Diffusion by Imitation: The Importance of Targeting Agents☆

Nikolas Tsakas[a]

[a]*Department of Economics, Universidad Carlos III de Madrid*

Abstract

We study the optimal intervention of a planner who seeks to maximize the diffusion of an action in a circular network where agents imitate successful past behavior of their neighbors. We find that the optimal targeting strategy depends on two parameters: (i) the likelihood of the action being more successful than its alternative and (ii) and planner's patience. More specifically, when planner's preferred action has high probability of being more successful than its alternative, then the optimal strategy for an infinitely patient planner is to concentrate all the targeted agents in one connected group; whereas when this probability is low it is optimal to spread them uniformly around the network. Interestingly, for a very impatient planner, the optimal targeting strategy is exactly the opposite. Our results highlight the importance of knowing a society's exact network structure for the efficient design of targeting strategies, especially in settings where the agents are positionally similar.

JEL classification: D83, D85, H23, M37
Keywords: Social Networks, Targeting, Diffusion, Imitation.
First Version: December 2012

1. Introduction

1.1. Motivation

The importance of social interactions for the diffusion of innovations, ideas and behavior are topics that have attracted a lot of research interest (see Jackson, 2008). Recent technological advances have made possible the collection and analysis of data related to the structure of relationships inside societies, as well as the rules guiding the behavior of their members. The appropriate usage of this information can provide useful tools for the effective diffusion of products, technologies and ideas in societies.

In this paper, we describe the optimal intervention of an interested party (from now on called *planner*) who seeks to maximize the diffusion of a given action in

☆I am indebted to Antonio Cabrales, Matthew Jackson, Hamid Sabourian, Marco Celentani, Ignacio Ortuño, Antoine Loeper, Angel Hernando, Daniel Garcia, Adrien Vigier, Jesper Rüdiger, Marcin Dziubinski, Ronald Peeters, Natalia Fabra, George Deltas, Isabel Melguizo, Anil Yildizparlak and the participants at Tilburg University GSS Seminar, UC3M Student Workshop, ENTER Jamboree 2013, GSE Workshop in Dynamic Decisions and CRETE 2013 for fruitful comments and suggestions which have improved significantly the quality of this paper. All remaining errors are my sole responsibility.

Email address: `ntsakas@eco.uc3m.es` (Nikolas Tsakas)

a society where the agents imitate successful past behavior of their neighbors. In practice, the planner can be either a firm who seeks to spread its product in a new market, or a social planner who seeks to promote the use of a new technology, or even a political party that seeks to propagate its ideology. We illustrate how beneficial the knowledge about society's network structure may be for the efficient design of marketing and social influence strategies.[1]

There are several reasons why economic agents adopt simple behavioral rules, such as the imitation of successful past behavior. For example, they often need to make decisions without knowing the potential gains or losses of their possible choices. Additionally, when these situations arise with high frequency and the agents' computational capabilities are limited, then they tend to rely on information received by past experience of others, rather than experimenting themselves.[2] These arguments are also supported by a recent, but growing, empirical and experimental literature which provides strong evidence about the fact that in several decision problems the agents, indeed, tend to imitate those who have been particularly successful (see Apesteguia et al., 2007, Conley and Udry, 2010, Bigoni and Fort, 2013).

Most of the existing literature on targeting in networks has focused on the importance of those agents who exhibit some kind of centrality (see Ballester et al., 2006). Having a high or a low number of connections (see Galeotti and Goyal, 2009, Chatterjee and Dutta, 2011), or diffusing information to many others who are poorly connected (see Galeotti et al., 2011) are some usual characteristics of influential agents. The importance of these characteristics is obvious and beyond doubt. Nevertheless, we show that there is another important factor that affects diffusion significantly. This is whether the targeted agents are concentrated all together, in the sense that they are connected between them, or they are spread around the network. Notice that, in order to use this additional tool, one should have information about the exact structure of the network. Each one of these strategies may be optimal depending on certain parameters, with the most important of them being the patience of the planner.

Throughout our analysis we highlight the differences between the optimal targeting strategies of an *infinitely patient planner*, i.e. one who cares about the diffusion of her preferred action in the long run, and an *impatient planner*, i.e. one who cares about the diffusion of her preferred action in the short run. To our knowledge, this is the first paper that determines both short and long run optimal targeting strategies and provides a clear comparison between them. Interestingly enough, we find a sharp contrast between the optimal targeting strategies in the two cases, which persists for all values of the other parameters. This comparison is important in several realistic scenarios, since different targeting strategies may be appropriate for different time horizons.

[1] In broader terms, the importance of social networks in efficient marketing design has been studied extensively and in several different disciplines. For an extensive list see Galeotti and Goyal (2009) and references therein.

[2] These are some of the reasons why imitation has been subject to extensive theoretical study in different environments (see Ellison and Fudenberg, 1993, 1995, Vega-Redondo, 1997, Eshel et al., 1998, Schlag, 1998, Alós-Ferrer and Weidenholzer, 2008, Duersch et al., 2012).

1.2. Results

Formally, we consider a finite population of behaviorally homogeneous agents located around a circle. In each period, all agents choose simultaneously between two alternative actions. The stage payoff that each action yields is uncertain and depends on a random shock, which is common for all the agents who have chosen the same action in that period.[3] Shocks are independent across actions and across periods. There are no strategic interactions between agents. After making their decisions, all agent observe the chosen actions and the realized payoffs of their two immediate neighbors. Subsequently, they update their choice myopically, imitating the action that yielded the highest payoff within their neighborhood in the preceding period. Notice that, an agent who does not observe any of her neighbors choosing the alternative action never changes her choice.

A problem that fits well our model is that of the diffusion of agricultural technologies.[4] Farmers' harvests depends mostly on common factors such as the weather and the fertility of the land. Moreover, it is normal to assume that they are aware of the technologies and crops used by their neighbors, as well as the payoffs they receive. In particular, Conley and Udry (2010) show that farmers tend to imitate those who have been very successful in the past, whereas as pointed out by Ellison and Fudenberg (1993) the farmers's technology decisions are guided mostly by short–term considerations, especially when capital markets are poorly developed or malfunctioning.

The planner is interested in maximizing the diffusion of her preferred action in the population. The planner can be either infinitely patient, therefore interested in the diffusion of the action in the long run; or impatient, therefore interested in the diffusion of the action after just one period. She is assumed to know the structure of the network, as well as how agents behave and she can intervene in society by enforcing a change at the initial choice of a subset of the population. Ideally, she would like to target the whole population, but doing so in reality would be extremely costly. Hence, our goal is to identify the planner's optimal targeting strategy given the number of agents she is able to target.

Observe that none of the agents exhibits any kind of centrality. In fact, none of them has any positional advantage or disadvantage compared to the rest of the population. Despite this fact, we find that expected diffusion changes substantially depending on the subset of the population that has been initially targeted by the planner. This highlights the significance for the planner of knowing the exact structure of the network.

We show that the optimal targeting strategy depends on two parameters: (i) the likelihood of the planner's preferred action being more successful than its alternative and (ii) and planner's patience. In fact, we observe a sharp contrast between the optimal strategies of an infinitely patient planner and that of a very impatient one. More specifically, when planner's preferred action has high probability of being more successful than its alternative, then the optimal targeting strategy for an infinitely

[3]This assumption facilitates the tractability of results and does not seem to affect the main insights

[4]see also Ellison and Fudenberg (1993)

patient planner is to concentrate all the targeted agents in one connected group;[5] whereas when this probability is low it is optimal to spread them uniformly around the network. Interestingly, for a very impatient planner, the optimal targeting strategy is exactly the opposite.

The intuition is relatively simple and depends on the fact that in the long run only one of the two actions survives. Therefore, when the action is likely to be successful, then an infinitely patient planner wants to prevent its disappearance due to a few consecutive negative shocks in the first periods. For this reason she prefers to concentrate them all together. To the contrary, if the action is unlikely to be successful, then the optimal strategy for the planner is to try and take advantage of a possible sequence of successful shocks during the first period. By concentrating all the agents together, she would only manage to make her preferred action disappear more slowly, since for its diffusion it would be needed a large number of consecutive successful shocks, which is rather unlikely to happen.

For the impatient planner the arguments are reversed. When the action is likely to be successful, then the planner wants to make it visible to as many agents as possible, therefore she should spread the initial adopters around the society. On the other hand, if the action is more likely to be unsuccessful, then the planner wants to prevent as many of the initial adopters as possible from observing the alternative action, therefore she should concentrate them all together.

Finally, we extend our analysis in many different directions. We discuss the optimal strategies of planners with intermediate levels of patience, thus intending to identify how the transition between the two extreme cases occurs. Moreover, we quantify the practical meaning of infinite patience by characterizing the expected waiting time before convergence occurs. We observe that, for those cases in which the planner's optimal strategy is to concentrate all the initial adopters in one group the process is slowed down substantially. In addition to this, we discuss what happens if we allow for inertia and we show that our results remain unchanged. This extension captures many realistic features, as the existence of switching costs and some forms of conformity. Finally, we repeat our analysis for the linear and the star network, as an attempt to identify the effect of centrality to our results.

1.3. Related Literature

The role of influential agents on social networks has been studied in different disciplines, such as computer science (see Kempe et al., 2003, Richardson and Domingos, 2002) and marketing (see Kirby and Marsden, 2006), as well as in economics. Intuitively, a crucial feature is the centrality of the agents, which depicts either the number of immediate neighbors they have or how important they are for the connectivity of the network (see Ballester et al., 2006).

Other environments similar to ours have been studied in physics, mathematics and computer science, especially in the areas of cellular automata and voter models. The most similar paper to ours is Bagnoli et al. (2001), where they study the long-run behavior and the phase transition of a system with similar characteristics as ours. A crucial difference of our work is that they assume the initial conditions to be random, given that they refer to initial positions of particles. Hence, they do

[5] In the circle network, a connected group is a segment of the circle.

not focus on identifying which of these initial conditions would be the optimal ones, which is the main focus of the present work.

In economics literature, the paper which is the most closely related to voter models is by Ortuño (1993). The author considers a standeard voter model setting, where the agents are located in a two dimensional infinite lattice and they probabilistically adopt the choice made by one of their neighbors. This is the only paper where centrality does not play a role. The agents are homogeneous in preferences and location and there exists a planner who seeks the diffusion of a technology in the society and can choose between targeting a single connected segment of agents, or spread them in the population uniformly. The author concludes that these two choices lead to the same probability of diffusion. The result is guided mainly by the infinite size of the lattice. In our paper, the population is finite and we do not restrict the potential choices of the planner. This features allow us to obtain more general results. More recently, Yildiz et al. (2011) generalize the standard voter model by introducing "stubborn" agents, i.e. who never change their choice. Similarly to us, the authors also discuss the problem of optimal placement of stubborn agents, when trying to maximize their impact on the long run expected choices of agents.

Our paper is more closely related to Galeotti and Goyal (2009). Their research question is similar, however they focus on "word-of-mouth" communication and on social conformism, which are mechanisms that disregard the performance of a product. They look for the optimal influence strategy of a firm who seeks to maximize the diffusion of its product in a society. Apart from focusing on different learning mechanisms, there are two more main differences with our work. First, they focus completely on the short run optimal targeting strategy, using a two–period model. Second, in their case the network is represented as a degree distribution, hence each agent meets a fixed number of agents every period, but those are randomly drawn by a population mass. This approach disregards potential information about the exact structure of the network, which we show to be important for the planner. Moreover, there remains another open question, which is what would happen if the degree distribution tends to become uniform and furthermore whether we could do better by knowing the exact formation of the network. Both of these questions can be tackled in our environment with promising results. They show that, depending on the learning mechanism, agents with low or high number of connections should be targeted, underlining again the important role of centrality.

In a recent paper, Galeotti et al. (2011) show that, in a setting where information transmission is strategic (in contrast to what happens here), influential agents are those who diffuse information to many others, who themselves are poorly connected. Another closely related paper is the one by Chatterjee and Dutta (2011). They study the optimal behavior of a firm that seeks to diffuse a technology in a society with network structure, focusing mostly on the linear and the circular network. The fact that part of the population consists of perfectly rational agents and that the structure of the network is common knowledge changes significantly the dynamics of the system. Also, the firm is allowed to target only a single agent. They find that a firm which produces a good quality product will want to place target a node that maximizes the decay centrality, whereas a firm producing a bad quality product will want to target the agent with the highest number of connections. There is an apparent analogy between these results and our, in the sense that in both cases the optimal choice depends on the quality of the technology, which in our case could be approximated loosely by the value of p.

Regarding agents' behavior, we focus on imitation of successful past behavior. Recently, there is a growing empirical literature (see Apesteguia et al., 2007, Conley and Udry, 2010, Bigoni and Fort, 2013) that provides empirical evidence on the adoption of this behavior in real environments. In theoretical framework, Vega-Redondo (1997) has shown that a Cournot economy, where the agents follow this rule, converges to the Walrasian equilibrium. Moreover, Alós-Ferrer and Weidenholzer (2008), Eshel et al. (1998) and Fosco and Mengel (2011) study coordination games and public good games respectively, played between neighbors, where the agents imitate their most successful neighbor. Their focus is mostly on the characterization of stochastically stable configurations.

In a general framework, the current analysis builds upon the work on 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). These articles study different learning mechanisms in environments where the agents face common individual problems and there are no strategic interactions. They mainly discuss conditions under which efficient actions spread to the whole population. Especially, Ellison and Fudenberg (1993) use an environment very similar to ours, in the sense that the agents repeatedly choose between two alternative technologies whose payoff depends on a random shock, which is common for all the agents who use the same technology. Apart from those similarities in the setting, there are several differences regarding the role of the network and other details of the model, but the main difference is that their focus is not on the characterization of the optimal intervention in favor of one of the two technologies.

The rest of the paper is organized as follows. In section 2 we define formally the model. Section 3 contains the characterization of the optimal targeting strategy for the impatient planner, whereas in Section 4 we analyze the optimal targeting strategy for the infinitely patient planner. In Section 5 we discuss briefly some extensions and in Section 6 we conclude. An extensive study of the extensions can be found in Appendix A. All proofs can be found in Appendix B.

2. The Model

2.1. The Agents

There is a finite set of agents $N = \{1, ..., n\}$, mentioned as *population* of the network. At $\tau = 1, 2, ...$, each agent $i \in N$ chooses between two alternative actions, $a_i^\tau \in \{A, B\}$. Each action yields random payoff. The payoff of agent i is independent of the other agents' choices. Therefore, interactions among agents are not strategic and their connections represent only informational exchange. Moreover, agents who choose the same action at a given period receive equal payoffs.[6] The payoffs of both actions change in each period, with the realizations being independent across periods. Action B yields strictly higher payoff than action A with probability $p \in$

[6]This assumption does not affect the main intuitions and is imposed mainly in order to facilitate the tractability of the results.

(0, 1), while action A yields strictly higher payoff than B with probability q. For the derivation of the main results, we focus on the case where $q = 1 - p$. Notice that, for $q = 1 - p$, the probability of both actions yielding exactly equal payoffs is zero, which is assumed only in order to avoid unnecessary tie-breaks and does not affect the results. In the section of extensions we relax this assumption and we allow for different values of q.[7] The ratio between p and q turns out to be crucial for our analysis, therefore, we define $r = \frac{p}{q}$. From now on, we will say that there is a *success* (*failure*) in period τ if action B (A) yielded higher payoff in this period. [8]

The *planner* is an agent, outside of the population, who seeks to maximize the diffusion of action B in the population. She can do so by changing to her favor the choice of a subset of the population before the beginning of the first period. Optimally, she would like to affect the whole population, but in reality this would be extremely costly. This cost enters implicitly by assuming that the cardinality of the subset she can affect is fixed exogenously.[9] More specifically, given that at period $\tau = 0$ all the agents are choosing action A, the planner can target $t \leq n$ agents from the population and make them choose B at period $\tau = 1$. After that, she cannot affect the society anymore. The goal of this paper is to characterize the planner's optimal targeting strategy.

The planner can be either impatient or infinitely patient. A planner is called *impatient* if she cares about the diffusion of her preferred action after only one period. Similarly, a planner is called *infinitely patient* if she cares only about the diffusion of her preferred action in the long–run. We find that the optimal behavior of an impatient planner is exactly the opposite than that of an infinitely patient planner. In the section of Extensions, we discuss as well some intermediate levels of planner's patience.

2.2. The Network

A social network is represented by a family of sets $\mathcal{N} := \{N_i \subseteq N \mid i = 1, ..., n\}$, with N_i, called i's neighborhood, denoting the set of agents observed by agent i. We assume N_i to contain i as well. In the main part of our analysis, we examine the *circular* network, where each agent interacts with her two immediate neighbors, i.e. $N_i = \{i-1, i, i+1\}$ for $i = 2, \ldots, n-1$, whereas $N_1 = \{n, 1, 2\}$ and $N_n = \{n-1, n, 1\}$. The current structure imposes an undirected network, because for all $i, j \in N$, $j \in N_i$ if and only if $i \in N_j$. In this setting, the network structure describes the flow of information in the network, in the sense that each agent $i \in N$ can observe the action and the realized payoff of her neighbors, $j \in N_i$.

2.3. The Behavior

At the end of each period, the agents observe the actions and realized payoffs of their neighbors. Subsequently, they have the opportunity to revise their choices.

[7] Different values of q allow us to introduce introduce the possibility of inertia. Later on, we use this feature to capture scenarios where either the two actions may have the same realized payoffs, or there are externalities between the actions of one's neighbors, or switching costs.

[8] Later on, with some abuse of terminology, we will also define the success and failure in a random walk, in a similar way.

[9] In the section of Extensions in the Appendix we endogenize t and we discuss the returns of investment for different values of it

Revisions happen simultaneously for all agents. We assume that each agent $i \in N$ can observe the choices and the realized payoffs of her neighbors in the previous period. According to these observations, she revises her choice by imitating the most successful action within her neighborhood in the preceding period. Notice that, an agent never switches to an action that she did not observe, i.e. that neither her nor any of her neighbors chose in the previous period. Moreover, if an action disappears from the population, then it never reappears.

The important aspect of this myopic behavior is that the agents discard most of the available information. They ignore whatever has happened before the previous period, hence they are unable to form beliefs about the underline payoff distributions of their alternative choices.

2.4. The Problem

After the action of the planner, the population consists of s agents choosing action A (from now on called *non–adopters*) and t agents choosing action B (from now on called *adopters*); obviously $s+t = n$. We call *group* a sequence of neighboring agents who all choose the same action and are surrounded by agents choosing the opposite action. The population is formed of m groups of neighboring agents who choose action A, with sizes $\{s_1, s_2, \ldots, s_m\}$, where $\sum_{k=1}^{m} s_k = s$ and analogously m groups of neighboring agents who all choose action B, with sizes $\{t_1, t_2, \ldots, t_m\}$, where $\sum_{k=1}^{m} t_k = t$.[10] We mention these groups as groups *of type* A and *of type* B respectively. The numbering of the groups is based on their size in increasing order, $s_1 \leq s_2 \leq \cdots \leq s_m$ and $t_1 \leq t_2 \leq \cdots \leq t_m$. With some abuse of notation we also use $s_1, s_2, \ldots, s_m$ and $t_1, t_2, \ldots, t_m$ to name the groups.

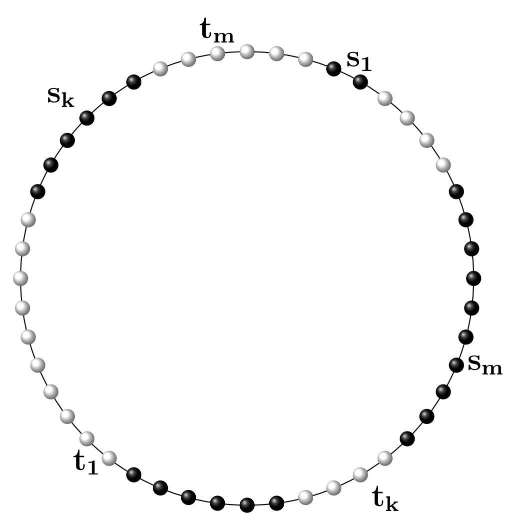


Figure 1: Example of an initial configuration: White nodes represent agents choosing action B and black nodes agents choosing action A.

Our goal is to find the optimal size of all s_k and t_k for $k = 1, \ldots, m-1$,[11] their optimal position (if it matters), as well as the optimal number of groups, m.

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

[11] s_m and t_m must be equal to $s_m = s - \sum_{k=1}^{m-1} s_k$ and $t_m = t - \sum_{k=1}^{m-1} t_k$ respectively.

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. This does not cause important limitations in our results, since the underlying ideas remain the same. We refer the reader to the analysis of the linear network for a more detailed motivation regarding this assumption. Formally:

Assumption 1 (A1). *s_i and t_i are even numbers for all $i \in \{1, \ldots, m\}$.*

3. Results for the Impatient Planner

In this section, we study the optimal targeting strategy of a planner who cares about maximizing the expected number of agents choosing action B after exactly one period.[12] The following figure (Figure 2) shows the two possible configurations after one period. White dots represent the agents who choose initially action B and black dots those agents who choose initially action A. Observe that only those agents who are on the boundary of a group can change their choice. In fact, for m denoting the total number of groups, in case action B is more successful in the first period, then there will be $2m$ additional adopters in the next period; whereas, in case action A is more successful, the number of adopters will decrease by $2m$. The probabilities of ending in each of the two possible states is p and $q = 1 - p$ respectively.

Figure 2: Initial configuration and the two possible configurations after one period.

Hence we can construct the objective function of the impatient planner:

$$EN_B(1) = t + 2mp - 2m(1-p) = t + 2m(2p-1) = t + \frac{2m}{1-p}(r-1)$$

It is easy to see that the optimal targeting strategy depends on the probability p of action B being more successful. Namely, for $p > 1/2$ the objective function is strictly increasing in m.[13] Therefore, it is optimal to have as many groups as possible. On the other hand, if $p < 1/2$ the objective function is strictly decreasing

[12]Implicitly, the planner is assumed to be risk neutral.

[13]If $q \neq 1 - p$, we would simply have to replace $2p - 1$ with $p - q$.

in m and therefore it is optimal to locate all initial adopters in one group. We state formally this result in the following proposition:

Proposition 1. *Under (A1), then for an impatient planner*

- *If $p > 1/2$, the optimal targeting strategy is to spread the initial adopters to as many groups as possible, i.e.*
  - *If $t < s$, then $m = \frac{t}{2}$, with $t_1 = \cdots = t_m = 2$.*
  - *If $t > s$, then $m = \frac{s}{2}$, with $s_1 = \cdots = s_m = 2$.*
- *If $p < 1/2$, the optimal targeting strategy is to concentrate all the initial adopters in one group, i.e. $m = 1$ and $t_1 = t$.*

Observe that, as long as the planner creates the maximum number of groups, the allocation of the agents inside these groups is not important.

Intuitively, this result suggests the following. If an action is likely to be successful, then the planner should try to make it directly visible to as many non–adopters as possible. In doing so, upon a highly probable successful realization, she will manage to attract the maximum number of additional adopters. On the contrary, if an action is unlikely to be successful, then the planner prefers to prevent most of the initial adopters from observing the opposite action. As a result, even upon an unsuccessful realization, most of the initial adopters will not observe the alternative action, therefore they will not revise their choice in the second period. As we will see, this optimal strategy changes sharply if the planner is infinitely patient.

4. Results for an Infinitely Patient Planner

In this section, we study the optimal behavior of an infinitely patient planner, i.e. who cares only about the diffusion of her preferred action in the long run. A crucial feature of this setting is that such a planner disregards completely the speed of the procedure. Before beginning our analysis, it is useful to state the following two prior results.

4.1. Preliminaries

4.1.1. Diffusion when Agents Imitate-the-Best Neighbor

The present behavioral rule constitutes a special case of an "Imitate-the-Best Neighbor" rule, applied in a setting of individual decision–making under uncertainty without strategic interactions between agents. Agents observe the choices of their neighbors and the payoff those choices yield. Subsequently, they revise their choices repeatedly according to these observations. In particular, they do so by imitating the action that yielded the highest payoff within their neighborhood in the preceding period.

In such a setting, it turns out that every connected network converges with probability one to a steady state where all the agents are choosing the same action (see Tsakas, 2013). Moreover, any of the actions can be the one to survive in the long run. This is based on the fact that all actions are vulnerable to a sequence of negative shocks, which can lead to their disappearance. Given that an action which disappears from the network never reappears, it turns out that only one of them survives at the end.

