ISSN 1755-5361

University of Essex

Department of Economics

Discussion Paper Series

No. 747 February 2014

Constrained Interactions and Social Coordination

Mathias Staudigl and Simon Weidenholzer

Constrained Interactions and Social Coordination*

Mathias Staudigl† Simon Weidenholzer‡

This Version: June 2013

Abstract

We consider a co-evolutionary model of social coordination and network formation where agents may decide on an action in a 2×2 - coordination game and on whom to establish costly links to. We find that a payoff dominant convention is selected for a wider parameter range when agents may only support a limited number of links as compared to a scenario where agents are not constrained in their linking choice. The main reason behind this result is that whenever there is a small cluster of agents playing the efficient strategy other player want to link up to those players and choose the efficient action.

Keywords: Coordination Games, Equilibrium Selection, Learning, Network Formation.
JEL Classification Numbers: C72, D83.

1 Introduction

In many situations people can benefit from coordinating on the same action. Typical examples include common technology standards (e.g. Blue-ray Disc vs. HD DVD), or the choice of common legal standards (e.g. driving on the left versus the right side of the road). These situations give rise to coordination games with multiple strict Nash equilibria. A broad range of global and local interaction models (see e.g. Kandori, Mailath, and Rob (1993), Kandori and Rob (1995), Young (1993), Blume (1993, 1995), Ellison (1993, 2000), or Alós-Ferrer and Weidenholzer (2007)) finds that in coordination games (potentially) inefficient risk dominant conventions will emerge in the long run when agents use myopic best response rules and occasionally make mistakes.[1] The main reason behind this result is that risk dominant strategies perform well in a world of uncertainty, where there is the possibility of misscoordination, and will eventually take over the entire population.

In this paper, we present a model where agents use best response learning to decide on an action in a 2×2-coordination game and to directly choose the set of their opponents. We model this by assuming that agents decide on whom to maintain (costly) links to, thereby giving rise to a model of non-cooperative network formation à la Bala and Goyal (2000).

We study a process of directed network formation where only the active side is involved in the link formation process. As an example consider the case of Twitter where the relationship between followers and followee does not have to be symmetric. [2] In the context of Twitter the baseline coordination game may be thought of as e.g. adopting one of multiple opinions, following a fashion trend, being part of the same subculture, etc. Another example is provided by the the open software development (OSD) phenomenon where the action choice may refer to the usage of a particular programming language and the linking decision refers to incorporating parts of other programmers' code in one's project.[3]

*We are indebted to Andrea Galeotti, Matthew O. Jackson, Manfred Nermuth, and Fernando Vega-Redondo for helpful comments and suggestions. We also would like to thank Tim Hellmann for insightful discussions. Financial support from the Vienna Science and Technology Fund (WWTF) under project fund MA 09-017 is gratefully acknowledged.

†Institute for Mathematical Economics (IMW), Bielefeld University, Postfach 100131 33501 Bielefeld (Germany), E-mail: Mathias.Staudigl@uni-bielefeld.de

‡Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ (United Kingdom), Email: sweide@essex.ac.uk

[1] See Eshel, Samuelson, and Shaked (1998) and Alós-Ferrer and Weidenholzer (2008) for local interaction models showing that imitation learning may lead to efficient outcomes.

[2] Indeed, Kwak, Lee, Park, and Moon (2010) have shown that only 22.1% of the relationships on Twitter are reciprocal.

[3]

We focus on situations where agents do not take into account any payoff that is (possibly) received through passive links when deciding on their action in the coordination game and on their own links. We have two interpretations in mind. Firstly, the payoff of an interaction may only be received by the party that initiated the link. Goyal and Vega-Redondo (2005) put forward the interpretation of such networks as a peer networks, i.e. networks where influence is uni-directional. Such peer networks may, e.g. arise in the context of peer groups or fashion trends where it seems natural that influence is asymmetric. While in many situations also the passive side of an interaction will benefit, one can think of many scenarios where the payoff received from passive interactions is at least on a different level of magnitude as the active interaction, and hence, only plays a minor role when deciding on one's own strategy. For instance, in the OSD example (but also in other settings) it is clear that using others' ideas will advance the own project. However, it is not clear to which degree the passive side knows about - let alone benefits from - its ideas being used.

Our second interpretation allows for payoff also to be received by the active side, provided that this passive payoff does not depend on the action chosen by the opponent. For instance, in the context of Twitter it seems plausible that not only the follower receives some payoff but also the following benefits through having influence or simply feeling important. However, it is also likely that only the number of followees, irrespective of their action choice, determines the followees passive payoff.[4] In our setting where action and link are chosen at the same time agents take the linking decision of other agents as given and, thus under myopic best response learning, do not take into account how their action choice influences their popularity.[5]

We motivate our study of constrained interactions by the observations that in many circumstances the set of agents a typical economic agent links up to is fairly small compared to the overall population. For instance, in the context of the aforementioned twitter example Kwak, Lee, Park, and Moon (2010) show that over 90% of users follow less than 100 people. **Mathias, can you check whether your reading of fig 1 confirms this?** Such constrained interactions will typically arise in situations where there are limitations on the amount of time agents can socialize, where the marginal benefit of socializing is decreasing, and/or the marginal cost of socializing is increasing. Further, the linking technology itself may impose some technological constraints. For instance, Twitter - in order to ensure reliability of the service and discourage spamming - currently only allows certain users to follow more than 2000 others.[6] We further remark that, especially in large populations, constrained interactions impose weaker assumptions on the information agents need to have on the network.

The key feature of constrained interactions is that agents will have to carefully decide on whom to establish one of their links to. Due to the coordination nature of the game, agents will always first try to link up to agents choosing the same action as they do. Only after that they will consider linking up to agents using different actions. Note that as passive payoffs either do not enter agents' payoffs or can not be influenced by the agents the only determinant of agents' future behavior is the current distribution of actions in the population.

When interactions are sufficiently constrained, already a (relatively) small number of agents choosing the payoff dominant strategy enables agents -by linking up to those agents- to secure themselves the highest possible payoff. This has two important consequences: (i) When at the inefficient convention, already a small number of agents playing the payoff dominant action will prompt other players to link up to them and switch to the payoff dominant action. The smaller the constraint the smaller will be the number of individuals required to change to the payoff dominant action. (ii) When all players choose the efficient action a rather large group of players has to change actions. For, if only some agents switch actions it is a best response to sever the ties to those players and link up to the remaining agents using the efficient action. The smaller the constraint, the smaller will be the number of payoff dominant from which one moves back to the payoff dominant convention, and the larger will be the number of players that have to switch to the

[4]Likewise, under discriminatory pricing in telecommunication an agent might care whether she calls somebody with the same or a different telecommunication provider. However, as a recipient of call it should not matter whether it originated on- or off- net.

[5]In a sequential model where first actions and then links are chosen agents would take into account how their action choice will impact their popularity in the link formation stage.

[6]See http://support.twitter.com/articles/68916-following-rules-and-best-practices# for details.

risk dominant action for a transition to occur. Thus, constrained interactions make the efficient convention easier to reach and more difficult to escape.

In this paper we provide a full characterization of the set of long run outcomes under constrained interactions. Essentially, efficient networks will be selected if the number of maximally allowed links is low and/or linking cost are high. Conversely, risk dominant network configurations are only selected if the number of maximally allowed links is high and linking costs are low. We extend the focus of our model by providing a partial characterization of the set of long run outcomes in general $m \times m$ games, an analysis of convex payoff functions, a discussion of heterogenous constraints, and a discussion on frictions in link formation.

1.1 Related Literature

The present work is closely related to the recent literature on social coordination and network formation. As in the present paper, in Goyal and Vega-Redondo (2005) agents may unilaterally decide on whom to maintain links to. In Jackson and Watts (2002) the consent of both parties is needed to form a link, which stipulates the use of Jackson and Wolinsky's (1996) concept of pairwise stability. Further, in Goyal and Vega-Redondo (2005) agents may change strategies and links at the same time whereas in Jackson and Watts (2002) an updating agent may either decide on her action or on a link.[7]

Apart from considering constrained interactions, the only difference between our model and the work by Goyal and Vega-Redondo (2005) is that in their main model the payoff received by agents is on both sides of the link (the active and the passive one) whereas in our setup the payoff is only on the active side (or at least the passive payoff does not depend on the actions used by the opponents,). Goyal and Vega-Redondo (2005) shortly discuss the implications of having only active links in their extensions section.[8] Hojman and Szeidl (2006) also present a model with uni-directional payoff flows. Their analysis focuses on situations where agents receive payoffs from all path-connected neighbors. Goyal and Vega-Redondo (2005) show that, regardless whether the payoff is only on the active side or on both sides of an interaction, for relatively low costs of link formation the risk dominant complete network is selected whereas for relatively high costs of link formation the payoff dominant convention is selected. The main reason for this result is that if costs are low agents obtain a positive payoff from linking to other players irrespective of their action and a complete network will form. This generates endogenously a model of global interactions where the risk dominant action is uniquely selected. If costs of forming links are however high agents may not want to support links to agents using a different strategy, which renders the advantage of the risk dominant action obsolete. Hellmann (2007) provides a partial characterization of the long run outcomes in the Goyal and Vega-Redondo setting where payoff is also enjoyed by the passive party and where agents are constrained in the number of active links. The focus of his study is, however, only on situations where complete networks may form, i.e. where everybody is either actively or passively connected to everybody else. Jackson and Watts (2002) show that, in the context of pairwise stable networks, for low linking costs the risk dominant convention is selected whereas for high linking costs the payoff dominant and the risk dominant conventions are both selected. The main reason behind this discrepancy to the non-cooperative approach is that the nature of transition from one convention to another is different.[9] This is also why Jackson and Watts (2002), in their discussion of a constrained interaction scenario, do not find any relevant effect of constrained interactions on the predictions of their model (with of course the exception being the number of links agents form).

A different branch in the literature analyzes models where agents in addition to their strategy choice in the coordination game may choose among several locations where the game is played (see e.g. Oechssler (1997), Ely (2002), or Bhaskar and Vega-Redondo (2004)).[10] In these models, the most likely scenario will

[7] See also Staudigl (2011) for a model with asynchronous updating using the logit response dynamics and see Corten and Buskens (2010) for a theoretical and experimental analysis of a model with mutual consent in the link formation and sequential updating of links and actions.

[8] See also the survey of **?** for a more detailed analysis.

[9] In Jackson and Watts (2002) this transition is stepwise: starting with a connected component of size two, other players mutating will join one-by-one and we gradually reach the other convention. In Goyal and Vega-Redondo (2005) and in the present paper, once sufficiently many players play one action, all other players will immediately follow.

