Fondazione Eni Enrico Mattei Working Papers
Fondazione Eni Enrico Mattei

Year 2010 *Paper* 512

Stochastic Stability in the Best Shot Game

Leonardo Boncinelli
Università degli Studi di Siena

Paolo Pin
Università degli Studi di Siena

Stochastic Stability in the Best Shot Game

Abstract

The best shot game applied to networks is a discrete model of many processes of contribution to local public goods. It has generally a wide multiplicity of equilibria that we refine through stochastic stability. In this paper we show that, depending on how we define perturbations, i.e. the possible mistakes that agents can make, we can obtain very different sets of stochastically stable equilibria. In particular and non-trivially, if we assume that the only possible source of error is that of an agent contributing that stops doing so, then the only stochastically stable equilibria are those in which the maximal number of players contributes.

Stochastic Stability in the Best Shot Game

Leonardo Boncinelli[*] Paolo Pin[†]

July 12, 2010

Abstract

The best shot game applied to networks is a discrete model of many processes of contribution to local public goods. It has generally a wide multiplicity of equilibria that we refine through stochastic stability. In this paper we show that, depending on how we define perturbations, i.e. the possible mistakes that agents can make, we can obtain very different set of stochastically stable equilibria. In particular and non–trivially, if we assume that the only possible source of error is that of an agent contributing that stops doing so, then the only stochastically stable equilibria are those in which the maximal number of players contributes.

JEL classification code: C72, C73, D85, H41.

Keywords: networks, best shot game, stochastic stability.

[*]Dipartimento di Economia Politica, Universitá degli Studi di Siena (Italy). email: boncinelli@unisi.it

[†]Dipartimento di Economia Politica, Universitá degli Studi di Siena (Italy). I acknowledge support from the project Prin 2007TKLTSR "Computational markets design and agent–based models of trading behavior". email: pin3@unisi.it

1 Introduction

In this paper we consider a stylized game of contribution to a discrete local public good where the range of externalities is defined by a network. With a small probability players may fail to play their best response and we analyze which equilibria are most stable to such errors. In particular, we show that the nature of the mistake has a fundamental role in determining the characteristics of such stable equilibria.

Let us start with an example.

Example 1. Ann, Bob, Cindy, Dan and Eve live in a suburb of a big city and they all have to take private cars in order to reach downtown every working day. They could share the car but they are not all friends together: Ann and Eve do not know each other but they both know Bob, Cindy and Dan, who also don't know each other. The network of relations is shown in Figure 1. In a one–shot equilibrium (the first working day) they will end up sharing cars. Any of our characters would be happy to give a lift to a friend, but we assume here that non–linked people do not know each other and would not offer each other a lift. No one would take the car if a friend is doing so, but someone would be forced to take it if none of her/his friends is doing so. There is a less congestioned equilibrium in which Ann and Eve take the car (and the other three take somehow a lift), and a more polluting one in which Bob, Cindy and Dan take their car (offering a lift to Ann and Eve, who will choose one of them).

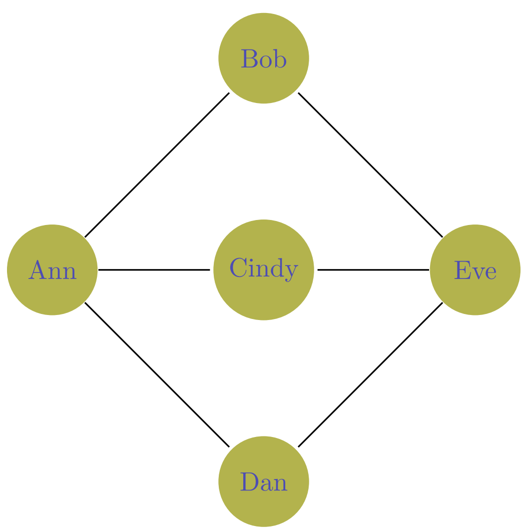


Figure 1: Five potential drivers in a network of relations.

Imagine to be in the less congestioned equilibrium. Now suppose that, even if they all agreed on how to manage the trip, in the morning Ann finds out that her car's engine is broken and she cannot start it. She will call her three friends, who are however not planning to take the car and will not be able to offer her a lift. As Ann does not know Eve, and Eve will be the only one left with a car, Ann will have to wait for her own car to be repaired before she can reach her workplace. Only if both cars of Ann and Eve break down, then Bob, Cindy and Dan will take their cars, and we will shift to the inefficient equilibrium. It is easy to see that if we start instead from the congestioned equilibrium, then we need three cars to break down before we can induce Ann and Eve to get their own. In this sense the *bad* equilibrium is more *stable*, as it needs a less likely event in order to be changed with another equilibrium. □

In this paper we analyze the best shot game:[1] in a fixed exogenous network of binary relations, each node (player) may or may not provide a local public good. The best response for each player is to provide the good if and only if no one of her neighbors is doing so.
In the previous example we have described an equilibrium where each player can take one of two actions: take or not take the car. Then we have included a possible source of *error*: the car may break down and one should pass from action 'take the car' to action 'not take the car'. Clearly we can also imagine a second type of error, e.g. if a player forgets that someone offered her/him a lift and takes her/his own car anyway. We think however that there are situations in which the first type of error is the only plausible one, as well as there can be cases in which the opposite is true, and finally cases where the two are both likely, possibly with different probabilities.

What we want to investigate in the present paper is how the likelihood of different kinds of error may influence the likelihood of different Nash Equilibria. Formally, we will analyze stochastic stability (Young, 1998) of the Nash equilibria of the best shot game, under different assumptions on the perturbed Markov chain that allows the agents to make errors.

What we find is that, if only errors of the type described in the example are possible, that is players can only make a mistake by not providing the public good even if that is their best response in equilibrium, then the only stochastically stable Nash equilibria are those that maximize the number of contributors. If instead the other type of error (i.e. provide the good even if it is a dominant action to free ride) is the only one admitted, or it is admitted with a relatively high probability, then every Nash equilibrium is stochastically stable.

The best shot game is very similar to the local public goods game of Bramoullé and Kranton (2007): they motivate their model with a story of neighborhood farmers, with

[1] This name for exactly the same game comes from Galeotti et al. (2010), but it stems back to the non–network application of Hirshleifer (1983).

reduced ability to check each others' technology (this is the network constraint), who can invest in experimenting a new fertilizer. They assume that the action set of players is continuous on the non–negative numbers (how much to invest in the new risky fertilizer), they define stable equilibria as those that survive small perturbations, and they find that stable equilibria are *specialized* ones, in which every agent either contributes an optimal amount (which is the same for all contributors) or free rides, so that their stable equilibria look like the equilibria of the discrete best shot game.

The main difference between our setup and the one of Bramoullé and Kranton (2007) is that in the best shot games that we study actions are discrete, errors, even if rare, are therefore more dramatic and the concept of stochastic stability naturally applies. We think that our model, even if stark, offers a valid intuition of why typical problems of congestion are much more frequently observed in some coordination problems with local externalities. Most of these problems deal with discrete choices. Traffic is an intuitive and appealing example,[2] while others are given in the introduction of Dall'Asta et al. (2010). In such complex situations we analyze those equilibria which are more likely to be the outcome of convergence, under the effect of local positive externalities and the possibility of errors.

In next section we formalize the best shot game. Section 3 describes the general best response dynamics that we apply to the game. In Section 4 we introduce the possibility of errors thus obtaining a perturbed dynamics, and we present the main theoretical analysis of the effects of different perturbation schemes. Finally, a brief discussion is in Section 5.

2 Best Shot Game

We consider a finite set of agents I of cardinality n. Players are linked together in a fixed exogenous network which is undirected and irreflexive; this network defines the range of a local externality described below. We represent such network through a $n \times n$ symmetric matrix G with null diagonal, where $G_{ij} = 1$ means that agents i and j are linked together (they are called *neighbors*), while $G_{ij} = 0$ means that they are not. We indicate with N_i the set of i's neighbors (the number of neighbors of a node is called its *degree* and is also its number of links). A *path* between two nodes i and j is an ordered set of nodes $(i, h_1, h_2 \dots h_\ell, j)$ such that $G_{ih_1} = 1$, $G_{h_1h_2} = 1$, …, $G_{h_\ell j} = 1$. A *connected* subset $J \subseteq I$

[2] Economic modelling of traffic have shown that simple assumptions can easily lead to congestion, even when agents are rational and utility maximizers (see Arnott and Small, 1994). Moreover, if we consider the discretization of the choice space, the motivation for the Logit model of McFadden (1973) were actually the transport choices of commuting workers.

is such that, for any $i, j \in J$, there is a path between i and j where all the elements of the path are members of J. Finally, a subset $H \subseteq I$ is *surrounding* a subset $J \subseteq I$ if, for any $h \in H$, we have that $h \notin J$ and there is $j \in J$ such that $G_{jh} = 1$.