In our case, this result ensures that only one of the two actions will survive in the long run and that both actions have a positive probability to be the ones to succeed it. Hence, the optimal strategy for an infinitely patient planner is the one that maximizes the probability that action B gets diffused to the whole population in the long run. We define this probability as follows:

Definition 1. *$P_B(\cdot)$ is the probability that action B will be diffused to the whole population in the long run.*

This probability will depend not only on the size of the population n, the number of targeted agents t and the probability of success p, but also on the choice of the planner about which agents to target. Notice as well that maximizing this probability is equivalent to maximizing the expected number of agents choosing action B in the long run; a remark that may seem obvious, but it will clarify the analogy between our short run and long run analysis.

4.1.2. Results on Random Walks with Absorbing Barriers

A technical result which turns out to be particularly useful comes from Kemeny and Snell (1960). It refers to the properties of a finite one–dimensional random walk with absorbing barriers. The endpoints of a random walk are called absorbing barriers if upon reaching one of the endpoints the random walk eventually stays there forever. Specifically, the authors compute the random walk's probabilities of absorption at each one of the two absorbing states as follows:

Lemma 1 (Kemeny and Snell (1960))**.** *Consider a random walk with states $\{0, 1, \ldots, n\}$, where both barriers 0 and n are absorbing. If the probability of moving to the right (from i to $i+1$) is p, the probability of moving to the left (from i to $i-1$) is q, and $r = \frac{p}{q}$, then the probability of absorption at state n, when starting from state i is:*

$$P_n(i) = \begin{cases} \frac{r^n - r^{n-i}}{r^n - 1} & \text{if } p \neq q \text{ (or equivalently } r \neq 1) \\ \frac{i}{n} & \text{if } p = q \text{ (or equivalently } r = 1) \end{cases} \tag{1}$$

Analogously, the probability of absorption at the state 0 is $P_0(i) = 1 - P_n(i)$.

For the moment, $p \neq q$ is equivalent to$p \neq \frac{1}{2}$. Later on we extend our analysis to show that the results are completely analogous in the general case.

To help us understand how this result can be used to express the current diffusion process we consider a linear network (see figure below), with agents named $\{1, 2, \ldots, n\}$, where each agent has two neighbors, except of agents 1 and n who have one neighbor each. In period $\tau = 1$, agents 1 to i choose action B and the rest choose action A. Hence, every period only two agents may revise their choice (for example in the first period those are the agents i and $i+1$). The border fluctuates until either agent n chooses B, or agent 1 chooses A. The position of the right border of adopters is following a random walk with absorbing barriers 0 and n.[14]

[14]In order that agent 1 chooses A, the left barrier must be set at location equal to 0.

Hence, we can use the result stated above to describe the probability of diffusion for each one of the two actions. We call a random walk *successful* (*unsuccessful*) if it ends up in the absorbing state where all the nodes included in the walk choose action $B(A)$.

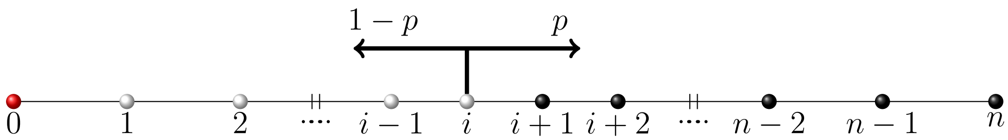


Figure 3: Initial configuration of a random walk with absorbing barriers 0 and n.

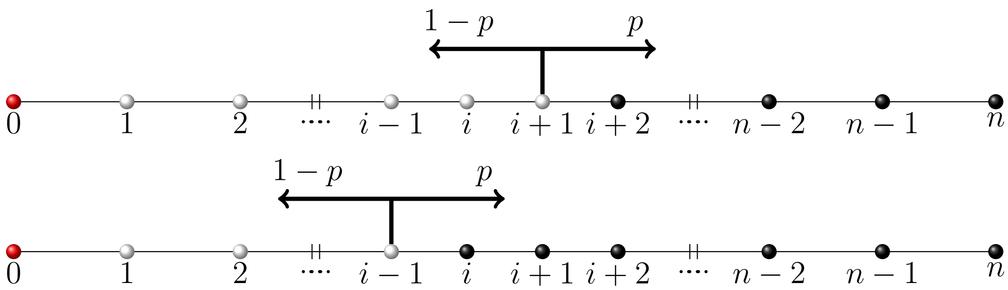


Figure 4: The two possible configurations after one period.

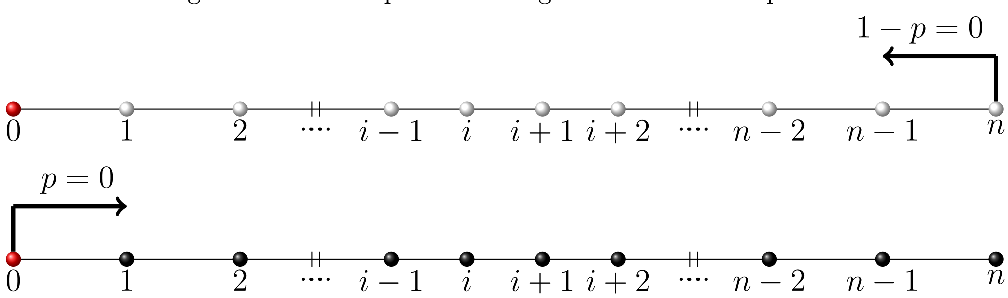


Figure 5: The two possible configurations after absorption.

This result is particularly helpful for our analysis, because any initial targeting choice induces a stochastic process that can be expressed as a sequence of conditionally independent random walks with absorbing barriers, similar to the one described above. Despite having multiple borders between groups, all of them fluctuate synchronously, because the payoffs for each action are perfectly correlated and therefore all agents in the boundaries make the same choice in each period. For example, considering the beginning of the process, each border fluctuates until either smallest group of adopters, with size t_1 or the smallest group of non–adopters, with size s_1 disappears. This process can be represented by the random walk that is shown in the following figure. Upon success or failure of the first walk, the process starts fluctuating according to a new random walk that depends on the smaller still existing groups of each type.

4.2. Main Results

Not surprisingly, the ratio $r = \frac{p}{q} = \frac{p}{1-p}$, which describes the likelihood of action B being more successful than action A, is a crucial parameter. However, surprisingly enough, this is the only parameter that affects the optimal targeting strategy and

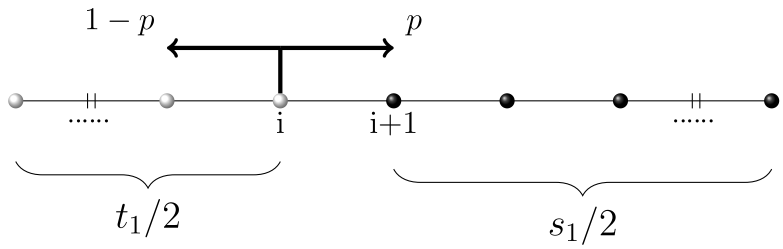


Figure 6: The first random walk

more specifically whether r is higher or lower than 1.[15] The most interesting result, though, is that the optimal strategy of the infinitely patient planner is in complete contrast to that of the impatient planner. In particular, we observe that, for $r > 1$ the optimal targeting strategy of the planner is to concentrate all the targeted agents in a single group, whereas for low values of $r < 1$, the optimal strategy is to spread them as much as possible in the population, splitting them in as many groups as possible and as symmetrically as possible. Observe that, the results are not just different, but are exactly opposite to those found for the impatient planner. In fact, the optimal strategy for an infinitely patient planner is the worst possible strategy for an impatient planner and vice versa.

For a better exposition of the general results, we split the problem into three sub-problems. First, we consider the case where the groups are restricted to be symmetric. Then, we consider the asymmetric case where the planner can target up to two groups and finally we consider the general asymmetric case.

4.2.1. Symmetric cases

First, we consider the symmetric case, where the groups of agents choosing the same action are restricted to have equal sizes, namely $s_1 = \cdots = s_m = \frac{s}{m}$ and $t_1 = \cdots = t_m = \frac{t}{m}$. Assuming no problems of divisibility we find the optimal number of groups, m. As we have mentioned already, the optimal targeting strategy depends only on the ratio r. In particular, when $r > 1$ it is optimal to concentrate all the initial adopters in one group, whereas when $r < 1$, it is optimal to spread them in as many groups as possible, i.e. $m = \min\{s/2, t/2\}$. Formally:

Proposition 2. *Under (A1) and given $s_1 = \cdots = s_m = \frac{s}{m}$ and $t_1 = \cdots = t_m = \frac{t}{m}$, then for an infinitely patient planner:*

- *If $r > 1$, then* $\arg\max_m P_B(m|s,t,n,r) = 1$
- *If $r < 1$, then* $\arg\max_m P_B(m|s,t,n,r) = \min\{s/2, t/2\}$

All proofs can be found in the Appendix B.

Intuitively, this proposition suggests that when the probability of success is high, it is beneficial to concentrate all initial adopters together. This prevents the disappearance of the preferred action upon the realization of a sequence of negative

[15]Even though, at the moment $r > 1$ is identical to $p > 1/2$, we keep this notation because it facilitates the extension to cases where $q \neq 1 - p$.

shocks during the first periods. The opposite strategy is optimal when the probability of success is low. Then the planner wants to take advantage of some potential good shocks during the first periods, which will spread the action to as many agents as possible.

4.2.2. Asymmetric Cases

We turn our attention towards the more general asymmetric cases. At first, let us consider the case where the planner is restricted to target at most two groups of each type, with sizes s_1, s_2 and t_1, t_2 respectively. Then the process can be described as shown in the figure below. Recall that $s_1 \leq s_2$ and $t_1 \leq t_2$.

It turns out that the optimal decision depends completely on the value of r. More specifically, if $r > 1$ it is optimal to concentrate all the agents in one group, while if $r < 1$ the optimal choice is to have two groups of equal size for each action.

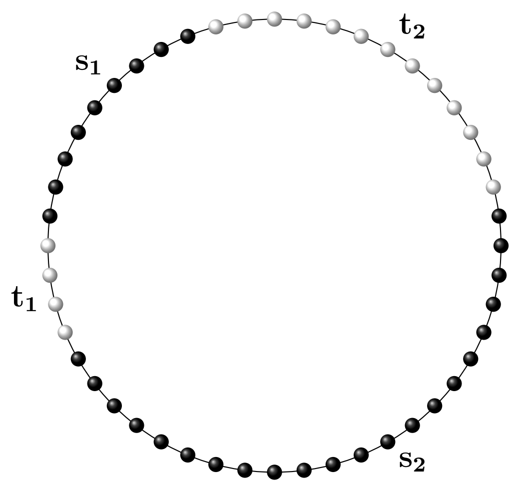


Figure 7: Example of an initial configuration with two groups of each type

Proposition 3. *Under (A1) and given $m \leq 2$, then for an infinitely patient planner:*

- *If $r > 1$, the optimal targeting strategy is to concentrate all the initial adopters in one group.*
- *If $r < 1$, then the optimal targeting strategy is to split the initial adopters into two as equal as possible groups, and locate them in the population as symmetrically as possible, i.e. $s_2 - s_1 \leq 2$ and $t_2 - t_1 \leq 2$*

The intuition is similar to that of the symmetric case. However, an interesting finding is that this result does not hold for all restrictions on m. For example, if we restrict the number of groups to be not greater than three, then it is not optimal to split the agents into three equal groups of each type. Hence, it is not the case that we always prefer symmetric compared to asymmetric configuration. Notice that, this would be a sufficient condition for the proof of our main result, however it does not always hold. Nevertheless, this does not affect our general result which is stated below.

The two propositions help us construct the main theorem of the paper which describes the optimal targeting strategy in the general case of m initial groups of each type, allowing them to have different sizes. The result is in line with the previous findings and suggests that when $r > 1$, then the optimal choice is to concentrate all the initial adopters in one group; whereas when $r < 1$, then the optimal choice is to spread them uniformly to the population in as many and as equal groups as possible. Namely:

Theorem 1. *Under (A1), for an infinitely patient planner*

- *If $r > 1$, the optimal targeting strategy is to concentrate all the initial adopters in one group, i.e. $t_m = t$ and $t_1 = \cdots = t_{m-1} = 0$ for any m.*
- *If $r < 1$, the optimal targeting strategy is to spread the initial adopters in as many groups as possible and locate these groups as symmetrically as possible, i.e.*
  - *If $t < s$, then $m = \frac{t}{2}$, with $t_1 = \cdots = t_m = 2$ and $s_m - s_1 \leq 2$,*
  - *If $t > s$, then $m = \frac{s}{2}$, with $s_1 = \cdots = s_m = 2$ and $t_m - t_1 \leq 2$*

As it has been already mentioned, the importance of this result lies in the complete contrast between the optimal strategy of an infinitely patient planner in comparison to an impatient one. An infinitely patient planner prefers to protect from some initial negative shocks an action which is more likely to be successful, whereas she prefers to spread as much as possible an action which is more likely to be unsuccessful, trying to take advantage of a few positive shocks in the first periods. When the probability of success is low, she knows that by concentrating all the targeted agents together, a lot of positive shocks will be needed in order to capture the whole population, which is rather improbable for an action that is expected to be often unsuccessful.

5. Extensions

In this section, we briefly present the results of several extensions that clarify specific questions regarding the process of interest. One can find an analytical discussion of those extensions in Appendix A.

A natural question that arises from the contrast between the optimal behavior of an impatient and an infinitely patient planner is what happens for intermediate levels of patience. First of all, we observe that the expected diffusion is very sensitive to small changes in the initial configuration and therefore it becomes particularly hard to construct a general strategy for all intermediate levels of patience. Nevertheless, we discuss the optimal targeting strategy of a planner who cares about the diffusion of the action after 3 periods and we obtain a very enlightening result. Namely, if $r > 1$, then in some cases the planner prefers to spread the initial adopters in groups consisting of four, instead of two, agents. Whereas, if $r < 1$, she always needs to compare between the two extreme cases, i.e. concentrating all of them in one group or spread them to as many groups as possible. This result provides a useful starting point to understand how the optimal targeting strategy changes as the planner becomes more patient.

Furthermore, we discuss what happens if the choice of t becomes endogenous. We find that under some simple and intuitive conditions, if $r > 1$ then the expected profits of the planner have decreasing returns to scale in t. We also provide a condition under which the function of expected profits attains a strictly interior maximum, which means that for the planner it is not optimal anymore to target all the agents in the population, even if she has no budget constraints.

Another crucial aspect over which we would like to shed more light is the practical meaning of infinite patience. In reality, no planner can wait literally infinitely many periods, so we try to identify the expected time of diffusion and how the planner's choices that maximize the expected diffusion affect this expected time. Not surprisingly, we find that a larger number of groups leads to faster diffusion and therefore the optimal strategy of the planner for $r > 1$ has the drawback of maximizing also the expected waiting time before diffusion occurs.. On the other hand, for $r < 1$ the optimal strategy is also the one that leads to the fastest expected time of diffusion. An interesting feature, which is in line with standard results in the analysis of random walks, is that the expected time of diffusion explodes as r gets very close to 1. In general we find that for different configurations the expected time until diffusion occurs may vary substantially and therefore one must be very careful when acting as an infinitely patient planner.

Moreover, we discuss cases where inertia is possible. This generalization allows us to capture some realistic scenarios, which include the possibility of both actions having equally good realizations, the existence of switching costs and some forms of conformity. Such settings can be captured by allowing $q \neq 1 - p$, where q is the probability of action A being more successful than action B, and we explain why our results are not affected by this feature.

Finally, given that our analysis is based on the complete absence of centrality features, we also discuss optimal targeting strategy of the planner for some simple structures to give a notion of how our results would be affected under the presence of central agents. First, we study the linear network and we find that for $r > 1$ it is optimal to target one of the two corners, whereas for $r < 1$ it is optimal to target the agents located around the center. In this part of our analysis, we also drop the assumption of groups having an even number of agents and we show why dropping this assumption complicates our analysis without providing additional insights. We discuss also briefly the star network, in which it becomes apparent the vast importance of very central agents.

6. Conclusion

We have analyzed the optimal intervention of a planner who seeks to maximize the diffusion of an action, in a circular network where the agents imitate successful past behavior of their neighbors. It turns out that there is room for strategic targeting of initial adopters even in this case where all the agents are completely identical. Knowledge about the exact structure of the network, and not only its degree distribution, can be beneficial for a planner. We find that the optimal decision depends almost completely on two parameters. On the likelihood r of the preferred action being more successful and on how patient the planner is. Changes in these two parameters lead to completely opposite optimal behavior.

A very important parameter that we have disregarded completely from our analysis is the risk aversion of the planner. We have assumed the planner to be risk neutral, caring only about the expected number of adopters. For a risk averse planner, we would expect the optimal behavior to contain more dispersed targets than for the risk neutral one, but this is definitely an open question for future research.

The current paper constitutes a first attempt to explore targeting possibilities in networks where agents imitate successful behavior. Given that our network structure is relatively simple, a natural extension would be to discover which of the current features are still present and which of these are changing when we pass to more general network structures. It is apparent that centrality features that arise in more complex networks will play an important role, however the exact characteristics still remain to be studied.

Appendix A. Extensions

Appendix A.1. Intermediate patience - the τ-patient planner

The sharp contrast between optimal targeting strategies of impatient versus infinitely patient planner makes plausible the question of how this transition occurs as the planner becomes more patient. Intermediate levels of patience may be defined in several alternative ways, of which we choose the following:

Definition 2. *The planner is τ-*patient *if she cares about maximizing the expected number of agents choosing action B at period τ, denoted by $EN_B(\tau)$.*

Although, at first sight, this definition of impatience might seem unusual, it covers all the important intuitions and at the same time it can be easily extended to more complicated cases. For example, the qualitative results would be very similar if the planner cared about the sum of discounted expect number of agents choosing her preferred action from period one up to period τ.

First of all, we analyze the case where the planner is 2-patient, so that she cares only about the expected number of adopters after two periods. Notice that, after two periods, there are four possible states the society can be at.[16] Notice, also, that the number of groups may have decreased after the first period. This is because, after the first period, the groups consisting of only two agents who chose the action that was unsuccessful during that period will disappear. For a visual representation, look at the configuration occurring after failure at Figure 2, where the bottom left group of two agents choosing action B disappears after a failure for action B in the first period. Hence, we can construct easily the objective function of a 2-patient planner as:

$$EN_B(2) = t + 2(2p-1)\{2m - [p\alpha_2 + (1-p)\beta_2]\}$$

where α_2 is the number of groups consisting of two agents choosing A and β_2 is the number of groups consisting of two agents choosing action B.

Notice that there are two opposing forces. We analyze them for $p > 1/2$, since for $p < 1/2$ the analysis is exactly the opposite. For $p > 1/2$, on the one hand, we would like to have as many groups as possible, since this increases the expected number of agents choosing B. On the other hand, we would like not to have groups

[16] In general after τ-periods there are 2^τ possible histories, so there are 2^τ possible configurations.

with size two, since this enters with a negative sign. Hence, as long as we can create more groups, with size larger or equal than four agents, it is preferred to do so. Now, what would happen if in order to create one extra group, we would have to create one or more groups with size two, for each one of the actions. The most difficult condition to be satisfied arises if we are left with groups of size no larger than four agents for both actions. In this case, in order to create one additional group, we would need to substitute one group of four agents with two groups of two agents for each type. Therefore, we would need to increase the number of both α_2 and β_2 by two. The necessary condition such that we prefer to have one extra group would be:

$$2m - [p\alpha_2 + (1-p)\beta_2] \leq 2(m+1) - [p(\alpha_2+2) + (1-p)(\beta_2+2)]$$

With some straightforward calculations, we see that this condition is satisfied always with equality, hence we are indifferent between having this one additional group or not. Given that this is the toughest condition to satisfy, this means that in all the other cases it is strictly preferred to have one more group. Concluding, for a 2-patient planner it is always preferred, at least weakly, to have one extra group. Therefore, her optimal behavior is identical to the one of the 1-patient planner.

The intuition is basically the same for the 3-patient planner. Notice that, during the first two periods it is possible that also groups of four agents disappear, which may affect the total number of groups in the third period. The objective function of the planner becomes:

$$EN_B(3) = t + 2(2p-1)\{3m - 2[p\alpha_2 + (1-p)\beta_2] - [p^2\alpha_4 + (1-p)^2\beta_4]\}$$

where α_4 (β_4) is the number of groups consisting of four agents who choose action A (B) and α_2, β_2 are as defined above.

We observe that, in this case it would be preferable not to have groups consisting neither of two nor or four agents. Hence, there is an ambiguous trade–off between having more groups and those additional groups consisting of four or less agents. One can easily observe that the creation of groups consisting of four agents has very small negative effect, compared to the positive effect of the addition of a group. This means that the planner prefers to have two groups with four agents, rather than one with eight. However, there are still more cases to analyze. The case captured in the figure below is the one that imposes the toughest condition to satisfy. In particular, the question is whether the planner prefers to break two groups of four agents each into four groups of two agents each.

The condition is the following and it is not satisfied for any p:

$$\begin{aligned}
&3m - 2[p\alpha_2 + (1-p)\beta_2] - [p^2\alpha_4 + (1-p)^2\beta_4] \leq \\
&\leq 3(m+1) - 2[p(\alpha_2+2) + (1-p)(\beta_2+2)] - [p^2(\alpha_4-1) + (1-p)^2(\beta_4-1)] \Leftrightarrow \\
&\Leftrightarrow 3 - 4p - 4(1-p) + p^2 + (1-p)^2 \geq 0
\end{aligned}$$

In a similar manner, we need to check whether other possible ways of creating more groups are beneficial. This case turns up to be the only one where the 3-patient planner prefers not to have more groups. The second stricter condition arises by

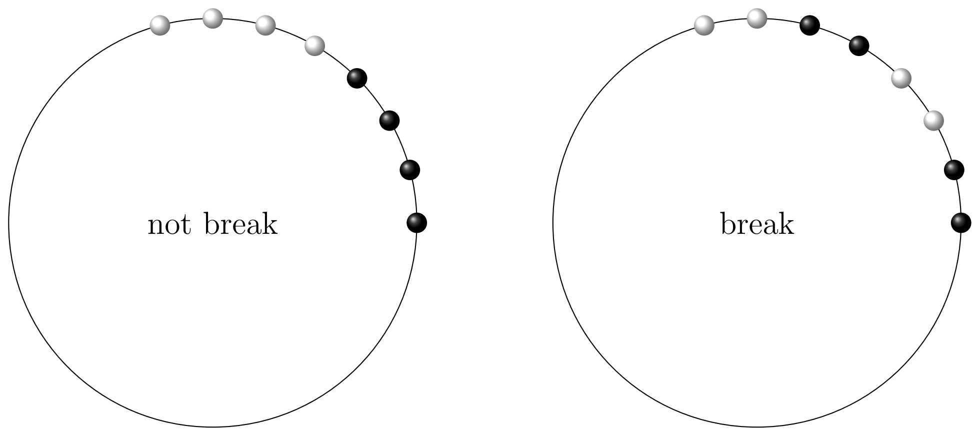


Figure A.8: Configuration where a 3-patient planner prefers not to create an additional group if $r > 1$

the configuration depicted in the following figure. In this case, it turns out that the planner is indifferent between the two alternative configurations. All the other conditions are strictly satisfied.

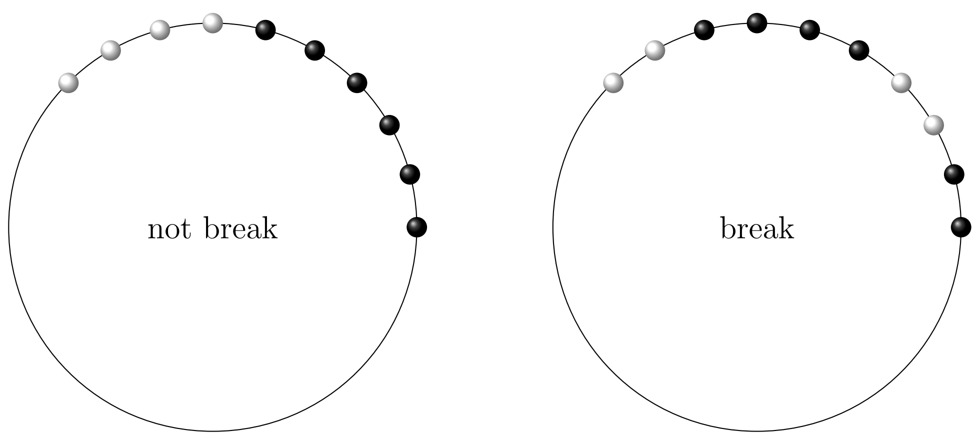


Figure A.9: Configuration where a 3-patient planner is indifferent between creating or not an additional group, if $r > 1$.

