + \frac{1}{4l}$ adjust downwards by $\frac{1}{4l}$ and stop consuming. Agents between $c + \frac{1}{4l}$ and $\theta + \frac{1}{2l}$ adjust downwards by the same amount, which is enough to keep them consuming. However, agents between $\theta + \frac{1}{2l}$ and $c + 3\frac{1}{4l}$ adjust by $-3\frac{1}{4l}$ and stop consuming. Those between $c + 3\frac{1}{4l}$ and $\theta + 2\frac{1}{2l}$ again consume, and so on (see Figure 6 below

illustrating the case $l = 4$). The measure of agents consuming in round two is given by

$$\Pr(o_i^2 > c) = \delta + \frac{1}{2} - \sum_{k=1}^l \left[c + (2k-1)\frac{1}{4l} - \theta - (k-1)\frac{1}{2l} \right] = \delta(1+l) + \frac{1}{4},$$

which generalizes expression (7), and is increasing in l .

***opinions: to be completed

Finally, when the surplus is negative, only negative messages are transmitted, and the set of consumers shrinks every period. When $\delta < -\frac{1}{4l}$, consumption stops completely after the first round. When $\delta > -\frac{1}{4l}$, so that some agents consume, $o_i^2 \sim U[c, \theta + \frac{1}{4l}]$, implying that

$$E(s_i^2 | o_i^2 > c, \theta) = \theta - E(o_i^2 | o_i^2 > c, \theta) = \frac{\delta}{2} - \frac{1}{8l}$$

and

$$\Pr(o_i^2 > c) = \delta + \frac{1}{4l}.$$

These expressions also confirm the proposition. ■

Since the more refined communication technology allows for more precise information to be transmitted regarding the unknown quality θ , it is not surprising that the effects identified in the benchmark case by Propositions 1 and 2 are weakened. As proposition 4 shows, expected disappointment declines, and the measure of consumers changes in the direction required by efficiency. In particular, as communication becomes more accurate, a lower quality advantage δ of the new product is enough to prevent the set of consumers from shrinking. Similarly, a smaller quality difference in favor of the old product is enough to drive out the new product immediately. Nevertheless, as long as $l < \infty$, so that communication has some limitation, there is expected disappointment in the population, which causes negative messages to be sent and leads to a decline in the measure of consumers from round 1 to round 2. Our results from the benchmark generalize.²³

Before turning to other extensions, we note that, interestingly, the one effect that becomes stronger and not weaker by making communication more precise is the non-monotonicity of the probability of consuming with respect to the initial beliefs x_{0i} . In contrast to the benchmark case, this derivative can change sign several times already in round 2. As Figure 6 illustrates, with more precise communication, even very optimistic agents might stop consuming a slightly superior new product immediately after the first round of learning. (In the Figure, $c < \frac{1}{16}$, and $l = 4$, so that opinions are adjusted by multiples of $\frac{1}{16}$, according to (10).)

²³ As before, communication helps the disappointment to become smaller, so that these effects will decline as learning occurs over several rounds. Since communication is more precise, we expect the processes described in Proposition 2 to converge faster.

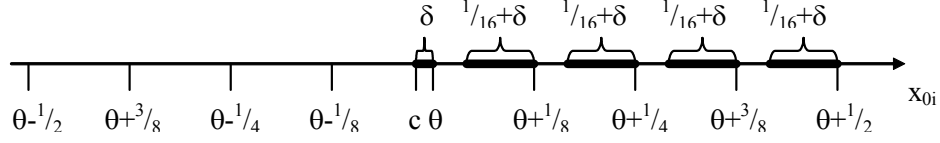


Figure 6: Set of second-round consumers (bold segments) when $l = 4$.

5 Consumption experience as a signal

In this section, we relax Assumption 1 by allowing for the fact that consumption might not reveal fully the quality of a product. We show that the finding of negative "post decision surprise" of Harrison and March (1984), which in our model results in negative messages sent, survives in this setting, but expected disappointment decreases with the variance of the signal that consumption provides.

We now assume that if an agent decides to consume the new product, she receives a signal, which is of at least the same quality as her prior x_{0i} . In particular, we assume that $x_{1i} = \theta + \varepsilon_{1i}$, where $\varepsilon_{1i} \sim U(-a, a)$, and $0 \leq a \leq \frac{1}{2}$ is a known constant (the benchmark corresponds to $a = 0$). Let $\Delta x_i \equiv x_{1i} - x_{0i}$. An agent's posterior opinion, conditional on having consumed the good, is as follows:

$$o_i(\theta|x_{0i}, x_{1i}) = \left\{ \begin{array}{ll} \frac{1}{2}(x_{1i} + x_{0i} - a + \frac{1}{2}) & \text{if } \Delta x_i > \frac{1}{2} - a \\ x_{1i} & \text{if } \frac{1}{2} - a > \Delta x_i > a - \frac{1}{2} \\ \frac{1}{2}(x_{1i} + x_{0i} + a - \frac{1}{2}) & \text{if } a - \frac{1}{2} > \Delta x_i \end{array} \right\}$$

Clearly, $o_i(\theta|x_{0i}, x_{1i}) > o_i(\theta|x_{0i})$ ($o_i(\theta|x_{0i}, x_{1i}) < o_i(\theta|x_{0i})$) when $\Delta x_i > 0$ ($\Delta x_i < 0$). For given Δx_i , if the signal x_{1i} is precise enough (so that $a < \frac{1}{2} - |\Delta x_i|$), it will override the first signal in the agent's inference on θ .

The surprise of agent i is then given by

$$s_i = \left\{ \begin{array}{ll} \frac{1}{2}(\Delta x_i - a + \frac{1}{2}) & \text{if } \Delta x_i > \frac{1}{2} - a \\ \Delta x_i & \text{if } \frac{1}{2} - a > \Delta x_i > a - \frac{1}{2} \\ \frac{1}{2}(\Delta x_i + a - \frac{1}{2}) & \text{if } a - \frac{1}{2} > \Delta x_i \end{array} \right\}$$

As before, if everyone buys the product, one can check that

$$E(s_i|\theta) = 0. \tag{12}$$

Because on average people hold correct beliefs, the average surprise is zero. How does expected surprise change from (12) when only those with $x_{0i} > c$ consume the new product? The following Proposition provides the answer.