Each player can take one of two actions, $x_i \in \{0, 1\}$ with x_i denoting i's action. Action 1 is interpreted as contribution, and an agent i such that $x_i = 1$ is called *contributor*. Similarly, action 0 is interpreted as defection, and an agent i such that $x_i = 0$ is called *defector*.[3] We will consider only pure strategies. A state of the system is represented by a vector $\mathbf{x}$ which specifies each agent's action, $\mathbf{x} = (x_1, \ldots, x_i, \ldots, x_n)$. The set of all states is denoted with X.

Payoffs are not explicitly specified. We limit ourselves to the class of payoffs that generate the same type of best reply functions.[4] In particular, if we denote with b_i agent i's best reply function that maps a state of the system into a utility maximizer, then:

$$b_i(\mathbf{x}) = \begin{cases} 1 & \text{if } x_j = 0 \text{ for all } j \in N_i \ , \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

We introduce some further notation in order to simplify the following exposition. We define the set of *satisfied agents* at state $\mathbf{x}$ as $S(\mathbf{x}) = \{i \in I : \ x_i = b_i(\mathbf{x})\}$. Similarly, the set of *unsatisfied agents* at state $\mathbf{x}$ is $U(\mathbf{x}) = I \setminus S(\mathbf{x})$. We also refer to the set of contributors as $C(\mathbf{x}) = \{i \in I : \ x_i = 1\}$, and to the set of defectors as $D(\mathbf{x}) = \{i \in I : \ x_i = 0\}$. We also define intersections of the above sets: the set of *satisfied contributors* is $S^C(\mathbf{x}) = S(\mathbf{x}) \cap C(\mathbf{x})$, the set of *unsatisfied contributors* is $U^C(\mathbf{x}) = U(\mathbf{x}) \cap C(\mathbf{x})$, the set of *satisfied defectors* is $S^D(\mathbf{x}) = S(\mathbf{x}) \cap D(\mathbf{x})$, and the set of *unsatisfied defectors* is $U^D(\mathbf{x}) = U(\mathbf{x}) \cap D(\mathbf{x})$. Finally, given any pair of states $(\mathbf{x}, \mathbf{x}')$ we indicate with $K(\mathbf{x}, \mathbf{x}') = \{i \in I : \ x_i = x'_i\}$ the set of agents that keep the same action in both states, and we indicate with $M(\mathbf{x}, \mathbf{x}') = I \setminus K(\mathbf{x}, \mathbf{x}')$ the set of agents whose action is modified between the states.

The above game is called *best shot game*. A state $\mathbf{x}$ is a pure strategy Nash equilibrium of the best shot game if and only if $S(\mathbf{x}) = I$ and consequently $U(\mathbf{x}) = \emptyset$. We will call all

[3] As will be clear below, we are dealing with a *local* public good game, so probably *free rider* would be a more suitable term than *defector*. Nevertheless, in the public goods game also "defector" is often used.

[4] Note that it would be very easy to define specific payoffs that generate the best reply defined by (1): imagine that the cost for contributing is c and the value of a contribution, either from an agent herself and/or from one of her neighbors (players are satiated by one unit of contribution in the neighborhood), is $V > c > 0$. There are however many other payoff functions that could have the same best reply function (see Bramoullé et al. (2010) for other examples). As we consider the whole class we are not entering a discussion about welfare (i.e. aggregate payoffs – that could differ between the specific cases), but discuss only the issue of *congestion* (i.e. the aggregate number of contributors).

the possible Nash equilibria in pure strategies, given a particular network, as $\mathcal{N} \subseteq X$.

The set $\mathcal{N}$ is always non–empty but typically very large. It is an NP–hard problem to enumerate all the elements of $\mathcal{N}$,[5] and to identify, among them, those that maximize and minimize the set $C(\mathbf{x})$ of contributors. For extensive discussions on this point see Dall'Asta et al. (2009) and Dall'Asta et al. (2010). Here we provide two examples, the second one illustrates how even very homogeneous networks may display a large variability of contributors in different equilibria.

Example 2. Figure 2 shows two of the three possible Nash equilibria of the same 5–nodes network, where the five characters of our introductory example have now a different network of friendships. □

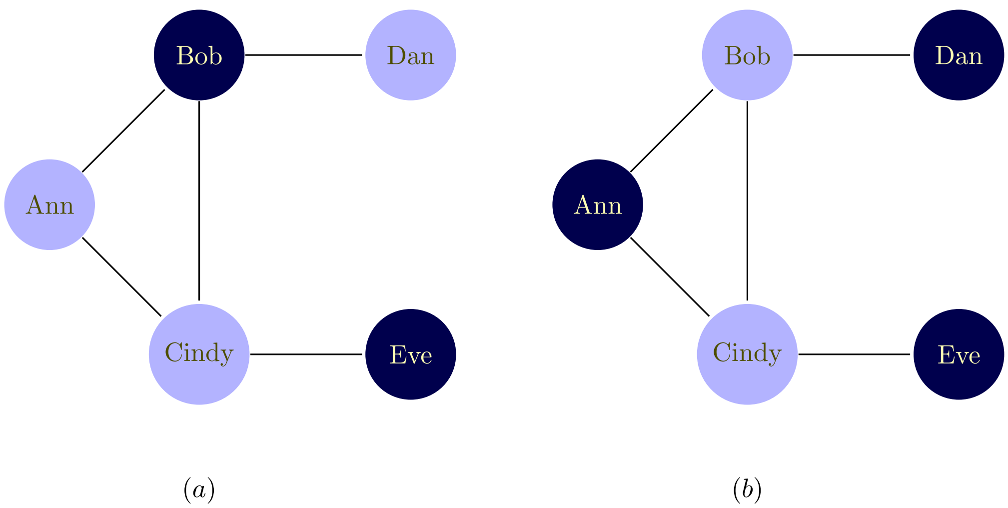


Figure 2: Two Nash equilibria for a 5–nodes network. The dark blue stands for contribution, while the light blue stands for defection.

Example 3. Consider the particular regular random network, of 20 nodes and homogeneous degree 4, that is shown in Figure 3. The relatively small size of this network allows us to count all its Nash equilibria. There exist 132 equilibria: 1 with 4 contributors (Figure 3, left), 17 with 5 contributors, 81 with 6 contributors, 32 with 7 contributors, 1 with 8 contributors (Figure 3, right). □

[5] In particular, all maximal independent sets can be found in time $O(3^{n/3})$ for a graph with n vertices (Tomita et al., 2006).

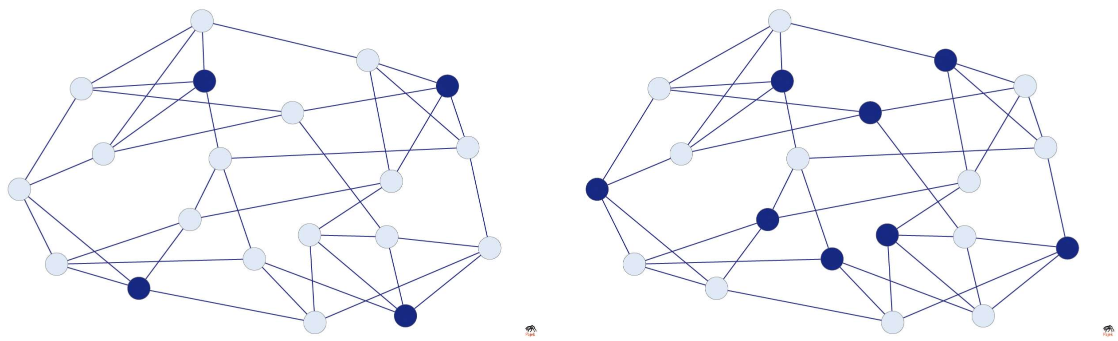


Figure 3: Two Nash equilibria for the same regular random network of 20 nodes and degree 4. The dark blue stands for contribution, while the light blue stands for defection. Picture is obtained by means of the software Pajek (http://pajek.imfm.si/).

3 Unperturbed Dynamics

We imagine a dynamic process in which the network G is kept fixed, while the actions $\mathbf{x}$ of the nodes change.

At each time, which is assumed discrete, every agent best replies to the previous state of the system with an i.i.d. positive probability $\beta \in (0,1)$, while with the complementary probability $(1-\beta)$ her action remains the same. If we denote with $\mathbf{x}$ the current state and with $\mathbf{x}'$ the state at next time, we can then formalize as follows:

$$x_i' = \begin{cases} b_i(\mathbf{x}) & \text{with i.i.d. probability } \beta \ , \\ x_i & \text{with i.i.d. probability } 1-\beta \ . \end{cases} \tag{2}$$

By so doing, a Markov chain (X, T) turns out to be defined, where X is the finite state space and T is the transition matrix[6] resulting from the individual update process in (2). We note that T depends on β.

