Technical Appendix

Proof of Lemma 1

An individual maximizes her expected utility given her information set I_n. If the expected utility from adoption is greater than 0, then she will adopt; otherwise she will reject. In other words, she will choose to adopt if $E[(V-(1-t_n)|I_n]>0$. Since t_n does not depend on the information an individual has, the inequality is equivalent to $E(V|I_n)>1-t_n$. And $E(V|I_n)=P(V=1|I_n)\times 1+P(V=0|I_n)\times 0=P(V=1|I_n)$. Therefore, for any individual, given the information set I_n she has, if the conditional (posterior) probability that the value is one is greater than the disutility (transportation cost as 1-t_n), the individual will choose to adopt. Otherwise, she will choose to reject. □

Relax *t* to 1-*p* < *t* < ½ and ½ < *t* < *p* in Friend-networks

We start with the case for ½ < t < p. If the first person sees H, then she calculates P(V=1| H) = p > 1-t. If the first person sees L, then she calculates P(V=1|L) = 1- p < 1 - t. Therefore, the first person will adopt when seeing an H signal and reject when seeing an L signal. If the second person sees AH, then P(V=1 | AH) > p > 1 - t, and she will adopt as well. If the second person sees AL, then P(V=1 | AL) = ½ > 1 - t, then she will adopt as well. This means that no matter what the second signal is, the second person will adopt as long as the first person adopts. Then a cascade will have happen.

If the second person sees RH, then she calculates P(V=1|RH)= ½ > 1-t, then she will adopt. If the second person sees RL, then she calculates P(V=1|RL) = $\frac{(1-p)^2}{(1-p)^2+p^2}$ < 1-p < 1-t, then she will reject. If the third person sees RA, she knows that it must be that the second person sees H, and the third person's decision rule is like the 1st person's. To see this,

P(V=1|RAH) = $\frac{(1-p)p^2}{(1-p)p^2+p(1-p)^2}$ = p > 1-t, and she will adopt.

P(V=1|RAL) = $\frac{p(1-p)^2}{(1-p)p^2+p(1-p)^2}$ = 1-p < 1-t, and she will reject.

If the third person sees RRH, P(V=1|RRH) = $\frac{p(1-p)^2}{(1-p)p^2 + p(1-p)^2}$ = 1-p < 1-t, and she will reject. It can be easily verified that P(V=1|RRL) < 1-t and the third person will reject as well. Therefore, a cascade will happen.

Compared to the case when t = ½, an up cascade will be more likely to happen in the case when ½ < t < p. This is because t= ½ is the case when people are most neutral to a product in terms of preferences. Two consecutive adoptions or rejections will give rise to cascade when t = ½. But when ½ < t < p and people have upward biased preference for the product, as long as the person in the odd number sequence adopts, an up cascade will happen.

Symmetrically, when 1-p < t < ½ , it can be similarly prove that as long as the person in the odd number sequence rejects, a rejection cascade will happen.

Proof of Lemma 3

Suppose an "up" information cascade starts at the n^{th} person who adopts the product. Based on the definition of cascades, all later people will choose to adopt regardless of their signals. For the $(n+1)^{th}$ person, according to Lemma 1, a cascade will continue on this person iff t_{n+1} > 1 - P(V=1| x_n=A, s_{n+1}=L). Similarly, for the $n+2^{th}$ person, a cascade will continue on her iff t_{n+2} > 1 - P(V=1| x_n=x_{n+1}=A, S_n+2=L); and for the $n+k^{th}$ person, a cascade will continue on her iff t_{n+k} > 1 - P(V=1| x_n=x_{n+1}=…=x_{n+k-1}=A, S_{n+k}=L). Let P_1= P(V=1| x_n=A, s_{n+1}=L), P_2= P(V=1| x_n=x_{n+1}=A, S_n+2=L), …, P_k= P(V=1| x_n=x_{n+1}=…=x_{n+k-1}=A, S_{n+k}=L). Therefore, P[t_{n+1} > 1 - P(V=1| x_n=A, s_{n+1}=L)]= P(V=1| x_n=A, s_{n+1}=L)=P_1, …, P[t_{n+k} > 1 - P(V=1| x_n=x_{n+1}=…=x_{n+k-1}=A, S_{n+k}=L)]=P(V=1| x_n=x_{n+1}=…=x_{n+k-1}=A, S_{n+k}=L)=P_k. Thus, the probability that a cascade starts at the n^{th} person in a stranger-network equals to $P(cascade\ at\ n) = \prod_{k=1}^{\infty} P_k$. It can be easily verified that $P_1 < P_2 < \ldots < P_k$, thus $P_1^k \le P(cascad\ after\ n) \le P_k^k$. Since $0 < P_k < 1\ \forall\ k \in N$, thus $\lim_{k\to\infty} P_1^k = \lim_{k\to\infty} P_k^k = 0$. Therefore, $\lim_{k\to\infty} \prod_{k=1}^{\infty} P_k = 0$ according to squeeze theorem. That is, the probability of a cascade happens in a stranger-network converges to zero as the number of people n increases to infinite. □

