Stochastic Stability in Best Shot Network Games

Leonardo Boncinelli∗ Paolo Pin†

January 11, 2012

Abstract

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

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

Keywords: networks, best shot game, stochastic stability.

∗Corresponding author. Dipartimento di Scienze Economiche, Università degli Studi di Pisa, Via Cosimo Ridolfi 10, 56124 Pisa, Italy. Tel.: +39 050 221 6219, fax: +39 050 221 6384, email: l.boncinelli@ec.unipi.it.

†Dipartimento di Economia Politica e Statistica, Università degli Studi di Siena, Italy. Email: paolo.pin@unisi.it.
1 Introduction

In this paper we will look at a stylized game of contribution to a discrete local public good where the range of externalities is defined by a network. With 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 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: 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 friend a lift, but we presume here that non-linked people do not know each other and would not offer one another a lift. No one would take the car if a friend is doing so, but someone would be forced to take it if all her friends are not doing so. There is a less congested equilibrium in which Ann and Eve take the car (and the other three somehow get a lift), and a more polluting one in which Bob, Cindy and Dan take their car (offering Ann and Eve a lift, who will choose one of them).

![Figure 1: Five potential drivers in a network of relations.](image)

Imagine being in the less congested equilibrium. Now suppose that, even if they all agreed on how to manage the trip, in the morning Ann finds out that her car engine is broken and
cannot start it. She then calls her three friends, who are however not planning to take their cars and can not offer her a lift. As Ann does not know Eve, and Eve is 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 Ann and Eve’s cars break down, then Bob, Cindy and Dan would take their cars, and we shift to the equilibrium with more pollution. It is easy to see that if we start, instead, from the congested equilibrium, then we need three cars to break down before we can induce Ann and Eve to get their own. Therefore, it is intuitively established that the equilibrium with more pollution is more stable, since it is relatively easier to reach from the other equilibrium (two car engines would have to break down) than to leave for the other equilibrium (involving three car engines breaking down).

Recently, the literature on games played in networks has put attention on how network structure shapes outcomes (see Ballester et al. 2006, Bramoullé and Kranton 2007 – hereafter denoted by BK – and Bramoullé et al. 2011 – hereafter denoted BKD). This is particularly important for games of strategic substitutes, where networks easily allow a variety of Nash equilibria that are not present when interactions are global (see BK and BKD, who generalize the analysis and relate the emergence of a multiplicity of equilibria to the degree of substitutability between agents’ actions).

The strategic interaction that we analyze is rather close to that in BK. In particular, we take into consideration best shot network games: 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 Example 1 we described an equilibrium where each player can take one of two actions: take or not take her car. Then we have included a possible source of error: the car may break down and one is forced to change action from “take the car” to “not take the car”. Clearly we can also imagine a second type of error, e.g., if a player forgets that someone offered her a lift and takes her 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 this 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 Nash equilibria of the best shot game played in networks, under different assumptions on the perturbed Markov chain that allows agents to make errors.

---

1 Best shot games played in networks are considered in Galeotti et al. (2010), but the term best shot game goes back to the non-network application of Hirshleifer (1983).
What we find is that, if only errors of the type described in the example are possible – i.e., players can only make a mistake by not providing the public good even if that is their best response – 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 is admitted with a relatively high probability, then every Nash equilibrium is stochastically stable. The following example provides an intuition.

**Example 2.** Consider a hub and spoke network (i.e., a star) like the one depicted in Figure 2. If the individual at the center contributes, no one else does, since each individual on a spoke has the hub as a neighbor and this describes an equilibrium. Similarly, having every individual at the end of a spoke contributing and the hub not contributing is also a Nash equilibrium. Let us first consider the case in which errors affect only non-contributing agents that suddenly start contributing. From the first equilibrium, if an individual at the end of a spoke starts contributing, then by best response the hub will stop contributing, and after that all other individuals at the end of the spokes will find it convenient to switch to contribution. From the second equilibrium, if the hub starts contributing, then all other individuals will switch to defection. Therefore, a single error is sufficient to pass from one equilibrium to the other, and viceversa. Intuitively, the two equilibria are the same in terms of stability, and indeed they are both stochastically stable.

Let us now consider the other type of error, i.e., suppose that only contributors can be hit by errors that will induce them to switch to defection. From the first equilibrium, if the hub switches from contributing to not contributing, every individual on a spoke will switch to contributing by best response. From the second equilibrium, the hub will find it convenient to start contributing only when all the individuals at the end of a spoke will have stopped contributing. This means that a single error is sufficient for the system to pass from the first equilibrium to the second one, while 3 errors are required to move the system from the second equilibrium to the first one. This intuitively means that the second equilibrium – the one where the number of contributing agents is maximum – is more stable, and indeed it is the unique stochastically stable state.

As already anticipated, the best shot network game is very similar to BK’s local public goods game: they motivate their model with a story of neighborhood farmers, with 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 non-negative real numbers (how much to invest in the new risky fertilizer), and they
employ a notion of stability based on Nash tâtonnement.\textsuperscript{2} Theorem 2 in BK says that stable equilibria constitute a very specific subset of specialized equilibria.\textsuperscript{3} More precisely, a Nash equilibrium is stable to Nash tâtonnement if and only if it is specialized and every defector is connected to at least two contributors. Thus tâtonnement stability has a strong selecting power in BK’s setting, since many equilibria and specialized equilibria are not stable.

In the best shot network games that we study actions are discrete, errors, even if rare, are therefore more dramatic and the concept of stochastic stability naturally applies. Comparing our results with those in BK, we observe that Nash tâtonnement stability seems to share some features with stochastic stability under contributors’ mutations. They both select the same equilibria on many networks such as stars. They both select equilibria with relatively more contributors. But they are clearly not equivalent. In particular, a key problem with Nash tâtonnement is that stable equilibria often fail to exist. In contrast, and to be noted, stochastically stable equilibria always exist.

We think that our model, even if stark, offers 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

\textsuperscript{2}We note that since Nash tâtonnement relies only on best replies, then BK’s results apply to a more general class of games than public goods provision in networks. We comment on this in Section 2 when presenting best shot network games.