We can describe the optimal behavior of a 3-patient planner as follows. For $p > 1/2$, if $t \lll s$, then optimally $t_1 = t_2 = \cdots = t_m = 2$ and $m = t/2$. For $s \lll t$, optimally $s_1 = s_2 = \cdots = s_m = 2$ and $m = s/2$. However, if $s \approx t$ then if t is a multiple of 4 the optimal targeting strategy is to choose $t_1 = t_2 = \cdots = t_m = 4$ and $m = t/4$ and if t is not a multiple of 4, then $t_1 = 2$, $t_2 = \cdots = t_m$ and $m = \frac{t+2}{4}$ and the same for s.

Analogously, for $p < 1/2$, the objective function of the planner decreases in the number of groups until $m = t/4$ and increases afterwards. Hence, we need to compare the cases where $m = 1$ and $m = \min\{s/2, t/2\}$. Comparing the two cases for $s \approx t$, we get the condition:

$$3 \geq 3\frac{t}{2} - 2p\frac{s}{2} - 2(1-p)\frac{t}{2}$$

which is satisfied whenever $t \geq 6$. Hence, it is not optimal to target one group only when $t \approx s$ and $t < 6$.

The case of the 3-patient planner provides very useful insights about the change in the planner's optimal behavior as she becomes more patient. On the one hand, for actions with high probability of success ($p > 1/2$) it seems that this transition happens smoothly, since it becomes optimal for the planner to target larger and larger groups. On the other hand, when the action has low probability of success ($p < 1/2$) this transition happens suddenly, meaning that for low values of t it is optimal to target as many groups as possible, whereas for high values of t it is optimal to target only one group.

Continuing the analysis for more patient planners would give more complicated results which would not be easily tractable. This is because the results would depend vastly on the exact initial position of each group and not only on their number and sizes. A complete characterization for any value of planner's patience would be an interesting and illuminating extension of our paper.

Appendix A.2. Returns to Scale of Investment

In this section, we endogenize the choice of the number of initial adopters t and the returns to scale of investment for an infinitely patient planner. We define explicitly the expected profits of the planner as follows:

$$E\Pi(t) = \pi P_B(t) - c(t)$$

where $P_B(t)$ is the probability of successful diffusion, which depends only on t if we assume that the planner has chosen the optimal targeting strategy. The constant π is a fixed benefit that the planner receives if action B captures the whole population and $c(t)$ is a a strictly increasing and weakly convex cost function. As in most of our analysis, we consider $E\Pi(t)$, $P_B(t)$ and$c(t)$ as continuous and twice differentiable functions, of which we are interested only at the integer values of t.

For $r > 1$, we show that the expected profits have decreasing returns to scale. Moreover, if the cost function is linear and the marginal cost of adding an initial adopter is neither two small, nor too large, then there exists an interior choice of initial adopters which maximizes the expected profits. This result is very intuitive if we recall that targeting additional agents increases the probability of successful diffusion, but never ensures it. Therefore, for a given number of initial adopters, by targeting one more the increase in probability may be so little, that the cost for convincing this agent exceeds the expected increase in profits. We summarize these results in the following proposition.

Proposition 4. *For $r > 1$, the expected profits' function of the planner, $E\Pi(t)$, given her optimal targeting strategy and cost function $c(t)$ such that $c'(t) > 0$ and $c''(t) \geq 0$:*

- *Has decreasing returns to scale in t,*
- *If $c(t) = kt$, $k \in \mathbb{R}_+$, then it has strictly interior maximum if and only if $\frac{\pi \ln r}{2(r^{\frac{n}{2}}-1)} < k < \frac{\pi \ln r}{2(r^{\frac{n}{2}}-1)} r^{\frac{n}{2}}$.*

If k is larger than the upper bound then the derivative is always negative, which means that targeting any agent is too expensive and therefore the optimal choice is $t = 0$. If k is lower than the lower bound then the derivative is always positive, which means that targeting an agent is very cheap and therefore under the absence of initial budget constraint the optimal choice is $t = n$.

For $r < 1$, the result is more ambiguous because targeting more agents affects also the optimal number of groups. Nevertheless, if we slightly modify the argument, we can show that for a fixed number of groups it is possible that the expected profits have increasing returns to scale when $r < 1$ and $t < s$. In particular, this is the case if the cost function is homogeneous of degree one, which holds for example when the cost function is linear. Hence, given that splitting the agents into more groups yields even higher expected profits, at least when the cost function is linear the expected profits have increasing returns to scale when $r < 1$ and $t < s$.

Appendix A.3. Expected Waiting Time before Absorption

We have analyzed the different optimal targeting strategies, given the planner's level of patience. However, in practice, our definition of long–run may have very different characteristics depending on the parameters of the problem. In order to complement our previous analysis, we turn our attention towards the identification of the expected waiting time before the diffusion of one of the two actions in the whole population.

Our goal is twofold. On the one hand, we want to quantify the meaning of "infinite patience" of a planner, by characterizing the performance of her optimal choices with respect to the expected waiting time before the population converges to a steady state. On the other hand, we explore potential trade-offs between maximizing the probability of successful diffusion and minimizing the expected waiting time before diffusion occurs.

Not surprisingly, we find that diffusion is expected to occur faster as the number of groups increases. This result is very intuitive, because having more groups (of equal sizes) implies that the size of each group is going to be smaller; and when the groups are smaller then fewer failures are sufficient to make them disappear. feature that leads to faster convergence. This means that, for $p < \frac{1}{2}$ the targeting strategy that maximizes the probability of diffusion, minimizes also the expected time until absorption, therefore there is no trade–off between the two aspects. Moreover, we see that convergence occurs fast even in absolute terms. On the other hand, for $p > \frac{1}{2}$, our optimal targeting strategy is the one with the maximum expected waiting time, compared to all the symmetric configurations, which for some values of p close to $\frac{1}{2}$ may be unrealistically high in absolute terms. Therefore, there is an important trade–off for the planner.

Furthermore, for the optimal initial configurations if $p > \frac{1}{2}$ or $p < \frac{1}{2}$ with $t > s$then the expected waiting time increases as t increases. However, it decreases if $p < \frac{1}{2}$ with $t < s$. For $p > \frac{1}{2}$, this happens because an increase in t is associated with an increased size of the unique group of each type. For $p < \frac{1}{2}$ with $t > s$, increasing t leaves unaffected the number of groups (because this depends on s) and increases their size. Whereas, for $p < \frac{1}{2}$ and $t > s$, increasing t leads to an increase in the number of groups and a decrease in their size.

In addition to this, in most of the cases we see that the expected waiting time becomes larger for values of p close to $\frac{1}{2}$ and decreases sometimes up to one or two orders of magnitude as p approaches either 0 or 1. We see also that the larger the difference between the values of s and t, the more extreme the effect in waiting time of changes in the values of p. The problem becomes particularly important when the planner targets a single group of initial adopters, because then the differences may become extreme (for example look at Figure B.25 in Appendix B). In addition to this, in absolute terms these changes increase significantly as the population increases.

To establish our results on a concrete manner, we use the following result, which comes from Kemeny and Snell (1960). Namely:

Lemma 2 (Kemeny and Snell (1960))**.** *Consider a random walk with states* $\{0, 1, \ldots, n\}$*, where both barriers* 0 *and* n *are absorbing. If the probability of moving to the right (from* i *to* $i+1$*) is* p*, the probability of moving to the left (from* i *to* $i-1$*) is* q *and* $r = \frac{p}{q}$*, then, starting from state* i*, the expected waiting time before absorption occurs is:*

$$\tau_{(i,n-i)} = \begin{cases} \frac{1}{p-q}[n\frac{r^n - r^{n-i}}{r^n - 1} - i] & \text{if } p \neq \frac{1}{2} \\ i(n-i) & \text{if } p = \frac{1}{2} \end{cases} \tag{A.1}$$

Recall that we focus on the case where $q = 1 - p$. Now, we are ready to characterize the expected waiting time for the planner's optimal targeting strategies.

Appendix A.3.1. Expected waiting times of optimal configurations

For $r > 1$, recall that the optimal choice for the infinitely patient planner is to target a single group of adopters. In this case the process can be summarized as a random walk with $n/2$ steps and two absorbing barriers, starting from $i = t/2$. Hence the expected waiting time is equal to:

$$\tau(t, s|r > 1) = \tau_{(\frac{t}{2}, \frac{s}{2})} = \frac{1}{2p-1}\left[\frac{n}{2}\left(\frac{r^{n/2} - r^{s/2}}{r^{n/2} - 1}\right) - \frac{t}{2}\right] = \frac{1}{2p-1}\left[\frac{s}{2} - \frac{n}{2}\left(\frac{r^{s/2} - 1}{r^{n/2} - 1}\right)\right]$$

The following proposition summarizes the properties of the above expression:

Remark 1. For $r > 1$, the expected waiting time before absorption, $\tau(t, s|r > 1)$, is

1. Strictly increasing in n,
2. Strictly concave in s and in t,
3. Given p and n, attains interior maximum $t^* = n - s^* = n - 2\frac{\ln(r^{n/2}-1) - \ln(\ln r) - \ln(n/2)}{\ln r}$
4. $\lim_{r \to 1^+} t^* = \frac{n}{2}$ and $\lim_{r \to 1^+} \tau_{(t^*, s^* = \frac{n}{2})} = \frac{n^2}{16}$
5. $\lim_{r \to +\infty} t^* = 0$ and $\lim_{r \to +\infty} \tau_{(t^*, s^* = n-2)} = \frac{n}{2} - 1$.

◁

Notice, that t^* may not be an even integer number. In this case we have to compare the two closest even integers to t^*, in order to identify the size of t that would give the maximum expected time before absorption occurs. As r diverges to infinity, t^* approaches 0, which is not in the domain of t, hence we check the expected time at $t = 2$.

When p is close to $\frac{1}{2}$, the relation between the expected time of absorption, τ, and the probability of success, p, is a bit more sensitive to the different parameter values. This makes it hard to provide a general result. However, we have analyzed numerically every possible circular network with up to one million agents and we have found that, for all the values of p when we can target sufficiently many agents and for values of p sufficiently far away from $1/2$ the expected time of absorption is decreasing in p. Sufficiently far away has different meaning depending on the size of the network. For populations larger than 30 agents is less than 0.6, for populations larger than 100 agents is below 0.55 and for populations larger than 1000 it gets so close to $1/2$ that in practice it is true everywhere. Moreover, this decrease may be as big as one or two orders of magnitude. The Figures B.25 and B.26 (in Appendix B) depict the expected time of absorption for different values of $s = n - t$, in a population with 200 agents.

For $r < 1$ we must differentiate between the cases where $t < s$ and $t > s$. We concentrate in cases where there is no problem of divisibility, i.e. $s_i = s/m$ and $t_i = t/m$ respectively. For $t < s$, after some simplifications, the expected waiting time is equal to:

$$\tau(t, s|r < 1, t < s) = \tau_{(1,\frac{s}{t})} = \frac{1}{2p-1}\left[\frac{s}{n-s} - \frac{n}{n-s}\left(\frac{r^{\frac{s}{n-s}} - 1}{r^{\frac{n}{n-s}} - 1}\right)\right]$$

Remark 2. For $r < 1$ and $t \leq s$, the expected waiting time, $\tau(t, s|r < 1, t < s)$, is

1. Strictly decreasing in n,
2. Strictly decreasing in t and strictly increasing in s,
3. $\lim_{r \to 0} \tau = 1$ and $\lim_{r \to 1^-} \tau = \frac{s}{n-s}$,
4. $\tau(t = \frac{n}{2}, s = \frac{n}{2}) = \tau_{(1,1)} = 1$ and $\lim_{s \to n} \tau = -\frac{1}{2p-1}$

◁

The first two results are very intuitive. Increasing $t(s)$, decreases (increases) $\frac{s}{t}$ and therefore decreases (increases) the length of the random walk associated with the process. Decreasing (Increasing) the length of the walk, decreases (increases) the expected time before absorption occurs.

More interesting is the part (4) of Remark 2 which shows that the minimum and the maximum expected waiting time does not depend on the size of the population, but only on the probability of success. This result is very useful, because it provides a natural upper bound for the expected waiting time, which does not explode as the population grows.

For the relation between τ and p we repeat the same numerical analysis we did before for $r > 1$ and for all circular networks with up to one million agents. We find that for $s > \frac{n}{2}$ the expected waiting time is strictly increasing in p. For, $s = \frac{n}{2}$ is stable and equal to 1. The behavior is qualitatively identical irrespectively of the size of the population. In Figures B.27 and B.28 in Appendix B we show the typical relation for different values of t.

Notice that, from parts (2) and (3) of Remark 2, we find that the maximum expected time of absorption occurs as p approaches $\frac{1}{2}$ and $t = 2$ and is equal to $\frac{n}{2} - 1$. This is exactly equal to the previous case, described in Remark 1, where $r \to \infty$ and $t = 2$, which is the fastest among the configurations that maximize the expected waiting time. This shows, that comparing worst case scenarios for $r < 1$

and $r > 1$, we find that the slowest of all the worst case scenarios when $r < 1$ is still faster than the fastest of the worst case configurations for $r > 1$. This provides a notion of how much faster are the optimal configurations for $r < 1$ with respect to those for $r > 1$.

For $t > s$, the intuitions are analogous to the previous case. The expected waiting time before absorption is equal to:

$$\tau(t, s|r < 1, t > s) = \tau_{(\frac{t}{s},1)} = \frac{1}{2p-1}\left[1 - \frac{n}{s}\frac{r-1}{r^{\frac{n}{s}}-1}\right]$$

Remark 3. For $r < 1$ and $t \geq s$, the expected waiting time, $\tau(t, s|r < 1, t > s)$, is:

1. Strictly increasing in n,
2. Strictly increasing in t and strictly decreasing in s,
3. $\lim\limits_{r\to 0} \tau = \lim\limits_{r\to 1^-} \tau = \frac{t}{s}$
4. $\lim\limits_{s\to 0} \tau = +\infty$ and $\lim\limits_{s\to \frac{n}{2}} \tau = 1$

$\triangleleft$

As expected, the two first results are the contrary to those of Remark 2. Their justification is a straightforward reversal of the argument used there. An important issue that arises here is that the expected time is not always increasing as p increases. In fact, it seems to be increasing in the beginning up to some point and then decreasing. The point where it gets maximized varies a lot as we alter p, s and t. Again, a typical behavior can be found in Figures B.29 and B.30. The shape of the figures is very similar for all networks we checked (all networks with up to one million agents). An interesting feature is expected waiting is minimized as p approaches 0 and 1/2 with the two limits being always exactly equal.

Appendix A.3.2. Comparison between configurations with different number of groups

Now, we turn our attention to the comparison between configurations consisting of different number of groups. We analyze the simplest of these cases, where all the groups are symmetric. Recall that, when groups are symmetric, then the optimal strategy for an infinitely patient planner is, if $r > 1$, to locate them in a single group and if $r < 1$ to spread them to as many groups as possible. Moreover, even when we allow for asymmetric configurations, the optimal choices again tend to be symmetric. However, it is not clear yet which is the effect of this choice on the expected waiting time before absorption occurs. By Lemma 2, this is equal to:

$$\tau_{(\frac{t}{2m},\frac{s}{2m})} = \frac{1}{2p-1}\left[\frac{s}{2m} - \frac{n}{2m}\left(\frac{r^{\frac{s}{2m}}-1}{r^{\frac{n}{2m}}-1}\right)\right]$$

Proposition 5. *For all $r \neq 1$ and for symmetric configurations, i.e. $s_1 = \cdots = s_m = \frac{s}{m}$ and $t_1 = \cdots = t_m = \frac{t}{m}$, the expected waiting time is strictly decreasing in the number of groups, m.*

Numerical examples can be found in Figures B.31 and B.32, which show that this decrease has an exponential shape. Even though we have restricted our attention to symmetric configurations, this proposition provides an interesting intuition regarding the trade-off between maximizing the probability of successful diffusion and minimizing the expected waiting time before diffusion occurs. On the one hand, for $r > 1$, the choice that maximizes the probability of successful diffusion is the one with the longest expected waiting time. On the other hand, for $r < 1$, the planner's choice that maximizes the probability of successful diffusion is also the one minimizing the expected waiting time until diffusion. Theoretically, these findings are in line with standard results on finite random walks with absorbing barriers, where increasing equally the number of steps in both directions leads to an increase in the expected waiting time before absorption, but it has an ambiguous effect on the probability of absorption in each one of the two barriers.

In our problem, this result stresses out the fact that the strategy of concentrating all the initial adopters in one group when r is high, despite maximizing the probability of successful diffusion, may slow down the procedure significantly. This should be taken into account by a planner when trying to quantify the meaning of being infinitely patient. In practical terms, the waiting time may be so long that even a very patient planner would find it unrealistic to wait until diffusion occurs. Nevertheless, when r is low, the optimal choice for the planner is also the one which minimizes the expected waiting time. This facilitates the decision of the planner since there is no trade-off between the two characteristics of the procedure.

Appendix A.4. $q \neq 1 - p$: Conformity, Switching Cost and Possibility of Inertia

Until now, we have focused on the case where $q = 1 - p$. In practice this meant that if two neighbors were using different actions at some period, then in the next period either they would both use action B, which would happen with probability p, or would both use action A, which would happen with probability $1 - p$. This assumption rules out several realistic problems, which we mention here.

Nevertheless, our whole analysis, for both the patient and the impatient planner, is based not on the value of p itself, but on the value of r, which we defined as the relative likelihood of action B being more successful than action A, more formally as $r = \frac{p}{q}$. Despite the fact that, for simplicity reasons, we concentrated on the case where $q = 1 - p$, this need not be the case necessarily. Our results remain unchanged if we substitute the condition $p > \frac{1}{2}$ ($p < \frac{1}{2}$ respectively) with $p > q$ ($p < q$ respectively). In both cases, the conditions can be summarized by the value of r and in fact whether $r > 1$ or $r < 1$. In this extension we discuss three realistic scenarios, where relaxing this assumption would be necessary and therefore our results are appropriate.

Appendix A.4.1. Possibility of Inertia

With our previous description we have ruled out the possibility of both actions yielding the same payoff. A plausible modification would be to allow both actions to be equally successful in some states of nature. This would lead to a combination of configurations with transition probabilities as shown in Figure 9.
Focusing on i and $i + 1$, who would be the only agents (in this part of the network) who could change their decision, we would obtain the following possible cases:
As we have already mentioned, the results are still valid if we define $r = \frac{p}{q}$.

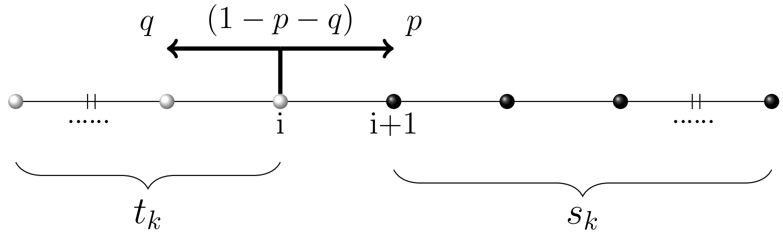


Figure A.10: Transition probabilities under the possibility of inertia

Figure A.11: The left figure shows agents' i and $i+1$ choices at period τ. The other three figures show the choices of the agents at period $\tau+1$, after (ii) success, (iii) failure, (iv) draw at period τ.

Appendix A.4.2. Switching Cost

In the example of technology adoption, we have disregarded the effect of switching costs. Implementing a new technology assumes the purchase of new machinery, or effort of learning how to use the new technology efficiently. Therefore, a farmer in order to decide to pay this cost must have observed the new technology to have been sufficiently better than the one she already uses. Sufficiently better in this setting can be translated as follows. Agent i who uses action B at period τ changes to action A if the payoff of action A at that period was sufficiently higher, i.e. $\pi_A^\tau > \pi_B^\tau + c$, where π_B^τ is the payoff of B at that period and c the switching cost. Let us call q the probability of this realization. Analogously, agent $i+1$ who uses action A, changes to action B if $\pi_B^\tau > \pi_A^\tau + c$, which occurs with probability p. Now, if $|\pi_A - \pi_B| < c$ which happens with probability $1-p-q$,[17] then both agents keep using the same action. The alternative action did not seem to be successful enough to convince them to abandon their current technology. This scenario is along the same lines with those have already mentioned. Hence, our results are suitable also under the presence of switching costs.

Appendix A.4.3. Conformity

Our mechanism is also suitable to describe cases where the choices of the agents do not depend only on the performance of the actions, but also on the number of their neighbors (including themselves) who used each action in the previous period. Think of the following mechanism:

- Those agents who observed at least one agent (including themselves) choosing action B in period τ will choose action B in period $\tau+1$, with probability p_1,
- Those agents who observed at least two agents (including themselves) choosing action B in period τ will choose action B in period $\tau+1$, with probability p_2,

[17] We assume that the payoffs are defined such that the switching cost is high enough to ensure that $p+q<1$.

- Those agents who observed exactly three agents (including themselves) choosing action B in period τ will choose action B in period $\tau+1$, with probability p_3.

This mechanism introduces a notion of *conformity* in the network, since the more of your neighbors chose one action in the previous period, the more probable is that you choose it in the next period. For example, notice that those agents who observe only one of the two alternative actions satisfy all the three conditions, therefore they will choose the same action again with probability $p_1 + p_2 + p_3 = 1$. Also, those agents who observed two agents choosing B will do the same in the next period with probability $p_1 + p_2$. This probability is larger than the one of those who observed only one agent choosing B in the previous period, which is p_1. The dynamics of the network can be summarized by the following figure.

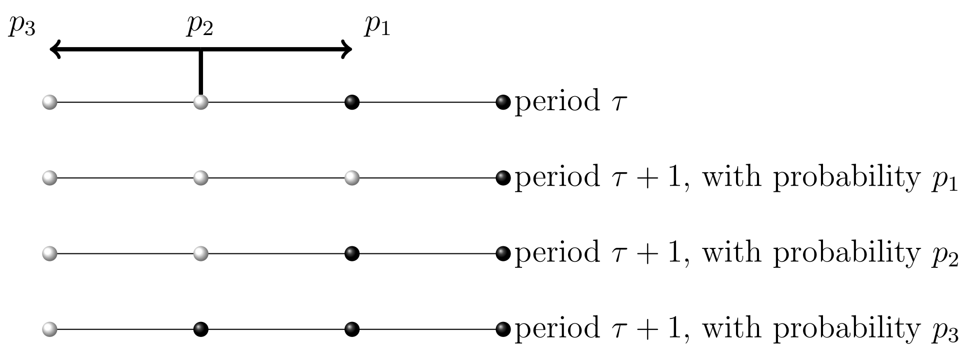


Figure A.12: Possible configurations and transition probabilities under the presence of conformity.

To be in line with our previous notation, if we define $p_1 = p$, $p_3 = q$ and $p_2 = 1 - p - q$, we see that our results still hold under the presence of *conformity*.

Appendix A.5. The Line

In this section we turn our attention towards the *linear* network. The only difference between the line and the circle is that in the line the agents 1 and n are not connected between them, so they have only one neighbor. Formally, $N_i = \{i-1, i, i+1\}$ for $i = 2, \ldots, n-1$, whereas $N_1 = \{1, 2\}$ and $N_n = \{n-1, n\}$.

This structure introduces a notion of centrality in the network. Mainly, this is not because of the fact that some agents have different number of neighbors; but because there is only one path that connects indirectly each two agents. Hence, the agents located closer to the center of the line act as hubs for the transmission of information through the network. This new feature has interesting interesting implications which we are worth discussing.

We consider only the case where there is a single group of initial adopters, with size t and is surrounded by two groups of non–adopters with sizes s_1 and s_2 respectively (see figure below). First, let us first construct $P_B(s_1|s, t, n, r)$ for different values of s_1, s_2 and t. Notice that the only independent variable is s_1 since $s_2 = s - s_1$. Without loss of generality, we consider only the cases where $s_1 \leq s_2$. By symmetry, the rest of the cases are completely analogous.

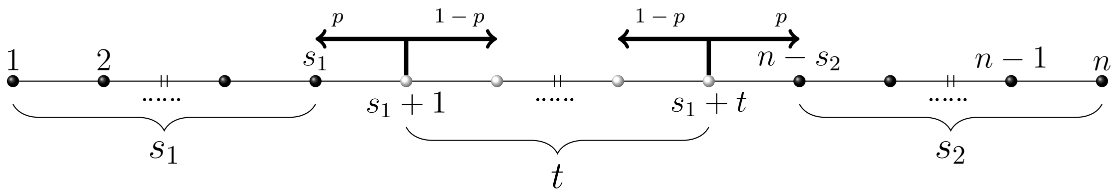


Figure A.13: A linear network with one group of initial adopters.

