

Word of Mouth when Talk is Cheap: Information Diffusion and Optimal Targeting

Bartosz Redlicki*

June 2015

Abstract

This paper studies strategic information transmission in a population which consists of two types of consumers: enthusiasts and skeptics of a new product. A single producer targets one consumer, who is the only one to observe the quality of the product. Information about the quality is then diffused in a chain of communication by cheap talk. I show that truthful diffusion of information is possible in equilibrium under certain conditions, in particular if the enthusiasts' and the skeptics' preferences are sufficiently aligned. There are also multiple equilibria in which some babbling occurs. I apply the undefeatedness refinement to show which equilibria are plausible. Finally, I find that although in most cases it is optimal for the producer to target an enthusiast, a skeptic may be an optimal target if the preferences of enthusiasts and skeptics are sufficiently misaligned. The model fits well a common marketing strategy in which access to a new product is initially limited in order to stimulate word of mouth.

Keywords: strategic communication, cheap talk, word of mouth, information diffusion, optimal targeting.

JEL Codes: D82, D83, M31.

*Faculty of Economics, University of Cambridge. E-mail: bar43@cam.ac.uk.

1 Introduction

It is a common marketing strategy for companies to release a new product to a limited group of people. Probably the most famous example of such a limited release is the launch of Gmail on 1 April 2004. Access to Google’s new product was initially given to only a thousand people outside of the company and each of them was allowed to invite two friends. Interestingly, the main reason for the limited release was not marketing but technology; in fact, Google did not have enough server capacity that could handle millions of people joining the service. The strategy turned out to be hugely successful. According to Georges Harik, who was responsible for the majority of new products of Google at the time, “Everyone wanted it even more. It was hailed as one of the best marketing decisions in tech history, but it was a little bit unintentional.”¹ Google later used the same invitation-only marketing strategy when it launched Google+ in 2011, i.e. people could join the service only by receiving an invitation from the company or from another member.² The invitation-only strategy was also used by the social networking site Pinterest when it launched in 2010³ and by the music streaming site Spotify when it entered the United States market in 2011.⁴ Invitation-only strategies are not the only examples of limited release marketing. For instance, Pottermore, an interactive website aimed at fans of *Harry Potter* novels and movies, was initially made available only to a selected group of people who successfully completed a number of tasks online.⁵ Furthermore, movie distributors often use the so-called platform release of a movie, in which the movie is typically first shown in a small number of cinemas in selected cities and only later it is expanded to other cinemas in other areas.

What the above marketing strategies have in common is that access to a new product is initially limited in order to stimulate *word of mouth* (although, of course, the limited release strategy often also has other aims such as testing the functioning of the product). In his review of academic marketing research on “word of mouth”, Nyilasy (2006) points out that there is strong consensus what this phrase means. He supports using the definition proposed by Arndt (1967): “Oral, person-to person communication between a receiver and a communicator whom the receiver perceives as non-commercial, concerning a brand, a product or a service.” What is important about this definition is that it specifies that the content of word of mouth is commercial because it concerns brands, products, and services, but the communicators’ motivation is not commercial, or at least

¹<http://time.com/43263/gmail-10th-anniversary/>

²<http://edition.cnn.com/2011/TECH/web/06/30/google.plus.invites/>

³<http://www.bbc.co.uk/news/technology-19197531>

⁴http://www.nytimes.com/2011/07/18/technology/start-ups-try-to-beckon-users-by-invitation-only.html?_r=0

⁵<http://www.theguardian.com/books/2012/apr/14/pottermore-jk-rowling-harry-potter>

it is perceived not to be (Nyilasy 2006). In other words, communicators talk not because they are employed by the company but because they have their own desire to talk.

One of the reason why a limited release stimulates word of mouth so effectively is that it gives the selected people a feeling of being a member of an exclusive circle. As Campbell (2013) points out, this exclusivity may induce people to engage in word-of-mouth communication for two reasons. First, it may be because of their desire to make themselves appear knowledgeable to others. Second, it may be because people like to see themselves as knowledgeable, a motive often referred to as ego-enhancement (Nyilasy 2006) or self-enhancement (Baumeister 1998).

The desires to make oneself appear knowledgeable to others and see oneself as knowledgeable are particularly strong in the case of limited release marketing but are not the only motivations for engaging in word of mouth. People use word of mouth also in order to persuade others. Marketing research shows that that people report this motivation for word of mouth in a wide range of domains, including purchase decisions and health behaviours (Berger 2014). Another motivation to share information by word of mouth is altruism, i.e. the intention to help the receiver make a better purchase decision. Indeed, altruistic motives were often reported by people in the study by Sundaram et al. (1998). The main motivation for engaging in word of mouth as a receiver is information acquisition, especially if other sources of information are unavailable or the decision is risky, important, and complex (Berger 2014).

Another important aspect of word-of-mouth communication is the role of the social network. Empirical evidence by Johnson Brown and Reingen (1987) shows that if word of mouth occurs, it is more likely to occur in a strong tie (i.e. between friends, relatives, or close neighbours) than in a weak tie. The same study reveals that the more homophilous the tie (in terms of occupation, education, age, sex, etc.), the more likely it is to be used for communication by word of mouth.

This paper develops a model for analysing information diffusion by cheap talk in a heterogeneous population of consumers in a setting where only one consumer observes the state of the world at the beginning of the game. It can be interpreted as a stylised model of a marketing campaign which involves a limited release of a new product followed by spread of information by word of mouth. The model is consistent with the empirical observations about people's motivations to engage in word of mouth and the role of the social network.

In my model, the population consists of two types of consumers: enthusiasts and skeptics of a new product. The two types differ in their willingness to adopt the product, i.e. for a given state of the world, enthusiasts optimally take a higher action than skeptics. Apart from the consumers, there is a single producer who targets one of the consumers. The targeted consumer learns the quality of the product, which is given by a binary state of the world. Subsequently, he decides

whether to adopt the product, be neutral to it, or reject it, and sends a cheap-talk message about the quality to a single other consumer, with whom he is matched through an exogenous matching process. Information about the quality is then diffused in a chain of communication by cheap talk. Whether a particular consumer gets to talk to an enthusiast or a skeptic depends on the distribution of the types in the population and on the strength of assortativity of matching. Upon receiving a message, each consumer forms a posterior belief about the quality of the product and needs to make two decisions. First, the consumer decides whether to adopt the product, be neutral to it, or reject it. However, since the product is not released yet, all consumers except for the first one cannot verify its quality. The adoption decision is not observed by other consumers. Second, he needs to decide what cheap-talk message to send to another consumer. The message is costless and may or may not convey the sender's true beliefs about the quality of the product. Each consumer is not aware of the existence of the product unless he receives a message. One important feature of my model is an externality property: consumers' payoffs depend not only on their own adoption decisions but also on whether other consumers adopt the product. This externality property provides an incentive to communicate strategically.

This paper has two main aims: (1) identify how much information can be truthfully diffused by cheap talk in the population, and (2) identify an optimal target for the producer. As far as the second aim is concerned, I am particularly interested in whether it is ever optimal for the producer to target a skeptic rather than an enthusiast.

The paper has three main economic insights. First, there exists an equilibrium in which each consumer communicates truthfully regardless of whether they speak to a consumer of their own type or of the other type - I call it the *universally truthful equilibrium*. This equilibrium exists if and only if some conditions on parameters of the model are satisfied: the externality property is at a moderate level, the populations is not too dominated by either of the two types of consumers, the assortativity of matching is not too strong, and the probability of a breakdown of the communication chain is sufficiently low. As it turns out, these conditions can be satisfied only if the preferences of the two types of consumers are sufficiently aligned. More precisely, the preferences of a consumer of one type about the adoption decision of a consumer of the other type must be sufficiently close to the latter's preferences about his own adoption decision.

Second, there is a multiplicity of equilibria in this setting. Apart from the universally truthful equilibrium, there exist other equilibria such as the *homophilically truthful equilibrium* and the *universally babbling equilibrium*. In fact, these two equilibria exist for any parameter values. In the former, consumers communicate truthfully only to consumers of their own type but send uninformative messages (i.e. they "babble") to consumers of the other type. In the latter, consumers babble

regardless of whom they speak to. In order to resolve the problem of multiplicity, I extend the undefeated equilibrium refinement proposed by Mailath et al. (1993) and apply it to this setting. In the class of perfect Bayesian equilibria that I discuss in the paper, only the universally truthful equilibrium and the homophilically truthful equilibrium are undefeated.

Third, I answer the question whether it is ever optimal for the producer to target a skeptic rather than an enthusiast. I show that targeting a skeptic is never optimal in the universally truthful equilibrium but it is optimal in the homophilically truthful equilibrium under some conditions. If the population is dominated by enthusiasts who have a high prior belief about the quality of the product, the producer may want to target a skeptic in order to disinform the enthusiasts and thereby ensure that they do not change their positive attitude towards the product. On the other hand, if the population is dominated by skeptics who have a low prior belief about the quality of the product, the producer may want to target a skeptic in order to educate them about the quality of the product.

The paper is organised as follows. The rest of Section 1 reviews some related literature. Section 2 describes the setup of the model. Section 3 specifies the equilibrium concept and provides existence results of several equilibria. Section 4 introduces an extended version of the undefeated equilibrium refinement and applies it to the model, thus decreasing the number of equilibria. Section 5 studies the question of optimal targeting. Section 6 concludes.

1.1 Related literature

This paper is a natural extension of the cheap talk model by Crawford and Sobel (1982). Instead of a single sender and a single receiver, there is a communication chain of agents, each of whom is both a sender and a receiver, but only the first agent in the chain observes the underlying decision-relevant state of the world. Rather than using the uniform-quadratic framework, which is most commonly used in models of cheap talk, I develop a simple framework with a binary state of the world and two types of agents, each of whom can take one of three actions. This simplification allows me to study a tractable model while maintaining a heterogeneity of biases of agents.

My paper, therefore, naturally relates to the literature on cheap talk, which builds on the seminal contribution by Crawford and Sobel (1982). The research in this field has studied various extensions and applications of the cheap talk model, such as communication with multiple receivers (e.g., Farrell and Gibbons 1989, Goltsman and Pavlov 2011), multi-stage communication (e.g., Krishna and Morgan 2004, Golosov et al. 2014), communication in a network setting (e.g., Hagenbach and Koessler 2010, Galeotti et al. 2013), and others. The literature on cheap talk is too extensive to

attempt to summarise it here, so I refer the reader to surveys by Farrell and Rabin (1996), Krishna and Morgan (2008), and Sobel (2009). The paper has also elements in common with the literature on word of mouth (e.g., Campbell 2013, Campbell et al. 2013), dynastic communication (e.g., Anderlini and Lagunoff 2005, Lagunoff 2006, Anderlini et al. 2012), rumours (e.g., Banerjee 1993, Bloch et al. 2014), diffusion and social learning (e.g., Chatterjee and Dutta 2015) and optimal targeting (e.g., Galeotti and Goyal 2009).

Below I discuss papers which are most related to mine: Ambrus et al. (2013), Anderlini et al. (2012), Bloch et al. (2014), Campbell (2013), and Campbell et al. (2013).

Ambrus et al. (2013) study communication between a sender and a receiver via a chain of intermediators. Unlike in my paper, only the final receiver takes an action. The final receiver's action affects the payoffs of all players, which provides them an incentive to communicate strategically. This kind of setting often occurs in hierarchical organisations. Ambrus et al. (2013) investigate whether intermediated communication can increase information transmission. They conclude that intermediation can only decrease information transmission in pure-strategy perfect Bayesian Nash equilibria, but once mixed strategies are allowed for, there exist equilibria which improve information transmission.

Anderlini et al. (2012) investigate strategic information in an infinite sequence of agents. Their model is similar to mine in that each agent receives information from his predecessors, decides what (if any) information to pass on to future individuals by cheap talk, and takes a hidden action for himself. However, unlike in my paper, each agent observes some information about the state of the world on his own. The focus of the paper is on social learning. In one of their main results, Anderlini et al. (2012) show that if the preferences are not perfectly aligned, then there does not exist an equilibrium where full learning occurs, i.e. such that the agents' posterior beliefs place full weight on the true state of the world.

Bloch et al. (2014) study how rumours spread in a network. Like in my paper, one agent learns the true state of the world and spreads the word to other agents. However, the agent is randomly selected rather than targeted by a principal. In addition, the communication protocol in their model is not cheap talk. Instead, the agents can only choose whether to transmit the message or not, but they cannot transform it. The network consists of unbiased and biased agents and its structure is common knowledge among all agents. The unbiased agents want the collective decision to match the true state of the world while the biased ones prefer a particular decision regardless of the state of the world. The main results of the paper are that agents block messages which come from parts of the network that contain many biased agents, and that biased agents might be better off by decreasing their number.

Campbell (2013) investigates, like me, how information about a product spreads by word of mouth but the setup is very different. The diffusion by word of mouth is modelled as a percolation process on a random graph. Agents do not communicate strategically; instead, they are assumed to send messages about the existence of the product with a probability which depends on the individual’s valuation of the product and its price. The focus of the paper is on the impact of word of mouth on the elasticity of demand for the product, comparative statics of pricing with respect to the characteristics of the network, and optimal targeting for advertising purposes. However, unlike in my paper, optimal targeting is analysed in the context of the position in the network.

Finally, my paper shares some features with Campbell et al. (2013). Like mine, their paper studies a firm’s marketing strategy of restricting access to information about its product in order to stimulate word of mouth. However, they focus primarily on the self-enhancement motive of consumers: each consumer’s utility is an increasing function of others’ belief that he is a high (i.e. more knowledgeable) type. Furthermore, consumers share information by sending costly and verifiable messages rather than by cheap talk. Campbell et al. (2013) conclude that the firm optimally chooses to restrict the access to information about the product to low-type consumers in order to spur word of mouth. It is also beneficial for the firm to commit not to use advertising because it crowds out word-of-mouth communication between consumers.

2 Model

Players. There is a single producer, P , and an infinite set of consumers, N . The set of consumers is referred to as population. The producer releases a new product, the existence of which the consumers are initially not aware of. The population consists of two types of consumers: *enthusiasts* who are more willing to adopt the new product, and *skeptics* who are less willing to do so. Formally, the type of consumer is denoted by $\tau \in T = \{H, L\}$, where H-types are enthusiasts and L-types are skeptics. The proportion of enthusiasts in the population is common knowledge and is denoted by ρ .

State of the world. The state of the world is binary, $s \in S = \{s_1, s_2\}$, and unknown both to the producer and to the consumers. The state of the world can be interpreted as the consumers’ perception of the quality of the new product (s_1 is good quality and s_2 is bad quality). The producer and the consumers have a common prior belief that $s = s_1$ with probability π . This common prior is common knowledge.

Timeline. The timeline of the game is as follows.

- In period $t = 1$, the producer targets one of the consumers, denoted by c_1 , who gets to observe the true state of the world, s , and takes an action, α_1 .
- In periods $t \in \{2, 3, \dots\}$, consumer c_{t-1} is matched with another consumer, c_t , from the population and communicates the state of the world to him by sending a cheap-talk message, m_{t-1} . Following the communication, consumer c_t forms a posterior belief about the state of the world, β_t , and takes an action, α_t . Consumer c_t does not observe the true state of the world before sending his message nor before taking his action.