Proposition 5 *On the domain $a \in [0, \frac{1}{2}]$, $\delta \in (-\frac{1}{2}, \frac{1}{2})$ we have $E(s_i|x_{0i} > c) < 0$, $\frac{\partial|E(s_i|x_{0i} > c)|}{\partial\delta} < 0$ and $\frac{\partial|E(s_i|x_{0i} > c)|}{\partial a} < 0$.*

Proof. See Appendix A.3. ■

The first two statements simply generalize the properties derived in the benchmark case. There is disappointment for all values of a , and the disappointment is reduced if the new product yields a higher surplus δ so that more agents will consume it. The third result is more interesting: as a increases for given δ , so that experience provides a less precise signal of the true quality, the expected disappointment in the population declines. To see the intuition, recall that average disappointment is caused by a lack of positive surprises from the agents with x_{0i} below θ who did not consume the good.²⁴ This effect is reduced if some of the agents with x_{0i} above θ can have positive surprises with some probability. But the more precise the second signal, the higher this probability,²⁵ making the lack of positive surprises from the lower end of the distribution less salient and decreasing average disappointment. Disappointment is thus smallest when $a = \frac{1}{2}$ (both signals have the same quality). When $a = 0$, as in the benchmark model, quality is perfectly revealed, and expected disappointment reaches its maximum of $\frac{1}{2}(\delta - \frac{1}{2})$.

6 Perfect communication with a threshold

In this section, we modify the rules governing communication in the model (Assumption 2). We assume that experiences x_{1i} can be communicated precisely, but communication occurs only if the surprise of agent i was larger than a fixed “communication threshold” \bar{s} . Formally, we assume that

$$m_{ij} = \begin{cases} 0 & \text{if } -\bar{s} \leq s_i \leq \bar{s} \\ x_{1i} & \text{otherwise} \end{cases}. \quad (13)$$

This seems realistic: many of us would find a surprising experience a more interesting topic for conversation than an unsurprising one. Remembering past experiences might also be easier if the experience was sufficiently different from what we expected.²⁶

Compared to previous sections, now surprises do not affect the content of the messages directly, only the probability that they are sent. Furthermore, we no longer need to assume that the receiver knows something about the initial expectation of the sender for social learning to be informative. Because the experience itself is communicated, messages will always contain information. Finally, note that with messages given by (13), learning is trivial when $x_{1i} = \theta$. To make the model interesting, we assume that the signal provided by the consumption experience is only as good as the prior x_{0i} . Thus, we will consider the case of $a = \frac{1}{2}$, in the notation of the previous section.

We show that our basic result that the expected message received is negative even if the new product is better holds in this set-up as well. The reason, however, is quite different from before. Note that we now have two selection effects. First, a subset of agents defined by c is selected to

²⁴Note that if $a < \delta$, those who did not consume the good would have gotten a signal higher than x_{0i} with probability 1, had they consumed it.

²⁵For example, when $\delta < 0$, the only way *anyone* can have a positive surprise with some probability is if a is large enough to exceed $|\delta|$.

²⁶A more complicated formulation would make the precision of the signal received from a network partner an increasing function of $|s_i|$. This might be justified for example if people talk about bigger surprises in greater length, making it possible to transmit more precise information.

consume the good (selection in participation). Next, of the potential messages reporting about the consumption experience, only a subset defined by \bar{s} is actually sent (selection in communication). None of these distortions affects the content of the message. Those who do send a message will reveal their experience, x_{1i} , which is an unbiased signal of θ . Nevertheless, as we show below, the average message will still be biased downwards. This is a consequence of the interaction of the two selection effects.

When $a = \frac{1}{2}$, results in the previous section imply that $s_i = \frac{1}{2}(x_{1i} - x_{0i})$. Agents will be positively (negatively) surprised if their experience is higher (lower) than their prior. The two selection effects are illustrated in Figure 7, which shows the joint distribution of the signals x_{1i} and x_{0i} . The lines with positive slopes are “iso-surprise” lines (pairs (x_{0i}, x_{1i}) which yield a given value of s_i), with increasing surprise as we move up and left on the graph. The vertical line at $x_{0i} = c$ shows the selection in participation: only agents with noises to the right of this line will send messages. The upward sloping lines $x_{1i} = x_{0i} + 2\bar{s}$ and $x_{1i} = x_{0i} - 2\bar{s}$ show the selection in communication: above the upper line, the positive surprise of the agents will be high enough to make them talk, while below the lower line, agents will be disappointed enough to talk. Hence, messages will only be sent by agents with noises in areas A and B, with messages from area A typically being the high messages, while messages from area B typically being the low messages. As can be seen from the figure, for any c and \bar{s} , more agents are removed from area A than from area B. When $\bar{s} > 0$, this implies that on average lower messages are sent, i.e. the expected message sent is below θ .

Note that since the signals x_{1i} and x_{0i} are independent, if $\bar{s} = 0$, any subset of agents send messages that are accurate on average, so that selection in participation is irrelevant. Although with $\bar{s} = 0$ there would still be more disappointed agents, with perfect communication disappointment per se does not matter, as messages are determined by experience (i.e. whether an agent was lucky (high x_{1i}) or unlucky (low x_{1i})). When the threshold is positive, $\bar{s} > 0$, x_{0i} and x_{1i} jointly determine the probability of sending a message. This makes selection in participation, which is based on x_{0i} , relevant for the expected content of the message, which is based on x_{1i} .

Similarly, with no selection in participation, selection in communication does not matter. If $\delta \geq \frac{1}{2}$, the distribution of messages will be symmetric around θ for any value of \bar{s} , so that selection in communication has no effect. Hence, it is the interaction of the two selection effects which causes messages to be negatively biased.

These arguments are developed more formally in the following proposition.