It is easy to check that the Markov chain (X, T) satisfies the following property, which formalizes the idea that all and only the unsatisfied agents have the possibility to change action:[7]

$$T_{\mathbf{x}\mathbf{x}'} > 0 \text{ if and only if } M(\mathbf{x}, \mathbf{x}') \subseteq U(\mathbf{x}) \quad . \tag{3}$$

[6] $T_{\mathbf{x}\mathbf{x}'}$ denotes the probability to pass from state $\mathbf{x}$ to state $\mathbf{x}'$.

[7] It is the generalized property (3) that we exploit in all the following propositions. Our results on the unperturbed dynamics hold with any transition matrix satisfying that property. Note also that the Markov chain defined in (3), and hence in (2), is aperiodic because, as $M(\mathbf{x}, \mathbf{x}) = \emptyset$ for all $\mathbf{x}$, then $T_{\mathbf{x}\mathbf{x}} > 0$ for all $\mathbf{x}$.

We introduce some terminology from Markov chain theory following Young (1998). A state $\mathbf{x}'$ is called *accessible* from a state $\mathbf{x}$ if there exists a sequence of states, with $\mathbf{x}$ as first state and $\mathbf{x}'$ as last state, such that the system can move with positive probability from each state in the sequence to the next state in the sequence. A set $\mathcal{E}$ of states is called *ergodic set* (or *recurrent class*) when each state in $\mathcal{E}$ is accessible from any other state in $\mathcal{E}$, and no state out of $\mathcal{E}$ is accessible from any state in $\mathcal{E}$. If $\mathcal{E}$ is an ergodic set and $\mathbf{x} \in \mathcal{E}$, then $\mathbf{x}$ is called *recurrent*. Let $\mathcal{R}$ denote the set of all reccurrent states of (X, T). If $\{\mathbf{x}\}$ is an ergodic set, then $\mathbf{x}$ is called *absorbing*. Equivalently, $\mathbf{x}$ is absorbing when $T_{\mathbf{x}\mathbf{x}'} = 1$. Let $\mathcal{A}$ denote the set of all absorbing states of (X, T). Clearly, an absorbing state is recurrent, hence $\mathcal{A} \subseteq \mathcal{R}$.

In the next two propositions we show that in our setup the set $\mathcal{N}$ of Nash equilibria is equivalent to all and only the absorbing states (Proposition 1), and that there are no other recurrent states (Proposition 2).

Proposition 1. *$\mathcal{A} = \mathcal{N}$.*

Proof. We prove double inclusion, first we show that $\mathcal{N} \subseteq \mathcal{A}$.
Suppose $\mathbf{x} \in \mathcal{N}$. Since by (3) we have that $T_{\mathbf{x}\mathbf{x}'} > 0$ with $\mathbf{x}' \neq \mathbf{x}$ only if $U(\mathbf{x}) \neq \emptyset$, then $T_{\mathbf{x}\mathbf{x}'} = 0$ for any $\mathbf{x}' \neq \mathbf{x}$, hence $T_{\mathbf{x}\mathbf{x}} = 1$ and $\mathbf{x}$ is absorbing.

Now we show that $\mathcal{A} \subseteq \mathcal{N}$.
By contradiction, suppose $\mathbf{x} \notin \mathcal{N}$. Then $U(\mathbf{x}) \neq \emptyset$. Consider a state $\mathbf{x}'$ where $x_i' = x_i$ if $i \in S(\mathbf{x})$, and $x_i' \neq x_i$ otherwise. We have that $\mathbf{x}' \neq \mathbf{x}$ and, by (3), that $T_{\mathbf{x}\mathbf{x}'} > 0$, hence $T_{\mathbf{x}\mathbf{x}} < 1$ and $\mathbf{x}$ is not absorbing. $\square$

Proposition 2. *$\mathcal{A} = \mathcal{R}$.*

Proof. The first inclusion $\mathcal{A} \subseteq \mathcal{R}$ follows from the definitions of $\mathcal{A}$ and $\mathcal{R}$.

Now we show that $\mathcal{R} \subseteq \mathcal{A}$.
We prove that every element $\mathbf{x}$ which is not in $\mathcal{A}$ is also not in $\mathcal{R}$. Suppose that $\mathbf{x} \notin \mathcal{A}$. We identify, by means of a recursive algorithm, a state $\hat{\mathbf{x}}$ such that $\hat{\mathbf{x}}$ is accessible from $\mathbf{x}$, but $\mathbf{x}$ is not accessible from $\hat{\mathbf{x}}$. This implies that $\mathbf{x} \notin \mathcal{R}$.

By Proposition 1 we know that $\mathcal{A} = \mathcal{N}$. Then $\mathbf{x} \notin \mathcal{N}$ and we have that $U(\mathbf{x}) \neq \emptyset$. If $U^C(\mathbf{x}) \neq \emptyset$, we define $\mathbf{x}' \equiv \mathbf{x}$ and we go to Step 1, otherwise we jump to Step 2.

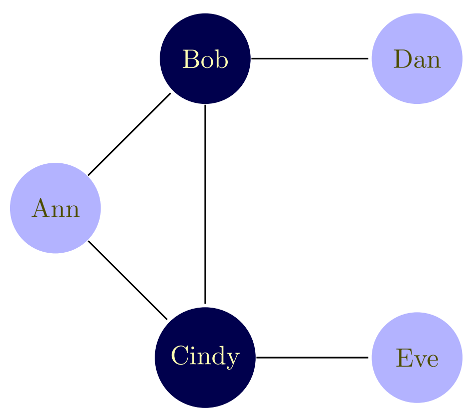
Step 1. We take $i \in U^C(\mathbf{x}')$ and we define state $\mathbf{x}''$ such that $x_i'' \equiv 0 \neq x_i' = 1$ and $x_j'' \equiv x_j'$ for all $j \neq i$.
Note that $||U^C(\mathbf{x}'')|| < ||U^C(\mathbf{x}')||$. This is because of two reasons: first of all, $i \in U^C(\mathbf{x}')$ and $i \notin U^C(\mathbf{x}'')$; the second is that $U^C(\mathbf{x}'') \subseteq U^C(\mathbf{x}')$, otherwise two neighbors contribute in $\mathbf{x}''$ and do not contribute in $U^C(\mathbf{x}')$, but that is not possible because $C(\mathbf{x}'') \subset C(\mathbf{x}')$. Moreover,

by (3) we have that $T_{\mathbf{x}'\mathbf{x}''} > 0$.
We redefine $\mathbf{x}' \equiv \mathbf{x}''$. Then, if $U^C(\mathbf{x}') = \emptyset$ we pass to Step 2, otherwise we repeat Step 1.

Step 2. We know that $U^C(\mathbf{x}') = \emptyset$. We take $i \in U^D(\hat{\mathbf{x}})$ and we define state $\mathbf{x}''$ such that $x''_i \equiv 1 \neq x'_i = 0$ and $x''_j \equiv x'_j$ for all $j \neq i$.
Note that $||U^D(\mathbf{x}'')|| < ||U^D(\mathbf{x}')||$. This is because of two reasons: first of all, $i \in U^D(\mathbf{x}')$ and $i \notin U^D(\mathbf{x}'')$; the second is that $U^D(\mathbf{x}'') \subseteq U^D(\mathbf{x}')$, otherwise two neighbors do not contribute in $\mathbf{x}''$ and do contribute in $U^C(\mathbf{x}')$, but that is not possible because $D(\mathbf{x}'') \subset D(\mathbf{x}')$.
We also note that still $U^C(\mathbf{x}'') = U^C(\mathbf{x}') = \emptyset$, since only i has become contributor and all i's neighbors are defectors.
By (3) we have that $T_{\mathbf{x}'\mathbf{x}''} > 0$. Finally, if $U^D(\mathbf{x}'') \neq \emptyset$ we redefine $\mathbf{x}' \equiv \mathbf{x}''$ and repeat Step 2, otherwise it means that $\hat{\mathbf{x}} = \mathbf{x}''$ and we have reached the goal of the algorithm.

The sequence of states we have constructed shows that $\hat{\mathbf{x}}$ is accessible from $\mathbf{x}$.
Since $U(\mathbf{x}') = \emptyset$, we have that $T_{\hat{\mathbf{x}}\hat{\mathbf{x}}} = 1$ by (3), and hence $\mathbf{x}$ is not accessible from $\mathbf{x}'$. □

An immediate corollary of Propositions 1 and 2 is that $\mathcal{R} = \mathcal{A} = \mathcal{N}$.