The process continues until the communication chain is exogenously broken, which happens with probability $\epsilon \in [0, 1]$ in each period t . Therefore, consumers form a chain of communication $C = \{c_1, c_2, c_3, \dots\}$, whose length depends on whether and when the communication breaks down. Apart from his own type, every consumer c_t observes the types of consumers c_{t-1} and c_{t+1} , i.e. the types of consumers with whom he communicates. However, consumer c_t does not observe any messages other than m_{t-1} and m_t , nor any actions other than α_t . Consumer c_t also does not observe his position in the communication chain, t .

Communication takes the form of cheap talk, i.e. messages are costless and convey information which is not verifiable. Consumer c_t 's beliefs about the state of the world s are given by $\beta_t(s)$.

The matching mechanism works as follows: with probability κ , the matching of consumers is perfectly assortative, and with probability $1 - \kappa$, the matching is random and depends solely on the distribution of types in the population.⁶ Hence, an enthusiast c_t is matched with another enthusiast c_{t+1} with probability $\kappa + (1 - \kappa)\rho$, and is matched with a skeptic c_{t+1} with probability $(1 - \kappa)(1 - \rho)$. Conversely, a skeptic c_t is matched with another skeptic c_{t+1} with probability $\kappa + (1 - \kappa)(1 - \rho)$ and with an enthusiast c_{t+1} with probability $(1 - \kappa)\rho$. Parameter κ thus serves as a measure of assortativity of the matching mechanism.⁷

Strategies. The producer needs to make only one decision: choose which consumer to target. I denote her *targeting strategy* by $\varphi \in N$. Later, I focus on equilibria such that all enthusiasts have identical strategies and all skeptics have identical strategies, which means that the producer is effectively choosing between targeting any enthusiast and any skeptic. Hence, I can then denote the targeting strategy by $\varphi \in T = \{H, L\}$.

⁶Therefore, the choice of the receiver is non-strategic. This assumption is consistent with the definition of word of mouth by Arndt (1967), which implies that consumers' motivation to engage in word of mouth is not commercial.

⁷The possibility of assortative matching is a reflection of empirical evidence which shows that the more homophilous the tie, the more likely it is to be used for communication by word of mouth (Johnson Brown and Reingen 1987).

Each consumer c_t needs to make two decisions: choose what action α_t to take and choose what message m_t to send to consumer c_{t+1} . Therefore, each consumer has an *action strategy* and a *communication strategy*. The action strategy for consumer c_1 is a function $\sigma_1 : S \rightarrow A_1$, and for all other consumers it is a function $\sigma_t : M_{t-1} \times T_{t-1} \rightarrow A_t$. The communication strategy for the targeted consumer is a function $\mu_1 : S \times T_2 \rightarrow M_1$, and for all other consumers it is a function $\mu_t : M_{t-1} \times T_{t-1} \times T_{t+1} \rightarrow M_t$. An *action strategy profile* $\sigma = (\sigma_i)_{i \in N}$ specifies action strategies for all consumers. A *communication strategy profile* $\mu = (\mu_i)_{i \in N}$ specifies communication strategies for all consumers.

Let us now define the action set, A_i , and the message set, M_i . Each of the two types of consumers can take one of three actions: *adopt* (adopting the new product), *neutral* (being neutral towards the new product), and *reject* (rejecting the new product). Therefore, we have $A_t = \{\text{adopt}, \text{neutral}, \text{reject}\} \forall c_t$. I assume that each consumer can send one of three messages: “ s_1 ” (i.e. “the product is good”), “ s_2 ” (i.e. “the product is bad”), and “ \emptyset ” (i.e. “I don’t know the quality of the product”). Hence, the message set is given by $M_t = \{”s_1”, ”s_2”, ”\emptyset”\} \forall c_t$.⁸ I assume that a consumer does not have an option of not sending a message or he would never choose it even if he had such an option.⁹

Consumers’ payoffs. I assume that each consumer receives a payoff not only from his own action but also from the actions of the consumer that he talks to and all consumers who take an action later. That is, consumer c_t ’s payoff is a function of all the actions α_{t+i} , with $i = 0, 1, 2, \dots$. In principle, each consumer also receives a payoff from the actions of the preceding consumers but it is assumed that he cannot affect their actions because they have already made their decisions. Therefore, they do not enter the consumer’s payoff function. I assume that the consumers’ payoffs are such that if the state were known to be s_1 (i.e. good quality), then an enthusiast would optimally adopt the product and a skeptic would optimally be neutral towards it. On the other hand, if the state were known to be s_2 (i.e. bad quality), an enthusiast would be neutral and a skeptic would reject. Regardless of the state of the world, an enthusiast would never reject the product and a skeptic would never adopt the product. Consumers have the same preferences regarding their own actions and the actions of others.

⁸In cheap-talk games, messages have no literal meaning; instead, their meaning is established by use in equilibrium. However, for greater clarity of the exposition of the model, I pretend that the messages have these literal meanings.

⁹This is a reflection of the consumer’s desire to make himself appear knowledgeable to others or his desire to see himself as knowledgeable, i.e. the “self-enhancement” motive (Baumeister 1998). These desires could be easily incorporated in the model by assuming that a consumer obtains a fixed positive payoff from sending any message, which would outweigh any potential negative consequences of sending the message.

The payoffs which reflect the preferences described in the previous paragraph are shown in Table 1 and Table 2.

$\alpha_{t+i} \setminus s$	s_1	s_2
<i>adopt</i>	x_1	y_1
<i>neutral</i>	x_2	y_2
<i>reject</i>	0	y_3

Table 1: H-type consumer c_t 's payoffs from α_{t+i} , where $i = 0, 1, 2, \dots$

$\alpha_{t+i} \setminus s$	s_1	s_2
<i>adopt</i>	y_3	0
<i>neutral</i>	y_2	x_2
<i>reject</i>	y_1	x_1

Table 2: L-type consumer c_t 's payoffs from α_{t+i} , where $i = 0, 1, 2, \dots$

As we shall see later in the model, it may happen that a consumer does not get to know the true state of the world and has to choose his action based on his prior belief, π . We call this action his *pooling* action. An enthusiast's pooling action is:

$$pool_H = \begin{cases} \textit{adopt} & \text{if } \pi > \frac{y_2 - y_1}{x_1 - x_2 + y_2 - y_1} \\ \textit{neutral} & \text{otherwise} \end{cases} . \quad (2.1)$$

A skeptic's pooling action is:

$$pool_L = \begin{cases} \textit{neutral} & \text{if } \pi > \frac{x_1 - x_2}{x_1 - x_2 + y_2 - y_1} \\ \textit{reject} & \text{otherwise} \end{cases} . \quad (2.2)$$

A consumer's payoffs from the actions of other consumers are discounted. The discount factor is $\delta(1 - \epsilon)$, where $\delta \in [0, 1]$ and $\epsilon \in [0, 1]$. I interpret δ as an *externality parameter*, which describes the strength of externality from one consumer's action to another consumer. Furthermore, the discount factor accounts for the possibility that the communication chain is broken with probability ϵ in any period t .

I can now specify the total payoff function of consumer c_t . Let us denote consumer c_t 's payoff from consumer $c_{t'}$'s action by $u_t(\alpha_{t'})$. Formally, the payoff to consumer c_t is:

$$u_t(\alpha_t) + \sum_{t'=t+1}^{\infty} (\delta(1-\epsilon))^{t'-t} u_t(\alpha_{t'}). \quad (2.3)$$

Let us denote consumer c_t 's expected payoff by $U_t(\sigma, \mu, \beta)$.

In the paragraphs below, I discuss our assumptions on the parameters in Tables 1 and 2.

Assumption 1. I assume that $x_1, x_2, y_1, y_2, y_3 > 0$.

Assumption 1 simply states that when a product is of good quality (i.e. $s = s_1$), the worst outcome for an enthusiast is if he or another consumer rejects the product. Conversely, when the product's quality is bad (i.e. $s = s_2$), the worst outcome for a skeptic is if he or another consumer adopts this product .

Assumption 2. I assume that $x_1 > x_2 > 0$.

Assumption 2 states that when the product is good, an enthusiast most prefers the product to be adopted, which is preferred to neutrality and rejection. Conversely, when the product is bad, a skeptic most preferred action is a rejection of the product, which is seen as better than neutrality or adoption.

Assumption 3. I assume that $y_2 > y_1$ and $y_2 > y_3$ and $y_1 \geq y_3$.

Assumption 3 describes an enthusiast's preferences when the product is bad and a skeptic's preferences when the product is good. It says that an enthusiast's most preferred action towards a bad product is neutrality and a skeptic's most preferred action towards a good product is also neutrality. Note that I do not specify which of y_1 and y_3 is greater. That is, I do not make any assumption whether an enthusiast would prefer adoption or rejection of a bad product; all I say is that he most prefers neutrality towards it. Similarly, I do not make any assumption whether a skeptic prefers adoption or rejection of a good product.

Producer's payoffs. Finally, let us turn to the producer's payoffs. I denote the producer's payoff from any consumer c_t 's action by $u_P(\alpha_t)$. I assume that the producer receives a payoff of a from the action *adopt*, b from *neutral*, and 0 from *reject*, where $a > b > 0$. In other words, the producer values adoption the most but a consumer's neutrality is better than rejection. The

producer's discount factor is $(1 - \epsilon)$ and reflects the possibility that the communication chain may be broken in any period t with probability $\epsilon \in [0, 1]$. Formally, the producer's payoff is:

$$\sum_{t=1}^{\infty} (1 - \epsilon)^{t-1} u_P(\alpha_t). \quad (2.4)$$

Let us denote the producer's expected payoff by $U_P(\sigma, \mu, \beta, \varphi)$.

3 Equilibrium

3.1 Equilibrium concept

For now, I can exclude the producer's targeting strategy from the discussion because it does not affect the consumers' actions and messages. In Sections 3 and 4, I focus on the equilibrium of the game of information diffusion and postpone the discussion of optimal targeting until Section 5. Since the formation of the communication chain is exogenous to the strategies of consumers, I describe the set of players of the information diffusion game by the chain $C = \{c_1, c_2, c_3, \dots\}$ rather than by the set N . Therefore, the information diffusion game is defined by consumers $C = \{c_1, c_2, c_3, \dots\}$, their action and communication strategies, and their payoff functions. It is a sequential game of incomplete information, so a natural solution concept is a perfect Bayesian equilibrium (PBE) in pure strategies.¹⁰

The triple (σ, μ, β) describes the action and communication strategy profiles of all consumers, and their beliefs about the state of the world. Consumer c_1 's individual expected payoff upon observing state s is expressed as $U_1(\sigma, \mu, \beta \mid s)$. Consumer c_t 's (where $t > 1$) individual expected payoff upon receiving message m_{t-1} is expressed as $U_t(\sigma, \mu, \beta \mid m_{t-1})$. I define a perfect Bayesian equilibrium $(\sigma^*, \mu^*, \beta^*)$ in the standard way: for $t = 1$, for any σ_1 and μ_1 , and for any s and τ_{t+1} , $U_1(\sigma^*, \mu^*, \beta^* \mid s) \geq U_1(\sigma_{-1}^*, \sigma_1, \mu^*, \beta^* \mid s)$ and $U_1(\sigma^*, \mu^*, \beta^* \mid s, \tau_{t+1}) \geq U_1(\sigma^*, \mu_{-1}^*, \mu_1, \beta^* \mid s, \tau_{t+1})$, and for each $t > 1$, for any σ_t and μ_t , and for any m_{t-1} , τ_{t-1} , and τ_{t+1} , $U_t(\sigma^*, \mu^*, \beta^* \mid m_{t-1}, \tau_{t-1}) \geq U_t(\sigma_{-t}^*, \sigma_t, \mu^*, \beta^* \mid m_{t-1}, \tau_{t-1})$ and $U_t(\sigma^*, \mu^*, \beta^* \mid m_{t-1}, \tau_{t-1}, \tau_{t+1}) \geq U_t(\sigma^*, \mu_{-t}^*, \mu_t, \beta^* \mid m_{t-1}, \tau_{t-1}, \tau_{t+1})$. In the PBE, beliefs of all players, β^* , are consistent with the equilibrium play and are updated according to Bayes' rule wherever possible.

¹⁰Mixed-strategy equilibria in which the sender randomises between messages with the same equilibrium meaning are not plausible if we postulate that players prefer messages that are short, simple, and straightforward. Farrell (1993) argues that: "For example, if type t wants (and is expected) to reveal himself, and if both the English sentences, 'I am t ,' and 'I am either u or v ,' are interpreted in equilibrium as meaning 'I am t ,' then S [the sender] will prefer the former. This suggests that it is hard to sustain mixed-strategy equilibria in which S randomizes over many messages with the same equilibrium meaning".

Like in any other game of cheap talk, we can expect multiple equilibria to arise. For simplicity and to keep the discussion brief, throughout the paper, I make a restriction to a class of equilibria which satisfies Assumption 4. Before I state the assumption, let me denote the beliefs of consumer c_t about the state of the world by $\omega_t \in \Omega = \{s_1, s_2, \emptyset\}$, where $\omega_t = s_1$ is equivalent to $\beta_t(s_1) = 1$, $\omega_t = s_2$ is equivalent to $\beta_t(s_1) = 0$, and $\omega_t = \emptyset$ is equivalent to $\beta_t(s_1) \in (0, 1)$. Then, I can write $\sigma_t : \Omega_t \times T_{t-1} \rightarrow A_t$ for consumer c_t 's action strategy (for any $t \geq 1$) and $\mu_t : \Omega_t \times T_{t+1} \rightarrow M_t$ for consumer c_t 's communication strategy (for any $t \geq 1$).

Assumption 4. For a given type of consumer c_{t+1} , τ_{t+1} , any consumer c_t 's communication strategy in equilibrium $e \equiv (\sigma, \mu, \beta)$ is either

$$(a) \mu_t(\omega_t, \tau_{t+1}) = \begin{cases} "s_1" & \text{if } \omega_t = s_1 \\ "s_2" & \text{if } \omega_t = s_2, \text{ or} \\ "\emptyset" & \text{if } \omega_t = \emptyset \end{cases}$$

$$(b) \mu_t(\omega_t, \tau_{t+1}) = m_t^b.$$

By making Assumption 4, I only consider equilibria in which each consumer's communication strategy towards a consumer of type τ_{t+1} takes one of two extreme forms: either he completely truthfully reveals his beliefs about the state of the world or his communication is completely uninformative. Note that under this restriction, each consumer will only be able to hold one of three beliefs: $\beta_t(s_1) = 1$, $\beta_t(s_1) = 0$, and $\beta_t(s_1) = \pi$ (prior beliefs), so the belief denoted by $\omega_t = \emptyset$ is equivalent to $\beta_t(s_1) = \pi$. The restriction to only these two extreme cases reduces the number of possible equilibria and makes the analysis easier and clearer. One possible interpretation of it is that it is a reflection of a limited cognitive or coordination ability of people, which results in them not being able to coordinate on an equilibrium that contains a less extreme protocol of communication. Intuitively, in any equilibrium, each person either completely trusts what he hears from another person or he completely ignores it.