Proposition 6 *For any $-\frac{1}{2} < \delta < \frac{1}{2}$ and $0 < \bar{s} < \frac{1}{2}$, $E(m_{ij}|\theta) < \theta$.*

Proof. Write the expected message as

$$E(m_{ij}|\theta) = E(x_{1i}|\theta, m_{ij} \neq 0) = E(x_{1i}|s_i > \bar{s}) \Pr(s_i > \bar{s}) + E(x_{1i}|s_i < -\bar{s}) \Pr(s_i < -\bar{s})$$

Consider first the case where each agent i consumes the good, so that $\delta \geq \frac{1}{2}$:

$$E(x_{1i}|s_i > \bar{s}) \Pr(s_i > \bar{s}) = \int_{\theta - \frac{1}{2}}^{\theta + \frac{1}{2} - 2\bar{s}} \int_{2\bar{s} + x_{0i}}^{\theta + \frac{1}{2}} x_{1i} dx_{1i} dx_{0i} = \frac{1}{12} (4\bar{s} + 1) (2\bar{s} - 1)^2$$

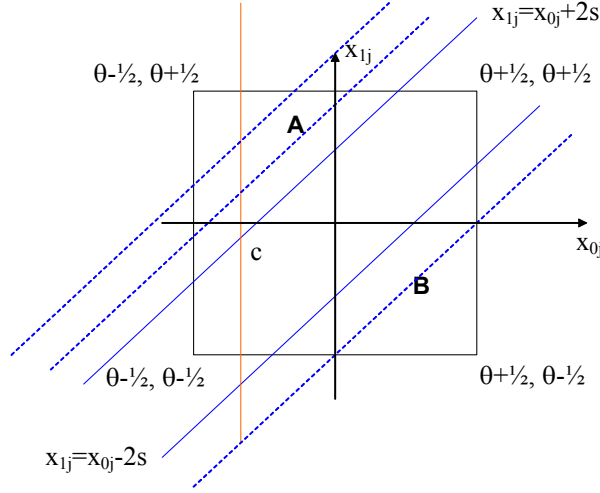


Figure 7: Perfect communication with a threshold

$$E(x_{1i}|s_i < \bar{s}) \Pr(s_i < \bar{s}) = \int_{\theta+2\bar{s}-\frac{1}{2}}^{\theta+\frac{1}{2}} \int_{\theta-\frac{1}{2}}^{x_{0i}-2\bar{s}} x_{1i} dx_{1i} dx_{0i} = -\frac{1}{12} (4\bar{s} + 1) (2\bar{s} - 1)^2$$

Hence, in this case $E(m_{ij}|\theta) = \theta$.

We will now show that as we increase c , $E(x_{1i}|\theta, m_{ij} \neq 0)$ falls monotonically, so that $E(m_{ij}|\theta)$ is always below θ . From our figure, it is clear that as we increase c , the marginal change of $E(x_{1i}|\theta, m_{ij} \neq 0)$ is

$$\frac{\partial E(x_{1i}|\theta, m_{ij} \neq 0)}{\partial c} = -\chi_{\theta-\frac{1}{2} < c < \theta+\frac{1}{2}-2\bar{s}} \int_{c+2\bar{s}}^{\theta+\frac{1}{2}} x_{1i} dx_{1i} + \chi_{\theta+2\bar{s}-\frac{1}{2} < c < \theta+\frac{1}{2}} \int_{\theta-\frac{1}{2}}^{c-2\bar{s}} x_{1i} dx_{1i}$$

where χ is the indicator function, and the first term is the change in the expectation due to the change in area A and the second term is the effect of the change in area B . Note that the first term is always negative, while the second term is always positive in the relevant interval. We can rewrite this expression as follows:

$$\begin{aligned} \frac{\partial E(x_{1i}|\theta, m_{ij} \neq 0)}{\partial c} &= -\chi_{\theta-\frac{1}{2} < c < \theta+\frac{1}{2}-2\bar{s}} \frac{1}{2} \left(\frac{1}{4} - (2\bar{s} - \delta)^2 \right) + \chi_{\theta+2\bar{s}-\frac{1}{2} < c < \theta+\frac{1}{2}} \frac{1}{2} \left((-\delta - 2\bar{s})^2 - \frac{1}{4} \right) \\ &= -\chi_{\theta-\frac{1}{2} < c < \theta+2\bar{s}-\frac{1}{2}} \frac{1}{2} \left(\frac{1}{4} - (2\bar{s} - \delta)^2 \right) + \chi_{\theta+2\bar{s}-\frac{1}{2} < c < \theta+\frac{1}{2}-2\bar{s}} \frac{1}{2} \left(-8\delta\bar{s} - \frac{1}{4} \right) \\ &\quad + \chi_{\theta+\frac{1}{2}-2\bar{s} < c < \theta+\frac{1}{2}} \frac{1}{2} \left((-\delta - 2\bar{s})^2 - \frac{1}{4} \right) \end{aligned}$$

One can check that all three terms are negative for any decision cut-off $c \in [\theta - \frac{1}{2}, \theta + \frac{1}{2}]$ and any conversation cut-off, $0 < \bar{s} < \frac{1}{2}$, so that $E(m_{ij}|\theta) \leq \theta$ for all δ . ■

7 Conclusion

This paper studied a simple model of social learning which incorporates two realistic features: subjective communication and self-selection of social contacts. Self-selection implies that consumers of a product with unknown quality will be disappointed on average. Subjective communication means that messages sent by these disappointed consumers will tend to be negative. We analyzed how these negative messages affected other agents' opinions and likelihood of consuming a new product.

In a benchmark model with close contacts, limited subjective communication and fully revealing experience, we showed that average disappointment declines over time, but the overwhelmingly negative messages lead to a decrease in the measure of consumers. Thus, subjective communication helps efficiency by driving out inferior new products. Indeed, if the quality advantage of the old product is large, the limited amount of information transmitted here is enough to drive out the new product immediately after the first round. However, learning hurts efficiency by causing some agents to stop consuming a superior new product. Consumption of a new product that is only slightly better than the old one can cease almost completely.