Proof of Proposition 1

The general idea of this proof consists of two parts. The first part of proof is that when $n = 2$, $P(correct)_s < P(correct)_f$. Here a correct inference is defined as that the n^{th} individual correctly infers the quality of the product. The n^{th} individual could face 2^n different information sets, so we use I_k to represent the k^{th} possible information set for the n^{th} individual. Mathematically, the probability of making a correct inference is:

$$P(correct) = P(V=1)\sum_{k=1}^{2^n} P(I_k \mid V=1)P(V=1 \mid I_k) + P(V=0)\sum_{k=1}^{2^n} P(I_k \mid V=0)P(V=0 \mid I_k) =$$

$\sum_{k=1}^{2^n} P(I_k \mid V=1)P(V=1 \mid I_k)$. The second part of proof is to prove that the probability of making a correct inference in a stranger-network, $P(correct)_s$ converges to 1 when $N \to \infty$ but that in a friend-network $P(correct)_f$ converges to a value strictly less than 1 and both these two probabilities are monotonic increasing in N. Therefore there must be one point when these two curves intersect with each other.

To begin with, we are proving that for the second individual, the probability of making a correct inference is higher in friend-networks than that in stranger-networks. From Section 3.2 and 3.3, it can be shown that the probability for the second individual to make a correct inference in friend and stranger-networks are: $P(correct)_f = \frac{p^4 + (1-p)^4}{p^2 + (1-p)^2} + p(1-p)$, and $P(correct)_s =$

$\frac{24p^4 - 48p^3 + 40p^2 - 16p + 3}{(4p^2 - 2p + 1)(4p^2 - 6p + 3)}$. It can be verified that $P(correct)_{2,f} > P(correct)_{2,s}$ when $½ < p < 1$. To improve readability as well to save pages, we omit this proof. The detailed technical proof is available from the authors upon request.

The next step is to prove that the probability of making a correct inference in a stranger-network converges to c as $N \to \infty$. Several conditions are required for the proof. First is that the probability of making a correct inference $P(correct)_s$ is monotonic increasing in N. The second is that $P(correct)_s$ is bounded. Therefore, according to the monotone convergence theorem, $P(correct)_s$ converges to c. To prove that $P(correct)_s$ is bounded in the range of (0, 1):

Recall that $P(correct)_s = P(V=1)\sum_{k=1}^{2^n} P(I_k \mid V=1)P(V=1 \mid I_k) + P(V=0)\sum_{k=1}^{2^n} P(I_k \mid V=0)P(V=0 \mid I_k) =$

$\sum_{k=1}^{2^n} P(I_k \mid V=1)P(V=1 \mid I_k)$. Because $0 < \sum_{k=1}^{2^n} P(I_k \mid V=1)P(V=1 \mid I_k) < \sum_{k=1}^{2^n} P(I_k \mid V=1) = 1$, we get

$0 < \sum_{k=1}^{2^n} P(I_k \mid V=1)P(V=1 \mid I_k) < 1$.

Next is to prove that $P(correct)_s$ is monotonic increasing in N. For the n^{th} individual, she has 2^n possible information sets in $\boldsymbol{I}_n$. For the next individual -- the $(n+1)^{th}$ individual, she has 2^{n+1} possible information sets in $\boldsymbol{I}_{n+1}$. Suppose one information set for the n^{th} person is $I_k = \{X_{n-1}, s_n\}$, where X_{n-1} is the actions made by previous n-1 people. Then, the $(n+1)^{th}$ individual's corresponding information is $I_m=\{X_{n-1}, x_n, s_{n+1}\}$. where x_n is the action of the n^{th} person, either adopt or reject. It can be proved that $P(correct, I_k) < P(correct, I_m)$, where

$P(correct, I_k) = P(V=1)P(I_k \mid V=1)P(V=1 \mid I_k) + P(V=0)P(I_k \mid V=0)P(V=0 \mid I_k)$, and

$P(correct, I_m) = P(V=1)P(I_m \mid V=1)P(V=1 \mid I_m) + P(V=0)P(I_m \mid V=0)P(V=0 \mid I_m)$.[1] Then similar results can be obtained for the rest of $2^n - 1$ possible information sets for the n^{th} individual and the corresponding $2^{n+1}-2$ sets for the $(n+1)^{th}$ individual.

From above, we know that $\lim_{n\to\infty} P(Correct)_s = c$. Now we need to prove that $c = 1$. To prove $\lim_{n\to\infty} P(Correct)_s = 1$, according to Lobel (2009), when the preference is unbounded ($t_n \in [0, 1]$), social beliefs about the true state of the world will converge to 1 and here we replicate the proof. It can be shown that $\lim_{n\to\infty} P(Correct) =$

[1] The proof is quite lengthy. To improve readability as well to save pages, we omit this proof. The detailed technical proof is available from the authors upon request.

