The Spread of Free-Riding Behavior in a Social Network*

Dunia López-Pintado

Universidad Autónoma de Barcelona

November 4, 2007

Abstract

We study a model where agents, located in a social network, decide whether to exert effort or not in experimenting with a new technology (or acquiring a new skill, innovating, etc.). We assume that agents have strong incentives to free ride on their neighbors' effort decisions. In the static version of the model efforts are chosen simultaneously. In equilibrium, agents exerting effort are never connected with each other and all other agents are connected with at least one agent exerting effort. We propose a mean-field dynamics in which agents choose in each period the best response to the last period's decisions of their neighbors. We characterize the equilibrium of such a dynamics and show how the pattern of free riders in the network depends on properties of the connectivity distribution.

JEL classification: C45, C73, D00, D83, D85, H41.

1   Introduction

There are many situations in which agents' actions depend crucially on the actions of their social contacts (e.g., the decision of adopting a new technology, voting in a referendum, engaging in civil disorder, etc.). A common ingredient in all these situations is that a certain behavior spreads across the social network due to some "contagion process" (namely, the more agents choosing a certain action, the more it becomes appealing for another neighbor in the social network to do so as well). In this paper, however, we examine the polar framework, that is, situations where agents want to anti-coordinate with their neighbors. A context where this generally applies is the case of a local public good which is non-excludable along social contacts. A specific example fitting this framework is the case of spillovers from "individual innovation" (e.g., friends benefiting from the research into a product of other friends, firms learning from related firms the benefits of a new technology and so on). Another plausible context is that in which agents can develop a certain skill (valuable to all agents) provided they exert some effort. Nevertheless they can also free ride on their (already skilled) neighbors instead.[1]

More precisely, this paper analyzes a social-network model in which the two following assumptions hold. First, agents make a binary decision, that is, whether or not to exert effort. Second, incentives are such that whenever an agent has someone in his neighborhood exerting effort, he free rides and decides not to do so.

The first part of the paper presents a static model where agents choose their actions simultaneously. We characterize the set of Nash equilibria of this

---

[1] Local public goods have been recently analyzed in a network model by Bramoulle and Kranton (2006).

simple game. To do so, we follow the arguments exposed by Bramoulle and Kranton (2006) in their (more general) model.[2] We obtain (as they do) that whenever the set of agents exerting effort forms a maximal independent set this determines a Nash equilibrium. In other words, in equilibrium, agents exerting effort are never connected with each other and all other agents are connected with at least one agent exerting effort. Although in our framework this result is straightforward, it serves as a natural starting point.

The second (and more innovative) part of this paper, studies a dynamic model in which agents choose a myopic best response in each period to the actions taken by their neighbors in the previous period. We analyze a mean-field version of this dynamics which assumes that every period neighbors are chosen randomly from the population. In this case, the main property of the network is its connectivity distribution (i.e., the distribution of the number of neighbors of each agent). This approach raises a whole new set of research questions. For instance, how do the properties (such as the mean and variance) of the connectivity distribution affect the pattern of free-riding behavior? More precisely, are more dense networks better or worse for the diffusion of free-riding behavior? Also, how does the variance of the connectivity distribution affect this phenomenon?

Our analysis of this dynamic model yields the following main insights. First, there exists a unique globally stable state of the dynamics. In other words, although the identities of those agents who free ride might change from one period to another, there exists a unique value of the fraction of individuals free riding sustainable in equilibrium. We also provide some evidence that the higher the density of the network the higher the fraction of free riders in equilibrium. Also, comparing networks with the same average connectivity, we claim that the higher their variance the lower the equilibrium fraction of free riders.

---

[2] As discussed in some detail in Section 5, Bramoulle and Kranton (2006) assume agents can choose a level of effort in $[0, +\infty)$.

The paper is organized as follows. In the next section, we present the static model and characterize the Nash equilibria of the game. In Section 3, we describe the dynamic mean-field model and characterize the globally stable state of the dynamics. In Section 4, we provide simulations of the process on random networks with scale-free and exponential connectivity distributions. Finally, in Section 5, we discuss the results and compare our approach with the related literature.

2 The Static Approach

2.1 The Model

Consider a finite set of agents $N = \{1, ..., n\}$. Each agent $i \in N$ chooses an action a_i from the set $A = \{0, 1\}$, where $a_i = 1$ can be interpreted as the decision of putting effort (e.g., the decision to research a new product). Let $a = (a_1, ..., a_n)$ denote an action profile.