In an extension with neighbors instead of close contacts, we showed that larger neighborhoods imply that a larger measure of consumers choose the new product, regardless of its quality. We described four effects of having neighbors (*experienced-neighbor*, *inexperienced-neighbor*, *misleading-neighbor*, and *learning-from-neighbors*), which together suggest that more "diverse" neighborhoods are more favorable to the diffusion of a superior new product, but also that these same neighborhoods are slower in driving out inferior new products.

We showed that the results of the benchmark case are robust to allowing more accurate communication, with an interesting pattern of non-monotonic consumption likelihood emerging in this case. We also showed that the basic phenomenon driving our results, expected disappointment and negative messages, remains if we allow for the consumption experience to provide an imperfect signal of the true quality, or if subjective communication is modeled as perfect communication subject to a surprise-threshold.

Appendix

A.1 Proof of Lemma 1

To simplify the notation, denote $W^* = [x_{0j} - \omega, x_{0j}]$, $W^{**} = [x_{0j}, x_{0j} + \omega]$ and $W = W^* \cup W^{**}$. For $m_{ij} \neq 0$, we can then write

$$\begin{aligned}
 o_j^2(\theta|x_{0j}, m_{ij}) &= o_j^2(\theta|x_{0j}, \text{sign}(s_i), x_{0i} \in W) \\
 &= o_j^2(\theta|x_{0j}, \text{sign}(s_i), x_{0i} \in W^{**}) \Pr(x_{0i} \in W^{**}|x_{0j}, \text{sign}(s_i)) \\
 &+ o_j^2(\theta|x_{0j}, \text{sign}(s_i), x_{0i} \in W^*) \Pr(x_{0i} \in W^*|x_{0j}, \text{sign}(s_i)).
 \end{aligned} \tag{14}$$

(i) If $x_{0j} > c + \omega$, then $m_{ij} \neq 0$ necessarily. In this case,

$$\Pr(x_{0i} \in W^{**} | x_{0j}, \text{sign}(s_i)) = \Pr(x_{0i} \in W^* | x_{0j}, \text{sign}(s_i)) = \frac{1}{2}. \quad (15)$$

Given a positive message $m_{ij} = 1$, we may use $o_j^2(\theta | x_{0i} \in W^{**}, s_i > 0, x_{0j}, x_{0i}) = \frac{x_{0i} + x_{0j} + \frac{1}{2}}{2}$ to find

$$\begin{aligned} o_j^2(\theta | x_{0i} \in W^{**}, s_i > 0, x_{0j}) &= \int_{x_{0j}}^{x_{0j} + \omega} \frac{x_{0i} + x_{0j} + \frac{1}{2}}{2} \Pr(x_{0i}) dx_{0i} = \frac{1}{\omega} \int_{x_{0j}}^{x_{0j} + \omega} \frac{x_{0i} + x_{0j} + \frac{1}{2}}{2} dx_{0i} \\ &= \frac{1}{4} + x_{0j} + \frac{\omega}{4}, \end{aligned} \quad (16)$$

and $o_j^2(\theta | x_{0i} \in W^*, s_i > 0, x_{0j}, x_{0i}) = x_{0i} + \frac{1}{4}$ to find

$$o_j^2(\theta | x_{0i} \in W^*, s_i > 0, x_{0j}) = \frac{1}{\omega} \int_{x_{0j} - \omega}^{x_{0j}} x_{0i} + \frac{1}{4} dx_{0i} = \frac{1}{4} + x_{0j} - \frac{1}{2}\omega. \quad (17)$$

Substituting (15-17) into (14), we get

$$o_j^2(\theta | x_{0j}, m_{ij} = 1) = \frac{1}{2} \left(\frac{1}{4} + x_{0j} + \frac{1}{4}\omega \right) + \frac{1}{2} \left(\frac{1}{4} + x_{0j} - \frac{1}{2}\omega \right) = x_{0j} - \frac{1}{8}\omega + \frac{1}{4}.$$

By the same logic, we find

$$\begin{aligned} o_j^2(\theta | x_{0j}, m_{ij} = -1) &= \frac{1}{2} o_j^2(\theta | x_{0i} \in W^{**}, s_i < 0, x_{0j}) + \frac{1}{2} o_j^2(\theta | x_{0i} \in W^*, s_i < 0, x_{0j}) \\ &= \frac{1}{2} \left[\frac{1}{\omega} \int_{x_{0j}}^{x_{0j} + \omega} x_{0i} - \frac{1}{4} dx_{0i} + \frac{1}{\omega} \int_{x_{0j} - \omega}^{x_{0j}} \frac{x_{0i} + x_{0j} - \frac{1}{2}}{2} dx_{0i} \right] \\ &= x_{0j} + \frac{1}{8}\omega - \frac{1}{4} \end{aligned}$$

(ii) If $c < x_{0j} < c + \omega$, then for $m_{ij} \neq 0$,

$$\Pr(x_{0i} \in W^{**} | x_{0j}, \text{sign}(s_i)) = \frac{\omega}{\omega + x_{0j} - c} \quad (18)$$

$$\Pr(x_{0i} \in W^* | x_{0j}, \text{sign}(s_i)) = \frac{x_{0j} - c}{\omega + x_{0j} - c} \quad (19)$$

Given a positive message $m_{ij} = 1$, we again find $o_j^2(\theta | x_{0i} \in W^{**}, s_i > 0, x_{0j}, x_{0i}) = \frac{x_{0i} + x_{0j} + \frac{1}{2}}{2}$ and

$$o_j^2(\theta | x_{0i} \in W^{**}, s_i > 0, x_{0j}) = \frac{1}{4} + x_{0j} + \frac{1}{4}\omega. \quad (20)$$

We also have $o_j^2(\theta | x_{0i} \in W^*, s_i > 0, x_{0j}, x_{0i}) = x_{0i} + \frac{1}{4}$, which now implies

$$o_j^2(\theta | x_{0i} \in W^*, s_i > 0, x_{0j}) = \frac{1}{x_{0j} - c} \int_c^{x_{0j}} x_{0i} + \frac{1}{4} dx_{0i} = \frac{1}{4} + \frac{x_{0j} + c}{2}. \quad (21)$$