\textsuperscript{3}Specialized equilibria are those in which every agent either contributes an optimal amount (which is the same for all contributors) or free rides.
appealing example,\textsuperscript{4} while others are given in the introduction of Dall’Asta et al. (2011). 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 the next section we will formalize the best shot network game. Section 3 describes the general best response dynamics that we apply to the game. In Section 4 we will introduce the possibility of errors thus obtaining a perturbed dynamics, and the main theoretical analysis of the effects of different perturbation schemes. Section 5 provides some examples with the aim of providing an intuitive understanding of how stochastic stability refines equilibria in best shot network games; another example allows us to comment on welfare, and finally we will discuss some attempts to generalize our results. A brief closing discussion is given in Section 6. Finally, to stress the robustness of our findings, in the Appendix we will provide a simple alternative model, restricted to a specific payoff function and to a logit-response updating of strategies, in which only the equilibria with the largest contribution are selected by stochastic stability.

2 Best Shot Network 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 a 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, i.e., $N_i = \{ j \in N : G_{ij} = 1 \}$. The number of neighbors of a node is called its degree. A path between two nodes $i$ and $j$ is an ordered set of nodes $(i, h_1, h_2 \ldots h_\ell, j)$ such that $G_{ih_1} = 1$, $G_{h_1h_2} = 1$, $\ldots$, $G_{h_\ell j} = 1$.

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.\textsuperscript{5} We

\textsuperscript{4}Economic modelling of traffic has 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.

\textsuperscript{5}With our choice of terminology we are suggesting an interpretation of the strategic interaction under study in terms of contribution to a local public good and, in such respect, the term free rider might be used instead of the term defector. However, we remark that different interpretations might be advanced, since the best reply functions that we adopt in the following arise in a class of games that is larger than public goods.
will consider only pure strategies. A state of the system is represented by a vector \( x \) which specifies each agent’s action, \( x = (x_1, \ldots, x_i, \ldots, x_n) \). 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.\(^6\) 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:

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

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

The above game is called best shot network game.\(^7\) A state \( x \) is a pure strategy Nash equilibrium of the best shot network game if and only if \( S(x) = I \) and consequently \( U(x) = \emptyset \).

Theorem 1 in BK shows that Nash equilibria are such that contributors constitute a maximal independent set of the graph. We will call all 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

---

\(^6\)As discussed in Footnote 5, there are many payoff functions that yield the best reply defined by (1) (see BKD for examples). One instance is the following. Imagine that the cost for contributing is \( c \), and the value of either being a contributor or having a contributing agent in her own neighborhood is \( \alpha > c > 0 \). The benefit of contribution can be seen as having access to a local public good. In this interpretation, agents can be, but not necessarily are, satiated by one unit of contribution (see also Example 9).

\(^7\)Here we are using the term “game” with a slight abuse of language. Strictly speaking, a game is defined as a collection of players, strategies, and payoffs. Many different games with radically different welfare properties share the same best reply functions. See also the discussion in Footnote 5 and Footnote 6.
to enumerate all the elements of $\mathcal{N}$,\textsuperscript{8} 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. (2011). 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 3.** Figure 3 shows two of the three possible Nash equilibria of the same 5-node network, where the five characters of our introductory example now have a different network of friendships.

![Figure 3: Two Nash equilibria for a 5-node network. Dark blue stands for contribution, while light blue stands for defection.](image)

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

\textsuperscript{8}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 4: Two Nash equilibria for the same regular random network of 20 nodes and degree 4. Dark blue stands for contribution, while light blue stands for defection. This picture was 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 $x$ 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 $x$ the current state and with $x'$ the next state, we can then formalize as follows:

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

(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\(^9\) resulting from the individual update process in (2). We note that $T$ depends on $\beta$.

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 of changing action:\(^{10}\)

\(^9\) $T_{xx'}$ denotes the probability of passing from state $x$ to state $x'$.

\(^{10}\) It is the property stated in (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 any Markov chain whose transition matrix satisfies such property is aperiodic because $M(x, x) = \emptyset$ for all $x$, which by (3) implies that $T_{xx} > 0$ for all $x$. 
\[ T_{xx'} > 0 \text{ if and only if } M(x, x') \subseteq U(x). \] (3)

We introduce some terminology from Markov chain theory following Young (1998). A state \( x' \) is called accessible from a state \( x \) if a sequence of states exists, with \( x \) as first state and \( x' \) as last state, such that the system can move with positive probability from each state to the next state. 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 \( x \in \mathcal{E} \), then \( x \) is called recurrent. Let \( \mathcal{R} \) denote the set of all recurrent states of \((X, T)\). If \( \{x\} \) is an ergodic set, then \( x \) is called absorbing. Equivalently, \( x \) is absorbing when \( T_{xx} = 1 \). 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 of Nash equilibria – denoted by \( \mathcal{N} \) – coincides with set of 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 \( x \in \mathcal{N} \). Since by (3) we have that \( T_{xx'} > 0 \) with \( x' \neq x \) only if \( U(x) \neq \emptyset \), then \( T_{xx} = 0 \) for any \( x' \neq x \), hence \( T_{xx} = 1 \) and \( x \) is absorbing.

Now we show that \( \mathcal{A} \subseteq \mathcal{N} \). By contradiction, suppose \( x \notin \mathcal{N} \). Then \( U(x) \neq \emptyset \). Consider a state \( x' \) where \( x'_i = x_i \) if \( i \in S(x) \), and \( x'_i \neq x_i \) otherwise. We have that \( x' \neq x \) and, by (3), that \( T_{xx'} > 0 \), hence \( T_{xx} < 1 \) and \( x \) 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 \( x \) which is not in \( \mathcal{A} \) is also not in \( \mathcal{R} \). Suppose that \( x \notin \mathcal{A} \). We identify, by means of a recursive algorithm, a state \( \hat{x} \) such that \( \hat{x} \) is accessible from \( x \), but \( x \) is not accessible from \( \hat{x} \). This implies that \( x \notin \mathcal{R} \).

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

**Step 1.** We take \( i \in U^C(x') \) and we define state \( x'' \) such that \( x''_i \equiv 0 \neq x'_i = 1 \) and \( x''_j \equiv x'_j \) for all \( j \neq i \). Note that \( ||U^C(x'')|| \leq ||U^C(x')|| \). This is because of two reasons: first of all, \( i \in U^C(x') \) and \( i \notin U^C(x'') \); the second is that \( U^C(x'') \subseteq U^C(x') \), otherwise two neighbors contribute in \( x'' \) and do not contribute in \( U^C(x') \), but that is not possible because...
$C(x'') \subset C(x')$. Moreover, by (3) we have that $T_{x'x''} > 0$. We redefine $x' \equiv x''$. Then, if $U^C(x') = \emptyset$ we pass to Step 2, otherwise we repeat Step 1.