In the rest of this section, I present several equilibria (in the class of equilibria satisfying Assumption 4) with the property that any two consumers c_t and $c_{t'}$ of the same type (i.e. $\tau_t = \tau_{t'}$) have the same communication strategy, i.e. $\mu_t(\omega_t, \tau_{t+1}) = \mu_{t'}(\omega_{t'}, \tau_{t'+1}) \forall c_t, c_{t'}$ such that $\tau_t = \tau_{t'}$. This property simply states that all consumers of type H have the same communication strategy and all consumers of type L have the same communication strategy. That is, for given beliefs about the state of the world and a given type of the receiver of their message, two consumers of the same type would send the same message.

3.2 Universally truthful equilibrium

Consumer c_t is said to communicate *universally truthfully* if his communication strategy reveals to consumer c_{t+1} the true beliefs that he holds upon observing the state of the world s (for consumer c_1) or upon receiving a message m_{t-1} from his predecessor (for consumers c_t where $t > 1$) regardless of c_{t+1} 's type, τ_{t+1} . A triple $(\sigma^*, \mu^*, \beta^*)$ constitutes a *universally truthful equilibrium* (UTE) if the conditions of a perfect Bayesian equilibrium are satisfied and all consumers communicate universally truthfully.

Definition 1. A triple $(\sigma^*, \mu^*, \beta^*)$ is a universally truthful equilibrium (UTE) if it is a perfect Bayesian equilibrium of the game and the equilibrium communication strategy profile μ^* is: $\mu_1^*(s_i, \tau_2) = "s_i"$ for $i = 1, 2$, and $\mu_t^*(m_{t-1}, \tau_{t-1}, \tau_{t+1}) = m_{t-1}$ for $t > 1$.

I am interested in analysing whether there exists a UTE in the information diffusion game. If it does exist, then I would like to know under what conditions, i.e. for what values of parameters δ (strength of externality), ρ (proportion of consumers of type H in the population), κ (assortativity of matching), and ϵ (probability of breakdown of communication).

Proposition 1. *A universally truthful equilibrium exists if and only if the following conditions are satisfied:*

- (a) $\delta \in [\delta^*, \delta^{**}]$ for $\rho < \frac{1}{2}$ and $\delta \in [\delta^\dagger, \delta^{\dagger\dagger}]$ for $\rho \geq \frac{1}{2}$;
- (b) $\rho \in [\rho^*, \rho^{**}]$;
- (c) $\kappa \in [0, \kappa^*]$ for $\rho < \frac{1}{2}$ and $\kappa \in [0, \kappa^\dagger]$ for $\rho \geq \frac{1}{2}$;
- (d) $\epsilon \in [0, \epsilon^*]$.

The proof is provided in the Appendix.

In the UTE, the equilibrium beliefs of each consumer c_t ($t > 1$) are naturally $\beta_t(s_1 | m_{t-1} = "s_1") = 1$ and $\beta_t(s_1 | m_{t-1} = "s_2") = 0$. An H-type consumer c_t 's equilibrium action strategy is $\sigma_t^*(m_{t-1}, \tau_{t-1}) = \text{adopt}$ if $m_{t-1} = "s_1"$ and $\sigma_t^*(m_{t-1}, \tau_{t-1}) = \text{neutral}$ if $m_{t-1} = "s_2"$, whereas an L-type consumer c_t 's equilibrium action strategy is $\sigma_t^*(m_{t-1}, \tau_{t-1}) = \text{neutral}$ if $m_{t-1} = "s_1"$ and $\sigma_t^*(m_{t-1}, \tau_{t-1}) = \text{reject}$ if $m_{t-1} = "s_2"$.¹¹

¹¹Like in any cheap-talk game, messages have no literal meaning in this game but their meaning is established by use in equilibrium. Therefore, there exist multiple (outcome-equivalent) UTEs in which different permutations of messages are used to convey the meanings established in this equilibrium. For instance, there exists a UTE in which the literal meaning of messages is reversed in equilibrium: the equilibrium meaning of a message "s₁" is "the state of the world is s₂" and the equilibrium meaning of a message "s₂" is "the state of the world is s₁". If we postulate that players prefer to use messages that are straightforward, then the most reasonable UTE is the one in which the meaning of messages in equilibrium corresponds to their literal meaning.

Since consumer c_1 truthfully communicates s to consumer c_2 and all subsequent consumers communicate messages which they have received, a message " \emptyset " is never sent in the UTE. In a PBE, out-of-equilibrium beliefs of consumer c_t upon receiving message $m_{t-1} = "\emptyset"$, denoted by $\beta_t(s_1 \mid m_{t-1} = "\emptyset")$, are not restricted by Bayes' rule. Hence, these out-of-equilibrium beliefs can take any form as long as the consumers' best response given these beliefs is to communicate fully truthfully. For instance, the UTE is sustained by out-of-equilibrium beliefs of the following form: for each c_t (where $t > 1$), either $\beta_t(s_1 \mid m_{t-1} = "\emptyset") = 0$ or $\beta_t(s_1 \mid m_{t-1} = "\emptyset") = 1$. This specification of out-of-equilibrium beliefs sustains the UTE because it means that sending a message " \emptyset " is equivalent to sending one of the two messages which are used on the equilibrium path, " s_1 " and " s_2 ".

Let me provide some intuition for the conditions in Proposition 1.

First, for the UTE to exist, the strength of externality of the actions of consumers further down the communication chain needs to be at a moderate level. If parameter δ were too low, a consumer c_t of type τ_t would be able to profitably deviate from the equilibrium by sending a message $m_t = "s_i"$ to a consumer c_{t+1} of type $\tau_{t+1} \neq \tau_t$ when in fact $m_{t-1} = "s_j"$, where $i \neq j$ and $i = 1$ if $\tau_t = H$ and $i = 2$ if $\tau_t = L$. If it were too high, then a consumer c_t of type τ_t would be able to profitably deviate from the equilibrium by sending a message $m_t = "s_i"$ to a consumer c_{t+1} of type $\tau_{t+1} = \tau_t$ when in fact $m_{t-1} = "s_j"$, where $i \neq j$ and $i = 1$ if $\tau_t = H$ and $i = 2$ if $\tau_t = L$. In other words, the strength of externality cannot be too low because then each consumer would have an incentive to lie to a consumer of the other type, and it cannot be too high because then each consumer would have an incentive to lie to a consumer of their own type (in each case, an H -type consumer's profitable lie would be to send " s_1 " upon receiving " s_2 ", and an L -type consumer's profitable lie would be to send " s_2 " upon receiving " s_1 ").

Second, the proportions of types in the population of consumers need to be sufficiently close to each other, i.e. the population cannot be dominated by any of the two types. If there were too many L -type consumers in the population, then an H -type consumer would be able to profitably deviate from the equilibrium by lying to an L -type consumer (he would send " s_1 " upon receiving " s_2 " and thus induce actions *neutral* from all subsequent L -types). Conversely, if there were too many H -type consumers in the population, an L -type consumer would be able to profitably deviate by lying to an H -type consumer (he would be send " s_2 " upon receiving " s_1 " and thus induce actions *neutral* from all subsequent H -types).

Third, the assortativity of the matching mechanism needs to be sufficiently weak. If κ were too high, then each consumer would be able to profitably deviate by lying to a consumer of the other type. The reason is that, under strong assortativity of matching, a consumer who communicates

with a consumer of the other type would realise that consumers further down the chain will more likely be of the other type too. Hence, an H -type consumer would be able to profitably deviate by sending " s_1 " upon receiving " s_2 " and thus inducing actions *neutral* from all subsequent L -types. An L -type consumer would be able to profitably deviate by sending " s_2 " upon receiving " s_1 " and thus inducing actions *neutral* from all subsequent H -types.

Fourth, the probability of a breakdown in communication needs to be sufficiently small. A consumer has an incentive to reveal his true beliefs to a consumer of the other type if and only if there is a sufficiently high probability that a consumer of his own type will later learn these beliefs.

Corollaries 1 and 2 provide more insight into the conditions under which the UTE exists.

Corollary 1. *A universally truthful equilibrium exists only if*

$$(a) \frac{y_2 - y_1}{y_2 - y_3} > \frac{1 - \rho}{\rho} \text{ for } \rho < \frac{1}{2},$$

$$(b) \frac{y_2 - y_1}{y_2 - y_3} > \frac{\rho}{1 - \rho} \text{ for } \rho \geq \frac{1}{2}.$$

Corollary 1 describes the relationship between parameters y_1 , y_2 , y_3 , and ρ . This relationship is a necessary condition for the existence of the UTE. The meaning of this result is as follows: the closer ρ is to either 0 or 1, the lower y_1 needs to be compared to y_3 . In other words, as the population of consumers becomes more dominated by any of the two types, the higher payoff an H -type consumer must obtain from *reject* relative to *adopt* when $s = s_2$, and the higher payoff an L -type consumer must obtain from *adopt* relative to *reject* when $s = s_1$.

Corollary 2. *A universally truthful equilibrium exists only if $y_3 > y_1$.*

Corollary 2 gives a more general result than Corollary 1. It shows that in the UTE it must be that y_3 is greater than y_1 . This result means that an H -type consumer must obtain a higher payoff from the action *reject* than from the action *adopt* when $s = s_2$ and an L -type consumer must obtain a higher payoff from the action *adopt* than from the action *reject* when $s = s_1$; otherwise, the UTE does not exist. Intuitively, an enthusiast must prefer a bad product to be rejected rather than adopted (however, his most preferred action is neutrality). At the same time, a skeptic must prefer a good product to be adopted rather than rejected (again, his most preferred action is neutrality).

The underlying message of Corollary 2 is that for the UTE to exist, preferences of an enthusiast and a skeptic cannot be too far from each other. Given the setup of our model, a skeptic would like to reject a bad product while an enthusiast would like a skeptic to be neutral towards a bad product. This means that the preferences about a skeptic's action are not fully aligned. A necessary condition for the existence of the UTE is that these preferences are not too misaligned: an enthusiast cannot rank a skeptic's rejection of a bad product too low. That is, in the eyes of an enthusiast, a rejection

of a bad product must be better than an adoption of a bad product. Conversely, in the eyes of a skeptic, an adoption of a good product must be preferred to a rejection of a good product.

3.3 Homophilically truthful equilibrium

In this subsection I look at another possible perfect Bayesian equilibrium of the game, in which a consumer communicates truthfully if only if he talks to a consumer of his own type.

Formally, consumer c_t is said to communicate *homophilically truthfully* if his communication strategy reveals the true beliefs that he holds upon observing the state of the world s or upon receiving a message from his predecessor if and only if the type of consumer c_{t+1} is the same as his, i.e. $\tau_{t+1} = \tau_t$. If $\tau_{t+1} \neq \tau_t$, then consumer c_t 's communication strategy is uninformative, i.e. he sends the same “babbling” message $m_t^b \in \{”s_1”, ”s_2”, ”\emptyset”\}$ regardless of what he really observes. A triple $(\sigma^*, \mu^*, \beta^*)$ constitutes a *homophilically truthful equilibrium* (HTE) if the conditions of a perfect Bayesian equilibrium are satisfied and all consumers communicate homophilically truthfully.

Definition 2. A triple $(\sigma^*, \mu^*, \beta^*)$ is a homophilically truthful equilibrium (HTE) if it is a perfect Bayesian equilibrium of the game and the equilibrium communication strategy profile μ^* is:

$$\mu_1^*(s_i, \tau_2) = \begin{cases} ”s_i” & \text{for } i = 1, 2 \text{ if } \tau_2 = \tau_1; \\ m_1^b & \text{for } i = 1, 2 \text{ if } \tau_2 \neq \tau_1 \end{cases};$$

$$\mu_t^*(m_{t-1}, \tau_{t-1}, \tau_{t+1}) = \begin{cases} m_{t-1} & \text{if } \tau_{t-1} = \tau_t = \tau_{t+1} \\ ”\emptyset” & \text{if } \tau_{t-1} \neq \tau_t = \tau_{t+1} \\ m_t^b & \text{if } \tau_t \neq \tau_{t+1} \end{cases} \quad \text{for } t > 1.$$

Inuitively, in this equilibrium, a consumer reveals what he knows about the state of the world if he talks to a consumer of his own type (“ \emptyset ” means “I don’t know the state of the world” or, more precisely, “I hold my prior beliefs about the state of the world”), but babbles if he talks to a consumer of the other type. It turns out that the HTE always exists in this model setup.

Proposition 2. A homophilically truthful equilibrium exists for any values of parameters $\delta \in [0, 1]$, $\rho \in [0, 1]$, $\kappa \in [0, 1]$, and $\epsilon \in [0, 1]$.

In the HTE, the equilibrium beliefs of each consumer c_t ($t > 1$) are $\beta_t^*(s_1 \mid \tau_{t-1} = \tau_t, m_{t-1} = ”s_1”) = 1$, $\beta_t^*(s_1 \mid \tau_{t-1} = \tau_t, m_{t-1} = ”s_2”) = 0$ and $\beta_t^*(s_1 \mid \tau_{t-1}, m_{t-1}) = \pi$ if $\tau_{t-1} \neq \tau_t$ or $m_{t-1} = ”\emptyset”$. An H-type consumer c_t 's equilibrium action strategy is $\sigma_t^*(m_{t-1}, \tau_{t-1}) = \text{adopt}$

if $\tau_{t-1} = \tau_t$ and $m_{t-1} = "s_1"$, $\sigma_t^*(m_{t-1}, \tau_{t-1}) = neutral$ if $\tau_{t-1} = \tau_t$ and $m_{t-1} = "s_2"$, and $\sigma_t^*(m_{t-1}, \tau_{t-1}) = pool_H$ if $\tau_{t-1} \neq \tau_t$ or $m_{t-1} = "\emptyset"$. An L-type consumer c_t 's equilibrium action strategy is $\sigma_t^*(m_{t-1}, \tau_{t-1}) = neutral$ if $\tau_{t-1} = \tau_t$ and $m_{t-1} = "s_1"$ and $\sigma_t^*(m_{t-1}, \tau_{t-1}) = reject$ if $\tau_{t-1} = \tau_t$ and $m_{t-1} = "s_2"$, and $\sigma_t^*(m_{t-1}, \tau_{t-1}) = pool_L$ if $\tau_{t-1} \neq \tau_t$ or $m_{t-1} = "\emptyset"$.¹²

As far as out-of-equilibrium beliefs are concerned, note that in a communication between consumers c_{t-1} and c_t where $\tau_{t-1} = \tau_t$ (for all t) all messages in M are used in equilibrium, so there is no need to specify out-of-equilibrium beliefs for this case. However, in a communication between c_{t-1} and c_t where $\tau_{t-1} \neq \tau_t$, only the "babbling" message m_{t-1}^b is used in equilibrium. The HTE is sustained if each consumer c_t (where $\tau_{t-1} \neq \tau_t$) holds the following out-of-equilibrium beliefs: for $\tau_t = L$, $\beta_t(s_1 | \tau_{t-1}, m_{t-1}) = \pi$ for all $m_{t-1} \neq m_{t-1}^b$ or $\beta_t(s_1 | \tau_{t-1}, m_{t-1}) = 0$ for all $m_{t-1} \neq m_{t-1}^b$, and for $\tau_t = H$, $\beta_t(s_1 | \tau_{t-1}, m_{t-1}) = \pi$ for all $m_{t-1} \neq m_{t-1}^b$ or $\beta_t(s_1 | \tau_{t-1}, m_{t-1}) = 1$ for all $m_{t-1} \neq m_{t-1}^b$. In other words, if an L-type consumer c_t takes his pooling action or action *reject* upon receiving an out-of-equilibrium message, then the H-type consumer c_{t-1} is weakly worse off by deviating from the equilibrium message to the out-of-equilibrium message. Conversely, if an H-type consumer c_t takes his pooling action or action *reject* upon receiving an out-of-equilibrium message, then the L-type consumer c_{t-1} is weakly worse off by deviating from the equilibrium message to the out-of-equilibrium message.