$\frac{(-\overline{\alpha}+\underline{\beta})\overline{\beta}(-1+\underline{\alpha})}{-\overline{\alpha}\underline{\alpha}\overline{\beta}+\overline{\alpha}\overline{\beta}+\overline{\alpha}\underline{\beta}\underline{\alpha}-\overline{\alpha}\underline{\beta}\overline{\beta}-\underline{\beta}\underline{\alpha}+\underline{\beta}\underline{\alpha}\overline{\beta}}$, where $\overline{\alpha}$ and $\underline{\alpha}$ are the upper and lower bound of the support of preference t, and $\overline{\beta}$ and $\underline{\beta}$ are the upper and lower bound of the support of private signal *s*. In a stranger-network, $\overline{\alpha}$ = 1 and $\underline{\alpha}$ = 0 while in a friend-network, $\overline{\alpha}$ = $\underline{\alpha}$ = ½. In both networks, $\overline{\beta}$<1 and $\underline{\beta}$ >0. Then in a stranger-network, it can be shown that $\lim_{n\to\infty}P(Correct)_s = 1$ when $\overline{\alpha}$ = 1 and $\underline{\alpha}$ = 0, while in a friend-network, $\lim_{n\to\infty}P(Correct)_f = c'$ where $c' < 1$ when $\overline{\alpha}$ = $\underline{\alpha}$ = ½. □

Proof of Proposition 2

Expected Sales in a Friend-network

A cascade will happen after two people choose the same action in a friend-network. We use this property of a friend-network to calculate the expected sales in a friend-network. Here we assume that the quality is high, *V*=1. Results can be similarly obtained for *V*=0 using the same algorithm.

Let P_1 be the probability that two people choose two different actions when *V*=1 (either AR or RA), thus P_1 = P(AR|*V*=1) + P(RA|*V*=1) = $p(1-p)$. Let P_2 be the probability that two people both choose to adopt when *V*=1 (AA), thus P_2 = P(AA|*V*=1) = $p^2 + ½\, p(1-p)$. Let P_0 be the probability that two people both choose to reject when *V*=1 (RR), thus P_0=P(RR|*V*=1) = $(1-p)^2 + ½\, p(1-p)$.

When there are even number (2×*i*) of people in a friend-network, the expected sales is:

$$sales_f^{2\times i} = P_1 \times P_0 + 2P_1^2 \times P_0 + 3P_1^3 \times P_0 + 4P_1^4 \times P_0 + ... + (i-1)P_1^{i-1} \times P_0 + iP_1^i + (i+1)P_1^{i-1}P_2 + (i+2)P_1^{i-2}P_2 + ... + (i+i)P_2 = iP_1^i + \sum_{j=1}^{i-1} j \times P_1^j \times P_0 + \sum_{j=i+1}^{2\times i} j \times P_1^{2\times i-j} P_2$$

When there are odd number (2×*i*-1) of people in a friend-network, the expected sales is:

$$sales_f^{2\times i\text{-}1} = P_1 \times P_0 + 2P_1^2 \times P_0 + 3P_1^3 \times P_0 + 4P_1^4 \times P_0 + ... + (i\text{-}2)P_1^{i\text{-}2} \times P_0 + (i\text{-}1)P_1^{i\text{-}1}(1\text{-}p) + iP_1^{i\text{-}1}p + (i+1)P_1^{i\text{-}2}P_2$$

$$+ (i+2)P_1^{i\text{-}3}P_2 + ... + (2\times i\text{-}1)P_2 = (i\text{-}1)P_1^{i\text{-}1}(1\text{-}p) + iP_1^{i\text{-}1}p + \sum_{j=1}^{i-2} j \times P_1^j \times P_0 + \sum_{j=i+1}^{2\times i\text{-}1} j \times P_1^{2\times i\text{-}1\text{-}j} \times P_2$$

Expected Sales in a Stranger-network

Suppose the product quality is high. Given n people in a stranger-network, there are 2^n possible combinations of choices made by these n people (each person can choose either to adopt or reject, and there are n people). The expected sales in is obtained by first multiplying the number of 'As' (adoptions) in a possible choice combination C_i of these n people by the conditional probability of such a choice to happen given V=1, then summating across all the 2^n possibilities: $sales_s = \sum_{i=1}^{2^n}[k_i \times P(C_i \mid V = 1)]$. Here k_i is the number of adoptions in choice C_i, $P(C_i \mid V = 1)$ is the conditional probability for C_i to occur given that V=1. For example, there are two people in the network. Thus AA (both adopt) is a possible choice made by these two people. k_i thus equals to 2 in this case because two people choose to adopt in the choice set. The probability of AA to happen in the stranger-network given that V=1 is P(AA|V=1) = $[p^2 + (1\text{-}p)^2]^2[\frac{p}{p^2 + 3(1\text{-}p)^2} + \frac{1\text{-}p}{3p^2 + (1\text{-}p)^2}]$. We calculate the conditional probability for each possible choice made by two people, P(AA|V=1), P(AR|V=1), P(RA|V=1) and P(RR|V=1). Then we multiply each probability by the number of adoptions in each choice, and then sum them up to get the expected sales for a high quality product in a network of two people.

To get the expected sales at n, in a more general term, the conditional probability of a specific choice C_i made by these n people given that V=1 is: $P(C_{i,n}|V=1)$ =

$$\begin{cases} P(C_{i,n-1} \mid V = 1) \times P(V = 1 \mid C_{i,n-1}, s_n) \times P(s_n \mid V = 1) & \textit{if the } n^{th} \textit{ person adopts} \\ P(C_{i,n-1} \mid V = 1) \times [1 - P(V = 1 \mid C_{i,n-1}, s_n)] \times P(s_n \mid V = 1) & \textit{if the } n^{th} \textit{ person rejects} \end{cases}$$

where $C_{i,n-1}$ indicates the choices made by the previous n-1 people in choice $C_{i,n}$. For example, if $C_{i,3}$ ={AAR} for the first three people, then $C_{i,3-1}$ is {AA} made by the previous two people. s_n is the quality signal the n^{th} person observes, either H or L.