Each agent interacts only with a subset of the population. In particular, agents form a network g, where $g_{ij} = 1$ if agent j benefits from the result of agent i's effort, and $g_{ij} = 0$ otherwise. We assume that the network is undirected and therefore if i is connected with j, then j is also connected with i. That is, $g_{ij} = 1$ if and only if $g_{ji} = 0$. In addition, reflexive links are ruled out and therefore $g_{ii} = 0$. Let $N_i \subseteq N$ be the set of neighbors of i. Formally, $N_i = \{j \in N, \text{ s.t. } g_{ij} = 1\}$ and let k_i, namely i's connectivity, be the cardinality of N_i.

We make the following assumptions regarding the utility functions of agents. First, individuals care about the actions of their neighbors, but not of individuals further away in the graph. Second, only the overall number of neighbors exerting effort matters and not their identities. Finally, we assume an extreme version of substitutability of agents' effort, where agents have incentives to exert effort only if no neighbor has already done so. Formally, let $U_i(a_i, a_{-i})$ be

agent i's utility function or analogously let $u_i : A \times \{0, 1, 2, ..., d_i\} \rightarrow \mathbb{R}$, where $u_i(a_i, d_i)$ is the utility individual i enjoys if he chooses action a_i and d_i neighbors have decided to exert effort. That is, given action profile a, $d_i = \sum_{j \in N_i} a_j$ and

$$U_i(a_i, a_{-i}) = u_i(a_i, d_i).$$

We assume that

$$u_i(1, 0) > u_i(0, 0) \tag{1}$$

which implies that an agent has incentives to exert effort if no neighbor is doing so, and

$$u_i(1, d) < u_i(0, d) \tag{2}$$

for $1 \leq d \leq k_i$, which implies that if at least one neighbor is exerting effort, an agent would not want to do so.

The above utility expression allows us to particularize the standard notion of Nash equilibrium as follows. An action profile a^* is said to be a *Nash equilibrium* for the game if, for all $i \in N$

$$u_i(a_i^*, d_i^*) \geq u_i(a_i, d_i^*), \forall a_i \in A.$$

A Nash equilibrium is said to be strict if every player gets a strictly higher utility with her current strategy than he would with any other strategy.

2.2 The Result

Before presenting the first result of the paper, let us define formally the concept of a *maximal independent set* of the network g. First, given the network g, we say that I is an independent set of agents if no two agents belonging to I are connected (i.e., $\forall i, j \in I \ g_{ij} = 0$). An independent set is maximal when it is not a proper set of any other independent set.

Proposition 1 *The strategy profile $a^* = (a_1^*, \ldots a_n^*)$ is a Nash Equilibrium if and only if the set of agents choosing action 1 form a maximal independent set of the network g.*

The proof is straightforward and uses an analogous argument to the one proposed by Bramoulle and Kranton (2006). Let I^* be the set of agents choosing 1 in a Nash equilibrium a^*. Due to condition (2) $\forall i, j \in I^*$, $g_{ij} = 0$ and therefore I^* is an independent set. Moreover, due to condition (1), every agent in $N \backslash I^*$ is connected to at least one agent in I^*, which implies that I^* is also maximal.

Proposition 1 provides a characterization of the Nash equilibria of the game. Notice that, for any graph, there are potentially many Nash equilibria. This does not allow us to relate the network architecture with the number of free riders in equilibrium in a precise way. We next propose a concept which is relevant for the analysis of the stability of the equilibrium states.

Following Bramoulle and Kranton (2006) a *maximal independent set of order r* is defined as a maximal independent set I such that any individual not in I is connected to at least r individuals in I. Unlike a maximal independent set which always exists, a maximal independent set of order r, where $r > 1$ might not even exist (e.g., the complete network has no maximal independent set of order higher than 1). It seems rather intuitive that this concept allows us to rank the equilibria in terms of their stability. We conjecture that an equilibrium characterized by a maximal independent set of a higher order is "more stable".

To see this, consider the following heuristic argument. Assume a dynamics where every period an individual is selected to update his action, and then chooses a best response to the action taken by his neighbors in the previous period. Assume that individuals can mutate and switch actions. Moreover, for simplicity, assume a mutation from action 1 to action 0 is significantly more likely than the reverse mutation. It seems intuitive that equilibria where agents choosing 1 are forming a maximal independent set of a high order, will be more stable than those corresponding to maximal independent sets of lower orders. Notice that individuals choosing 0 would have many neighbors choosing 1 and therefore, many of them would have to simultaneously mutate in order

6

to destabilize this equilibrium. This implies, in particular, that a network with many starlike patterns and a state where the peripheral agents in the stars are the ones exerting effort is a "very stable state".

To formally analyze this problem, however, one must specify the updating and mutation dynamics carefully and test for the robustness of the results to small changes in the assumptions. Moreover, given a large and complex network it is computationally hard to find maximal independent sets of a certain order and furthermore we might still have multiplicity of the set of stable states. Therefore, to proceed we have decided to analyze a mean-field version of the dynamics instead, hoping that this will allow us relate the collective outcomes with other, more tangible, properties of the network architecture. This approach is explained in the next section.