Substituting (18-21) into (14) gives:

$$o_j^2(\theta|x_{0j}, m_{ij} = 1) = \frac{\omega}{\omega + x_{0j} - c} \left(\frac{1}{4} + x_{0j} + \frac{1}{4}\omega \right) + \frac{x_{0j} - c}{\omega + x_{0j} - c} \left(\frac{1}{4} + \frac{x_{0j} + c}{2} \right).$$

For bad news ($m_{ij} = -1$), $o_j^2(\theta|x_{0i} \in W^{**}, s_i < 0, x_{0j}, x_{0i}) = x_{0i} - \frac{1}{4}$, and

$$o_j^2(\theta|x_{0i} \in W^{**}, s_i < 0, x_{0j}) = x_{0j} + \frac{\omega}{2} - \frac{1}{4}.$$

Moreover, $o_j^2(\theta|x_{0i} \in W^*, s_i < 0, x_{0j}, x_{0i}) = \frac{x_{0i} + x_{0j} - \frac{1}{2}}{2}$, and

$$o_j^2(\theta|x_{0i} \in W^*, s_i < 0, x_{0j}, x_{0i}) = \frac{1}{4}c + \frac{3}{4}x_0 - \frac{1}{4}.$$

Therefore, from (14),

$$o_j^2(\theta|x_{0j}, m_{ij} = -1) = \frac{\omega}{\omega + x_{0j} - c} \left(x_{0j} + \frac{\omega}{2} - \frac{1}{4} \right) + \frac{x_{0j} - c}{\omega + x_{0j} - c} \left(\frac{1}{4}c + \frac{3}{4}x_0 - \frac{1}{4} \right).$$

For $m_{ij} = 0$, x_{0i} must be in $W' = [x_{0j} - \omega, c]$. We find $o_j^2(x_{0i}|x_{0i} \in W', x_{0j}) = \frac{x_{0j} - \omega - c}{2}$, and $o_j^2(\theta|x_{0i} \in W', x_{0j}, x_{0i}) = \frac{x_{0j} + x_{0i}}{2}$, which implies that

$$o_j^2(\theta|x_{0j}, m_{ij} = 0) = o_j^2(\theta|x_{0i} \in W', x_{0j}) = \frac{1}{4}(3x_{0j} - \omega + c).$$

(iii) If $c - \omega < x_{0j} < c$ then non-zero messages can only come from $W'' = [c, x_{0j} + \omega]$. Using $o_j^2(\theta|x_{0i} \in W'', s_i > 0, x_{0j}, x_{0i}) = \frac{x_{0i} + x_{0j} + \frac{1}{2}}{2}$, we get

$$o_j^2(\theta|x_{0j}, m_{ij} = 1) = \frac{1}{x_{0j} + \omega - c} \int_c^{x_{0j} + \omega} \frac{x_{0i} + x_{0j} + \frac{1}{2}}{2} dx_{0i} = \frac{1}{4}(c + \omega + 3x_{0j} + 1).$$

For bad news, $o_j^2(\theta|x_{0i} \in W'', s_i < 0, x_{0j}, x_{0i}) = x_{0i} - \frac{1}{4}$, and

$$o_j^2(\theta|x_{0j}, m_{ij} = -1) = \frac{1}{2}c + \frac{1}{2}x_0 + \frac{1}{2}\omega - \frac{1}{4}.$$

Finally, $m_{ij} = 0$ implies $x_{0i} \in W' = [x_{0j} - \omega, c]$ as in case (ii), therefore we again get

$$o_j^2(\theta|x_{0j}, m_{ij} = 0) = \frac{1}{4}(3x_{0j} - \omega + c).$$

A.2 Proof of Proposition 3

Consider first the case of $\delta > \frac{1}{4}$. The following table describes the relevant ranges of x_{0j} , the messages received, and the resulting decisions.

Range	Restriction	Possible messages	Tries New?
1.	$x_{0j} < c - \omega$	$\{0\}$	No
2.	$c - \omega < x_{0j} < c + \frac{\omega}{3}$	$\left\{ \begin{array}{l} 1 \text{ with prob. } \frac{x_{0j} + \omega - c}{2\omega} \\ 0 \text{ otherwise} \end{array} \right\}$	Only if $m_{ij} = 1$
3.	$c + \frac{\omega}{3} < x_{0j} < c + \omega$	$\left\{ \begin{array}{l} 0 \text{ with prob. } \frac{c - (x_{0j} - \omega)}{2\omega} \\ -1 \text{ otherwise} \end{array} \right\}$	Yes
4.	$c + \omega < x_{0j} < \theta - \omega$	$\{1\}$	Yes
5.	$\theta - \omega < x_{0j} < \theta + \omega$	$\left\{ \begin{array}{l} 1 \text{ with prob. } \frac{\theta - (x_{0j} - \omega)}{2\omega} \\ -1 \text{ otherwise} \end{array} \right\}$	Yes
6.	$\theta + \omega < x_{0j}$	$\{-1\}$	Yes

The *experienced-neighbor* and the *inexperienced-neighbor* effects explained in the text appear in range 2. The measure of agents consuming in round 2 can be computed as

$$\Pr(o_j^2(m_{ij}) > c) = \int_{c-\omega}^{c+\frac{\omega}{3}} \frac{x_{0j} - (c - \omega)}{2\omega} dx_{0j} + \frac{1}{2} + \delta - \frac{\omega}{3} = \frac{1}{2} + \delta + \frac{1}{9}\omega,$$

which is higher than the $\frac{1}{2} + \delta$ we saw in section 2. The difference is larger, the larger the neighborhood.

When $\frac{1}{4} - \frac{\omega}{8} < \delta < \frac{1}{4}$, the above table still applies, but the interpretation changes slightly: we now also have a *learning-from-neighbors* effect, for x_{0j} between θ and $c + \frac{1}{4}$ in range 5.

