

# Network Formation and Social Coordination

Sanjeev Goyal\*      Fernando Vega-Redondo†

First Version: November 1999.

This version: December 2003.

---

\*Department of Economics, University of Essex & Econometric Institute, Erasmus University Rotterdam. E-mail address: sgoyal@essex.ac.uk

†Departamento de Fundamentos del Análisis Económico and Instituto Valenciano de Investigaciones Económicas, Universidad de Alicante, 03071 Alicante, Spain & Departament de Economia i Empresa, Universitat Pompeu Fabra, 08005 Barcelona, Spain. E-mail address: vega@merlin.fae.ua.es.

An earlier version of this paper, dating from November 1999, was circulated under the title, 'Learning, network formation and coordination'. We would like to thank the associate editor for useful comments. We are also indebted to Yann Bramoulle, Bhaskar Dutta, Matt Jackson, Massimo Morelli, Ben Polak, Karl Schlag, Mark Stegeman, Steve Tadelis, and participants in presentations at Alicante, Amsterdam, Center(Tilburg), ISI (Delhi), Marseille, The Roy Seminar (Paris), Stanford, Toulouse, Yale, Econometric Society World Congress (Seattle), and Game Theory World Congress (Bilbao). Vega-Redondo acknowledges support by the Spanish Ministry of Education, CICYT Project no. 970131.

## Abstract

This paper develops a simple model to examine the interaction between partner choice and individual behavior in games of coordination. An important ingredient of our approach is the way we model partner choice: we suppose that a player can establish ties with other players by unilaterally investing in costly pair-wise links. In this context, individual efforts to balance the costs and benefits of links are shown to lead to a unique equilibrium interaction architecture. The dynamics of network formation, however, has powerful effects on individual behavior: if costs of forming links are below a certain threshold then players coordinate on the risk-dominant action, while if costs are above this threshold then they coordinate on the efficient action. These findings are robust to modifications in the link formation process, different specifications of link formation costs, alternative models of mutations as well as the possibility of interaction among indirectly connected players.

**Keywords:** networks, links, coordination games, equilibrium selection, risk dominance, efficiency

# 1 Introduction

In recent years, several authors have examined the role of interaction structure – different terms like network structure, neighborhood influences, and peer group pressures, have been used – in explaining a wide range of social and economic phenomena. This includes work on social learning and adoption of new technologies, evolution of conventions, collective action, labor markets, and financial fragility.<sup>1</sup> The research suggests that the structure of interaction can be decisive in determining the nature of outcomes. This leads us to examine the reasonableness/robustness of alternative structures within a model in which the social network is itself an object of study and *co*-evolves with the other dimensions of agents' choice.

More specifically, in the present paper we apply this approach to the following problem: the influence of link formation on individual behavior in games of coordination.<sup>2</sup> There is a group of players, who have the opportunity to play a  $2 \times 2$  coordination game with each other. This game has two pure-strategy Nash equilibria, one of them Pareto efficient (but risk-dominated) and the other risk dominant (but inefficient). Two players can only play with one another if they have 'link' between them. These links are made on individual initiative. They are also costly to form, in the sense that it takes effort and resources to create and maintain them. A link permits several interpretations; examples include communication links (with messages sent from one person to another), investments of time and effort by two persons in building a common understanding of a research problem, or travel by one person to the location of another to carry out some joint project.

The link decisions of different players define a network of social interaction. In addition to the choice of links, each player has to select an action that she must use in all the games that she engages in. Thus, given the incentives of individuals to form (or destroy) their links, twin processes of link and action adjustment unfold that *jointly* determine the social outcome. We are interested in the nature of networks that emerge and the effects of link formation on social coordination. We mostly focus on a setting where links as well as actions in the coordination game are chosen by individuals on an independent basis. (The idea that links can be one-sided is closer in spirit to the first and third examples of links given above.) This approach of one-sided links allows us to explore the implications of link formation for social coordination as part of a non-cooperative game, which facilitates the exposition greatly.<sup>3</sup>

---

<sup>1</sup>See e.g., Allen and Gale (1998), Bala and Goyal (1998), Chwe (2000), Coleman (1966), Ellison and Fudenberg (1993), Ellison (1993), Granovetter (1974), Haag and Lagunoff (1999), and Morris (2000), among others.

<sup>2</sup>Many games of interest have multiple equilibria. The study of equilibrium selection (which manifests itself most sharply in coordination games) therefore occupies a central place in game theory. We discuss the contribution of our paper to this research in greater detail below.

<sup>3</sup>In a subsequent section we elaborate on alternative formulations of link formation and argue that the main insights are robust to a variety of modifications (see Section 4).

We start with a consideration of the static problem. Here we find that a variety of networks – including the complete network, the empty network and partially connected networks – can be supported at Nash equilibria of the static (strategic-form) game induced. Moreover, the society can coordinate on different actions and conformism as well as diversity with regard to actions of individuals is possible at equilibrium. The immediate counterpart of this multiplicity of Nash equilibria is that any (best-response) learning dynamics must also have multiple rest points, which in turn motivates an examination of the stochastic stability of different outcomes.

To this end, we propose a dynamic model in which, at regular intervals, individuals choose links and actions to maximize (myopically) their respective payoffs. Occasionally, they also make errors or experiment. Our interest is in the nature of long run outcomes, when the probability of these errors is small. This leads to clear-cut predictions, both concerning the architecture of networks as well as the nature of social coordination.

