University of Cyprus
Department of Economics

Working Paper 10-2015

Optimal Influence under Observational Learning

Nikolas Tsakas

Department of Economics, University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus
Tel.: +357-22893700, Fax: +357-22895028, Web site: http://www.ucy.ac.cy/econ/en

Optimal Influence under Observational Learning

Nikolas Tsakas*

Department of Economics, University of Cyprus

December 1, 2015

Abstract

We study the optimal targeting problem of a firm that seeks to maximize the diffusion of a product in a society where agents learn from their neighbors. The firm can seed the product to a subset of the population and our goal is to find which is the optimal subset to target. We provide a condition that characterizes the optimal targeting strategy for any network structure. The key parameter in this condition is the agents' decay centrality, which takes into account how close an agent is to others, in a way that distant agents are weighted less than closer ones.

JEL Classification: D83, D85, H23, M37

Keywords: Social Networks, Targeting, Diffusion, Observational Learning.

*Correspondence Address: University of Cyprus, Department of Economics, PO Box 20537, 1678 Nicosia, Cyprus. E–mail: tsakas.nikolaos@ucy.ac.cy

1. Introduction

The role of influential agents in the diffusion of products and ideas has been subject to extended research in several different fields (see for example Domingos and Richardson, 2001; Kempe et al., 2003, 2005; Galeotti and Goyal, 2009; Kirby and Marsden, 2006, and references therein). Recent technological advances have made possible the collection of previously unobservable data related to particular individual characteristics. Based on such observations, researchers, firms and other interested parties are able to identify the structure and the kind of social interactions between the members of a society.

Observing these interactions is particularly useful for firms that want to determine effective targeting strategies at the individual level, attempting to maximize the spread of their products in a market. Given that targeting agents is costly, the effective selection of targets becomes crucial. The targeting problem becomes even tougher because the firms may not be know ex–ante the relative quality of their product (compared to those of their competitors). This is a commonly observed issue, since it is normal that a firm can determine the level of satisfaction of the consumers only after the product has been circulated in the market. There are several other economic examples that motivate the study of this problem. Some of them are for instance the diffusion of agricultural technologies in rural areas, or the entry of a computer software or a mobile phone in a new market.

In this paper, we want to identify the optimal targeting strategy of a firm that is ex–ante uncertain about the relative quality of its product, while the agents learn the quality of a product as soon as they have observed someone consuming it. In an environment where agents choose repeatedly between two alternative products, if the product is of higher quality than its alternative, then it will get diffused to the whole population; hence the firm would like diffusion to occur as fast as possible. To the contrary, if the product is of lower quality, then it will survive only for a finite number of periods, and in particular until every agent has observed the alternative at least once. Therefore, the firm would like to protect the early consumers of the product from observing the alternative. It turns out that the most important parameter for identifying the optimal targets is agents' decay centrality, which is a measure of centrality that takes into account how close an agent is to others, but in a way that very distant agents are weighted less than closer ones (see Jackson, 2008).

To the best our knowledge, this is the first paper that characterizes completely the targeting strategy for a firm that is able to target multiple agents, allowing for any network structure and relative patience of the firm.

1.1. Setting and Results

We consider two interested parties, from now on called *the firms* A and B, who seek to maximize the diffusion of two competing actions, from now on called *products* A and B respectively, in a social network with finitely many agents. Both firms introduce their products to the market simultaneously using fixed advertising budgets. We focus on the case where only firm B allocates its budget strategically.

In particular, firm A employs a uniform advertising strategy that makes every agent aware of its product. One can think of this as advertising the product on television or social media etc. To the contrary, firm B advertises its product to a strategically targeted subset of the population. The aim of this paper is to characterise the optimal targeting strategy of firm B. Targeting intensity is assumed to be such that every agent who gets targeted by firm B is convinced to choose its product once, whereas the rest of the agents are convinced to choose product A once. In addition to this, both products are considered to be sufficiently useful to the agents, so that each of them consumes one unit of either of them in each period thereafter.

After the targeting phase, the agents choose repeatedly between products A and B and they observe the qualities of the products chosen by themselves and by their immediate neighbors. The agents choose the product that they have observed to be of higher quality –a comparison that is possible only after having observed both products being chosen within their neighborhood–. The quality of each product remains unchanged throughout the process; therefore the agents may revise their choice only the period after they have observed both products for the first time. They do not revise earlier because they are have not observed the quality of the alternative and they do not revise later because they already know which product is of higher quality.

The analysis covers the case where firm B is ex–ante uncertain about whether its product is the one of higher quality and has just some limited information that determines the probability that this is the case. Therefore, the diffusion of its product is uncertain and the targeting strategy has to be designed taking this factor of uncertainty into account. In practice, this information is related to the relative performance of the product and the level of satisfaction of the consumers, which cannot be known to the firm unless both product have already been circulated.

We provide a necessary and sufficient condition that characterizes the optimal targeting strategy for any network structure. More specifically, in the particular case that the firm can target only one agent, then it is optimal to target the agent with the maximum decay centrality. In the general case where the firm can target a larger subset of the population the optimal strategy combines two features: On the one hand, it should intend to maximize the decay centrality of the set of targeted agents, so as to capture the population quickly in case the product is of higher quality. On the other hand, it should intend to minimize the decay centrality of the targeted set's complement, so that the product survives as long as possible if it is of lower quality. If the firm is extremely impatient, then decay centrality coincides with degree centrality, whereas if the firm is extremely patient then decay centrality coincides with closeness.

Given that decay centrality is a measure that depends crucially on the exact network structure, we study separately the case of the circle as an attempt to provide a more intuitive picture of this measure. In this case, we identify the shape of the targeted set, which should consist of one large group of connected agents and several small ones spread uniformly around the network. The exact number of groups depends on the particular characteristics of each environment and for this reason we provide some partial results and numerical examples. In particular, if the firm is sufficiently optimistic about the product, then it prefers to spread the targeted agents as uniformly as possible around the society; whereas if the firm is sufficiently pessimistic, then it prefers to concentrate all the targeted agents together. This result is true for any level of patience and thus it is in partial contrast with settings where the agents cannot learn the true quality of the products (see Tsakas, 2014).

1.2. Related Literature

The problem of diffusion of innovations has been studied very extensively and from different perspectives (see Jackson, 2008; Goyal, 2007; Vega-Redondo, 2007; Peyton Young, 2009). This interest has led to an extended literature dealing with optimal influence in networks, which focuses on different network structures and behavioral rules (see Galeotti and Goyal, 2009; Chatterjee and Dutta, 2011; Goyal and Kearns, 2012; Tsakas, 2014; Campbell, 2013; Ortuño, 1993). The particular problem of optimal targeting has also been explored in marketing[1] and computer science research (see Kempe et al., 2003, 2005). A general finding is that computing a set of nodes to optimize diffusion is generally a hard problem, which is a potentially limiting hurdle in terms of implementing the methods outlined here.

A common denominator among most of the results is that optimal seeds are characterized by some form of centrality, defined in each context analogously. In what might be the most closely related paper (in terms of the setup) Banerjee et.al (2013) define the measures of "communication" and "diffusion" centrality, which are based on probabilistic transmission from one person to another. This

[1] There is a long literature starting from Bass (1969) focusing on diffusion in marketing and network targeting. For instance see also Bell and Song (2007).

implicitly defines a decay to the level of diffusion, given that the further a node is from the seed, the less likely is to be affected. In the current setup instead, transmission is perfect, and so the "decay" parameter actually arises from the discounting of profits rather than the probabilistic transmission. This difference provides a new perspective that is interesting for two reasons. First, because decay centrality, contrary to communication centrality (and other similar probabilistic measures) is based on shortest paths, rather than on multiple probabilistic paths. Therefore, it provides a new foundation for decay centrality relative to the previous discussion regarding its original definition (mainly an alternative to closeness centrality). Second, it is based on the discounted time to reach various nodes, rather than the expected reach of a diffusion process.

In broad terms, our paper studies a similar question to Galeotti and Goyal (2009) and Goyal and Kearns (2012). Galeotti and Goyal (2009) study the maximization problem of a firm that seeks the diffusion of its product in a society, under word–of–mouth communication and social conformism. Apart from the different structure, the analysis focus exclusively on the short–run strategies and on the degree distribution of the agents, assuming that agents meet randomly. To the contrary, we provide a general analysis for any network and different levels of firm's farsightedness. Similarly, Goyal and Kearns (2012) study competition between two firms that distribute resources in order to maximize the long–run diffusion of a product in a society. The authors focus more on the efficient allocation of resources and on the effect of budget asymmetries on the equilibrium allocations.[2]

In a more closely related paper, Chatterjee and Dutta (2011) analyze a similar question to ours as well, however in their framework the firm knows exactly the quality of the product and the agents are either completely naive, in the sense that they adopt the product as soon as they observe it, or fully rational bayesian maximizers. Nevertheless, in their paper one of the parameters that is crucial for the identification of optimal targeting strategies is the decay centrality of the agents, which turns out to be the case also in this paper. The most closely related paper is the one of Tsakas (2014), where it is analyzed a similar problem of targeting agents. The main difference is that uncertainty regarding the quality of the product persists forever, leading the agents to revise their choices repeatedly. This absence of learning leads to significant changes in the optimal targeting strategies.