The intuition behind the result in Proposition 2 is simple. A consumer c_{t-1} has an incentive to reveal his true beliefs about the state of the world to a consumer c_t of type τ_{t-1} because their preferences are perfectly aligned and, given the equilibrium strategies, any message which c_{t-1} sends to c_t will not affect the action taken by any consumer c_{t+i} (where $i > 0$) of type $\tau_{t+i} \neq \tau_{t-1}$. In turn, consumer c_t is better off by believing the message from c_{t-1} . Babbling is a part of this equilibrium for the usual reason: if a receiving consumer ignores the sender's message, then the sender cannot do better than send an uninformative message, and if the sender sends an uninformative message, then the receiver's best response is to ignore the message.

3.4 Universally babbling equilibrium

The universally truthful equilibrium and the homophilically truthful equilibrium are not the only equilibria of the game that satisfy Assumption 4 and the property that any two consumers c_t and $c_{t'}$ of the same type (i.e. $\tau_t = \tau_{t'}$) have the same communication strategy.

Consider for instance an equilibrium in which each consumer's strategy is to send an uninfor-

¹²Again, since the messages have no literal meaning in this game, there exist multiple HTEs in which different permutations of messages are used to convey the meanings established in this equilibrium.

mative message regardless of which type of consumer he talks to. Consumer c_t 's communication strategy is *universally babbling* if he sends the same “babbling” message $m_t^b \in \{”s_1”, ”s_2”, ”\emptyset”\}$ regardless of what message he has received and regardless of whether $\tau_t = \tau_{t+1}$ or $\tau_t \neq \tau_{t+1}$. A triple $(\sigma^*, \mu^*, \beta^*)$ constitutes a *universally babbling equilibrium* (UBE) if the conditions of a perfect Bayesian equilibrium are satisfied and all consumers’ communication strategy is universally babbling.

Definition 3. A triple $(\sigma^*, \mu^*, \beta^*)$ is a universally babbling equilibrium (UBE) if it is a perfect Bayesian equilibrium of the game and the equilibrium communication strategy μ^* is:

$$\begin{aligned} \mu_1^*(s_i, \tau_2) &= m_1^b \text{ for } i = 1, 2 \text{ for any } \tau_2, \\ \mu_t^*(m_{t-1}, \tau_{t-1}, \tau_{t+1}) &= m_t^b \text{ for any } m_{t-1}, \tau_{t-1}, \tau_{t+1} \text{ for all } t > 1, \end{aligned}$$

where $m_1^b, m_t^b \in \{”s_1”, ”s_2”, ”\emptyset”\}$.

Proposition 3 states an obvious result about the existence of a UBE.

Proposition 3. *A universally babbling equilibrium exists for any values of parameters $\delta \in [0, 1]$, $\rho \in [0, 1]$, $\kappa \in [0, 1]$, and $\epsilon \in [0, 1]$.*

In the UBE, the equilibrium beliefs of each consumer c_t (where $t > 1$) are $\beta_t(s_1 \mid \tau_{t-1}, m_{t-1} = m_{t-1}^b) = \pi$. That is, upon observing the equilibrium message, each player maintains his prior beliefs about the state of the world. Consequently, each consumer c_t 's ($t > 1$) equilibrium action strategy is $\sigma_t^*(m_{t-1}, \tau_{t-1}) = pool_L$ if $\tau_t = L$ and $\sigma_t^*(m_{t-1}, \tau_{t-1}) = pool_H$ if $\tau_t = H$.¹³

The out-of-equilibrium beliefs upon receiving a message $m_{t-1} \neq m_{t-1}^b$ need to sustain the UBE. The UBE is sustained if each consumer c_t holds the following out-of-equilibrium beliefs: for $\tau_t = L$, $\beta_t(s_1 \mid \tau_{t-1}, m_{t-1}) = \pi$ for all $m_{t-1} \neq m_{t-1}^b$, while for $\tau_t = H$, $\beta_t(s_1 \mid \tau_{t-1}, m_{t-1}) = \pi$ for all $m_{t-1} \neq m_{t-1}^b$. These beliefs ensure that each consumer c_{t-1} cannot be better off by sending an out-of-equilibrium message.

3.5 Multiplicity of equilibria

In general, the multiplicity of equilibria arises for several reasons. One of them is the already mentioned multiplicity of permutations of meanings of messages. Besides, the multiplicity arises

¹³There exists a multiplicity of universally babbling equilibria because any message in M can serve as the babbling message m_t^b . It could also be that consumers use different messages as their babbling messages when talking to a consumer of their own type and to a consumer of the other type.

due to the possibility that a consumer may have different communication strategies towards a consumer of his own type and of the other type. Therefore, even in the class of equilibria satisfying Assumption 4, there exist multiple equilibria other than the UTE, HTE, and UBE. For instance, under certain conditions, there exists a heterophilically truthful equilibrium, i.e. an equilibrium in which each consumer c_t reveals his true beliefs about the state of the world if and only if the type of consumer c_{t+1} is other than his own, and if the type of consumer c_{t+1} is the same as c_t 's, then consumer c_t 's communication strategy is uninformative.

The multiplicity of equilibria does not mean that any communication strategy profile can constitute an equilibrium for some parameter values. Consider, for instance, the following communication strategy profile μ :

$$\mu_1(s_i, \tau_2) = \begin{cases} m_1^b & \text{for } i = 1, 2 \text{ for any } \tau_2 \text{ if } \tau_1 = L \\ "s_i" & \text{for } i = 1, 2 \text{ for any } \tau_2 \text{ if } \tau_1 = H \end{cases},$$

$$\mu_t(m_{t-1}, \tau_{t-1}, \tau_{t+1}) = \begin{cases} m_t^b & \text{for any } m_{t-1}, \tau_{t-1}, \tau_{t+1} \text{ if } \tau_t = L \\ m_{t-1} & \text{if } \tau_{t-1} = \tau_t = \tau_{t+1} \text{ and } \tau_t = H \\ " \emptyset " & \text{if } \tau_{t-1} \neq \tau_t = \tau_{t+1} \text{ and } \tau_t = H \end{cases} \quad \text{for all } t > 1.$$

This communication strategy profile implies that each consumer c_t of type $\tau_t = L$ sends a babbling message regardless of whom he speaks to and each consumer c_t of type $c_t = H$ sends a message that reveals his true beliefs about the state regardless of whom he speaks to. In other words, L -type consumers always babble and H -type consumers always communicate truthfully. This cannot be an equilibrium for the following reason: a consumer c_t of type $\tau_t = H$ has no incentive to send a message $m_t = "s_2"$ to a consumer c_{t+1} of type $\tau_{t+1} = L$ because, given the babbling strategy of L -type consumers, there is zero probability that a consumer c_{t+i} (where $i > 1$) of type $\tau_{t+i} = H$ will learn that $s = s_2$. Consumer c_t can then profitably deviate by sending a message $m_t = "s_1"$ and thus induce consumer c_{t+1} to take action *neutral*.

Nevertheless, there is a multiplicity of equilibria in the game and I address this problem in Section 4.

4 Equilibrium Selection

4.1 Undefeated equilibrium

A multiplicity of equilibria often occurs in games of incomplete information and the game presented in this paper is no exception, as discussed in Section 3. This multiplicity significantly limits the usefulness of a model in predicting outcomes. In order to address this problem, several refinements of equilibrium concepts have been proposed, e.g., the intuitive criterion (Cho and Kreps 1987), the divine equilibrium (Banks and Sobel 1987), the neologism-proof equilibrium (Farrell 1985, Farrell 1993), and the announcement-proof equilibrium (Matthews et al. 1993). In this paper, I apply an equilibrium refinement proposed by Mailath et al. (1993), the *undefeated equilibrium*. The motivation for the choice of this particular refinement is that I believe that it captures well the reasons why some of the perfect Bayesian equilibria of the game in this paper seem implausible.¹⁴

The undefeated equilibrium refinement postulates that a receiver of an out-of-equilibrium message should interpret it as the sender's attempt to signal his preference to switch to another equilibrium in which this message is sent on the equilibrium path. The receiver's out-of-equilibrium beliefs in the original equilibrium need to be consistent with this interpretation of the receiver; otherwise, the original equilibrium is said to be defeated by the equilibrium which the sender attempts to switch to. An equilibrium is undefeated if there is no equilibrium that defeats it.

More formally, the idea behind the refinement is as follows. Consider a communication game with a single sender and a single receiver. The sender possesses payoff-relevant private information, $\omega \in \Omega$, which is referred to as his "type". Suppose that $e \equiv (\mu, \sigma, \beta)$ is the original equilibrium of the game and that there exists another equilibrium $e' \equiv (\mu', \sigma', \beta')$. Let m' be a message which is not used in e . Let K be the set of types which send a message m' on the equilibrium path of e' and suppose that K is non-empty. Let $\Omega^+ \subseteq \Omega$ be the set of types which strictly prefer e' to e and send the message m' in e' , and by $\Omega^0 \subseteq \Omega$ the set of types which are indifferent between e' and e and send the message m' in e' . Suppose that all types in K belong either to Ω^+ or Ω^0 but at least some belong to Ω^+ . Then, in the original equilibrium e , upon receiving the out-of-equilibrium message m' , the receiver should hold a belief which is consistent with the fact that a type in Ω^+ would send m' with probability 1 and a type in Ω^0 would send it with a positive probability (allowing for randomising between strategies in the two equilibria), but it would not be sent by a type outside of

¹⁴There is also some experimental support for the use of the undefeated equilibrium concept. Schmidt and Buell (2014) conducted experiments in which they compared the predictive power of the undefeated refinement and the intuitive criterion refinement for decisions of people in operations management. They conclude that the undefeated refinement is more predictive than the intuitive criterion, and the result is particularly pronounced among participants who report a high level of understanding of the game played in the experiment.

K .¹⁵

4.2 Undeclared equilibrium in a communication chain

The undeclared equilibrium refinement by Mailath et al. (1993) was designed for a signalling game of one sender and one receiver and it cannot be directly applied to the game presented in this paper, which consists of a chain of senders and receivers. Therefore, I redefine the refinement so that it can be applied in our model.

I denote the set of pure-strategy perfect Bayesian equilibria of the information diffusion game G by $PBE(G)$. Let $\omega_t \in \Omega = \{s_1, s_2, \emptyset\}$ denote the private information of consumer c_t , which specifies his beliefs about the state of the world, s , and is his “type” (not to be confused with type $\tau_t \in T = \{H, L\}$). With an abuse of notation, I write $U_t(e, \omega_i)$ for the expected payoff to consumer c_t of type ω_t associated with an equilibrium $e \equiv (\mu, \sigma, \beta)$.

Definition 4. An equilibrium $e' \equiv ((\mu'_i)_{c_t}, (\mu_i)_{C \setminus \{c_t\}}, (\sigma'_i)_{c_{t+1}}, (\sigma_i)_{C \setminus \{c_{t+1}\}}, (\beta'_i)_{c_{t+1}}, (\beta_i)_{C \setminus \{c_{t+1}\}}) \in PBE(G)$ defeats $e \equiv (\mu, \sigma, \beta) \in PBE(G)$ if for some $c_t \in C$, $\exists m'_t \in M$ such that

- (1) $\forall \omega_t \in \Omega : \mu_t(\omega_t) \neq m'_t$;
- (2) $\exists \omega_t^* \in \Omega$ such that $\mu'_t(\omega_t^*) = m'_t$ and $\forall \omega_t \neq \omega_t^* : \mu'_t(\omega_t) \neq m'_t$;
- (c) $U_t(e', \omega_t^*) > U_t(e, \omega_t^*)$;
- (d) $\beta_{t+1}(\omega_t^* | m'_t) \neq 1$.

This version of the undeclared equilibrium refinement postulates, like the original refinement, that a receiver of an out-of-equilibrium message should interpret it as the sender’s attempt to signal his preference to switch to an alternative equilibrium in which the message is sent on the equilibrium path. Roughly speaking, the refinement captures the idea that a sender might have some leadership in the choice of equilibrium. For simplicity, I make two stronger requirements in this version of undeclared equilibrium. First, in order to make the refinement suitable for our model, I only consider deviations by a single type from the equilibrium path rather than by a set of types. Second, given that I consider deviations by a single type only, I must require that a sender would signal his preference to switch to an alternative equilibrium only if he strictly prefers it to the original one; otherwise, the new version would be inconsistent with the original definition of the undeclared equilibrium.

¹⁵See Perez-Richet (2014) for a similar discussion of a refinement by Umbhauer (1994), the *consistent forward induction equilibrium*, which differs only slightly from the undeclared equilibrium. Umbhauer’s (1994) refinement is different in that it does not require every type in K to weakly prefer e' to e .

Let me provide a more formal explanation of the refinement in the context of the game discussed in this paper. Suppose that e is the original equilibrium of the game and that there exists another equilibrium e' which differs in that some consumer c_t has a different communication strategy (and, hence, consumer c_{t+1} must have a different action strategy and different beliefs). Like any other consumer in C , consumer c_t possesses payoff-relevant private information, $\omega_t \in \Omega = \{s_1, s_2, \emptyset\}$, which specifies his beliefs about the state of the world, s . Let m'_t be a message which is not used in e by consumer c_t . Suppose that type ω_t^* of consumer c_t is the only type to send message m'_t in equilibrium e' and that he is strictly better off in e' than in e . Then, in the original equilibrium e , upon receiving the out-of-equilibrium message m'_t , consumer c_{t+1} should believe that consumer c_t is of type ω_t^* with probability 1. If it is not the case, i.e. if $\beta_{t+1}(\omega_t^* | m'_t) \neq 1$, then we say that equilibrium e is defeated by equilibrium e' . It is possible that equilibrium e' is itself defeated by another equilibrium, which is also a feature of the original undefeated equilibrium refinement by Mailath et al. (1993).¹⁶ An equilibrium is undefeated if there is no other equilibrium which defeats it.