First, we show that, provided the costs of link formation are not too high,<sup>4</sup> any network architecture that is robust enough to be observed a significant fraction of time in the long run (i.e. occurs at so-called stochastically stable states) must be complete.<sup>5</sup> (Figure 1a gives an example of a complete network in a society with 4 players, where a filled circle lying on the edge near a player indicates that this player has formed, or supports, that link.) This implies that partially connected networks, even if they define Nash equilibria of the social game (and thus rest-points of the perturbation-free best-response dynamics), are just ephemeral situations in the long run.

Secondly, we also find that in the long-run states (where the social network is complete), players always coordinate on the same action, i.e. social conformism obtains. However, the specific nature of coordination sharply depends on the costs of link formation. There is a threshold value in the interior of the payoff range such that, if the costs of link formation are below the threshold, players coordinate on the risk-dominant action. In contrast, if those costs are above that threshold, players coordinate on the efficient action at all stochastically stable states. This is the content of our main result, Theorem 3.1. In sum, therefore, our analysis reveals that, even though the eventual architecture of the social network is the same (i.e. complete) in all “robust” cases, the *process* of network formation (i.e. the dynamics by which links are created and destroyed *out of equilibrium*) has crucial implications for the nature of social coordination. Specifically, it leads to very different conclusions concerning the strategy choice selected in the long run, as the magnitude of linking costs changes. We elaborate on some aspects of these results and sketch the intuition underlying them.

---

<sup>4</sup>Of course, if the linking cost is higher than the maximum payoff in the coordination game, only the empty network can prevail.

<sup>5</sup>In a complete network, every pair of players is directly linked.

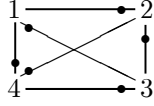


Figure 1a  
Complete Network

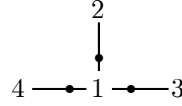


Figure 1b  
Center-sponsored Star

*First*, we stress that the *dynamics* of link formation play a crucial role in the model. Despite the fact that the only architecture that is stochastically stable (within the interesting parameter range) is the complete one, players' behavior in the coordination game is *different* depending on the costs of *forming* links. But if the network were to remain fixed throughout, standard arguments indicate that the risk-dominant action must prevail in the long run (cf. Kandori, Mailath and Rob, 1993). This serves to highlight the fact that, indeed, it is the link formation *process* that, by allowing for the *co*-evolution of the links and actions, plays a decisive role in shaping individual behavior in our model.

*Second*, we want to develop some intuition on the sharp relationship found between the *costs* of forming links and the corresponding behavior displayed by players in the coordination game. On the one hand, note the obvious fact that, if the cost of forming links is small and the gross payoffs to be earned in the game are positive,<sup>6</sup> players wish to be linked with everyone irrespective of the actions they choose. Hence, from an individual perspective, the relative attractiveness of different actions is quite *insensitive* to what is the network structure faced by any given player at the time of revising her choices. In essence, a player must make her fresh choices as if she were in a complete network. In this case, therefore, the risk dominant and inefficient convention prevails since, under complete connectivity, this convention is harder to destabilize (through mutations) than the efficient but risk-dominated one. By contrast, if costs of forming links are high, individual players choose to form links only with those who lead to substantial gross payoffs. This, in turn, leads to more selective linking decisions by players and a reduction in their strategic uncertainty, consequently facilitating the emergence of the efficient action.

*Third*, we elaborate on the role of cost-bearing in link formation. In our model, links are one-sided, i.e. they are taken at the initiative of one player, who also incurs its cost. This brings in the issue of externalities in the link formation

---

<sup>6</sup>The role played by our assumption that the game payoffs are positive (or at least non-negative) is discussed in Subsection 2.2. In Section 4, we contemplate alternative variations of the model that may dispense with it.

process and the potential for free-rider problems. But perhaps more interestingly, it also has an important bearing on the different vulnerability to change displayed by the very different ways of supporting a given architecture.<sup>7</sup> To fix ideas, consider a state where the social network is complete and all players choose a common action. What is the underlying pattern of links that makes some such state more fragile to a particular set of mutations? A moment's reflection suggests that the particular state of that kind which is most fragile is the one where the mutant players induce the strongest externalities (and thus incentives to change) on the rest of the players. This happens when, collectively, all of the mutants support (i.e. have active) links to the remaining players. If, for concreteness, the mutants are indexed from 1 to  $k$  and the other players from  $k+1$  to  $n$ , some such (complete) network would be one where every player  $i$  supports a link to *every* other player with higher index. In fact, our analysis yields the insight that such a highly *asymmetric* pattern of connections enhances the fragility of otherwise stationary states and thus must be at the origin of the paths of least-resistance that underlie the notion of stochastic stability.<sup>8</sup> This, again, serves to illustrate the interplay between network structure and action choice that is at the heart of our analysis.

We now place the paper and the results in the context of the literature. Traditionally, sociologists have held the view that individual actions, and in turn aggregate outcomes, are in large part determined by interaction structure. By contrast, economists have tended to focus on markets, where social ties and the specific features of the interaction structures between agents are typically not important. In recent years, economists have examined in greater detail the role of interaction structure and found that it plays an important role in shaping important economic phenomena (see the references given above, and also Granovetter, 1985). This has led to a study of the processes through which the structure emerges. The present paper is part of this general research program.