The probability of diffusion can be expressed as the product of two random walks with absorbing barriers. The first walk describes the procedure until either the group of type B disappears, or the smaller group of type A disappears. For $r \neq 1$ the first walk is depicted in Figure 13. The probability of success in this walk depends on whether t is an odd or an even number. In case the first walk is unsuccessful, action B disappears from the population. In case it is successful, then it is pursued by the random walk in Figure 14, which is the same for t being odd or even.

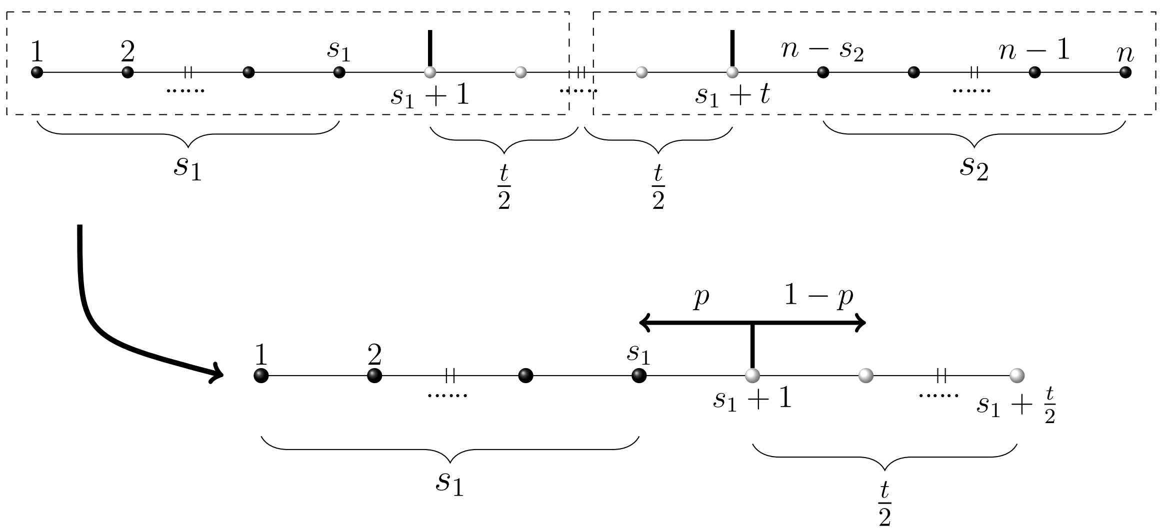


Figure A.14: The first random walk.

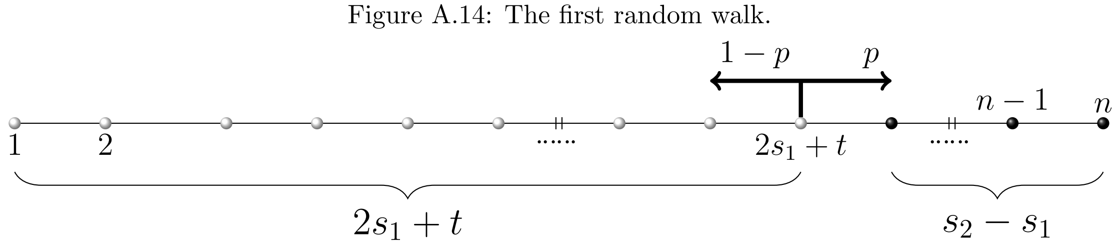


Figure A.15: The random walk that follows successful absorption in the first walk.

Hence, the probability of diffusion of action B, for $r \neq 1$ is:

$$P_B(s_1|s,t,n,r) = \begin{cases} \frac{r^{(s_1+\frac{t+1}{2})}-r^{s_1}}{r^{(s_1+\frac{t+1}{2}}-1)} \frac{r^n-r^{(n-t-2s_1)}}{r^n-1} & \text{if } t \text{ odd} \\ \frac{r^{(s_1+\frac{t}{2})}-r^{s_1}}{r^{(s_1+\frac{t}{2})}-1} \frac{r^n-r^{(n-t-2s_1)}}{r^n-1} & \text{if } t \text{ even} \end{cases} \tag{A.2}$$

and for $r = 1$:

$$P_B(s_1|s,t,n,r=1) = \begin{cases} \frac{(t+1)(t+2s_1)}{(t+2s_1+1)n} & \text{if } t \text{ odd} \\ \frac{t}{n} & \text{if } t \text{ even} \end{cases} \tag{A.3}$$

We see that the probability depends slightly on whether t is an odd or an even number. This is because an odd number of initial adopters provides one additional step before the disappearance of action B from the society.

Proposition 6. *If the number t of initial adopters is **even** then:*

- *If $r < 1$ then target the middle, i.e. $s_1 = s_2$ if s even, or $s_1 = s_2 - 1$ if s odd.*
- *If $r > 1$ then target a corner, i.e. $s_1 = 0$.*
- *If $r = 1$ then the probability does not depend on s_1.*

Proposition 7. *If the number t of initial adopters is **odd** then:*

- *If $r < 1$ then target the middle, i.e. $s_1^* = s_2$ if s even, or $s_1^* = s_2 - 1$ if s odd.*
- *If $r = 1$ then target the middle, i.e. $s_1^* = s_2$ if s even, or $s_1^* = s_2 - 1$ if s odd.*
- *If $r > 1$ then $s_1^* \in \{\lfloor g(r,t) \rfloor, \lceil g(r,t) \rceil\}$[18], for $g(r,t) = \frac{\ln[r^{1/2}+(r-1)^{1/2}]}{\ln r} - \frac{t}{2}$. If $P_B(s_1 = \lfloor g(r,t) \rfloor) > P_B(s_1 = \lceil g(r,t) \rceil)$ then $s_1^* = \lfloor g(r,t) \rfloor$ and vice versa.*

These two propositions clarify the difference between having groups with odd or even number of agents. We see that for an odd number of agents, the exposition of the results is slightly more complicated, without providing additional insights. The following corollary provides some more concrete results regarding the cases where t is odd.

Corollary 1. *For t being an **odd** number*

1. *If $t = 1$ it is never optimal to target the corner.*
2. *For $t \geq 3$ it is optimal to target the corner whenever $r \geq 1.618$.*
3. *For all $t \geq 3$, there exists $\hat{r} \leq 1.618$ such that for $r > \hat{r}$ it is optimal to target the corner.*

In particular, it seems that the results are substantially different than for t even only when the number of targeted agents is small. In particular, if we can target only one agent, we never want her to be in the corner of the network. This happens because when we target only one agent, she is never safe for more than one period, meaning that a failure in the first period leads to the disappearance of the action. After a positive shock an agent located in the corner can affect the choice of only one additional agent, instead of two. Nevertheless, the problem becomes unimportant when the number of targeted agents is sufficiently large or r is sufficiently high.

[18] $\lfloor g(r,t) \rfloor = \max\{m \in \mathbb{Z} | m \leq g(r,t)\}$ is the floor function of g and $\lceil g(r,t) \rceil = \max\{m \in \mathbb{Z} | m \geq g(r,t)\}$ is the ceiling function of g

Appendix A.6. The Star

The *star* network is a very special case, because there is a unique agent -*the center*- who performs as hub for the information transmission in the network. She observes the actions and outcomes of all the other -*peripheral*- agents, while everyone else observes only her. Formally, $N_1 = \{1, \ldots, n\} = N$ while $N_i = \{1, i\}$ for all $i \in \{2, \ldots, n\}$. This extreme form of centrality turns out to be crucial in the present setting, making always optimal to target the central agent (independently of the value of r).

Namely, if the central agent (call it "agent 1", for simplicity) is targeted together with l more peripheral agents, then $P_B(1, l) = \frac{p}{1-p(1-p)}$,[19] which is bounded below by p and does not depend on l, as long as $l > 0$. On the other hand, if the central agent is not targeted, then for any number l' of targeted peripheral agents $P_B(not1, l') = pP_B(1, l') = \frac{p^2}{1-p(1-p)}$ which is bounded above by p and again does not depend on l'.

Hence, it is apparent that it is always optimal to target the central agent. Targeting, also, a peripheral agent increases the probability of diffusion, because it "secures" the action from disappearing in case of a failure in the first period. Targeting more than one peripheral agents does not improve the chances of successful diffusion, since all of them would transmit to the central agent information that she is already aware of.

Appendix B. Proofs

Proof of Proposition 2. Under (A1), $\frac{s}{m}$ and $\frac{t}{m}$ are even numbers. Then, the process is equivalent to having a line of $\frac{n}{2m}$ agents, consisting of one group of $\frac{t}{2m}$ adjacent agents choosing B and another group of $\frac{s}{2m}$ adjacent agents choosing A.

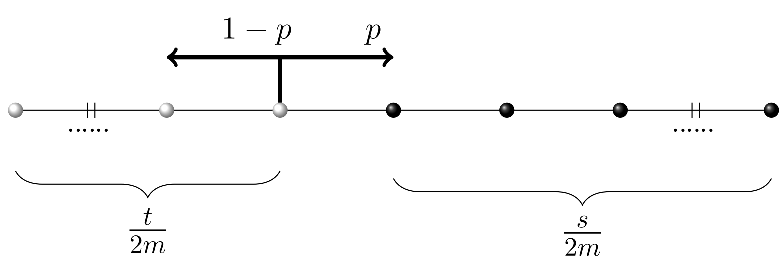


Figure B.16: The random walk that describes the process in the symmetric case.

By Lemma 1, the probability of successful diffusion becomes:

$$P_B(m|s, t, n, r) = \frac{r^{\frac{n}{2m}} - r^{\frac{s}{2m}}}{r^{\frac{n}{2m}} - 1}$$

Despite the fact, that we are interested only in the integer values of m, t and n, the function $P_B(\cdot)$ is well-defined and smooth for all $r \neq 1$ and $m \geq 1$. Hence, we can

[19] $P_B(1, l) = p + (1-p)P_B(not1, l) = p + (1-p)pP_B(1, l) \Rightarrow P_B(1, l) = \frac{p}{1-p(1-p)}$

check its monotonicity by differentiating over m.

$$
\begin{aligned}
\frac{dP_B}{dm} &= \frac{\left[r^{\frac{n}{2m}} \ln r\left(-\frac{n}{2m^2}\right) - r^{\frac{s}{2m}} \ln r\left(-\frac{s}{2m^2}\right)\right]\left(r^{\frac{n}{2m}}-1\right) - \left(r^{\frac{n}{2m}} - r^{\frac{s}{2m}}\right)\left[r^{\frac{n}{2m}} \ln r\left(-\frac{n}{2m^2}\right)\right]}{\left(r^{\frac{n}{2m}}-1\right)^2} \\
&= \frac{\ln r}{\left(r^{\frac{n}{2m}}-1\right)^2}\left[\left(-\frac{n}{2m^2} r^{\frac{n}{2m}} + \frac{s}{2m^2} r^{\frac{s}{2m}}\right)\left(r^{\frac{n}{2m}}-1\right) - \left(r^{\frac{n}{2m}} - r^{\frac{s}{2m}}\right)\left(-\frac{n}{2m^2} r^{\frac{n}{2m}}\right)\right] \\
&= \frac{\ln r}{\left(r^{\frac{n}{2m}}-1\right)^2}\left(-\frac{n}{2m^2} r^{\frac{2n}{2m}} + \frac{s}{2m^2} r^{\frac{s+n}{2m}} + \frac{n}{2m^2} r^{\frac{n}{2m}} - \frac{s}{2m^2} r^{\frac{s}{2m}} + \frac{n}{2m^2} r^{\frac{2n}{2m}} - \frac{n}{2m^2} r^{\frac{s+n}{2m}}\right) \\
&= \frac{\ln r}{2m^2\left(r^{\frac{n}{2m}}-1\right)^2}\left(s r^{\frac{s+n}{2m}} + n r^{\frac{n}{2m}} - s r^{\frac{s}{2m}} - n r^{\frac{s+n}{2m}}\right) \\
&= \frac{\ln r}{2m^2\left(r^{\frac{n}{2m}}-1\right)^2}\left[s r^{\frac{s}{2m}}\left(r^{\frac{n}{2m}}-1\right) - n r^{\frac{n}{2m}}\left(r^{\frac{s}{2m}}-1\right)\right] \\
&= \frac{\ln r}{2m^2\left(r^{\frac{n}{2m}}-1\right)^2}\left(r^{\frac{n}{2m}}-1\right)\left(r^{\frac{s}{2m}}-1\right)\left(\frac{s r^{\frac{s}{2m}}}{r^{\frac{s}{2m}}-1} - \frac{n r^{\frac{n}{2m}}}{r^{\frac{n}{2m}}-1}\right)
\end{aligned}
$$

If we call $\frac{s}{2m} = s'$ and $\frac{n}{2m} = n'$, then the following lemma helps us conclude the argument.

Lemma 3. *$f(x) = \frac{2mxr^x}{r^x-1}$ is strictly increasing for $x \geq 1$, for all $r \neq 1$ and $m \geq 1$*

Proof. Let $r \neq 1$ and $m \geq 1$, then

$$
\begin{aligned}
\frac{df}{dx} &= \frac{2m}{(r^x-1)^2}[(r^x + xr^x \ln r)(r^x-1) - xr^x(r^x)\ln r] \\
&= \frac{2m}{(r^x-1)^2}(r^{2x} - r^x - xr^x \ln r) = \frac{2mr^x}{(r^x-1)^2}(r^x - 1 - x\ln r) > 0 \text{ for all } x \geq 1
\end{aligned}
$$

To show this, we define $g(x) = r^x - 1 - x\ln r$, which is strictly increasing for $x \geq 1$ because $\frac{dg}{dx} = r^x \ln r - \ln r = \ln r(r^x - 1) > 0$. So it attains minimum for $x = 1$, which is $g(1) = r - 1 - \ln r$. Moreover, $g(1) > 0$ for all $r \neq 1$ because it holds that $h(r) = r - 1 - \ln r > 0$ for all $r \neq 1$. This is because $\frac{dh}{dr} = 1 - \frac{1}{r}$ is strictly positive when $r > 1$ and strictly negative when $r < 1$. So, h attains global minimum for $r = 1$, the $h(1) = 0$. Hence, $g(x) > 0$ for all $x \geq 1$, which means that also $\frac{df}{dx} > 0$ for all $x \geq 1$ and this concludes the argument. □

By Lemma 3, given that $n > s$, we get that $\left(\frac{sr^{\frac{s}{2m}}}{r^{\frac{s}{2m}}-1} - \frac{nr^{\frac{n}{2m}}}{r^{\frac{n}{2m}}-1}\right) < 0$ always, so we can conclude that $\frac{dP_B}{dm} < 0$ if $r > 1$ and $\frac{dP_B}{dm} > 0$ if $r < 1$. Hence, for $r > 1$ the $P_B(m|s,t,n,r)$ is decreasing in m, so $\arg\max_m P_B(m|s,t,n,r) = 1$, i.e. the optimal choice is to target a single group of initial adopters. On the other hand, for $r < 1$, P is increasing in m, so we would like to split the initial adopters in as many groups as possible, i.e. $\arg\max_m P_B(m|s,t,n,r) = \min\{s/2, t/2\}$. □

Proof of Proposition 3. First, we have to construct the probability of successful diffusion. For $r \neq 1$, the process again can be described as a sequence of random walks with absorbing barriers. At the beginning, we have a random walk of $(s_1+t_1)/2$ nodes, starting from node $t_1/2$, until it disappears either t_1 or s_1. In case of successful absorption we get a random walk of $n/2$ nodes starting from the node $(t+2s_1)/2$. Otherwise, in case of unsuccessful absorption we get a random walk of $n/2$ nodes as well, but starting from node $(t-2t_1)/2$. Notice that (A1) solves all the problems of divisibility.

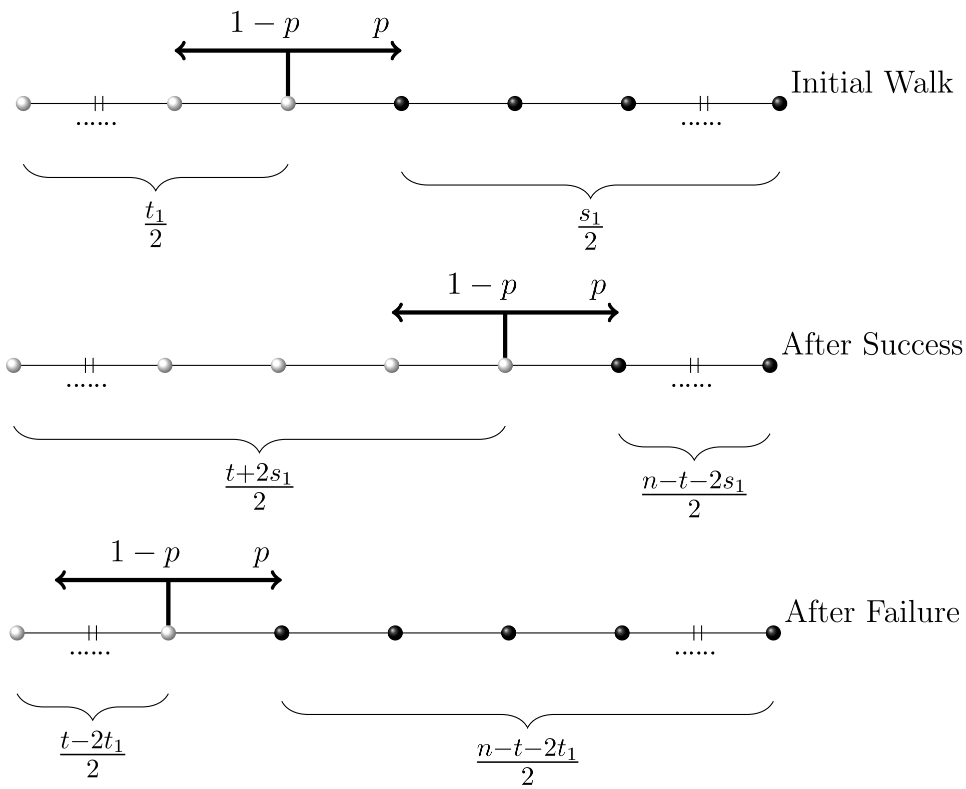


Figure B.17: The random walks that describe the process in the asymmetric case with two groups.

Hence, by Lemma 1, the probability of successful diffusion for $r \neq 1$ is:

$$P_B(s_1, t_1|s, t, n, r) = \frac{r^{\frac{s_1+t_1}{2}} - r^{\frac{s_1}{2}}}{r^{\frac{s_1+t_1}{2}} - 1} \frac{r^{\frac{n}{2}} - r^{\frac{n-t-2s_1}{2}}}{r^{\frac{n}{2}} - 1} + \frac{r^{\frac{s_1}{2}} - 1}{r^{\frac{s_1+t_1}{2}} - 1} \frac{r^{\frac{n}{2}} - r^{\frac{s+2t_1}{2}}}{r^{\frac{n}{2}} - 1}$$

Now, we compute the derivatives with respect to t_1 and s_1. As usually, we are only interested in integer points, but the function P is a well-behaved smooth function for $r \neq 1$, so we can study its monotonicity.

$$
\begin{aligned}
\frac{\partial P_B}{\partial s_1} &= \left[\frac{r^{\frac{n}{2}}\left(r^{\frac{t_1}{2}}-1\right)}{r^{\frac{n}{2}}-1}\right]\left[\frac{\left(\frac{\ln r}{2}r^{\frac{s_1}{2}}+\frac{\ln r}{2}r^{\frac{-s_1-t_1}{2}}\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)-\left(r^{\frac{s_1}{2}}-r^{\frac{-s_1-t_1}{2}}\right)\frac{\ln r}{2}r^{\frac{s_1+t_1}{2}}}{\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\right]+ \\
&+\left(\frac{r^{\frac{n}{2}}-r^{\frac{s+2t_1}{2}}}{r^{\frac{n}{2}}-1}\right)\left[\frac{\frac{\ln r}{2}r^{\frac{s_1}{2}}\left(r^{\frac{s_1+t_1}{2}}-1\right)-\left(r^{\frac{s_1}{2}}-1\right)\frac{\ln r}{2}r^{\frac{s_1+t_1}{2}}}{\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\right]= \\
&=\frac{r^{\frac{n}{2}}\left(r^{\frac{t_1}{2}}-1\right)\ln r}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(r^{\frac{2s_1+t_1}{2}}+r^{\frac{-t+t_1}{2}}-r^{\frac{s_1}{2}}-r^{\frac{-t-s_1}{2}}-r^{\frac{s_1+t_1}{2}}+r^{\frac{-t+t_1}{2}}\right)+ \\
&+\frac{\left(r^{\frac{n}{2}}-r^{\frac{s+2t_1}{2}}\right)\ln r}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(r^{\frac{2s_1+t_1}{2}}-r^{\frac{s_1}{2}}-r^{\frac{2s_1+t_1}{2}}+r^{\frac{s_1+t_1}{2}}\right)= \\
&=\frac{\ln r r^{\frac{n}{2}}\left(r^{\frac{t_1}{2}}-1\right)\left[2r^{\frac{-t+t_1}{2}}-r^{\frac{s_1}{2}}-r^{\frac{-t-s_1}{2}}+\left(r^{\frac{n}{2}}-r^{\frac{s+2t_1}{2}}\right)\left(r^{\frac{s_1+t_1}{2}}-r^{\frac{s_1}{2}}\right)\right]}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}= \\
&=\frac{\ln r\left(r^{\frac{t_1}{2}}-1\right)\left(2r^{\frac{n-t+t_1}{2}}-r^{\frac{n+s_1}{2}}-r^{\frac{n-t-s_1}{2}}+r^{\frac{n+s_1}{2}}-r^{\frac{s+s_1+2t_1}{2}}\right)}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}= \\
&=\frac{\ln r\left(r^{\frac{t_1}{2}}-1\right)r^{\frac{s-s_1}{2}}\left(2r^{\frac{s_1+t_1}{2}}-1-r^{s_1+t_1}\right)}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}= \\
&=\frac{\ln r\left(r^{\frac{t_1}{2}}-1\right)r^{\frac{s-s_1}{2}}\left[-\left(r^{\frac{s_1+t_1}{2}}-1\right)^2\right]}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}=-\frac{\ln r\left(r^{\frac{t_1}{2}}-1\right)r^{\frac{s-s_1}{2}}}{2\left(r^{\frac{n}{2}}-1\right)}\begin{cases}<0 & \text{if } r>1\\ >0 & \text{if } r<1\end{cases}
\end{aligned}
$$

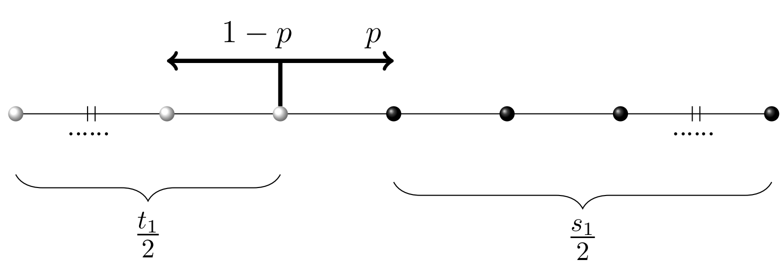
Analogously for t_1 we have:

$$\begin{aligned}
\frac{\partial P_B}{\partial t_1} &= \frac{\ln r\left[(r^{\frac{n+s_1}{2}}-r^{\frac{s-s_1}{2}})r^{\frac{t_1}{2}}(r^{\frac{s_1+t_1}{2}}-1)-(r^{\frac{n+s_1}{2}}-r^{\frac{s-s_1}{2}})(r^{\frac{t_1}{2}}-1)r^{\frac{s_1+t_1}{2}}\right]}{2(r^{\frac{n}{2}}-1)(r^{\frac{s_1+t_1}{2}}-1)^2}+ \\
&+\frac{\ln r\left[(r^{\frac{s_1}{2}}-1)(-2r^{\frac{s+2t_1}{2}})(r^{\frac{s_1+t_1}{2}}-1)-(r^{\frac{s_1}{2}}-1)(r^{\frac{n}{2}}-r^{\frac{s+2t_1}{2}})r^{\frac{s_1+t_1}{2}}\right]}{2(r^{\frac{n}{2}}-1)(r^{\frac{s_1+t_1}{2}}-1)^2}= \\
&=\frac{\ln r}{2(r^{\frac{n}{2}}-1)(r^{\frac{s_1+t_1}{2}}-1)^2}\left(r^{\frac{n+s_1}{2}}-r^{\frac{s-s_1}{2}}\right)\left(r^{\frac{s_1+2t_1}{2}}-r^{\frac{t_1}{2}}-r^{\frac{s_1+2t_1}{2}}+r^{\frac{s_1+t_1}{2}}\right)+ \\
&+\frac{\ln r}{2(r^{\frac{n}{2}}-1)(r^{\frac{s_1+t_1}{2}}-1)^2}\left(r^{s_1/2}-1\right)\left(-2r^{(s+s_1+3t_1)/2}+2r^{\frac{s+2t_1}{2}}-r^{\frac{n+s_1+t_1}{2}}+r^{\frac{s+s_1+3t_1}{2}}\right)= \\
&=\frac{\ln r\left(r^{\frac{s_1}{2}}-1\right)\left(r^{\frac{n+s_1+t_1}{2}}-r^{\frac{s+t_1-s_1}{2}}-2r^{\frac{s+s_1+3t_1}{2}}+2r^{\frac{s+2t_1}{2}}-r^{\frac{n+s_1+t_1}{2}}+r^{\frac{s+s_1+3t_1}{2}}\right)}{2(r^{\frac{n}{2}}-1)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}= \\
&=\frac{\ln r\left(r^{\frac{s_1}{2}}-1\right)r^{\frac{s+t_1}{2}}}{2r^{\frac{s_1}{2}}\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(2r^{\frac{s_1+t_1}{2}}+r^{\frac{2s_1+2t_1}{2}}-2r^{\frac{2s_1+2t_1}{2}}-1\right)= \\
&=\frac{\ln r\left(r^{\frac{s_1}{2}}-1\right)r^{\frac{s+t_1}{2}}\left[-\left(r^{\frac{s_1+t_1}{2}}-1\right)^2\right]}{2r^{\frac{s_1}{2}}\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}=-\frac{\ln r\left(r^{\frac{s_1}{2}}-1\right)r^{\frac{s+t_1}{2}}}{2r^{\frac{s_1}{2}}\left(r^{\frac{n}{2}}-1\right)}\begin{cases}<0 & \text{if } r>1\\ >0 & \text{if } r<1\end{cases}
\end{aligned}$$

Hence, given that $0 \le s_1 \le s_2$ and $0 \le t_1 \le t_2$ we conclude that for $r > 1$ the optimal targeting strategy is $(s_1, t_1) = (0, 0)$, whereas for $r < 1$ it is $s_2 - s_1 \le 2$ and $t_2 - t_1 \le 2$.