[10] A related idea has been recently formulated by Jackson and Watts (2010) who study (potentially asymmetric) games

be the emergence of payoff dominant conventions. The reason behind this result is that agents using risk dominant strategies may no longer prompt their interaction partners to switch strategies but instead to simply move away. In this sense, agents can vote by their feet which allows them to coordinate at efficient outcomes. If, however, one is prepared to identify free mobility with low linking costs, this leaves a puzzle to explain: multiple location models select the payoff dominant convention, while the unconstrained network approach favors the risk dominant convention. The main reason for this discrepancy lies in the fact that the multiple location models typically consider average payoffs whereas network models consider additive payoffs. The additive payoff structure implies that all links are valuable for sufficiently low linking costs, giving rise to the risk dominant convention. On the contrary, in multiple locations models the number of potential opponents does not matter and players will always prefer to interact with a small number of players choosing the payoff dominant strategy than with a large number choosing the inefficient strategy. Note that in our constrained links scenario a similar mechanism is at work. If the number of permitted links is relatively small only a small fraction of agents using the payoff dominant action will prompt other agents to link up with them and switch to the payoff dominant action. Dieckmann (1999) and Anwar (2002) present multiple location models where each location is subject to a capacity constraint, thus limiting movements between them. In the multiple location context these constraints on interactions imply that efficient conventions will no longer be selected. Instead, the most likely scenario will be the coexistence of conventions. Thus, in the multiple location context constrained interactions impede efficiency whereas in the network approach they might facilitate efficiency.

Constrained interactions have also been recently received some attention in the wider literature on social networks. Lucier, Rogers, and Immorlica (2011) exhibit a model where rational agents in addition to action choice in a prisoner dilemma may decide through mutual link formation on the set of (anonymous) neighbors. In their setting constrained interactions may bring along cooperative behavior in stationary equilibrium. In this model passive connections act as social capital which is lost upon defection. The constraint on the number of active interactions ensures that defectors can not rebuild social capital quickly which enhances the relative benefits of cooperation. Galeotti and Goyal (2010) study a model where agents can decide on the amount of information privately collected and in addition can connect to other agents to access their information. The returns to information are increasing and concave and substitutable across agents. While agents are ex-ante identical, in equilibrium they will endogenously differentiate themselves in those who directly collect information and those who access this information through the network. The proportion of individuals who directly collect information turns out to be relatively small and the induced networks will display a core-periphery property, and hence display a particular form of constrained interactions.

2 Model Setup

Our model is set in the following environment. We consider N agents who play a 2×2 symmetric coordination game against each other. In addition to choosing an action in the coordination game agents can choose their interaction partners.

Each player i can choose an action $a_i \in \{A, B\}$ in the coordination game. We denote by $u(a_i, a_j)$ the payoff agent i receives from interacting with agent j. The following table describes the payoffs of the coordination game.

	A	B
A	a	c
B	d	b

We assume that $a > d$ and $b > c$ so that, (A, A) and (B, B), are strict Nash equilibria. We further assume that $b > a$ so that the latter equilibrium is payoff-dominant. Further, we assume that $a + c > d + b$ so that the equilibrium (A, A) is risk dominant in the sense of Harsanyi and Selten (1988), i.e. A is the unique best response against an opponent playing both strategies with equal probability. Note that this assumption together with payoff dominance of (B, B) implies $c > d$. We further assume that $a > c$. That is, we

where players in addition to their strategy may choose with whom to interact. This choice of interaction partners is modelled through a cooperative process where players will rematch if they can form a matching group with strictly higher payoffs for all of its members.

restrict our analysis to coordination games where an A-player prefers playing against another A-player over playing against a B-player.[11] Thus, the payoffs in our coordination game are ordered in the following way, $b > a > c > d$.

In addition to their choice in the coordination game, agents can decide on whom to link to. If player i forms a link to player j we write $g_{ij} = 1$ and we write $g_{ij} = 0$ if player i does not form a link to player j. Players may not be linked to themselves, i.e. we have $g_{ii} = 0$ for all $i \in I$. The linking decision of agent i can be summarized by an N-tuple $g_i = (g_{i1}, g_{i2}, \dots, g_{iN}) \in \mathcal{G}_i = \{0,1\}^N$ where g_{ii} is always zero. We denote by $g = (g_i)_{i \in I}$ the network induced by the link decisions of all agents. A pure strategy of an agent consists of her action choice in the coordination game, $a_i \in \{A, B\}$, *and* of her linking decisions, i.e. $s_i = (a_i, g_i) \in \mathcal{S}_i = \{A, B\} \times \mathcal{G}_i$. A strategy profile is a tuple $s = (s_i)_{i \in I} \in \prod_{i \in I} \mathcal{S}_i = \mathcal{S}$. We denote by $d_i^{out} = \sum_j g_{ij}$ the out-degree and by $d_i^{in} = \sum_j g_{ji}$ the in-degree of player i , i.e. the number of players i is actively and the number of players she is passively linked to. Further, we denote by m the number of A-players at a given strategy profile s. Conversely, the number of B-players is given by $N - m$.[12]

We assume that the utility of an agent is given by the sum of payoffs she receives from interacting with each of her neighbors minus a cost of γ for each link sustained. So, given a strategy profile $s = (s_i)_{i \in I}$ the total payoff for player i is given by

$$U_i(s_i, s_{-i}) = \sum_{j=1}^{N} g_{ij} u(a_i, a_j) - \gamma d_i^{out}. \tag{1}$$

This specification is equivalent to one where, in addition to the payoff above, agents also receive payoff which depends on the number of passive connections

$$\tilde{U}_i(s_i, s_{-i}) = \sum_{j=1}^{N} g_{ij} u(a_i, a_j) + -\gamma d_i^{out} + p(d_i^{in}). \tag{2}$$

where $p(\cdot)$ is some increasing function that measures the payoff received from passive interactions. Note that an agent can not influence the number of passive connections, d_i^{in}. Thus, from the point of a best responding agent the formulation (2) is equivalent to (1).

We focus on a scenario where agents may only support a limited number of links k, i.e. $d_i \leq k$ with $1 \leq k \leq N - 1$.[13,14] We assume that agents are not constrained in the number of links they may receive.

In the following, we denote by $\overrightarrow{a}[k] = \{s \in \mathcal{S} | a_i = a \text{ and } d_i = k \quad \forall i \in I\}$ the set of monomorphic states where all agents choose the same action a and each agent supports k links and refer to states in this set as conventions.

We consider a model of *noisy best response learning* in discrete time à la Kandori, Mailath, and Rob (1993) and Young (1993). Each period $t = 0, 1, 2, \dots$ there is a positive independent probability $\lambda \in (0, 1)$ that any given agent receives the opportunity to update her strategy.[15] When such a revision opportunity arises we assume that with probability $1 - \varepsilon$ each agent chooses a strategy (i.e. an action in the base game *and* the set of agents she links to) that would have maximized her payoff in the previous period. More formally, in period t agent i chooses

$$s_i(t) \in \arg\max_{s_i \in \mathcal{S}_i} U_i(s_i, s_{-i}(t-1))$$

where $s_{-i}(t-1)$ is the strategy profile used by all other agents except i in the previous period. If multiple strategies are suggested we assume that agents choose one at random. With the remaining probability

[11] In Proposition 2 we provide sufficient conditions for a payoff dominant action to be selected in a large class of two player games which includes coordination games where $c \geq a$.

[12] Of course, m and $N - m$ will depend on the strategy profile s.

[13] If $k = N - 1$ we have unconstrained interactions, resembling the one sided active links model in the extensions section of Goyal and Vega-Redondo (2005).

[14] Alternatively, one could also model constraints in the number of links by introducing a kinked payoff function, as e.g. in Jackson and Watts (2002). In that case the constraint would come along endogenously.

[15] I.e. we are considering a model of positive inertia, where there is positive probability that a certain fraction of the population does not receive a revision opportunity.

$\varepsilon \in (0,1)$ an updating agent ignores the prescription of the adjustment process and chooses a strategy, consisting of links and payoffs, at random. i.e. she makes a mistake or mutates. We assume that ε is independent across agents, time, and payoffs.

First, note that since passive links do not enter the agents' decision problem the agents' action and linking choice will only depend on the distribution of actions in the population. Since in our model agents may choose both their actions and the set of agents they want to link up to simultaneously, one has to take into account her optimal linking decision when analyzing under which conditions an agent will choose a particular action. The decision problem can therefore be split in two parts: First, determine the optimal set of links for both actions, A and B given the distribution of play in the population. And second, decide which of the two actions to play, given the optimal set of links. We solve the first part of this problem by introducing the concept of a *link optimized payoff function*, for short LOP. The LOP, which we denote by $v(a_i, m)$, of an agent i in period t with action $a_i \in \{A, B\}$ is given by the maximally attainable payoff when linking up optimally given that m agents play A (and $N-m$ agents play B) in period t, i.e.

$$v(a_i, m) = \max_{g_i \in \mathcal{G}_i} \tilde{U}\left((a_i, g_i), m\right),$$

where $\tilde{U}((a_i, g_i), m)$ denotes the payoff received by agent i when playing $s_i = (a_i, g_i)$ when there are in total m A-players in the population. Given the LOPs, we can then solve the second part of our problem, which consists of finding the optimal action. Here, we consider the following myopic best response rule where an agent i with action $a_{i,t-1}$ in period $t-1$, given that m_{t-1} agents chose A in period $t-1$, chooses his action in period t in the following way.[16]

- If $a_{i,t-1} = A$ switch to B if $v(B, m_{t-1}-1) > v(A, m_{t-1})$, randomize between A and B if $v(B, m_{t-1}-1) = v(A, m_{t-1})$, and stay with A otherwise.
- If $a_{i,t-1} = B$ switch to A if $v(A, m_{t-1}+1) > v(B, m_{t-1})$, randomize between A and B if $v(A, m_{t-1}+1) = v(B, m_{t-1})$, and stay with B otherwise.

The process defined above gives rise to a finite state time–homogeneous Markov chain with stationary transition probabilities. The *limit invariant distribution* (as the rate of experimentations tends to zero) $\mu^* = \lim_{\epsilon \to 0} \mu(\epsilon)$ exists and is an invariant distribution of the process without mistakes (see e.g. Freidlin and Wentzell (1988), Kandori, Mailath, and Rob (1993), Young (1993), or Ellison (2000)). It singles out a stable prediction of the original process, in the sense that, for any ϵ small enough, the play approximates that described by μ^* in the long run.

Definition 1. The states in the support of μ^*, $S = \{\omega \in \Omega \mid \mu^*(\omega) > 0\}$ are called *Long Run Equilibria (LRE)* or stochastically stable states.

In particular, we will be using a lemma based on Ellison's (2000) Radius-Coradius Theorem and on methods developed by Freidlin and Wentzell (1988) to identify the set S.[17]

3 Constrained Interactions

Note that, since we are considering coordination games with $a > c$ and $b > d$, all agents will first try to link up to other agents using the same action and only then may consider linking up to other agents using a different action.[18] However, whether agents will indeed link up to agents using a different action or not will depend on the relative magnitude of the linking cost γ. First, if linking costs are relatively low, $0 \leq \gamma \leq d$, all agents will first connect to agents of their own kind and will only then fill up the remaining slots with

[16] This decision rule is similar to Sandholm (1998). We deviate from Sandholm's (1998) original rule, which prescribes agents to stay put in case of ties, by imposing random tie breaking. We have also considered the case where agents stay put at their current action in case of a payoff tie and have found that the long run predictions remains unaffected.

[17] See Appendix A for details.