(c)

Figure 4: A non-Nash (non-absorbing) state for the same network of Figure 2. Here Bob and Cindy are contributing, while Ann, Dan and Eve are not.

Example 4. Consider the network from Figure 2: both states (a) and (b) shown there are absorbing, as they are Nash equilibria. Consider now the new state (c) on the same network, shown in Figure 4: the satisfied nodes here are only the defectors Ann, Dan and Eve. Both states (a) and (b) are accessible from state (c), but through different paths. To reach (a) from

(c), the unsatisfied contributor Cindy should turn to defection, so that Eve would become (the only) unsatisfied and would be forced to become a contributor. To reach (b) from (c), both the unsatisfied contributors Bob and Cindy should simultaneously turn to defection, then all the five nodes would be unsatisfied. If we now turn to contribution exactly Ann, Dan and Eve, we reach state (b). $\square$

The following Lemma 3 and Lemma 4 adapt the results in Lemma 2 of Dall'Asta et al. (2010) to our setup, as the dynamics employed there is different from ours. Both of them play an important role in the analysis of the perturbed dynamics that we develop in Section 4.

Lemma 3 states that if we start from a Nash equilibrium, we impose to an agent i a change from contribution to defection, and we let the dynamics T operate, then in no way agents that are neither agent i, nor neighbors of agent i, will ever change their action. In other words, under the above conditions the best reply dynamics is restricted to the neighborhood of agent i.

LEMMA **3.** *Suppose* $\mathbf{x} \in \mathcal{N}$ *and* $x_i = 1$. *Define* $\mathbf{x}'$ *such that* $\mathbf{x}'_i = 0$ *and* $\mathbf{x}'_j = \mathbf{x}_j$ *for all* $j \neq i$. *Then, for every state* $\mathbf{x}''$ *that is accessible from* $\mathbf{x}'$ *through* T *we have that if* $\mathbf{x}''_j \neq \mathbf{x}'_j$ *then either* $j = i$ *or* $j \in N_i$.

Proof. Call $J \equiv U(\mathbf{x}')$, J contains i (by assumption) and all and only nodes $j \in N_i$ (hence $\mathbf{x}'_j = 0$) such that there exists no $k \in N_j$, such that $x'_k = 1$. Define $H \equiv \{h \in I : \; h \in S(\mathbf{x}'),\ \exists j \in U(\mathbf{x}')$ such that $h \in N_j\}$, that is the set of all satisfied agents that are neighbors of some unsatisfied agent in $\mathbf{x}'$. In the topology of the network, J is a connected subset and H is the surrounding set of J, so that any path between any node $j \in J$ and $k \in I \setminus (J \cup H)$ contains at least a node $h \in H$. Note that for any $h \in H$ we have that $x'_h = 0$, and then h has at least a contributing neighbor $k \in I \setminus (J \cup H)$. Then h is satisfied independently of the action of any $j \in J$. As T allows only unsatisfied agents to update their action, this proves that for any $\mathbf{x}''$ that is accessible from $\mathbf{x}'$ through T, $U(\mathbf{x}'') \subseteq U(\mathbf{x}') = J$. As J contains only i and neighbors of i the statement is proven. $\square$

Lemma 4 states that if we start from a Nash equilibrium, we impose to an agent i a change from defection to contribution, and we let the dynamics T operate, then in no way agents that are neither agent i, nor neighbors of agent i, nor neighbors of neighbors of agent i, will ever change their action. In other words, under the above conditions the best reply dynamics is restricted to the neighborhood of neighbors of agent i.

Lemma 4. *Suppose* $\mathbf{x} \in \mathcal{N}$ *and* $x_i = 0$. *Define* $\mathbf{x}'$ *such that* $\mathbf{x}'_i = 1$ *and* $\mathbf{x}'_j = \mathbf{x}_j$ *for all* $j \neq i$. *Then, for every state* $\mathbf{x}''$ *that is accessible from* $\mathbf{x}'$ *through* T *we have that if* $\mathbf{x}''_j \neq \mathbf{x}'_j$ *then either* $j = i$ *or* $j \in N_i$ *or* $j \in N_k$ *for some* $k \in N_i$.

Proof. Call $J' \equiv U(\mathbf{x}')$, J' contains i (by assumption) and all and only nodes $j' \in N_i$ such that $\mathbf{x}'_{j'} = 1$. Call now J the set of all those nodes $j \in I$ such that there exists a node $j' \in J' \setminus \{i\}$ for which $j' \in N_j$ (hence $\mathbf{x}'_j = 0$) and such that there exists no $k \in N_j \setminus J'$, such that $x'_k = 1$. We can now define, as in the proof of Lemma 3, the subset $H \subseteq I$ surrounding J. The proof proceeds analogously. As J contains only i and neighbors of neighbors of i the statement is proven. □

4 Perturbed Dynamics

Given the multiplicity of Nash equilibria, we are uncertain about the final outcome of (X, T), that depends in part on the initial state and in part on the realizations of the probabilistic passage from states to states. In order to obtain a sharper prediction, which is also independent of the initial state, we introduce a small amount of perturbations and we use the techniques developed in economics by Foster and Young (1990), Young (1993), Kandori et al. (1993). Since the way in which perturbations are modeled has in general important consequences on the outcome of the perturbed dynamics (see Bergin and Lipman, 1996), we consider three specific perturbation schemes, each of which has its own interpretation and may better fit a particular application.

We introduce perturbations by means of a *regular perturbed Markov chain* (Young, 1993, see also Ellison, 2000), that is a triple $(X, T, (T^\epsilon)_{\epsilon \in (0,\bar{\epsilon})})$ where (X, T) is the unperturbed Markov chain and:

1. (X, T^ϵ) is an ergodic Markov chain, for all $\epsilon \in (0, \bar{\epsilon})$;
2. $\lim_{\epsilon \to 0} T^\epsilon = T$;
3. there exists a *resistance* function $r : X \times X \to \mathbb{R}^+ \cup \{\infty\}$ such that for all pairs of states $\mathbf{x}, \mathbf{x}' \in X$,

$$\begin{cases} \lim_{\epsilon \to 0} \dfrac{T^\epsilon_{\mathbf{x}\mathbf{x}'}}{\epsilon^{r(\mathbf{x},\mathbf{x}')}} \text{ exists and is strictly positive} & \text{if } r(\mathbf{x}, \mathbf{x}') < \infty \ ; \\ T^\epsilon_{\mathbf{x}\mathbf{x}'} = 0 \text{ for sufficiently small } \epsilon & \text{if } r(\mathbf{x}, \mathbf{x}') = \infty \ . \end{cases}$$

The resistance $r(\mathbf{x}, \mathbf{x}')$ is part of the definition and can be interpreted informally as the *amount of perturbations* required to move the system from $\mathbf{x}$ to $\mathbf{x}'$ with a single application of T^ϵ. It defines a weighted directed network between the states in X, where the weight of the passage from $\mathbf{x}$ to $\mathbf{x}'$ is equal to $r(\mathbf{x},\mathbf{x}')$. If $r(\mathbf{x},\mathbf{x}')=0$, then the system can move from state $\mathbf{x}$ directly to state $\mathbf{x}'$ in the unperturbed dynamics, that is $T_{\mathbf{x}\mathbf{x}'}>0$. If $r(\mathbf{x},\mathbf{x}')=\infty$, then the system cannot move from $\mathbf{x}$ directly to $\mathbf{x}'$ even in the presence of perturbations, that is $T^\epsilon_{\mathbf{x}\mathbf{x}'}=0$ for ϵ sufficiently small.

Even if T and r are defined on all the possible states of X, we can limit our analysis to the absorbing states only, which are all and only the recurrent ones (Proposition 2). This technical procedure is illustrated in Young (1998) and simplifies the complexity of the notation, without loss of generality. Given $\mathbf{x},\mathbf{x}'\in\mathcal{A}$, we define $r^*(\mathbf{x},\mathbf{x}')$ as the minimum sum of the resistances between absorbing states over any path starting in $\mathbf{x}$ and ending in $\mathbf{x}'$.