Next, we relate the paper to work in economics. The paper contributes to two research areas: network formation games and equilibrium selection/coordination problems. In earlier work on network formation it is assumed that the sole concern of players is whom they connect to – i.e. the only strategic considerations are associated to their linking decisions (see e.g., Aumann and Myerson (1989), Bala and Goyal (2000), Dutta, van den Nouweland and Tijs (1995), and

---

<sup>7</sup>Note, for example, that there are  $2^{n(n-1)/2}$  strategy profiles that support (with non-redundant links) a complete network with  $n$  players. Clearly, these strategy profiles allow for wide variation in the number of links formed by individual players (and hence also a wide range of payoffs).

<sup>8</sup>This observation is related to some recent work by Albert, Jeong, and Barabási (2000) on the error and attack tolerance displayed by different network arrangements. Specifically, these authors show that the wide dispersion in the distribution of links in many complex networks (e.g. the World-Wide Web) makes them rather fragile to targeted attack although very tolerant to unguided error. In our case, where mutation probabilities are conceived as very small, the “attack fragility” is the dominant consideration and this lends to networks with unequal distribution of links their key role in the analysis.

Jackson and Wolinsky (1996)). By contrast, the present paper presents a unified framework in which the emergence of social networks *and* the behavior of linked players can be jointly studied.<sup>9</sup>

Next we outline briefly the relationship of our paper to the literature on equilibrium selection in games. In many games of interest, multiple equilibria arise naturally, and so recent years have seen a considerable amount of research on equilibrium selection/coordination.<sup>10</sup> An important finding of this work is that interaction structure (i.e. the social network) matters and that, by varying it, the rate of change as well as the long run outcome can be significantly altered.<sup>11</sup> This underscores the importance of endogenizing the social network, i.e., examining the circumstances under which different interaction patterns emerge. From a methodological point of view, a natural way to do this is by assessing the stochastic stability of the different networks arising at Nash equilibria. This is the route undertaken in the present paper, where we adapt the techniques customarily used in the evolutionary literature to the present scenario (where players choose not only actions but partners as well).

Somewhat more specifically, the present approach is related in spirit to that subbranch of recent evolutionary literature where players are allowed to move among a fixed set of locations.<sup>12</sup> The basic insight flowing from it is that, if individuals can separate/insulate themselves easily from those who are playing an inefficient action (e.g., the risk-dominant action), then efficient “enclaves” will be readily formed and eventually attract the “migration” of others (who will therefore turn to playing efficiently). In a rough sense, one may be inclined to identify *easy* mobility with *low* costs of forming links. However, the considerations involved in each case turn out to be very different, as is evident from the stark contrast between our conclusions and those of the mobility literature (recall the above summary). There are two main reasons for this contrast. First, in our case, players do not *indirectly* choose their pattern of interaction with others by moving across a *pre-specified* network of locations (as in the case of player mobility). Rather, they construct *directly* their interaction network (with no

---

<sup>9</sup>In independent work, Droste, Gilles and Johnson (1999), Jackson and Watts (1999) and Skyrms and Pemantle (2000), study endogenous network formation. The first and third paper have a model of link formation based on individual incentives and are more directly related to our paper. The primary difference between these papers and our paper pertains to the timing of actions and links. We assume that the two are simultaneous, while the earlier papers assume that links and actions are revised one at a time, taking the other as given. This difference leads to different conclusions. We further discuss the issue of timing of actions and links in Subsection 4.6 below.

<sup>10</sup>One strand of this work considers dynamic models. This work includes Blume (1993), Canning (1992), Ellison (1993), Kandori, Mailath and Rob (1993), and Young (1993), among others. For a consideration of this same equilibrium selection problem from a different (“educative”) perspective, the reader may refer to the work of Harsanyi and Selten (1988) or the more recent paper by Carlson and van Damme (1993).

<sup>11</sup>See, for example, Ellison (1993), Goyal (1996), Lee and Valentinyi (2000), and Morris (2000), and Robson and Vega-Redondo (1996), among others.

<sup>12</sup>See e.g. Ely (1996), Mailath, Samuelson and Shaked (1994), Oechssler (1997), or Bhaskar and Vega-Redondo (2000), among others.

exogenous restrictions) by choosing those agents with whom they want to play the game. Second, the cost of link formation is paid per link formed and thus becomes truly effective only if it is high enough. In a heuristic sense, we may say that it is precisely the restricted “mobility” that high costs induce which helps insulate (and thus protect) the individuals who are choosing the efficient action. If the link-formation costs are too low, the extensive interaction this facilitates may have the unfortunate consequence of rendering risk-dominance considerations decisive.

The rest of this paper is organized as follows. Section 2 describes the framework. Section 3 presents the results for the basic model. Section 4 explores a robustness of our findings with respect to a number of changes in the model such as modifications in the link formation process, different specifications of link formation costs, alternative models of mutations as well as the possibility of interaction among indirectly connected players. Section 5 concludes.

## 2 The Model

### 2.1 Networks

Let  $N = \{1, 2, \dots, n\}$  be a set of players, where  $n \geq 3$ . We are interested in modelling a situation where each of these players can choose the subset of other players with whom to play a fixed bilateral game. Formally, let  $g_i = (g_{i1}, \dots, g_{i,i-1}, g_{i,i+1}, \dots, g_{in})$  be the set of links formed by player  $i$ . We suppose that  $g_{ij} \in \{1, 0\}$ , and say that player  $i$  forms a link with player  $j$  if  $g_{ij} = 1$ . The set of link options is denoted by  $\mathcal{G}_i$ . Any player profile of link decisions  $g = (g_1, g_2 \dots g_n)$  defines a directed graph, called a *network*. Abusing notation, the network will also be denoted by  $g$ .