When $\frac{1}{4} - \frac{9}{8}\omega < \delta < \frac{1}{4} - \frac{\omega}{8}$, the above table applies, except that within range 5, for $\theta - \omega < x_{0j} < c + \frac{1}{4} - \frac{\omega}{8}$, the new product will only be consumed if $m_{ij} = 1$. We have the *experienced-neighbor* and the *inexperienced-neighbor* effects in range 2, *misleading-neighbors* effects in range 5, and the *learning-from-neighbors* effect in range 6. We find the measure of consumers as

$$\begin{aligned} \Pr(o_j^2(m_{ij}) > c) &= \int_{c-\omega}^{c+\frac{\omega}{3}} \frac{x_{0j} + \omega - c}{2\omega} dx_{0j} + \int_{c+\frac{\omega}{3}}^{\theta-\omega} dx_{0j} + \int_{\theta-\omega}^{c+\frac{1}{4}-\frac{\omega}{8}} \frac{\theta - (x_{0j} - \omega)}{2\omega} dx_{0j} + \int_{c+\frac{1}{4}-\frac{\omega}{8}}^{\theta+\frac{1}{2}} dx_{0j} \\ &= 2\delta + \frac{1}{4} + \left(\frac{1}{8} + \frac{1}{9}\right)\omega - \left(\frac{1}{4} - \frac{9\omega}{8} - \delta\right)^2 \frac{1}{4\omega}. \end{aligned}$$

One can check that this is larger than the $2\delta + \frac{1}{4}$ of the benchmark case, so that the positive neighborhood effects dominate.

When $\omega < \delta < \frac{1}{4} - \frac{9}{8}\omega$, agents in range 5 will only consume the new product if they get a positive message, and within range 6, agents with $\theta + \omega < x_{0j} < c + \frac{1}{4} - \frac{\omega}{8}$ will never consume it. We observe the *experienced-neighbor* effect (range 2), the *misleading-neighbor* effect (range 5), and the *learning-from-neighbors* effect (range 6). We find the measure of consumers to be

$$\begin{aligned}
\Pr(o_j^2(m_{ij}) > c) &= \int_{c-\omega}^{c+\frac{\omega}{3}} \frac{x_{0j} + \omega - c}{2\omega} dx_{0j} + \int_{c+\frac{\omega}{3}}^{\theta-\omega} dx_{0j} + \int_{\theta-\omega}^{\theta+\omega} \frac{\theta - (x_{0j} - \omega)}{2\omega} dx_{0j} + \int_{c+\frac{1}{4}-\frac{\omega}{8}}^{\theta+\frac{1}{2}} dx_{0j} \\
&= \left(\frac{1}{8} + \frac{1}{9}\right)\omega + 2\delta + \frac{1}{4}
\end{aligned}$$

which is higher than the $2\delta + \frac{1}{4}$ of the benchmark case.

When $0 < \delta < \omega$, we find

Range	Restriction	Message	Tries?
1.	$x_{0j} < c - \omega$	$\{0\}$	No
2.	$c - \omega < x_{0j} < \theta - \omega$	$\left\{ \begin{array}{l} 1 \text{ with prob. } \frac{x_{0j} + \omega - c}{2\omega} \\ 0 \text{ otherwise} \end{array} \right\}$	Only if $m_{ij} = 1$
3.	$\theta - \omega < x_{0j} < c + \frac{\omega}{3}$	$\left\{ \begin{array}{l} -1 \text{ with prob. } \frac{x_{0j} + \omega - \theta}{2\omega} \\ 1 \text{ with prob. } \frac{\theta - c}{2\omega} \\ 0 \text{ otherwise} \end{array} \right\}$	Only if $m_{ij} = 1$
4.	$c + \frac{\omega}{3} < x_{0j} < c + \omega$	$\left\{ \begin{array}{l} -1 \text{ with prob. } \frac{x_{0j} + \omega - \theta}{2\omega} \\ 1 \text{ with prob. } \frac{\theta - c}{2\omega} \\ 0 \text{ otherwise} \end{array} \right\}$	Only if $m_{ij} = 1$ or $m_{ij} = 0$
5.	$c + \omega < x_{0j} < \theta + \omega$	$\left\{ \begin{array}{l} 1 \text{ with prob. } \frac{\theta - (x_{0j} - \omega)}{2\omega} \\ -1 \text{ otherwise} \end{array} \right\}$	Only if $m_{ij} = 1$
6.	$\theta + \omega < x_{0j} < c + \frac{1}{4} - \frac{\omega}{8}$	$\{-1\}$	No
7.	$c + \frac{1}{4} - \frac{\omega}{8} < x_{0j}$	$\{-1\}$	Yes

We see the *experienced-neighbor* effect in range 2 and 3, the *inexperienced-neighbor* effect in range 4 (for $x_{0j} \in [\theta, c + \omega]$), the *misleading-neighbor* effect in range 5, and the *learning-from-neighbors* effect in range 7. The measure of consumers is given by

$$\begin{aligned}
\Pr(o_j^2(m_{ij}) > c) &= \int_{c-\omega}^{\theta-\omega} \frac{x_{0j} + \omega - c}{2\omega} dx_{0j} + \int_{\theta-\omega}^{c+\frac{\omega}{3}} \frac{\theta - c}{2\omega} dx_{0j} + \int_{c+\frac{\omega}{3}}^{c+\omega} \frac{\omega - x_{0j} + \theta}{2\omega} dx_{0j} \\
&\quad + \int_{c+\omega}^{\theta+\omega} \frac{\theta - (x_{0j} - \omega)}{2\omega} dx_{0j} + \int_{c+\frac{1}{4}-\frac{\omega}{8}}^{\theta+\frac{1}{2}} dx_{0j} \\
&= \left[\frac{\omega - c}{2\omega}(\theta - c) + \frac{(\theta - c)(c + \theta - 2\omega)}{4\omega} \right] + \frac{(c - \theta + \frac{4}{3}\omega)(\theta - c)}{2\omega} + \left[\frac{\theta + \omega}{3} - (2c + \frac{4}{3}\omega) \frac{1}{12} \right] \\
&\quad + \left[\frac{\theta + \omega}{2\omega}(\theta - c) - \frac{(\theta - c)(\theta + c + 2\omega)}{4\omega} \right] + \theta + \frac{1}{2} - (c + \frac{1}{4} - \frac{\omega}{8})
\end{aligned}$$