Given $\mathbf{x}\in\mathcal{A}$, an $\mathbf{x}$-tree on $\mathcal{A}$ is a subset of $\mathcal{A}\times\mathcal{A}$ that constitutes a tree rooted at $\mathbf{x}$.[8] We denote such $\mathbf{x}$-tree with $F_\mathbf{x}$ and the set of all $\mathbf{x}$-trees with $\mathcal{F}_\mathbf{x}$. The *resistance of an* $\mathbf{x}$*-tree*, denoted with $r^*(F_\mathbf{x})$, is defined to be the sum of the resistances of its edges, that is:

$$r^*(F_\mathbf{x}) \equiv \sum_{(\mathbf{x},\mathbf{x}')\in F_\mathbf{x}} r^*(\mathbf{x},\mathbf{x}') \ .$$

Finally, the *stochastic potential* of $\mathbf{x}$ is defined to be

$$\rho(\mathbf{x}) \equiv \min\{r^*(F_\mathbf{x}):\ F_\mathbf{x}\in\mathcal{F}_\mathbf{x}\} \ .$$

A state $\mathbf{x}$ is said *stochastically stable* (Foster and Young, 1990) if $\rho(\mathbf{x})=\min\{\rho(\mathbf{x}):\ \mathbf{x}\in\mathcal{A}\}$. Intuitively, stochastically stable states are those and only those states that the system can occupy after very long time has elapsed in the presence of very small perturbations.[9]

We first consider two extreme types of perturbations in Subsections 4.1, 4.2. Then we address cases that lie in between those extrema in Subsection 4.3.

4.1 Perturbations affect only the agents that are playing action 0

We assume that every agent playing action 0 can be hit by a perturbation which makes her switch action to 1. Each perturbation occurs with an i.i.d. positive probability $\epsilon\in(0,1)$. No agent playing action 1 can be hit by a perturbation. We define a transition matrix $P^{0,\epsilon}$ – that we call perturbation matrix – starting from individual probabilities, in the same way

[8] By *tree* we will refer only to this structure between absorbing states, and in no way to the topology of the underlying exogenous undirected network on which the best shot game is played.

[9] For a formal statement see Young (1993).

as we defined T from (2). We indicate with $\mathbf{x}$ the state prior to perturbations and with $\mathbf{x}'$ the resulting state:

$$x_i' = \begin{cases} \qquad\qquad x_i & \text{if } x_i = 1\ , \\ \begin{cases} 1 & \text{with i.i.d. probability } \epsilon\ , \\ 0 & \text{with i.i.d. probability } 1-\epsilon \end{cases} & \text{if } x_i = 0\ . \end{cases} \tag{4}$$

The perturbation matrix $P^{0,\epsilon}$ collects the probabilities to move between any two states in X when the individual perturbation process is as in (4). We assume our perturbed Markov chain to be such that first T applies and then errors can occur through $P^{0,\epsilon}$. We now check that $(X, T, (T^{0,\epsilon})_{\epsilon\in(0,\bar{\epsilon})})$, with $T^{0,\epsilon} = P^{0,\epsilon}T$, is indeed a regular perturbed Markov chain.[10]

1. $(X, T^{0,\epsilon})$ is ergodic for all positive ϵ: this can be seen applying the last sufficient condition for ergodicity in Fudenberg and Levine (1998, appendix of Chapter 5), once we take into account that i) $\mathcal{A} = \mathcal{R}$ by Proposition 2, and ii) $r^*(\mathbf{x}, \mathbf{x}') < \infty$ for all $\mathbf{x}, \mathbf{x}' \in \mathcal{A}$ by the following Lemma 5.

2. $\lim_{\epsilon\to 0} T^{0,\epsilon} = T$, since $\lim_{\epsilon\to 0} P^{0,\epsilon}$ is equal to the identity matrix.

3. The resistance function[11] is

$$r_0(\mathbf{x}, \mathbf{x}') = \begin{cases} ||S^D(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}')|| & \text{if } S^C(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}') = \emptyset \\ \infty & \text{otherwise} \end{cases} \tag{5}$$

   In fact,

   (a) if $r_0(\mathbf{x}, \mathbf{x}') = \infty$, then $S^C(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}') \neq \emptyset$ and there is no way to go from $\mathbf{x}$ to $\mathbf{x}'$, since no satisfied agent can change in the unperturbed dynamics and no contributor can be hit by a perturbation, hence $T^{\epsilon}_{\mathbf{x}\mathbf{x}'} = 0$ for every ϵ;

   (b) if $r_0(\mathbf{x}, \mathbf{x}') < \infty$, then $T^{0,\epsilon}_{\mathbf{x}\mathbf{x}'}$ has the same order of $\epsilon^{r(\mathbf{x},\mathbf{x}')}$, when ϵ approaches zero; in fact, the agents in $U(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}')$ can change independently with probability β when T is applied, the agents in $S^D(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}')$ can change independently

[10] We follow Samuelson (1997) when we derive the perturbed transition matrix $T^{0,\epsilon}$ by post-multiplying the unperturbed transition matrix T with the perturbations matrix $P^{0,\epsilon}$. If we exchange the order in the matrix multiplication some details should change (since matrix multiplication is not commutative), but all our results would still hold (as we iterate $T^{0,\epsilon}$ and matrix multiplication is associative).

[11] The necessary assumption for our results is to have a regular perturbed Markov chain whose resistance function is as in (5). By deriving it from an individual perturbation process (defined in (4)) and from an individual update process (defined in (2)), we show that there exists at least one significant case satisfying this property.

with probability ϵ when $P^{0,\epsilon}$ is applied (and only then), and no agent is left since $S^C(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}') = \emptyset$.

In the next remark we provide a lower bound for the resistance to move between Nash equilibria under this perturbation scheme, and we then use such remark in Proposition 6.

Remark 1. *When (5) holds, $r_0^*(\mathbf{x}, \mathbf{x}') \geq 1$ for all $\mathbf{x}, \mathbf{x}' \in \mathcal{N}$.*

The following lemma, which is of help in the proof of Proposition 6, shows that under this perturbation scheme any two absorbing states are connected through a sequence of absorbing states, with each step in the sequence having resistance 1.

Lemma **5.** *When (5) holds, for all $\mathbf{x}, \mathbf{x}' \in \mathcal{A}$, $\mathbf{x} \neq \mathbf{x}'$, there exists a sequence $\mathbf{x^0}, \ldots, \mathbf{x^s}, \ldots, \mathbf{x^k}$, with $\mathbf{x^s} \in \mathcal{A}$ for $0 \leq s \leq k$, $\mathbf{x^0} = \mathbf{x}$ and $\mathbf{x^k} = \mathbf{x}'$, such that $r_0^*(\mathbf{x^s}, \mathbf{x^{s+1}}) = 1$ for $0 \leq s < k$.*

Proof. Since $\mathbf{x} \neq \mathbf{x}'$, we have that $k \geq 1$. We set $\mathbf{x^0} = \mathbf{x}$. Suppose $\mathbf{x}^s$ is an element of the sequence, and take $i_s \in C(\mathbf{x}') \cap D(\mathbf{x}^s)$.

We define state $\tilde{\mathbf{x}}$ such that $\tilde{x}_{i_s} \equiv 1 \neq x^s_{i_s} = 0$ and $\tilde{x}_j \equiv x^s_j$ for all $j \neq i_s$. Note that $r_0(\mathbf{x^s}, \tilde{\mathbf{x}}) = 1$. We define state $\tilde{\mathbf{x}}'$ such that $\tilde{x}'_j \equiv 0$ for all $j \in N_{i_s}$ and $\tilde{x}'_k \equiv \tilde{x}_k$ for any other node k. We define state $\mathbf{x}^{s+1}$ such that $x^{s+1}_k = b_k(\tilde{\mathbf{x}}')$ for all $k \in N_j$, $j \in N_{i_s}$, and $x^{s+1}_\ell \equiv \tilde{x}'_\ell$ for any other node ℓ. By Lemma 4 and Proposition 1, $\mathbf{x}^{s+1} \in \mathcal{A}$. We note that $\mathbf{x}^{s+1}$ is obtained from $\tilde{\mathbf{x}}$ applying only the unperturbed dynamics T, hence the resistance $r_0^*(\mathbf{x}^s, \mathbf{x}^{s+1}) = 1$.

Note that, since $i_s \notin D(\mathbf{x^{s+1}})$ and $(j \in N_{i_s} \Rightarrow j \in D(\mathbf{x}'))$, then neither node i_s nor any of her neigbors is in the set $C(\mathbf{x}') \cap D(\mathbf{x^t})$, for all $t \geq s+1$. As the network is finite, this sequence reaches $\mathbf{x}'$ in a finite number k of steps. □

Next proposition provides a characterization of stochastically stable states under (5). We use a known result by Samuelson (1994) to obtain stochastic stability from single-mutation connected neighborhoods of absorbing sets.