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

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

An immediate corollary of Proposition 1 and Proposition 2 is that $R = A = N$.

![Network Diagram](c)

Figure 5: A non-Nash (non-absorbing) state for the same network of Figure 3. Here Bob and Cindy are contributing, while Ann, Dan and Eve are not.

**Example 5.** Consider the network from Figure 3: 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 5: the satisfied nodes here are only Ann, Dan and Eve, who are all defecting. 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 unsatisfied and hence inclined to switch to contribution. To reach (b) from (c), both unsatisfied contributors Bob and Cindy should simultaneously turn to defection, so that all five nodes would become unsatisfied and, if we Ann, Dan and Eve switch to contribution, then state (b) is reached.

The following Lemma 3 and Lemma 4 adapt the results in Lemma 2 of Dall’Asta et al. (2011) 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 is developed in Section 4.

Lemma 3 states that if we start from a Nash equilibrium, we force agent \( i \) to switch from contribution to defection, and we let 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 \( x \in N \) and \( x_i = 1 \). Define \( x' \) such that \( x'_i = 0 \) and \( x'_j = x_j \) for all \( j \neq i \). Then, for every state \( x'' \) that is accessible from \( x' \) through \( T \) we have that if \( x''_j \neq x'_j \) then either \( j = i \) or \( j \in N_i \).

**Proof.** We first prove two intermediate results.

**Result 1.** If \( j \) is a satisfied contributor at \( x' \), then she will never change action under dynamics \( T \), i.e., \( x''_j = x'_j \) at every \( x'' \) that is accessible from \( x' \). In fact, all \( j \)'s neighbors are not contributing, and no one of them will start doing so under dynamics \( T \) since they have \( j \) as a contributing neighbor. So \( j \) will never become unsatisfied.

**Result 2.** If \( j \) is a defector and at least one of her neighbors is a satisfied contributor at \( x' \), then \( j \) will never change action under dynamics \( T \), i.e., \( x''_j = x'_j \) at every \( x'' \) that is accessible from \( x' \). Agent \( j \) is satisfied since she has at least one contributing neighbor. Since one of her contributing neighbors is satisfied, by Result 1 such a neighbor will never change action. So \( j \) will never become unsatisfied.

We now show that at \( x' \) any agent \( j \neq i, j \notin N_i \) is either a satisfied contributor or a defector with a satisfied contributing neighbor. Then Result 1 and Result 2 apply and the proof is completed. Suppose \( j \neq i, j \notin N_i \), and \( x'_j = 1 \). Agent \( j \) is satisfied at \( x \) – since \( x \) is a Nash equilibrium – and so \( j \) must be satisfied at \( x' \) since only agent \( i \) has changed action. Suppose \( j \neq i, j \notin N_i \), and \( x'_j = 0 \). Since \( x \) is a Nash equilibrium and \( x' \) differs only for \( i \)'s action, then some of \( j \)'s neighbors must be contributing at \( x' \). Since all agents belonging to \( N_i \) are not contributing at \( x \), and at \( x' \) as well, then such \( j \)'s contributing neighbor can not belong to \( N_i \cup \{ i \} \), and so she must be satisfied at \( x' \).
Lemma 4 states that if we start from a Nash equilibrium, we force agent $i$ to switch from contribution to defection, and we let 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 $x \in \mathcal{N}$ and $x_i = 0$. Define $x'$ such that $x'_i = 1$ and $x'_j = x_j$ for all $j \neq i$. Then, for every state $x''$ that is accessible from $x'$ through $T$ we have that if $x''_j \neq x'_j$ then either $j = i$ or $j \in N_i$ or $j \in N_k$ for some $k \in N_i$.

**Proof.** The proof proceeds similarly to the proof of Lemma 3. We show that at $x'$ any agent $j \neq i$, $j \notin N_i$, $j \notin N_k$ for any $k \in N_i$, is either a satisfied contributor or a defector with a satisfied contributing neighbor. Then Result 1 and Result 2 apply and the proof is completed. Suppose $j \neq i$, $j \notin N_i$, $j \notin N_k$ for any $k \in N_i$, and $x'_j = 1$. Agent $j$ is satisfied at $x$ – since $x$ is a Nash equilibrium – and so $j$ must be satisfied at $x'$ since only agent $i$ has changed action. Suppose $j \neq i$, $j \notin N_i$, $j \notin N_k$ for any $k \in N_i$, and $x'_j = 0$. Since $x$ is a Nash equilibrium and $x'$ differs only for $i$’s action, then some of $j$’s neighbors must be contributing at $x'$. Moreover, such a contributing neighbor can not belong to $N_i \cup \{i\}$, and so she must be satisfied at $x'$.

## 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 state to state. 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 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 triple $(X, T, (T^\epsilon)_{\epsilon \in (0, \bar{\epsilon})})$ where $(X, T)$ is the unperturbed Markov chain and:

1. $(X, T^\epsilon)$ is an ergodic Markov chain, for all $\epsilon \in (0, \bar{\epsilon})$;
2. \( \lim_{\epsilon \to 0} T^\epsilon = T; \)

3. there exists a \textit{resistance} function \( r : X \times X \to \mathbb{R}^+ \cup \{\infty\} \) such that for all pairs of states \( x, x' \in X \),

\[
\begin{cases}
\lim_{\epsilon \to 0} \frac{T^{\epsilon}_{xx'}}{r^{\epsilon}(x,x')} \text{ exists and is strictly positive} & \text{if } r(x, x') < \infty; \\
T^{\epsilon}_{xx'} = 0 \text{ for sufficiently small } \epsilon & \text{if } r(x, x') = \infty.
\end{cases}
\]

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

For any \( \epsilon > 0 \), the Markov chain \( (X, T^\epsilon) \) has a unique invariant distribution \( \mu^\epsilon \), which can be thought of as giving the probability of observing each of the states after a very long time. The limit distribution \( \mu^* \) is defined by \( \mu^* = \lim_{\epsilon \to 0} \mu^\epsilon \). A state \( x \) is said to be \textit{stochastically stable} if it receives positive probability in the limit distribution, i.e., \( \mu^*(x) > 0 \). Intuitively, stochastically stable states are those and only those states that the system can occupy after evolution has been going on for a very long time in the presence of a tiny amount of perturbations. In the following we present a characterization of stochastic stability mainly due to Foster and Young (1990), Young (1993) and Young (1998).