□

Proof of Theorem 1. For the case of $r > 1$ we proceed by induction. First, we recall the result by Proposition 2, which states that if we can target up to two groups, then the optimal choice is to concentrate all the initial adopters in one group. Remember also that $s_1 \le s_2 \le s_3$ and $t_1 \le t_2 \le t_3$. Now suppose that we can target up to three groups ($m \le 3$). Then again at first we are interested in the two smallest groups of each type and we have the following random walk.



The system fluctuates in this direction until either s_1 or t_1 disappears. Depending on the successful or unsuccessful absorption of this process we get one of the following two configurations (Figure A.17), with only two groups remaining of each type.

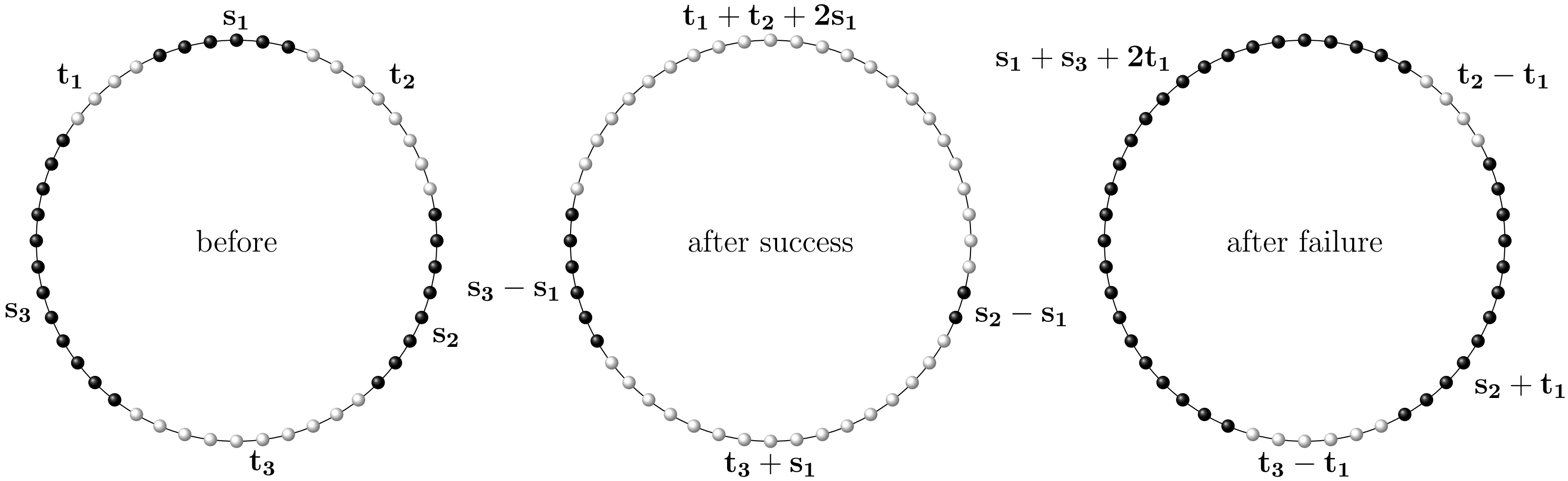


Figure B.18: Configurations with 3 groups of adopters.

By Proposition 2, in both of these cases we know that the optimal choice would be to eliminate one of the two groups of initial adopters. Hence, we would like to choose s_2 and t_2 (as functions of s_1 and t_1 respectively), in such a way that the probability of diffusion is maximized in both of these cases. Recalling that $s_1 \leq s_2 \leq s_3$ and $t_1 \leq t_2 \leq t_3$, we see that this can be achieved if $s_2 = s_1$ and $t_2 = t_1$, where the optimal s_1 and t_1 remained to be determined. Notice that, by construction, $s_3 = s - s_1 - s_2$ and $t_3 = t - t_1 - t_2$.
So now, we can rewrite the probability of diffusion as a function of s_1 and t_1 only.

$$P_B(s_1, t_1|s, t, n, r, m = 3) = \frac{r^{\frac{t_1+s_1}{2}} - r^{\frac{s_1}{2}}}{r^{\frac{t_1+s_1}{2}} - 1} \frac{r^{\frac{n}{2}} - r^{\frac{n-t-3s_1}{2}}}{r^{\frac{n}{2}} - 1} + \frac{r^{\frac{s_1}{2}} - 1}{r^{\frac{t_1+s_1}{2}} - 1} \frac{r^{\frac{n}{2}} - r^{\frac{s+3t_1}{2}}}{r^{\frac{n}{2}} - 1}$$

Like before we study the monotonicity of the function with respect to s_1 and t_1

$$
\begin{aligned}
\frac{\partial P_B}{\partial s_1} &= \frac{\ln r}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left[\left(r^{\frac{t_1}{2}}-1\right)\left(r^{\frac{n+s_1}{2}}+2r^{\frac{s-2s_1}{2}}\right)+\left(r^{\frac{n}{2}}-r^{\frac{s+3t_1}{2}}\right)r^{\frac{s_1}{2}}\right]\left(r^{\frac{s_1+t_1}{2}}-1\right)- \\
&-\frac{\ln r}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left[\left(r^{\frac{t_1}{2}}-1\right)\left(r^{\frac{n+s_1}{2}}-r^{\frac{s-2s_1}{2}}\right)+\left(r^{\frac{n}{2}}-r^{\frac{s+3t_1}{2}}\right)\left(r^{\frac{s_1}{2}}-1\right)\right]r^{\frac{s_1+t_1}{2}}= \\
&=\frac{\ln r}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(r^{\frac{t_1}{2}}-1\right)\left[\left(r^{\frac{n+s_1}{2}}+2r^{\frac{s-2s_1}{2}}\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)-\left(r^{\frac{n+s_1}{2}}-r^{\frac{s-2s_1}{2}}\right)r^{\frac{s_1+t_1}{2}}\right]+ \\
&+\frac{\ln r}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(r^{\frac{n}{2}}-r^{\frac{s+3t_1}{2}}\right)\left[r^{\frac{s_1}{2}}\left(r^{\frac{s_1+t_1}{2}}-1\right)-\left(r^{\frac{s_1}{2}}-1\right)r^{\frac{s_1+t_1}{2}}\right]= \\
&=\frac{\ln r\left(3r^{\frac{s+2t_1-s_1}{2}}-3r^{\frac{s+t_1-s_1}{2}}+2r^{\frac{s-2s_1}{2}}-2r^{\frac{s+t_1-2s_1}{2}}+r^{\frac{s+s_1+3t_1}{2}}-r^{\frac{s+s_1+4t_1}{2}}\right)}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}= \\
&=\frac{\ln r}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(r^{\frac{t_1}{2}}-1\right)\left(3r^{\frac{s+t_1-s_1}{2}}-2r^{\frac{s-2s_1}{2}}-r^{\frac{s+s_1+3t_1}{2}}\right)= \\
&=\frac{r^{\frac{s}{2}}\ln r}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(r^{\frac{t_1}{2}}-1\right)\left(3r^{\frac{t_1-s_1}{2}}-2r^{\frac{-2s_1}{2}}-r^{\frac{s_1+3t_1}{2}}\right)= \\
&=\frac{r^{\frac{s}{2}}\ln r}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(r^{\frac{t_1}{2}}-1\right)\left[-\frac{\left(r^{\frac{s_1+t_1}{2}}-1\right)^2\left(r^{\frac{s_1+t_1}{2}}+2\right)}{r^{s_1}}\right]= \\
&=-\frac{r^{\frac{s}{2}}\ln r\left(r^{\frac{t_1}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}+2\right)}{2\left(r^{\frac{n}{2}}-1\right)r^{s_1}}<0 \text{ for } r>1
\end{aligned}
$$

Hence the optimal choice is $s_1=s_2=0$ and $s_3=s$.

Analogously for t_1 we get the following:

$$\frac{\partial P_B}{\partial t_1} = \frac{\ln r}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(r^{\frac{n+s_1}{2}}-r^{\frac{s-s_1}{2}}\right)\left[r^{\frac{t_1}{2}}\left(r^{\frac{s_1+t_1}{2}}-1\right)-r^{\frac{s_1+t_1}{2}}\left(r^{\frac{t_1}{2}}-1\right)\right]+$$

$$+\frac{\ln r}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(r^{\frac{s_1}{2}}-1\right)\left[-3r^{\frac{s+3t_1}{2}}\left(r^{\frac{s_1+t_1}{2}}-1\right)-\left(r^{\frac{n}{2}}-r^{\frac{s+3t_1}{2}}\right)r^{\frac{s_1+t_1}{2}}\right]=$$

$$=\ln r\frac{\left(r^{\frac{n+s_1}{2}}-r^{\frac{s-2s_1}{2}}\right)\left(r^{\frac{s_1+t_1}{2}}-r^{\frac{t_1}{2}}\right)+\left(r^{\frac{s_1}{2}}-1\right)\left(3r^{\frac{s+3t_1}{2}}-2r^{\frac{s+s_1+4t_1}{2}}-r^{\frac{n+s_1+t_1}{2}}\right)}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}=$$

$$=\frac{\ln r\left(r^{\frac{s_1}{2}}-1\right)}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(3r^{\frac{s+3t_1}{2}}-2r^{\frac{s+s_1+4t_1}{2}}-r^{\frac{s+t_1-2s_1}{2}}\right)=$$

$$=\frac{\ln r\left(r^{\frac{s_1}{2}}-1\right)r^{\frac{s}{2}}r^{\frac{t_1}{2}}}{2\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(3r^{\frac{2t_1}{2}}-2r^{\frac{s_1+3t_1}{2}}-r^{\frac{-2s_1}{2}}\right)=$$

$$=\frac{\ln r\left(r^{\frac{s_1}{2}}-1\right)r^{\frac{s}{2}}r^{\frac{t_1}{2}}}{2r^{s_1}\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}\left(3r^{\frac{2s_1+2t_1}{2}}-2r^{\frac{3s_1+3t_1}{2}}-1\right)=$$

$$=\frac{\ln r\left(r^{\frac{s_1}{2}}-1\right)r^{\frac{s}{2}}r^{\frac{t_1}{2}}\left(r^{\frac{s_1+t_1}{2}}-1\right)\left(2r^{2\left(\frac{s_1+t_1}{2}\right)}-r^{\frac{s_1+t_1}{2}}+1\right)}{2r^{s_1}\left(r^{\frac{n}{2}}-1\right)\left(r^{\frac{s_1+t_1}{2}}-1\right)^2}<0 \text{ for } r>1.$$

Notice that $2r^{2(\frac{s_1+t_1}{2})}-r^{\frac{s_1+t_1}{2}}+1=r^{2(\frac{s_1+t_1}{2})}+r^{\frac{s_1+t_1}{2}}+(r^{\frac{s_1+t_1}{2}}-1)^2$ which is strictly positive for s_1 and t_1 positive. Hence, P_B is always decreasing in t_1, and given that $t_1=t_2$ the optimal choice is $t_1=t_2=0$ and $t_3=t$. This concludes the argument for the case where $m=3$. We will generalize this argument by induction.

Formally, given that the argument holds for $m=3$, it suffices to show that if it holds for $m=k-1\geq 3$ then it holds as well for $m=k$.

At the beginning of the process we care only about the two smallest groups of each type s_1 and t_1 and the system fluctuates as in the previous cases until one of the two disappears. Figure A.18 shows the possible configurations after the disappearance of either s_1 or t_1. The location of the groups around the network comes without loss of generality.

Given that the argument holds for $k-1$ groups then we know that $s_1=\cdots=s_{k-1}$ and $t_1=\cdots=t_{k-1}$. Therefore, we only need to find the optimal s_1 and t_1. The probability of diffusion becomes:

$$P_B(s_1,t_1|s,t,n,r,m=k)=\frac{r^{\frac{t_1+s_1}{2}}-r^{\frac{s_1}{2}}}{r^{\frac{t_1+s_1}{2}}-1}\frac{r^{\frac{n}{2}}-r^{\frac{n-t-ks_1}{2}}}{r^{\frac{n}{2}}-1}+\frac{r^{\frac{s_1}{2}}-1}{r^{\frac{t_1+s_1}{2}}-1}\frac{r^{\frac{n}{2}}-r^{\frac{s+kt_1}{2}}}{r^{\frac{n}{2}}-1}$$

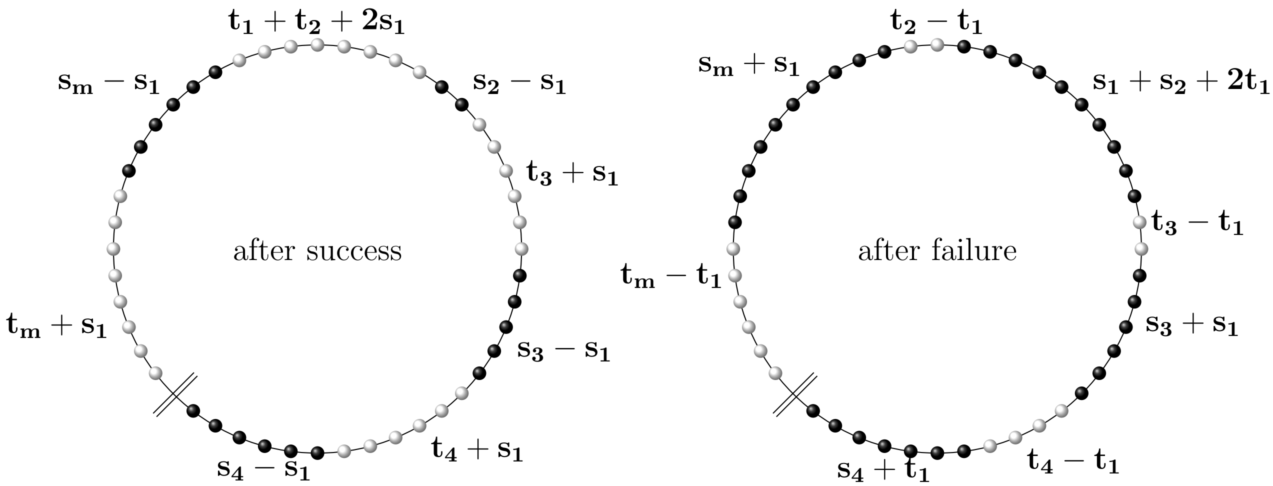


Figure B.19: Configurations after the disappearance of s_1 or t_1 with m groups.

By some calculations which are omitted because they are identical to the case where $m = 3$, we get:

$$\frac{\partial P_B}{\partial s_1} = \frac{r^{\frac{s}{2}}\left(r^{\frac{t_1}{2}} - 1\right)\ln r}{2r^{(k-1)\frac{s_1}{2}}\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2}\left[kr^{\frac{s_1+t_1}{2}} - r^{k\frac{s_1+t_1}{2}} - (k-1)\right] \leq 0 \text{ for } r > 1$$

and equality holds only if $s_1 = t_1 = 0$. For the argument to hold we need $kr^{\frac{s_1+t_1}{2}} - r^{k\frac{s_1+t_1}{2}} - (k-1)$ to be negative. So, let $x = r^{\frac{s_1+t_1}{2}}$ and take the function $f(x) = kx - x^k - (k-1)$ for some $k \geq 3$ and $x \geq 0$. Now, $\frac{df}{dx} = k - kx^{k-1}$ is positive if $x < 1$ and negative if $x > 1$, hence f attains global max at $x = 1$ equal to $f(1) = k - 1^k - (k-1) = 0$, hence $f(x) < 0$ for all $x \neq 1$. Now given that $x = r^{\frac{s_1+t_1}{2}}$, with $r > 1$ and $s_1, t_1 \geq 0$ the function is always strictly negative and becomes equal to zero only when $s_1 = t_1 = 0$. So, the optimal choice is $s_1 = \cdots = s_{k-1} = 0$ and $s_k = s$.

Analogously for t_1 we get that:

$$\frac{\partial P_B}{\partial s_1} = \frac{r^{\frac{s}{2}}r^{\frac{t_1}{2}}\left(r^{\frac{s_1}{2}} - 1\right)\ln r}{2r^{(k-1)\frac{s_1}{2}}\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_1+t_1}{2}} - 1\right)^2}\left[kr^{(k-1)\frac{s_1+t_1}{2}} - (k-1)r^{k\frac{s_1+t_1}{2}} - 1\right] \leq 0 \text{ for } r > 1$$

again equality holds only when $s_1 = t_1 = 0$ and to ensure the result we need that $kr^{(k-1)\frac{s_1+t_1}{2}} - (k-1)r^{k\frac{s_1+t_1}{2}} - 1 \leq 0$ for all s_1 and t_1 with equality holding only in case they are both equal to zero. As before, let $x = r^{\frac{s_1+t_1}{2}}$ and define the function $g(x) = kx^{k-1} - (k-1)x^k - 1$. Then $\frac{df}{dx} = k(k-1)x^{k-2} - k(k-1)x^{k-1}$ which is strictly negative for $x > 1$ and strictly positive for $x < 1$, then g attains unique maximum at $x = 1$ equal to $g(1) = 0$. So $g(x) < 0$ for all $x \neq 1$. Given again that $x = r^{\frac{s_1+t_1}{2}}$ then $x = 1$ only if $s_1 = t_1 = 0$. So again the optimal choices are $t_1 = \cdots = t_{k-1} = 0$

and $t_k = t$, which completes the inductive argument.

Hence, when $r > 1$, for any possible number of groups m, the optimal choice is to concentrate all the initial adopters in one group, i.e. $s_1 = \cdots = s_{m-1} = 0$ and $s_m = s$, as well as $t_1 = \cdots = t_{m-1} = 0$ and $t_m = t$.

Now, we turn our attention towards the case where $r < 1$. We tackle this case in a different way. Namely, we construct an upper bound for the probability of successful convergence and we show that for the same configurations that this upper bound is maximized, the actual probability is equal to this upper bound. Hence this has to be the maximum value of the probability as well.

In order to proceed, we need to construct the upper bound for the value of the probability of successful diffusion of action B. We solve it first for $t < s$ and then for $s < t$.

Let $t < s$, then allowing for the existence of $m = \frac{t}{2}$ groups, the network will have the form of the following figure. Notice that, the fact that s_i can have size equal to zero, allows us to construct any possible configuration. For example, in the following figure (Figure A.19), if $s_1 = 0$ then the two groups next to s_1 merge to one group with four agents. According to this structure, the network will initially follow a random walk with $\frac{s_1}{2}$ black steps and one white. By Lemma 1, the probability of success in this first walk is equal to $\frac{r^{\frac{s_1}{2}+1} - r^{\frac{s_1}{2}}}{r^{\frac{s_1}{2}+1} - 1}$.

In Figure A.19, we see as well, how the network will look like if the first walk is successful. Unsuccessful absorption in the first walk leads to the disappearance of action B from the network, because all t's have the same size.

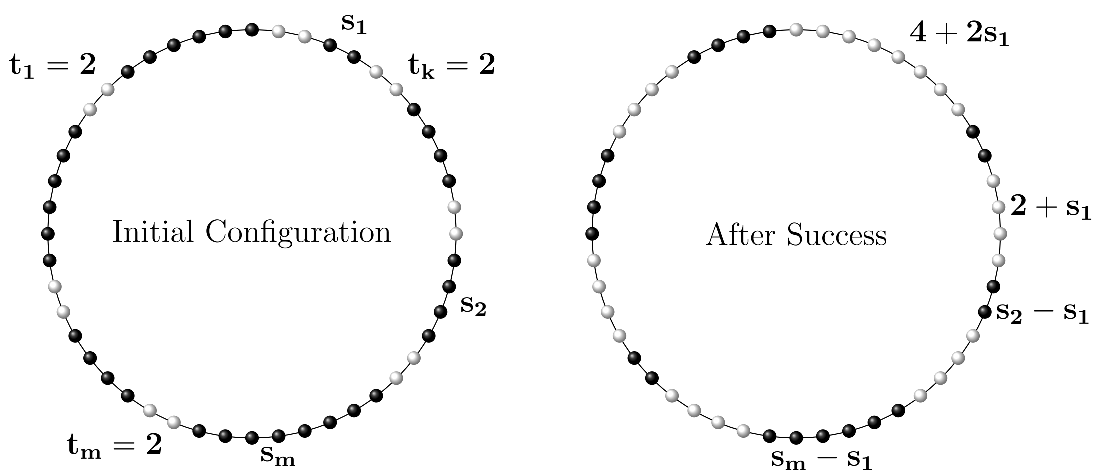
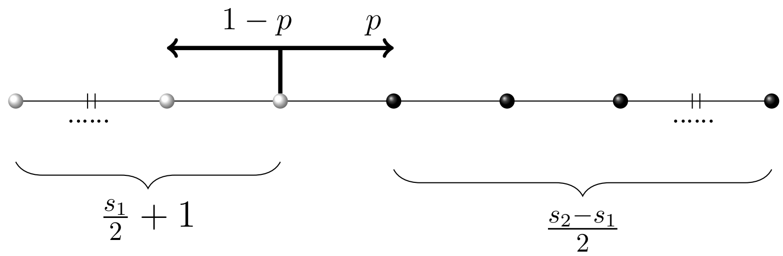


Figure B.20: General Initial Configuration and Result After Success, for $t < s$.

After success, the network will move according to the following random walk. The probability of success in this walk is equal to $\frac{r^{\frac{s_2}{2}+1} - r^{\frac{s_2 - s_1}{2}}}{r^{\frac{s_2}{2}+1} - 1}$.

In Figure A.20, we depict the two possible configurations that arise after success or failure in the second walk. It is important to notice that the probability of successful diffusion after two successes is obviously weakly lower than one and it is strictly lower than one, as long as $s_m - s_1 > 2$, where s_m is the size of the largest



group and s_1 is the size of the smallest one.