Finally, consider the case of an inferior new product ($\delta < 0$). The various possibilities are given in

the following table:

Range	Restriction	Possible messages	Tries New?
1.	$x_{0j} < c - \omega$	$\{0\}$	No
2.	$c - \omega < x_{0j} < c + \frac{\omega}{3}$	$\{0, -1\}$	No
3.	$c + \frac{\omega}{3} < x_{0j} < c + \omega$	$\left\{ \begin{array}{l} 0 \text{ with prob. } \frac{c - (x_{0j} - \omega)}{2\omega} \\ -1 \text{ otherwise} \end{array} \right\}$	Only if $m_{ij} = 0$
4.	$c + \omega < x_{0j} < c + \frac{1}{4} - \frac{\omega}{8}$	$\{-1\}$	No
5.	$c + \frac{1}{4} - \frac{\omega}{8} < x_{0j}$	$\{-1\}$	Yes

If $-\frac{1}{4} - \frac{\omega}{8} > \delta$, only the *inexperienced-neighbor* effect is at work (in range 3), and we find the measure of consumers to be

$$\Pr(o_j^2 > c) = \int_{c + \frac{\omega}{3}}^{c + \omega} \frac{c - (x_{0j} - \omega)}{2\omega} dx_{0j} = \frac{1}{9}\omega.$$

If $-\frac{1}{4} - \frac{\omega}{8} < \delta$, both the *inexperienced-neighbor* effect and the *learning-from-neighbor* effect (range 5) operates, and we find

$$\Pr(o_j^2 > c) = \frac{1}{9}\omega + \int_{c + \frac{1}{4} - \frac{\omega}{8}}^{\theta + \frac{1}{2}} dx_{0j} = \left(\frac{1}{9} + \frac{1}{8}\right)\omega + \delta + \frac{1}{4},$$

which is also larger than in section 2.

A.3 Proof of Proposition 5

We first compute the density function of Δx_i , $g(\Delta x_i | x_{0i} > c)$:

if $2a - \frac{1}{2} < \delta$,

$$g(\Delta x_i | x_{0i} > c) = \left\{ \begin{array}{l} \frac{a - \Delta x_i + \delta}{a(1+2\delta)} \text{ if } \delta + a > \Delta x_i > \delta - a \\ \frac{2}{2\delta + 1} \text{ if } \delta - a > \Delta x_i > a - \frac{1}{2} \\ \frac{a + \Delta x_i + \frac{1}{2}}{a(1+2\delta)} \text{ if } a - \frac{1}{2} > \Delta x_i > -a - \frac{1}{2} \\ 0 \text{ otherwise} \end{array} \right\}$$

if $2a - \frac{1}{2} > \delta$,

$$g(\Delta x_i | x_{0i} > c) = \left\{ \begin{array}{l} \frac{a - \Delta x_i + \delta}{a(1+2\delta)} \text{ if } \delta + a > \Delta x_i > a - \frac{1}{2} \\ \frac{1}{2a} \text{ if } a - \frac{1}{2} > \Delta x_i > \delta - a \\ \frac{a + \Delta x_i + \frac{1}{2}}{a(1+2\delta)} \text{ if } \delta - a > \Delta x_i > -a - \frac{1}{2} \\ 0 \text{ otherwise} \end{array} \right\}$$

Expected surprise can now be computed as follows:

if $0 \leq a < \frac{1}{4} - \frac{|\delta|}{2}$,

$$\begin{aligned} E(s_i|x_{0i} > c, \theta) &= \int_{\delta-a}^{a+\delta} \Delta x_i \frac{a - \Delta x_i + \delta}{a(1+2\delta)} d\Delta x_i + \int_{a-\frac{1}{2}}^{\delta-a} \Delta x_i \frac{2}{1+2\delta} d\Delta x_i + \int_{-a-\frac{1}{2}}^{a-\frac{1}{2}} \frac{1}{2} \left(\Delta x_i + a - \frac{1}{2} \right) \frac{a + \Delta x_i + \frac{1}{2}}{a(1+2\delta)} d\Delta x_i \\ &= \frac{1}{3} \frac{a^2}{\delta + \frac{1}{2}} + \frac{1}{2} \left(\delta - \frac{1}{2} \right) \end{aligned}$$

if $\frac{1}{4} - \frac{\delta}{2} \leq a < \frac{1}{4} + \frac{\delta}{2}$ (implying that $\delta > 0$),

$$\begin{aligned} E(s_i|x_{0i} > c, \theta) &= \int_{\frac{1}{2}-a}^{a+\delta} \frac{1}{2} \left(\Delta x_i - a + \frac{1}{2} \right) \frac{a - \Delta x_i + \delta}{a(1+2\delta)} d\Delta x_i + \int_{\delta-a}^{\frac{1}{2}-a} \Delta x_i \frac{a - \Delta x_i + \delta}{a(1+2\delta)} d\Delta x_i + \int_{a-\frac{1}{2}}^{\delta-a} \Delta x_i \frac{2}{1+2\delta} d\Delta x_i \\ &\quad + \int_{-a-\frac{1}{2}}^{a-\frac{1}{2}} \frac{1}{2} \left(\Delta x_i + a - \frac{1}{2} \right) \frac{a + \Delta x_i + \frac{1}{2}}{a(1+2\delta)} d\Delta x_i \\ &= -\frac{1}{24} \frac{(2\delta - 1)((\frac{1}{2} - \delta + 3a)^2 - 3a(4 - a))}{a(2\delta + 1)} \end{aligned}$$

if $\frac{1}{4} + \frac{\delta}{2} \leq a < \frac{1}{4} - \frac{\delta}{2}$ (implying that $\delta < 0$),

$$\begin{aligned} E(s_i|x_{0i} > c, \theta) &= \int_{a-\frac{1}{2}}^{a+\delta} \Delta x_i \frac{a - \Delta x_i + \delta}{a(1+2\delta)} d\Delta x_i + \int_{\delta-a}^{a-\frac{1}{2}} \frac{1}{2} \left(\Delta x_i + a - \frac{1}{2} \right) \frac{1}{2a} d\Delta x_i \\ &\quad + \int_{-a-\frac{1}{2}}^{\delta-a} \frac{1}{2} \left(\Delta x_i + a - \frac{1}{2} \right) \frac{a + \Delta x_i + \frac{1}{2}}{a(1+2\delta)} d\Delta x_i \\ &= \text{etc.} \end{aligned}$$