Specifically, the network  $g$  has the set of players  $N$  as its set of *vertices* while its set of arrows,  $\Gamma \subset N \times N$ , is defined as follows:  $\Gamma = \{(i, j) \in N \times N : g_{ij} = 1\}$ . Graphically, the link  $(i, j)$  may be represented as an edge between  $i$  and  $j$ , a filled circle lying on the edge near agent  $i$  indicating that this agent has formed (or supports) that link. Every link profile  $g \in \mathcal{G}$  has a unique representation in this manner. Figure 2 below depicts an example. In it, player 1 has formed links with players 2 and 3, player 3 has formed a link with player 1, while player 2 has formed no links.<sup>13</sup>

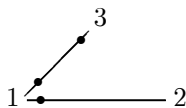


Figure 2

<sup>13</sup>Since agents choose strategies independently of each other, two agents may simultaneously initiate a two-way link, as seen in the figure.

Given a network  $g$ , we say that a pair of players  $i$  and  $j$  are directly linked if at least one of them has established a link with the other one, i.e. if  $\max\{g_{ij}, g_{ji}\} = 1$ . To describe the pattern of players' links, it is useful to define a modified version of  $g$ , denoted by  $\bar{g}$ , that is defined as follows:  $\bar{g}_{ij} = \max\{g_{ij}, g_{ji}\}$  for each  $i$  and  $j$  in  $N$ . Note that  $\bar{g}_{ij} = \bar{g}_{ji}$  so that the index order is irrelevant. We refer to  $g_{ij}$  as an active link for player  $i$  and a passive link for player  $j$ .

We say there is a *path* between  $i$  and  $j$  if either  $\bar{g}_{ij} = \max\{g_{ij}, g_{ji}\} = 1$  or there exist agents  $j_1, \dots, j_m$  distinct from each other and  $i$  and  $j$  such that  $\bar{g}_{i,j_1} = \dots = \bar{g}_{j_k, j_{k+1}} = \dots = \bar{g}_{j_m, j} = 1$ . A sub-graph  $g' \subset g$  is called a *component* of  $g$  if for all  $i, j \in g'$ ,  $i \neq j$ , there exists a path in  $g'$  connecting  $i$  and  $j$ , and for all  $i \in g'$  and  $j \in g$ ,  $g_{ij} = 1$  implies  $g'_{ij} = 1$ . A network with only one component is called connected. On the other hand, a network (or a component) is said to be complete if every pair of nodes in it is connected by a link in either direction (recall Figure 1a). Finally, a network is called minimally connected if the removal of any single link renders it disconnected. A simple example of such a network is provided by the center-sponsored star of Figure 1b.

## 2.2 Social Game

Individuals located in a social network play a  $2 \times 2$  symmetric game in strategic form with common action set. The set of partners of player  $i$  depends on her location in the network. In the basic model we assume that two individuals can play a game if, and only if, they have a direct link between them.

We now describe the bilateral game that is played between any two partners. The set of actions is  $A = \{\alpha, \beta\}$ . For each pair of actions  $a, a' \in A \times A$  the payoffs to the players are given by the following table, with the pay-offs to the row player given first.

	2	$\alpha$	$\beta$
1		$d, d$	$e, f$
		$f, e$	$b, b$

Table I

We shall assume that the game is one of coordination with two pure strategy equilibria,  $(\alpha, \alpha)$  and  $(\beta, \beta)$ . Without loss of generality we will also assume that  $(\alpha, \alpha)$  is the efficient equilibrium. Finally, in order to focus on the interesting case, we will assume that there is a conflict between efficiency and risk-

dominance. These considerations are summarized in the following restrictions on the payoffs.<sup>14</sup>

$$d > f; \quad b > e; \quad d > b; \quad d + e < b + f. \quad (1)$$

An important feature of our approach is that links are *costly*. Specifically, every agent who establishes a link with some other player incurs a cost  $c > 0$ . Thus, we suppose that the cost of forming each link is independent of the number of links being established and is the same across all players.

In the basic model we assume that links are *one-sided*. This aspect of the model allows us to use standard solution concepts from non-cooperative game theory in addressing the issue of link formation. We shall suppose that the payoffs in the bilateral game are all positive and, therefore, no player has any incentive to refuse links initiated by other players. There are different ways in which the assumption of positive payoffs in the coordination game can be relaxed. One route is to dispense with any restriction on payoffs but suppose that, when player  $i$  supports a link to player  $j$ , the payoff (which may be negative) flows only to  $i$ . This formulation may be interpreted as reflecting a model of peer groups and fashion, where asymmetric flow of influence seems a natural feature. Another possible route to tackle possibly negative payoffs would be to maintain the bilateral nature of payoffs but give players the option to refuse the links initiated by others. We discuss a variety of alternative formulations of the link formation process in section 4.