3   A Dynamic Mean-Field Approach

3.1   The Model

We propose a dynamic process where every period individuals (chosen randomly from the population) have the opportunity of revising their strategy. Two complementary approaches are considered. On the one hand, we study the so-called mean-field version of the model, where agents interact every period with neighbors which are a random draw from the population. In other words, it is as if at each period the network were generated randomly. On the other hand, we run simulations of the dynamic model on fixed networks (but randomly generated) and, in particular, compare the long-run outcomes for two networks with different connectivity distributions (a scale-free and an exponential distribution). The first approach is analyzed in detail next, whereas the second approach will be partially tackled in Section 4. A more extensive simulation study will be postponed for future research.

For the mean-field version of the model, the network is characterized by

the distribution of the number of neighbors, or connectivity, that each agent has. In particular, let $P(k)$ be the fraction of agents in the population with k neighbors. That is, $0 \leq P(k) \leq 1$ for all $k \geq 1$ and $\sum_{k \geq 1} P(k) = 1$.

We consider the following continuous-time dynamics to describe the evolution of agents' choices through time. At each time t an agent revises his strategy at a rate $\lambda \geq 0$ and chooses a myopic-best response. This implies that an agent will choose action 1 if and only if none of his neighbors is doing so. Agents are characterized by their connectivity which remains fixed throughout the dynamics. The set of neighbors, however, is chosen randomly from the population every period. Note that which neighbor is chosen is also a function of the recipients connectivity. To make this idea precise, consider the following notation. Let $\rho_k(t)$ be the proportion of agents with connectivity k that are choosing action 1 at time t. The probability that any given link points to an agent choosing 1 at time t is denoted by $\theta(t)$ and can be calculated as

$$\theta(t) = \frac{1}{\langle k \rangle} \sum_{k \geq 1} k P(k) \rho_k(t) \tag{3}$$

where $\langle k \rangle$ is the average connectivity which can be computed as

$$\langle k \rangle = \sum_{k \geq 1} k P(k).$$

Notice that $\frac{k P(k)}{\langle k \rangle}$ corresponds to the usual calculation of the probability of the connectivity of an agent conditional on that agent being at the end of a randomly chosen link in the network.

The mean-field dynamic equation represents the evolution of the fraction of agents with connectivity k choosing 1. Specifically,

$$\frac{d\rho_k(t)}{dt} = -\rho_k(t)\lambda(1 - (1 - \theta(t))^k) + (1 - \rho_k(t))\lambda(1 - \theta(t))^k \tag{4}$$

where λ is the rate of strategy revision and $(1 - \theta(t))^k$ is the probability of having no neighbor choosing 1 and therefore deciding to choose 1 yourself. Notice that, eq. (4) says the following: the variation of the relative density of agents with connectivity k choosing 1 at time t equals the proportion of agents

8

with connectivity k choosing 0 that switch to 1 at time t minus the proportion of agents with connectivity k choosing 1 that switch to 0 at time t.

Note that, eq. (4) is a deterministic approximation of the stochastic dynamics described in words above. This approximation is appropriate when dealing with large populations as described in Benaim and Weibull (2003). Specifically, they show that if the deterministic population flow remains forever in some subset of the state space, then the stochastic process will remain in the same subset space for a very long time with a probability arbitrarily close to one, granted that the population is large enough. Thus, hereafter, we will assume that the population is infinite.

After simplifications of eq. (4) we find that

$$\frac{d\rho_k(t)}{dt} = \lambda(-\rho_k + (1 - \theta(t))^k) \tag{5}$$

and, imposing the stationary condition $\frac{d\rho_k(t)}{dt} = 0$ we obtain the equation valid for the behavior of the system at large times,

$$\rho_k(\theta) = (1 - \theta)^k. \tag{6}$$

Then, substituting eq. (6) in eq. (3) we find that in equilibrium θ^* is the solution of the following fixed-point equation

$$\theta = H(\theta) \equiv \frac{1}{\langle k \rangle} \sum_{k \geq 1} kP(k)(1 - \theta)^k. \tag{7}$$

Once we know θ^* we can solve for the overall fraction of individuals choosing 1 in the stationary state as

$$\rho^*(\theta) = \sum_{k \geq 1} P(k)(1 - \theta^*)^k. \tag{8}$$

3.2   The Results

In this section we present the main results of the paper. We want to study how the properties of the connectivity distribution affect the mean-field equilibrium outcomes. There are two relevant values in equilibrium. On the one hand, θ^* is

9

the number of links that point to an agent exerting effort. On the other hand, ρ^* is the fraction of agents exerting effort.