[18] If we were considering coordination games with $c > a$, A-players will first link up to B players and only then consider players of their own kind. This would give rise to different dynamics, where the payoff of playing A (weakly) increases the fewer A-agents there are. Note, however, that since (B, B) is a payoff dominant Nash equilibrium it will eventually be optimal to switch to B if the number of A-players is sufficiently small.

agents using different actions. Consequently, the LOPs of an A-player and of a B-player, when confronted with a distribution of play $(m, N-m)$, are given by

$$v(A,m) = a\min\{k, m-1\} + c(k - \min\{k, m-1\}) - \gamma k$$
$$v(B,m) = b\min\{k, N-m-1\} + d(k - \min\{k, N-m-1\}) - \gamma k.$$

Now consider the intermediate cost scenario where $d \leq \gamma \leq c$. In this case B-players will only link up to other B-players whereas A- players will first link up to all other A-players and only then will also link up to B-players, yielding

$$v(A,m) = a\min\{k, m-1\} + c(k - \min\{k, m-1\}) - \gamma k$$
$$v(B,m) = (b-\gamma)\min\{k, N-m-1\}.$$

For high linking costs, $c \leq \gamma \leq a$, agents will only interact with agents using the same action and we obtain

$$v(A,m) = (a-\gamma)\min\{k, m-1\}$$
$$v(B,m) = (b-\gamma)\min\{k, N-m-1\}.$$

For each of these three scenarios, we can now identify conditions under which an A-player will switch to B with positive probability, i.e. $v(B, m-1) \geq v(A,m)$ and under which a B-player will switch to A, i.e. $v(A, m+1) \geq v(B,m)$ (again with positive probability). Depending on the relationship between m, N, and k, we have to analyze four sub-cases for each of our three cost scenarios.[19] We report our findings in Table 1.[20]

Switching thresholds for A-players			
$v(B,m-1) \geq v(A,m)$	$k > m-1$ $k > N-m$	$k \leq m-1$ $k > N-m$	$k \leq N-m$
$0 \leq \gamma \leq d$	$m \leq \frac{(N-1)(b-d)-k(c-d)}{a+b-c-d} + 1 := \psi_1^\ell$	$m \leq N - \frac{a-d}{b-d}k := \psi_2^\ell$	$a.s.$
$d \leq \gamma \leq c$	$m \leq \frac{(N-1)(b-\gamma)-k(c-\gamma)}{a+b-c-\gamma} + 1 := \psi_1^m$	$m \leq N - \frac{a-\gamma}{b-\gamma}k := \psi_2^m$	$a.s.$
$c \leq \gamma \leq a$	$m \leq \frac{(N-1)(b-\gamma)}{a+b-2\gamma} + 1 := \psi_1^h$	$m \leq N - \frac{a-\gamma}{b-\gamma}k := \psi_2^h$	$a.s.$
Switching thresholds for B-players			
$v(A,m+1) \geq v(B,m)$	$k > m$ $k > N-m-1$	$k \leq m$ $k > N-m-1$	$k \leq N-m-1$
$0 \leq \gamma \leq d$	$m \geq \frac{(N-1)(b-d)-k(c-d)}{a+b-c-d} := \psi_1^\ell - 1$	$m \geq N-1-\frac{a-d}{b-d}k := \psi_2^\ell - 1$	$n.s.$
$d \leq \gamma \leq c$	$m \geq \frac{(N-1)(b-\gamma)-k(c-\gamma)}{a+b-c-\gamma} := \psi_1^m - 1$	$m \geq N-1-\frac{a-\gamma}{b-\gamma}k := \psi_2^m - 1$	$n.s.$
$c \leq \gamma \leq a$	$m \geq \frac{(N-1)(b-\gamma)}{a+b-2\gamma} := \psi_1^h - 1$	$m \geq N-1-\frac{a-\gamma}{b-\gamma}k := \psi_2^h - 1$	$n.s.$

Table 1: Where "a.s." means that a player always switches to the other action and "n.s." means that a player never switches to the other action.

With the help of Table 1 we are now able to characterize the set of absorbing sets, i.e. sets of states that can not be left under the dynamics without mistakes.

Lemma 1. *The sets $\overrightarrow{A}[k]$ and $\overrightarrow{B}[k]$ are the only absorbing sets.*

Proof. In each of the above cases, we find that if it is optimal for an agent to remain at her action then it is optimal for agents using a different action to switch. Consider any state $s \notin \overrightarrow{A}[k] \cup \overrightarrow{B}[k]$ and give revision

[19] In the first sub-case, with $k > m-1$ and $k > N-m$ neither A- nor B- players may fill all their slots with agents of their own kind. In the second sub-case, $k > m-1$ and $k \leq N-m$, A-players do not find enough A-players to fill up all their slots whereas B-players can fill up all their slots with other B-players. In the third case, with $k \leq m-1$ and $k > N-m$, A-players will only link to other A players whereas B-players can not fill up all their slots with other B-agents. In the remaining case, with $k \leq m-1$ and $k \leq N-m$, both A- and B- players will link up only to agents of their own kind.

[20] For the referees' convenience we report the derivation in an appendix which is not intended for publication.

opportunity to agent i with action a_i. If the agent remains at her action we know that all subsequent agents will either switch to that action or remain at that action and we arrive at a state in $\overrightarrow{A}[k] \cup \overrightarrow{B}[k]$. If the revising agent i switches to the other action we give a revision opportunity to agents who also choose a_i. Those agents will all switch to the other action and we arrive at a state in $\overrightarrow{A}[k] \cup \overrightarrow{B}[k]$. Further, consider the states $\overrightarrow{A}[k] \cup \overrightarrow{B}[k]$ and note that under our best response process ties are broken randomly. This implies that agents are indifferent between having links to, say, agent i and agent j. It follows that for each pair of states $s, s' \in \overrightarrow{A}[k]$ (and also for each pair in $\overrightarrow{B}[k]$) there is a positive probability of moving from s to s' without mistakes, i.e. all states in $\overrightarrow{A}[k]$ (and all sates in $\overrightarrow{B}[k]$) form an absorbing set. ■

We are now able to state our main theorem which characterizes the set of long run equilibria.

Theorem 1. *Under constrained interactions we have that,*

- *for low linking costs, $0 \leq \gamma \leq d$, there exist two thresholds, $\underline{k}^\ell$ and $\bar{k}^\ell$, with $\bar{k}^\ell \geq \underline{k}^\ell > \frac{N-1}{2}$, such that (i) for $k < \underline{k}^\ell$, $S = \overrightarrow{B}[k]$, (ii) for $k \in [\underline{k}^\ell, \bar{k}^\ell]$, $S = \overrightarrow{A}[k] \cup \overrightarrow{B}[k]$, and (iii) for $k > \bar{k}^\ell$, $S = \overrightarrow{A}[k]$;*
- *for intermediate linking costs, $d \leq \gamma \leq c$, there exist two thresholds, $\underline{k}^m$ and $\bar{k}^m$, with $\bar{k}^m \geq \underline{k}^m > \frac{N-1}{2}$, such that (i) for $k < \underline{k}^m$, $S = \overrightarrow{B}[k]$, (ii) for $k \in [\underline{k}^m, \bar{k}^m]$, $S = \overrightarrow{A}[k] \cup \overrightarrow{B}[k]$, and (iii) for $k > \bar{k}^m$, $S = \overrightarrow{A}[k]$;*
- *for high linking costs, $c \leq \gamma \leq a$, there exists a thresholds k^h, with $k^h > \frac{N-1}{2}$, such that (i) for $k \leq k^h$, $S = \overrightarrow{B}[k]$ and (ii) for $k > \bar{k}^h$, $S = \overrightarrow{A}[k] \cup \overrightarrow{B}[k]$;*

We have relegated the derivation of the thresholds and the proof of Theorem 1 to the appendix. Let us provide some technical intuition for our result using the case when agents only support links less than half of the other agents, $k < \frac{N-1}{2}$. First, consider the A-convention. Assume that k players mutate to action B and choose any linking strategy.[21] By severing all links to A-players, linking up to the B-players, and switching to B, the remaining A-players can ensure themselves the maximally attainable payoff. Thus, k mutations are sufficient to move to the B-convention. Consider the B-convention and assume that x players have mutated to B. We want to understand how large x has at least got to be for a transition to occur. To this end, note that if after the mutations $k+1$ B-players are left, by the above observation, all players will link up to those players and either switch to or remain at B. Thus, we would remain at the B convention. Thus, after the mutations less than $k+1$ B-players have to be left. This requires at least $N-1-k$ mutations. If $k < \frac{N-1}{2}$ the number of required mutation to leave the efficient convention is larger than the number to enter it, the efficient convention is unique LRE. Note that we only used the property that if a player can fill all of her links with B-players she will for sure switch. For a given game, this switch will however occur with less B-players. In the proof we provide necessary and sufficient conditions for the transitions to occur and are able to provide a complete characterization of the set of LRE.

We now proceed to contrast our results under constrained interactions to the unconstrained interaction case. For the sake of concreteness, we focus our discussion on one particular case where the thresholds take a rather simple form and remark that the qualitative insights are not altered by this restriction.

Remark 1. In the proof of Theorem 1 we obtain explicit expressions for the thresholds $\underline{k}^\ell$, $\bar{k}^\ell$, $\underline{k}^m$, $\bar{k}^m$, and k^h. In the non-generic case (where $\psi_1^\ell, \psi_1^m, \psi_1^h \notin \mathbb{Z}$) and N is odd these are given by

$$\underline{k}^\ell = \bar{k}^\ell = \frac{N-1}{2}\left(\frac{b-a}{c-d}+1\right), \quad \underline{k}^m = \bar{k}^m = \frac{N-1}{2}\left(\frac{b-a}{c-\gamma}+1\right), \quad \text{and} \quad k^h = N-1.$$

Setting $k = N-1$ in the above thresholds reveals that under unconstrained interaction we have a linking cost threshold $\gamma^* = a + c - b$, such that (i) $S = \overrightarrow{A}[N-1]$ if $\gamma \leq \gamma^*$ and (ii) $S = \overrightarrow{B}[N-1]$ if $\gamma > \gamma^*$.

Note that the linking cost threshold in the unconstrained interaction scenario γ^* lies in the range (d, c). Thus, under unconstrained interactions the risk dominant convention is selected for (relatively) low linking cost and the payoff-dominant convention is only selected for relatively high linking costs. This is in stark

[21] Note that since passive payoffs do not enter the decision problem of the remaining agents the mutation can involve any linking strategy.

contrast to the case of constrained interactions, where the payoff dominant convention may be selected even for low linking costs. In particular, note that we have $\underline{k}^\ell, \underline{k}^m, k^h > \frac{N-1}{2}$. Thus, if agents may support at most links to half of the population, $k \leq \frac{N-1}{2}$, the payoff dominant convention is always selected, regardless of the level of linking costs. In Figure 1 we highlight our results by plotting the parameter combinations under which either of the two conventions is LRE for general linking costs $0 \leq \gamma \leq a$ and $1 \leq k \leq N-1$ permitted links. Note that the right border of Figure 1 corresponds to the unconstrained interaction scenario. In contrast to this unconstrained interaction case, the efficient convention is selected for a quite large range of parameter combinations.