In a more general framework, our work builds upon the study of learning from neighbors,(see Banerjee, 1992; Banerjee and Fudenberg, 2004; Bala and Goyal, 1998; Chatterjee and Xu, 2004; Ellison and Fudenberg, 1993, 1995; Gale and Kariv, 2003) where most of the papers focus mainly conditions under which efficient actions spread to the whole population and not on optimal influence strategies. Large part of this literature, including the present paper, assumes that agents adopt

[2] There is a growing literature on competitive strategic influence focusing on a number of different environments, see for instance Bimpikis et al. (2015); Chasparis and Shamma (2001); Grabisch et al. (2015); Fazeliy et al. (2015).

simple behavioral rules; an assumption that is also supported by a recent but growing empirical literature (see Apesteguia et al., 2007; Conley and Udry, 2010; Bigoni and Fort, 2013).

2. Model

2.1. The Network

There is a finite set of agents N, who are connected through a social network. The set has cardinality n, typical elements i and j and is mentioned as *population* of the network or *market*. A social network is represented by a family of sets $\mathcal{N} := \{N_i \subseteq N \mid i = 1, \dots, n\}$, with N_i denoting the set of agents who can observe and be observed by i. Throughout the paper N_i is called i's neighborhood and is assumed to contain i.[3] Its cardinality, $|N_i|$, is called i's degree and the agents who belong to N_i are called i's neighbors. In the present setting, the network structure describes the channels of communication in the population and does not impose strategic interactions. More specifically, each agent $i \in N$ observes the choices and their associated outcomes of all $j \in N_i$. We focus on undirected networks, where $j \in N_i$ if and only if $i \in N_j$, nevertheless with appropriate modifications the results for directed networks would look very similar.

A path between agents i and j is a sequence $i_1, ..., i_K$ such that $i_1 = i$, $i_K = j$ and $i_{k+1} \in N_{i_k}$ for $k = 1, ..., K-1$. The geodesic distance, $d(i, j)$, between two agents is the length of the shortest path between them. More generally, the geodesic distance, $d(T, j)$, between a set of agents T and an agent outside of the set, $j \in T^C := N \setminus T$, is defined as the minimum among all the distances between some agent $i \in T$ and agent $j \in T^C$, i.e. $d(T, j) := \min_{i \in T} d(i, j)$. [4]

We say that two agents are connected if there exists a path between them. The network is connected if every pair of agents is connected. We focus on connected networks, nevertheless for disconnected networks the analysis would be identical for each of their connected components.[5]

2.2. The Problem

There are two firms that produce products A and B. With some abuse of notation we use the same name to refer to the firms and to the products that they produce. The two firms introduce

[3]This assumption is not usual, however in this setting it is necessary since N_i describes the set of agents whose choices can be observed by i. Therefore, since it is reasonable to assume that one can observe her own choices, she should be contained in her own neighborhood.

[4]Throughout the paper, T^C denotes complement of T.

[5]A connected component is a non-empty sub-network N' such that (i) $N' \subset N$, (ii) N' is connected and (iii) if $i, j \in N'$ and $j \in N_i$, then $j \in N_i'$.

their products simultaneously to a new market using fixed advertising budgets that are determined exogenously. We focus on the case where only firm B allocates this budget strategically. For this reason, throughout the paper unless stated explicitly, whenever we mention "the firm" we refer to firm B.

In particular, firm A allocates its budget uniformly to the whole population, thus making every agent aware of its product. An example of such strategy could be a firm that advertises the product on television or social media etc. To the contrary, firm B targets a subset of the population to advertise the product to, i.e. $T \subset N$ with cardinality $t \leq n$ which is determined exogenously. We assume that targeting intensity is such that every agent who gets targeted by firm B is convinced to choose its product once, whereas the rest of the agents are convinced to choose product A once. Moreover, both products are of sufficiently high quality, so that each agent consumes one unit of either of them in each period thereafter.

More specifically, targeting occurs at $\tau = -1$, at $\tau = 0$ each agent consumes product B if she has been targeted and product A otherwise. After that, in each period $\tau = 1, \ldots,$ each agent consumes one unit of either product A or B, whichever of the two she has observed to be of higher quality. The relative qualities of the two products are ex–ante unknown to the firms, therefore firm B has to condition its targeting strategy on this factor of uncertainty. Uncertainty is summarized in one parameter that determines the probability of product B being of higher quality than product A, i.e. $p := \mathbb{P}[q_B - q_A \equiv \Delta q > 0] \in [0, 1]$. Our results cover also the limiting cases where there is no uncertainty regarding the relative qualities.

The objective of firm B is to target the subset of the population that will maximize the discounted expected total consumption, from now on called *sales*, of its product. We refer to this as *optimal targeting strategy* of the firm. Formally, let $\mathcal{T}$ be the collection of subsets of N with cardinality t and $S_\tau(T|\Delta q)$ be the firm's sales function, in period τ, for a targeting strategy $T \in \mathcal{T}$ and given the relative quality Δq. Therefore, the firm seeks the targeting strategy $T \in \mathcal{T}$ that maximizes the following payoff function:

$$\Pi(T) := E_{\Delta q}\left[\sum_{\tau=1}^{+\infty} \delta^{\tau-1} S_\tau(T)\right] = p \sum_{\tau=1}^{+\infty} \delta^{\tau-1} S_\tau(T|\Delta q > 0) + (1-p) \sum_{\tau=1}^{+\infty} \delta^{\tau-1} S_\tau(T|\Delta q < 0)$$

We intentionally refrain from introducing prices in the model, since the focus of the paper is on studying optimal strategies related to the position of agents within a social network. Needless to say that it would be a very interesting extension to see the extent to which pricing strategies could be combined with targeting in the development of a complete advertising strategy.

We still need to define how agents choose which product to consume from period $\tau = 1$ onwards. First, it is important to note that despite the fact that the agents are initially aware of either only product A or both products A and B they do not know their qualities. This is a reasonable assumption since advertising focuses on convincing a consumer to purchase a product, not necessarily explaining in depth its actual features. This is where the role of the social network comes into play. Each period, the agents observe the products chosen by their neighbors (and themselves of course), as well as their qualities. Therefore, upon observing both products, they are able to compare them and choose the one they have observed to be of higher quality.

Notice that, as long as no one in their neighbourhood has ever chosen the alternative product they stick to their initial choice. The first time they observe both products they decide whether to switch or not and they make the same choice from the next period onwards. Hence, in this setting, despite the fact that they are allowed to switch as many times as they want, they end up switching at most once.

One might argue how is it possible that the consumers are able to identify the relative qualities immediately upon observing both products, while the firms could not do so before spreading them to the market. There is a number of reasons why this might be the case. For example, it is a common practice for firms to refrain from announcing all the features of their new products before circulating them, exactly in order to hide crucial information from their competitors. Furthermore, for particular products relative quality might be related with compatibility issues, which cannot be easily determined before the circulation of both products. For instance, consider two competing computer softwares that it turns out that one is able to run files created in both of them, while the other one is not. Once again, in this case a firm might not be able to have this information beforehand.

These arguments might raise incentives to the firms to wait for the entrance of their competitor before entering the market. Nevertheless, such incentives would be similar for both firms, leading them to take a strategic decision of optimal entry time. This is another interesting extension, that lies outside the scope of the paper. In any case, our analysis is also able to capture the extreme cases where the firms know exactly the relative qualities of their products and there is no uncertainty during targeting.

As it becomes apparent, there are two opposite forces that will be driving the results. On the one hand, if product B is of higher quality then it will be diffused with certainty to the whole population after a finite number of periods. Hence, given the presence of discounting, the firm would like to capture the whole population as quickly as possible. On the other hand, if its quality is lower, then it will unavoidably disappear after finitely many periods. In this case, the firm would want to make the product survive for as many periods as possible.

3. Results: Targeting only one agent

First, we study the case where firm B can target only one agent. In this case, it is apparent that if the product is of lower quality then the payoff of the firm is zero, irrespectively of the targeting strategy. Therefore, the firm focuses on maximizing the discounted sum of sales in case the product is of higher quality. The parameter that determines the optimal targeting strategy is the decay centrality of the agents, which is defined as follows (Jackson, 2008):

Definition 1. Consider a decay parameter $\delta \in (0,1)$. The decay centrality of an agent $i \in N$ is defined as $\sum\limits_{j \neq i} \delta^{d(i,j)}$.