Even if \( T \) and \( r \) are defined on all the possible states of \( X \), we can limit our analysis to absorbing states only, since we know from Proposition 2 that no other recurrent state exists. This technical procedure is illustrated in Young (1998) and simplifies the complexity of the notation, without loss of generality. Given \( x, x' \in \mathcal{A} \), we define \( r^*(x, x') \) as the minimum sum of the resistances between absorbing states over any path starting in \( x \) and ending in \( x' \).

Given \( x \in \mathcal{A} \), an \( x \)-tree on \( \mathcal{A} \) is a subset of \( \mathcal{A} \times \mathcal{A} \) that constitutes a tree rooted at \( x \).\(^{11}\) We denote such \( x \)-tree with \( F_x \) and the set of all \( x \)-trees with \( \mathcal{F}_x \). The \textit{resistance of an \( x \)-tree}, denoted with \( r^*(F_x) \), is defined to be the sum of the resistances of its edges, that is:

\[
r^*(F_x) = \sum_{(x,x') \in F_x} r^*(x, x').
\]

\(^{11}\)By \textit{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.
Finally, the *stochastic potential* of $x$ is defined to be

$$
\rho(x) \equiv \min \{ r^*(F_x) : F_x \in \mathcal{F}_x \}.
$$

A state $x$ is proven to be stochastically stable (Foster and Young, 1990) if and only if

$$
\rho(x) = \min \{ \rho(x) : x \in \mathcal{A} \}.
$$

Intuitively, stochastic stability selects those states that are easiest to reach from other states, with “easiest” interpreted as requiring the fewest mutations (as measured by the stochastic potential).

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. In the Appendix we consider a specific alternative model based on logit-response.

### 4.1 Perturbations only affect 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 as we defined $T$ from (2). We indicate with $x$ the state prior to perturbations and with $x'$ the resulting state:

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

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.$^{12}$

1. $(X, T^{0,\epsilon})$ is ergodic for all positive $\epsilon$: 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) $r^*(x, x') < \infty$ for all $x, x' \in \mathcal{A}$ by the following Lemma 5.

$^{12}$We are following Samuelson (1997) when we derive the perturbed transition matrix $T^{0,\epsilon}$ by post-multiplying the unperturbed transition matrix $T$ with the perturbation matrix $P^{0,\epsilon}$. If we exchange 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).
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\(^{13}\) is

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

In fact,

(a) if \( r_0(x, x') = \infty \), then \( S^C(x) \cap M(x, x') \neq \emptyset \) and there is no way to go from \( x \) to \( x' \), since no satisfied agent can change in the unperturbed dynamics and no contributor can be hit by a perturbation, hence \( T_{xx'}^\epsilon = 0 \) for every \( \epsilon \);

(b) if \( r_0(x, x') < \infty \), then \( T_{xx'}^{0,\epsilon} \) has the same order of \( \epsilon r^{(x,x')} \), when \( \epsilon \) approaches zero; in fact, the agents in \( U(x) \cap M(x, x') \) can change independently with probability \( \beta \) when \( T \) is applied, the agents in \( S^D(x) \cap M(x, x') \) can change independently with probability \( \epsilon \) when \( P^{0,\epsilon} \) is applied (and only then), and no agent is left since \( S^C(x) \cap M(x, x') = \emptyset \).

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 a remark in Proposition 6.

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

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

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

We define state \( \tilde{x} \) 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 \( r_0(x^s, \tilde{x}) = 1 \). We define state \( \tilde{x}' \) 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 \( x^{s+1} \) such that \( x^{s+1}_k = b_k(\tilde{x}') \) for all \( k \in N_j, j \in N_{i_s} \), and

\(^{13}\)The necessary assumption for our results is having a regular perturbed Markov chain whose resistance function is as in (5). By deriving it from an individual perturbation process – which is defined in (4) – and from an individual update process – which is defined in (2) – we show that at least one significant case satisfying this property exists.
\(x_{\ell}^{s+1} \equiv \tilde{x}'_{\ell}\) for any other node \(\ell\). By Lemma 4 and Proposition 1, \(x^{s+1} \in \mathcal{A}\). We note that \(x^{s+1}\) is obtained from \(\tilde{x}\) applying only the unperturbed dynamics \(T\), hence the resistance \(r_0^*(x^s, x^{s+1}) = 1\).

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

The 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 \(x\) is stochastically stable in \((X, T, (T^{0, \epsilon})_{\epsilon \in (0, \bar{\epsilon}}))\) if and only if \(x \in \mathcal{N}\).

**Proof.** We first show that \(x \in \mathcal{N}\) implies \(x\) stochastically stable. Theorem 2 in Samuelson (1994) implies that if \(x'\) is stochastically stable, \(x \in \mathcal{A}\), \(r^*(x', x)\) is equal to the minimum resistance between recurrent states, then \(x\) is stochastically stable. Since at least one recurrent state must be stochastically stable, Proposition 2 implies that there must exist an absorbing state \(x'\) that is stochastically stable. For any \(x \in \mathcal{N}\), if \(x = x'\) we are done. If \(x \neq x'\), then by Proposition 1 we can use Lemma 5 to say that there exists a finite sequence of absorbing states from \(x'\) to \(x\), 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 \(x\).

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

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 only affect the agents that are playing action 1**

Now we are going to 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 \(x\) the state prior to perturbations and with \(x'\) the resulting state, and we formalize the individual perturbation process as follows:
\[ x'_i = \begin{cases} 
 x_i & \text{if } x_i = 0, \\
 0 & \text{with i.i.d. probability } \epsilon, \\
 1 & \text{with i.i.d. probability } 1 - \epsilon, 
\end{cases} \quad (6) \]

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 \( \epsilon \) 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
\[ r_1(x, x') = \begin{cases} ||S^C(x) \cap M(x, x')|| & \text{if } S^D(x) \cap M(x, x') = \emptyset, \\
\infty & \text{otherwise.} \end{cases} \quad (7) \]

In fact, analogously to what happens for Subsection 4.1,

(a) if \( r_1(x, x') = \infty \), then \( S^D(x) \cap M(x, x') \neq \emptyset \) and there is no way to go from \( x \) to \( x' \), since no satisfied agent can change in the unperturbed dynamics and no defector can be hit by a perturbation, hence \( T^{1,\epsilon}_{xx'} = 0 \) for every \( \epsilon \);

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

This remark plays the same role as Remark 1.