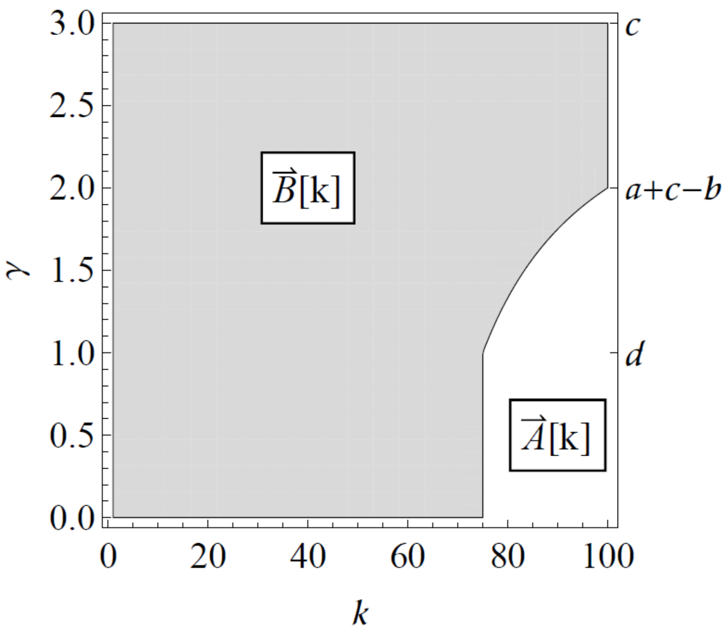

Figure 1: LRE in the game $[a, c, d, b] = [4, 3, 1, 5]$ with $N = 101$. Note that the line segment separating the two the selection regions is $\underline{k}^\ell = \bar{k}^\ell$ for $\gamma \in [0, d]$ and $\underline{k}^m = \bar{k}^m$ for $\gamma \in [d, c]$.

4 Extensions

4.1 General games

It is interesting to note that, by Theorem 1, we always have $S = \overrightarrow{B}[k]$ for $k \leq \frac{N-1}{2}$. This implies that if agents may only support links to less than half of the population we will always observe efficient outcomes in the long run. Note that this insight is not confined to the class of 2×2-coordination games, but can be easily generalized to $r \times r$ games in the presence of a payoff dominant strategy. To this end, let us consider a two-player symmetric game with action set $\mathcal{A} = \{a^1, \ldots, a^r\}$.

We say that an action a is (uniquely) payoff dominant if it is a symmetric NE with the highest pure-action payoff, i.e. $u(a, a) > u(a', a'')$ for all $a', a'' \neq a$ and $u(a, a) > u(a', a)$ for all $a' \neq a$. We use Morris, Rob, and Shin's (1995) notion of $\frac{1}{2}$-dominance to generalize the concept of risk dominance to $r \times r$-games. In this sense, a strategy a is said to be $\frac{1}{2}$-dominant if a is the unique best response against any mixed strategy putting at least probability $\frac{1}{2}$ on the pure strategy a. Further, we denote by $\underline{u} = \min\{u(a, a') | a, a' \in \mathcal{A}\}$ the lowest payoff in the game. Then,

Proposition 2. *Assume that* $0 \leq \gamma \leq \underline{u}$. *For* N *sufficiently large we have*

(i) If a *is* $\frac{1}{2}$*-dominant and* $k = N-1$ *we have that* $S = \overrightarrow{a}[N-1]$.

(ii) If a *is payoff dominant and* $k \leq \frac{N-1}{2}$ *we have that* $S = \overrightarrow{a}[k]$.

We have relegated the proof into the appendix. Note that this partial characterization also includes coordination game with $c > a$ (see footnote 11). The intuition behind the first part is that under unconstrained interactions all links will form and we obtain a model of global interactions where the $\frac{1}{2}$-dominant strategy is selected. The second part exploits the idea that if agents may only support links to k other agents, then agents who choose the payoff dominant action and link up to k other agents choosing the payoff dominant action will receive the highest possible payoff. If $k \leq \frac{N-1}{2}$ then less than half of the population playing the payoff dominant action will cause other players to shift. This essentially implies that the payoff dominant convention becomes easier to reach than to leave.

4.2 Convex Linking Costs

We will now consider the case where the cost of forming links is convex in the degree of a player. In this setting the number of links an agent maximally supports arises endogenously. Here we find that if the cost function is sufficiently curved, so that agents will not link to everybody in the population, the payoff dominant convention is selected. However, for sufficiently less curved cost functions agents will link to everybody in the population and the selection of the risk dominant convention remains.

We assume that the cost of forming d links is given by $\phi(d)$, with $\phi(0) = 0$, $\phi'(\cdot) > 0$, and $\phi''(\cdot) > 0$. Thus, the payoff of an agent is given by

$$U_i(s_i, s_{-i}) = \sum_{j=1}^{N} g_{ij} u(a_i, a_j) - \phi(d_i).$$

Let $d_{x|y}$ denote the outdegree of a typical $x \in \{A, B\}$ player going to the pool of $y \in \{A, B\}$ players. Then $d_x = d_{x|A} + d_{x|B}, x \in \{A, B\}$ is the total outdegree of a typical $x \in \{A, B\}$-player. The LOPs of an A-player and of a B-player are given by:

$$\begin{aligned} v(A, m) &= a d^*_{A|A} + c d^*_{B|A} - \phi(d^*_A), \\ v(B, m) &= d d^*_{B|A} + b d^*_{B|B} - \phi(d^*_B) \end{aligned}$$

where $d^*_{x|y}(m)$ is the optimal out-degree of an $x \in \{A, B\}$ player with $y \in \{A, B\}$ players, when there are in total m A-players in the population. The optimal number of A links of an A-player, $d^*_{A|A}(m)$, is the highest number of links such that the creation of an additional link to an A-player will not weakly increase the utility of the player. Thus, it is given by

$$d^*_{A|A}(m) := \max \{x \in \{0, 1 \ldots, m-1\} | a \geq \phi(x) - \phi(x-1)\}.$$

Now let us attend to the optimal number of B-links of an A-player. An A-player with $d^*_{A|A}(m)$ links to A-players will establish links to B-players as long as each additional link carries a positive net utility, i.e. $d^*_{A|A}(m)$ is characterized by

$$d^*_{A|B}(m) := \max \left\{x \in \{0, 1 \ldots, N-m\} | c \geq \phi\left(d^*_{A|A}(m) + x\right) - \phi\left(d^*_{A|A}(m) + x - 1\right)\right\}.$$

Now note that if $d^*_{A|A}(m) < m-1$ we have that $\phi(d_{A|A}(m)+1) - \phi(d_{A|A}(m)) > a > c$ as $d_{A|A}(m)$ is the largest integer for which $\phi(d_{A|A}(m)) - \phi(d_{A|A}(m)-1)$ holds and (by convexity) $\phi(x+1) - \phi(x)$ is increasing in x. Thus, if an A- agent does not form links to all other A-agents, he will never consider linking up to a B-agent and we have $d^*_A(m) = d^*_{A|A}(m)$. Conversely, if $d_{A|A}(m) = m-1$, additional links to B-players are possible.

The optimal linking strategy of a B-player can be characterized by

$$d^*_{B|B}(m) := \max \{x \in \{0, 1 \ldots, N-m-1\} | b \geq \phi(x) - \phi(x-1)\}$$

and

$$d^*_{B|A}(m) := \max \left\{x \in \{0, 1 \ldots, m\} | d \geq \phi\left(d^*_{B|B}(m) + x\right) - \phi\left(d^*_{B|B}(m) + x - 1\right)\right\}.$$

Now note that if $\phi(N-1) - \phi(N-2) \leq d$ a B-player will establish links to all other A players. Since $d < c < a < a$, this tells us that the B-player will also establish links to all other B-players and that also A-players will link up to everybody. Thus, we have $d^*_{A|A}(m) = m-1$, $d^*_{A|B}(m) = N-m$, $d^*_{B|B}(m) = N-m-1$, and $d^*_{B|A}(m) = m$, implying global interactions. Another observation is that if $\phi(x) - \phi(x+1) > b$, so that B-players establish at most x links to other B-players, no player will establish more than x links to other players, i.e. $d^*_A(m), d^*_B(m) \leq x$ for all $m \in \{0, 1, \ldots N\}$. These two observations allow us to provide the following partial characterization of LRE (the proof of which can be found in the appendix).

Proposition 3. *Under convex linking costs we have for N sufficiently large*

(i) If $\phi(N-1) - \phi(N-2) \leq d$ we have that $S = \overrightarrow{A}[N-1]$.

(ii) If $\phi\left(\frac{N-1}{2}\right) - \phi\left(\frac{N+1}{2}\right) > b$ we have that $S = \overrightarrow{B}[d^*_{B|B}(0)]$.

4.3 Heterogenous Constraints

We will now shortly explore a variant of our model that allows for heterogeneous constraints. Each agent i is assumed to be able to maximally support k_i links. We denote by n_k the number of agents with k links and by $\vec{n} = (n_1, n_2, \ldots, n_{N-1})$ the distribution of thresholds in the population.

One can easily extend our Theorem 1 to this framework: If all agents hav e thresholds smaller than the ones identified in the Theorem then the payoff dominant outcome will be the unique long run equilibrium. A more interesting result may arise when some agents are only able to connect to a small subset and others may connect to almost everybody else. To study this idea in more detail consider a population of $N = 101$ players and the case of zero linking costs, $\gamma = 0$. Suppose that the threshold distribution is given by

$$n_1 = 2, n_j = 1, \text{ for } j = 2, \ldots, 100$$

First, consider the risk dominant conventions where all players support all of their links. Assume that, e.g. the player with the most links mutates to B. When given revision opportunity the players with $k_i = 1$ want to link up to this player and switch to B. Now we have three B-players. In a next step the $k_i = 2$ player will switch to B and link up to them. Iterating this argument and by the construction of the threshold distribution we are able to reach the payoff dominant convention with just one mutation.

Interestingly, this observation does not imply that the payoff dominant convention is unique long run equilibrium. When the risk dominant action has a "high degree of risk dominance" it might be that both the risk- and the payoff- dominant conventions are long run equilibria. To see this point consider the game $[a, c, d, b] = [199, 198, 0, 200]$. While B is still payoff dominant A is a best response whenever 1% or more a player's opponents use it. Consider now the payoff dominant convention and assume that one of the $k_i = 1$ agents mutates to A. The $k_i = 100$ player is linked to everybody and will switch to B. Now the $k_i = 99$ player faces two B-players, implying that he will at least link to one of them. Thus, he has one of his 99 opponents playing B. Consequently, he will switch too. We can iterate this argument to show that players will one-by-one switch to A. This implies that in this example both conventions are long run equilibria.

In this example heterogenous constraints allowed the payoff dominant action to spread in a way that - starting with players with few allowed links and ending with those with many - the constraints of the various players were reached. However, at the same time heterogenous constraints also made the payoff dominant convention vulnerable to being taken over by risk dominant actions in certain games. For, agents with many links were not able to avoid the influence of A-players. The direction of transition was however in the opposite direction, starting with players with many links and ending with players with only a few links. Note that, in more general settings, which convention is selected in the long run will depend on the particular constraint distribution and the parameters of the game (beyond their implications for payoff- and risk dominance.)