Proposition **6.** *When (5) holds, a state $\mathbf{x}$ is stochastically stable in $(X, T, (T^{0,\epsilon})_{\epsilon \in (0,\bar{\epsilon})})$ if and only if $\mathbf{x} \in \mathcal{N}$.*

Proof. We first show that $\mathbf{x} \in \mathcal{N}$ implies $\mathbf{x}$ stochastically stable. Theorem 2 in Samuelson (1994) implies that if $\mathbf{x}'$ is stochastically stable, $\mathbf{x} \in \mathcal{A}$, $r^*(\mathbf{x}', \mathbf{x})$ is equal to the minimum resistance between recurrent states, then $\mathbf{x}$ is stochastically stable. Since at least one recurrent state must be stochastically stable, Proposition 2 implies that there must exist an absorbing state $\mathbf{x}'$ that is stochastically stable. For any $\mathbf{x} \in \mathcal{N}$, if $\mathbf{x} = \mathbf{x}'$ we are done. If

$\mathbf{x} \neq \mathbf{x}'$, then by Proposition 1 we can use Lemma 5 to say that there exists a finite sequence of absorbing states from $\mathbf{x}'$ to $\mathbf{x}$, where the resistance between subsequent states is always 1. Remark 1, together with Propositions 1 and 2, implies that 1 is the minimum resistance between recurrent states. A repeated application of Theorem 2 in Samuelson (1994) shows that each state in the sequence is stochastically stable, and in particular the final state $\mathbf{x}$.

It is trivial to show that $\mathbf{x}$ stochastically stable implies $\mathbf{x} \in \mathcal{N}$. By contradiction, suppose $\mathbf{x} \notin \mathcal{N}$. Then, by Propositions 1 and 2, $\mathbf{x} \notin \mathcal{R}$, and hence cannot be stochastically stable. $\square$

Proposition 6 tells us that, under the perturbation scheme considered in this subsection, stochastic stability turns out to be ineffective in selecting among equilibria.

4.2 Perturbations affect only the agents that are playing action 1

Now we assume that every agent playing action 1 is hit by an i.i.d. perturbation with probability $\epsilon \in (0, 1)$, while no agent playing action 0 is susceptible to perturbations. Again we indicate with $\mathbf{x}$ the state prior to perturbations and with $\mathbf{x}'$ the resulting state, and we formalize the individual perturbation process as follows:

$$x_i' = \begin{cases} x_i & \text{if } x_i = 0 \ , \\ \begin{cases} 0 & \text{with i.i.d. probability } \epsilon \ , \\ 1 & \text{with i.i.d. probability } 1 - \epsilon \end{cases} & \text{if } x_i = 1 \ . \end{cases} \tag{6}$$

We denote with $P^{1,\epsilon}$ the perturbations matrix resulting from (6). We check that $(X, T, (T^{1,\epsilon})_{\epsilon \in (0,\bar{\epsilon})})$, with $T^{1,\epsilon} = P^{1,\epsilon} T$, is indeed a regular perturbed Markov chain.

1. $(X, T^{1,\epsilon})$ is ergodic for all positive ϵ by the same argument of the corresponding point in the previous Subsection 4.1, once we replace Lemma 5 with Lemma 7.

2. $\lim_{\epsilon \to 0} T^{1,\epsilon} = T$, since $\lim_{\epsilon \to 0} P^{1,\epsilon}$ is equal to the identity matrix.

3. The resistance function is

$$r_1(\mathbf{x}, \mathbf{x}') = \begin{cases} ||S^C(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}')|| & \text{if } S^D(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}') = \emptyset \\ \infty & \text{otherwise} \end{cases} \tag{7}$$

   In fact, analogously to what happens for Subsection 4.1,

   (a) if $r_1(\mathbf{x}, \mathbf{x}') = \infty$, then $S^D(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}') \neq \emptyset$ and there is no way to go from $\mathbf{x}$ to $\mathbf{x}'$, since no satisfied agent can change in the unperturbed dynamics and no defector can be hit by a perturbation, hence $T^{\epsilon}_{\mathbf{x}\mathbf{x}'} = 0$ for every ϵ;

(b) if $r_1(\mathbf{x}, \mathbf{x}') < \infty$, then $T^{0,\epsilon}_{\mathbf{x}\mathbf{x}'}$ has the same order of $\epsilon^{r(\mathbf{x},\mathbf{x}')}$ when ϵ approaches zero; in fact, the agents in $U(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}')$ can change independently with probability β when T is applied, the agents in $S^C(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}')$ can change independently with probability ϵ when $P^{1,\epsilon}$ is applied (and only then), and no agent is left since $S^D(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}') = \emptyset$.

This remark plays the same role of Remark 1.

Remark **2.** *When (7) holds,* $r_1^*(\mathbf{x}, \mathbf{x}') \geq 1$ *for all* $\mathbf{x}, \mathbf{x}' \in \mathcal{N}$*.*

The following lemma shows that the resistance between any two absorbing states is equal to the number of contributors that must change to defection. This result is less trivial than it might appear: it shows that there is no possibility that by changing only some of the contributors to defectors, the remaining ones are induced to change by the unperturbed dynamics. The lemma could be proven directly from Lemma 3, but we find more intuitive the argument below.

Lemma **7.** *When (7) holds, for all* $\mathbf{x}, \mathbf{x}' \in \mathcal{A}$*,* $r_1^*(\mathbf{x}, \mathbf{x}') = ||C(\mathbf{x}) \cap D(\mathbf{x}')||$*.*

Proof. We first show that $r_1^*(\mathbf{x}, \mathbf{x}') \geq ||C(\mathbf{x}) \cap D(\mathbf{x}')||$. By contradiction, suppose $r^*(\mathbf{x}, \mathbf{x}') < ||C(\mathbf{x}) \cap D(\mathbf{x}')||$. Then, some $i \in C(\mathbf{x}) \cap D(\mathbf{x}')$ must switch from contribution to defection along a path from $\mathbf{x}$ to $\mathbf{x}'$ by best reply to the previous state. This requires that in the previous state there must exist some $j \in N_i$ that contributes. However, $j \in D(\mathbf{x})$ and j can never change to contribution as long as i is a contributor, neither by best reply nor by perturbation when (7) holds.

We now show that $r_1^*(\mathbf{x}, \mathbf{x}') \leq ||C(\mathbf{x}) \cap D(\mathbf{x}')||$. Define state $\tilde{\mathbf{x}}$ such that $\tilde{x}_i \equiv 0 \neq x_i = 1$ for all $i \in C(\mathbf{x}) \cap D(\mathbf{x}')$, and $\tilde{x}_i \equiv x_i$ otherwise. Note that $r(\mathbf{x}, \tilde{\mathbf{x}}) = ||C(\mathbf{x}) \cap D(\mathbf{x}')||$. Note also that $b_i(\tilde{\mathbf{x}}) = 1$ for all $i \in C(\mathbf{x}') \cap D(\tilde{\mathbf{x}})$. This means that $r(\tilde{\mathbf{x}}, \mathbf{x}') = 0$, and therefore $r^*(\mathbf{x}, \mathbf{x}') \leq r(\mathbf{x}, \tilde{\mathbf{x}}) + r(\tilde{\mathbf{x}}, \mathbf{x}') = ||C(\mathbf{x}) \cap D(\mathbf{x}')||$. □

We now use Lemma 7 to relate algebraically the resistance to move from $\mathbf{x}$ to $\mathbf{x}'$ to the resistance to come back from $\mathbf{x}'$ to $\mathbf{x}$.

Lemma **8.** *When (7) holds, for all* $\mathbf{x}, \mathbf{x}' \in \mathcal{A}$*,* $r^*(\mathbf{x}, \mathbf{x}') = r^*(\mathbf{x}', \mathbf{x}) + ||C(\mathbf{x})|| - ||C(\mathbf{x}')||$*.*

Proof. From Lemma 7 we know that $r^*(\mathbf{x}, \mathbf{x}') = ||C(\mathbf{x}) \cap D(\mathbf{x}')||$. Note that $||C(\mathbf{x}) \cap D(\mathbf{x}')|| = ||C(\mathbf{x})|| - ||C(\mathbf{x}) \cap C(\mathbf{x}')||$. Always from Lemma 7 we also know that $r^*(\mathbf{x}, \mathbf{x}') = ||C(\mathbf{x}') \cap D(\mathbf{x})|| = ||C(\mathbf{x}')|| - ||C(\mathbf{x}) \cap C(\mathbf{x}')||$, from which $||C(\mathbf{x}) \cap C(\mathbf{x}')|| = C(\mathbf{x}') - r^*(\mathbf{x}', \mathbf{x})$, which substituted in the former equality gives the desired result. □

We are now ready to provide a characterization of stochastically stable states under (7).