Since $P(V=1|\ C_{i,n-1},s_n)= \frac{P(V=1)P(C_{i,n-1}\mid V=1)P(s_n\mid V=1)}{\sum_{j=0}^{1}P(V=j)P(C_{i,n-1}\mid V=j)P(s_n\mid V=j)}$, we can express $P(C_{i,n}|V=1)$ in terms of $P(C_{i,n-1}|V=1)$ and $P(C_{i,n-1}|V=0)$, and $P(C_{i,n-1}|V=1)$ in terms of $P(C_{i,n-2}|V=1)$ and $P(C_{i,n-2}|V=0)$, and so forth. Based on these recursive relations, we could express $P(C_{i,n}|V=1)$ in terms of $P(C_{i,1}|V=1)$ and $P(C_{i,1}|V=0)$.

Next, by multiplying the number of adoptions in $C_{i,n}$ by the conditional probability $(C_{i,n}\ |V=1)$, we get the expected sales in one specific choice, $C_{i,n}$. After summating the expected sales across all the 2^n possible choices, we get the expected sales in a stranger-network with size n. Because of the exponentially growing number of items in the expression (2^n conditional probabilities for a network size of n), it becomes computationally intensive and thus extremely complex to analytically solve the expected sales as n increases. So we use numerical simulation to explore the results. Notice that the computer run-time and memory required scale with at 2^n -- it doubles at each increment of number of people. So we use the supercomputing facilities to solve the computation and memory constraints. When n = 40, it takes about 20 days run-time and when n = 50 it will take about 65 *years*. So the maximum N we provide in Figure 2 is 40.

Based on the algorithm described above for $sale_s$ and $sales_f$, we calculate the sales in these two networks when n ranges from 2 to 40 and p ranges from 0.5 to 1. The equilibrium outcome (curve AB in Figure 2) is the threshold of network size $\overline{N}$ when a stranger-network generates higher sales than a friend-network does. As shown in Figure 2, the larger the network size is, the more likely the stranger-network generates higher sales than the friend-network does. And the higher p is, the smaller $\overline{N}$ is.□

Proof of Proposition 3

We assume that in both networks, the expert is neutral to the product, t_e = ½, and this is common knowledge. First consider the case of a friend-network. The proof is to show that after adding an expert into the beginning of networks, the third and later people could be worse off in a friend-network. The same as in the basic model, friends have the same preference for the product, t = ½ for all non-experts. The signal accuracy for the expert's is p_e and the non-experts' is p, where $½ < p < p_e < 1$. Here subscript e indicates the case with an expert. Then the probability of the expert making a correct inference is $P(correct)_e = P(H|V{=}1)P(V{=}1|H) + P(L|V{=}1)P(V{=}1|L) = p_e^2 + (1-p_e)^2$.

Because the second person knows that the expert has better information than she does, she will follow the expert's action no matter what signal she observes. Technically, this means that $P(V{=}1|\ x_e{=}A,\ s_2{=}L) > ½$, which equals to $\frac{p_e(1-p)}{p_e(1-p)+(1-p_e)p} > 1/2$. Since $p_e > p$, $\rightarrow p_e(1-p) > (1-p_e)p \rightarrow 2p_e(1-p) > p_e(1-p) + (1-p_e)p \rightarrow \frac{p_e(1-p)}{p_e(1-p)+(1-p_e)p} > 1/2$. Therefore, we have proven that the second person will follow the expert's action no matter what signal she observes.

Therefore, a cascade starts even sooner than the case without an expert in the friend-network! All the later individuals will thus make the same decision as the expert does and no more useful information will aggregate after the expert. And the probability of everyone making a correct quality inference is $P(correct)_e = p_e^2 + (1-p_e)^2$.

In contrast, if there is no expert in the friend-network, the probability of the third person making a correct quality inference is $P(correct) = -(32p^6 - 96p^5 + 106p^4 - 52p^3 - 9p^2 + 19p - 6)/(3\times(5p^2 - 5p + 2))$. It can be shown that there exists p satisfying $P(correct) > P(correct)_e$ when $½ < p < p_e < 1$. For instance, when p =0.7, and p_e=0.75, $P(correct)$=.6545 while $P(correct)_e$=.625. Therefore, the third person is better off in the network without an expert than in the network with an expert. For the fourth and later decision makers, because their probability of making a correct decision is no smaller than the third individual, they will also be better off without an expert.[2]

Now consider the stranger-network. The roadmap consists of two parts. First, the second person is more likely to make a correct quality inference in the stranger-network with an expert than in the network without an expert. Second, assume that for the nth person, the probability of making a correct quality inference is greater in the network with an expert than in the network without an expert, then it also holds for the (n+1)th person in the stranger-network.