4.4 Frictions in Link Adaptation

An important factor behind results is the ability of agents to sever their current links, change their action, and link up to players using the payoff dominant action. This allows them to i) focus all of their interactions on payoff dominant players and ii) avoid interacting with players using the risk dominant action. In the following we explore what happens if there are frictions in this link adaptation process. We focus on a natural kind of friction in link adaptation where agents can not change all of their links at the same time.[22]

To study this issue let us consider the case of zero linking costs, $\gamma = 0$, and a large population. First, assume that each agent may support two links, one of which is changeable at the same time as the action. It is still true that only the only the sets $\overrightarrow{A}[K]$ and $\overrightarrow{B}[K]$ are the only absorbing states. Further, due to random tie breaking in case of payoff ties, any state in $\overrightarrow{A}[K]$ (and any state in $\overrightarrow{B}[K]$) can be reached from any other state without mutations. Consider the risk dominant conventions and assume that the process is in a state where agents $3, \ldots, N$ have all of their links to agents 1 and 2. If those two agents mutate to B, the remaining agents will follow and we reach the convention $\overrightarrow{B}[K]$. Interestingly, it is also possible to access the convention $\overrightarrow{A}[K]$ at the cost of two mutations: Consider the efficient convention $\overrightarrow{B}[K]$ and assume that the network is in a state with a linking configuration as above. Assume agents 1 and 2 mutate to A. All other agents can at most change one link and, thus, face at best a configuration with one A- and one B- players. Thus, by risk dominance, both conventions are accessible from each other by one mutation, implying that both are long run equilibria.

It is easy to come up with examples where the prediction may change depending on the degree of frictions in link adaptation. To see this point, assume that agents may support up to six links and consider the game $[a, c, d, b] = [7, 6, 0, 9]$. In this game A is a best response if at least three in ten opponents use it. First, consider the case when only one link is changeable at a time. Consider the efficient convention $\overrightarrow{B}[K]$ and consider a state where agents $4, \ldots, N$ support links to agents $1, \ldots, 3$. Assume that those three agents mutate to A. Since all other agents can only change one link they will at least have two A-opponents. Thus, it is optimal for them to switch, establishing that the risk dominant convention can be reached with three mutations. Conversely, consider the risk dominant convention and assume that three agents mutate to B. Every player will have at least three of her opponents playing A. Thus, playing B can not be optimal for any player, implying that the A-players will remain and the B-players will switch back. Thus, the risk dominant convention cannot be left with three mutations, establishing that it is the unique long run equilibrium. Note that if, however, five of the six links are changeable the prediction reverts to the payoff dominant convention. To see this point note that linking up to B-players and choosing B is a unique best response whenever there are five or more B-players in the population. Since, agents can switch five links at a time, they can focus a sufficiently high number of their interactions on payoff dominant players and the same logic that underlies our main result applies.[23]

5 Conclusion

We have presented a model of social coordination and network formation where agents may only support a limited number of links. We find that under sufficiently constrained interactions a population of myopic players will learn to coordinate on efficient convention in the long run. This is in sharp contrast to the unconstrained interaction case where they may get stuck at risk dominant conventions.

In this note we have concentrated on a situation were agents face no constraints with respect to the number of incoming links. However, there are many situations where this is not the case, e.g. when socialising is costly. We have also focused on interactions where (any potential) payoff enjoyed by the passive party does not depend on the actions used by the active parties. A natural extension would, thus, be to include these two features in our model. A possible starting point to model such situations might be the "two sided links through independent decisions" model discussed in the extension section of Goyal and Vega-Redondo (2005) where a link is only formed if both parties involved offer to form it. One difficulty that would arise in such a setting is that under constrained interactions there might actually be agents who receive more link requests than free slots. Thus, in a setting where passive links are also constrained

[22] An alternative form of friction might occur if forming new links (or changing existing) links is costly. Our analysis of this scenario revealed that for a low level of switching costs constrained interactions still foster payoff dominant outcomes. For larger levels of switching costs the predictions might, however, be reversed. The mechanism that underlies this finding, albeit more complicated, is very similar to the mechanism outlined here.

[23] One can also check that for two links changeable the risk dominant convention is selected, for four links changeable the payoff dominant one is selected, and for three links changeable both are selected.

one would also have to model how active link requests and (free) passive links are matched. We think that these points deserve further attention and leave them as a topic for further research.

Appendix A Review of techniques

We refer to the process without mistakes ($\epsilon = 0$) as the *unperturbed* process and call the process with mistakes ($\epsilon > 0$) the *perturbed* process. Since $P(\epsilon)$ is strictly positive for $\epsilon > 0$, the perturbed process always has a unique strictly positive invariant distribution $\mu(\epsilon) \in \Delta(\Omega)$.

Ellison (2000) presents a powerful method to determine the set of LRE which is based on a characterization by Freidlin and Wentzell (1988). Let X and Y be two absorbing sets of the unperturbed process and let $c(X,Y) > 0$ be the minimal number of mistakes needed for a direct transition from X to Y (i.e. the cost of transition). Define a path P of length $\ell(P)$ from X to Y as a finite sequence of absorbing sets $P = \{X = S_0, S_1, ..., S_{\ell(P)} = Y\}$ and let $S(X,Y)$ be the set of all paths from X to Y. The *cost* of the path is given by the sum of its transition costs

$$c(P) = \sum_{k=1}^{l(P)} c(S_{k-1}, S_k).$$

The minimal number of mistakes required for a (possibly indirect) transition from X to Y is

$$C(X,Y) = \min_{P \in S(X,Y)} c(P).$$

The *radius* of an absorbing set X is defined as

$$R(X) = \min\{C(X,Y) | Y \text{ is an absorbing set, } Y \neq X\},$$

i.e. the minimal number of mistakes needed to leave X. The *coradius* of X is defined as

$$CR(X) = \max\{C(Y,X) | Y \text{ is an absorbing set, } Y \neq X\},$$

i.e. the maximal number of mistakes needed to reach X. In the proof we will make use of the following lemma:

Lemma 4. *Let X and Y be two absorbing sets. Then:*

(a) If $R(X) > CR(X)$, then $S = X$.

(b) If there are only two absorbing sets and if $R(X) = CR(X)$, then $S = X \cup Y$.

Part (i) of the lemma is simply the Radius-Coradius theorem of Ellison (2000). Part (ii) can be proved using the methods characterized by Freidlin and Wentzell (1988). To this end, define an X-tree as a directed tree such that the set of nodes are the absorbing sets, and the tree is directed into the root X. The cost of a tree is the sum of the costs of transition for each edge. A state is then LRE if and only if it is the root of a minimum cost tree. If there are only two absorbing sets we have that the cost of an X-tree is $c(Y,X) = CR(X) = R(Y)$ and the cost of a Y-tree is given by $c(X,Y) = CR(Y) = R(X)$. If $R(X) = CR(X)$ we have that both the X-tree and the Y-tree are of minimal cost, implying $S = X \cup Y$.

Appendix B Proofs

Proof of Theorem 1

Using the switching thresholds reported in Table 1 we can identify the set of long run equilibria S. We will analyze the low, the intermediate, and the high cost scenario in turn. For each of these three cases the proof proceeds in two steps: i) to determine the *Radius* and the *Coradius* of the absorbing sets and ii) apply part (b) of Lemma 4 to identify the set of LRE.

B.1 The low cost case

B.1.1 Computing the Radius and Coradius

Transition from $\overrightarrow{A}[k]$ to $\overrightarrow{B}[k]$.

First consider the transition from $\overrightarrow{A}[k]$ to $\overrightarrow{B}[k]$. Take a state in $\overrightarrow{A}[k]$. We denote by m^{AB} the remaining number of A-players after the necessary number of mutations towards action B have occurred. Thus, m^{AB} is the maximum number of A-players such that the transition from $\overrightarrow{A}[k]$ to $\overrightarrow{B}[k]$ occurs with positive probability. Hence, $N - m^{AB}$ is the minimum number of B-players making B a best-response. Attending to Table 1 we see that whenever $m \leq N-k$ this transition always occurs. Thus, we know that $m^{AB} \geq N-k$ must be true. For, in case the remaining number of A-players after the mutations is lower or equal than $N-k$ we already know that A players always switch. Now consider the case where $m^{AB} > N-k$. In this case the $N - m^{AB}$ B-players can not fill their k slots with fellow B-players. There are two possibilities we have to consider: (i) $m^{AB} < k+1$ and (ii) $m^{AB} \geq k+1$.

If $m^{AB} < k+1$, A-players will link up to both kinds of players after the mutations have happened. It follows from Table 1 that in this case $m^{AB} = \lfloor \psi_1^\ell \rfloor$. If $m^{AB} \geq k+1$, then the m^{AB} remaining A-players will only link up to other A-players. Attending Table 1, we see that $m^{AB} = \lfloor \psi_2^\ell \rfloor$. First let us look at the inequality $\lfloor \psi_1^\ell \rfloor < k+1$. Since, $k \in \mathbb{Z}$ this holds if $\psi_1^\ell < k+1$, which translates into $k > \frac{(N-1)(b-d)}{a+b-2d}$. Second, we see that $\lfloor \psi_2^\ell \rfloor \geq k+1$ if $\psi_2^\ell \geq k+1$, which translates into $k \leq \frac{(N-1)(b-d)}{a+b-2d}$. Recalling that m^{AB} has to be larger or equal to $N-k$, we have that

$$m^{AB} = \begin{cases} \max\left\{\lfloor \psi_2^\ell \rfloor, N-k\right\} & \text{if } k \leq \frac{(N-1)(b-d)}{a+b-2d} \\ \max\{\lfloor \psi_1^\ell \rfloor, N-k\} & \text{if } k > \frac{(N-1)(b-d)}{a+b-2d} \end{cases} .$$

One can check that $\lfloor \psi_2^\ell \rfloor \geq N-k$ and that $\lfloor \psi_1^\ell \rfloor \geq N-k$ whenever $k \geq \frac{(N-1)(a-c)}{a+b-2c}$. Since $\frac{(N-1)(a-c)}{a+b-2c} < \frac{(N-1)(b-d)}{a+b-2d}$, this is always the case in the relevant range. Thus, we have that

$$m^{AB} = \begin{cases} \lfloor \psi_2^\ell \rfloor & \text{if } k \leq \frac{(N-1)(b-d)}{a+b-2d} \\ \lfloor \psi_1^\ell \rfloor & \text{if } k > \frac{(N-1)(b-d)}{a+b-2d} \end{cases} .$$

For the transition from $\overrightarrow{A}[k]$ to $\overrightarrow{B}[k]$ it has to be the case that there are at most m^{AB} A-players in the population. Thus, $N - m^{AB}$ A-players must switch from A to B, establishing $R(\overrightarrow{A}[k]) = CR(\overrightarrow{B}[k]) = N - m^{AB}$.