Proposition 9. *When (7) holds, a state $\mathbf{x}$ is stochastically stable in $(X, T, (T^{1,\epsilon})_{\epsilon\in(0,\bar{\epsilon})})$ if and only if $\mathbf{x} \in \arg\max_{\mathbf{x}'\in\mathcal{N}} ||C(\mathbf{x}')||$.*

Proof. We first prove that only a state in $\arg\max_{\mathbf{x}'\in\mathcal{N}} ||C(\mathbf{x}')||$ may be stochastically stable. Ad absurdum, suppose $||C(\mathbf{x})|| \notin \arg\max_{\mathbf{x}'\in\mathcal{N}} ||C(\mathbf{x}')||$ and $\mathbf{x}$ is stochastically stable. There must exist $\mathbf{x}'$ such that $||C(\mathbf{x}')|| > ||C(\mathbf{x})||$. Take an $\mathbf{x}$-tree $F_{\mathbf{x}}$. Consider the path in $F_{\mathbf{x}}$ going from $\mathbf{x}'$ to $\mathbf{x}$, that is the unique $\{(\mathbf{x^0}, \mathbf{x^1}), \dots, (\mathbf{x^{k-1}}, \mathbf{x^k})\}$ such that $\mathbf{x^0} = \mathbf{x}'$, $\mathbf{x^k} = \mathbf{x}$, and $(\mathbf{x^i}, \mathbf{x^{i+1}}) \in F_{\mathbf{x}}$ for all $i \in \{0, k-1\}$. We now modify $F_{\mathbf{x}}$ by reverting the path from $\mathbf{x}'$ to $\mathbf{x}$, so we define $F_{\mathbf{x}'} \equiv (F_{\mathbf{x}} \setminus \{(\mathbf{x^i}, \mathbf{x^{i+1}}) : \ i \in \{0, k-1\}\}) \cup \{(\mathbf{x^{i+1}}, \mathbf{x^i}) : \ i \in \{0, k-1\}\}$, which is indeed an $\mathbf{x}'$-tree. It is straightforward that $r_1^*(F_{\mathbf{x}'}) = r_1^*(F_{\mathbf{x}}) - \sum_{i=0}^{k-1} r_1^*(\mathbf{x^i}, \mathbf{x^{i+1}}) + \sum_{i=0}^{k-1} r_1^*(\mathbf{x^{i+1}}, \mathbf{x^i})$. Applying Lemma 8 we obtain $r_1^*(F_{\mathbf{x}'}) = r_1^*(F_{\mathbf{x}}) + \sum_{i=0}^{k-1}(||C(\mathbf{x^{i+1}})|| - ||C(\mathbf{x^i})||)$, that simplifies to $r_1^*(F_{\mathbf{x}'}) = r^*(F_{\mathbf{x}}) + ||C(\mathbf{x})|| - ||C(\mathbf{x}')||$. Since $||C(\mathbf{x}')|| > ||C(\mathbf{x})||$, then $r_1^*(F_{\mathbf{x}'}) < r_1^*(F_{\mathbf{x}})$. In terms of stochastic potentials, this implies that $\rho(\mathbf{x}') < \rho(\mathbf{x})$, against the hypothesis that $\mathbf{x}$ is stochastically stable.

We now prove that any state in $\arg\max_{\mathbf{x}'\in\mathcal{N}} ||C(\mathbf{x}')||$ is stochastically stable. Since at least one stochastically stable state must exist, from the above argument we conclude that there exists $\mathbf{x} \in \arg\max_{\mathbf{x}'\in\mathcal{N}} ||C(\mathbf{x}')||$ that is stochastically stable. Take any other $\mathbf{x}' \in \arg\max_{\mathbf{x}'\in\mathcal{N}} ||C(\mathbf{x}')||$. Following exactly the same reasoning as above we obtain that $\rho(\mathbf{x}') = \rho(\mathbf{x})$. Since $\rho(\mathbf{x})$ is a minimum, $\rho(\mathbf{x}')$ is a minimum too, and $\mathbf{x}'$ is hence stochastically stable. □

Previous proposition is the main point of this work: it provides a characterization of stochastically stable equilibria which is much more refined than the one obtained in Proposition 6. Next section analyzes the stability of this result, and generalizes to a wider class of possible sources of errors.

4.3 Perturbations affect all agents

We have obtained different results about stochastic stability in the extreme cases when perturbations hit exclusively agents playing either contribution or defection. We are now interested in understanding what happens when we allow both types of perturbation. In particular, we assume that every agent playing action 1 is hit by an i.i.d. perturbation with probability $\epsilon \in (0, 1)$, and every agent playing action 0 is hit by an i.i.d. perturbation

ϵ^m, where m is a positive real number.[12] Formally, with $\mathbf{x}$ denoting the state prior to perturbations and $\mathbf{x}'$ the resulting state:

$$x_i' = \begin{cases} \begin{cases} 1 & \text{with i.i.d. probability } \epsilon^m \ , \\ 0 & \text{with i.i.d. probability } 1-\epsilon^m \end{cases} & \text{if } x_i = 0 \ , \\ \\ \begin{cases} 0 & \text{with i.i.d. probability } \epsilon \ , \\ 1 & \text{with i.i.d. probability } 1-\epsilon \end{cases} & \text{if } x_i = 1 \ . \end{cases} \tag{8}$$

We denote with $P^{m,\epsilon}$ the perturbations matrix resulting from (8).

We check that $(X, T, (T^{m,\epsilon})_{\epsilon \in (0,\bar{\epsilon})})$, with $T^{m,\epsilon} = P^{m,\epsilon} T$, is indeed a regular perturbed Markov chain.

1. $(X, T^{m,\epsilon})$ is ergodic because if a state $\mathbf{x}'$ is accessible from a state $\mathbf{x}$ in $(X, T^{0,\epsilon})$ or in $(X, T^{1,\epsilon})$, then the same is true in $(X, T^{m,\epsilon})$.

2. $\lim_{\epsilon \to 0} T^{m,\epsilon} = T$, since $\lim_{\epsilon \to 0} P^{m,\epsilon}$ is equal to the identity matrix.

3. The resistance function is

$$r_m(\mathbf{x}, \mathbf{x}') = ||S^C(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}')|| + m||S^D(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}')|| \tag{9}$$

   In fact, the agents in $U(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}')$ can change independently with probability β when T is applied; the agents in $S^C(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}')$ can change independently with probability ϵ when $P^{m,\epsilon}$ is applied (and only then); and the agents in $S^D(\mathbf{x}) \cap M(\mathbf{x}, \mathbf{x}')$ can change independently with probability ϵ^m when $P^{m,\epsilon}$ is applied (and only then).

The usual kind of remark sets a lower bound to the resistance between any two absorbing states.

Remark **3.** *When (9) holds,* $r_m^*(\mathbf{x}, \mathbf{x}') \geq \min\{1, m\}$ *for all* $\mathbf{x}, \mathbf{x}' \in \mathcal{N}$.

We are ready for the last result: what happens when all agents are affected by perturbations.

Proposition **10.** *When (9) holds:*

[12] We might have used η^{m_0} for the probability that a perturbation hits a defector and η^{m_1} for the probability that a perturbation hits a contributor. Here we adopt, without loss of generality, the normalization that $\epsilon \equiv \eta^{m_0}$ and $m \equiv m_1/m_0$.

1. *if* $m \leq 1$*, then a state* $\mathbf{x}$ *is stochastically stable in* $(X, T, (T^{k,\epsilon})_{\epsilon \in (0,\bar{\epsilon})})$ *if and only if* $\mathbf{x} \in \mathcal{N}$*;*

2. *if* $m \geq \max_{i \in I}\{||N_i||\}$*, then a state* $\mathbf{x}$ *is stochastically stable in* $(X, T, (T^{m,\epsilon})_{\epsilon \in (0,\bar{\epsilon})})$ *if and only if* $\mathbf{x} \in \arg\max_{\mathbf{x}' \in \mathcal{N}} ||C(\mathbf{x}')||$.

Proof. Suppose $m \leq 1$. Following the proof of Lemma 5, we obtain that when (9) holds, for all $\mathbf{x}, \mathbf{x}' \in \mathcal{A}$, $\mathbf{x} \neq \mathbf{x}'$, there exists a sequence $\mathbf{x^0}, \ldots, \mathbf{x^s}, \ldots, \mathbf{x^k}$, with $\mathbf{x^s} \in \mathcal{A}$ for $0 \leq s \leq k$, $\mathbf{x^0} = \mathbf{x}$ and $\mathbf{x^k} = \mathbf{x}'$, such that $r^*(\mathbf{x^s}, \mathbf{x^{s+1}}) = m$ for $0 \leq s < k$. This result and Remark 3 allow us to follow the proof of Proposition 6 and use m instead of 1, obtaining the same result as in Proposition 6.