First, we are showing that, with an expert in the network, the second person is more likely to make a correct quality inference than the case without an expert. The second person's probability of making a correct quality inference without an expert is $P(correct)$ = $P(A_1H_2|V=1)P(V=1|A_1H_2)+P(A_1L_2|V=1)P(V=1|A_1L_2)+P(R_1H_2|V=1)P(V=1|R_1H_2)+P(R_1L_2|V=1)P(V=1|R_1L_2) = \frac{24p^4-48p^3+40p^2-16p+3}{(4p^2-2p+1)(4p^2-6p+3)}$.

When there is an expert in the network, the second person's probability of making a correct quality inference is $P(correct)_e = \frac{6p^2p_e^2-6p^2p_e+p^2-6pp_e^2+6pp_e-p+p_e^2-p_e}{(2pp_e-p_e-p)(2pp_e-p_e-p+1)}$.

So $P(correct)_e - P(correct) = \frac{-2p(p-1)(2p^2-2p+p_e)(2p^2-2p-p_e+1)}{(4p^2-2p+1)(4p^2-6p+3)(2pp_e-p_e-p)(2pp_e-p_e-p+1)}$.

It can be verified that $P(correct)_e - P(correct) > 0$ when $½ < p < p_e < 1$.[3]

Now we are showing that, if for the nth person in the stranger-network, $P_{n,e}(correct) > P_n(correct)$, then for the (n+1)th person, it also holds that $P_{n+1,e}(correct) > P_{n+1}(correct)$. To facilitate the proof, we write out the nth person's information into two parts: the social information X, which includes all the actions made by previous n-1 people, and private signal s_n.

---

[2] Because the fourth and later decision makers observe more information than the third one, their probability of making a correct inference is no smaller than the third one. To improve readability as well to save pages, we omit this proof. The detailed technical proof is available from the authors upon request.

[3] To improve readability as well to save pages, we omit this proof. The detailed technical proof is available from the authors upon request.

Given any information set $\{X_k, s_n\}$ that the n[th] person could possibly observe, for the (n+1)[th] person, the corresponding social information she observes is either $\{X_k, A\}$ (if the n[th] person adopts), or $\{X_k, R\}$ (if the (n+1)[th] person rejects). Thus, we could write the corresponding information set the (n+1)[th] person observes as $\{ X_k, x_n, s_{n+1}\}$, where x_n indicates the action made by the n[th] person, either A or R.

Since $P_{n,e}(correct) > P_n(correct)$, this means that

$$\sum_k P(X_k, s_n \mid V=1)_e P(V=1 \mid X_k, s_n)_e + \sum_k P(X_k, s_n \mid V=0)_e P(V=0 \mid X_k, s_n)_e >$$
$$\sum_k P(X_k, s_n \mid V=1) P(V=1 \mid X_k, s_n) + \sum_k P(X_k, s_n \mid V=0) P(V=0 \mid X_k, s_n)$$

or equivalently, $\sum_k \frac{P(X_k, s_n | V=1)_e{}^2 + P(X_k, s_n | V=0)_e{}^2}{\sum_j P(X_k, s_n | V=j)_e} > \sum_k \frac{P(X_k, s_n | V=1)^2 + P(X_{n-1}, s_n | V=0)^2}{\sum_j P(X_k, s_n | V=j)}$, or

equivalently,

$$\sum_k \frac{P(X_k \mid V=1)_e{}^2 P(s_n \mid V=1)_e{}^2 + P(X_k \mid V=0)_e{}^2 P(s_n \mid V=0)_e{}^2}{\sum_j P(X_k \mid V=j)_e P(s_n \mid V=j)_e} >$$
$$\sum_k \frac{P(X_k \mid V=1)^2 P(s_n \mid V=1)^2 + P(X_k \mid V=0)^2 P(s_n \mid V=0)^2}{\sum_j P(X_k \mid V=j) P(s_n \mid V=j)}$$

Also we can write the probability of the (n+1)[th] person making a correct quality inference as follows:

$$P_{n+1}(correct) = \sum_k P(X_k, x_n, s_{n+1} \mid V=1) P(V=1 \mid X_k, x_n, s_{n+1}) + \sum_k P(X_k, x_n, s_{n+1} \mid V=0) P(V=0 \mid X_k, x_n, s_{n+1})$$

$$= \sum_k \frac{P(X_k \mid V=1)^2 P(x_n \mid X_k)^2 P(s_{n+1} \mid V=1)^2}{\sum_j P(X_k \mid V=j) P(x_n \mid X_k) P(s_{n+1} \mid V=j)} + \sum_k \frac{P(X_k \mid V=0)^2 P(x_n \mid X_k)^2 P(s_{n+1} \mid V=0)^2}{\sum_j P(X_k \mid V=j) P(x_n \mid X_k) P(s_{n+1} \mid V=j)}$$

$$= \sum_k \frac{P(X_k \mid V=1)^2 P(x_n \mid X_k) P(s_{n+1} \mid V=1)^2}{\sum_j P(X_k \mid V=j) P(s_{n+1} \mid V=j)} + \sum_k \frac{P(X_k \mid V=0)^2 P(x_n \mid X_k) P(s_{n+1} \mid V=0)^2}{\sum_j P(X_k \mid V=j) P(s_{n+1} \mid V=j)}$$

$$= \sum_k \frac{P(X_k \mid V=1)^2 P(s_{n+1} \mid V=1)^2 [(P(A \mid X_k) + P(R \mid X_k)]}{\sum_j P(X_k \mid V=j) P(s_{n+1} \mid V=j)}$$