Every player  $i$  is obliged to choose the same action in the (possibly) several bilateral games that she is engaged in. This assumption is natural in the present context: if players were allowed to choose a different action for every two-person game they are involved in, this would make the behavior of players in any particular game insensitive to the network structure. Thus, combining the former considerations, the strategy space of a player can be identified with  $S_i = \mathcal{G}_i \times A$ , where  $\mathcal{G}_i$  is the set of possible link decisions by  $i$  and  $A$  is the common action space of the underlying bilateral game.<sup>15</sup>

We can now present the payoffs of the social game. Given the link decisions of players,  $g = (g_1, g_2, \dots, g_n)$ , let  $N(i; g) \equiv \{j \in N : g_{ij} = 1\}$  be the set of agents in the induced network with whom player  $i$  has *established* links, while  $\nu(i; g) \equiv |N(i; g)|$  is its cardinality. Similarly, denote by  $N(i; \bar{g}) \equiv \{j \in N : \bar{g}_{ij} = 1\}$  the set of agents with whom player  $i$  is directly connected (by active

---

<sup>14</sup>Our results extend in a natural way in case the risk-dominant equilibrium is also efficient, i.e., if  $d + e > b + f$ . In particular, players coordinate on the  $(\alpha, \alpha)$  equilibrium, which is risk-dominant as well as efficient, in the long run.

<sup>15</sup>In our formulation, players choose links and actions in the coordination game at the same time. An alternative formulation would be to have players choose links first and then choose actions, contingent on the nature of the network observed. Finally another alternative, considered by the literature, is to postulate that the action and every link are revised separately (cf. Footnote 9 and Subsection 4.6).

or passive links), while  $\nu(i; \bar{g}) \equiv |N(i; \bar{g})|$  stands for the cardinality of this set. Then, given the strategies of other players,  $s_{-i} = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$ , the payoff to a player  $i$  from playing some strategy  $s_i = (g_i, a_i)$  is given by:

$$\Pi_i(s_i, s_{-i}) = \sum_{j \in N(i; \bar{g})} \pi(a_i, a_j) - \nu(i; g) \cdot c. \quad (2)$$

We note that the individual payoffs are aggregated across the games played by him. In much of earlier work, e.g. Kandori, Mailath and Rob (1993) or Ellison (1993), the distinction between average or total payoffs was irrelevant since the size of the neighborhood was given. In our model, however, where the number of games an agent plays is endogenous, we want to explicitly account for the influence of the size of the neighborhood and thus choose the aggregate-payoff formulation.<sup>16</sup>

These payoff expressions allow us to particularize the standard notion of Nash Equilibrium to each of the two alternative scenarios. Thus, for the model with direct links, a strategy profile  $s^* = (s_1^*, \dots, s_n^*)$  is said to be a *Nash equilibrium* if, for all  $i \in N$ ,

$$\Pi_i(s_i^*, s_{-i}^*) \geq \Pi_i(s_i, s_{-i}^*), \quad \forall s_i \in S_i. \quad (3)$$

On the other hand, a Nash equilibrium in either scenario will be called *strict* if every player gets a strictly higher payoff with her current strategy than she would with any other strategy. The set of Nash equilibria will be denoted by  $S^*$  and that of strict Nash equilibria by  $S^{**}$ .

### 2.3 Dynamics

Time is discrete, and denoted by  $t = 1, 2, 3, \dots$ . At each  $t$ , the state of the system is given by the strategy profile  $s(t) \equiv [(g_i(t), a_i(t))]_{i=1}^n$  specifying the action played, and links established, by each player  $i \in N$ . At every period  $t$ , there is a positive independent probability  $p \in (0, 1)$  that any given individual gets a chance to revise her strategy. If she receives this opportunity, we assume that she selects a new strategy

$$s_i(t) \in \arg \max_{s_i \in S_i} \Pi_i(s_i, s_{-i}(t-1)). \quad (4)$$

Time is modeled discretely,  $t = 1, 2, 3, \dots$ . At each  $t$ , the state of the system is given by the strategy profile  $s(t) \equiv [(g_i(t), a_i(t))]_{i=1}^n$  specifying the action played, and links established, by each player  $i \in N$ . At every period  $t$ , there is a positive independent probability  $p \in (0, 1)$  that any given individual gets a

---

<sup>16</sup>When players seek to maximize average payoffs, the size of the interaction group plays no essential role and we conjecture that at least some efficient state (i.e. a state where all players choose action  $\alpha$ ) must be stochastically stable. The intuition for this conjecture is that when neighborhood size is irrelevant per se it should be particularly easy to destabilize inefficient states.

chance to revise her strategy. If she receives this opportunity, we assume that she selects a new strategy

$$s_i(t) \in \arg \max_{s_i \in S_i} \Pi_i(s_i, s_{-i}(t-1)). \quad (5)$$

That is, she selects a *myopic* best response to what other players chose in the preceding period.<sup>17</sup> If there are several strategies that fulfill (4), then any one of them is taken to be selected with, say, equal probability. This strategy revision process defines a simple Markov chain on  $S \equiv S_1 \times \dots \times S_n$ . In our setting, which will be seen to display multiple strict equilibria, there are several absorbing states of the Markov chain.<sup>18</sup> This motivates the examination of the relative robustness of each of them.

To do so, we rely on the approach proposed by Kandori, Mailath and Rob (1993), and Young (1993). We suppose that, occasionally, players make mistakes, experiment, or simply disregard payoff considerations in choosing their strategies. Specifically, we assume that, conditional on receiving a revision opportunity, a player chooses her strategy at random with some small “mutation” probability  $\epsilon > 0$ . For any  $\epsilon > 0$ , the process defines a Markov chain that is aperiodic and irreducible and, therefore, has a unique invariant probability distribution. Let us denote this distribution by  $\mu_\epsilon$ . We analyze the form of  $\mu_\epsilon$  as the probability of mistakes becomes very small, i.e. formally, as  $\epsilon$  converges to zero. Define  $\lim_{\epsilon \rightarrow 0} \mu_\epsilon = \hat{\mu}$ . When a state  $s = (s_1, s_2, \dots, s_n)$  has  $\hat{\mu}(s) > 0$ , i.e. it is in the support of  $\hat{\mu}$ , we say that it is *stochastically stable*. Intuitively, this reflects the idea that, even for infinitesimal mutation probability (and independently of initial conditions), this state materializes a *significant* fraction of time in the long run.