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

The following lemma shows that the resistance to go from an absorbing state \( x \) to an absorbing state \( x' \) is equal to the number of contributors at \( x \) that are defectors at \( x' \). While it is clear that those agents must switch action in the passage from \( x \) to \( x' \), this result shows that there is no possibility that even some of them can make such a change in the unperturbed dynamics: each of them must be hit by a perturbation in order to switch to defection. The lemma could be proven directly from Lemma 3, but we find that the argument below is more instructive.
Lemma 7. When (7) holds, for all \( x, x' \in \mathcal{A} \), \( r_1^*(x, x') = ||C(x) \cap D(x')|| \).

Proof. We first show that \( r_1^*(x, x') \geq ||C(x) \cap D(x')|| \). By contradiction, suppose \( r_1^*(x, x') < ||C(x) \cap D(x')|| \). Then, some \( i \in C(x) \cap D(x') \) must switch from contribution to defection along a path from \( x \) to \( x' \) by best reply to the previous state. This requires that some \( j \in N_i \) is contributing at the previous state. However, \( j \in D(x) \) 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 \( r_1^*(x, x') \leq ||C(x) \cap D(x')|| \). Define state \( \tilde{x} \) such that \( \tilde{x}_i \equiv 0 \neq x_i = 1 \) for all \( i \in C(x) \cap D(x') \), and \( \tilde{x}_i \equiv x_i \) otherwise. Note that \( r(x, \tilde{x}) = ||C(x) \cap D(x')|| \). Note also that \( b_i(\tilde{x}) = 1 \) for all \( i \in C(x') \cap D(\tilde{x}) \). This means that \( r(\tilde{x}, x') = 0 \), and therefore \( r_1^*(x, x') \leq r(x, \tilde{x}) + r(\tilde{x}, x') = ||C(x) \cap D(x')|| \). \qed

The above lemma plays a crucial role in the derivation of the results in this section. We can think of it as a consequence of Lemma 3. In fact, Lemma 3 states that when a contributor is hit by a perturbation and switches to defection, then the adjustment due to the unperturbed dynamics is restricted to immediate neighbors, so that any other contributor (who cannot be an immediate neighbor) remains unaffected. When instead a contributor is hit by a perturbation and switches to defection, then we cannot exclude that neighbors of neighbors are affected by the adjustment dynamics (see Lemma 4). This should intuitively motivate why Lemma 7 holds when the individual perturbation process is as in (6), but not when it is as in (4).

We now use Lemma 7 to relate algebraically the resistance to move from \( x \) to \( x' \) and the resistance to come back from \( x' \) to \( x \).

Lemma 8. When (7) holds, for all \( x, x' \in \mathcal{A} \), \( r_1^*(x, x') = r_1^*(x', x) + ||C(x)|| - ||C(x')|| \).

Proof. From Lemma 7 we know that \( r_1^*(x, x') = ||C(x) \cap D(x')|| \). Note that \( ||C(x) \cap D(x')|| = ||C(x)|| - ||C(x) \cap C(x')|| \). Always from Lemma 7 we also know that \( r_1^*(x, x') = ||C(x') \cap D(x)|| = ||C(x')|| - ||C(x) \cap (x')|| \), from which \( ||C(x) \cap C(x')|| = C(x') - r_1^*(x, x') \), which substituted in the former equality gives the desired result. \qed

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

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

Proof. We first prove that only a state belonging to \( \arg \max_{x' \in \mathcal{N}} ||C(x')|| \) may be stochastically stable. Ad absurdum, suppose \( x \notin \arg \max_{x' \in \mathcal{N}} ||C(x')|| \) and \( x \) is stochastically stable.
A state \( x' \) such that \( ||C(x')|| > ||C(x)|| \) must exist. Take an \( x \)-tree \( F_x \). Consider the path in \( F_x \) going from \( x' \) to \( x \), that is the unique \( \{(x^0, x^1), \ldots, (x^{k-1}, x^k)\} \) such that \( x^0 = x' \), \( x^k = x \), and \( (x^i, x^{i+1}) \in F_x \) for all \( i \in \{0, k - 1\} \). We now modify \( F_x \) by reverting the path from \( x' \) to \( x \), so we define \( F_{x'} = (F_x \setminus \{(x^i, x^{i+1}) : i \in \{0, k - 1\}\}) \cup \{(x^{i+1}, x^i) : i \in \{0, k - 1\}\} \), which is indeed an \( x' \)-tree. It is straightforward that \( r^*_1(F_{x'}) = r^*_1(F_x) - \sum_{i=0}^{k-1} r^*_1(x^i, x^{i+1}) + \sum_{i=0}^{k-1} r^*_1(x^{i+1}, x^i) \). Applying Lemma 8 we obtain \( r^*_1(F_{x'}) = r^*_1(F_x) + \sum_{i=0}^{k-1} (||C(x^{i+1})|| - ||C(x^i)||) \), that simplifies to \( r^*_1(F_{x'}) = r^*_1(F_x) + ||C(x')|| - ||C(x)|| \). Since \( ||C(x')|| > ||C(x)|| \), then \( r^*_1(F_{x'}) < r^*_1(F_x) \). In terms of stochastic potentials, this implies that \( \rho(x') < \rho(x) \), against the hypothesis that \( x \) is stochastically stable.

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

The 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. The next section analyzes the stability of this result, and generalizes it to a wider class of possible sources of error.

### 4.3 Perturbations affect all agents

We obtained different results about stochastic stability in the extreme cases when perturbations hit 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 with probability \( \epsilon^m \), where \( m \) is a positive real number.\(^\text{14}\) Formally, with \( x \) denoting the state prior to perturbations and \( x' \) the resulting state:

\(^{14}\)We might have used \( \eta^{m_0} \) for the probability that a perturbation hits a defector and \( \eta^{m_1} \) 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 \( m \equiv m_1/m_0 \).
\[ x'_i = \begin{cases} 
1 \text{ with i.i.d. probability } \epsilon^m, & \text{if } x_i = 0, \\
0 \text{ with i.i.d. probability } 1 - \epsilon^m, & \text{if } x_i = 0, \\
0 \text{ with i.i.d. probability } \epsilon, & \text{if } x_i = 1, \\
1 \text{ with i.i.d. probability } 1 - \epsilon, & \text{if } x_i = 1.
\] (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 \( x' \) is accessible from a state \( x \) 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

\[ r_m(x, x') = ||S^C(x) \cap M(x, x')|| + m||S^D(x) \cap M(x, x')||. \] (9)

In fact, agents in \( U(x) \cap M(x, x') \) can change independently with probability \( \beta \) when \( T \) is applied; agents in \( S^C(x) \cap M(x, x') \) can change independently with probability \( \epsilon \) when \( P^{m,\epsilon} \) is applied (and only then); and agents in \( S^D(x) \cap M(x, x') \) can change independently with probability \( \epsilon^m \) 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, \( r^*_m(x, x') \geq \min\{1, m\} \) for all \( x, x' \in \mathcal{N} \).