Transition from $\overrightarrow{B}[k]$ to $\overrightarrow{A}[k]$

Now consider the transition from $\overrightarrow{B}[k]$ to $\overrightarrow{A}[k]$. We denote by m^{BA} the minimal number of A-players so that the remaining B-players have A as their best-response. We can infer from Table 1 that B-players will never switch if $k \leq N-m-1$. Thus, for B-players to switch $m^{BA} > N-k-1$ must be true. Now there are two remaining possibilities: $m^{BA} < k$ or $m^{BA} \geq k$. We can infer from Table 1 that $m^{BA} = \lceil \psi_1^\ell \rceil - 1$ in the first case, and $m^{BA} = \lceil \psi_2^\ell \rceil - 1$ in the second case. We then have $\lceil \psi_1^\ell \rceil - 1 < k$ if $\psi_1^\ell \leq k$, which gives $k \geq \frac{N(b-d)}{a+b-2d} + \frac{a-c}{a+b-2d}$. Likewise, $\lceil \psi_2^\ell \rceil - 1 \geq k$ if $\psi_2^\ell > k$, which yields $k < \frac{N(b-d)}{a+b-2d}$.

It remains to see what happens for $k \in \left[\frac{N(b-d)}{a+b-2d}, \frac{N(b-d)}{a+b-2d} + \frac{a-c}{a+b-2d}\right)$. We claim that in this range $m^{BA} \geq k$, and prove the claim via contradiction. Assume that in this range for k it is the case that $m^{BA} < k$. Then the new A-agents will also have to link up to B-agents. For a B-player to switch we must have $m \geq m^{BA} \geq \lceil \psi_1^\ell \rceil - 1$. Thus, $k > m^{BA} \geq \lceil \psi_1^\ell \rceil - 1$ must be true. We have that $k > \lceil \psi_1^\ell \rceil - 1$ if $k \geq \psi_1^\ell$, which yields that $k \geq \frac{N(b-d)}{a+b-2d} + \frac{(a-c)}{a+b-2d}$. This contradicts $k \in \left[\frac{N(b-d)}{a+b-2d}, \frac{N(b-d)}{a+b-2d} + \frac{a-c}{a+b-2d}\right)$, proving the claim. It follows that $m^{BA} \geq k$. Now assume that we have exactly k A-players. Attending the LOP we see that a B-player will switch to A if $v(A, k+1) \geq v(B, k)$, which yields $k \geq \frac{(N-1)(b-d)}{a+b-2d}$. Hence, for all $k \geq \frac{(N-1)(b-d)}{a+b-2d}$ indeed k mutations are sufficient for a transition. Thus, since $m^{BA} \geq k$ and m^{BA} is the minimal number of players involved in a transition, we conclude that $m^{BA} = k$. Summing up, provided that $m^{BA} > N - k - 1$, we find that

$$m^{BA} = \begin{cases} \lceil \psi_2^\ell \rceil - 1 & \text{if } k < \frac{N(b-d)}{a+b-2d} \\ k & \text{if } \frac{N(b-d)}{a+b-2d} \leq k < \frac{N(b-d)}{a+b-2d} + \frac{a-c}{a+b-2d} \\ \lceil \psi_1^\ell \rceil - 1 & \text{if } k \geq \frac{N(b-d)}{a+b-2d} + \frac{a-c}{a+b-2d} \end{cases}.$$

Now observe that $\lceil \psi_2^\ell \rceil - 1 = k$ if $k < \psi_2^\ell \leq k+1$ which holds for $k \in \left[\frac{(N-1)(b-d)}{a+b-2d}, \frac{N(b-d)}{a+b-2d}\right)$. Thus, we have

$$m^{BA} = \begin{cases} \lceil \psi_2^\ell \rceil - 1 & \text{if } k < \frac{(N-1)(b-d)}{a+b-2d} \\ k & \text{if } \frac{(N-1)(b-d)}{a+b-2d} \leq k < \frac{N(b-d)}{a+b-2d} + \frac{a-c}{a+b-2d} \\ \lceil \psi_1^\ell \rceil - 1 & \text{if } k \geq \frac{N(b-d)}{a+b-2d} + \frac{a-c}{a+b-2d} \end{cases}.$$

Further, note that $\lceil \psi_1^\ell \rceil - 1 = k$ if $k < \psi_1^\ell \leq k+1$ which translates to $k \in \left[\frac{(N-1)(b-d)}{a+b-2d}, \frac{N(b-d)}{a+b-2d} + \frac{a-c}{a+b-2d}\right)$. Hence, for k in this range we have $k = \lceil \psi_1^\ell \rceil - 1$. Thus,

$$m^{BA} = \begin{cases} \lceil \psi_2^\ell \rceil - 1 & \text{if } k < \frac{(N-1)(b-d)}{a+b-2d} \\ \lceil \psi_1^\ell \rceil - 1 & \text{if } k \geq \frac{(N-1)(b-d)}{a+b-2d} \end{cases}.$$

Finally, we see that $\lceil \psi_2^\ell \rceil - 1 > N - k - 1$ and that $\lceil \psi_1^\ell \rceil - 1 > N - k - 1$ whenever $k > \frac{(N-1)(a-c)}{a+b-2c}$. Note that as $\frac{(N-1)(a-c)}{a+b-2c} < \frac{(N-1)(b-d)}{a+b-2d}$[24] this is always the the range where $k \geq \frac{(N-1)(b-d)}{a+b-2d}$. It follows that m^{BA} is indeed the minimal number of A-players required for the transition from $\overrightarrow{B}[k]$ to $\overrightarrow{A}[k]$, establishing $CR(\overrightarrow{A}[k]) = R(\overrightarrow{B}[k]) = m^{BA}$.

B.1.2 Identifying the LRE

We will now identify the set of LRE for the ranges of k identified above. We start with the case where $k < \frac{(N-1)(b-d)}{a+b-2d}$. Note that we have $R(\overrightarrow{B}[k]) = \lceil \psi_2^\ell \rceil - 1$ and $CR(\overrightarrow{B}[k]) = N - \lfloor \psi_2^\ell \rfloor$. If $\psi_2^\ell \in \mathbb{Z}$ we have that $CR(\overrightarrow{B}[k]) = N - (N - \frac{a-d}{b-d}k) = \frac{a-d}{b-d}k$. As $b > a$ it follows that $CR(\overrightarrow{B}[k]) < k$. Further, it is easy to verify that in this case $R(B[k]) > k$. Hence, $R(\overrightarrow{B}[k]) > k > CR(\overrightarrow{B}[k])$, which implies by Lemma 4 that $S = \overrightarrow{B}[k]$ if $\psi_2^\ell \in \mathbb{Z}$. Likewise, consider the case $\psi_2^\ell \notin \mathbb{Z}$. We know that $CR(\overrightarrow{B}[k]) = N - \lfloor N - \frac{a-d}{b-d}k \rfloor = N - (N - \lceil \frac{a-d}{b-d}k \rceil) = \lceil \frac{a-d}{b-d}k \rceil$. As $b > a$, it follows that $CR(\overrightarrow{B}[k]) \leq k$. Now, consider $R(\overrightarrow{B}[k]) = \lfloor \psi_2^\ell \rfloor = \lceil \psi_2^\ell \rceil - 1$. We have $R(\overrightarrow{B}[k]) > k$, whenever $\lceil \psi_2^\ell \rceil - 1 > k$ which is the case whenever $\psi_2^\ell - 1 > k$. As above, $k < \frac{(N-1)(b-d)}{a+b-2d}$ implies $\psi_2^\ell - 1 > k$. Thus, if $k < \frac{(N-1)(b-d)}{a+b-2d}$ we have that $R(\overrightarrow{B}[k]) > k \geq CR(\overrightarrow{B}[k])$, establishing that also $S = \overrightarrow{B}[k]$ if $\psi_2^\ell \notin \mathbb{Z}$.

Now consider the case where $k = \frac{(N-1)(b-d)}{a+b-2d}$. First, note that in this case $\psi_1^\ell - 1 = \psi_2^\ell - 1 = k$. Thus, $\psi_1^\ell, \psi_2^\ell \in \mathbb{Z}$, and we see that $R(\overrightarrow{B}[k]) = k$ and $CR(\overrightarrow{B}[k]) = N - k - 1$. Consequently, $R(\overrightarrow{B}[k]) > CR(\overrightarrow{B}[k])$ whenever $k = \frac{(N-1)(b-d)}{a+b-2d} > \frac{N-1}{2}$. Since $b > a$ this inequality always holds which shows that $S = \overrightarrow{B}[k]$ in this case.

Finally we consider the case where $k > \frac{(N-1)(b-d)}{a+b-2d}$. Now, we have that $R(\overrightarrow{A}[k]) = CR(\overrightarrow{B}[k]) = N - \lfloor \psi_1^\ell \rfloor$ and $CR(\overrightarrow{A}[k]) = R(\overrightarrow{B}[k]) = \lceil \psi_1^\ell \rceil - 1$. Here, we have to distinguish two cases, $\psi_1^\ell \in \mathbb{Z}$ and $\psi_1^\ell \notin \mathbb{Z}$.

Case 1: $\psi_1^\ell \in \mathbb{Z}$
We have $R(\overrightarrow{A}[k]) = CR(\overrightarrow{B}[k]) = N - \psi_1^\ell$ and $CR(\overrightarrow{A}[k]) = R(\overrightarrow{B}[k]) = \psi_1^\ell - 1$. Therefore, $S = \overrightarrow{B}[k]$ whenever $\psi_1^\ell > \frac{N+1}{2}$, which translates into

$$k < \frac{N-1}{2}\left(\frac{b-a}{c-d} + 1\right) \equiv k^\ell.$$

[24] This inequality can be rewritten as $(b-a)(a+b-c-d) > 0$ and, thus, holds for our parameters.

Likewise, we see that $S = \overrightarrow{A}[k]$ if $\psi_1^\ell < \frac{N+1}{2}$ which is the case if $k > k^\ell$. If $k = k^\ell$ we have that $R(\overrightarrow{A}[k]) = CR(\overrightarrow{A}[k]) = R(\overrightarrow{B}[k]) = CR(\overrightarrow{B}[k])$ and, thus, $S = \overrightarrow{A}[k] \cup \overrightarrow{B}[k]$.

Finally, let us check whether the local threshold identified for $k > \frac{(N-1)(b-d)}{a+b-2d}$ is indeed also a global thresholds. Recall that for $k \leq \frac{(N-1)(b-d)}{a+b-2d}$ we have that $S = \overrightarrow{B}[k]$. Now note that $k^\ell > \frac{(N-1)(b-d)}{a+b-2d}$.[25] Thus, we have (for the entire range of k) that $S = \overrightarrow{B}[k]$ if $k < k^\ell$. We can summarize our results for $\psi_1^\ell \in \mathbb{Z}$:

$$S = \begin{cases} \overrightarrow{B}[k] & \text{if } k < k^\ell, \\ \overrightarrow{A}[k] \cup \overrightarrow{B}[k] & \text{if } k = k^\ell, \\ \overrightarrow{A}[k] & \text{if } k > k^\ell. \end{cases}$$

Thus, we have identified our two thresholds in the theorem as $\underline{k}^\ell = \bar{k}^\ell = k^\ell$.