### 3 Evolving networks and social coordination

We first characterize the Nash equilibrium of the social game. We then provide a complete characterization of the set of stochastically stable social outcomes.

---

<sup>17</sup>We are implicitly assuming that players have complete information on the network structure as well as on the profile of actions. This assumption simplifies the strategy choice significantly in a setting where a player can potentially play with everyone else in the society. Another important simplification derives from the assumption that players are fully myopic and hold static expectations. This may be justified if the adjustment of the process (say, due to inertia of choice) is quite slow and players are *relatively* impatient. If these conditions do not apply, however, the dynamics of the process should be significantly affected, e.g. the formation of a mutually beneficial link may often involve a strategic *tour de force* and the consequent delays familiar in public-good contexts (cf. Bliss and Nalebuff (1984)). For an evolutionary approach to modelling these issues, the reader is referred to the work of Blume (1995) and Lagunoff and Matsui (1995).

<sup>18</sup>We note that the set of absorbing states of the Markov chain coincides with the set of strict Nash equilibria of the one-shot game.

### 3.1 Equilibrium outcomes

Our first result concerns the nature of networks that arise in equilibria. If costs of link formation are low ( $c < e$ ), then a player has an incentive to link up with other players irrespective of the actions the other players are choosing. On the other hand, when costs are quite high (specifically,  $b < c < d$ ) then everyone who is linked must be choosing the efficient action. This, however, implies that it is attractive to form a link with every other player and we get the complete network again. Thus, for relatively low and high costs, we should expect to see the complete network. In contrast, if costs are at an intermediate level ( $f < c < b$ ), a richer set of configurations is possible. On the one hand, since  $c > f (> e)$ , the link formation is only worthwhile if other players are choosing the same action. On the other hand, since  $c < b (< d)$ , coordinating at either of the two equilibria (in the underlying coordination game) is better than not playing the game at all. This allows for networks with two disconnected components in equilibria. The former considerations are reflected by the following result, whose proof is given in Appendix.

**Proposition 3.1** *Suppose (1) and (2) hold. (a) If  $c < \min\{f, b\}$ , then an equilibrium network is complete. (b) If  $f < c < b$ , then an equilibrium network is either complete or can be partitioned into two complete components.<sup>19</sup> (c) If  $b < c < d$ , then an equilibrium network is either empty or complete. (d) If  $c > d$ , then the unique equilibrium network is empty.*

Next, we characterize the Nash equilibria of the static game. First, we introduce some convenient notation. On the one hand, recall that  $g^e$  denotes the empty network characterized by  $g_{ij}^e = 0$  for all  $i, j \in N$  ( $i \neq j$ ). We shall say that a network  $g$  is essential if  $g_{ij}g_{ji} = 0$ , for every pair of players  $i$  and  $j$ . Also, let  $G^c \equiv \{g : \forall i, j \in N, \bar{g}_{ij} = 1, g_{ij}g_{ji} = 0\}$  stand for the set of complete and essential networks on the set  $N$ . Analogously, for any given subset  $M \subset N$ , denote by  $G^c(M)$  the set of complete and essential subgraphs on  $M$ . Given any state  $s \in S$ , we shall say that  $s = (g, a) \in S^h$  for some  $h \in \{\alpha, \beta\}$  if  $g \in G^c$  and  $a_i = h$  for all  $i \in N$ . More generally, we shall write  $s = (g, a) \in S^{\alpha, \beta}$  if there exists a partition of the population into two subgroups,  $N^\alpha$  and  $N^\beta$  (one of them possibly empty), and corresponding components of  $g$ ,  $g^\alpha$  and  $g^\beta$ , such that: (i)  $g^\alpha \in G^c(N^\alpha)$ ,  $g^\beta \in G^c(N^\beta)$ ; and (ii)  $\forall i \in N^\alpha, a_i = \alpha$ ;  $\forall i \in N^\beta, a_i = \beta$ . With this notation in hand, we may state the following result.

**Proposition 3.2** *Suppose (1) and (2) hold. (a) If  $c < \min\{f, b\}$ , then the set of equilibrium states  $S^* = S^\alpha \cup S^\beta$ . (b) If  $f < c < b$ , then  $S^\alpha \cup S^\beta \subset S^* \subset S^{\alpha, \beta}$ , the first inclusion being strict for large enough  $n$ . (c) If  $b < c < d$ , then  $S^* = S^\alpha \cup \{(g^e, (\beta, \beta, \dots, \beta))\}$ . (d) If  $c > d$ , then  $S^* = \{g^e\} \times A^n$ .*

Parts (a) and (c) are straightforward; we therefore elaborate on the coexistence equilibria identified in part (b). In these equilibria, there are two unconnected

<sup>19</sup>Our parameter conditions allow both  $f < b$  and  $b < f$ . If the latter inequality holds, Part (b) of Proposition 3.1 (and also that of Proposition 3.2 below) applies trivially.

groups, with each group adopting a common action (different in each group). The strategic stability of this configuration rests on the appeal of ‘passive’ links. A link such as  $g_{ij} = 1$  is paid for by player  $i$ , but both players  $i$  and  $j$  derive payoffs from it. In a mixed equilibrium configuration, the links in each group must be, roughly, evenly distributed. This means that all players enjoy some benefits from passive links. In contrast, if a player were to switch actions, then to derive the full benefits of this switch, she would have to form (active) links with everyone in the new group. This lowers the incentives to switch, a consideration which becomes decisive if the number of passive links is large enough (hence the requirement of large  $n$ ).