Let us point out that in the mean-field version of the dynamics individuals could, in principle, switch actions (from 0 to 1 and vice-versa) at any moment in time given that the network is randomly generated every period. Therefore, the concept of stationary states only refers to stationary values of θ^* and ρ^*.

The next result addresses the issue of existence and uniqueness of the equilibrium.

Proposition 2 *There exists a unique θ^* and ρ^* in equilibrium. Furthermore, the equilibrium is globally stable.*

Proof. The value of θ^* in equilibrium is the solution of eq. (7). Denote by $\widetilde{H}(\theta) \equiv H(\theta) - \theta$ and notice that $\widetilde{H}(\theta)$ is a continuous and decreasing function of θ.[3] Also note that $\widetilde{H}(0) = 1$ and $\widetilde{H}(1) = -1$. Thus, there exists a unique solution θ^* of equation (7). Given θ^*, we can find the value of ρ^* which is therefore also unique. To conclude, let us show that θ^* is globally stable in the following sense. Starting from any initial value of agents choosing 1, the dynamics converges to a state where $\theta = \theta^*$ (and thus, $\rho = \rho^*$). Notice that

$$\frac{d\theta(t)}{dt} = \frac{1}{\langle k \rangle} \sum_{k \geq 1} kP(k)\frac{d\rho_k(t)}{dt}$$

and substituting $\frac{d\rho_k(t)}{dt}$ for its value determined by eq. (5) we obtain that

$$\frac{d\theta(t)}{dt} = \frac{1}{\langle k \rangle} \sum_{k \geq 1} kP(k)\lambda(-\rho_k + (1 - \theta(t))^k)$$

or equivalently

$$\frac{d\theta}{dt} = \lambda(H(\theta(t)) - \theta(t))$$

which proves the result. ■

We next concentrate on the comparison of higher and lower density networks. To further simplify the question assume that networks are homogeneous.

---

[3]Note that $\frac{d\widetilde{H}(\theta)}{d\theta} = -\frac{1}{\langle k \rangle} \sum_{k \geq 1} kP(k)k(1 - \theta)^{k-1} - 1$ is negative for all $0 \leq \theta \leq 1$.

That is, all agents have the same connectivity. Let $\theta^*(k)$ and $\rho^*(k)$ represent the values of θ and ρ in equilibrium when the network is homogeneous and with connectivity k. The following result holds.

Proposition 3 *Consider two homogeneous networks with connectivities k and $\overline{k}$ respectively, and such that $k \leq \overline{k}$. In equilibrium $\theta^*(\overline{k}) \leq \theta^*(k)$ and $\rho^*(\overline{k}) \leq \rho^*(k)$.*

Proof. Notice that $(1 - \theta)^{\overline{k}} \leq (1 - \theta)^k$ for all $\theta \in [0, 1]$. Then, the solution of eq. (7) is larger for the network with connectivity k than for the network with connectivity $\overline{k}$. This implies that $\overline{\theta}^* \leq \theta^*$. In this case $\theta^* = \rho^*$ and $\overline{\theta}^* = \overline{\rho}^*$, which is why $\overline{\rho}^* \leq \rho^*$. ∎

This result is intuitive because if agents have higher connectivity they observe a larger fraction of the population and therefore, the chances of having at least one neighbor exerting effort are very high. Hence, more free riding is plausible in equilibrium.[4] To further investigate this question let us consider two connectivity distributions $P(k)$ and $\overline{P}(k)$, where $\overline{P}(k)$ First Order Stochastic Dominates (FOSD) $P(k)$. Our next result suggests that under some conditions the fraction of free riders is higher when the connectivity distribution is $\overline{P}(k)$.

Proposition 4 *Let $\overline{P}(k)$ First Order Stochastic Dominate $P(k)$ and let k_m be the minimum connectivity in the network. If*

$$\frac{1}{\langle k \rangle} \sum_{k \geq 1} k\overline{P}(k)e^{\frac{-k}{k_m}} \geq 1 - e^{\frac{-1}{k_m}} \tag{9}$$

then

$$\overline{\theta}^* \leq \theta^*.$$

First of all, it is straightforward to show that for sufficiently large values of k_m condition (9) holds. This result again suggests that the higher the average

---

[4]Notice that in this context the concept of free riding simply refers to the fact that individuals do not exert effort (or provide the public good) because a neighbor does so, but we are not claiming any welfare implications derived from this behavior.

connectivity in the network the larger the fraction of free riders in the popula-
tion. As formally explained in footnote 5, the order between $\overline{\rho}^*$ and ρ^* cannot
be precisely established and might depend on additional properties of $P(k)$
and $\overline{P}(k)$. However, given the result and intuition derived for homogeneous
networks we conjecture that, in most cases, $\overline{\rho}^* \leq \rho^*$.