if $\frac{1}{4} + \frac{|\delta|}{2} \leq a$,

$$\begin{aligned} E(s_i|x_{0i} > c, \theta) &= \int_{\frac{1}{2}-a}^{a+\delta} \frac{1}{2} \left(\Delta x_i - a + \frac{1}{2} \right) \frac{a - \Delta x_i + \delta}{a(1+2\delta)} d\Delta x_i + \int_{a-\frac{1}{2}}^{\frac{1}{2}-a} \Delta x_i \frac{a - \Delta x_i + \delta}{a(1+2\delta)} d\Delta x_i \\ &\quad + \int_{\delta-a}^{a-\frac{1}{2}} \frac{1}{2} \left(\Delta x_i + a - \frac{1}{2} \right) \frac{1}{2a} d\Delta x_i + \int_{-a-\frac{1}{2}}^{\delta-a} \frac{1}{2} \left(\Delta x_i + a - \frac{1}{2} \right) \frac{a + \Delta x_i + \frac{1}{2}}{a(1+2\delta)} d\Delta x_i \\ &= -\frac{1}{24} \left[\frac{16(a - \frac{1}{2})^3}{a(2\delta + 1)} - \frac{3}{a} \left(\delta - \frac{1}{2} \right) \right] \end{aligned}$$

The results in the Proposition follow.

References

- [1] Alba, J.W., and H. Marmorstein (1987): "The Effects of Frequency Knowledge On Consumer Decision Making," *Journal of Consumer Decision Making* 14, 14-25.
- [2] Adongo, P.B., J.F. Phillips, B. Kajihara, C. Fayorsey, C. Debpuur, and F. Binka (1997): "Cultural Factors Constraining the Introduction of Family Planning Among the Kassena-Nankana of Northern Ghana," *Social Science and Medicine* 45(12), 1789-1804.
- [3] Banerjee, A., and D. Fudenberg (2004): "Word-of-mouth Learning," *Games and Economic Behavior* 46, 1-22.
- [4] Bernardi, L. (2003): "Channels of Social Influence on Reproduction," *Population Research and Policy Review* 22, 527-555.
- [5] Chamley, C.P. (2004): *Rational Herds. Economic Models of Social Learning*, Cambridge University Press: Cambridge, UK.
- [6] DeMarzo, P.M., D. Vayanos, and J. Zwiebel (2003): "Persuasion Bias, Social Influence, and Unidimensional Opinions," *The Quarterly Journal of Economics*, August, 909-968.
- [7] Ellison, G., and D. Fudenberg (1993): "Rules of Thumb for Social Learning," *Journal of Political Economy* 101(4), 612-643.
- [8] Ellison, G., and D. Fudenberg (1995): "Word-of-Mouth Communication and Social Learning," *The Quarterly Journal of Economics*, February, 93-125.
- [9] Entwisle, B., R.R. Rindfuss, D.K. Guilkey, A. Chamratrithirong, S.R. Curran, and Y. Sawangdee (1996): "Community and Contraceptive Choice in Rural Thailand: A Case Study of Nang Rong," *Demography* 33(1), 1-11.
- [10] Hendricks, K., and D. Kovenock (1989): "Asymmetric Information, Information Externalities, and Efficiency: The Case of Oil Exploration," *RAND Journal of Economics* 20, 164-182.
- [11] Israel, M. (2004): "Services as Experience Goods: An Empirical Examination of Consumer Learning in Automobile Insurance," forthcoming in the *American Economic Review*.
- [12] Januszewski, S. I. (2004): "The Role of Service Quality and Expectations in Explaining Customer Complaints," working paper, University of California at San Diego.
- [13] Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica* 47 (2), 263-291.
- [14] Kohler, H.-P. (1997): "Learning in Social Networks and Contraceptive Choice," *Demography* 34(3), 369-383.

- [15] Kőszegi, B., and M. Rabin (2005): "A Model of Reference-Dependent Preferences," working paper, University of California at Berkeley.
- [16] Kremer, M., and T. Miguel (2004): "Networks, Social Learning and Technology Adoption: The Case of Deworming Drugs in Kenya," working paper, Harvard University.
- [17] Lynch, J.G. Jr., Marmorstein, H., and M.F. Weigold (1988): "Choices from Sets Including Remembered Brands: Use of Recalled Attributes and Prior Overall Evaluations," *Journal of Consumer Research* 15, 169-184.
- [18] Manski, C.F. (2004): "Social Learning from Private Experiences: The Dynamics of the Selection Problem," *Review of Economic Studies* 71, 443-458.
- [19] Morris, S., and H.S. Shin (2002): "Global Games: Theory and Applications," *in: Advances in Economics and Econometrics*, edited by M. Dewatripont, L. Hansen and S. Turnovsky, Cambridge University Press: Cambridge, UK.
- [20] Munshi, K., and J. Myaux (2002): "Development as a Process of Social Change: An Application to the Fertility Transition," working paper, University of Pennsylvania.
- [21] Oliver, R.C. (1980): "A Cognitive Model on the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research* 17, 460-469.
- [22] Romani, M. (2003): "Love Thy Neighbour? Evidence form Ethnic Discrimination in Information Sharing within Villages in Cote d'Ivoire," *Journal of African Economics* 12(4), 533-563.
- [23] Shiller, R.J. (1995): "Conversation, Information, and Herd Behavior," *American Economic Review (AEA Papers and Proceedings)* 85(2), 181-185.
- [24] Szymanski, D.M., and Henard, D.H. (2004): "Customer Satisfaction: A Meta-Analysis of the Empirical Evidence," *Journal of the Academy of Marketing Science* 29(1), 16-35.
- [25] Van den Steen, E. (2004): "Rational Overoptimism (and Other Biases)," *American Economic Review* 94(4), 1141-1151.