The above result indicates that, whenever the cost of links is not very high (i.e. not above the maximum payoff attainable in the game), a wide range of outcomes can arise in equilibrium. For example, under the parameter configurations allowed in Parts (a) and (c), states where either of the two actions is homogeneously chosen by the whole population can arise in equilibrium. On the other hand, if  $f < c < b$ , states where neither action homogeneity nor full connectedness obtains can arise in equilibrium. The model, therefore, raises a fundamental issue of equilibrium selection.

## 3.2 Dynamics

This section addresses the problem of equilibrium selection by using the techniques of stochastic stability. As a first step in this analysis, we establish convergence of the unperturbed dynamics for the relevant parameter range.

Let  $\bar{S}$  denote the set of absorbing states of the unperturbed dynamics. In view of the postulated adjustment process, it follows that there is an one-to-one correspondence between  $\bar{S}$  and the class of *strict* Nash equilibria of the social game, i.e.  $\bar{S} = S^{**}$ . Proposition 3.2 characterizes *all* Nash equilibria of this game. But, clearly, if  $c < b$ , every Nash equilibrium is strict, while if  $b < c < d$ , only the Nash equilibria in  $S^\alpha$  are strict. Finally, no strict Nash equilibrium exists if  $c > d$ . So the next result focuses on the case where  $c < d$ .

**Proposition 3.3** *Suppose (1)-(2) hold and  $c < d$ . Then, starting from any initial strategy configuration, the best response dynamics converges to a strict Nash equilibrium of the social game, with probability one.*

The proof of the above result is given in Appendix. This result delimits the set of states that can potentially be stochastically stable since, obviously, every such state must be a limit point for the unperturbed dynamics. Let the set of *stochastically stable states* be denoted by  $\hat{S} \equiv \{s \in S : \hat{\mu}(s) > 0\}$ . The following result summarizes our analysis.

**Theorem 3.1** *Suppose (1) and (2) hold. There exists some  $\bar{c} \in (e, b)$  such that if  $c < \bar{c}$  then  $\hat{S} = S^\beta$  while if  $\bar{c} < c < d$  then  $\hat{S} = S^\alpha$ , provided  $n$  is large enough.<sup>20</sup> Finally, if  $c > d$  then  $\hat{S} = \{g^e\} \times A^n$ .*

Recall that a social outcome is stochastically stable if it lies in the support of the limit distribution,  $\hat{\mu}$ . In order to determine this support we use the techniques introduced in Kandori *et al.*(1993) and Young (1993). They can be summarized as follows. Fix some state  $s \in \bar{S}$ . An  $s$ -tree is a directed graph on  $\bar{S}$  whose root is  $s$  and such that there is a *unique* (directed) path joining *any* other  $s' \in \bar{S}$  to  $s$ . For each arrow  $s' \rightarrow s''$  in any given  $s$ -tree, a “cost” is defined as the minimum number of simultaneous mutations that are required for the transition from  $s'$  to  $s''$  to be feasible through the ensuing operation of the unperturbed dynamics alone. The cost of the tree is obtained by adding up the costs associated with all the arrows of a particular  $s$ -tree. The stochastic potential of  $s$  is defined as the minimum cost across all  $s$ -trees. Then, a state  $s \in \bar{S}$  is seen to be stochastically stable if it has the lowest stochastic potential across all  $s \in \bar{S}$ .

In our framework, individual strategies involve both link-formation and action choices. This richness in the strategy space leads to a corresponding wide variety in the nature of (strict) Nash equilibria of the social game. There are two facets of this variety: (a) we obtain three different types of equilibria in terms of action configuration:  $S^\alpha$ ,  $S^\beta$  and  $S^{\alpha,\beta}$ ; and (b) there are a large number of strategy profiles that support the complete connectivity prevailing at equilibrium configurations - recall Footnote 7. This proliferation of equilibria leads us to develop a simple relationship between the different profiles. In particular, we consider strategy profiles within the sets  $S^h$  ( $h = \alpha, \beta$ ) and show that states in each of these sets can be connected by a chain of single-mutation steps, each such step followed by a suitable operation of the best-response dynamics. Let distance between two networks  $g$  and  $g'$  be defined as follows:  $d(g, g') = d(g', g) \equiv \sum_{i,j} |g_{i,j} - g'_{i,j}|/2$ . In words, this distance is simply a measure of the number of links that are different across the two networks. With this metric in place, we have:

**Lemma 3.1** *For each  $s \in S^h$ ,  $h = \alpha, \beta$ , there exists an  $s$ -tree restricted to  $S^h$  such that for all arrows  $s' \rightarrow s''$  in it,  $d(g', g'') = 1$ , where  $g'$  and  $g''$  are the networks respectively associated to  $s'$  and  $s''$ .*