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

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

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

2. if \( m \geq \max_{i \in I} \{||N_i||\} \), then a state \( x \) is stochastically stable in \((X, T, (T^{m,\epsilon})_{\epsilon \in (0,\bar{\epsilon})})\) if and only if \( x \in \arg\max_{x' \in \mathcal{N}} ||C(x')|| \).
Proof. Suppose \( m \leq 1 \). Following the proof of Lemma 5, we obtain that when (9) holds, for all \( x, x' \in A \), \( x \neq x' \), there exists a sequence \( x^0, \ldots, x^s, \ldots, x^k \), with \( x^s \in A \) for \( 0 \leq s \leq k \), \( x^0 = x \) and \( x^k = x' \), such that \( r^*_m(x^s, x^{s+1}) = m \) 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 \( x, x' \in A \), \( r^*_m(x, x') = ||C(x) \cap D(x')|| \). Therefore Lemma 8 and Proposition 9 apply here too, and the result is obtained.

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

Proposition 10 tells us what happens when \( m \) is sufficiently high or sufficiently low, but is silent for an intermediate range of values, i.e., when \( m \in (1, \max_{i \in I} \{||N_i||\}) \). In the next section we provide some examples that, besides giving an intuitive understanding of how our model works, also identify the set of stochastically stable states for the whole range of values, thus showing that several cases are possible.

5 Examples and Discussion

In the next examples we apply stochastic stability to very simple instances of the best shot network game. In particular, we are interested in showing how equilibria selected by stochastic stability change when the relative likelihood of contributors’ mistakes and defectors’ mistakes – summarized by parameter \( m \) – changes. Therefore, we work with an individual perturbation process as in (8). Even if these examples are simple, they may convey the main intuition behind the asymmetry of our results, and provide 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 6. This example generalizes Example 2 with respect to the error structure and the number of nodes. Consider a star with \( n + 1 \) nodes. Clearly, there are two Nash equilibria: the center equilibrium, where only the central node contributes, and the peripheral equilibrium, where all and only the peripheral nodes contribute. To pass from the center equilibrium to
the peripheral equilibrium we either need the central node to stop contributing (this happens with probability $\epsilon$) or at least one peripheral node to start contributing (this happens with probability $\epsilon^m$). So the resistance is $\min\{1, m\}$. To pass from the peripheral equilibrium to the center equilibrium we either need all peripheral nodes to stop contributing (this happens with probability $\epsilon^m$) or the central node to start contributing (this happens with probability $\epsilon^m$). So the resistance is $\min\{n, m\}$.

When $m \leq 1$ both equilibria are stochastically stable, when instead $m > 1$ only the peripheral equilibrium (where the number of contributors is highest) is stochastically stable.

In the above example, when $m$ is in between 1 and $m$ (the range of values for which Proposition 10 does not make predictions) stochastic stability selects the equilibrium with the largest contribution. The following example provides a different case.

**Example 7.** 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 $NE_3$). To pass from $NE_2$ to $NE_3$ we either need at least one of the central nodes to start contributing (with probability $\epsilon^m$), or both Ann and Eve to stop contributing (with probability $\epsilon^2$). So the resistance is $\min\{2, m\}$. To pass from $NE_3$ to $NE_3$ we either need at least one between Ann and Eve to start contributing (with probability $\epsilon^m$), or altogether Bob, Cindy and Eve to stop contributing (with probability $\epsilon^3$). So the resistance is $\min\{3, m\}$. 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 $\min\{3, m\}$, while the one from $NE_2$ to $NE_3$ is 2.

In Example 7, when $m$ is in between 1 and 3 (that again is the range of values for which Proposition 10 does not apply) stochastic stability does not always select the same outcome: when $1 < m \leq 2$ both equilibria are stochastically stable, and when $2 < m < 3$ only the equilibrium with largest contribution is stochastically stable.

In the previous two examples the set of stochastically stable states depends on $m$ in a very simple and dramatic way: there are two ranges of values of $m$, the lowest one in which all equilibria are stochastically stable, and the highest one in which only the equilibria with the largest number of contributors are stochastically stable. This is not always the case, since there may exist intermediate occurrences, as the next example illustrates.

**Example 8.** Consider the network at the top of Figure 6. All its four Nash equilibria are depicted below, and at the bottom there is a table where the $ij$ entry contains the resistance
to pass from the state in row $i$ to the state in column $j$. To compute the stochastic potential of every Nash equilibrium is rather easy but boring: when $m \leq 1$ the stochastic potential (i.e., the tree with minimum resistance) is $3m$ for all Nash equilibria, that are hence all stochastically stable; when $1 < m \leq 2$ the stochastic potential of $NE5$ and $NE4$ is $2 + m$ while the stochastic potential of $NE3_1$ and $NE3_2$ is $1 + 2m$, and hence both and only $NE5$ and $NE4$ are stochastically stable; finally, when $m > 2$ the stochastic potential of $NE5$ is $4$, the stochastic potential of $NE4$ is $2 + m$, the stochastic potential of $NE3_1$ and $NE3_2$ is $1 + 2m$, and hence $NE5$ is the only stochastically stable state. A more intuitive explanation of the result is the following: whatever $m \geq 1$ the passages from $NE3_1$ to $NE5$ and from $NE3_2$ to $NE4$ have both resistance 1; when $m \leq 2$ the passages from $NE4$ to $NE5$ and from $NE5$ to $NE4$ have both resistance $m$ (as in Example 7), so that $NE5$ and $NE4$ are equally difficult to reach from other states; when $m > 2$ the passage from $NE4$ to $NE5$ has resistance 2 while the passage from $NE5$ to $NE4$ has resistance $\min\{3, m\} > 2$ (as in Example 7), so that $NE5$ is easiest to reach from other states.