Proof. We must prove that $H_{\overline{P}(k)}(\theta) \leq H_{P(k)}(\theta)$ for all values of θ in the
interval $[\overline{\theta}^*, 1]$ because if this is so then $\overline{\theta}^* \leq \theta^*$ also holds. Let us first show that
$g(k) = k(1-\theta)^k$ is decreasing with respect to $k \geq k_m$ given that $\theta \in [1-e^{\frac{-1}{k_m}}, 1]$.
Notice that

$$\frac{dg}{dk} = (1-\theta)^k(1 + kln(1-\theta))$$

which is negative if and only if

$$1 + kln(1-\theta) \leq 0$$

or equivalently

$$\frac{-1}{ln(1-\theta)} \leq k.$$

Denote by θ_m the value of θ satisfying

$$\frac{-1}{ln(1-\theta_m)} = k_m.$$

Then

$$\theta_m = 1 - e^{\frac{-k}{k_m}}.$$

It is straightforward to show that if $\frac{-1}{ln(1-\theta_m)} \leq k$ then it is also the case
that $\frac{-1}{ln(1-\theta)} \leq k$ for all $\theta \geq \theta_m$. Therefore, $g(k)$ is decreasing for all $k \geq k_m$
given that $\theta \in [\theta_m, 1]$. Since $\overline{P}(k)$ FOSD $P(k)$ and $g(k)$ is decreasing then

$$H_{\overline{P}(k)}(\theta) \leq H_{P(k)}(\theta) \tag{10}$$

for $\theta \in [\theta_m, 1]$ because

$$\sum_{k \geq 1} k(1-\theta)^k \overline{P}(k) \leq \sum_{k \geq 1} k(1-\theta)^k P(k)$$

and since $\langle k \rangle \leq \langle \overline{k} \rangle$ then

$$\frac{1}{\langle \overline{k} \rangle} \sum_{k \geq 1} k \overline{P}(k)(1-\theta)^k \leq \frac{1}{\langle k \rangle} \sum_{k \geq 1} kP(k)(1-\theta)^k.$$

Condition (9) implies that $\theta_m < H_{\overline{P}(k)}(\theta_m)$ and therefore $\theta_m \leq \overline{\theta}^*$ which in turn implies that condition (10) also holds for $\theta \in [\overline{\theta}^*, 1]$. This completes the proof.[5] ∎

Consider two networks with the same average connectivity, but where one is a Mean Preserving Spread (MPS) of the other. Which network has a larger fraction of free-riders in equilibrium? The next result partially addresses this question.

Proposition 5 *Let $P(k)$ be a Mean Preserving Spread of $\overline{P}(k)$ and let k_m be the minimum connectivity in the network. If*

$$\frac{1}{\langle k \rangle} \sum_{k \geq 1} k \overline{P}(k) e^{\frac{-2k}{k_m}} \geq 1 - e^{\frac{-2}{k_m}} \tag{11}$$

then

$$\overline{\theta}^* \leq \theta^*.$$

Again, notice that, condition (11) holds for sufficiently large values of k_m. The proof of this result resembles the proof of Proposition 4 and is described next.

Proof. We must prove that $H_{\overline{P}(k)}(\theta) \leq H_{P(k)}(\theta)$ for all values of θ in the interval $[\overline{\theta}^*, 1]$ because if this is so then $\overline{\theta}^* \leq \theta^*$ also holds. Let us first show that $g(k) = k(1-\theta)^k$ is convex with respect to $k \geq k_m$ given that $\theta \in [1 - e^{\frac{-2}{k_m}}, 1]$. Notice that

$$\frac{d^2 g}{d^2 k} = (1-\theta)^k (2 + kln(1-\theta))ln(1-\theta)$$

---

[5] To actually compare the values ρ^* and $\overline{\rho}^*$ one would have to compare $\sum_{k \geq 1} P(k)(1-\theta^*)^k$ and $\sum_{k \geq 1} \overline{P}(k)(1-\overline{\theta}^*)^k$. Since $\overline{P}(k)$ FOSD $P(k)$ and $(1-\theta^*)^k$ is decreasing with respect to k then $\sum_{k \geq 1} P(k)(1-\theta^*)^k \geq \sum_{k \geq 1} \overline{P}(k)(1-\theta^*)^k$ but notice that $(1-\overline{\theta}^*)^k \geq (1-\theta^*)^k$ which is why the comparison between ρ^* and $\overline{\rho}^*$ might depend on further details of $P(k)$ and $\overline{P}(k)$.