4.3 Selecting undefeated equilibria

In this section, I investigate which equilibria of the information diffusion game are undefeated within the class of equilibria which satisfy Assumption 4. In order to make the analysis more interesting, I assume that the message set, M_i , consists of more than three messages, contrary to the earlier assumption that $M_i = \{s_1, s_2, \emptyset\}$. With only three messages, any equilibrium such that all three messages are used on the equilibrium path would be trivially undefeated because there would be no messages which are not used in that equilibrium but are used in some other equilibrium. The results on the existence of equilibria presented in Section 3 are not affected by the assumption that the message space consists of more than three messages. Therefore, I now assume that the message set $M_i = \{s_1, s_2, \emptyset, \dots\}$ is countably infinite.

First, consider any equilibrium in which a consumer c_t sends an uninformative message (i.e. babbles) to a consumer of his own type, c_{t+1} . Intuitively, any such equilibrium does not seem reasonable because c_t can only gain by revealing his true beliefs about the state of the world to c_{t+1} . On the one hand, this allows consumer c_{t+1} to take his optimal action given the state of the world, which is also the optimal action from the perspective of consumer c_t . On the other hand, there

¹⁶As Mailath et al. (1993) rightly point out, “using an ‘implausible’ equilibrium to argue the implausibility of another equilibrium may seem suspicious”. However, according to them, this should not be seen as a concern because “[t]he defeating equilibrium is not an object of interest in itself, but rather merely a vehicle to assure the consistency of beliefs that are recommended to a player who observes an out-of-equilibrium message”.

is no risk that consumer c_{t+1} later sends a message which will hurt consumer c_t . This is because their preferences are perfectly aligned, so whatever message consumer c_{t+1} sends is also optimal for consumer c_t . Undefeatedness captures well this reasoning and rejects any such equilibrium, as stated in Proposition 4.

Proposition 4. *Any equilibrium $e \equiv (\mu, \sigma, \beta)$ in which consumer c_t 's communication strategy is $\mu_t(\omega_t, \tau_{t+1} \mid \tau_t = \tau_{t+1}) = m_t^b \forall \omega_t$ is defeated.*

The proof is in the Appendix. Briefly speaking, the reasoning behind this result is as follows. If there exists an equilibrium $e \equiv (\mu, \sigma, \beta)$ in which some consumer c_t of type τ_t sends the same “babbling” message $m_t^b \in \{”s_1”, ”s_2”, ”\emptyset”\}$ to consumer c_{t+1} of the same type, $\tau_{t+1} = \tau_t$, then there also exists an equilibrium e' which differs from e only in that consumer c_t 's strategy is to truthfully reveal his own beliefs about the state of the world to consumer c_{t+1} . Since consumer c_t always sends the same babbling message to c_{t+1} in equilibrium e , there must exist a message m_t' which is not sent by c_t to c_{t+1} in equilibrium e but is sent in equilibrium e' . If consumer c_{t+1} 's beliefs upon receiving m_t' are such that he would choose a different action and/or message than those which he chooses in e , then consumer c_t would strictly benefit from inducing these beliefs if he had the same beliefs about the state of the world. Therefore, by sending message m_t' , consumer c_t with these beliefs can signal his preference for equilibrium e' . Consumer c_{t+1} 's out-of-equilibrium beliefs in e upon receiving message m_t' should be consistent with this reasoning. However, it turns out that the out-of-equilibrium beliefs of c_{t+1} that are necessary to sustain equilibrium e are inconsistent with this reasoning. Hence, equilibrium e is defeated by equilibrium e' .

The result in Proposition 4 rules out all equilibria in which any consumer babbles to a consumer of his own type, including the universally babbling equilibrium (UBE). Therefore, among the equilibria discussed in Section 3, the only candidates for undefeated equilibria are the universally truthful equilibrium (UTE) and the homophilically truthful equilibrium (HTE).

Proposition 5. *The universally truthful equilibrium is undefeated if, for any out-of-equilibrium message m_t' , any consumer c_{t+1} 's belief is $\beta_{t+1}(s_2 \mid m_t') = 1$ for $\tau_t = H$ and $\beta_{t+1}(s_1 \mid m_t') = 1$ for $\tau_t = L$.*

The intuition behind Proposition 5 is that consumer c_t of type $\tau_t = H$ can obtain a strictly higher payoff by switching to another equilibrium e' only if his beliefs are $\omega_t = s_2$. If his beliefs are $\omega_t = s_1$, then he obtains the highest possible payoff in equilibrium e because all consumers of type H and L take optimal actions from his perspective. Similarly, consumer c_t of type $\tau_t = L$ can obtain a strictly higher payoff by switching to another equilibrium e' only if his beliefs are

$\omega_t = s_1$. Therefore, if consumer c_{t+1} , upon receiving an out-of-equilibrium message m'_t , forms a belief $\beta_{t+1}(s_2 | m'_t) = 1$ for $\tau_t = H$ and $\beta_{t+1}(s_1 | m'_t) = 1$ for $\tau_t = L$, then the universally truthful equilibrium is undefeated. This is a sufficient condition which ensures that the UTE is undefeated regardless of exact parameter values of the model.

Proposition 6. *The homophilically truthful equilibrium is undefeated.*

In brief, the homophilically truthful equilibrium cannot be defeated by an equilibrium that differs from the HTE only in that some consumer c_t 's communication strategy is to send an uninformative message to consumer c_{t+1} of type $\tau_{t+1} = \tau_t$. This is for two reasons. First, the preferences of consumers c_t and c_{t+1} are perfectly aligned, so consumer c_t can only lose out on the actions of consumers of his own type by not revealing his true beliefs. Second, given that any consumer $c_{t'}$ such that $t' > t$ and $\tau_{t'} = \tau_t$ will send a babbling message to a consumer $c_{t'+1}$ such that $\tau_{t'+1} = \tau_t$, consumer c_t cannot affect the actions of any consumer of the other type. Therefore, he cannot strictly benefit by babbling to c_{t+1} .

The HTE also cannot be defeated by an equilibrium that differs from the HTE only in that some consumer c_t 's communication strategy is to reveal truthfully his beliefs about the state of the world to consumer c_{t+1} of type $\tau_{t+1} \neq \tau_t$. The reason for this is that such an equilibrium simply does not exist. Given that any consumer $c_{t'}$ such that $t' > t$ and $\tau_{t'} \neq \tau_t$ will send a babbling message to a consumer $c_{t'+1}$ such that $\tau_{t'+1} = \tau_t$, consumer c_t cannot affect the actions of any consumer of his own type by revealing his true beliefs to consumer c_{t+1} . Since he can only affect the actions of consumers of the other type than his own, consumer c_t is not able to credibly communicate his beliefs about the state of the world. Intuitively, such an equilibrium would break down because consumer c_t of type $\tau_t = H$ ($\tau_t = L$) would have an incentive to overstate (understate) his beliefs about the state of the world while talking to consumer c_{t+1} of type $\tau_{t+1} = L$ ($\tau_{t+1} = H$).

I conclude that only two equilibria are undefeated in the class of equilibria satisfying Assumption 4 and with the property that any two consumers c_t and $c_{t'}$ of the same type (i.e. $\tau_t = \tau_{t'}$) have the same communication strategy. These two equilibria are the universally truthful equilibrium and the homophilically truthful equilibrium.

5 Optimal Targeting

In this section I analyse the optimal targeting strategy for the producer. At the beginning of the game, the producer can choose whether to allow a consumer of type $\tau_1 = H$ or $\tau_1 = L$ to learn the true state of the world, s . The message about the state of the world then spreads by cheap

talk via a communication chain, as described in Section 3. I study the conditions under which it is optimal to target each of the two types of the consumer given that the producer's objective is to maximise the expected value of consumers' actions. I focus on equilibria in which all consumers of the same type have the same strategies, so it is sufficient to consider only two possible options of the producer: targeting a consumer c_1 of type $\tau_1 = L$ and c_1 of type $\tau_1 = H$. The following question seems particularly interesting: is it ever optimal to target a consumer c_1 of type $\tau_1 = L$? To keep it brief, I consider only the two undefeated equilibria: the universally truthful equilibrium and the homophilically truthful equilibrium.

5.1 Optimal targeting in the universally truthful equilibrium

It is easy to see that in the universally truthful equilibrium it is always optimal to target consumer c_1 of type $\tau_1 = H$. Proposition 7 states this result.

Proposition 7. *In the universally truthful equilibrium, it is optimal to target consumer c_1 of type $\tau_1 = H$ for all possible values of π , ρ , κ , and ϵ .*

The expected payoff in period 1 is greater when an enthusiast is targeted. If consumers are matched randomly (i.e. $\kappa = 0$), then expected payoffs from actions in periods 2 and later are not affected by who is targeted. Therefore, it is optimal for the producer to target an enthusiast because it maximises the expected payoff in period 1. If the matching is positively assortative (i.e. $\kappa > 0$), then targeting an enthusiast becomes even more attractive as it increases the probability that an enthusiast - rather than a skeptic - receives the message in any future period.

5.2 Optimal targeting in the homophilically truthful equilibrium

Which type is the optimal target in the homophilically truthful equilibrium depends on the pooling actions of the two types of consumers. An enthusiast can take one of two actions once he knows that the sender's message is uninformative: *adopt* or *neutral*. Similarly, a skeptic can take action *neutral* or *reject*. This gives us four combinations of pooling actions: case 1 in which $pool_H = adopt$ and $pool_L = neutral$, case 2 in which $pool_H = adopt$ and $pool_L = reject$, etc. Proposition 8 states the conditions under which it is optimal for the producer to target a skeptic.

Proposition 8. *It is optimal for the producer to target a skeptic if:*

- (a) $\rho \rightarrow 1$, $pool_H = adopt$, $a > a^*$, and $\pi < \pi^*$;
- (b) $\rho \rightarrow 0$, $pool_L = reject$, $b > b^*$, and $\pi > \pi^\dagger$.

The intuition for part (a) of Proposition 8 is as follows. If the pooling action of an enthusiast is *adopt*, then targeting a skeptic guarantees a payoff of a from each enthusiast who happens to receive a message about the product. The payoff of a is guaranteed because, in this equilibrium, babbling is bound to occur before any enthusiast gets a message. As the proportion of enthusiasts in the population approaches 1, targeting an enthusiast effectively means that only enthusiasts receive messages about the product, regardless of the strength of assortativity of matching. In that case, by targeting an enthusiast, the producer obtains an expected payoff of $\pi a + (1 - \pi)b$ in each period of the communication chain. Targeting a skeptic is optimal if the probability of the product being good, π , is not too high and if the producer's benefit from an adoption of the product, a , is sufficiently high (relative to b). These two conditions ensure that a is sufficiently greater than $\pi a + (1 - \pi)b$.

Part (b) can be explained using a similar argument. If the pooling action of a skeptic is *reject*, then targeting a skeptic ensures that the producer does not get a payoff of only 0 from every skeptic who happens to receive a message about the product. This is because if a skeptic is targeted, then at least some skeptics learn the true quality of the product and are neutral with positive probability π . As the proportion of skeptics in the population approaches 1 (i.e. the proportion ρ approaches 0), targeting a skeptic effectively gives an expected payoff of πb in every period, while targeting an enthusiast would mean a payoff of 0 from every skeptic. If the payoff b is sufficiently high (relative to a), then the potential benefit from enthusiasts adopting the product is not high, and so the producer can gain more by preventing at least some of the skeptics from rejecting the product. In the limit, as the proportion ρ converges to 0, if the payoff b is sufficiently high (relative to a), then the producer is better off by targeting a skeptic.

These two cases illustrate the main motivations for targeting a skeptic. Case (a) shows that a producer may want to target a skeptic in order to disinform the enthusiasts and thereby ensure that they do not change their “default” positive attitude towards the product. On the other hand, case (b) shows that a producer may want to target a skeptic in order to inform the skeptics or, in other words, educate them about the qualities of the product.

In most cases, however, it is optimal for the producer to target an enthusiast rather than a skeptic. In particular, if the assortativity of matching is very strong or if the probability of a breakdown of the communication chain is very high, then an enthusiast is the optimal target regardless of the values of other parameters.

Proposition 9. *If $\kappa \rightarrow 1$, it is optimal for the producer to target an enthusiast for any values of parameters $\pi \in [0, 1]$, $\rho \in [0, 1]$ and $\epsilon \in [0, 1]$.*

Under perfect assortativity, only consumers of the targeted consumer's type receive messages about the product. Note that the expected payoff to the producer from the action of an enthusiast who gets to know s is higher than from the action of a skeptic who gets to know s , regardless of the values of parameters. Therefore, it is optimal to target an enthusiast in this case.

Proposition 10. *If $\epsilon \rightarrow 1$, it is optimal for the producer to target an enthusiast for any values of parameters $\pi \in [0, 1]$, $\rho \in [0, 1]$ and $\kappa \in [0, 1]$.*

If the probability of a breakdown of the communication chain is very high, then most likely only one consumer will learn about the existence of the product. Therefore, again, since the expected payoff to the producer from the action of an enthusiast who gets to know s is higher than from the action of a skeptic who gets to know s regardless of the values of parameters, it is optimal for the producer to target an enthusiast.

6 Conclusion

The main contribution of this paper is to study how information diffuses by cheap talk along a communication chain of heterogeneous agents in which only the first agent observes the true state of the world. The paper can explain, for instance, diffusion of information about a product via word of mouth in a population of consumers. In particular, my model fits well a common marketing strategy, often referred to as a limited release of a product, in which access to a new product is initially restricted to a selected group of consumers.

My analysis shows that truthful diffusion of information about a product is possible if several conditions are satisfied: the externality property is at a moderate level, the populations is not too dominated by either of the two types of consumers, the assortativity of matching is not too strong, and the probability of a breakdown of the communication chain is sufficiently low. I show that these conditions can be satisfied only if the preferences of the two types of consumers are sufficiently aligned. Truthful diffusion is not the only outcome that can arise in equilibrium. In fact, there are multiple other equilibria, even if we make a restriction to a class of equilibria such that all consumers either fully truthfully reveal their beliefs about the state of the world or send an uninformative message. However, by applying an extended version of the undefeated equilibrium refinement, I show that, in this class of equilibria, two equilibria are plausible: one in which consumers communicate truthfully regardless of whom they talk to (the universally truthful equilibrium) and one in which consumers communicate truthfully only if they talk to a consumer of their own type (the homophilically truthful equilibrium). As far as the question of optimal targeting is

concerned, targeting a skeptic is optimal for the producer only if the preferences of the two types of consumers are sufficiently misaligned as it is never optimal to target a skeptic in the universally truthful equilibrium. In the homophilically truthful equilibrium, the producer finds it optimal to target a skeptic in two special cases.

The model can be extended in several ways. For instance, if we interpret the actions taken by consumers as their willingness to pay for the product, then we can study the pricing strategy of the producer and its interplay with the targeting strategy. Another potential extension is to introduce a tree network structure by allowing consumers to talk to more than one other consumer. An interesting topic for future research is also to study a similar scenario under the assumption that consumers do not observe the types of consumers with whom they communicate. It could be that a consumer observes a noisy signal of another consumer's type or, in a more extreme case, he has to form a belief about the other consumer's type based only on the information about the population distribution and assortativity. It would be interesting how the results differ from those presented in this paper.