In the current setting the decay parameter δ coincides with the discount factor of the firm and $d(i,j)$ is the geodesic distance between agents i and j. It is the same kind of centrality that determines the optimal targeting strategy in Chatterjee and Dutta (2011), despite the fact that the settings differ significantly. Intuitively, this is a measure of centrality that takes into account how close an agent is to others, but in a way that very distant agents are weighted less than closer ones. In a sense, it captures how quickly a product of high quality can become visible to the whole population. Agents with high decay centrality are expected to be those who combine a higher number of neighbors with a small geodesic distance from the most isolated agents in the network. The following theorem characterizes the optimal strategy for a firm that targets only one agent ($t = 1$).

Theorem 1. *For all $(\delta, p) \in (0,1)^2$, the optimal strategy for the firm is to target the agent $i \in N$ with the maximum decay centrality.*

All proofs can be found in the Appendix.

Notice that the exact optimal strategy depends on the structure of the network and it cannot be easily identified for general networks. As it is also mentioned by Chatterjee and Dutta (2011), this is easily identified in simple structures such as the star, where the center is the optimal target, and the line where the median is the one maximizing decay centrality.

An obviously crucial parameter is the value of the discount factor δ. Intuitively, the lower the value of δ the more important becomes the degree of the targeted agent. If δ gets very close to 0, then the firm cares only about the payoff of the first period, which comes from the sales to agents who have geodesic distance equal to one from the targeted agent. Hence, it is optimal to target the agent with the maximum degree. This result is formally stated in the following corollary, whose proof is omitted, since it follows directly from the proof of the theorem.

Corollary 1. *If $\delta = 0$, then the optimal strategy for the firm is to target the agent $i \in N$ with the maximum degree, $|N_i|$.*

Notice that, this is also true for positive, but sufficiently small values of δ.

In the other extreme case where $\delta = 1$ the result is not that obvious. First, observe that the payoffs of the firm become infinitely large irrespectively of the targeting strategy. Nevertheless, we can determine which will the optimal targeting strategy by considering the differences in payoffs between alternative strategies, despite the fact that these differences will only be finitely large. In order to obtain a formal result we should slightly modify the original objective function of the firm in the following way: Consider that the firm cares equally about the expected sales of the next P periods, i.e. $\Pi(T) = E_{\Delta q}\left[\sum_{\tau=1}^{P} S_\tau(T)\right]$, P can be any finite number, as long as it is sufficiently large to ensure total diffusion for some targeting strategy.[6] As long as P is sufficiently large the difference in payoffs generated by different targeting strategies remains unchanged and for P being finite this difference is not negligible. It turns out that in this case the optimal strategy for the firm is to target the agent with the maximum closeness (or equivalently the minimum farness), which are both formally defined below.

Definition 2. The farness of an agent $i \in N$ is defined as the sum of the geodesic distances from each other agent in the network, i.e. $\sum_{j \neq i} d(i,j)$, and its closeness is defined as the inverse of the farness.

Proposition 1. *If $\delta = 1$, then in the modified problem the optimal strategy for the firm is to target the agent with the maximum closeness (or equivalently with the minimum farness).*

In the following section, the introduced concepts are extended to cases where the firm can target several agents, rather than a single one. As we will see below, the results are slightly modified, but the intuition remains similar. An important feature is that the agents who belong to the optimal targeted set might be different from those who would be optimal as individual targets.

4. Results: Targeting several agents

The problem becomes more complicated if the firm targets several agents. In this case, we can no longer ignore the payoffs of the firm when the product is of lower quality. This happens because when the firm can target many agents, it can distribute them in such a way that some of them do not observe the quality of the alternative product immediately, because none of their neighbors uses it initially. Before proceeding to the formal statement it is necessary to slightly modify the previous

[6]For this to happen it suffices that P is larger than the value of the largest geodesic distance between any two agents in the network, which is usually called diameter of the network.

definitions, so as to be adequate for groups of multiple agents. The generalized definitions are based on Everett and Borgatti (1999). Recall that, the geodesic distance between a group of agents T and an agent $j \notin T$ is defined as the minimum distance between some agent $i \in T$ and agent j. Namely, $d(T, j) := \min_{i \in T} d(i, j)$. Also, $\mathcal{T}$ denotes the collection of subsets of N with cardinality t and $T^C = N \setminus T$ is the complement of T.

Definition 3. The *group degree* of a set $T \in \mathcal{T}$, denoted by $|N_T|$, is defined as the number of agents not belonging to T who have at least one neighbor in T, or equivalently the cardinality of the union of neighborhoods of all agents in T, i.e. $|N_T| := \left| \bigcup_{i \in T} N_i \right|$.

The importance of an agent's degree lies in the fact that many others observe its actions. When targeting multiple agents, what is important is the total number of agents that can observe at least one of the targeted agents. Therefore, the parameter of interest is not the sum of neighbors of all targeted agents per se, but the sum of different agents who are directly connected with some agent in the targeted set.

Definition 4. The *group farness* of a set $T \in \mathcal{T}$ is defined as the sum of the distances between T and each other agent outside T, i.e. $\sum_{j \in T^C} d(T, j)$. Its *group closeness* is defined as the inverse of the group farness.

The definition is totally analogous to the case of one agent, taking into account that the focus now is on the necessary steps to reach an agent outside the set starting from any agent in the set.

Definition 5. Consider a decay parameter $\delta \in (0, 1)$. The *group decay centrality* of set $T \in \mathcal{T}$ is defined as $\sum_{j \in T^C} \delta^{d(T,j)}$.

This definition is a generalization of decay centrality. Given that the focus is on sets of multiple agents, the important parameters are the lengths of shortest paths that start from some agent inside the set and end at any agent outside it. Intuitively, this means that the aim of a targeting strategy is to affect the number of periods needed before every agent learns the quality of both products. Ideally, this number should be minimized in case the product is of higher quality and be maximized otherwise.

Theorem 2. *For $(\delta, p) \in (0, 1)^2$, among all $T \in \mathcal{T}$ (subsets of N with t agents) the optimal strategy is to target the set that maximizes the following expression:*

$$p \sum_{j \in T^C} \delta^{d(T,j)} - (1-p) \sum_{i \in T} \delta^{d(T^C,i)}$$

The above expression is very intuitive and the idea is similar to the case with only one targeted agent. In words, the firm seeks to target the set that maximizes a linear combination of two opposite features: On the one hand, the firm wants to maximize the group decay centrality of the targeted set, so as to capture the whole population as quickly as possible in case it is of higher quality. On the other hand, it wants to minimize the group decay centrality of the targeted set's complement, so that the product will survive for as many periods as possible in case it is of lower quality. The level of importance of each one of the two factors is determined by the probability p of the product being of higher quality. If the product is very likely to be better, then the firm cares almost exclusively about minimizing the number of periods needed for the product to be observed by everyone in the network. Whereas, if the product is very likely to be bad, then the firm cares about protecting the targeted agents from observing the alternative product for as many periods as possible. The following corollaries summarize the results for the extreme values of p and δ. The proofs are omitted since they follow directly from the objective function of the firm and the proof of Theorem 2. The trade–off will become even clearer in the following subsection that we analyze the particular case of the circle network.

Corollary 2. *For $p = 1$, among all $T \in \mathcal{T}$ the optimal strategy for the firm is to target the set:*

- *with the maximum group degree, if $\delta = 0$.*
- *with the maximum closeness, if $\delta = 1$.*
- *with the maximum decay centrality, if $\delta \in (0, 1)$.*

The above corollary would describe the optimal targeting strategy if we were assuming the agents to learn the quality of product A in the targeting period; since in that case if product B was of lower quality they would all switch to product A from period $\tau = 1$ onwards. Once again, we observe that this captures only a particular case of our setting.

Corollary 3. *For $p = 0$, among all $T \in \mathcal{T}$ the optimal strategy for the firm is to target the set:*

- *whose complement, T^C, has the minimum group degree, if $\delta = 0$.*
- *whose complement, T^C, has the minimum closeness, if $\delta = 1$.*
- *whose complement, T^C, has the minimum decay centrality, if $\delta \in (0, 1)$.*

Corollary 4. *For $p \in (0, 1)$, among all $T \in \mathcal{T}$ the optimal strategy for the firm is to target the set:*

- *that maximizes $p\,|N_T| - (1-p)\,|N_{T^C}|$, if $\delta = 0$.*
- *that minimizes $p \sum_{j \in T^C} d(T, j) - (1-p) \sum_{i \in T} d(T^C, i)$, if $\delta = 1$.*

5. Example: The Circle

It has already been mentioned that decay centrality depends much on the exact topology of the network. In this section, we intend to provide an intuitive image of the set that maximizes decay centrality in a simple network structure; the circle.[7] The fact that all the agents not only have the same degree, but they are also identical in any measure of centrality provides additional interest to the results regarding targeting of several agents.[8]

Before proceeding to the results we need to introduce some more notation. If firm B targets t agents, then initially the population consists of s agents choosing product A and t agents choosing product B, where $s+t=n$. We call *group* a sequence of neighboring agents who all choose the same product and are surrounded by agents choosing the alternative product. The population is formed of g groups of neighboring agents who choose product A, with sizes $\{s_1, s_2, \dots, s_g\}$, where $\sum_{k=1}^{g} s_k = s$ and analogously g groups of neighboring agents who all choose product B, with sizes $\{t_1, t_2, \dots, t_g\}$, where $\sum_{k=1}^{g} t_k = t$.[9] We mention these groups as being *of type* A and *of type* B respectively. The numbering of the groups is based on their size in increasing order, $s_1 \leq s_2 \leq \dots \leq s_g$ and $t_1 \leq t_2 \leq \dots \leq t_g$. With some abuse of notation we also use $s_1, s_2, \dots, s_g$ and $t_1, t_2, \dots, t_g$ to name the groups. Our goal is to find the optimal size of all s_k and t_k for $k = 1, \dots, g$, their optimal position (if it matters), as well as the optimal number of groups, g.