which is positive if and only if

$$2 + k\,ln(1 - \theta) \leq 0$$

or equivalently

$$\frac{-2}{ln(1 - \theta)} \leq k.$$

Denote by θ_m the value of θ satisfying

$$\frac{-2}{ln(1 - \theta_m)} = k_m.$$

Then

$$\theta_m = 1 - e^{\frac{-2k}{k_m}}.$$

It is straightforward to show that if $\frac{-2}{ln(1-\theta_m)} \leq k$ then it is also the case that $\frac{-2}{ln(1-\theta)} \leq k$ for all $\theta \geq \theta_m$. Therefore, $g(k)$ is convex for all $k \geq k_m$ given that $\theta \in [\theta_m, 1]$. Since $P(k)$ is a MPS of $\overline{P}(k)$ and $g(k)$ is convex then

$$H_{\overline{P}(k)}(\theta) \leq H_{P(k)}(\theta) \tag{12}$$

for $\theta \in [\theta_m, 1]$. Condition (11) implies that $\theta_m < H_{\overline{P}(k)}(\theta_m)$ and therefore $\theta_m \leq \overline{\theta}^*$ which in turn implies that condition (12) holds also for $\theta \in [\overline{\theta}^*, 1]$. This completes the proof.[6] ∎

We conjecture that generally θ^* and ρ^* are aligned in the sense that if θ^* (ρ^*) is higher for one connectivity distribution than for another, so is ρ^* (θ^*). We have not been able to show this analytically which is why we study some examples in the next section. However, the intuition for why the fraction of agents exerting effort is higher in high-variance networks (provided that θ^* is also high for high variance networks as proved in Proposition 5) is the following. Notice that a high variance network has a significant number of hubs

---

[6] To actually compare the values ρ^* and $\overline{\rho}^*$ one would have to compare $\sum_{k \geq 1} P(k)(1 - \theta^*)^k$ and $\sum_{k \geq 1} \overline{P}(k)(1 - \overline{\theta}^*)^k$. Since $P(k)$ is a MPS of $\overline{P}(k)$ and $(1 - \theta^*)^k$ is convex with respect to k then $\sum_{k \geq 1} P(k)(1 - \theta^*)^k \geq \sum_{k \geq 1} \overline{P}(k)(1 - \theta^*)^k$ but notice that $(1 - \overline{\theta}^*)^k \geq (1 - \theta^*)^k$ which is why the comparison between ρ^* and $\overline{\rho}^*$ might depend on further details of $P(k)$ and $\overline{P}(k)$.

(i.e., nodes with very high connectivity) and a significant fraction of nodes with low connectivity. Thus, there are high chances that a low connectivity node ends up connected with a high connectivity node. Obviously high connectivity nodes are going to often free ride from one of their neighbors and therefore often choose 0, whereas the peripheral agents connected to them would often choose 1. Given Proposition 5, the number of links pointing to an agent choosing 1 is higher in a high variance network, therefore the only way ρ^* could be reversed (i.e., higher for low variance networks) would be if the hubs of the high variance networks are the ones exerting effort. Nevertheless, this is not generally going to be the case given the argument stated above.

3.3 An Example

Consider networks where agents are of two types: high connectivity agents and low connectivity agents. In particular, assume that m is the average connectivity of the network and the connectivity distribution is given by:

$$
P_{m,v}(k) = \begin{cases} \frac{1}{2} & \text{if } k = m + v \\[2mm] \frac{1}{2} & \text{if } k = m - v \end{cases} \tag{13}
$$

where v indicates how far from the mean are the two connectivities. We substitute this connectivity distribution in eq.(7) and (8) to analyze the collective outcomes for different values of the parameters m and v determine. We concentrate on the question of how the variance of the connectivity distribution determines the fraction of free riders in equilibrium. To do so, we consider a fixed value of m and observe what happens with the equilibrium outcome as v takes values from 0 to approximately $m - 1$. In fig.1 we represent the fraction of agents exerting effort in equilibrium ρ^* as a function of v. Each graph represents a different value for the average connectivity ($m = 10$, $m = 100$ and $m = 1000$, in the first, second and third graph respectively). Notice that qualitatively all curves are very similar. That is, the higher the variance of the network (higher v) the higher the fraction of agents choosing 1 and thus the
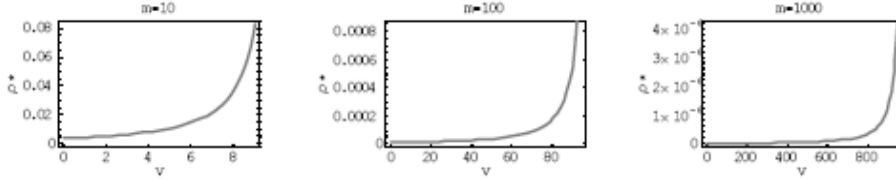
15

Figure 1: Representation of the fraction of agents exerting effort in equilibrium (ρ^*) as a function of v, for the values of $m = 10$, $m = 100$ and $m = 1000$.

lower the fraction of free-riders. Also, comparing one graph with the other, we find that going from a low to a high average connectivity network decreases the number of agents exerting effort and thus increases the fraction of free riders.