The previous example shows that a range of values for $m$ (in the specific example when $1 < m \leq 2$) where neither all Nash equilibria nor all and only the Nash equilibria with largest contribution are stochastically stable may exist.

We now turn our attention to welfare. Example 9 allows us to make some comments.

Example 9. Consider the network at the top of Figure 7. It has four Nash equilibria, that are depicted below. Suppose that the payoff accruing to individual $i$ in state $x$ is a function of the total number of contributing agents in $N_i \cup \{i\}$. In particular, we assume that a single contributing agent brings a benefit of 1, two contributing agents bring a benefit of $1 + \alpha$ (with $0 < \alpha < 1$), three contributing agents bring a benefit of $1 + \alpha + \alpha^2$, four contributing agents bring a benefit of $1 + \alpha + \alpha^2 + \alpha^3$.\footnote{In this example, four is the largest number of contributors in $N_i \cup \{i\}$ (for nodes B and C). For more general cases we might think of a longer sum of terms where any additional contributing agent brings a benefit that is $\alpha$ times the benefit of the previous additional contributing agent.} We set the individual cost of contribution $c$ such that $\alpha < c < 1$. This makes the best reply function to be as in (1).

Let us compute for every equilibrium the sum of utilities over all individuals, that we refer to as welfare.

\[
W(NE2) = 7 - 2c, \\
W(NE3_1) = W(NE3_2) = 7 + \alpha - 3c, \\
W(NE5) = 7 + 2\alpha + 2\alpha^2 - 5c.
\]
We may refer to \( W(x) = \sum_{i \in N} u_i(x) \) as the welfare of state \( x \). Since \( \alpha - c < 0 \), then we have that the welfare in \( NE2 \) is always larger than the welfare in \( NE3_1 \) and \( NE3_2 \). We now compare \( W(NE2) \) with \( W(NE5) \): 
\[
W(NE5) > W(NE2) \text{ if and only if } 2\alpha + 2\alpha^2 - 3c > 0.
\]
So, if for instance \( \alpha = 0.8 \) and \( c = 0.9 \), then the Nash equilibrium with maximum welfare is the one with the largest number of contributors. We also observe that in this case welfare is not monotone in the number of contributing agents; in fact, welfare decreases moving from \( W(NE2) \) to \( W(NE3_1) \) (or \( W(NE3_2) \)), while it increases moving from \( W(NE3_1) \) (or \( W(NE3_2) \)) to \( W(NE5) \), where it reaches its maximum. Finally we observe that, if \( \alpha = 0.6 \) and \( c = 0.7 \), then \( W(NE2) > W(NE5) \).

The above example shows that depending on the payoff structure we may have different and possibly non-trivial evaluations of equilibria in terms of welfare. However, a simple result can be stated for the special case in which no other benefit is received for any additional contributor in \( N_i \cup \{i\} \) other than the first one. In particular, suppose that the utility to have at least one contributor in \( N_i \cup \{i\} \) is equal to 1 for all individuals, while it is 0 otherwise, and that the individual homogeneous cost for contribution is \( c < 1 \). In such a case, the welfare of any equilibrium is equal to \( n - c||C(x)|| \), from which we can see that welfare decreases monotonically in the number of contributors in equilibrium. Here Proposition 9 and Proposition 10 (when \( m \) is large enough) predict maximally inefficient equilibria in the very long-run.

We end this section by providing a couple of attempts to generalize our results. A first attempt turns out to be successful. Consider the following modification of the best reply function described in (1), where matrix \( G^\ell \) denotes matrix \( G \) raised to power \( \ell \):
\[
b_i(x) = \begin{cases} 
1 & \text{if } C(x) \cap \{j \in I : j \neq i, G^\ell_{ij} = 1 \text{ for some } \ell \leq k\}, \\
0 & \text{otherwise.}
\end{cases}
\]
(10)
The above best reply function describes a situation in which an agent finds it convenient to contribute if and only if no one is contributing among agents that can be reached from her with at most \( k \) steps in network \( G \). Intuitively, this is a case in which the positive externality due to contribution spreads through the network and is not restricted to immediate neighbors. We note in fact that the best reply function in (1) corresponds to the special case \( k = 1 \). We are able to apply all our results to this modified best reply function after making a simple modification of the network of relationships. In particular, we construct a new network \( G' \) redefining the neighborhood relation among agents: \( G'_{ij} = 1 \) if and only if \( j \neq i \) and \( G^\ell_{ij} = 1 \) for some \( \ell \leq k \). We simply observe that (10) in \( G \) corresponds to (1) in \( G' \), so that we can use \( G' \) as underlying network and apply all the tools and results developed in this paper.

25
A second attempt to generalize our results turns out to be unsuccessful, but nevertheless we consider it useful to deepen our understanding of how the model works. Consider the following modification of the best reply function described in (1):

\[
    b_i(x) = \begin{cases} 
        1 & \text{if } C(x) \cap N_i \leq k, \\
        0 & \text{otherwise.} 
    \end{cases} \tag{11}
\]

In words, an agent finds it convenient to contribute if and only if at most \(k\) of her neighbors are contributing. The best reply function in (1) corresponds to \(k = 0\). The following is a counter-example showing that our results cannot be extended to the case \(k = 1\).

**Example 10.** Consider the network at the top of Figure 8. We assume that every agent finds it convenient to contribute if and only if at most one of her neighbors is contributing. The two Nash equilibria are depicted below in the figure. We work with an individual perturbation process like in (6). To pass from \(NE_3\) to \(NE_4\) it is sufficient that node \(C\) is hit by a perturbation and stops contributing, so that both node \(B\) and node \(D\) can start contributing under dynamics \(T\), and equilibrium \(NE_4\) is reached. To pass from \(NE_4\) to \(NE_3\) it is sufficient that node \(B\) (or \(D\)) is hit by a perturbation and stops contributing, so that under dynamics \(T\) node \(C\) can switch to contribution and, if she does, then node \(D\) (or \(B\)) can switch to defection, thus reaching equilibrium \(NE_3\). Therefore, the resistance from \(NE_3\) to \(NE_4\) and the resistance from \(NE_4\) to \(NE_3\) are both equal to 1, and this means that \(NE_3\) and \(NE_4\) are both stochastically stable. We note that there are 4 contributors in \(NE_4\), and only 3 contributors in \(NE_3\).