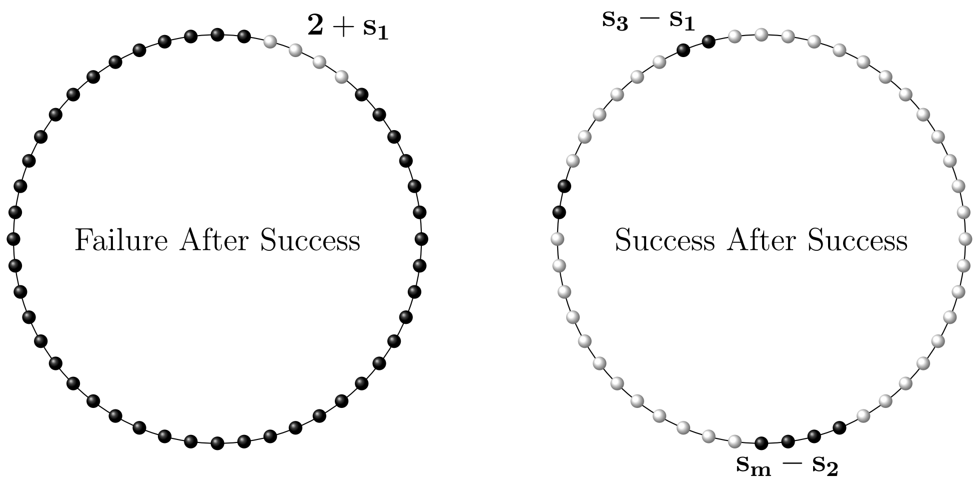


Figure B.21: Resulting Configurations given Failure or Success in the second random walk, given successful first walk, for $t < s$

Hence, we can construct the probability of successful diffusion, which is equal to:

$$P_B(\cdot) = \frac{r^{\frac{s_1}{2}+1} - r^{\frac{s_1}{2}}}{r^{\frac{s_1}{2}+1} - 1}\left[\frac{r^{\frac{s_2}{2}+1} - r^{\frac{s_2-s_1}{2}}}{r^{\frac{s_2}{2}+1} - 1}P_B(\cdot|s,s) + \frac{r^{\frac{s_2-s_1}{2}} - 1}{r^{\frac{s_2}{2}+1} - 1}\frac{r^{\frac{n}{2}} - r^{\frac{n-s_1}{2}-1}}{r^{\frac{n}{2}} - 1}\right]$$

where $P_B(\cdot|s,s)$ stands for the probability of diffusion of B after two successes in the first two random walks. Given that $P_B(\cdot|s,s) \leq 1$ we get the following upper bound of P_B, denoted by $\widetilde{P_B(\cdot)}$, which is equal to:

$$\widetilde{P_B(\cdot)} = \frac{r^{\frac{s_1}{2}+1} - r^{\frac{s_1}{2}}}{r^{\frac{s_1}{2}+1} - 1}\left[\frac{r^{\frac{s_2}{2}+1} - r^{\frac{s_2-s_1}{2}}}{r^{\frac{s_2}{2}+1} - 1} + \frac{r^{\frac{s_2-s_1}{2}} - 1}{r^{\frac{s_2}{2}+1} - 1}\frac{r^{\frac{n}{2}} - r^{\frac{n-s_1}{2}-1}}{r^{\frac{n}{2}} - 1}\right]$$

Before performing any calculations it is important to simplify the expression of $\widetilde{P_B(\cdot)}$. Specifically,

$$\widetilde{P_B(\cdot)} = \frac{r^{\frac{s_1}{2}+1} - r^{\frac{s_1}{2}}}{r^{\frac{s_1}{2}+1} - 1} \left[\frac{r^{\frac{s_2}{2}+1} - r^{\frac{s_2-s_1}{2}}}{r^{\frac{s_2}{2}+1} - 1} + \frac{r^{\frac{s_2-s_1}{2}} - 1}{r^{\frac{s_2}{2}+1} - 1} \frac{r^{\frac{n}{2}} - r^{\frac{n-s_1}{2}-1}}{r^{\frac{n}{2}} - 1} \right] =$$

$$= \frac{r-1}{r^{\frac{n}{2}} - 1} \left[\frac{\left(r^{\frac{s_1+s_2}{2}+1} - r^{\frac{s_2}{2}}\right)\left(r^{\frac{n}{2}} - 1\right) + \left(r^{\frac{n}{2}} - r^{\frac{n-s_1}{2}-1}\right)\left(r^{\frac{s_2}{2}} - r^{\frac{s_1}{2}}\right)}{\left(r^{\frac{s_1}{2}+1} - 1\right)\left(r^{\frac{s_2}{2}+1} - 1\right)} \right] =$$

$$= \frac{r-1}{r^{\frac{n}{2}} - 1} \left[\frac{\left(r^{\frac{n+s_2}{2}} + r^{\frac{n+s_2-s_1}{2}-1} - r^{\frac{s_2}{2}} - r^{\frac{n}{2}-1}\right)\left(r^{\frac{s_1}{2}+1} - 1\right)}{\left(r^{\frac{s_1}{2}+1} - 1\right)\left(r^{\frac{s_2}{2}+1} - 1\right)} \right] =$$

$$= \frac{r-1}{r^{\frac{n}{2}} - 1} \left[\frac{r^{\frac{n}{2}-1}\left(r^{\frac{s_2}{2}+1} - 1\right)}{r^{\frac{s_2}{2}+1} - 1} + \frac{r^{\frac{n+s_2-s_1}{2}-1} - r^{\frac{s_2}{2}}}{r^{\frac{s_2}{2}+1} - 1} \right] =$$

$$= \frac{r-1}{r^{\frac{n}{2}} - 1} \left[r^{\frac{n}{2}-1} + \frac{r^{\frac{n+s_2-s_1}{2}-1} - r^{\frac{s_2}{2}}}{r^{\frac{s_2}{2}+1} - 1} \right]$$

Notice that $s_2 = s - s_1 - s_3 - \cdots - s_m$, hence $\frac{\partial s_2}{\partial s_1} = -1$. And now we can differentiate $\widetilde{P_B(\cdot)}$ with respect to s_1.

$$\frac{\partial \widetilde{P_B(\cdot)}}{\partial s_1} = \frac{r-1}{r^{\frac{n}{2}} - 1} \left[\frac{\left(-r^{\frac{n+s_2-s_1}{2}-1} \ln r + \frac{\ln r}{2} r^{\frac{s_2}{2}}\right)\left(r^{\frac{s_2}{2}+1} - 1\right) - \left(r^{\frac{n+s_2-s_1}{2}-1} - r^{\frac{s_2}{2}}\right)\left(-\frac{\ln r}{2} r^{\frac{s_2}{2}+1}\right)}{\left(r^{\frac{s_2}{2}+1} - 1\right)^2} \right] =$$

$$= \frac{\ln r \,(r-1)\, r^{\frac{s_2}{2}}}{2\left(r^{\frac{n}{2}} - 1\right)\left(r^{\frac{s_2}{2}+1} - 1\right)^2} \left(2r^{\frac{n-s_1}{2}-1} - r^{\frac{n+s_2-s_1}{2}} - 1\right) > 0, \text{ for } r < 1.$$

The fact that the term $2r^{\frac{n-s_1}{2}-1} - r^{\frac{n+s_2-s_1}{2}} - 1$ is always negative is not obvious and is proven here. Substituting s_2, we can rewrite it as:

$$r^{-s_1-1}\left[2r^{\frac{n+s_1}{2}} - r^{\frac{n+s-s_3-\cdots-s_m}{2}+1} - r^{2\frac{s_1}{2}+1}\right]$$

If we denote $x = r^{\frac{s_1}{2}}$ then we get a polynomial of degree two with respect to x. The discriminant of this polynomial is equal to:

$$\Delta = 4r^{2\frac{n}{2}} - 4r^{\frac{n}{2}+\frac{s-s_3-\cdots-s_m}{2}+2} = 4r^{\frac{n}{2}}\left(r^{\frac{n}{2}} - r^{\frac{s-s_3-\cdots-s_m}{2}+2}\right) = 4r^{\frac{n}{2}}\left(r^{\frac{n}{2}} - r^{\frac{s_1+s_2}{2}+2}\right) < 0$$

Because $r < 1$ and $\frac{n}{2} > \frac{s_1+s_2}{2} + 2$ for $m \geq 3$. For $m = 2$ this holds with equality, but we have already analyzed this case. So, this polynomial has no roots and given that the factor of the quadratic term is negative $(-r)$, we can conclude that for $r < 1$ the polynomial is always negative. Therefore, $\widetilde{P_B(\cdot)}$ takes its maximum value when s_1 is maximized. For this value of s_1, the real probability of successful diffusion is equal to this upper bound as long as $s_m - s_1 \leq 2$. Therefore, remembering that

$\widetilde{P_B(\cdot)} \geq P_B(\cdot)$ always, it has to be that $P_B(\cdot)$ is also maximized for when both s_1 is maximized and $s_m - s_1 \leq 2$.

In case m divides s exactly, then the maximum of s_1 is equal to $\frac{s}{m}$ and the optimal choice is $s_1 = \cdots = s_m = \frac{s}{m}$. If m does not divide s exactly, then we have $s = mq + d$, where q is the quotient of the division and d is the remainder. In this case, P_B is maximized if we have $m - \frac{d}{2}$ groups of size $q = \frac{s-d}{m}$ and $\frac{d}{2}$ groups with size $q + 2 = \frac{s-d}{m} + 2$, so again the difference in the size of any two groups is no larger than four. We still remain to describe what is the optimal position of the groups that have the two additional agents. The result will become apparent after we analyze the case for $t > s$.

Now, we prove the result for $t > s$ in a completely analogous way. In this case, the initial configuration is as in the left part of Figure A.21. A success in the first random walk leads to the diffusion of action B, while a failure leads to a configuration as in the right part of the same figure. The probability of success in the first walk is $\frac{r^{\frac{t_1}{2}+1}-r}{r^{\frac{t_1}{2}+1}-1}$. Figure A.22 shows the possible configurations after successful or unsuccessful absorption in the second random walk, given unsuccessful absorption in the first one. The probability of success in the second walk is $\frac{r^{\frac{t_2}{2}+1}-r^{\frac{t_1}{2}+1}}{r^{\frac{t_2}{2}+1}-1}$. Therefore, we can construct again an upper bound for the probability of successful diffusion, equal to:

$$\widetilde{P_B(\cdot)} = \frac{r^{\frac{t_1}{2}+1} - r}{r^{\frac{t_1}{2}+1} - 1} + \frac{r-1}{r^{\frac{t_1}{2}+1} - 1}\left[\frac{r^{\frac{t_2}{2}+1} - r^{\frac{t_1}{2}}}{r^{\frac{t_2}{2}+1} - 1}\frac{r^{\frac{n}{2}} - r^{\frac{t_1}{2}-1}}{r^{\frac{n}{2}} - 1} + \frac{r^{\frac{t_1}{2}} - 1}{r^{\frac{t_2}{2}+1} - 1}\right]$$

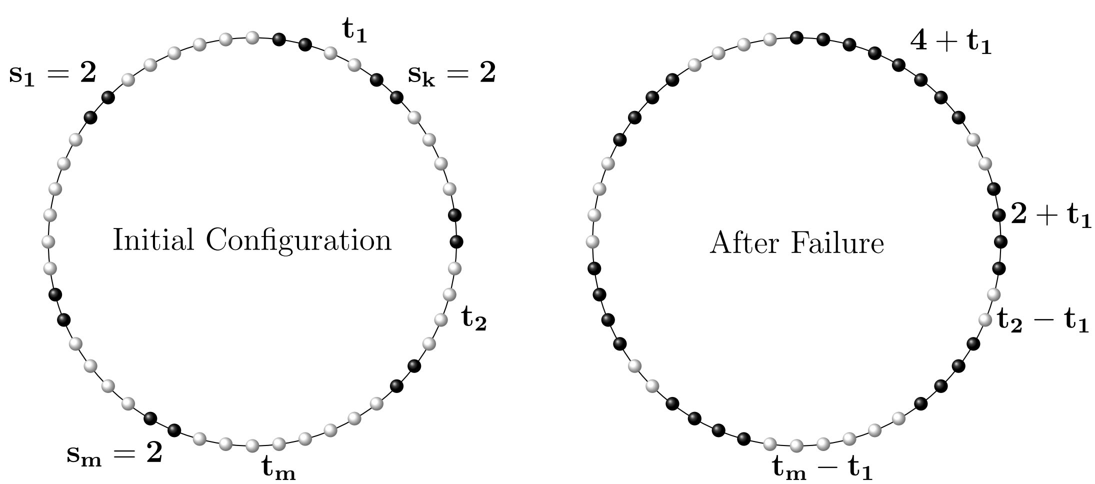


Figure B.22: General Initial Configuration and Result After Success, for $t > s$.

This expression can be transformed in a similar manner as before:

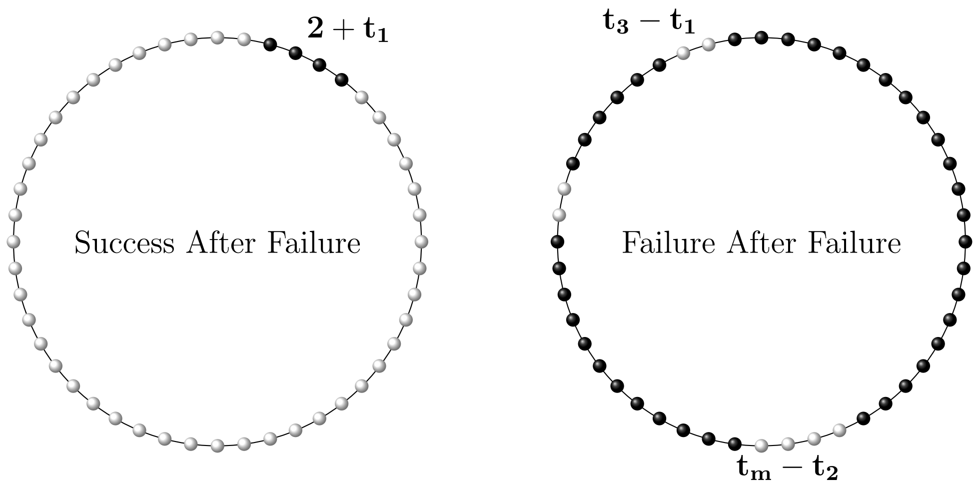


Figure B.23: Resulting configurations after success or failure in the second random walk, given failure in the first random walk, for $t > s$

$$\widetilde{P_B(\cdot)} = \frac{r^{\frac{t_1}{2}+1} - r}{r^{\frac{t_1}{2}+1} - 1} + \frac{r-1}{r^{\frac{t_1}{2}+1}-1}\left[\frac{r^{\frac{t_2}{2}+1} - r^{\frac{t_1}{2}}}{r^{\frac{t_2}{2}+1}-1}\frac{r^{\frac{n}{2}} - r^{\frac{t_1}{2}-1}}{r^{\frac{n}{2}}-1} + \frac{r^{\frac{t_1}{2}}-1}{r^{\frac{t_2}{2}+1}-1}\right] =$$
$$= \frac{1}{r^{n/2}-1}\left[\frac{(r^{\frac{t_2}{2}+1}-1)(r^{n/2}-1) - (r^{\frac{t_2}{2}+1} - r^{\frac{t_1}{2}+1})(r-1)}{r^{\frac{t_2}{2}+1}-1}\right] =$$
$$= \frac{1}{r^{n/2}-1}\left[r^{n/2} - r + (r-1)\frac{r^{\frac{t_1}{2}+1}-1}{r^{\frac{t_2}{2}+1}-1}\right]$$

Notice again, that the upper bound becomes equal to the actual probability if $t_m - t_1 \leq 2$, where t_m is the size of the largest group of type B and t_1 the smallest one.

Now, we can differentiate the expression with respect to t_1, remembering that $t_2 = t - t_1 - t_3 - \cdots - t_m$:

$$\frac{\partial \widetilde{P_B(\cdot)}}{\partial t_1} = \frac{(r-1)\ln r}{2(r^{n/2}-1)}\left[\frac{r^{\frac{t_1}{2}+1}(r^{\frac{t_2}{2}+1}-1) + (r^{\frac{t_1}{2}+1}-1)r^{\frac{t_2}{2}+1}}{(r^{\frac{t_2}{2}+1}-1)^2}\right] > 0, \text{ for all } r < 1.$$

The upper bound is increasing in t_1. For this value of t_1, the real probability of successful diffusion is equal to this upper bound as long as $t_m - t_1 \leq 2$. Therefore, remembering that $\widetilde{P_B(\cdot)} \geq P_B(\cdot)$ always, it has to be that $P_B(\cdot)$ is also maximized for when both t_1 is maximized and $t_m - t_1 \leq 2$.

In case m divides s exactly, then the maximum of t_1 is equal to $\frac{t}{m}$ and the optimal choice is $t_1 = \cdots = t_m = \frac{t}{m}$. If m does not divide t exactly, then we have $t = mq + d$, where q is the quotient of the division and d is the remainder. In this case, P_B is maximized if we have $m - \frac{d}{2}$ groups of size $q = \frac{t-d}{m}$ and $\frac{d}{2}$ groups with size $q + 2 = \frac{t-d}{m} + 2$, so again the difference in the size of any two groups is no larger than four.

To complete the proof we need to explain the optimal location of the groups which have the two additional agents. For the case where $t < s$ we need to notice that after successful absorption in the first random walk, now the network consists of $\frac{d}{2}$ groups of each type, where all the groups of type A have exactly two agents. Hence, we fall into the analysis of the case where $t > s$, where we would like the groups of type B to be as equal as possible. In order to succeed this we should have located the groups of type A with more agents as symmetrically as possible around the network.

An example can be illustrated in Figure A.23. We have targeted 14 out of 48 agents, having seven groups of two agents of type B, three groups of six agents and four groups of four agents of type A. After successful absorption in the first walk, there will be left only three groups of two agents of type A, which we want to be located as symmetrically as possible. For this reason we do not put two groups of six agents one next to the other in the initial configuration. However, notice that we cannot make the configuration arising after success totally symmetric, due to the restriction on the sizes of the groups. But again we want it to be as symmetric as possible, by maximizing the smallest group and minimizing its difference with the largest one. The argument for the case where $t > s$ is completely analogous.

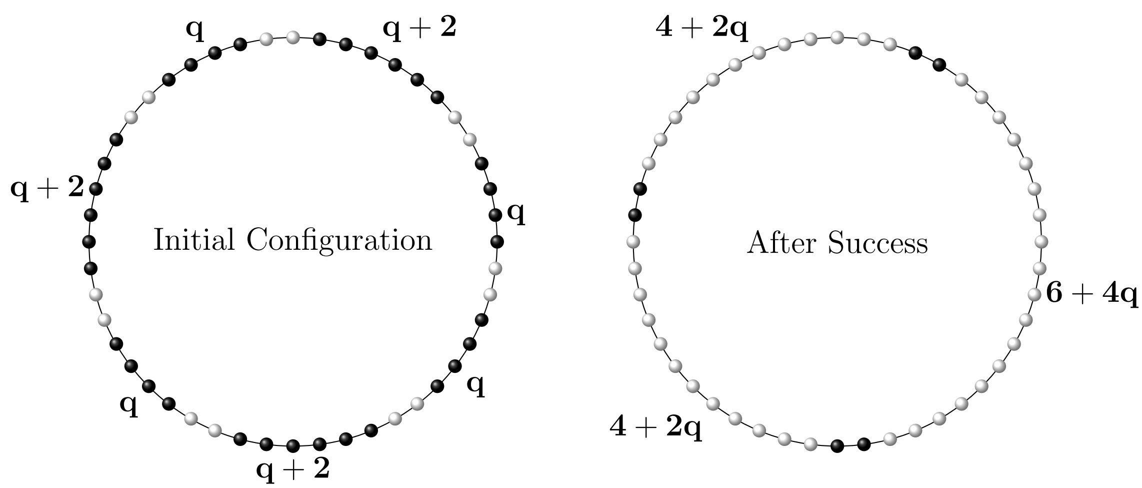


Figure B.24: Optimal Initial Configuration with $s = mq + d$, for $s > t$.

□

Proof of Proposition 4- Decreasing Returns to Scale. By Theorem 1, the optimal targeting strategy of the planner is to concentrate all initial adopters in one group. Therefore the expected profits' function has the following form, which we can differentiate twice:

$$E\Pi(t) = \pi \frac{r^{\frac{n}{2}} - r^{\frac{n-t}{2}}}{r^{\frac{n}{2}} - 1} - c(t) \Rightarrow \frac{d^2 E\Pi}{dt^2} = -\frac{\pi(\ln r)^2}{4(r^{\frac{n}{2}} - 1)} r^{\frac{n-t}{2}} - c''(t) < 0$$

Therefore, the expected profits' function is strictly concave in t and hence it has decreasing returns to scale in t.

If $c(t) = kt$, where $k \in \mathbb{R}_+$ is a constant then it is sufficient to see when the following equation has a solution.

$$\frac{dE\Pi}{dt} = \frac{\pi \ln r}{2(r^{\frac{n}{2}} - 1)} r^{\frac{n-t}{2}} - k = 0$$

It is apparent that this equation has a solution if and only if $\frac{dE\Pi}{dt}|_{t=0} < 0$ and $\frac{dE\Pi}{dt}|_{t=n} > 0$, which are satisfied when $\frac{\pi \ln r}{2(r^{\frac{n}{2}}-1)} < k < \frac{\pi \ln r}{2(r^{\frac{n}{2}}-1)} r^{\frac{n}{2}}$. One can even calculate the exact value of t which will be either the first integer to the right, or the first integer to the left of:

$$t^* = 2\frac{\ln(\frac{\pi}{2c}) + \ln(r^{\frac{n}{2}}) + \ln(\ln r) - \ln(r^{\frac{n}{2}} - 1)}{\ln r}$$

If k is larger than the upper bound then the derivative is always negative, and the optimal solution is $t = 0$. If k is lower than the lower bound then the derivative is always positive and the optimal solution is $t = n$.

$\square$

Proof of Remark 1. 1) τ strictly increasing in n:

$$\frac{\partial \tau}{\partial n} = \frac{1}{2p-1}\left[-\frac{(r^{s/2}-1)}{2}\frac{(r^{n/2}-1) - n\frac{\ln r}{2}r^{n/2}}{(r^{n/2}-1)^2}\right] = -\frac{(r^{s/2}-1)}{2(2p-1)(r^{n/2}-1)^2}\left[r^{n/2} - 1 - \frac{n \ln r}{2}r^{n/2}\right] > 0$$

because, for $r > 1$, $r^{n/2} - 1 - \frac{n \ln r}{2} r^{n/2}$ is strictly decreasing in r. For every n its maximum is attained for $r = 1$ and is equal to zero. Hence, the expression is negative for all $r > 1$. All the other terms are strictly positive, so the derivative is strictly positive and the mean waiting time is strictly increasing in n.

2) τ strictly concave in s:

$$\frac{\partial \tau}{\partial s} = \frac{1}{2p-1}\left[\frac{1}{2} - \frac{n}{4}\frac{r^{\frac{s}{2}} \ln r}{(r^{\frac{n}{2}} - 1)}\right] \Rightarrow \frac{\partial^2 \tau}{\partial s^2} = -\frac{n(\ln r)^2 r^{\frac{s}{2}}}{8(2p-1)(r^{\frac{n}{2}} - 1)} < 0$$

.

The other argument follows directly because $t = n - s \Rightarrow \frac{\partial^2 \tau}{\partial s^2} = \frac{\partial^2 \tau}{\partial t^2}$.

3) τ has interior maximum in t, therefore it has interior maximum in s:

We have already found that τ is strictly concave. Hence, we only need to ensure that there exists s^* such that $\frac{\partial \tau}{\partial s} = 0$.

At $s = 0$, $\frac{\partial \tau}{\partial s} = \frac{1}{2p-1}\left[\frac{1}{2} - \frac{n}{4}\frac{\ln r}{(r^{\frac{n}{2}}-1)}\right] > 0$. This holds because $\frac{n \ln r}{2(r^{\frac{n}{2}}-1)} < 1 \Leftrightarrow n \ln r - 2r^{\frac{n}{2}} + 2 < 0$, which is true for $r \geq 1$ because this latest expression, with respect to r, attains a unique maximum equal to zero for $r = 1$. Hence, it is strictly negative for all $r > 1$.

At $s = n$, $\frac{\partial \tau}{\partial s} = \frac{1}{2p-1}\left[\frac{1}{2} - \frac{n}{4}\frac{r^{\frac{n}{2}} \ln r}{(r^{\frac{n}{2}}-1)}\right] < 0$. This holds because $\frac{nr^{\frac{n}{2}} \ln r}{2(r^{\frac{n}{2}}-1)} > 1 \Leftrightarrow nr^{\frac{n}{2}} \ln r - 2r^{\frac{n}{2}} + 2 > 0$, which is true because this latest expression attains unique minimum equal to zero for $r = 1$. Hence, it is strictly positive for all $r > 1$.