One can also show that if instead of equilibrium value of ρ^* we represent the equilibrium value of θ^* the graphs look roughly the same.

4   Simulations

In this section we investigate the role of the variance of the connectivity distribution on the fraction of free riders through simulations of the dynamics on fixed networks. We generate two different random networks in terms of their connectivity distributions:[7] (i) a scale-free network with connectivity distribution $P(k) \propto k^{-2.5}$ for $k \geq 3$ and $P_{SF}(k) = 0$ otherwise (where $\propto$ means, equal up to a multiplicative constant) (ii) an exponential network $P(k) \propto e^{-\frac{k}{6}}$ for $k \geq 3$ and $P_E(k) = 0$ otherwise.[8] Each of these two networks have a total of 1000 nodes and an average connectivity of approximately 9, however, scale-free networks have significantly larger variance than exponential networks. The in-

---

[7] The main reason why the dynamics on random networks might reproduce the mean-field approximations is that in a random network the characteristics of any given node is unaffected by structural correlations.

[8] These networks were generated using the program *Pajek*, a software package for Large Network Analysis.

terest in the study of scale-free networks is enhanced by the empirical evidence that many paradigmatic examples of complex networks such as the WWW and the human sexual contact network, among others, are characterized by scale-free connectivity properties (see e.g., Barabasi et al., 2000 and Lijeros et al., 2001).

We consider the discrete version of the continuous dynamics used to derive the theoretical results. In this respect, we assume that, in every period one (and only one) agent is chosen to revise his strategy and chooses action 1 if and only if none of his neighbors doing so. The network remains fixed throughout the dynamics which marks a crucial difference between the mean-field approach and the simulation.

In fig. 2 below we represent the number of agents exerting effort (or public good providers, PGP) as a function of time, where $1 \leq t \leq 10000$ for the scale-free network (red) and the exponential network (blue). We considered two different initial conditions; one where the number of PGP is 0 and the other where this number is 1000 (i.e., all agents are initially choosing action 1). Although for clarity in the graph we have simply represented a sample of 12 simulations per network, we have obtained the same convergence result when considering a sample size of 100.

The main findings are the following. On the one hand, there is a unique long-run prediction which could suggest that for each network the maximal independent sets have all approximately the same size. Thus, as obtained in the mean-field analysis there is a unique equilibrium value of ρ (and θ). On the other hand, the number of agents exerting effort in the scale-free network is significantly higher than in the exponential network. This reinforces the idea already suggested by the mean-field approximations, namely, higher variance networks lead to a lower fraction of free riders in the population.
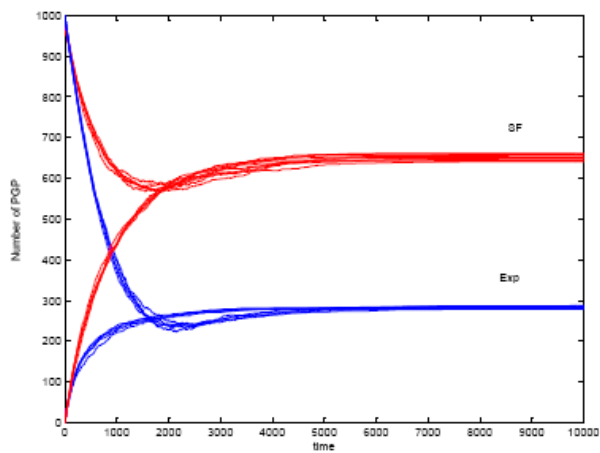
Figure 2: Representation of the number of agents exerting effort ρ (or public good providers, PGP) as a function of time $1 \leq t \leq 10000$ for a scale-free (SF, red in the graph) and an exponential (Exp, blue in the graph) network. The number of simulations are 12 per network and half start with an initial condition of $\rho_0 = 0$ and half with an initial condition of $\rho_0 = 1$.

5 Discussion

It is well known that the outcomes of many socioeconomic phenomena crucially depend on the properties of the social networks in which the agents are embedded. Some paradigmatic examples of processes where the social network is relevant are the diffusion of a new technology (e.g., Lopez-Pintado, 2004, 2006, Jackson and Yariv, 2006), the spread of information on job opportunities (e.g., Calvo-Armengol and Jackson, 2006), the uprising of political revolts (e.g., Chwe, 2000), or even the diffusion of an infectious disease (e.g., Pastor-Satorras and Vespignani, 2001). Whereas in most of this literature the behavior (or states) propagate in such a way that the more agents adopting the behavior the more it spreads, this paper studies a stylized model of precisely the polar case. That is, we study a process where agents want to anti-coordinate with their neighbors.