Case 2: $\psi_1^\ell \notin \mathbb{Z}$.
Now, consider the case where $\psi_1^\ell \notin \mathbb{Z}$. Recall that $R(\overrightarrow{A}[k]) = CR(\overrightarrow{B}[k]) = N - \lfloor \psi_1^\ell \rfloor = N - \lceil \psi_1^\ell \rceil + 1$. Consequently, $S = \overrightarrow{A}[k]$ if $\lceil \psi_1^\ell \rceil < \frac{N}{2} + 1$, $S = \overrightarrow{B}[k]$ if $\lceil \psi_1^\ell \rceil > \frac{N}{2} + 1$, and, $S = \overrightarrow{A}[k] \cup \overrightarrow{B}[k]$ if $\lceil \psi_1^\ell \rceil = \frac{N}{2} + 1$. As we are rounding up we now need to distinguish two subcases: i) N is odd and ii) N is even:

Subcase 2a: $\psi_1^\ell \notin \mathbb{Z}$ and N odd.
One sees that $\lceil \psi_1^\ell \rceil < \frac{N}{2} + 1$ if $\psi_1^\ell \leq \frac{N+1}{2}$.[26] Thus, $S = \overrightarrow{A}[k]$ for $k \geq k^\ell$. Likewise, we have that $\lceil \psi_1^\ell \rceil > \frac{N}{2} + 1$ if $\psi_1^\ell > \frac{N+1}{2}$.[27] Hence, $S = \overrightarrow{B}[k]$ for $k < k^\ell$.

Again, we can check that our local thresholds identified for $k > \frac{(N-1)(b-d)}{a+b-2d}$ is indeed also a global threshold. We have previously established that $k^\ell > \frac{(N-1)(b-d)}{a+b-2d}$. Thus, we globally have that $S = \overrightarrow{B}[k]$ if $k \leq k^\ell$. We can summarize our results for the case $\psi_1^\ell \notin \mathbb{Z}$ and N odd as follows:

$$S = \begin{cases} \overrightarrow{B}[k] & \text{if } k < k^\ell, \\ \overrightarrow{A}[k] & \text{if } k \geq k^\ell, \end{cases}$$

Now note that k^ℓ is the value of k that solves $\psi_1^\ell = \frac{N+1}{2}$. Thus, if N is odd we have $\psi_1^\ell \in \mathbb{Z}$. It follows that the case $k = k^\ell$ can not occur if $\psi_1^\ell \notin \mathbb{Z}$. Thus, the two thresholds in the statement of the theorem are given by $\underline{k}^\ell = \bar{k}^\ell = k^\ell$.

Subcase 2b: $\psi_1^\ell \notin \mathbb{Z}$ and N even.
We have that $\lceil \psi_1^\ell \rceil < \frac{N}{2} + 1$ if $\psi_1^\ell \leq \frac{N}{2}$. Thus, $S = \overrightarrow{A}[k]$ if $k \geq \tilde{k}^\ell$, where

$$\tilde{k}^\ell \equiv k^\ell + \frac{a+b-c-d}{2(c-d)}.$$

Similarly, we have that $\lceil \psi_1^\ell \rceil > \frac{N}{2} + 1$ if $\psi_1^\ell > \frac{N}{2} + 1$. Hence, $S = \overrightarrow{B}[k]$ if $k < k^\ell - \frac{a+b-c-d}{2(c-d)}$. Note, however, that at $k = k^\ell - \frac{a+b-c-d}{2(c-d)}$ we have that $\psi_1^\ell = \frac{N}{2} + 1$, which is impossible by our hypothesis that $\psi_1^\ell \notin \mathbb{Z}$. For the remaining interval $(k^\ell - \frac{a+b-c-d}{2(c-d)}, \tilde{k}^\ell)$ we have $\lceil \psi_1^\ell \rceil = \frac{N}{2} + 1$ and and, thus, $S = \overrightarrow{A}[k] \cup \overrightarrow{B}[k]$. Again, we can check that our local thresholds identified for $k > \frac{(N-1)(b-d)}{a+b-2d}$ are indeed also global threshold. We have previously established that $k^\ell > \frac{(N-1)(b-d)}{a+b-2d}$. Thus, also $\tilde{k}^\ell = k^\ell + \frac{a+b-c-d}{2(c-d)} > \frac{(N-1)(b-d)}{a+b-2d}$. We have that $k^\ell - \frac{a+b-c-d}{2(c-d)} > \frac{(N-1)(b-d)}{a+b-2d}$ if $N > 2 + \frac{2(a-d)}{b-a}$ and we have that $k^\ell - \frac{a+b-c-d}{2(c-d)} \leq \frac{(N-1)(b-d)}{a+b-2d}$ otherwise. Thus, it follows that globally we have $S = \overrightarrow{B}[k]$ if $k \leq \hat{k}^\ell$ where

$$\hat{k}^\ell \equiv \max\left\{k^\ell - \frac{a+b-c-d}{2(c-d)}, \frac{(N-1)(b-d)}{a+b-2d}\right\}.$$

[25] This inequality can be rewritten as $b(a+b-c-d) > a(a+b-c-d)$ and, hence, holds for our parameters.

[26] Note that if $\psi_1^\ell \leq \frac{N+1}{2}$ we have, since $\frac{N+1}{2} \in \mathbb{Z}$, that $\lceil \psi_1^\ell \rceil \leq \frac{N+1}{2}$ and, consequently, that $\lceil \psi_1^\ell \rceil < \frac{N}{2} + 1$.

[27] Note that $\frac{N+1}{2}$ is the largest integer smaller than $\frac{N}{2} + 1$ if N is odd. Since $\psi_1^\ell \notin \mathbb{Z}$ by hypothesis, it follows that whenever $\psi^\ell > \frac{N+1}{2}$ we get the desired inequality.

On the contrary, if $\underline{k}^\ell < k < \bar{k}^\ell$ we have $S = \overrightarrow{A}[k] \cup \overrightarrow{B}[k]$. Summarizing, the case where $\psi_1^\ell \notin \mathbb{Z}$ and N even, we have

$$S = \begin{cases} \overrightarrow{B}[k] & \text{if } k \leq \hat{k}^\ell \\ \overrightarrow{A}[k] \cup \overrightarrow{B}[k] & \text{if } \hat{k}^\ell < k < \tilde{k}^\ell \\ \overrightarrow{A}[k] & \text{if } k \geq \tilde{k}^\ell \end{cases} .$$

Finally, we remark that if $k = \tilde{k}^\ell$ we have that $\psi_1^\ell = \frac{N}{2}$ and therefore $\psi_1^\ell \in \mathbb{Z}$. Thus, this case can not occur for $\psi_1^\ell \notin \mathbb{Z}$. Now, consider $k = \hat{k}^\ell$. Note that $\hat{k}^\ell$ is either $k^\ell - \frac{a+b-c-d}{2(c-d)}$ or $\frac{(N-1)(b-d)}{a+b-2d}$. In either case we have $\psi_1^\ell \in \mathbb{Z}$. Thus, $k = \hat{k}^\ell$ can not occur either. It follows that the thresholds in the theorem are given by $\underline{k}^\ell = \hat{k}^\ell$ and $\bar{k}^\ell = \tilde{k}^\ell$.

Finally, note that in all subcases we have $\hat{k}^\ell \geq \underline{k}^\ell > \frac{(N-1)(b-d)}{a+b-2d} > \frac{N-1}{2}$. Thus, whenever $k \leq \frac{N-1}{2}$ we have $S = \overrightarrow{B}[k]$.

B.2 Intermediate costs

The proof of the intermediate cost case, $d \leq \gamma < c$, follows exactly the same steps as the low cost case and is omitted. We remark that the only difference between ψ_1^m and ψ_1^ℓ and between ψ_2^m and ψ_2^ℓ is that d is replaced by γ. Thus, also in the relevant thresholds for the intermediate cost case, k^m,$\underline{k}^m$, and $\bar{k}^m$, d is replaced by γ.

B.3 High costs

Finally, consider the high cost case, $c \leq \gamma \leq a$. Here, we can note that the only difference between ψ_1^h and ψ_1^ℓ and between ψ_2^h and ψ_2^ℓ is that c and d must be replaced by γ. In particular, we find that

$$m^{AB} = \begin{cases} \lfloor \psi_2^h \rfloor & \text{if } k \leq \frac{(N-1)(b-\gamma)}{a+b-2\gamma}, \\ \lfloor \psi_1^h \rfloor & \text{if } k > \frac{(N-1)(b-\gamma)}{a+b-2\gamma}, \end{cases} \quad \text{and} \quad m^{BA} = \begin{cases} \lceil \psi_2^h \rceil - 1 & \text{if } k < \frac{(N-1)(b-\gamma)}{a+b-2\gamma}, \\ \lceil \psi_1^h \rceil - 1 & \text{if } k \geq \frac{(N-1)(b-\gamma)}{a+b-2\gamma}. \end{cases}$$

This gives us that $R(\overrightarrow{A}[k]) = CR(\overrightarrow{B}[k]) = N - m^{AB}$ and $CR(\overrightarrow{A}[k]) = R(\overrightarrow{B}[k]) = m^{BA}$.

As before, we find that $S = \overrightarrow{B}[k]$ if $k \leq \frac{(N-1)(b-\gamma)}{a+b-2\gamma}$. Now, consider the case where $k > \frac{(N-1)(b-\gamma)}{a+b-2\gamma}$. Here, we have that $R(\overrightarrow{A}[k]) = CR(\overrightarrow{B}[k]) = N - \lfloor \psi_1^h \rfloor$ and $CR(\overrightarrow{A}[k]) = R(\overrightarrow{B}[k]) = \lceil \psi_1^h \rceil - 1$. First, consider the case where $\psi_1^h \in \mathbb{Z}$. Here, we have that $R(\overrightarrow{B}[k]) > CR(\overrightarrow{B}[k])$ if $\psi_1^h > \frac{N+1}{2}$. It is straightforward to check that, since $b > a$, this inequality always holds, and, thus, we have $S = \overrightarrow{B}[k]$ in this case. Thus, we have $k^h = N - 1$.

A similar argument establishes that we also have $S = \overrightarrow{B}[k]$ if $\psi_1^h \notin \mathbb{Z}$ and N is odd and that we also have $k^h = N - 1$ in this case.

Now, consider the case where $\psi_1^h \notin \mathbb{Z}$ and N is even. Here, we have $R(\overrightarrow{B}[k]) > CR(\overrightarrow{B}[k])$ if $\lceil \psi_1^h \rceil > \frac{N}{2} + 1$ which, in turn, translates into $N > 2 + \frac{2(a-\gamma)}{b-a}$. Thus, $S = \overrightarrow{B}[k]$ if $N > 2 + \frac{2(a-\gamma)}{b-a}$. Thus, in this case we have $k^h = N - 1$.

If, however, $N \leq 2 + \frac{2(a-\gamma)}{b-a}$ we have $\lceil \psi_1^h \rceil = \frac{N}{2} + 1$ and, thus, $S = \overrightarrow{A}[k] \cup \overrightarrow{B}[k]$ (provided that $k > \frac{(N-1)(b-\gamma)}{a+b-2\gamma}$). In this case we have that $k^h = \frac{(N-1)(b-\gamma)}{a+b-2\gamma}$.

Proof of Proposition 2

Let us start with part (i). For $0 \leq \gamma \leq \underline{u}$ and $k = N - 1$ we have that each agent (irrespective of his own action) will link up to all other agents in the population. Thus, we obtain a model of global interactions. It is well known (see e.g. Ellison (2000)) that under these premises $S = \overrightarrow{a}[N - 1]$ for N sufficiently large.