Using the two previous results and the fact that the derivative is continuous in $(0, n)$ for all $r > 1$, we can apply Bolzano theorem and conclude that there exists some s^* such that the derivative becomes equal to zero.

4)

$$\lim_{r\to 1^+} s^* = 2\lim_{r\to 1^+} \frac{\ln\left(r^{\frac{n}{2}}-1\right)-\ln(\ln r)-\ln\left(\frac{n}{2}\right)}{\ln r} = 2\lim_{r\to 1^+} \frac{\frac{n}{2}r^{\frac{n}{2}}\ln r - r^{\frac{n}{2}}+1}{\ln r(r^{\frac{n}{2}}-1)} =$$

$$= 2\lim_{r\to 1^+} \frac{\frac{n^2}{4}r^{\frac{n}{2}}\ln r}{r^{\frac{n}{2}}-1+\frac{n}{2}r^{\frac{n}{2}}\ln r} = \frac{n^2}{2}\lim_{r\to 1^+}\frac{\frac{n}{2}\ln r+1}{n+\frac{n^2}{4}\ln r} = \frac{n}{2}$$

$$\lim_{r\to 1^+} \tau_{(t=\frac{n}{2},s=\frac{n}{2})} = \lim_{r\to 1^+}\frac{1}{2p-1}\left[\frac{n}{4}-\frac{n}{2}\frac{1}{r^{\frac{n}{4}}+1}\right] \overset{\frac{0}{0}}{=} \lim_{p\to\frac{1}{2}^+}\frac{n^2 r^{\frac{n}{4}-1}}{16(r^{\frac{n}{4}}+1)^2}\frac{\partial r}{\partial p} = \frac{n^2}{16}$$

5)

$$\lim_{r\to+\infty} s^* = 2\lim_{r\to+\infty}\frac{\ln(r^{\frac{n}{2}}-1)-\ln(\ln r)-\ln\left(\frac{n}{2}\right)}{\ln r} = 2\lim_{r\to+\infty}\left[\frac{nr^{\frac{n}{2}}}{2(r^{\frac{n}{2}}-1)}-\frac{1}{\ln r}\right] = n$$

$$\lim_{r\to+\infty}\tau_{(t=2,s=n-2)} = \lim_{r\to+\infty}\frac{1}{2p-1}\left[\frac{n}{2}-1-\frac{n}{2}\frac{r^{\frac{n}{2}-1}-1}{r^{\frac{n}{2}}-1}\right] = \frac{n}{2}-1$$

□

Proof of Remark 2. 1)Call $s' = \frac{s}{n-s}$ and $n' = \frac{n}{n-s}$ and notice that $\frac{\partial s'}{\partial n} = \frac{\partial n'}{\partial n} = -\frac{s}{(n-s)^2}$. Now,

$$\frac{\partial \tau}{\partial n} = \frac{1}{2p-1}\left[\frac{\partial s'}{\partial n}-\frac{\partial n'}{\partial n}\left(\frac{r^{s'}-1}{r^{n'}-1}\right)-n'\frac{\frac{\partial s'}{\partial n}r^{s'}\ln r(r^{n'}-1)-\frac{\partial n'}{\partial n}r^{n'}\ln r(r^{s'}-1)}{(r^{n'}-1)^2}\right] =$$

$$= -\frac{s(r^{n'}-r^{s'})}{(2p-1)(n-s)^2(r^{n'}-1)^2}(r^{n'}-1-n'\ln r) < 0$$

Given that the last term is positive for $r<1$.
2) Here we use again s' and n' and notice now that $\frac{\partial s'}{\partial s} = \frac{\partial n'}{\partial s} = \frac{n}{(n-s)^2}$.
The calculation is identical as in part (1) and we get that:

$$\frac{\partial \tau}{\partial s} = \frac{n(r^{n'}-r^{s'})}{(2p-1)(n-s)^2(r^{n'}-1)^2}(r^{n'}-1-n'\ln r) > 0$$

3) For $r\to 0^+$ the calculation is straightforward. For $r\to 1^-$ we use the previous definition of s' and n'. It is easier to calculate the limit with respect to p as it goes to $\frac{1}{2}$, and applying three times L'Hopital's rule we get

$$\lim_{p\to\frac{1}{2}}\tau = -\frac{n'}{2}\frac{\frac{1}{(1/2)^2}[s'n'(s'-n')]}{2(1/2)^2 n'^2\frac{1}{(1/2)^2}} = s' = \frac{s}{n-s}$$

4) The first result is straightforward. For the second one we use again the definition of s' and n'. Notice that $s' = n'-1$ and take the limit with respect to n'. When $s\to n$, then $n'\to\infty$. Hence,

$$\lim_{s\to n}\tau = \lim_{n'\to\infty}\frac{1}{2p-1}\left(n'-1-n'\frac{r^{n'-1}-1}{r^{n'}-1}\right) = \lim_{n'\to\infty}\frac{1}{2p-1}\left[n'\frac{r^{n'}(r-1)}{r^{n'}-1}-1\right]=0$$

Because,

$$\lim_{n'\to\infty}\frac{n'r^{n'-1}}{r^{n'}-1} = \lim_{n'\to\infty}\frac{n'}{r-r^{1-n'}} = \lim_{n'\to\infty}\frac{1}{(1-n')r^{-n'}}=0$$

□

Proof of Remark 3. 1)

$$\frac{\partial\tau}{\partial n} = \frac{1}{2p-1}\left(-\frac{r-1}{s}\frac{(r^{\frac{n}{s}}-1)-\frac{n}{s}r^{\frac{n}{s}}\ln r}{(r^{\frac{n}{s}}-1)^2}\right) = -\frac{r-1}{s(2p-1)(r^{\frac{n}{s}}-1)^2}\left(r^{\frac{n}{s}}-1-\frac{n}{s}r^{\frac{n}{s}}\ln r\right)>0$$

Because the last term is negative for all $\frac{n}{s}>0$.

2)

$$\frac{\partial\tau}{\partial s} = \frac{n(r-1)(r^{\frac{n}{s}}-1-\frac{n}{s}r^{\frac{n}{s}}\ln r)}{(2p-1)s^2(r^{\frac{n}{s}}-1)^2}<0$$

3) For $r\to 0$ the proof is straightforward by substitution. For $r\to 1$ is easier to calculate the limit with respect to p, as $p\to\frac{1}{2}$ and by applying three times L'Hopital rule we get:

$$\lim_{r\to\frac{1}{2}^-}\frac{1}{2p-1}\left(1-\frac{n}{s}\frac{(r-1)}{(r^{\frac{n}{s}}-1)}\right) = -\frac{n}{2s}\frac{\frac{n}{s}(\frac{n}{s}-1)-\frac{n}{s}(\frac{n}{s}-1)+\frac{n}{s}(\frac{n}{s}-1)(\frac{n}{s}-2)}{2(\frac{1}{2})^2(\frac{n}{s})^2} = \frac{n}{s}-1=\frac{t}{s}$$

4) Remembering that $r=\frac{p}{1-p}$, both proofs are straightforward by substitution.

□

Proof of Proposition 5.

$$\frac{\partial\tau_{(\frac{t}{2m},\frac{s}{2m})}}{\partial m} = \frac{1}{2p-1}\left[-\frac{s}{2m^2}+\frac{n}{2m^2}\left(\frac{r^{\frac{s}{2m}}-1}{r^{\frac{n}{2m}}-1}\right)-\frac{n}{2m}\frac{\partial\left(\frac{r^{\frac{s}{2m}}-1}{r^{\frac{n}{2m}}-1}\right)}{\partial m}\right]<0 \text{ for all } r\neq 1$$

Which holds because of the following:

- The first part of the expression inside the brackets is

$$-\frac{s}{2m^2}+\frac{n}{2m^2}\left(\frac{r^{\frac{s}{2m}}-1}{r^{\frac{n}{2m}}-1}\right) = \frac{r^{\frac{s}{2m}}-1}{m}\left(\frac{n}{2m}\frac{1}{r^{\frac{n}{2m}}-1}-\frac{s}{2m}\frac{1}{r^{\frac{s}{2m}}-1}\right)\begin{cases}<0 & \text{if } r>1\\ >0 & \text{if } r<1\end{cases}$$

  To show this we need to define the function $f(x)=\frac{x}{r^x-1}$. Its derivative is $\frac{df}{dx}=\frac{r^x-1-r^x x\ln r}{(r^x-1)^2}$, whose denominator is positive for all $r\neq 1$, whereas the nominator is negative, because $\frac{\partial(r^x-1-r^x x\ln r)}{\partial r}=-x^2r^{x-1}\ln r$ which is positive for $r<1$ and negative for $r>1$, obtaining a maximum equal to zero for $r=1$. Hence, f is strictly decreasing in x for all $r\neq 1$, which makes the expression inside the parenthesis always negative (because $\frac{n}{2m}>\frac{s}{2m}$) and the sign depending only on $r^{\frac{s}{2m}}-1$ which is positive for $r>1$ and negative for $r<1$.

- Now, to find the sign of $-\frac{n}{2m}\frac{\partial\left(\frac{r^{\frac{s}{2m}}-1}{r^{\frac{n}{2m}}-1}\right)}{\partial m}$, notice that $\frac{r^{\frac{s}{2m}}-1}{r^{\frac{n}{2m}}-1} = 1 - \frac{r^{\frac{n}{2m}}-r^{\frac{s}{2m}}}{r^{\frac{n}{2m}}-1}$, hence:

$$-\frac{n}{2m}\frac{\partial\left(\frac{r^{\frac{s}{2m}}-1}{r^{\frac{n}{2m}}-1}\right)}{\partial m} = \frac{n}{2m}\frac{\partial\left(\frac{r^{\frac{n}{2m}}-r^{\frac{s}{2m}}}{r^{\frac{n}{2m}}-1}\right)}{\partial m}\begin{cases} <0 & \text{if } r>1 \\ >0 & \text{if } r<1 \end{cases}$$

  which holds by Proposition 1.

Therefore the whole expression inside the brackets is negative for $r > 1$ and positive for $r < 1$. This expression is multiplied by $(2p-1)$ which positive for $r > 1$ and negative for $r < 1$. This makes the derivative of τ with respect to m always negative.

□

Proof of Proposition 6. Despite being interesting only in integer values of s_1, s_2, t and n, this is a well-behaving smooth function for all $r \neq 1$. Hence, we can differentiate it with respect to s_1.

$$\frac{dP}{ds_1} = \frac{r^n\left[r^{\frac{t}{2}}-1\right]}{r^n-1}\left[\frac{\left(r^{s_1}\ln r + r^{-s_1-t}\ln r\right)\left(r^{s_1+\frac{t}{2}}-1\right) - \left(r^{s_1}-r^{-s_1-t}\right)\left(r^{s_1+\frac{t}{2}}\ln r\right)}{\left(r^{s_1+\frac{t}{2}}-1\right)^2}\right] =$$

$$= \frac{r^n\left(r^{\frac{t}{2}}-1\right)\ln r}{\left(r^n-1\right)\left(r^{s_1+\frac{t}{2}}-1\right)^2}\left[r^{2s_1+\frac{t}{2}} + r^{-\frac{t}{2}} - r^{s_1} - r^{-s_1-t} - r^{2s_1+\frac{t}{2}} + r^{-\frac{t}{2}}\right] =$$

$$= \frac{r^n\left(r^{\frac{t}{2}}-1\right)\ln r}{\left(r^n-1\right)\left(r^{s_1+\frac{t}{2}}-1\right)^2}\left(2r^{-\frac{t}{2}} - r^{s_1} - r^{-s_1-t}\right) =$$

$$= \frac{r^n\left(r^{\frac{t}{2}}-1\right)\ln r}{\left(r^n-1\right)\left(r^{s_1+\frac{t}{2}}-1\right)^2}\left[-r^{s_1}\left(r^{2s_1} - 2r^{s_1-\frac{t}{2}} + r^{-t}\right)\right] =$$

$$= -\frac{r^n\left(r^{\frac{t}{2}}-1\right)\ln r}{\left(r^n-1\right)\left(r^{s_1+\frac{t}{2}}-1\right)^2}\left(r^{s_1} - r^{-\frac{t}{2}}\right)^2 r^{-s_1}$$

If $r > 1$, then $\frac{dP}{ds_1} < 0$, so the optimal targeting decision is $s_1 = 0$, i.e. target one corner. Whilst, if $r < 1$, then $\frac{dP}{ds_1} > 0$ and recalling that $s_1 \leq s_2$, the optimal decision is $s_1 = s_2$ for s even, or $s_1 = s_2 - 1$ for s odd, i.e. to target the middle of the line. See also the following figures (Figure A.24).

For the case of $r = 1$, it is apparent that the $P(B|\cdot)$ does not depend on s_1, hence every decision yields the same result.

□

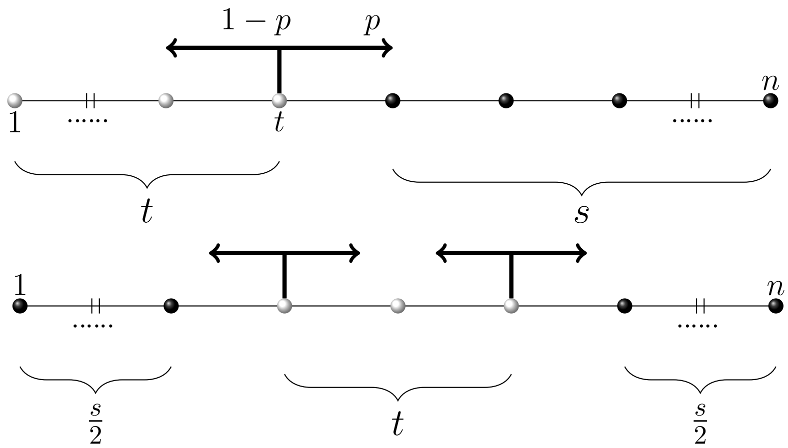


Figure B.25: Optimal choice for $p > 1/2$ (above) and for $p < 1/2$ (below)

Proof of Proposition 7.

$$\frac{dP}{ds_1} = \frac{r^n\left(r^{\frac{t+1}{2}}-1\right)}{r^n-1}\left[\frac{\left(r^{s_1}\ln r + r^{-s_1-t}\ln r\right)\left(r^{s_1+\frac{t+1}{2}}-1\right)-\left(r^{s_1}-r^{-s_1-t}\right)r^{s_1+\frac{t+1}{2}}\ln r}{\left(r^{s_1+\frac{t+1}{2}}-1\right)^2}\right] =$$

$$= \frac{r^n\left(r^{\frac{t+1}{2}}-1\right)\ln r}{\left(r^n-1\right)\left(r^{s_1+\frac{t+1}{2}}-1\right)^2}\left[r^{2s_1+\frac{t+1}{2}}+r^{\frac{-t+1}{2}}-r^{s_1}-r^{-s_1-t}-r^{2s_1+\frac{t+1}{2}}+r^{\frac{-t+1}{2}}\right] =$$

$$= \frac{r^n\left(r^{\frac{t+1}{2}}-1\right)\ln r}{\left(r^n-1\right)\left(r^{s_1+\frac{t+1}{2}}-1\right)^2}\left(2r^{\frac{1-t}{2}}-r^{s_1}-r^{-s_1-t}\right)$$

So, the sign of derivative will depend on the sign of $2r^{\frac{1-t}{2}} - r^{s_1} - r^{-s_1-t}$.

For $r < 1$, we can rewrite it as $r^{-s_1}(2r^{\frac{1}{2}}r^{s_1}r^{\frac{-t}{2}} - r^{2s_1} - r^{-t})$ which is negative because:

$$\begin{aligned} r<1 \Rightarrow r^{\frac{1}{2}} < 1 \Rightarrow 2r^{\frac{1}{2}}r^{s_1}r^{-\frac{t}{2}} < r^{s_1}r^{-\frac{t}{2}} \Rightarrow \\ \Rightarrow 2r^{\frac{1}{2}}r^{s_1}r^{-\frac{t}{2}} - r^{2s_1} - r^{-t} < r^{s_1}r^{-\frac{t}{2}} - r^{2s_1} - r^{-t} \Rightarrow \\ \Rightarrow 2r^{\frac{1}{2}}r^{s_1}r^{-\frac{t}{2}} - r^{2s_1} - r^{-t} < -\left(r^{s_1} - r^{-\frac{t}{2}}\right)^2 < 0 \end{aligned}$$

So in general for $r < 1$ we get $\frac{dP}{ds_1} > 0$, hence, as before, the optimal decision is $s_1 = s_2$ for s even, or $s_1 = s_2 - 1$ for s odd, i.e. to target the middle of the line.

For $r > 1$,we can rewrite it as $r^{-s_1}r^{-t}\left(2r^{s_1}r^{\frac{1+t}{2}} - r^{2s_1}r^t - 1\right)$ and naming $x = r^{s_1}$ the content of the parenthesis becomes a polynomial of degree 2, namely $-[r^t x^2 - 2r^{\frac{1+t}{2}}x + 1]$. Let us first calculate the roots of the polynomial which are

$$
\begin{aligned}
x_{1,2} &= \frac{2r^{\frac{1+t}{2}} \pm \sqrt{4r^{1+t} - 4r^t}}{2r^t} = r^{\frac{1-t}{2}} \pm (r-1)^{\frac{1}{2}} r^{-\frac{t}{2}} \\
\Rightarrow x_1 &= \left[r^{\frac{1}{2}} - (r-1)^{\frac{1}{2}}\right] r^{-\frac{t}{2}} \text{ and} \\
\Rightarrow x_2 &= \left[r^{\frac{1}{2}} + (r-1)^{\frac{1}{2}}\right] r^{-\frac{t}{2}}
\end{aligned}
$$

$$
\begin{aligned}
-\left(r^t x^2 - 2r^{\frac{1+t}{2}} x + 1\right) \geq 0 &\Leftrightarrow \left[r^{\frac{1}{2}} - (r-1)^{\frac{1}{2}}\right] r^{-\frac{t}{2}} \leq r^{s_1} \leq \left[r^{\frac{1}{2}} + (r-1)^{\frac{1}{2}}\right] r^{-\frac{t}{2}} \\
&\Leftrightarrow r^{\frac{1}{2}} - (r-1)^{\frac{1}{2}} \leq r^{s_1 + \frac{t}{2}} \leq r^{\frac{1}{2}} + (r-1)^{\frac{1}{2}} \\
&\Leftrightarrow \ln r^{s_1} \leq \ln\left(\frac{r^{\frac{1}{2}} - (r-1)^{\frac{1}{2}}}{r^{\frac{t}{2}}}\right) \\
&\Leftrightarrow s_1 \leq \frac{\ln\left[r^{\frac{1}{2}} + (r-1)^{\frac{1}{2}}\right]}{\ln r} - \frac{t}{2} = g(r,t)
\end{aligned}
$$

The left hand-side is always satisfied because $r^{\frac{1}{2}} - (r-1)^{\frac{1}{2}} < r^{\frac{1}{2}} < r^{s_1 + \frac{t}{2}}$.

So, for $s_1 \leq \frac{\ln\left[r^{\frac{1}{2}} + (r-1)^{\frac{1}{2}}\right]}{\ln r} - \frac{t}{2} = g(r,t)$ we find that $\frac{dP}{ds_1} > 0$, while for $s_1 \leq \frac{\ln\left[r^{\frac{1}{2}} + (r-1)^{\frac{1}{2}}\right]}{\ln r} - \frac{t}{2} = g(r,t)$ we find $\frac{dP}{ds_1} < 0$, hence the function has a global maximum at $s_1 = \frac{\ln[r^{1/2} + (r-1)^{1/2}]}{\ln r} - \frac{t}{2} = g(r,t)$, however, notice that in our problem s_1 has to be an integer, so in order to find the maximum we need to compare the two closest integers to s_1, namely $\lfloor g(r,t) \rfloor$ and $\lceil g(r,t) \rceil$. If $P_B(s_1 = \lfloor g(r,t) \rfloor) > P_B(s_1 = \lceil g(r,t) \rceil)$ then $s_1^* = \lfloor g(r,t) \rfloor$ and vice versa.

For $r = 1$, we have that $P_B(s_1; s, t, n, r = 1) = \frac{(t+1)(t+2s_1)}{(t+2s_1+1)n}$ so,

$$
\frac{dP}{ds_1} = \frac{t+1}{n} \frac{2(2s_1 + t + 1) - 2(2s_1 + t)}{(2s_1 + t + 1)^2} = \frac{2(t+1)}{n(2s_1 + t + 1)^2} > 0
$$

Hence, the optimal is to target the middle. □

Proof of Corollary 1. The proof of Corollary 1 comes directly after the proof of the following lemma (Lemma 4).

□

Lemma 4. *The function* $f(r) = \frac{\ln[r^{\frac{1}{2}} + (r-1)^{\frac{1}{2}}]}{\ln r}$ *has the following properties.*

1. $\lim_{r \to 1^+} f(r) = +\infty$,
2. $\lim_{r \to +\infty} f(r) = \frac{1}{2}$,
3. f *is strictly decreasing in* r,

Proof of Lemma 4.

$$(1)\ \lim_{r\to1^+} f(r) = \lim_{r\to1^+}\frac{\ln\left[r^{\frac12}+(r-1)^{\frac12}\right]}{\ln r} = \lim_{r\to1^+}\frac{\frac{1}{r^{\frac12}+(r-1)^{\frac12}}\left(\frac{1}{2r^{\frac12}}+\frac{1}{2(r-1)^{\frac12}}\right)}{\frac1r}\ \text{(L'Hopital)}$$

$$= \lim_{r\to1^+}\left\{\frac{r}{2r^{\frac12}\left[r^{\frac12}+(r-1)^{\frac12}\right]}+\frac{r}{2(r-1)^{\frac12}\left[r^{\frac12}+(r-1)^{\frac12}\right]}\right\}$$

$$= \lim_{r\to1^+}\frac{r+r\{2r^{\frac12}\left[r^{\frac12}+(r-1)^{\frac12}\right]\}}{4r^{\frac12}(r-1)^{\frac12}\left[r^{\frac12}+(r-1)^{\frac12}\right]^2} = \frac{3}{0^+} = +\infty$$

$$(2)\ \lim_{r\to+\infty} f(r) = \lim_{r\to+\infty}\frac{\ln\left[r^{\frac12}+(r-1)^{\frac12}\right]}{\ln r} = \lim_{r\to+\infty}\frac{\frac{1}{r^{\frac12}+(r-1)^{\frac12}}\left(\frac{1}{2r^{\frac12}}+\frac{1}{2(r-1)^{\frac12}}\right)}{\frac1r}$$

$$= \frac12\lim_{r\to+\infty}\left[\frac{r}{r+(r^2-r)^{1/2}}+\frac{r}{r+(r^2-r)^{1/2}-1}\right]\ \text{(L'Hopital)}$$

$$= \frac12\lim_{r\to+\infty}\left[\frac{1}{1+\frac{1}{2(r^2-r)^{1/2}(2r-1)}}+\frac{1}{1+\frac{1}{2(r^2-r)^{1/2}(2r-1)}}\right]$$

$$= \frac12\lim_{r\to+\infty}\frac{2}{1+\frac{2r-1}{2(r^2-r)^{1/2}}} = \frac12\lim_{r\to+\infty}\frac{2}{1+\frac{2r-2}{2(r^2-r)^{1/2}}+\frac{1}{2(r^2-r)^{1/2}}}$$

$$= \frac12\lim_{r\to+\infty}\frac{2}{1+\frac{r-1}{r^{\frac12}(r-1)^{\frac12}}+\frac{1}{2(r^2-r)^{1/2}}} = \frac12\left(\frac{2}{1+1+0}\right) = \frac12$$

$$(3)\ \frac{df}{dr} = \frac{d\left(\frac{\ln\left(r^{\frac12}+(r-1)^{\frac12}\right)}{\ln r}\right)}{dr} = \frac{\frac{\ln r}{r^{\frac12}+(r-1)^{\frac12}}\left[\frac{1}{2r^{\frac12}}+\frac{1}{2(r-1)^{\frac12}}\right]-\frac{\ln\left(r^{\frac12}+(r-1)^{\frac12}\right)}{r}}{(\ln r)^2}.$$

$$= \frac{\left[r^{\frac12}+(r-1)^{\frac12}\right]}{2\left[r^{\frac12}+(r-1)^{\frac12}\right]r^{\frac12}(r-1)^{\frac12}\ln r}-\frac{\ln\left(r^{\frac12}+(r-1)^{\frac12}\right)}{r(\ln r)^2}$$

$$= \frac{1}{2r^{\frac12}(r-1)^{\frac12}\ln r}-\frac{\ln\left(r^{\frac12}+(r-1)^{\frac12}\right)}{r(\ln r)^2}$$

$$< \frac{1}{2\ln r(r-1)}-\frac{\ln\left(r^{\frac12}+(r-1)^{\frac12}\right)}{r(\ln r)^2}$$

$$< \frac{1}{2\ln r(r-1)}-\frac{\ln\left(2r^{\frac12}\right)}{r(\ln r)^2} = \frac{1}{2\ln r(r-1)}-\frac{\ln 2}{r(\ln r)^2}-\frac{1}{2r\ln r}$$

$$= \frac{1}{2\ln rr(r-1)}-\frac{\ln 2}{r(\ln r)^2} = \frac{1}{r\ln r}\left[\frac{1}{2(r-1)}-\frac{\ln 2}{\ln r}\right]$$

$$= \frac{1}{2r(r-1)(\ln r)^2}\left[\ln r-2\ln 2(r-1)\right] < 0$$

because both $[\ln r]$ and $[2\ln 2\,(r-1)]$ are increasing functions, equal to zero for $r=1$ but the second one increases at a higher rate, because:

$$\frac{d\,(\ln r)}{dr} = \frac{1}{r} < 2\ln 2 = \frac{d\,[2\ln 2\,(r-1)]}{dr}$$

.