$$+\sum_k \frac{P(X_k \mid V=0)^2 P(s_{n+1} \mid V=0)^2[(P(A \mid X_k)+P(R \mid X_k)]}{\sum_j P(X_k \mid V=j)P(s_{n+1} \mid V=j)}$$

$$=\sum_k \frac{P(X_k \mid V=1)^2 P(s_{n+1} \mid V=1)^2}{\sum_j P(X_k \mid V=j)P(s_{n+1} \mid V=j)}+\sum_k \frac{P(X_k \mid V=0)^2 P(s_{n+1} \mid V=0)^2}{\sum_j P(X_k \mid V=j)P(s_{n+1} \mid V=j)}$$

$$=\sum_k \frac{P(X_k \mid V=1)^2 P(s_{n+1} \mid V=1)^2 + P(X_k \mid V=0)^2 P(s_{n+1} \mid V=0)^2}{\sum_j P(X_k \mid V=j)P(s_{n+1} \mid V=j)}$$

Since private signal s_n is i.i.d., and we already know that

$$\sum_k \frac{P(X_k \mid V=1)_e{}^2 P(s_n \mid V=1)_e{}^2 + P(X_k \mid V=0)_e{}^2 P(s_n \mid V=0)_e{}^2}{\sum_j P(X_k \mid V=j)_e P(s_n \mid V=j)_e} >$$
$$\sum_k \frac{P(X_k \mid V=1)^2 P(s_n \mid V=1)^2 + P(X_k \mid V=0)^2 P(s_n \mid V=0)^2}{\sum_j P(X_k \mid V=j)P(s_n \mid V=j)}$$

Thus we can also obtain

$$\sum_k \frac{P(X_k \mid V=1)_e{}^2 P(s_{n+1} \mid V=1)_e{}^2 + P(X_k \mid V=0)_e{}^2 P(s_{n+1} \mid V=0)_e{}^2}{\sum_j P(X_k \mid V=j)_e P(s_{n+1} \mid V=j)_e} >$$
$$\sum_k \frac{P(X_k \mid V=1)^2 P(s_{n+1} \mid V=1)^2 + P(X_k \mid V=0)^2 P(s_{n+1} \mid V=0)^2}{\sum_j P(X_k \mid V=j)P(s_{n+1} \mid V=j)}$$

or equivalently, $P_{n+1,e}(correct) > P_{n+1}(correct)$.□

Proof of Lemma 4

An individual maximizes her expected utility given her information set I_n. If the expected utility from adoption is greater than 0, then she will adopt; otherwise she will reject. In other words, she will choose to adopt if $E[(V-(1-t_n) \mid I_n] > 0$. Since t_n does not depend on the information an individual has, thus the inequality is equivalent to $E(V \mid I_n) > 1-t_n$.

$E(V \mid I_n) = \theta \cdot P(V=\theta \mid I_n) + 0 \cdot P(V=0 \mid I_n) = \theta \cdot P(V=\theta \mid I_n)$. Therefore, for any individual, given the information set I_n she has, if the product of θ and conditional (posterior) probability that the value is θ is greater than the disutility (transportation cost as 1- t_n), this individual will adopt. Otherwise, she will reject. □

Proof of Proposition 4

The roadmap of the proof consists of two parts. First, when n=2, $P(correct)_s < P(correct)_f$. Second, when $n \rightarrow \infty$, $\lim_{n\to\infty} P(Correct)_s > \lim_{n\to\infty} P(Correct)_f$. Because probabilities of making a correct choice are non-decreasing in both networks, therefore, there must exists $\overline{N}$, when $n < \overline{N}$, $P(correct)_s < P(correct)_f$, and when $n > \overline{N}$, $P(correct)_s > P(correct)_f$.

In the friend-network, to exclude uninteresting result, we only look at the range of t when $1 - \theta p < t_n < 1 - \theta(1 - p)$. This is because when $t_n < 1 - \theta p$, the first individual will reject the product no matter what private signals she observes. Thus the first individual's action does not reveal any information to later ones. Then the second individual will act as the first one. For the same reason, the second individual will reject regardless of her private signals. So the third, fourth and so forth. Then in the whole friend-network, everyone will reject. Similarly, when $t_n > 1 - \theta(1 - p)$, everyone in the friend-network will adopt.

We first look at $t = 1- ½ \theta$ for the friend-network. Similar results can be extended for $1 - \theta p < t_n < 1- ½ \theta$ and $1- ½ \theta \leq t_n < 1 - \theta (1 - p)$. When the first person sees H, he calculates $P(V= \theta|H) = p$. Because $\theta p > 1- t$ when $t = 1- ½ \theta$, the first person will adopt after seeing H. When the first person sees L, he calculates $P(V= \theta|L)=1-p$. Because $\theta(1-p) < 1- t$, the first person will reject after seeing L. When the second person sees AH, he calculates $P(V= \theta|AH) = \frac{\theta p^2}{p^2 + (1-p)^2} > ½ \theta$, thus he will adopt. When the second person sees AL, because $(V= \theta|AL)= ½ = 1-t$, the second person will flip a coin to decide. Similarly, we can get results for RH and RL. The second person will reject when seeing RL; she will reject or adopt with equal probabilities when seeing RH.