Now consider part (ii). Now $0 \leq \gamma \leq \underline{u}$ implies that each agent will form all of his k links. Since, (a, a) is a strict Nash equilibrium the set $\overrightarrow{a}[k]$ is absorbing. Further, as a is payoff dominant, we know that once we have k agents adopting it all other agents will switch to a. Hence, $CR(\overrightarrow{a}[k]) \leq k$. Conversely, note that if there are more (or exactly) $k+1$ a-players in the population we will for sure move back to $\overrightarrow{a}[k]$. Thus, to leave the basin of attraction of $\overrightarrow{a}[k]$ we need more than $N-k-1$ players to mutate to some other action, establishing $R(\overrightarrow{a}[k]) \geq N-1-k$. Thus, $S = \overrightarrow{a}[k]$ if $k \leq \frac{N-1}{2}$.

Proof of Proposition 3

The proof of part (i) follows by the observation that if $\phi(N-1) - \phi(N-2) \leq d$ all links will be formed and we have a model of global interaction where the risk dominant action is selected in a sufficiently large population.

Consider now the second part of the Proposition. Consider the set $S = \overrightarrow{B}[d^*_{B|B}(0)]$ where all agents play B and support $d^*_{B|B}(0)$ links. Note that these states form an absorbing set since (B, B) is a Nash equilibrium. Consider any state $\omega \notin D\left(\overrightarrow{B}[d^*_{B|B}(0)]\right)$. Note that if there are $d^*_{B|B}(0)$ B-players in the population all A-players will want to link up to those players and switch to the B. To see this more formally note that $d^*_{B|B}(0) \geq d^*_{A|A}(m) + d^*_{A|B}(m)$ for all $m \in \{0, 1, \ldots N\}$. Thus, since $b > a > c$, we have that a B-player receives a higher LOP than an A-player:

$$bd^*_{B|B}(0) > ad^*_{A|A}(m) + d^*_{A|B}(m).$$

Thus, all A-players will switch to B. This implies that with at most $d^*_{B|B}(m)$ mutations we can reach the state $\overrightarrow{B}\left[d^*_{B|B}(0)\right]$. Thus, $CR\left(\overrightarrow{B}[d^*_{B|B}(0)]\right) \leq d^*_{B|B}(0) < \frac{N-1}{2}$. Conversely, to leave the basin of attraction of of $\overrightarrow{B}[d^*_{B|B}(0)]$ we need to make sure that there are less than $d^*_{B|B}(0) + 1$ B-agents in the population. (Otherwise, the current B-agents will remain at B and the A agents will switch to B.) Thus, we need more than $N - d^*_{B|B}(0) - 1$ agents to mutate to A, establishing $R\left(\overrightarrow{B}[d^*_{B|B}(0)]\right) > N - d^*_{B|B}(0) - 1$. Thus, $S = \overrightarrow{B}[d^*_{B|B}(0)]$ if $d^*_{B|B}(0) \leq \frac{N-1}{2}$, which is the case if $\phi(\frac{N-1}{2}) - \phi(\frac{N+1}{2}) > b$.

References

Alós-Ferrer, C., and S. Weidenholzer (2007): "Partial Bandwagon Effects and Local Interactions," *Games and Economic Behavior*, 61, 1–19.

——— (2008): "Contagion and Efficiency," *Journal of Economic Theory*, 143, 251 – 274.

Anwar, A. W. (2002): "On the Co-existence of Conventions," *Journal of Economic Theory*, 107, 145–155.

Bala, V., and S. Goyal (2000): "A Noncooperative Model of Network Formation," *Econometrica*, 68, 1181 – 1229.

Bhaskar, V., and F. Vega-Redondo (2004): "Migration and the Evolution of Conventions," *Journal of Economic Behaviour and Organization*, 13, 397–418.

Blume, L. (1993): "The Statistical Mechanics of Strategic Interaction," *Games and Economic Behavior*, 5, 387–424.

——— (1995): "The Statistical Mechanics of Best-Response Strategy Revision," *Games and Economic Behavior*, 11, 111–145.

Corten, R., and V. Buskens (2010): "Co-evolution of conventions and networks: An experimental study," *Social Networks*, 32(1), 4–15.

Dieckmann, T. (1999): "The Evolution of Conventions with Mobile Players," *Journal of Economic Behaviour and Organization*, 38, 93–111.

Ellison, G. (1993): "Learning, Local Interaction, and Coordination," *Econometrica*, 61, 1047–1071.

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

Ely, J. C. (2002): "Local Conventions," *Advances in Theoretical Economics*, 2, 1–30.

Eshel, I., L. Samuelson, and A. Shaked (1998): "Altruists, Egoists, and Hooligans in a Local Interaction Model," *The American Economic Review*, 88, 157–179.

Freidlin, M., and A. Wentzell (1988): *Random Perturbations of Dynamical Systems, 2nd Ed.* Springer Verlag, New York.

Galeotti, A., and S. Goyal (2010): "The Law of the Few," *American Economic Review*, 100(4), 1468–92.

Goyal, S., and F. Vega-Redondo (2005): "Network Formation and Social Coordination," *Games and Economic Behavior*, 50, 178–207.

Harsanyi, J., and R. Selten (1988): *A General Theory of Equilibrium Selection in Games*. The MIT Press, Cambridge, MA.

Hellmann, T. (2007): "Social Coordination and Endogenous Network Decisions with Limited Contacts," *Master's thesis, Institute for Mathematical Economics and Department of Mathematics, University Bielefeld.*

Hojman, D., and A. Szeidl (2006): "Endogenous Networks, Social Games, and Evolution," *Games and Economic Behavior*, 55, 112–130.

Jackson, M. O., and A. Watts (2002): "On the Formation of Interaction Networks in Social Coordination Games," *Games and Economic Behavior*, 41, 265–291.

Jackson, M. O., and A. Watts (2010): "Social Games: Matching and the Play of Finnitely Repeated Games," *Games and Economic Behavior*, 70, 170–191.

Jackson, M. O., and A. Wolinsky (1996): "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71, 44–74.

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

Kandori, M., and R. Rob (1995): "Evolution of Equilibria in the Long Run: A General Theory and Applications," *Journal of Economic Theory*, 65, 383–414.

Kwak, H., C. Lee, H. Park, and S. Moon (2010): "What is Twitter, a Social Network or a News Media?," *Proceedings of the 19th international conference on World wide web, April 26-30, Raleigh, North Carolina, USA*.

Lucier, B., B. W. Rogers, and N. Immorlica (2011): "Cooperation in Anonymous Dynamic Social Networks," Mimeo.

Morris, S., R. Rob, and H. Shin (1995): "p-dominance and Belief Potential," *Econometrica*, 63, 145–157.

Oechssler, J. (1997): "Decentralization and the Coordination Problem," *Journal of Economic Behaviour and Organization*, 32, 119–135.

Sandholm, W. H. (1998): "Simple and Clever Decision Rules for a Model of Evolution," *Economics Letters*, 61, 165–170.

Staudigl, M. (2011): "Potential games in volatile Environments," *Games and Economic Behavior*, 72, 271–287.

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

Appendix C Not intended for publication

C.1 Derivation of the switching thresholds reported in Table 1

We only provide the computations for the switching thresholds of A-players. The switching thresholds of B-players can be computed analogously.

Case 1 First, consider the case of low linking costs, $0 \leq \gamma \leq d$. An A-player will switch to B with positive probability if

$$a \min\{k, m-1\} + c(k - \min\{k, m-1\}) \leq b \min\{k, N-m\} + d(k - \min\{k, N-m\})$$

Depending on the relationship between N, n, and k we obtain four subcases.

(1i) If $k > m-1$ and $k > N-m$ neither A- nor B- players may fill up all their slots with other agents of their own kind. An A-player will switch to B with positive probability if

$$a(m-1) + c(k-m+1) \leq b(N-m) + d(k-N+m),$$

i.e. if

$$m \leq \frac{(N-1)(b-d) - k(c-d)}{a+b-c-d} + 1 := \psi_1^\ell.$$

(1ii) If $k > m-1$ and $k \leq N-m$, A-players do not find sufficiently many other A-players to fill up all their slots, whereas B-players can fill up all their slots with other B-players. As b is the highest payoff in the base game, B-players will always earn the highest payoff whenever they may fill up all their slots, and so A-players always switch to B.[28]

(1iii) If $k \leq m-1$ and $k > N-m$, A-players will link only to other A-players whereas B-players can not fill up all their slots with agents of their own kind. An A-player may switch to B whenever

$$ak \leq b(N-m) + d(k-N+m)$$

i.e. if

$$m \leq N - k\frac{a-d}{b-d} := \psi_2^\ell.$$

(1iv) In the remaining case with $k \leq m-1$ and $k \leq N-m$ both A- and B- players will link up only to agents of their own kind. Here we find that A-players always have an incentive to switch to B.

Case 2 For intermediate linking costs, $d \leq \gamma \leq c$, B-players will no longer interact with A-players, but A-players will still interact with B-players. Consequently, an A-player switches to B with positive probability if

$$a \min\{k, m-1\} + c(k - \min\{k, m-1\}) - \gamma k \leq (b-\gamma) \min\{k, N-m\}$$

Let us again consider our four subcases:

(2i) When $k > m-1$ and $k > N-m$ an A-player will switch to B with positive probability if

$$m \leq \frac{(N-1)(b-\gamma) - k(c-\gamma)}{a+b-c-\gamma} + 1 := \psi_1^m.$$

(2ii) If $k > m-1$ and $k \leq N-m$ there are sufficiently many B-players so that choosing B always gives the highest payoff.

[28] Formally, $v(A, m) = a(m-1) + c(k-m+1) - \gamma k$ and $v(B, m) = (b-\gamma)k$. A-players will switch to B iff $b > a\frac{m-1}{k} + (1 - \frac{m-1}{k})c$. On the right-hand side we have a convex combination in $[c, a]$. Since $b > a > c$ the claim follows.

(2iii) Whenever $k \leq m-1$ and $k > N-m$, A-player will switch to B with positive probability if

$$m \leq N - \frac{a-\gamma}{b-\gamma}k := \psi_2^m.$$

(2iv) If $k \leq m-1$ and $k \leq N-m$ then A-players as well as B-players can completely isolate. Since the B-players earn always a higher payoff, all A-players will switch to B.

Case 3 Finally, we consider the case of high linking costs, $c \leq \gamma \leq a$. In this case any interaction between groups of agents choosing different actions is completely shut down. In this scenario, an A-player will switch to with positive probability B whenever

$$(a-\gamma)\min\{k, m-1\} \leq (b-\gamma)\min\{k, N-m\}.$$

(3i) If $k > m-1$ and $k > N-m$ we find that an A-player will switch to B with positive probability if

$$m \leq \frac{(N-1)(b-\gamma)}{a+b-2\gamma} + 1 := \psi_1^h.$$

(3ii) If $k > m-1$ and $k \leq N-m$ the same applies as in cases (1ii) or (2ii).

(3iii) If $k \leq m-1$ and $k > N-m$ we have that A-players will switch to B with positive probability whenever

$$m \leq N - \frac{a-\gamma}{b-\gamma}k := \psi_2^h.$$

(3iv) If $k \leq m-1$ and $k \leq N-m$ the same applies as in cases (1iv) or (2iv).