In order to avoid unnecessary complications in the calculations (which arise without the gain of any additional intuition) we assume that every group must have an *even* number of agents. Formally:

Assumption 1. [A1] s_i and t_i are even numbers for all $i \in \{1, \dots, g\}$. $\triangleleft$

First of all, one should notice that, for a given number of groups, the sizes of groups $t_1, \dots, t_g$ are important only when the product is of lower quality, whereas those of $s_1, \dots, s_g$ only when the product is of higher quality. This is a direct consequence of Theorem 2 and the fact that all the agents are positionally identical.

Proposition 2. *For a given number of targeted groups, denoted by g, the optimal configuration satisfies the following conditions:*

1. $s_g - s_1 \leq 2$

[7]In the circle network, each agent has exactly two immediate neighbors; Namely $N_i = \{i-1, i, i+1\}$ for $i = 2, \dots, n-1$, whereas $N_1 = \{n, 1, 2\}$ and $N_n = \{n-1, n, 1\}$.

[8]It is apparent that the decision is trivial when the firm can target only one agent.

[9]Notice that, the fact that the network is a circle and there exist exactly two products ensures that the number of groups is the same for both products.

2. $t_1 = \cdots = t_{g-1} = 2$ *and* $t_g = t - 2(g-1)$

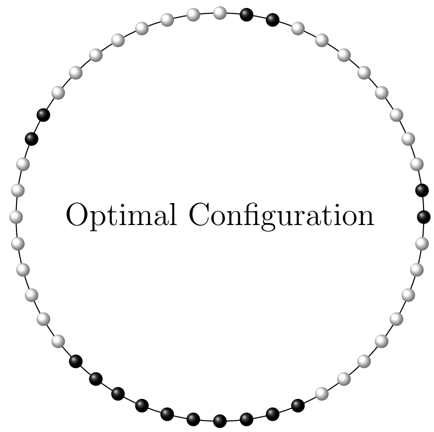

Figure 1: The optimal targeting strategy if the firm targets four groups.

The first condition captures the fact that the firm wants to maximize the decay centrality of the targeted set if the product is of higher quality. This is equivalent saying that the firm wants to make the product visible to everyone in the population as quickly as possible. Therefore, it locates the targeted agents in a way such that all groups of type B have almost the same size.[10]

The second condition captures the fact that the firm wants to minimize the decay centrality of the complement of the targeted set if the product is of lower quality. That is equivalent saying that it wants to maximize the number of periods until all the targeted agents have observed the alternative product. Figure 1 depicts a typical configuration when the number of targeted groups is fixed to four.

However, the above result assumes the number of groups to be fixed, which need not be the case. It is still necessary to compare configurations that satisfy the above conditions and consist of different number of groups. The following two propositions, accompanied by Figures 2 and 3 intend to clarify this comparison.

Proposition 3. *Assume that the firm can only target either one or two groups and that s is a multiple of 4.*[11] *Then for any $(p,\delta) \in (0,1)^2$ it is optimal for the firm to target two groups if and only if* $p \geq \frac{(1-\delta)\left(1-\delta^{\frac{t}{2}-1}\right)}{(1-\delta)\left(1-\delta^{\frac{t}{2}-1}\right)+\left(1-\delta^{\frac{s}{4}}\right)^2}$.

Proposition 3 states that for any level of patience a sufficiently optimistic firm (with p high enough) would prefer to spread the targeted agents, whereas a sufficiently pessimistic firm (with p

[10]Ideally, they should all be equally sized, but this may not be possible because of divisibility issues.

[11]The assumption regarding s is introduced only in order to solve divisibility issues that make the form of the bound less intuitive. Nevertheless, the results would be very similar if one drops this assumption.

low enough) would prefer to concentrate them altogether. This result is very intuitive, given that an optimistic firm cares more about capturing the whole population faster in case the product is of higher quality, rather than making the product survive for more periods in case it is of lower quality.

Lemma 1. *The function* $f(\delta, t) = \frac{(1-\delta)\left(1-\delta^{\frac{t}{2}-1}\right)}{(1-\delta)\left(1-\delta^{\frac{t}{2}-1}\right)+\left(1-\delta^{\frac{s}{4}}\right)^2}$ *:*

- *is increasing in* t.
- *if* $t \leq n/3$, *is decreasing in* δ.
- *if* $t \geq n/2$, *is initially decreasing and then increasing in* δ.

Lemma 1 provides an intuition about how the incentives of the firm change with different parameters. First, the more agents the firm can target, the larger becomes the range of probabilities for which it prefers to target one group. Targeting four additional agents in the same group makes diffusion occur two periods faster in case the product is of higher quality; whereas doing so when the agents belong to two different groups makes diffusion occur only one period faster. Therefore, the marginal benefit from targeting two groups becomes smaller, as t becomes larger.

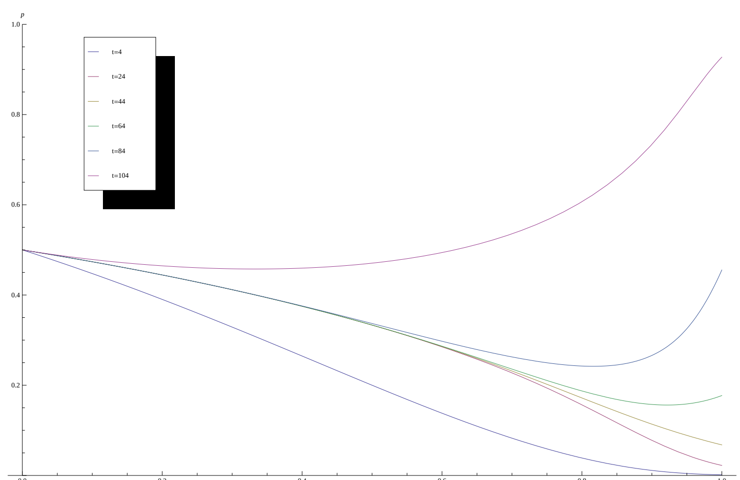

Figure 2: Choice between either one or two groups for different values of t, with $n = 112$. Given t, the firm prefers two groups over one for all p above the curve.

Before explaining the other two parts of the lemma, one should observe that for very high (low) values of p having a product of higher (lower) quality is much more likely, hence it is always better to target two groups (one group) irrespectively of the discount factor. For intermediate probability values, if t is sufficiently small the product survives only for a few periods in case it is of lower quality (even if the firm targets a single group), hence as the firm becomes more patient the potential payoff in case it is of higher quality becomes dominant for an increasing range of probabilities. The incentives are not so clear when t becomes sufficiently large, since the additional period that the product can survive if the firm targets only one group might become the dominant factor for characterizing the optimal targeting strategy.

Proposition 4. *Assume that the firm can target any number of groups g and that s is a multiple of $g!$,*[12] *then the optimal strategy for the firm is to target:*

- *One group if $p \leq \min\limits_{g \geq 1} f(\delta, t, g)$*
- *The maximum number of groups if $p \geq \max\limits_{g \geq 1} f(\delta, t, g)$*

where is defined as $f(\delta, t, g) = \frac{(1-\delta)\left(1-\delta^{\frac{t}{2}-g}\right)}{(1-\delta)\left(1-\delta^{\frac{t}{2}-g}\right)+\left[1-(g+1)\delta^{\frac{s}{2(g+1)}}+g\delta^{\frac{s}{2g}}\right]}$

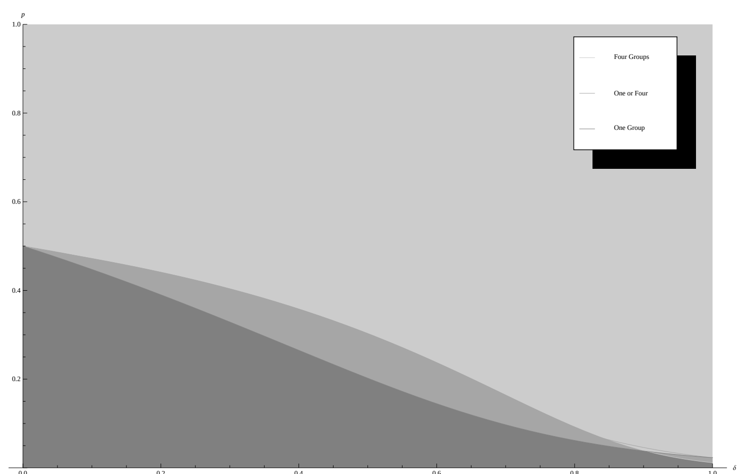

Figure 3: Choice of number of groups, given $t = 8$ and $n = 72$: (i) One group in the dark grey area, (ii) Max number of groups in the light grey area, (iii) compare the two in the intermediate area.