Hence, f is decreasing in r. $\square$

Appendix C. Figures

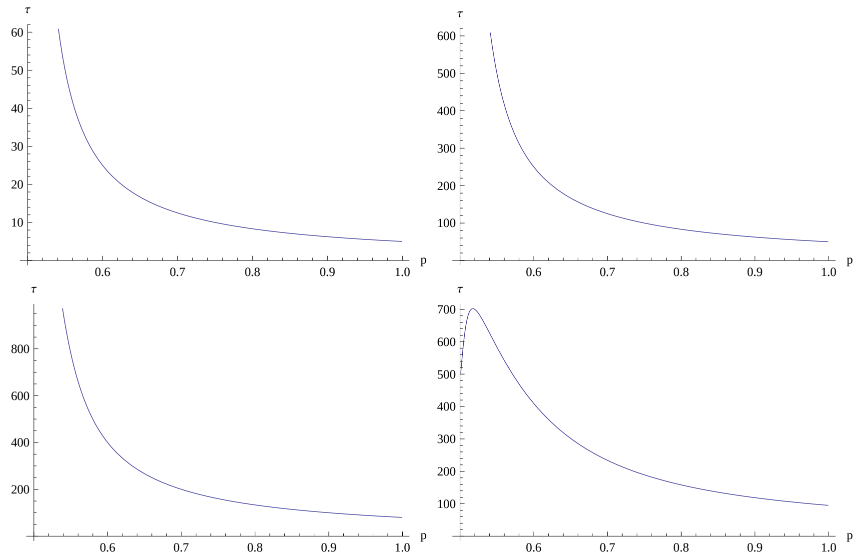


Figure C.26: $\tau(p)$ for $r > 1$, $n = 200$ and (i)$t = 190$, (ii)$t = 100$, (iii)$t = 40$, (iv)$t = 10$

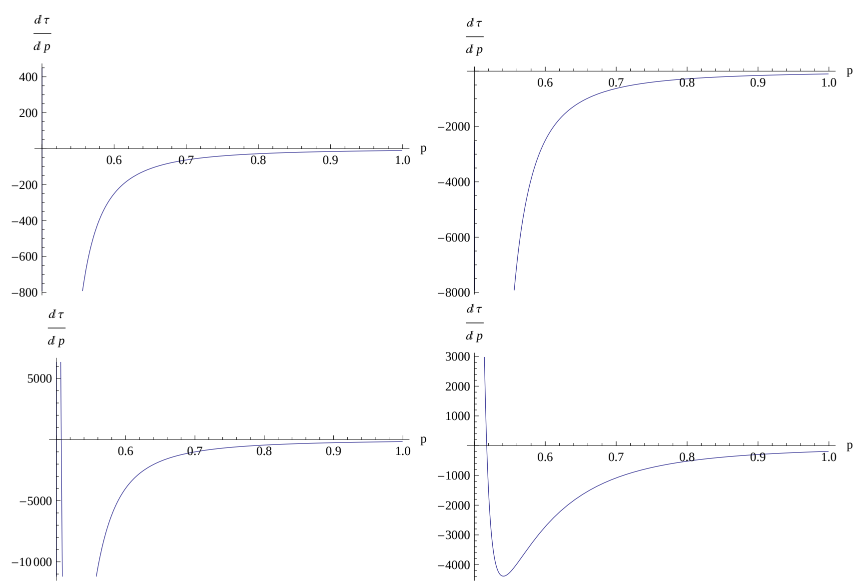


Figure C.27: $\frac{d\tau(p)}{dp}$ for $r > 1$, $n = 200$ and (i)$t = 190$, (ii)$t = 100$, (iii)$t = 40$, (iv)$t = 10$

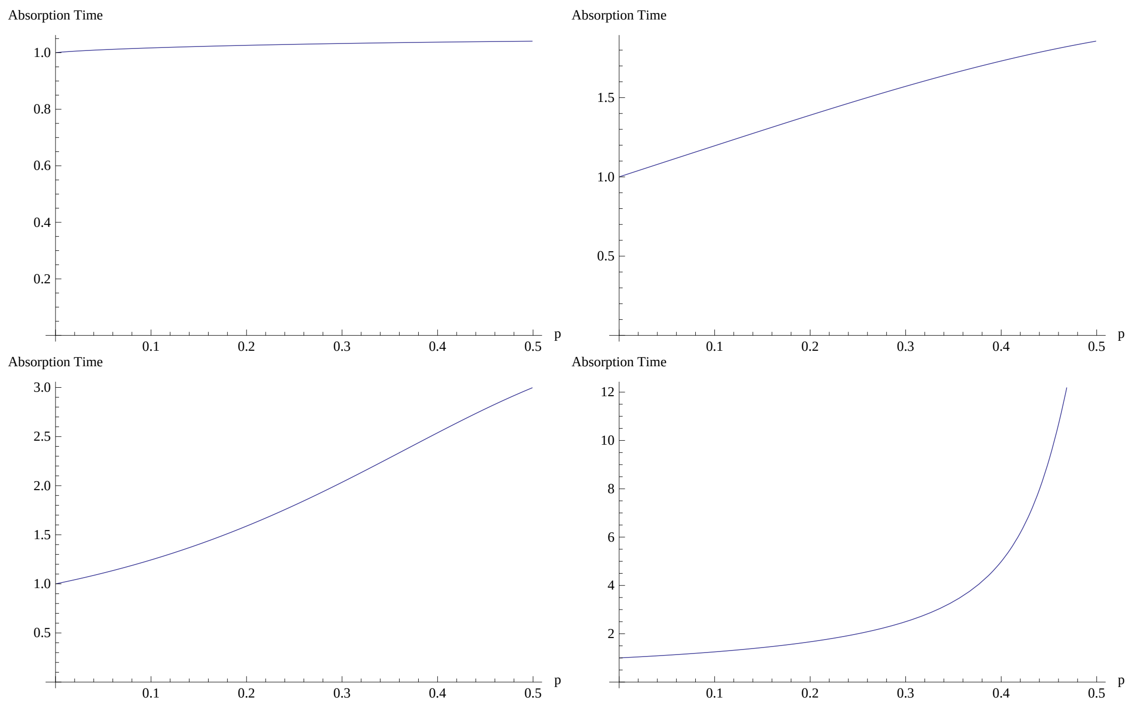


Figure C.28: $\tau(p)$ for $r < 1$, $t < s$, $n = 200$ and (i)$t = 98$, (ii)$t = 70$, (iii)$t = 50$, (iv)$t = 10$

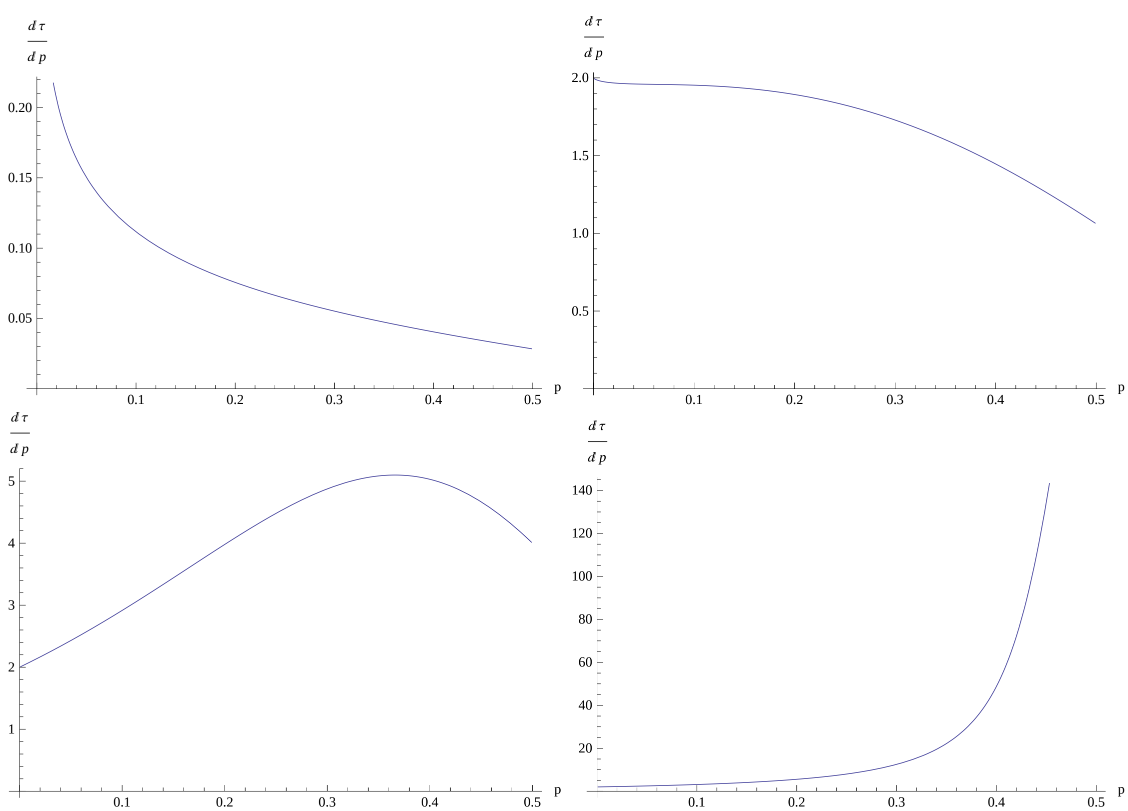


Figure C.29: $\frac{d\tau(p)}{dp}$ for $r < 1$, $t < s$, $n = 200$ and (i)$t = 98$, (ii)$t = 70$, (iii)$t = 50$, (iv)$t = 10$

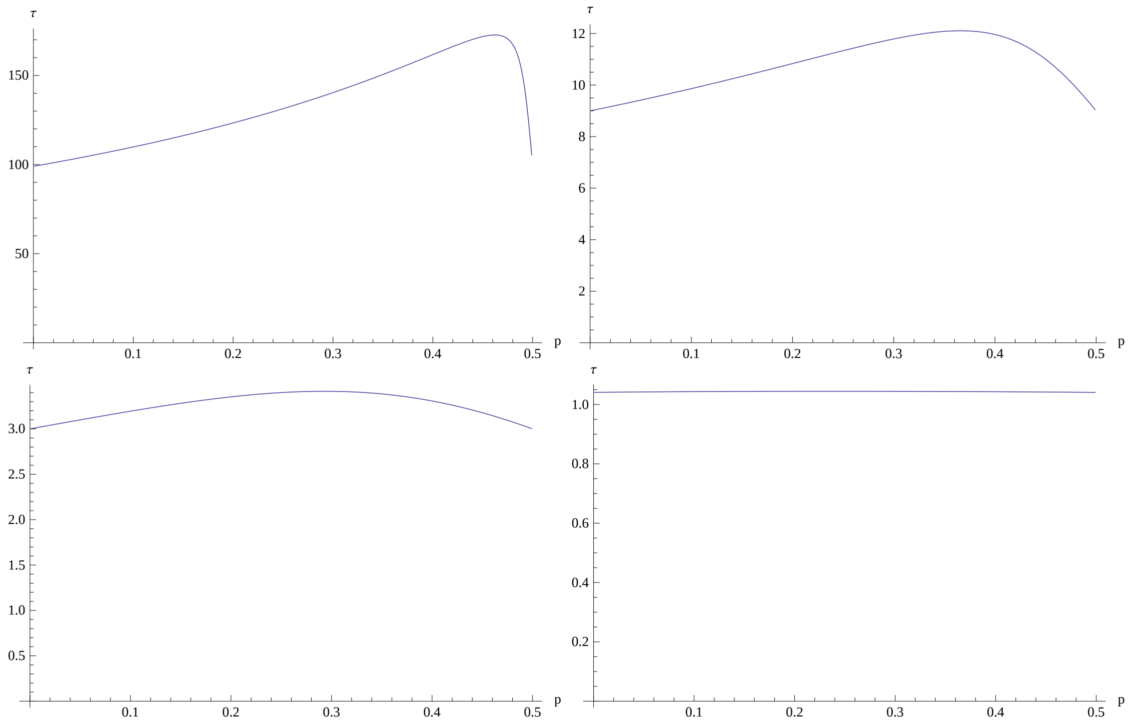


Figure C.30: $\tau(p)$ for $r < 1$, $t > s$, $n = 200$ and (i)$t = 198$, (ii)$t = 180$, (iii)$t = 150$, (iv)$t = 102$

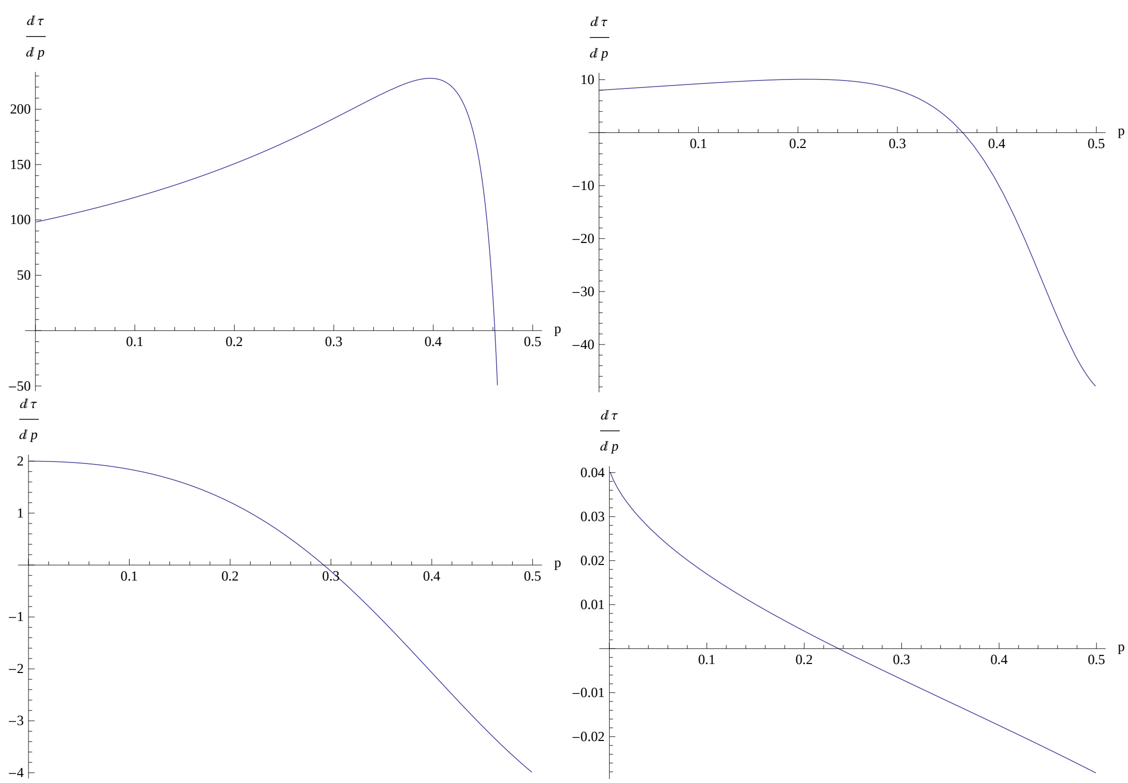


Figure C.31: $\frac{d\tau(p)}{dp}$ for $r < 1$, $t > s$, $n = 200$ and (i)$t = 198$, (ii)$t = 180$, (iii)$t = 150$, (iv)$t = 102$

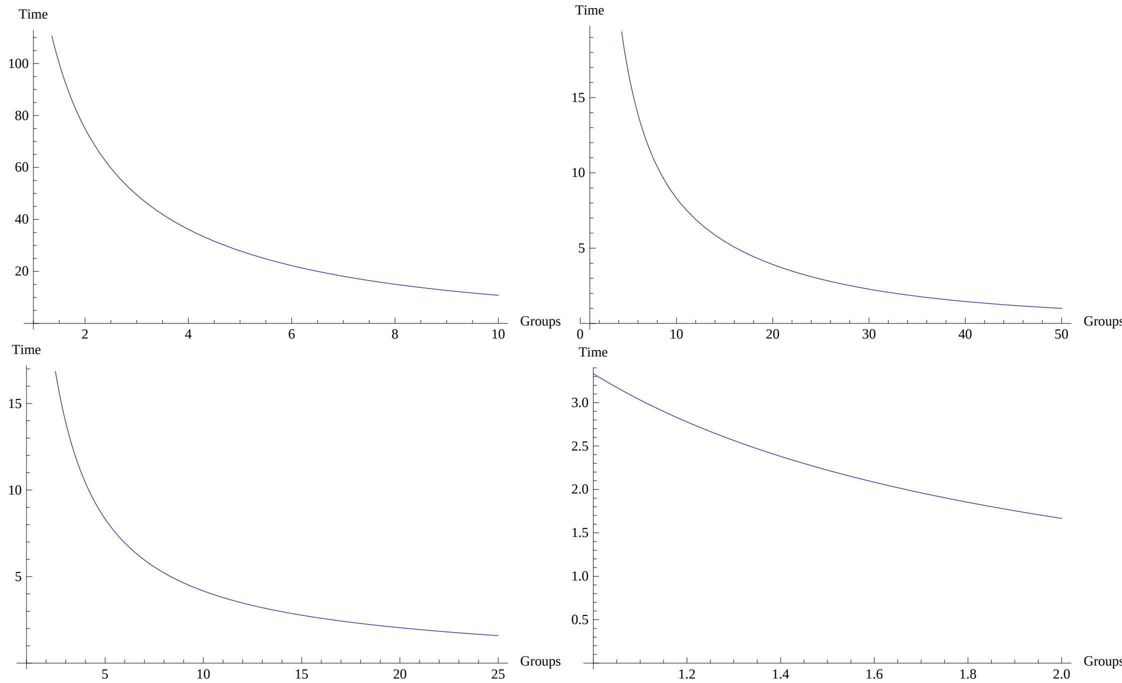


Figure C.32: $\tau(m)$ for $r = 0.2$, $n = 200$ and (i)$t = 180$, (ii)$t = 100$, (iii)$t = 50$, (iv)$t = 4$

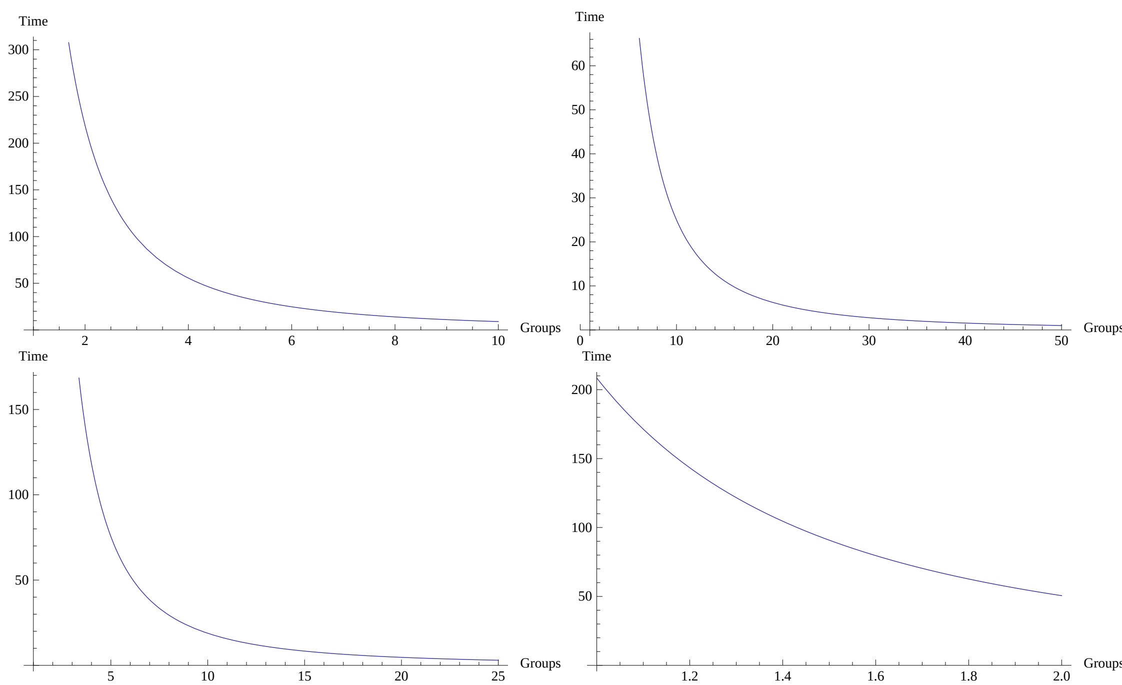


Figure C.33: $\tau(m)$ for $r = 0.501$, $n = 200$ and (i)$t = 180$, (ii)$t = 100$, (iii)$t = 50$, (iv)$t = 4$

References

Alós-Ferrer, C. & Weidenholzer, S. (2008). Contagion and efficiency. *Journal of Economic Theory* 143, 251–274.

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

Bagnoli, F., Boccara, N. & Rechtman, R. (2001). Nature of phase transition in a probabilistic cellular automaton with two absorbing states. *Physical Review E* 63(4), 046116.

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

Ballester,C., Calvo-Armengol, A. & Zenou, Y. (2006). Who's who in networks. Wanted: The key player *Econometrica* 74(5), 1403–1417.

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

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

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

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.

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.

Duersch, P., Oechssler, J. and Schipper, B.C. (2012). Unbeatable Imitation, *Games and Economic Behavior* 76(1), 88–96.

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.

Eshel, I., Samuelson, L. & Shaked A. (1998). Altruists, egoists and hooligans in a local interaction model. *American Economic Review* 88, 157–179.

Fosco, C. & Mengel, F. (2011). Cooperation through imitation and exclusion in networks. *Journal of Economic Dynamics and Control* 35(5), 641–658.

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

Galeotti, A. & Ghiglino, C. & Squintani, F. (2011). Strategic Information Transmission in Networks. *Working Paper*.

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

Gigerenzer, G. & Selten, R. (2002). *Bounded Rationality: The Adaptive Toolbox*. The MIT Press.

Kemeny, J.G. & Snell, J.L. (1960). *Finite Markov Chains*. Springer.

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.

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

Richardson, M. and Domingos, P. (2002). Mining Knowledge–Sharing Sites for Viral Martketing. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 61–70. Edmonton, Canada 2002.

Sákovics, J. & Steiner, J. (2012). Who matters in coordination problems?. *American Economic Review*, forthcoming.

Schlag, K. (1998). Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory* 78, 130–156.

Tsakas, N. (2013). Imitating the most successful neighbor in social networks. http://dx.doi.org/10.2139/ssrn.2276611.

Vega-Redondo, F. (1997). The evolution of Walrasian Behavior. *Econometrica* 65, 375–384.

Weibull, J. (1995). *Evolutionary Game Theory*. MIT Press, Cambridge, MA.

Yildiz, E., Acemoglu, D., Ozdaglar, A. E., Saberi, A. and Scaglione, A. (2011). Discrete Opinion Dynamics with Stubborn Agents.http://dx.doi.org/10.2139/ssrn.1744113