There are a few other papers that analyze this problem theoretically. For instance, Bramoulle (2007) addresses the case of anti-coordination games played among agents connected in a social network. Apart from the methodology used, the model proposed in Bramoulle (2007) differs from the model considered in this paper because he assumes that agents adopt an action whenever less than a certain fraction of their neighbors is doing so, whereas here we assume that one neighbor alone generates enough incentives to free ride.

Bramoulle and Kranton (2006) is perhaps the closest paper to this work, as they also analyze local public goods in terms of effort decisions among agents in a network. The main differences between their model and the model studied here are the following. First, Bramoulle and Kranton (2006) assume that agents choose a level of effort which is a continuous decision instead of a binary one. Second, Bramoulle and Kranton (2006) propose a utility function for each agent with the feature that the benefits each agent obtains from interacting with his neighbors are a concave function of the total level of effort exerted in the agent's neighborhood. In our model, the utility function is not necessarily concave on

19

benefits, but nevertheless we make the assumption that agents would only want to exert effort if no other neighbor does so, which in some sense restricts the problem even further. One of the disadvantages of not specifying the utility function is that we cannot engage in welfare comparisons of the equilibrium outcomes. The main novelty of our approach, however, is the methodology used in order to address questions that otherwise would be intractable. The mean-field approximations simplify the problem in a way so that we are able to compare network structures in terms of the properties of their connectivity distributions.

Finally, it is worth commenting on other works dealing with mean-field approximations of a network model such as Pastor-Satorras and Vespignani (2001), Lopez-Pintado (2004, 2006), Jackson and Yariv (2006), Rogers and Jackson (2007), etc. Most of these papers concentrate on the case where agents become more susceptible to a certain action the more neighbors adopt it. Lopez-Pintado (2004) studies a family of contagion processes with this feature, whereas Lopez-Pintado (2006) concentrates on a specific context where agents play a coordination game with each neighbor. Finally, Jackson and Yariv (2006) introduce a general diffusion of behavior model where not only coordination but also anti-coordination processes are considered. The current paper, however, can be viewed as a more detailed analysis of a specific scenario. It is worth mentioning that as Jackson and Yariv (2006) explain, the equilibrium outcomes of the mean-field dynamics could also be thought of as the symmetric Bayesian equilibrium outcomes of a Bayesian game where agents, characterized by their connectivities (types), simultaneously choose their actions knowing the connectivity distribution in the population, and assuming that connectivities are independently allocated throughout the network (see also Galeotti et al., 2005 for another use of this same approach).

References

[1] Benaim M. and J. Weibull, "Deterministic Approximation of Stochastic Evolution in Games", (2003) Econometrica 71, 873-903.

[2] Barabasi A., Albert R. and H. Jeong, "Scale-free Characteristics of Random Networks: the Topology of the World-Wide Web", (2000) Physica A 281, 2115.

[3] Bramoulle Y. and R. Kranton, "Public Goods in Networks" (2006) Journal of Economic Theory, forthcoming.

[4] Bramoulle Y., "Anti-Coordination Games and Social Interactions", (2007) Games and Economic Behavior 58, 30-49.

[5] Calvo-Armengol A. and M. Jackson, "The Effects of Social Networks on Employment and Inequality" (2004), American Economic Review 94 (3), 426-454.

[6] Chwe M. S-Y., "Communication and Coordination in Social Networks", (2000) The Review of Economic Studies 67, 1-16.

[7] Galeotti A., Goyal S., Jackson M. O., Vega-Redondo F. and L. Yariv, "Network Games", (2005) mimeo.

[8] Jackson M. O. and B. Rogers, "Relating Network Structures to Diffusion Properties through Stochastic Dominance", (2007) The B. E. Journal of Theoretical Economics 7(1) (Advances), 6.

[9] Jackson M.O. and L. Yariv, "Diffusion of Behavior and Equilibrium Properties in Network Games", (2006) American Economic Review (Papers and Proceedings), forthcoming.

[10] Lijeros F., Edling C. R., Amaral L. A. N., Stanley H. E., and Y. Aberg, "The Web of Human Sexual Contacts", (2001) Nature 441, 907-908.

[11] Lopez-Pintado D., "Contagion and Coordination in Random Networks", (2006) International Journal of Game Theory 34, 371-381.

[12] Lopez-Pintado D., "Diffusion in Complex Networks", (2004) IVIE WP-AD 2004-33, Universidad de Alicante.

[13] Pastor-Satorras P. and A. Vespignani, "Epidemic Dynamics and Endemic States in Complex Networks", (2001) Physical Review E 63, 066117.