6 Conclusions

The best shot network game is a very stark model, and clearly misses details of any specific real-world situation. We think however that, as the model in Schelling (1969) did 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. Moreover, best shot network games have been quite useful to advance our theoretical understanding of network games in general (see BK, Galeotti et al. 2010, Galeotti and Goyal 2010). Equilibrium multiplicity is typically very large in best shot network games, which naturally raises issues of stability and selection.

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 assumptions on the relative likelihood of different types of errors, as indeed occurs in our model (see the detailed discussion in Section 4). Blume (2003) focuses on finding sufficient conditions for errors, in a perturbed dynamics on a discrete action space, such that stochsatic stability always gives 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 it 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 requires agents to perform a task, that however may be hindered by the occurrence of some accident (e.g., a car engine may break down). Instead, it is more difficult to imagine some disturbance that forces agents to take the contributing action. In terms that are more familiar to economists, we may think of the cost $c$ of contribution as an actual amount of money that must be spent by an agent in order to contribute. Suppose now that at each time $t$ every agent receives an income equal to $y + \alpha_t$, with $y > c$ and $\alpha_t$ the time $t$ realization of an idiosyncratic shock with zero mean. With very small probability the realized shock can be negative enough for the actual income to be smaller than $c$ (for simplicity we assume no saving). This scenario suggests an underlying model where shocks may induce contributors not to contribute, but never lead non-contributors to contribute. Our counter-intuitive result is that, exactly in these cases where mostly contributors are hit by perturbations, the selected equilibrium will be the one with highest contribution, i.e., the number of contributors is the largest among equilibria. The welfare evaluation of these stochastically stable states depends of the underlying payoff structure and on the network of relationships, as discussed in Example 9.

For future research we will leave 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 network game, but possibly some of the results achieved in this paper could be generalized. For what concerns generalizations closely related to the model in this paper, we explore two attempts to extend our results to more general versions of the model. First, we show that if the benefit
of contribution flows for some steps through the network of relationships – instead of being limited to immediate neighbors – then a simple redefinition of neighborhood allows us to apply all our results. Second, we provide an example where we show that our selection results cannot be extended to models where agents find it convenient to contribute if and only if at most one neighbor is contributing – instead of no neighbor. Furthermore, in the Appendix we provide an alternative model, based on a specific payoff function but on a different updating of strategies, which behaves as in the case in which contributors are much more likely to be affected by errors than defectors. From a different and broader perspective, we think that further research might fruitfully explore which results about stability selection are robust to the change from a continuous action space to a discrete action space, and vice versa. In particular, BKD focus on 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 games with a discrete strategy space, to compare and possibly generalize their findings with the results that we obtain for stochastic stability.

Acknowledgements

We thank two anonymous referees for their useful comments, which helped us to improve the paper. All mistakes remain ours.

Appendix

A logit-response model of the best shot network game. In usual studies of stochastic stability, mutations often depend on the specific payoffs of the game. Since we work directly with best reply functions, mutations are payoff-independent in our model. In fact, we have considered different error models that depend on the action type, and not on payoffs.

A typical payoff-dependent error model used in the literature is the so-called logit-response mechanism, where player $i$ chooses action $x_i \in X_i$ when $i$’s opponents choose actions $x_{-i} \in X_{-i}$ with probability $p(x_i, x_{-i}) = e^{\beta u_i(x_i, x_{-i})} / \sum_{x'_i \in X_i} e^{\beta u_i(x'_i, x_{-i})}$. This choice function has a long tradition in psychology (see e.g., Thurstone 1927) and economics (see e.g., Blume 1993, McKelvey and Palfrey 1995, Blume 2003).

In the following we will exploit some recent results from the literature to show that equilibria with relatively more contributors are selected by stochastic stability not only if the error model is such that only contributors are hit by mutations, but also if we adopt a logit-response error model with specified payoffs. We think that this exercise further supports
our tempted conclusion that equilibria with relatively more contributors are evolutionary outcomes that are likely to be observed in practice in best shot network games.

BKD show that any game played in a network $G$ with strategies $x_i \in [0, \infty)$ and best reply $f_i(x_i, x_{-i}) = \max(0, 1 - \delta \sum_j G_{ij}x_j)$ turns out to be a potential game when payoffs are given by $u_i(x_i, x_{-i}) = x_i - \frac{1}{2}x_i^2 - \delta x_i(\sum_j G_{ij}x_j)$. The potential is then $\psi(x) = \sum_i(x_i - \frac{1}{2}x_i^2) - \frac{1}{2}\delta \sum_{i,j} G_{ij}x_ix_j$. Setting $\delta = 1$ and $x_i \in \{0, 1\}$, this implies that the best shot game played on network $g$ is a potential game when payoffs are given by $u_i(0, x_{-i}) = 0$ and $u_i(1, x_{-i}) = \frac{1}{2} - n_i(1)$, where $n_i(1)$ is the number of neighbors of $i$ who play 1. It is easy to check that these utilities indeed generate a best shot network game and that the potential is then $\psi(x) = \frac{1}{2}n(1) - n(11)$, where $n(1)$ is the total number of contributors and $n(11)$ is the total number of links between contributors.

We now make use of results described in Young (1998) linking potential games and stochastic stability. Consider the Markov chain such that at time $t + 1$, one individual $i$ is picked at random and plays 1 with probability $e^{\beta u_i(1, x_{t-1})} / (e^{\beta u_i(0, x_{t-1})} + e^{\beta u_i(1, x_{t-1})})$ and 0 otherwise. This is an ergodic Markov chain and the probability of state $x$ in the steady-state distribution is proportional to $e^{\beta \psi(x)}$. As $\beta$ tends to infinity the only states whose probability tends to a positive number are the maxima of the potential. These are the stochastically stable states for this game under logit-response error model. Here, $n(11) = 0$ in any equilibrium, so maxima of the potential are equilibria with largest $n(11)$, that is, equilibria with largest number of contributors.

Of course, the above model relies on one specific payoff function and one specific type of tremble, even if established in the literature. Nevertheless, it is quite remarkable that stochastic stability can give us the same answer through two very different routes.

References


Figure 6: (a) A network of relations, (b) all its Nash equilibria, and (c) from-row-to-column resistances.
Figure 7: (a) A network of relations, and (b) all its Nash equilibria.
Figure 8: (a) A network of relations, and (b) all its Nash equilibria.