This result is in total analogy with the proposition stated for two groups. It says that for sufficiently high values of p the firm always prefers to target the maximum number of groups, whereas for low values of p it always prefers to target a single group. The explanation is rather simple; If the product is very likely to be of higher quality, then the firm cares almost exclusively about capturing the population quickly, hence it finds optimal to target the maximum number of groups. To the contrary, if the product is more likely to be of lower quality, then the firm cares about protecting

[12]The factorial of a number g, is defined as $g! = 1 \cdot 2 \cdot \cdots \cdot g$.

the product from disappearance for as many periods as possible, hence it finds optimal to target a single group.

The above results show an interesting contrast with Tsakas (2014), where the patience of the firm could affect the optimal targeting strategy, given the values of p. In the current setting, we observe that p is the predominant parameter that determines the optimal number of groups to target. High values of p are always associated with higher number of groups. Furthermore, for certain values of t, higher patience is also associated with higher number groups for a larger range of probabilities, which is again in contrast to the results of Tsakas (2014).

The main reason behind these differences is that in the current setting uncertainty is resolved after the first period, which makes it easier for the firm to predict the possible histories and condition its targeting strategy on whichever of them is more likely. To the contrary, in Tsakas (2014) uncertainty is never resolved, which makes the set of possible histories much larger and the choice of the firm much more complicated and uncertain. Yet, both environments are able to cover different aspects of optimal influence under uncertainty, which combined can provide a better understanding that could be used in particular cases.

6. Conclusion

In this paper, we have analyzed the optimal strategy of a firm who seeks to maximize the diffusion of a product of uncertain quality in a society. We find that the crucial parameter in the current setup is the agents' decay centrality, which takes into account how close an agent is to other agents, but in a way that very distant agents are weighted less than closer ones.

An important aspect regarding the measure of decay centrality is that it depends vastly on the exact topology of the network. Until now, the question of identifying agents who maximize decay centrality in a network has been answered only in very simple network structures. The problem becomes even harder if we pass from individual decay centrality to group decay centrality. The systematic study of this problem can shed light to a number of environments where decay centrality seems to be crucial. Moreover, in the current paper we have focused on the case where only one of the competing firms targets agents strategically. The analysis of competition in influence strategies would be a natural extension of this paper.

Acknowledgements

I am indebted to Antonio Cabrales for his support and to the seminar participants at Universidad Carlos III de Madrid, University of Cyprus, University of Cologne, GSE Summer Forum, Workshop on Networks in Economics and Finance and RES Meeting for fruitful comments. Part of this project was carried out while the author was at SUTD–MIT International Design Center at Singapore University of Technology and Design supported by grant IDG31300110. Financial help of the Spanish Ministry of Science and Innovation under grant BES-2012-054732 is also greatfully acknowledged. All remaining errors are my sole responsibility.

References

Apesteguia, J., Huck, S. & Oechssler, J. (2007). Imitation – Theory and experimental evidence. *Journal of Economic Theory* 136, 217–235.

Bala, V. & Goyal, S. (1998). Learning from Neighbours. *Review of Economic Studies* 65, 595–621.

Banerjee, A. (1992). A Simple Model of Herd Behavior. *Quarterly Journal of Economics* 107, 797–817.

Banerjee, A., Chandrasekhar, A.G., Duflo, E. & Jackson, M. (2013). The Diffusion of Microfinance. *Science* 341 (6144).

Banerjee, A. & Fudenberg, D. (2004). Word-of-mouth learning. *Games and Economic Behavior* 46, 1–22.

Bass, F. (1969). A new product growth for model consumer durables. *Management Science* 15(5), 215–227.

Bell, D. & Song., S. (2007). Neighborhood Effects and Trial on the Internet: Evidence from Online Grocery Retailing. *Quantitative Marketing and Economics* 5(4), 361–400.

Bigoni, M. & Fort, M. (2013). Information and Learning in Oligopoly. *IZA Discussion Paper No. 7125.*

Bimpikis, K., Ozdaglar, A. & Yildiz, I. (2015). Competing over Networks, mimeo.

Campbell, A. (2013). Word-of-Mouth Communication and Percolation in Social Networks. *American Economic Review* 103(6), 2466–2498.

Chasparis, G. and Shamma, J. (2010). Control of preferences in social networks. In *49th IEEE Conference on Decision and Control (CDC)*, 6651–6656.

Chatterjee, K. & Dutta, B. (2011). Credibility and Strategic Learning in Networks. *Warwick Economic Research Papers* no.972.

Chatterjee, K. & Xu, S.H. (2004). Technology diffusion by learning from neighbors. *Advances in Applied Probability* 36(2), 355–376.

Conley, T. & Udry, C. (2010). Learning About a New Technology: Pineapple in Ghana. *American Economic Review* 100, 35–69.

Descartes, René (1886). *La Géométrie.*

Domingos, P. and Richardson, M. (2001). Mining the network value of customers. In *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 57–66.

Ellison, G. & Fudenberg, D. (1993). Rules of Thumb for Social Learning. *Journal of Political Economy* 101, 612–644.

——— (1995). Word-of-mouth communication and social learning. *Quarterly Journal of Economics* 109, 93–125.

Everett, M.G. & Borgatti, S.P. (1999). The centrality of groups and classes. *The Journal of Mathematical Sociology* 23(3), 181–201.

Fazeliy, A., Aiorlou, A. & Jadbabaie, A. (2015). Competitive Diffusion in Social Networks: Quality or Seeding?, mimeo.

Gale, D. & Kariv, S. (2003) Bayesian Learning in Social Networks. *Games and Economic Behavior* 45, 329–346.

Galeotti, A. & Goyal, S. (2009). Influencing the influencers: a theory of strategic diffusion. *RAND Journal of Economics* 40(3), 509–532.

Goyal, S. (2007). *Connections: An Introduction to the Economics of Networks.* Princeton University Press, Princeton, NJ.

Goyal, S. & Kearns, M. (2012). Competitive contagion in networks. In *Proceedings of the 44th symposium on Theory of Computing* 759–774, New York, NY.

Grabisch, A., Mandel, A., Rusinowska, A. & Tanimura, E. (2015). Strategic influence in social networks, CES Working Paper.

Kempe, D., Kleinberg, J. and Tardos, E. (2003). Maximizing the Spread of Influence in a Social Network. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 137–146, New York, NY.

Kempe, D., Kleinberg, J. and Tardos, E. (2005). Influential nodes in a diffusion model for social networks. In *ICALP 2005: Proceedings of the 32nd International Colloquium on Automata, Languages and Programming*, 1127–1138.

Kirby, J. & Marsden, P. (2006). *Connected marketing*. Elsevier.

Jackson, M.O. (2008). *Social and Economic Networks*. Princeton University Press, Princeton, NJ.

Ortuño, I. (1993), Multiple equilibria, local externalities and complexity, mimeo.

Peyton Young, H. (2009). Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence and Social Learning. *American Economic Review*, 99(5), 1899–1924.

Tsakas, N. (2013). Imitating the most successful neighbor in social networks. *Review of Network Economics*, 12(4), 403–435.

Tsakas, N. (2014). Diffusion by imitation: the importance of targeting agents. *mimeo*.

Vega-Redondo (2007). *Complex Social Networks*. Cambridge University Press, Cambridge.