Suppose now that $m \geq \max_{i \in I}\{||N_i||\}$. We show that the resistances are the same as in Subsection 4.2, i.e., for all $\mathbf{x}, \mathbf{x}' \in \mathcal{A}$, $r^*_m(\mathbf{x}, \mathbf{x}') = ||C(\mathbf{x}) \cap D(\mathbf{x}')||$. Therefore Lemma 8 and Proposition 10 apply here too, and the result is obtained.

Lemma 3 and Lemma 4 imply that there is only one way to change the agents in $C(\mathbf{x}) \cap D(\mathbf{x}')$ from contribution to defection, other than letting each of those agents be hit by a perturbation. This way is to let some neighbors of all the agents in $C(\mathbf{x}) \cap D(\mathbf{x}')$ be hit by a perturbation changing their action from defection to contribution. This amounts to having at least $||C(\mathbf{x}) \cap D(\mathbf{x}')||/\max_{i \in I}\{||N_i|\}$ perturbations, each of which costs m. Since $m \geq \max_{i \in I}\{||N_i|\}$ by assumption, this way of reaching $\mathbf{x}'$ from $\mathbf{x}$ has at least a cost of $||C(\mathbf{x}) \cap D(\mathbf{x}')||$. This shows that $r^*_m(\mathbf{x}, \mathbf{x}') = ||C(\mathbf{x}) \cap D(\mathbf{x}')||$. □

In the next examples we give the stochastically stable sets of the game, for the whole range of values of m, for a very simple network. Even if the example is simple, it may give a hint on the complexity of situations that may arise in general, for values of m between 1 and $\max_{i \in I}\{||N_i||\}$.

Example 5. Consider the network from introductory Figure 1. This network has two Nash equilibria, one in which the two peripheral nodes Ann and Eve contribute (call it NE_2), another one in which the three central nodes Bob, Cindy and Dan do so (call it NE_3). Imagine that we want to find the stochastically stable equilibria deriving from (8), and hence (9), as $\epsilon \to 0$.

To pass from NE_2 to NE_3 we need that at least one of the central nodes starts contributing (with probability ϵ^m) or both Ann and Eve stop contributing (with probability ϵ^2). This event happens with a probability of order $\min\{m, 2\}$, this number is the resistance from NE_2 to NE_3. To pass from NE_3 to NE_3 we need that at least one between Ann and Eve starts contributing (with probability ϵ^m) or altogether Bob, Cindy and Eve stop contributing (with probability ϵ^3). This event happens with a probability of order $\min\{m, 3\}$, the resistance

from NE_3 to NE_2.
Summing up, the two equilibria are both stochastically stable if $m \leq 2$, as in this case the two resistances are equal. If instead $m > 2$, then only NE_3 is stochastically stable, as the resistance from NE_3 to NE_2 is $\min\{m, 3\}$, while the one from NE_2 to NE_3 is 2. □

5 Discussion

The best shot game is a very stark model, and clearly misses the details of any specific real–world situation. We think however that, as the model in Schelling (1969) has done for the issue of residential segregation, this model is able to describe the backbone structure of incentives in many problems of local contribution, as discussed in the introduction. The best shot game has multiple equilibria, so a refinement is needed to understand which outcomes are likely to emerge.

As we deal with discrete actions, the natural candidate for selection is stochastic stability: it selects the equilibria that are more likely to be observed in the long run, in the presence of small errors occurring with a vanishing probability. It is well known (Bergin and Lipman, 1996) that different equilibria can be selected depending on the assumptions on the relative likelihood of different types of errors, as indeed it occurs in our model (see the detailed discussion in Section 4). Blume (2003) focuses on finding sufficient conditions for the errors, in a perturbed dynamic on a discrete action space, such that stability gives always the same prediction. This dependence of stochastically stable states on the type of perturbations is often interpreted as a limitation of the predictive efficacy of stochastic stability, since essentially any equilibrium can be selected by means of stochastic stability with proper assumptions on the errors. We think instead that what enriches the analysis is exactly this dependence of the selected equilibria on the nature of errors. Our model is in principle very general, but if we try to apply it to a particular situation, it can adapt itself to the object of analysis and give specific predictions (as has been done, for a very different model, by Ben-Shoham et al., 2004). In particular, we derive interesting results for the case in which errors that stop contribution are much more frequent than errors that make contribution arise. We think that this is a property of many real–world situations in which the action of contributing involves much more individual effort than the action of non contributing, but also requires some external factor to be carried out (such as a car in order to give a lift). However, such an external factor may accidentally be damaged (e.g. a car engine may break down). The counter–intuitive result is that, exactly in those cases, the selected equilibrium will be the one with highest congestion, i.e. the number of contributors is the highest among equilibria.[13]

We leave for future research the analysis of more general models of network games, in which the effort of a player is substitute to the effort of her neighbors (for a formal definition see Galeotti et al., 2010). This is clearly the case of the best shot game, but possibly some of the results achieved in this paper could be generalized. Bramoullé et al. (2010) consider a particular class of games in which actions are continuous and neighbors' efforts are linear substitutes, and they analyze asymptotic stability as defined in Weibull (1995). It would be interesting, in more general games with a discrete strategy space, to compare and possibly generalize their findings with the results that we obtain for stochastic stability.

References

Arnott, R. and K. Small (1994). "The Economics of Traffic Congestion". *American Scientist 82 (5)*, 446–455.

Ben-Shoham, A., R. Serrano, and O. Volijc (2004). "The evolution of exchange". *Journal of Economic Theory 114*, 310–328.

Bergin, J. and B. L. Lipman (1996). "Evolution with State-Dependent Mutations". *Econometrica 64*, 943–956.

Blume, L. (2003). "How noise matter". *Games and Economic Behavior 44*, 251–271.

Bramoullé, Y., M. D'Amours, and R. Kranton (2010). "Strategic Interaction and Networks". *Mimeo*.

Bramoullé, Y. and R. Kranton (2007). "Public Goods in Networks". *Journal of Economic Theory 135*, 478–494.

Dall'Asta, L., P. Pin, and A. Ramezanpour (2009). "Statistical Mechanics of maximal independent sets". *Physical Review E 80 (6)*, 061136.

[13] In many examples the most congestioned equilibria are also the most inefficient ones. Remember however that we discuss a whole class of games, all characterized by the best reply function described in (1). Not all of them may have the same simple threshold structure proposed in Footnote 4. It would be easy to design examples in which (1) is the best reply function, but the payoff functions from which it derives have a structure that is non–satiated, and in some cases the most inefficient equilibria (in terms of aggregate payoff) may not be those in which the number of contributors is maximized.

Dall'Asta, L., P. Pin, and A. Ramezanpour (2010). "Optimal equilibria of the best shot game". *Accepted for publication on the Journal of Public Economic Theory*.

Ellison, G. (2000). "Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step Evolution". *The Review of Economic Studies 67*, 17–45.

Foster, D. and H. P. Young (1990). "Stochastic Evolutionary Game Dynamics". *Theoretical Population Biology 38*, 219–232.

Fudenberg, D. and D. K. Levine (1998). *The Theory of Learning in Games*. Cambridge, Massachussets: the MIT Press.

Galeotti, A., S. Goyal, M. Jackson, F. Vega-Redondo, and L. Yariv (2010). "Network Games". *The Review of Economic Studies 77 (1)*, 218–244.

Hirshleifer, J. (1983). "From Weakest-Link to Best-Shot: The Voluntary Provision of Public Goods". *Public Choice 41 (3)*, 371–386.

Kandori, M., G. J. Mailath, and R. Rob (1993). "Learning, Mutation and Long Run Equilibria in Games". *Econometrica 61*, 29–56.

McFadden, D. (1973). "Conditional Logit Analysis of Qualitative Choice Behavior". *in P. Zarembka (ed.), Frontiers in Economics*.

Samuelson, L. (1994). "Stochastic Stability in Games with Alternative Best Replies". *Journal of Economic Theory 64*, 35–65.

Samuelson, L. (1997). *Evolutionary Games and Equilibrium Selection*. Cambridge, Massachusetts: the MIT Press.

Schelling, T. (1969). "Models of Segregation". *The American Economic Review 59 (2)*, 488–493.

Tomita, E., A. Tanaka, and H. Takahashi (2006). "The Worst-Case Time Complexity for Generating all Maximal Cliques and Computational Experiments". *Theoretical Computer Science 363*, 28–42.

Weibull, W. (1995). *Evolutionary Game Theory*. Cambridge, Massachusetts: the MIT Press.

Young, H. P. (1993). "The Evolution of Conventions". *Econometrica 61*, 57–84.

Young, H. P. (1998). *Individual Strategy and Social Structure*. Princeton: Princeton University Press.