Appendix

Proof of Proposition 1.

In any perfect Bayesian equilibrium, the following condition needs to hold for all consumers c_t :

$$U_t(\sigma^*, \mu^*, \beta^* \mid \omega_t, \tau_{t+1}) \geq U_t(\sigma^*, \mu_{-t}^*, \mu_t, \beta^* \mid \omega_t, \tau_{t+1}) \quad (6.1)$$

$\forall \omega_t, \forall \tau_{t+1}$, and $\forall \mu_t \neq \mu_t^*$.

Abusing notation, by $U_t^{\tau_t}(e, m_t \mid \omega_t, \tau_{t+1})$ I denote the expected payoff to consumer c_t of type τ_t from sending message m_t in the universally truthful equilibrium $e \equiv (\sigma^*, \mu^*, \beta^*)$. In the case of the UTE, condition (6.1) translates into eight conditions. For any consumer c_t of type $\tau_t = H$, the following four conditions need to hold:

$$U_t^H(e, "s_1" \mid \omega_t = s_1, \tau_{t+1} = H) \geq U_t^H(e, "s_2" \mid \omega_t = s_1, \tau_{t+1} = H) \quad (6.2)$$

$$U_t^H(e, "s_2" \mid \omega_t = s_2, \tau_{t+1} = H) \geq U_t^H(e, "s_1" \mid \omega_t = s_2, \tau_{t+1} = H) \quad (6.3)$$

$$U_t^H(e, "s_1" \mid \omega_t = s_1, \tau_{t+1} = L) \geq U_t^H(e, "s_2" \mid \omega_t = s_1, \tau_{t+1} = L) \quad (6.4)$$

$$U_t^H(e, "s_2" \mid \omega_t = s_2, \tau_{t+1} = L) \geq U_t^H(e, "s_1" \mid \omega_t = s_2, \tau_{t+1} = L) \quad (6.5)$$

Similarly, for any consumer c_t of type $\tau_t = L$, the following four conditions need to hold:

$$U_t^L(e, "s_1" \mid \omega_t = s_1, \tau_{t+1} = L) \geq U_t^L(e, "s_2" \mid \omega_t = s_1, \tau_{t+1} = L) \quad (6.6)$$

$$U_t^L(e, "s_2" \mid \omega_t = s_2, \tau_{t+1} = L) \geq U_t^L(e, "s_1" \mid \omega_t = s_2, \tau_{t+1} = L) \quad (6.7)$$

$$U_t^L(e, "s_1" \mid \omega_t = s_1, \tau_{t+1} = H) \geq U_t^L(e, "s_2" \mid \omega_t = s_1, \tau_{t+1} = H) \quad (6.8)$$

$$U_t^L(e, "s_2" \mid \omega_t = s_2, \tau_{t+1} = H) \geq U_t^L(e, "s_1" \mid \omega_t = s_2, \tau_{t+1} = H) \quad (6.9)$$

Inequality (6.2) states that an H-type consumer c_t must prefer to send to another H-type consumer c_{t+1} a message $m_t = "s_1"$ rather than $"s_2"$ when the state is indeed s_1 . The other inequalities are

interpreted in a similar fashion.

Let us first consider $U_t^H(e, "s_1" | \omega_t = s_1, \tau_{t+1} = H)$ for consumer c_t of type $\tau_t = H$. By sending a message $m_t = "s_1"$, the H-type consumer c_i gets a payoff of x_1 from the receiver's action. In the subsequent period, the H-type consumer c_{i+1} is matched with perfect assortativity with probability κ and is matched randomly with probability $1 - \kappa$. For shorter notation, I denote the discount factor $\delta(1 - \epsilon)$ by d , i.e. $d \equiv \delta(1 - \epsilon)$. The expected payoff to any H-type consumer c_i is:

$$\begin{aligned} U_t^H(e, "s_1" | \omega_t = s_1, \tau_{t+1} = H) &= x_1 + \kappa d U_t^H(e, "s_1" | \omega_t = s_1, \tau_{t+1} = H) + \\ &+ (1 - \kappa) d [\rho U_t^H(e, "s_1" | \omega_t = s_1, \tau_{t+1} = H) + \\ &+ (1 - \rho) U_t^H(e, "s_1" | \omega_t = s_1, \tau_{t+1} = L)]. \end{aligned} \quad (6.10)$$

Then, let us consider $U_t^H(e, "s_1" | \omega_t = s_1, \tau_{t+1} = L)$ for consumer c_t of type $\tau_t = H$:

$$\begin{aligned} U_t^H(e, "s_1" | \omega_t = s_1, \tau_{t+1} = L) &= x_1 + \kappa d U_t^H(e, "s_1" | \omega_t = s_1, \tau_{t+1} = L) + \\ &+ (1 - \kappa) \delta d [\rho U_t^H(e, "s_1" | \omega_t = s_1, \tau_{t+1} = H) + \\ &+ (1 - \rho) U_t^H(e, "s_1" | \omega_t = s_1, \tau_{t+1} = L)]. \end{aligned} \quad (6.11)$$

Having these two equations, I can solve for $U_t^H(e, "s_1" | \omega_t = s_1, \tau_{t+1} = H)$ and $U_t^H(e, "s_1" | \omega_t = s_1, \tau_{t+1} = L)$ using the simultaneous equations method. This gives us:

$$U_t^H(e, "s_1" | \omega_t = s_1, \tau_{t+1} = H) = \frac{x_1(1 - d\kappa - d(1 - \kappa)(1 - \rho)) + x_2 d(1 - \kappa)(1 - \rho)}{(1 - d\kappa)(1 - d)} \quad (6.12)$$

and

$$U_t^H(e, "s_1" | \omega_t = s_1, \tau_{t+1} = L) = \frac{x_2(1 - d\kappa - d(1 - \kappa)\rho) + x_1 d(1 - \kappa)\rho}{(1 - d\kappa)(1 - d)} \quad (6.13)$$

Following the same method, I derive the remaining expected payoffs which are needed to solve inequalities (6.2)-(6.9):

$$U_t^H(e, "s_2" | \omega_t = s_1, \tau_{t+1} = H) = \frac{x_2(1 - d\kappa - d(1 - \kappa)(1 - \rho)) + 0}{(1 - d\kappa)(1 - d)}, \quad (6.14)$$

$$U_t^H(e, "s_2" \mid \omega_t = s_2, \tau_{t+1} = H) = \frac{y_2(1 - d\kappa - d(1 - \kappa)(1 - \rho)) + y_3d(1 - \kappa)(1 - \rho)}{(1 - d\kappa)(1 - d)}, \quad (6.15)$$

$$U_t^H(e, "s_2" \mid \omega_t = s_2, \tau_{t+1} = H) = \frac{y_1(1 - d\kappa - d(1 - \kappa)(1 - \rho)) + y_2d(1 - \kappa)(1 - \rho)}{(1 - d\kappa)(1 - d)}, \quad (6.16)$$

$$U_t^H(e, "s_2" \mid \omega_t = s_1, \tau_{t+1} = L) = \frac{0 + x_2d(1 - \kappa)\rho}{(1 - d\kappa)(1 - d)}, \quad (6.17)$$

$$U_t^H(e, "s_2" \mid \omega_t = s_2, \tau_{t+1} = L) = \frac{y_3(1 - d\kappa - d(1 - \kappa)\rho) + y_2d(1 - \kappa)\rho}{(1 - d\kappa)(1 - d)}, \quad (6.18)$$

$$U_t^H(e, "s_1" \mid \omega_t = s_2, \tau_{t+1} = L) = \frac{y_2(1 - d\kappa - d(1 - \kappa)\rho) + y_1d(1 - \kappa)\rho}{(1 - d\kappa)(1 - d)}, \quad (6.19)$$

$$U_t^L(e, "s_1" \mid \omega_t = s_1, \tau_{t+1} = L) = \frac{y_2(1 - d\kappa - d(1 - \kappa)\rho) + y_3d(1 - \kappa)\rho}{(1 - d\kappa)(1 - d)}, \quad (6.20)$$

$$U_t^L(e, "s_2" \mid \omega_t = s_1, \tau_{t+1} = L) = \frac{y_1(1 - d\kappa - d(1 - \kappa)\rho) + y_2d(1 - \kappa)\rho}{(1 - d\kappa)(1 - d)}, \quad (6.21)$$

$$U_t^L(e, "s_2" \mid \omega_t = s_2, \tau_{t+1} = L) = \frac{x_1(1 - d\kappa - d(1 - \kappa)\rho) + x_2d(1 - \kappa)\rho}{(1 - d\kappa)(1 - d)}, \quad (6.22)$$

$$U_t^L(e, "s_1" \mid \omega_t = s_2, \tau_{t+1} = L) = \frac{x_2(1 - d\kappa - d(1 - \kappa)\rho) + 0}{(1 - d\kappa)(1 - d)}, \quad (6.23)$$

$$U_t^L(e, "s_1" \mid \omega_t = s_1, \tau_{t+1} = H) = \frac{y_3(1 - d\kappa - d(1 - \kappa)(1 - \rho)) + y_2d(1 - \kappa)(1 - \rho)}{(1 - d\kappa)(1 - d)}, \quad (6.24)$$

$$U_t^L(e, "s_2" \mid \omega_t = s_1, \tau_{t+1} = H) = \frac{y_2(1 - d\kappa - d(1 - \kappa)(1 - \rho)) + y_1d(1 - \kappa)(1 - \rho)}{(1 - d\kappa)(1 - d)}, \quad (6.25)$$

$$U_t^L(e, "s_2" \mid \omega_t = s_2, \tau_{t+1} = H) = \frac{x_2(1 - d\kappa - d(1 - \kappa)(1 - \rho)) + x_1d(1 - \kappa)(1 - \rho)}{(1 - d\kappa)(1 - d)}, \quad (6.26)$$

$$U_t^L(e, "s_1" \mid \omega_t = s_2, \tau_{t+1} = H) = \frac{0 + x_2d(1 - \kappa)(1 - \rho)}{(1 - d\kappa)(1 - d)}. \quad (6.27)$$

Simple calculation shows that inequality (6.2) holds for all possible values of parameters κ , d , ρ , x_1 , and x_2 (given that I assume $x_1 > x_2$). Using the same approach, I obtain that:

- inequalities (6.4), (6.7), and (6.9) hold for all possible values of the parameters,
- inequality (6.3) holds iff $\delta \leq \frac{y_2 - y_1}{(1 - \epsilon)[\kappa(y_2 - y_1) + (1 - \kappa)(1 - \rho)(2y_2 - y_1 - y_3)]}$,
- inequality (6.5) holds iff $\delta \geq \frac{y_2 - y_3}{(1 - \epsilon)[\kappa(y_2 - y_3) + (1 - \kappa)\rho(2y_2 - y_1 - y_3)]}$,
- inequality (6.6) holds iff $\delta \leq \frac{y_2 - y_1}{(1 - \epsilon)[\kappa(y_2 - y_1) + (1 - \kappa)\rho(2y_2 - y_1 - y_3)]}$,
- inequality (6.8) holds iff $\delta \geq \frac{y_2 - y_3}{(1 - \epsilon)[\kappa(y_2 - y_3) + (1 - \kappa)(1 - \rho)(2y_2 - y_1 - y_3)]}$.

A consumer c_t of type $\tau_t = H$ has no incentive to deviate from his equilibrium strategy in the UTE if and only if inequalities (6.2)-(6.5) hold, i.e. if and only if $\delta \in [\delta^*, \delta^{**}]$:

$$\frac{y_2 - y_3}{(1 - \epsilon)[\kappa(y_2 - y_3) + (1 - \kappa)\rho(2y_2 - y_1 - y_3)]} \leq \delta \leq \frac{y_2 - y_1}{(1 - \epsilon)[\kappa(y_2 - y_1) + (1 - \kappa)(1 - \rho)(2y_2 - y_1 - y_3)]}. \quad (6.28)$$

A consumer c_t of type $\tau_t = L$ has no incentive to deviate from his equilibrium strategy in the UTE if and only if inequalities (6.6)-(6.9) hold, i.e. if and only if $\delta \in [\delta^\dagger, \delta^{\dagger\dagger}]$:

$$\frac{y_2 - y_3}{(1 - \epsilon)[\kappa(y_2 - y_3) + (1 - \kappa)(1 - \rho)(2y_2 - y_1 - y_3)]} \leq \delta \leq \frac{y_2 - y_1}{(1 - \epsilon)[\kappa(y_2 - y_1) + (1 - \kappa)\rho(2y_2 - y_1 - y_3)]}. \quad (6.29)$$

In a truthful perfect Bayesian equilibrium, both conditions (6.28) and (6.29) need to hold. For $\rho < \frac{1}{2}$, (6.28) is both a necessary and a sufficient condition for the existence of a truthful equilibrium. For $\rho \geq \frac{1}{2}$, (6.29) is both a necessary and a sufficient condition for the existence of a truthful equilibrium.

The conditions in parts (b), (c), and (d) of Proposition 1 are obtained by rearranging conditions (6.28) and (6.29).

Proof of Corollary 1.

This follows directly from Proposition 1. First, let us consider $\rho < \frac{1}{2}$. For there to exist a δ such that

$$\frac{y_2 - y_3}{(1 - \epsilon)[\kappa(y_2 - y_3) + (1 - \kappa)\rho(2y_2 - y_1 - y_3)]} \leq \delta \leq \frac{y_2 - y_1}{(1 - \epsilon)[\kappa(y_2 - y_1) + (1 - \kappa)(1 - \rho)(2y_2 - y_1 - y_3)]},$$

it must be that

$$\frac{y_2 - y_3}{(1 - \epsilon)[\kappa(y_2 - y_3) + (1 - \kappa)\rho(2y_2 - y_1 - y_3)]} \leq \frac{y_2 - y_1}{(1 - \epsilon)[\kappa(y_2 - y_1) + (1 - \kappa)(1 - \rho)(2y_2 - y_1 - y_3)]}, \quad (6.30)$$

which simplifies to

$$\frac{y_2 - y_1}{y_2 - y_3} \geq \frac{1 - \rho}{\rho}. \quad (6.31)$$

Next, let us consider $\rho \geq \frac{1}{2}$. For there to exist a δ such that

$$\frac{y_2 - y_3}{(1 - \epsilon)[\kappa(y_2 - y_3) + (1 - \kappa)(1 - \rho)(2y_2 - y_1 - y_3)]} \leq \delta \leq \frac{y_2 - y_1}{(1 - \epsilon)[\kappa(y_2 - y_1) + (1 - \kappa)\rho(2y_2 - y_1 - y_3)]},$$

it must be that

$$\frac{y_2 - y_3}{(1 - \epsilon)[\kappa(y_2 - y_3) + (1 - \kappa)(1 - \rho)(2y_2 - y_1 - y_3)]} \leq \frac{y_2 - y_1}{(1 - \epsilon)[\kappa(y_2 - y_1) + (1 - \kappa)\rho(2y_2 - y_1 - y_3)]}, \quad (6.32)$$

which simplifies to

$$\frac{y_2 - y_1}{y_2 - y_3} \geq \frac{\rho}{1 - \rho}. \quad (6.33)$$

Proof of Corollary 2.