A. Mathematical Appendix

Proof of Theorem 1. It is obvious that if the product is of lower quality the sales will be zero irrespectively of the firm's targeting strategy. Therefore, we focus on the case that the product is of higher quality. Observing this, the firm aims to identify $T \in \mathcal{T}$ that maximizes $\sum_{\tau=1}^{+\infty} \delta^{\tau-1} S_\tau(T|\Delta q > 0)$, where in this case $\mathcal{T}$ consists of all the sets that contain only one agent, since the firm targets only one agent. Nevertheless, the objective of the firm can be restated slightly differently. Namely, instead of calculating the number of sales per period, equivalently one could calculate the total sales per agent of the population. This approach is much simpler, because we know that as soon as an agent observes the product of higher quality, she consumes it ever after. Therefore, in order to find the payoff of the firm from each agent, j, of the population is enough to find the number of periods that will pass before agent j observes the product. If the firm targets agent i, this number will coincide with the geodesic distance $d(i,j)$ from i to j. In particular, the payoff of the firm that targets agent i that comes from selling to agent j will be equal to $\sum_{\tau=d(i,j)}^{+\infty} \delta^{\tau-1} = \delta^{d(i,j)-1} \sum_{\tau=0}^{+\infty} \delta^{\tau} = \frac{1}{1-\delta}\delta^{d(i,j)-1}$. The total payoff of the firm that targets agent i (in case the product is of higher quality) will be equal to the sum of the previously calculated payoffs, for all $j \neq i$, that is $\frac{1}{\delta(1-\delta)} \sum_{j \neq i} \delta^{d(i,j)}$. Hence, the optimal strategy for the firm is to target agent $i \in N$ so as to satisfy that $\sum_{j \neq i} \delta^{d(i,j)} \geq \sum_{j \neq k} \delta^{d(k,j)}$, for all $k \in N$. By Definition 1, this is equivalent saying that the optimal strategy for the firm is to target the agent $i \in N$ with the maximum decay centrality. $\square$

Proof of Proposition 1. Similarly to the previous result, we restate the payoff of the firm, as being the sum of payoffs from each agent. If $\delta = 1$ and the product is of higher quality (once again the payoff is 0 if the product is of lower quality), then the payoff from some agent $j \in N$, when the firm targets agent i, is equal to 1 in all P periods except of the $d(i,j)$ first ones, because she has not observed the product yet. Recall that, in the modified version of the problem, P can be any finite number, as long as it is sufficiently large to ensure total diffusion for any targeting strategy. Therefore, the total payoff of the firm is equal to $\Pi(T = \{i\}) = \sum_{j \in N} [P - d(i,j)] = nP - \sum_{j \neq i} d(i,j)$. Observe that the first part of the right hand side does not depend on the targeting strategy. The optimal target, $i \in N$, needs to satisfy that $\Pi(T = \{i\}) \geq \Pi(T = \{k\})$ for all $k \in N$. This is equivalent saying that $\sum_{j \neq i} d(i,j) \leq \sum_{j \neq k} d(k,j)$ for all $k \in N$. By Definition 2 this means that the optimal strategy for the firm is to target the agent $i \in N$ with the maximum closeness (or equivalently, with the minimum farness). $\square$

Proof of Theorem 2. The proof is similar to that of Theorem 1. We need to determine the payoffs of the firm in case of the product being of higher or lower quality.

First, if the product is of higher quality then the payoff of the firm is the same as in the case where it could target only one agent, as long as we substitute the geodesic distance of each agent in the network from the targeted agent with its distance from the targeted set, $d(T,j)$. This happens because if the product is of higher quality then the number of periods that is needed before an agent becomes aware of it is equal to the minimum distance between that agent and someone who has been targeted. By definition, this is equal to its distance from the targeted set. Therefore, in that case the payoff of the firm would be equal to $\frac{1}{\delta(1-\delta)} \sum_{j \in T^C} \delta^{d(T,j)}$.

To the contrary, if the product is of lower quality then only the targeted agents will keep consuming the product, and this until they observe the alternative product. The number of periods needed before an agent $i \in T$ observed the alternative product is equal to its distance from the complement of the targeted set, denoted by T^C. Hence, the payoff of the firm from this agent will be equal to $\sum_{\tau=1}^{d(T^C,i)-1} \delta^{\tau-1} = \frac{1}{1-\delta}[1-\delta^{d(T^C,i)-1}] = \frac{1}{1-\delta} + \frac{1}{\delta(1-\delta)}\delta^{d(T^C,i)}$. The total payoff of the firm will be the sum of these payoffs for all agents $i \in T$ and will be equal to $\sum_{i\in T}[\frac{1}{1-\delta} - \frac{1}{\delta(1-\delta)}\delta^{d(T^C,i)}] = \frac{1}{1-\delta}t - \frac{1}{\delta(1-\delta)}\sum_{i\in T}\delta^{d(T^C,i)}$, where t is the number of targeted agents.

Therefore, the expected payoff for the firm will be equal to the following expression:

$$\Pi(T) = p\left[\frac{1}{\delta(1-\delta)}\sum_{j\in T^C}\delta^{d(T,j)}\right] + (1-p)\left[\frac{1}{1-\delta}t - \frac{1}{\delta(1-\delta)}\sum_{i\in T}\delta^{d(T^C,i)}\right] = \tag{1}$$

$$= \frac{1}{\delta(1-\delta)}\left[p\sum_{j\in T^C}\delta^{d(T,j)} - (1-p)\sum_{i\in T}\delta^{d(T^C,i)}\right] + (1-p)\frac{1}{1-\delta}t \tag{2}$$

As it becomes apparent the last part of the expression depends only on the size of the targeted set, therefore the optimal strategy for the firm is to target the set that maximizes the expression: $p\sum_{j\in T^C}\delta^{d(T,j)} - (1-p)\sum_{i\in T}\delta^{d(T^C,i)}$. □

Proof of Proposition 2. We first construct the payoffs from two arbitrary groups s_j and t_j of type A and type B respectively. A group s_j yields no payoff in case the product is of lower quality, since none of the agents who belong to this group is ever going to use product B. Hence, we can focus on payoffs in case the product is of higher quality. Notice that, in this group there are two agents who adopt the product after one period, two agents after two periods, and so on up to two agents who adopt the product after $\frac{s_j}{2}-1$ periods. This means that the payoff from this group is equal to:

$$2\frac{1}{1-\delta} + 2\frac{\delta}{1-\delta} + \cdots + 2\frac{\delta^{\frac{s_j}{2}-1}}{1-\delta} = \frac{2}{1-\delta}\sum_{i=1}^{s_j/2}\delta^{i-1} = \frac{2}{\delta(1-\delta)}\left(\frac{\delta}{1-\delta} - \frac{\delta^{\frac{s_j}{2}+1}}{1-\delta}\right) = \frac{2(1-\delta^{\frac{s_j}{2}})}{(1-\delta)^2}$$

. Suppose that $s_g - s_1 \geq 4$ and consider an alternative configuration such that $s'_g = s_g - 2$ and $s'_1 = s_1 + 2$. The payoffs from all the other groups are the same. Hence, $\Pi(s_1, s_2, \ldots, s_{g-1}, s_g) = \frac{2(1-\delta^{\frac{s_1}{2}})}{(1-\delta)^2} + \frac{2(1-\delta^{\frac{s_g}{2}})}{(1-\delta)^2} + \sum_{j=2}^{g-1}\frac{2(1-\delta^{\frac{s_j}{2}})}{(1-\delta)^2}$ and $\Pi(s'_1, s_2, \ldots, s_{g-1}, s'_g) = \frac{2(1-\delta^{\frac{s'_1}{2}})}{(1-\delta)^2} + \frac{2(1-\delta^{\frac{s'_g}{2}})}{(1-\delta)^2} + \sum_{j=2}^{g-1}\frac{2(1-\delta^{\frac{s_j}{2}})}{(1-\delta)^2}$. We will show that $\Pi(s'_1, s_2, \ldots, s_{g-1}, s'_g) \geq \Pi(s_1, s_2, \ldots, s_{g-1}, s_g)$ always.

$$\begin{aligned}
\Pi(s'_1, \ldots, s'_g) \geq \Pi(s_1, \ldots, s_g) &\Leftrightarrow 1 - \delta^{\frac{s_1}{2}+1} + 1 - \delta^{\frac{s_g}{2}-1} \geq 1 - \delta^{\frac{s_1}{2}} + 1 - \delta^{\frac{s_g}{2}} \Leftrightarrow \\
&\Leftrightarrow \delta^{\frac{s_1}{2}} + \delta^{\frac{s_g}{2}} - \delta^{\frac{s_1}{2}+1} - \delta^{\frac{s_g}{2}-1} \geq 0 \\
&\Leftrightarrow (\delta^{\frac{s_1}{2}} - \delta^{\frac{s_g}{2}-1})(1-\delta) \geq 0
\end{aligned}$$

Given that $\delta \in (0,1)$, this holds as long as $s_1 < s_g$. Therefore, it is optimal to choose $s_g - s_1 \leq 2$.