For the third person, if she observes AAL, she calculates P(V= θ|AAL) = $\frac{1}{3}(1+p)$.

Because $\frac{1}{3}$ θ (1+p) > ½ θ, she will adopt when see AAL. It is easy to verify that P(V= θ|AAH) > ½ as well, thus the third person will adopt when seeing AAH as well. Therefore, a cascade will happen when the third person see AA. Symmetric results can be easily obtained for RR. Therefore, when θ has been added into the model, the learning process is the same as that in the basic model of friend-network.

In the stranger-network, because $0 < \theta < 1$ and $t_n \sim U(0, 1)$, for those $t_n < 1- \theta$, they will get negative utility even the product is of high quality. Thus, they will reject no matter what information they receives. For those people, learning does not matter for them, for their "correct" choice is always to reject. We call these people as "stubborn" consumers. For "non-stubborn" consumers with $t_n > 1- \theta$, learning matters for them because the quality of the product will make a difference on valence of their utility. And we focus on the learning of "non-stubborn" consumers and their probability of making a correct inference. For "non-stubborn" consumers, a correct choice is to adopt when the product is of high quality and to reject when the product is of low quality. And we compare the probabilities of making a correct inference in the stranger-network with that in the friend-network.

In the stranger-network, if the first person sees H, she calculates P(V= θ|H) = p. If $\theta p > 1-$ t, she will adopt. Otherwise, she will reject. If the first person sees L, she calculates P(V= θ|L) = 1-p. If $\theta(1-p) > 1-t$, she will adopt. Otherwise, she will reject. The second person could possibly see four different sets of information, I_n. If the second person sees AH, she calculate P(V= θ|AH) = $\frac{p^2+(1-p)^2}{p^2+3(1-p)^2}$. If the second person sees AL, she calculate P(V= θ|AL) = $\frac{p^2+(1-p)^2}{3p^2+(1-p)^2}$. If the second person sees RH, she calculates P(V= θ|RH) = $\frac{p(2p^2\theta^2-2p\theta^2+\theta^2-1)}{4p^3\theta^2-6p^2\theta^2+3p\theta^2-1}$. And if the second person sees RL, she calculates = P(V= θ|RL)=

$\frac{(1-p)(1-\theta^2\{p^2+(1-p)^2]\}}{(1-p)(1-\theta^2\{p^2+(1-p)^2]\}+p[1-2\theta^2 p(1-p)]}$. And if $\theta \times P(V=\theta| I_n) > 1-t$, she will adopt. Otherwise, she will reject.

The probability of the second person (non-stubborn) making a correct quality inference in the stranger-network is $P(correct)_s$=

$P(V=\theta)\sum_{k=1}^{4} P(I_k \mid V=\theta)P(V=\theta \mid I_k)+P(V=0)\sum_{k=1}^{4} P(I_k \mid V=0)P(V=0 \mid I_k)$ = $(384p^{10}q^2 - 1920p^9q^2 + 4384p^8q^2 - 6016p^7q^2 + 5480p^6q^2 + 48p^6q - 32p^6 - 3448p^5q^2 - 144p^5q + 96p^5 + 1508p^4q^2 + 192p^4q - 136p^4 - 448p^3q^2 - 144p^3q + 112p^3 + 84p^2q^2 + 67p^2q - 58p^2 - 8pq^2 - 19pq + 18p + 3q - 3)/((4p^2 - 2p + 1)(4p^2 - 6p + 3)(4qp^3 - 6qp^2 + 3qp - 1)(4qp^3 - 6qp^2 + 3qp - q + 1))$

And the probability of the second person making a correct inference in the friend-network is $P(correct)_f = \frac{p^4+(1-p)^4}{p^2+(1-p)^2}+p(1-p)$. It can be verified that $P(correct)_s < P(correct)_f$ for the second person.

Now we are proving that $\lim_{n\to\infty}P(Correct)_f < \lim_{n\to\infty}P(Correct)_s$. Because we focus on non-stubborn consumers in the stranger-network, thus $1-\theta < t_n < 1$. Replace $\underline{\alpha}=1-\theta$, $\overline{\alpha}=1$ in

$\frac{(-\overline{\alpha}+\underline{\beta})\overline{\beta}(-1+\underline{\alpha})}{-\overline{\alpha}\underline{\alpha}\overline{\beta}+\overline{\alpha}\overline{\beta}+\overline{\alpha}\underline{\beta}\underline{\alpha}-\overline{\alpha}\underline{\beta}\overline{\beta}-\underline{\beta}\underline{\alpha}+\underline{\beta}\underline{\alpha}\overline{\beta}}$ (A6), we get $\lim_{n\to\infty}P(Correct)_s=1$. For the friend-network, replace $\underline{\alpha}=\overline{\alpha}=1-\frac{1}{2}\theta$, $\underline{\beta}=1-p, \overline{\beta}=p$ in eq. (A6), we get $-(p\times(2p-\theta))/((2p-1)(\theta-2))$. When $2(1-p) < \theta < 1$, eq. (A6) < 1. Therefore, we have proved that $\lim_{n\to\infty}P(Correct)_f < \lim_{n\to\infty}P(Correct)_s$. Because the probabilities of making a correct choice are non-decreasing in n, there must exists $\overline{N}$, when $n < \overline{N}$, $P(correct)_s < P(correct)_f$, and when $n > \overline{N}$, $P(correct)_s > P(correct)_f$. □