This follows from Corollary 1. If $\rho < \frac{1}{2}$, then $\frac{1-\rho}{\rho} > 1$, and hence from (6.33) I get $\frac{y_2-y_3}{y_2-y_1} > 1$. If $\rho \geq \frac{1}{2}$, then $\frac{\rho}{1-\rho} \geq 1$, and hence from (6.33) I obtain $\frac{y_2-y_3}{y_2-y_1} > 1$. Given that I assume that $y_2 > y_1$

and $y_2 > y_3$, the inequality $\frac{y_2 - y_3}{y_2 - y_1} > 1$ is equivalent to $y_1 > y_3$.

Proof of Proposition 4.

Consider any equilibrium $e \equiv (\mu, \sigma, \beta)$ such that there is a consumer c_t ($t \geq 1$) whose communication strategy is $\mu_1(s_i, \tau_2 \mid \tau_2 = \tau_1) = m_t^b$ for any $i = 1, 2$ or $\mu_t(m_{t-1}, \tau_{t-1}, \tau_{t+1} \mid \tau_{t+1} = \tau_t) = m_t^b$ for any m_{t-1}, τ_{t-1} . Suppose w.l.o.g. that consumer c_{t+1} 's equilibrium action is the same action that he would have taken if his belief were $\omega_{t+1} = s_i$ for some state s_i , i.e. $\sigma_{t+1}(m_t = m_t^b) = \sigma_{t+1}(\omega_{t+1} = s_i)$ with some abuse of notation.

Step 1: I show that for equilibrium e to be sustained, the belief of consumer c_{t+1} upon receiving an out-of-equilibrium message $m'_t \neq m_t^b$, $\beta_{t+1}(s_1 \mid m_t = m'_t)$, needs to be strictly positive, i.e. $\beta_{t+1}(s_1 \mid m_t = m'_t) > 0$.

Note that for equilibrium e to be sustained, the belief of consumer c_{t+1} upon receiving an out-of-equilibrium message $m'_t \neq m_t^b$, $\beta_{t+1}(s_1 \mid m_t = m'_t)$, needs to be such that c_{t+1} takes an action $\sigma_{t+1}(m_t = m_t^b)$ and sends a message $\mu_{t+1}(m_t = m_t^b)$, i.e. $\sigma_{t+1}(m_t = m'_t) = \sigma_{t+1}(m_t = m_t^b)$ and $\mu_{t+1}(m_t = m'_t) = \mu_{t+1}(m_t = m_t^b)$. To show this, suppose $\beta_{t+1}(s_1 \mid m_t = m'_t)$ is such that $\sigma_{t+1}(m_t = m'_t) \neq \sigma_{t+1}(m_t = m_t^b)$ and/or $\mu_{t+1}(m_t = m'_t) \neq \mu_{t+1}(m_t = m_t^b)$. Since consumer c_t has the same preferences as c_{t+1} , consumer c_t would profitably deviate by sending message m'_t if his beliefs were $\beta_t(s_1) = \beta_{t+1}(s_1 \mid m_t = m'_t)$. For $\sigma_{t+1}(m_t = m'_t) = \sigma_{t+1}(m_t = m_t^b)$ and $\mu_{t+1}(m_t = m'_t) = \mu_{t+1}(m_t = m_t^b)$ to hold, we need $\beta_{t+1}(s_i \mid m_t = m'_t)$ to be sufficiently high and necessarily $\beta_{t+1}(s_i \mid m_t = m'_t) > 0$. If $\beta_{t+1}(s_i \mid m_t = m'_t) = 0$, then consumer c_{t+1} would optimally take an action $\sigma_{t+1}(\omega_{t+1} = s_j) \neq \sigma_{t+1}(m_t = m_t^b)$.

Step 2: If equilibrium e exists, then there also exists an equilibrium

$$e' \equiv ((\mu'_i)_{c_t}, (\mu_i)_{C \setminus \{c_t\}}, (\sigma'_i)_{c_{t+1}}, (\sigma_i)_{C \setminus \{c_{t+1}\}}, (\beta'_i)_{c_{t+1}}, (\beta_i)_{C \setminus \{c_{t+1}\}})$$

such that

$$\mu'_t(\omega_t, \tau_{t+1}) = \begin{cases} \text{"}s_1\text{"} & \text{if } \omega_t = s_1 \\ \text{"}s_2\text{"} & \text{if } \omega_t = s_2 \\ \text{"}\emptyset\text{"} & \text{if } \omega_t = \emptyset \end{cases} \quad (6.34)$$

for some consumer c_t . If $\mu_{t+1}(\omega_{t+1}, \tau_{t+2}) = m_{t+1}^b \forall \omega_{t+1}$, then e' clearly is an equilibrium because preferences of c_t and c_{t+1} are perfectly aligned. In the only other possible case (given Assumption 4), if

$$\mu_{t+1}(\omega_{t+1}, \tau_{t+2}) = \begin{cases} "s_1" & \text{if } \omega_{t+1} = s_1 \\ "s_2" & \text{if } \omega_{t+1} = s_2, \\ "\emptyset" & \text{if } \omega_{t+1} = \emptyset \end{cases} \quad (6.35)$$

then e' is also an equilibrium because preferences of c_t and c_{t+1} are perfectly aligned and, therefore, if equilibrium strategy of c_{t+1} is given by (6.35), then the strategy in (6.34) can also be an equilibrium strategy.

Step 3: I show that equilibrium e' (or another outcome-equivalent equilibrium) defeats equilibrium e .

Case (a): In equilibrium e , suppose that either $m_t^b = "s_i"$ or $m_t^b = "\emptyset"$ and that consumer c_{t+1} 's equilibrium response, $\sigma_{t+1}(m_t = m_t^b)$, is the action which he would have taken in state s_i , i.e. $\sigma_{t+1}(m_t = m_t^b) = \sigma_{t+1}(\omega_{t+1} = s_i)$. Then, in equilibrium e , a message $m_t' = "s_j"$ (where $i \neq j$) is not sent from c_t to c_{t+1} regardless of c_t 's beliefs about the state of the world, given by $\omega_t \in \Omega = \{s_1, s_2, \emptyset\}$. Therefore, $\forall \omega_t \in \Omega : \mu_t(\omega_t) \neq "s_j"$. However, the message $m_t' = "s_j"$ is sent in equilibrium e' by type $\omega_t^* = s_j$ and no other type $\omega_t \in \Omega$. Furthermore, type $\omega_t^* = s_j$ is strictly better off in e' than in e . Therefore, consumer c_{t+1} 's beliefs should be $\beta_{t+1}(s_j | m_t' = "s_j") = 1$, which implies $\beta_{t+1}(s_i | m_t' = "s_j") = 0$. This is not consistent with the earlier requirement that $\beta_{t+1}(s_i | m_t = m_t^b) > 0$, so equilibrium e is defeated by equilibrium e' .

Case (b): In equilibrium e , suppose that $m_t^b = "s_j"$ and that consumer c_{t+1} optimally takes an action which he would have taken in state s_i , i.e. $\sigma_{t+1}(\omega_{t+1} = \emptyset) = \sigma_{t+1}(\omega_{t+1} = s_i)$. Then, I can construct an equilibrium that is outcome-equivalent to e' , denoted by e'' , in which equilibrium meanings of consumer c_t 's messages are reversed, i.e. $\mu_t''(\omega_t, \tau_{t+1}) = "s_j"$ if $\omega_t = s_i$ and $\tau_t = \tau_{t+1}$ where $i, j \in \{1, 2\}$ and $i \neq j$. Then, in equilibrium e , a message $m_t' = "s_i"$ is not sent from c_t to c_{t+1} . However, the message $m_t' = "s_i"$ is sent in equilibrium e'' by type $\omega_t^* = s_j$ and no other type $\omega_t \in \Omega$. Furthermore, type $\omega_t^* = s_j$ is strictly better off in e'' than in e . Therefore, consumer c_{t+1} 's beliefs should be $\beta_{t+1}(s_j | m_t' = "s_i") = 1$, which implies $\beta_{t+1}(s_i | m_t' = "s_i") = 0$. Again, this is not consistent with the earlier requirement that $\beta_{t+1}(s_i | m_t = m_t^b) > 0$, so equilibrium e is defeated by e'' .

Note that in this proof I actually do not need to extend the message space M beyond three messages.

Proof of Proposition 5.

First, note that equilibrium e cannot be defeated by an equilibrium e' which is outcome-equivalent to e but in which some consumer c_t uses a different set of messages from the message space M to convey the meanings: $\omega_t = s_1$, $\omega_t = s_2$, and $\omega_t = \emptyset$. From the equilibrium conditions that need to be satisfied in e , it follows that consumer c_t cannot be strictly better off by sending a message that is sent in e' but is not sent in e .

Second, consider the only other possibility given Assumption 4, i.e. consider any equilibrium, $e' \equiv (\mu', \sigma', \beta')$ which differs from the universally truthful equilibrium, $e \equiv (\mu, \sigma, \beta)$, in that the communication strategy of consumer c_t , is $\mu'_t(\omega_t, \tau_{t+1}) = m_t^b \forall \omega_t \in \{s_1, s_2\}$. On the equilibrium path, consumer c_{t+1} believes that consumer c_t knows the state of the world, s . However, consumer c_t of type $\tau_t = H$ with a belief $\omega_t = s_1$ cannot possibly be strictly better off in e' than in e because he gets the highest possible payoff in e . So the only type ω_t of consumer c_t which can benefit from moving away from e is $\omega_t = s_2$ (I omit $\omega_t = \emptyset$ because it cannot appear on the equilibrium path in e and e'). Similarly, the only type ω_t of consumer c_t which can benefit from moving away from e is $\omega_t = s_1$. Therefore, if consumer c_{t+1} 's out-of-equilibrium is $\beta_{t+1}(s_2 | m_t^b) = 1$ for $\tau_t = H$ and $\beta_{t+1}(s_1 | m_t^b) = 1$ for $\tau_t = L$, then equilibrium e is undefeated.

Proof of Proposition 6.

First, consider the homophilically truthful equilibrium e and take any two consumers c_t and c_{t+1} such that $\tau_t = \tau_{t+1}$. Then, given the equilibrium strategies of the HTE, no type ω_t of consumer c_t can strictly benefit from switching to an equilibrium e' which differs from the HTE only in that his communication strategy towards c_{t+1} is $\mu'_t(\omega_t, \tau_{t+1}) = m_t^b \forall \omega_t \in \{s_1, s_2\}$. First, the preferences of c_t , c_{t+1} , and all subsequent consumers of type $\tau_i = \tau_t$ are perfectly aligned, so there is no incentive to not truthfully reveal one's beliefs about the state of the world. Second, since any consumer $c_{t'}$ such that $t' > t$ and $\tau_{t'} = \tau_t$ will send a message $m_{t'}^b$ to a consumer $c_{t'+1}$ of type $\tau_{t'+1} \neq \tau_t$, consumer c_t cannot affect the action of any consumer of type $\tau_{t'+1} \neq \tau_t$, so there is no incentive for him to switch to any equilibrium e' for the purpose of affecting the actions of consumers of type $\tau_{t'+1} \neq \tau_t$. Since no type ω_t of consumer c_t can strictly benefit from switching to an equilibrium e' , equilibrium e' cannot defeat equilibrium e .

Second, consider the homophilically truthful equilibrium e and take any two consumers c_t and c_{t+1} such that $\tau_t \neq \tau_{t+1}$. Consider a potential equilibrium e' which differs from the HTE only in that consumer c_t is:

$$\mu'_t(\omega_t, \tau_{t+1}) = \begin{cases} "s_1" & \text{if } \omega_t = s_1 \\ "s_2" & \text{if } \omega_t = s_2 \\ " \emptyset " & \text{if } \omega_t = \emptyset \end{cases} \quad (6.36)$$

Communication strategies of all other consumers in e' are as in equilibrium e , i.e. consumers would truthfully reveal their beliefs to consumers of their own type but send a babbling message to consumers of the other type. Clearly, e' cannot be an equilibrium because consumer c_t of any type ω_t would have an incentive to send a message "s₁" if $\tau_{t+1} = L$ and "s₂" if $\tau_{t+1} = H$. Since there exists no such equilibrium e' , e cannot be defeated by e' .

Proof of Proposition 7.

The expected payoff to the producer is denoted by $U_P(\sigma, \mu, \beta, \varphi)$. Abusing notation, by $U_P(e, \varphi = H)$ I denote expected payoff to the producer from targeting a consumer c_1 of type $\tau_1 = H$ in equilibrium $e \equiv (\sigma, \mu, \beta)$, and by $U_P(e, \varphi = L)$ I denote expected payoff to the producer from targeting a consumer c_1 of type $\tau_1 = L$ in equilibrium $e \equiv (\sigma, \mu, \beta)$. The expected payoff from targeting an H-type consumer c_1 is

$$U_P(e, \varphi = H) = \pi a + (1 - \pi)b + \kappa(1 - \epsilon)U_P(e, \varphi = H) + (1 - \kappa)(1 - \epsilon)(\rho U_P(e, \varphi = H) + (1 - \rho)U_P(e, \varphi = L)). \quad (6.37)$$

The expected payoff from targeting an L-type consumer c_1 is

$$U_P(e, \varphi = L) = \pi b + \kappa(1 - \epsilon)U_P(e, \varphi = L) + (1 - \kappa)(1 - \epsilon)(\rho U_P(e, \varphi = H) + (1 - \rho)U_P(e, \varphi = L)). \quad (6.38)$$

By rearranging, I obtain:

$$U_P(e, \varphi = H) = \frac{\pi a + (1 - \pi)b + (1 - \kappa)(1 - \epsilon)(\rho U_P(e, \varphi = H) + (1 - \rho)U_P(e, \varphi = L))}{1 - \kappa(1 - \epsilon)}, \quad (6.39)$$

$$U_P(e, \varphi = L) = \frac{\pi b + (1 - \kappa)(1 - \epsilon)(\rho U_P(e, \varphi = H) + (1 - \rho)U_P(e, \varphi = L))}{1 - \kappa(1 - \epsilon)}. \quad (6.40)$$

Now, by subtracting (6.40) from (6.39) I obtain:

$$U_P(e, \varphi = H) - U_P(e, \varphi = L) = \frac{\pi(a - b) + (1 - \pi)b}{1 - \kappa(1 - \epsilon)} > 0 \quad (6.41)$$

for all possible values of parameters π , κ , and ϵ , given that I assume that $a > b$.

Proof of Proposition 8.