The second part of the proposition is proven in an analogous manner. We construct the payoffs from a group t_j in case the product is of lower quality. Observe that in case the product is of higher quality all initially targeted agents will yield a payoff equal to $\frac{1}{1-\delta}$ irrespectively of the targeting strategy. Each period, two additional targeted agents change to product A, hence the payoff associated to this group in case the product is of lower quality is equal to:

$$\Pi_{t_j} = (t_j - 2)\delta^{1-1} + (t_j - 4)\delta^{2-1} + \cdots + (t_j - t_j)\delta^{\frac{t_j}{2}-1} = \sum_{i=1}^{t_j/2}(t_j - 2i)\delta^{i-1} =$$

$$= t_j\left(\frac{1-\delta^{\frac{t_j}{2}}}{1-\delta}\right) - \frac{2}{(1-\delta)^2}\left[\left(1-\delta^{\frac{t_j}{2}}\right) - \frac{t_j}{2}\delta^{\frac{t_j}{2}}(1-\delta)\right] =$$

$$= t_j\frac{1}{1-\delta} - 2\frac{1-\delta^{\frac{t_j}{2}}}{(1-\delta)^2}$$

And the total payoff is equal to:

$$\Pi(t_1, \ldots, t_j) = \sum_{j=1}^{g}\left[t_j\frac{1}{1-\delta} - 2\frac{1-\delta^{\frac{t_j}{2}}}{(1-\delta)^2}\right] = \frac{t}{1-\delta} - \frac{2g}{(1-\delta)^2} + \frac{2}{(1-\delta)^2}\sum_{j=1}^{g}\delta^{\frac{t_j}{2}}$$

Therefore, the problem is equivalent to that of maximizing $\sum_{j=1}^{g}\delta^{\frac{t_j}{2}}$, given that $\sum_{j=1}^{g}t_j = t$. By substituting $t_g = t - \sum_{j=1}^{g-1}t_j$ and differentiating with respect to t_j for $j < g$ we get $\frac{\partial\left(\sum_{j=1}^{g}\delta^{\frac{t_j}{2}}\right)}{\partial t_j} = \frac{\ln\delta}{2}\left(\delta^{\frac{t_j}{2}} - \delta^{\frac{t_g}{2}}\right) < 0$. This means that the payoff decreases in the sizes of all groups, but the largest one. Hence, the optimal strategy is to choose $t_1 = \cdots = t_{g-1} = 2$ and $t_g = t - 2(g-1)$. □

Proof of Proposition 3. Let us first construct the objective functions that should be maximized in case the firm targets one or two groups based on Theorem 2. For targeting one group, the expression is:

$$p\left(2\delta + 2\delta^2 + \cdots + 2\delta^{\frac{s}{2}}\right) - (1-p)\left(2\delta + \cdots + 2\delta^{\frac{t}{2}}\right) = 2p\sum_{k=1}^{s/2}\delta^k - 2(1-p)\sum_{k=1}^{t/2}\delta^k$$

For targeting two groups, recall from Proposition 2 that they should be $s_1 = s_2 = s/2$ (given that s is a multiple of 4) and $t_1 = 2$, $t_2 = \frac{t}{2} - 2$. Therefore, the objective function is:

$$p\left(4\delta + 4\delta^2 + \cdots + 4\delta^{\frac{s}{4}}\right) - (1-p)\left(2\delta + \cdots + 2\delta^{\frac{t}{2}-1} + 2\delta\right) = 4p\sum_{k=1}^{s/2}\delta^k - 2(1-p)\sum_{k=1}^{\frac{t}{2}-1}\delta^k - 2(1-p)\delta$$

Hence, the firm should target two groups instead of one if and only if the following condition holds:

$$
\begin{aligned}
&4p\sum_{k=1}^{s/2}\delta^k - 2(1-p)\sum_{k=1}^{\frac{t}{2}-1}\delta^k - 2(1-p)\delta \geq 2p\sum_{k=1}^{s/2}\delta^k - 2(1-p)\sum_{k=1}^{t/2}\delta^k \Leftrightarrow \\
\Leftrightarrow &4p\left[\frac{\delta}{1-\delta} - \frac{\delta^{\frac{s}{4}+1}}{1-\delta}\right] - 2(1-p)\delta \geq 2p\left[\frac{\delta}{1-\delta} - \frac{\delta^{\frac{s}{2}+1}}{1-\delta}\right] - 2(1-p)\delta^{t2} \Leftrightarrow \\
\Leftrightarrow &2p\frac{\delta}{1-\delta} + 2p\frac{\delta^{\frac{s}{2}+1}}{1-\delta} - 4p\frac{\delta^{\frac{s}{4}+1}}{1-\delta} \geq 2(1-p)\left(\delta - \delta^{t2}\right) \Leftrightarrow \\
\Leftrightarrow &p\left(1-\delta^{\frac{s}{4}}\right)^2 \geq (1-p)(1-\delta)\left(1-\delta^{\frac{t}{2}-1}\right) \Leftrightarrow \\
\Leftrightarrow &p \geq \frac{(1-\delta)\left(1-\delta^{\frac{t}{2}-1}\right)}{(1-\delta)\left(1-\delta^{\frac{t}{2}-1}\right) + \left(1-\delta^{\frac{s}{4}}\right)^2}
\end{aligned}
$$

□

Proof of Lemma 1. We first prove that the function $f(\delta, t)$ is increasing in t. First of all, recall that $s = n - t$, therefore it depends on t. Although, t is an integer, this would be a smooth function for t being a real number, therefore we can work as if t was a real, but then focus on the integer values of t. Having clarified this, we can differentiate the function with respect to t and noticing that $\frac{\partial s}{\partial t} = -1$ we get $\frac{\partial f}{\partial t} = \frac{-\frac{1}{2}\delta^{\frac{t}{2}-1}\ln\delta(1-\delta)\left(1-\delta^{\frac{s}{4}}\right)^2 - \frac{1}{2}\left(1-\delta^{\frac{s}{4}}\right)\ln\delta(1-\delta)\left(1-\delta^{\frac{t}{2}-1}\right)}{\left[(1-\delta)\left(1-\delta^{\frac{t}{2}-1}\right)+\left(1-\delta^{\frac{s}{4}}\right)^2\right]^2} > 0$ for all t and $\delta \in [0, 1)$.

Now, we turn our attention to the behavior of the function with respect to δ. Attempting to evaluate the derivative of the function for intermediate values of δ is not easy. However, one can easily calculate that $f(\delta = 0, t) = \frac{1}{2}$, $\frac{\partial f(\delta,t)}{\partial \delta}\big|_{\delta=0} = -\frac{1}{4}$ both for all $t > 2$ and $\lim_{\delta \to 1} \frac{\partial f(\delta,t)}{\partial \delta} = -\frac{2s^2(s-t)(t-2)}{(s^2+8t-16)^2}$, which is negative whenever $t \leq s$ (equivalent to $t \leq n/2$) and positive otherwise.

Instead of working with the derivative, we proceed as follows: Observe that f lies always between 0 and 1. We will show that for any value $\hat{p} \in (0, 1)$ the function f is equal to this value at most once when $t > n/3$ (equivalently $t > s/2$) and at most twice when $t > n/2$ (equivalently $t > s$). For $t \leq n/3$, given that the f is strictly decreasing at 0, this means that the function should be always strictly decreasing. For $t > n/2$, f is strictly decreasing at 0 and strictly increasing very close to 1,[13] which means that it starts decreasing up to some point and then increases.

Consider the equation $\hat{p} = f(\delta, t) = \frac{(1-\delta)\left(1-\delta^{\frac{t}{2}-1}\right)}{(1-\delta)\left(1-\delta^{\frac{t}{2}-1}\right)+\left(1-\delta^{\frac{s}{4}}\right)^2}$. This expression can be rewritten as $\hat{p}\left(1-\delta^{\frac{s}{4}}\right)^2 = (1-\hat{p})(1-\delta)\left(1-\delta^{\frac{t}{2}-1}\right)$. Subsequently, if we define $\hat{r} = \frac{\hat{p}}{1-\hat{p}}$, then the equation becomes $\hat{r}\left(1-\delta^{\frac{s}{4}}\right)^2 = (1-\delta)\left(1-\delta^{\frac{t}{2}-1}\right)$, where $\hat{r} \in [0, \infty)$ and $\hat{r} > 1$ when $\hat{p} > 1/2$. Now, observe

[13] This argument holds because the limit of the derivative with respect to δ at 1 is strictly positive and the function is continuous in δ. Hence, there exists a sufficiently small neighborhood of δ in which the function is strictly increasing.

that using standard identities of algebra we get that $\left(1-\delta^{\frac{t}{2}-1}\right) = (1-\delta)\sum_{k=0}^{\frac{t}{2}-2}\delta^k$ and analogously $\left(1-\delta^{\frac{s}{4}}\right) = (1-\delta)\sum_{k=0}^{\frac{s}{4}-1}\delta^k$. Hence, we can simplify $(1-\delta)^2$ from both sides of the equation and get $\hat{r}\left(\sum_{k=0}^{\frac{s}{4}-1}\delta^k\right)^2 = \sum_{k=0}^{\frac{t}{2}-2}\delta^k$. With some appropriate calculations, which can be found at the end of the proof (see Lemma 2), we can rewrite the equation as follows:

$$\hat{r}\left(1+2\delta+3\delta^2+\cdots+\frac{s}{4}\delta^{\frac{s}{4}-1}+\cdots+2\delta^{\frac{s}{2}-3}+\delta^{\frac{s}{2}-2}\right) = 1+\delta+\cdots+\delta^{\frac{t}{2}-2}$$

Bringing all factors to the left hand side of the expression we can create a polynomial, for which we need to calculate its number of roots. We will do this using Descartes' Rule of Signs (Descartes, 1886), which states that "the number of positive roots of a polynomial is at most equal to the number of sign differences between consecutive nonzero coefficients, or is less than it by an even integers".