Mixed Network

In the mixed network, $t \sim U(1\text{-}a, a)$, where ½ < a ≤ 1. Parameter a captures the degree of heterogeneity in the network. The smaller a is, the more homogeneous the network is. The higher a is, the more heterogeneous the network is. When a is very close to ½, the network becomes very similar to the friend-network we examined before. When a equals to 1, the network becomes the stranger-network. The same as the basic modelling, p is the accuracy of private signals, where ½ < p < 1. First, let's look at how the first person makes a decision. If the first person sees a private signal H, then she infers that P(V=1|H)=p, and if the person sees L, then she infers that P(V=1|L)=1-p. Based on Lemma 1, given the information a person observes, I_n, this person will adopt when P(V=1| I_n) > 1- t_n and reject when P(V=1| I_n) < 1- t_n. Therefore, when observing H, the first person will adopt if p > 1 - t_n, otherwise she will reject. Similarly, when observing L, she will adopt when 1-p > 1- t_n, and otherwise she will reject.

Suppose the first person adopts. For the second person, because she is not sure the first person's preference, she can only infer the probability that the first person adopts when seeing an H, or equivalently, $P(p > 1\text{-}t_1) = \min(1, \frac{a-(1-p)}{2a-1})$ and the probability that the first person adopts when seeing an L, or equivalently, $P(1\text{-}p > 1\text{-}t_1) = \max(0, \frac{a-p}{2a-1})$. Combing the first person's action and her own signal, the second person calculates the probability that $P(V=1|I_2)$, then compares this probability with her preference and makes a decision.

In a more general term, we can write the probability that the n^{th} person adopts when observing information I_n as $P(\text{adopt}|I_n) = \max(0, \min(1, \frac{a-1+P(V=1|I_n)}{2a-1}))$ and $P(\text{reject}|\ I_n) = 1 - \max(0, \min(1, \frac{a-1+P(V=1|I_n)}{2a-1}))$. Here we rewrite I_n into two parts, the social information X_{n-1}, including all the previous actions made by the first n-1 people, and the private information s_n, the quality signal that the n^{th} person observes, either H or L. X_{n-1} have 2^{n-1} possible combinations. We use subscript k to indicate the k^{th} combination out of these 2^{n-1} possibilities: $X_{n-1,k}$.

Then the probability that the n^{th} person makes a correct quality inference is $P(correct)_n$ =

$$\sum_{k=1}^{2^{n-1}} \left\{ \begin{aligned} & P(V=1)P(X_{n-1,k}, s_n \mid V=1)\max(0,\min(1,\frac{a-1+P(V=1 \mid X_{n-1,k}, s_n)}{2a-1})) \\ & + P(V=0)P(X_{n-1,k}, s_n \mid V=0)[1-\max(0,\min(1,\frac{a-1+P(V=1 \mid X_{n-1,k}, s_n)}{2a-1}))] \end{aligned} \right\},$$

where $P(V=1|X_{n-1,k},s_n) = \dfrac{P(V=1)P(X_{n-1,k} \mid V=1)P(s_n \mid V=1)}{\sum_{j=0}^{1} P(V=j)P(X_{n-1,k} \mid V=j)P(s_n \mid V=j)}$, and $P(X_{n-1,k}|V=j)$ =

$$\begin{cases} P(X_{n-2,k} \mid V=j) \times \max(0,\min(1,\frac{a-1+P(V=1 \mid X_{n-2,k}, s_{n-1})}{2a-1})) \times P(s_{n-1} \mid V=j) & \text{if the } (n-1)^{th} \text{ adopts} \\ P(X_{n-2,k} \mid V=j) \times [1-\max(0,\min(1,\frac{a-1+P(V=1 \mid X_{n-2,k}, s_{n-1})}{2a-1}))] \times P(s_{n-1} \mid V=j) & \text{if the } (n-1)^{th} \text{ rejects} \end{cases}$$

where $X_{n-2,k}$ indicates the corresponding actions made by the first n-2 people given that the n^{th} person observes $X_{n-1,k}$. In other words, it is what the $(n-1)^{th}$ person observes from history actions. For example, if the third person observes that the first person adopts and the second person rejects, X_2={AR}, then for the 2^{nd} person, her corresponding social information is X_1={A}.

Since s_n could be either H or L, and X_{n-1} has 2^{n-1}elements, thus the n^{th} person could observe any of the 2^n possible information sets. As we mentioned in the proof of Proposition 2, it becomes extremely difficulty to analytically solve the probability that the n^{th} person makes a correct quality inference $P(correct)_n$. So we resort to numerical simulations. However, the run-time required scales with at 2^n - it doubles at each increment of number of agent. When n = 40, it takes about 20 days run-time and when n = 50 it will take about 65 years. So we provided a range of n (10, 20, 30 and 40), with different combinations of p (0.6, 0.7, 0.8 and 0.9) in Figure 4. The results show very similar patterns under different parameter values.

REFERENCES

Lobel, I. (2009), "Social Networks: Rational Learning and Information Aggregation", Ph.D. dissertation, Massachusetts Institute of Technology.