In the HTE, the expected payoff to the producer from targeting an H-type consumer c_1 is:

$$U_P(e, \varphi = H) = \pi\alpha + (1 - \pi)\beta + \kappa(1 - \epsilon)U_P(e, \varphi = H) + (1 - \kappa)(1 - \epsilon)(\rho U_P(e, \varphi = H) + (1 - \rho)V_P(L)), \quad (6.42)$$

where $V_P(L)$ is the expected payoff to the producer from a communication chain which starts with a consumer $c_t = L$ and in which all consumers take their pooling actions. Similarly, the expected payoff from targeting an L-type consumer c_1 is:

$$U_P(e, \varphi = L) = \pi\beta + \kappa(1 - \epsilon)U_P(e, \varphi = L) + (1 - \kappa)(1 - \epsilon)(\rho V_P(H) + (1 - \rho)U_P(e, \varphi = L)), \quad (6.43)$$

where $V_P(H)$ is the expected payoff from a communication chain which starts with a consumer $c_t = H$ and in which all consumers take their pooling actions. Let us expand the terms $V_P(L)$ and $V_P(H)$:

$$V_P(L) = u_P(pool_L) + \kappa(1 - \epsilon)V_P(L) + (1 - \kappa)(1 - \epsilon)(\rho V_P(H) + (1 - \rho)V_P(L)), \quad (6.44)$$

$$V_P(H) = u_P(pool_H) + \kappa(1 - \epsilon)V_P(H) + (1 - \kappa)(1 - \epsilon)(\rho V_P(H) + (1 - \rho)V_P(L)), \quad (6.45)$$

where $u_P(pool_L)$ and $u_P(pool_H)$ are the producer's payoffs from the pooling actions of L-type and an H-type consumers respectively.

Case 1: ρ converges to 1, i.e. $\rho \rightarrow 1$. Then, the value of $V_P(H)$ simplifies to:

$$\lim_{\rho \rightarrow 1} V_P(H) = \frac{u_P(pool_H)}{1 - (1 - \epsilon)}, \quad (6.46)$$

which, together with equations (6.42) and (6.43), allows us to obtain the following values of $U_P(e, \varphi = H)$ and $U_P(e, \varphi = L)$ for $\rho \rightarrow 1$:

$$\lim_{\rho \rightarrow 1} U_P(e, \varphi = H) = \frac{\pi a + (1 - \pi)b}{1 - (1 - \epsilon)}, \quad (6.47)$$

$$\lim_{\rho \rightarrow 1} U_P(e, \varphi = L) = \frac{\pi b + (1 - \epsilon)(1 - \kappa) \frac{u_P(pool_H)}{1 - \delta}}{1 - (1 - \epsilon)\kappa}. \quad (6.48)$$

Hence, for $\rho \rightarrow 1$, $U_P(e, \varphi = L)$ is greater than $U_P(e, \varphi = H)$ if and only if:

$$\lim_{\rho \rightarrow 1} U_P(e, \varphi = L) > \lim_{\rho \rightarrow 1} U_P(e, \varphi = H). \quad (6.49)$$

After substituting (6.47) and (6.48) into (6.49), and rearranging, I obtain:

$$\frac{\pi b + (1 - \epsilon)(1 - \kappa) \frac{u_P(pool_H)}{1 - \delta}}{1 - (1 - \epsilon)\kappa} > \frac{\pi a + (1 - \pi)b}{1 - (1 - \epsilon)}, \quad (6.50)$$

which simplifies to:

$$u_P(pool_H) > \frac{(1 - (1 - \epsilon)\kappa)(\pi a + (1 - \pi)b) - (1 - (1 - \epsilon))\pi b}{(1 - \epsilon)(1 - \kappa)}. \quad (6.51)$$

By substituting a and b for $u_P(pool_H)$ (the only possible values that $u_P(pool_H)$ can take) and rearranging, I obtain that if $u_P(pool_H) = a$ and $\pi < \frac{(1 - \epsilon) - (1 - \epsilon)\kappa}{1 - (1 - \epsilon)\kappa}$, then inequality (6.51) is satisfied if and only if $a > b \frac{(1 - (1 - \epsilon)\kappa)(1 - \pi) - (1 - (1 - \epsilon))\pi}{(1 - \epsilon)(1 - \kappa) - (1 - (1 - \epsilon)\kappa)\pi}$, which is always true because for $\pi < \frac{(1 - \epsilon) - (1 - \epsilon)\kappa}{1 - (1 - \epsilon)\kappa}$, $\frac{(1 - (1 - \epsilon)\kappa)(1 - \pi) - (1 - (1 - \epsilon))\pi}{(1 - \epsilon)(1 - \kappa) - (1 - (1 - \epsilon)\kappa)\pi}$ is greater than 1 for all possible parameter values given that I assume that $a > b$. If $u_P(pool_H) = a$ and $\pi > \frac{(1 - \epsilon) - (1 - \epsilon)\kappa}{1 - (1 - \epsilon)\kappa}$ or $u_P(pool_H) = b$, then there does not exist $b \in (0, a)$ which can satisfy inequality (6.51).

Case 2: ρ converges to 0, i.e. $\rho \rightarrow 0$. Then, the value of $V_P(L)$ simplifies to:

$$\lim_{\rho \rightarrow 0} V_P(L) = \frac{u_P(pool_L)}{1 - (1 - \epsilon)}, \quad (6.52)$$

which, together with equations (6.42) and (6.43), allows us to obtain the following values of $U_P(e, \varphi = H)$ and $U_P(e, \varphi = L)$:

$$\lim_{\rho \rightarrow 0} U_P(e, \varphi = H) = \frac{\pi a + (1 - \pi)b + (1 - \epsilon)(1 - \kappa) \frac{u_P(\text{pool}_L)}{1 - \delta}}{1 - (1 - \epsilon)\kappa}, \quad (6.53)$$

$$\lim_{\rho \rightarrow 0} U_P(e, \varphi = L) = \frac{\pi b}{1 - (1 - \epsilon)}. \quad (6.54)$$

Hence, for $\rho \rightarrow 0$, $U_P(e, \varphi = L)$ is greater than $U_P(e, \varphi = H)$ if and only if:

$$\lim_{\rho \rightarrow 0} U_P(e, \varphi = L) > \lim_{\rho \rightarrow 0} U_P(e, \varphi = H). \quad (6.55)$$

After substituting (6.53) and (6.54) into (6.55), and rearranging, I obtain:

$$\frac{\pi b}{1 - \delta} > \frac{\pi a + (1 - \pi)b + (1 - \epsilon)(1 - \kappa) \frac{u_P(\text{pool}_L)}{1 - \delta}}{1 - (1 - \epsilon)\kappa}, \quad (6.56)$$

which simplifies to:

$$u_P(\text{pool}_L) < \frac{(1 - (1 - \epsilon)\kappa)\pi b - (1 - (1 - \epsilon))(\pi a + (1 - \pi)b)}{(1 - \epsilon)(1 - \kappa)}. \quad (6.57)$$

By substituting b and 0 for $u_P(\text{pool}_L)$ (the only possible values that $u_P(\text{pool}_L)$ can take) and rearranging, I obtain that if $u_P(\text{pool}_L) = 0$ and $\pi > \frac{1 - (1 - \epsilon)}{1 - (1 - \epsilon) + 1 - (1 - \epsilon)\kappa}$, then inequality (6.57) is satisfied if and only if $b > \frac{-(1 - (1 - \epsilon))\pi}{(1 - (1 - \epsilon))(1 - \pi) - (1 - (1 - \epsilon)\kappa)\pi} a$ (and $a > b$, naturally). If $u_P(\text{pool}_L) = 0$ and $\pi < \frac{1 - (1 - \epsilon)}{1 - (1 - \epsilon) + 1 - (1 - \epsilon)\kappa}$ or $u_P(\text{pool}_L) = b$, then there does not exist $b \in (0, a)$ which can satisfy inequality (6.57).

Proof of Proposition 9.

If $\kappa \rightarrow 1$, the expected payoff to the producer from targeting a consumer c_1 of type $\tau_1 = H$ is:

$$\lim_{\kappa \rightarrow 1} U_P(e, \varphi = H) = \frac{\pi a + (1 - \pi)b}{1 - (1 - \epsilon)}. \quad (6.58)$$

Similarly, the expected payoff to the producer from targeting a consumer c_1 of type $\tau_1 = L$ is:

$$\lim_{\kappa \rightarrow 1} U_P(e, \varphi = L) = \frac{\pi b}{1 - (1 - \epsilon)}. \quad (6.59)$$

Clearly, the inequality $\lim_{\kappa \rightarrow 1} U_P(e, \varphi = H) > \lim_{\kappa \rightarrow 1} U_P(e, \varphi = L)$ holds for all $\pi \in (0, 1)$ and $\epsilon \in (0, 1)$ since we assume that $a > b$.

Proof of Proposition 10.

If $\epsilon \rightarrow 1$, the expected payoff to the producer from targeting a consumer c_1 of type $\tau_1 = H$ is:

$$\lim_{\epsilon \rightarrow 1} U_P(e, \varphi = H) = \pi a + (1 - \pi)b. \quad (6.60)$$

Similarly, the expected payoff to the producer from targeting a consumer c_1 of type $\tau_1 = L$ is:

$$\lim_{\epsilon \rightarrow 1} U_P(e, \varphi = L) = \pi b. \quad (6.61)$$

Clearly, the inequality $\lim_{\epsilon \rightarrow 1} U_P(e, \varphi = H) > \lim_{\epsilon \rightarrow 1} U_P(e, \varphi = L)$ holds for all $\pi \in (0, 1)$ since we assume that $a > b$.

References

- [1] Ambrus, A., E. Azevedo and Y. Kamada (2013): “Hierarchical Cheap Talk,” *Theoretical Economics*, 8, 233-261.
- [2] Anderlini, L., D. Gerardi and R. Lagunoff (2012): “Communication and Learning,” *Review of Economic Studies*, 79, 419-450.
- [3] Anderlini, L. and R. Lagunoff (2005): “Communication in dynastic repeated games: ‘White-washes’ and ‘coverups’,” *Economic Theory*, 26, 265-299.
- [4] Arndt, J. (1967): *Word of Mouth Advertising: A Review of Literature*, New York: Advertising Research Foundation.
- [5] Banerjee, A.V. (1993): “The Economics of Rumours,” *Review of Economic Studies*, 60, 309-327.
- [6] Banks, J.S. and J. Sobel (1987): “Equilibrium Selection in Signaling Games,” *Econometrica*, 55, 647-661.
- [7] Baumeister, Roy F. (1998): “The Self,” in D. T. Gibling, S. T. Fiske, and G. Lindzey (eds) *The Handbook of Social Psychology, Vol. 1*, (Hoboken, NJ: John Wiley & Sons).
- [8] Berger, J. (2014): “Word of mouth and interpersonal communication: A review and directions for future research,” *Journal of Consumer Psychology*, 24 (4), 586–607.
- [9] Bloch, F., G. Demange and R. Kranton (2014): “Rumors and Social Networks,” *Paris School of Economics, Working Paper 2014-2015*.

- [10] Campbell, A. (2013): “Word-of-Mouth Communication and Percolation in Social Networks,” *The American Economic Review*, 103 (6), 2466-2498.
- [11] Campbell, A., Mayzlin, D. and J. Shin (2013): “Managing Buzz,” working paper.
- [12] Chatterjee, K. and B. Dutta (2015): “Credibility and Strategic Learning in Networks,” working paper.
- [13] Cho, I. and D. Kreps (1987): “Signaling Games and Stable Equilibria,” *The Quarterly Journal of Economics*, 102 (2), 179-221.
- [14] Crawford, V. and J. Sobel (1982): “Strategic Information Transmission,” *Econometrica*, 50, 1431-1451.
- [15] Farrell, J. (1985): “Credible Neologisms in Games of Communication,” *Massachusetts Institute of Technology, Working Paper No. 386*.
- [16] Farrell, J. (1993): “Meaning and Credibility in Cheap-Talk Games,” *Games and Economic Behavior*, 5, 514-531.
- [17] Farrell, J. and R. Gibbons (1989): “Cheap Talk with Two Audiences,” *The American Economic Review*, 79 (5), 1214-1223.
- [18] Farrell, J. and M. Rabin (1996): “Cheap Talk,” *Journal of Economic Perspectives*, 10, 103-118.
- [19] Galeotti, A., C. Ghiglini and F. Squintani (2013): “Strategic Information Transmission Networks,” *Journal of Economic Theory*, 148, 1751-1769.
- [20] Galeotti, A. and S. Goyal (2009): “Influencing the Influencers: a Theory of Strategic Diffusion,” *RAND Journal of Economics*, 40 (3), 509-532.
- [21] Goltsman, M. and G. Pavlov (2011): “How to Talk to Multiple Audiences,” *Games and Economic Behavior*, 72, 100-122.
- [22] Golosov, M., Skreta, V., Tsyvinski, A. and A. Wilson (2014): “Dynamic Strategic Information Transmission,” *Journal of Economic Theory*, 151, 304-341.
- [23] Hagenbach, J. and F. Koessler (2010): “Strategic Communication Networks,” *Review of Economic Studies*, 77, 1072-1099.

- [24] Johnson Brown, J. and P.H. Reingen (1987): “Social Ties and Word-of-Mouth Referral Behavior,” *Journal of Consumer Research*, 14 (3), 350-362.
- [25] Krishna, V. and J. Morgan (2004): “The Art of Conversation: Eliciting Information from Experts through Multi-Stage Communication,” *Journal of Economic Theory*, 117, 147-179.
- [26] Krishna, V. and J. Morgan (2008): “Cheap Talk,” in Blume, L. and S. Durlauf (eds) *The New Palgrave Dictionary of Economics*, 2nd edition (New York: McMillan).
- [27] Lagunoff, R. (2006): “Credible Communication in Dynastic Government,” *Journal of Public Economics*, 90, 59-86.
- [28] Mailath, G., Okuno-Fujiwara, M. and A. Postlewaite (1993): “Belief-Based Refinements in Signalling Games,” *Journal of Economic Theory*, 60, 241-276.
- [29] Matthews, S. A., M. Okuno-Fujiwara and A. Postlewaite (1991): “Refining Cheap-Talk Equilibria,” *Journal of Economic Theory*, 55 (2), 247-273.
- [30] Nyilasy, G. (2006): “Word of mouth: what we really know - and what we don’t,” in Kirby, J. and P. Marsden (eds) *Connected Marketing: The Viral, Buzz and Word of Mouth Revolution* (Oxford: Butterworth-Heinemann).
- [31] Perez-Richet, E. (2014): “Interim Bayesian Persuasion: First Steps,” *American Economic Review: Papers & Proceedings*, 104(5), 469–474.
- [32] Schmidt, W., and R.W. Buell (2014): “Decision Making Under Information Asymmetry: Experimental Evidence on Belief Refinements,” *Harvard Business School Working Paper*, No. 15-001, July 2014.
- [33] Sundaram, D.S., Mitra, K. and C. Webster (1998): “Word-Of-Mouth Communications: a Motivational Analysis,” *Advances in Consumer Research*, 25, 527-531.
- [34] Sobel, J. (2009): “Signaling Games,” in Myers, R. A. (ed.) *Encyclopedia of Complexity and System Science* (Berlin: Springer-Verlag).
- [35] Umbhauer, G. (1994): “Forward Induction, Consistency, Preplay Communication and Epsilon Perturbations,” unpublished.