First, we consider the case where $t < n/3$ (or $t < s/2$). In this case the expression can be written as follows:

$$(\hat{r}-1)+(2\hat{r}-1)\delta+(3\hat{r}-1)\delta^2+\cdots+\left(\frac{t}{2}\hat{r}-1\right)\delta^{\frac{t}{2}-2}+\frac{t}{2}\hat{r}\delta^{\frac{t}{2}-1}+\cdots+\frac{s}{4}\hat{r}\delta^{\frac{s}{4}-1}+\cdots+2\hat{r}\delta^{\frac{s}{2}-3}+\hat{r}\delta^{\frac{s}{2}-2} = 0$$

One can easily see that all coefficients to the right of $\frac{t}{2}\hat{r}\delta^{\frac{t}{2}-1}$ are positive, as well as that the coefficients of the previous terms are increasing. Hence, if a coefficient $(k\hat{r}-1)$, then $[(k+1)\hat{r}-1]$ will also be positive. Therefore, if $\hat{r} \geq 1$ then all coefficients are positive and the polynomial has no root, whereas if $\hat{r} < 1$ then the first at most $\frac{t}{2}$ coefficients are negative and the rest are positive, leading to one sign difference. Therefore, this polynomial will have one root in the positive numbers. Notice that, this root need not necessarily be in the interval $(0,1)$, which means that the polynomial has at most one root in $(0,1)$. As we have already explained this means that if $t < n/3$ the function f is decreasing in $[0,1)$.

Second, we consider the case where $t > n/2$ (or $t > s$). In this case the expression can be written as follows:

$$(\hat{r}-1)+(2\hat{r}-1)\delta+(3\hat{r}-1)\delta^2+\cdots+\left(\frac{s}{4}\hat{r}-1\right)\delta^{\frac{s}{4}-1}+\cdots+(2\hat{r}-1)\delta^{\frac{s}{2}-3}+(\hat{r}-1)\delta^{\frac{s}{2}-2}-\delta^{\frac{s}{2}-1}-\delta^{\frac{t}{2}-2} = 0$$

Analogously to the previous case, one can see that the coefficients to the right of $\delta^{\frac{s}{2}-1}$ are all negative, however the coefficients of the previous terms are now first increasing and then decreasing. Hence, if $\hat{r} \geq 1$ all terms to the left of $(\hat{r}-1)\delta^{\frac{s}{2}-2}$ are positive and the polynomial has one positive root, whereas if $\hat{r} < 1$ we have to consider two different cases. if it also holds that $\frac{s}{4}\hat{r} \leq 1$ then all coefficients are negative and the polynomial has no root, whereas if $\frac{s}{4}\hat{r} > 1$ then the first coefficient is negative, but there exists at least one coefficient that is positive. This means that the polynomial has two sign changes, so either two or no positive roots (which again might not be in $(0,1)$). Having proven this, and given the values of the function and its derivative at 0, as well as their limits at 1, we can conclude that for $t > n/2$ the function is first decreasing and then increasing.

□

The following lemma is needed for the proof of Lemma 1:

Lemma 2. $\left(\sum_{k=0}^{n} \delta^k\right)^2 = \sum_{k=0}^{n-1}(k+1)(\delta^k+\delta^{2n-k})+(n+1)\delta^n = 1+2\delta+\cdots+(n+1)\delta^n+n\delta^{n+1}+\cdots+\delta^{2n}$

Proof. The lemma is proven by induction. First, we need to show that it holds for $n=1$. $LHS = \left(\sum_{k=0}^{1} \delta^k\right)^2 = (1+\delta)^2 = 1+2\delta+\delta^2$ and $RHS = \sum_{k=0}^{1-1}(k+1)(\delta^k+\delta^{2-k})+2\delta^1 = 1+2\delta+\delta^2$, so the statement holds. Then, we show that if it holds for n, then it also holds for $n+1$.

$$\left(\sum_{k=0}^{n+1} \delta^k\right)^2 = \left(\sum_{k=0}^{n} \delta^k + \delta^{n+1}\right)^2 = \left(\sum_{k=0}^{n} \delta^k\right)^2 + 2\delta^{n+1}\left(\sum_{k=0}^{n} \delta^k\right) + \delta^{2n+2} =$$

Since the statement is true for n, then the above expression is equal to

$$\begin{aligned}
&= \left[1+2\delta+\cdots+(n+1)\delta^n+n\delta^{n+1}+\cdots+\delta^{2n}\right] + \left(2\delta^{n+1}+\cdots+2\delta^{2n+1}\right) + \delta^{2n+2} = \\
&= 1+2\delta+\cdots+(n+1)\delta^n+(n+2)\delta^{n+1}+(n+1)\delta^{n+2}+\cdots+3\delta^{2n}+2\delta^{2n+1}+\delta^{2n+2} = \\
&= \sum_{k=0}^{n}(k+1)(\delta^k+\delta^{2(n+1)-k})+(n+2)\delta^{n+1}
\end{aligned}$$

and this concludes the argument. □

Proof of Proposition 4. Based on Proposition 2 and on the proof of Theorem 2, the expected payoff for the firm from targeting g groups (with some appropriate simplifications) is equal to

$$\Pi_g = 2g\frac{(2p-1)}{(1-\delta)^2} + t\frac{(1-p)}{1-\delta} - \frac{2p}{1-\delta)^2}\sum_{k=1}^{g}\delta^{\frac{s_k}{2}} + \frac{2(1-p)}{1-\delta)^2}\sum_{k=1}^{g}\delta^{\frac{t_k}{2}}$$

where $\sum_{k=1}^{g}\delta^{\frac{s_k}{2}} = g\delta^{\frac{s}{2g}}$ (when s is a multiple of g!) and $\sum_{k=1}^{g}\delta^{\frac{t_k}{2}} = (g-1)\delta+\delta^{\frac{t}{2}-g+1}$ Therefore the firm prefers to target $g+1$ groups instead of g if $\Pi_{g+1} - \Pi_g \geq 0$. Equivalently this means:

$$
\begin{aligned}
&2\frac{(2p-1)}{(1-\delta)^2}[(g+1)-g] - \frac{2p}{(1-\delta)^2}\left[\sum_{k=1}^{g+1}\delta^{\frac{s_k}{2}} - \sum_{k=1}^{g}\delta^{\frac{s_k}{2}}\right] + \frac{2(1-p)}{(1-\delta)^2}\left[\sum_{k=1}^{g+1}\delta^{\frac{t_k}{2}} - \sum_{k=1}^{g}\delta^{\frac{t_k}{2}}\right] \geq 0 \Leftrightarrow \\
&2\frac{(2p-1)}{(1-\delta)^2} - \frac{2p}{(1-\delta)^2}\left[(g+1)\delta^{\frac{s}{2(g+1)}} - g\delta^{\frac{s}{2g}}\right] + \frac{2(1-p)}{(1-\delta)^2}\left[g\delta + \delta^{\frac{t}{2}-g} - (g-1)\delta + \delta^{\frac{t}{2}-g+1}\right] \geq 0 \Leftrightarrow \\
&2p - 1 - p\left(g\delta^{\frac{s}{2(g+1)}} - g\delta^{\frac{s}{2g}}\right) + (1-p)\left(\delta + \delta^{\frac{t}{2}-g} - \delta^{\frac{t}{2}-g+1}\right) \geq 0 \Leftrightarrow \\
&p\left(2 - (g+1)\delta^{\frac{s}{2(g+1)}} - g\delta^{\frac{s}{2g}} - \delta - \delta^{\frac{t}{2}-g} + \delta^{\frac{t}{2}-g+1}\right) \geq 1 - \delta - \delta^{\frac{t}{2}-g} + \delta^{\frac{t}{2}-g+1} \Leftrightarrow \\
&p \geq \frac{(1-\delta)\left(1-\delta^{\frac{t}{2}-g}\right)}{(1-\delta)\left(1-\delta^{\frac{t}{2}-g+1}\right) + \left[1-(g+1)\delta^{\frac{s}{2(g+1)}} - g\delta^{\frac{s}{2g}}\right]} = f(\delta, t, g)
\end{aligned}
$$

If this condition holds, then the firm prefers to target $g+1$ groups instead of g. If this is true for any $g \geq 1$, which holds if $p \geq \max_{g\geq 1} f(\delta, t, g)$ then the firm should target the maximum number of groups. To the contrary, if this is not true for any $g \geq 1$, which happens if $p \leq \min_{g\geq 1} f(\delta, t, g)$ then the firm should target a single group. □