The proof of this lemma is given in Appendix. This lemma implies that, provided  $S^h \subset \bar{S}$ , the (restricted) tree established by Lemma 3.1 for any  $s \in S^h$  involves the minimum possible cost  $|S^h| - 1$ . This lemma also indicates that, in the language of Samuelson (1994),  $S^\alpha$  (if  $c < d$ ) and also  $S^\beta$  (if  $c < b$ ) are *recurrent sets*. This allows each of them to be treated as a single “entity” in the following two complementary senses: (i) if any state in one of these recurrent sets is stochastically stable, so is every other state in this same set; (ii)

---

<sup>20</sup>The proviso on  $n$  is simply required to deal with possible integer problems when studying the number of mutations needed for the various transitions.

in evaluating the minimum cost involved in a transition to, or away from, any *given* state in a recurrent set, the sole relevant issue concerns the minimum cost associated to a transition to, or away from, *some* state in that recurrent set. Using (i)-(ii), the analysis of the model can be greatly simplified. To organize matters, it is useful to consider different ranges of  $c$  separately.

Let us start with the case where  $0 < c < e$ , where the set of absorbing states  $\bar{S} = S^\alpha \cup S^\beta$ . Since, by Lemma 3.1, the sets  $S^\alpha$  and  $S^\beta$  are each recurrent, the crucial point here is to assess what is the minimum (mutation) cost across all path joining *some* state in  $S^h$  to *some* state in  $S^{h'}$  for each  $h, h' = \alpha, \beta, h \neq h'$ . Denote these costs by  $m^{h,h'}$  and let  $\lceil z \rceil$  stand for the smallest integer no smaller than any given  $z \in \mathcal{R}_+$ . With this notation in place, we state:

**Lemma 3.2** *Suppose that  $0 < c < e$ . Then,*

$$m^{\beta,\alpha} = \left\lceil \frac{b-e}{(d-f) + (b-e)}(n-1) \right\rceil; \quad m^{\alpha,\beta} = \left\lceil \frac{d-f}{(d-f) + (b-e)}(n-1) \right\rceil.$$

*Thus,  $m^{\beta,\alpha} > m^{\alpha,\beta}$ , for large enough  $n$ .*

The proof, given in the Appendix, reflects the standard considerations arising in much of the recent evolutionary theory where the *fixed* pattern of interaction involves every individual of the population playing with all others. Now, if costs are low ( $c < e$ ), such full connectivity is not just assumed but arises from players' own decisions, both at equilibrium (i.e. when the unperturbed best-response dynamics is at a rest-point) and away from it. In effect, this implies that the same basin-of-attraction considerations that privilege risk-dominance in the received approach also select for it in the present case.

We next examine the case where  $e < c < \min\{f, b\}$  where  $\bar{S} = S^\alpha \cup S^\beta$ . Now, since  $c > e$ , players who choose action  $\alpha$  no longer find it attractive to form links with other players who choose action  $\beta$ . This factor plays a crucial role in the analysis. The following result derives the relative magnitude of the minimum mutation costs.

**Lemma 3.3** *Suppose  $e < c < \min\{f, b\}$ . Then,*

$$m^{\beta,\alpha} = \left\lceil \frac{b-c}{(d-f) + (b-c)}(n-1) \right\rceil; \quad m^{\alpha,\beta} = \left\lceil \frac{d-f}{(d-f) + (b-e)}(n-1) \right\rceil.$$

*Thus, there is some  $\tilde{c}$ ,  $e < \tilde{c} \leq \min\{f, b\}$ , such that if  $c < \tilde{c}$  then  $m^{\beta,\alpha} - m^{\alpha,\beta} > 0$ , while if  $c > \tilde{c}$  then  $m^{\beta,\alpha} - m^{\alpha,\beta} < 0$ , for large enough  $n$ .*

The methods used to prove this lemma are quite general; we use them in establishing a number of other results required for the proof of Theorem 3.1. It is therefore useful to explain them in the text.

**Proof of Lemma 3.3:** Let  $s^\alpha$  and  $s^\beta$  be states in  $S^\alpha$  and  $S^\beta$ , respectively.

Step 1: Consider transitions from state  $s^\beta$  to state  $s^\alpha$  and let  $k$  be the number of mutations triggering it. If this transition is to take place after those many mutations, there must be some player currently choosing  $\beta$  who will then voluntarily switch to  $\alpha$ . Denote by  $q^h$  the number of active links this player *chooses* to support to players choosing  $h$  ( $h = \alpha, \beta$ ) and let  $r^h$  stand for the number of passive links she receives from players choosing  $h$  ( $h = \alpha, \beta$ ). If she chooses  $\alpha$ , her payoff is given by:

$$\pi_\alpha = r^\alpha d + r^\beta e + q^\alpha (d - c), \quad (6)$$

Note that since  $c > e$ ,  $q^\beta$  must equal zero. On the other hand, if the agent in question were to continue with  $\beta$ , her payoff would be equal to:

$$\pi_\beta = \hat{r}^\alpha f + \hat{r}^\beta b + \hat{q}^\alpha (f - c) + \hat{q}^\beta (b - c), \quad (7)$$

where  $\hat{q}^h$  and  $\hat{r}^h$  are interpreted as the active and passive links that would be chosen by the player if she adopts  $\beta$ . Clearly, we must have  $r^h = \hat{r}^h$  for each  $h = \alpha, \beta$ . Thus, if a switch to  $\alpha$  is to take place, it must be that

$$\pi_\alpha - \pi_\beta = (r^\alpha + q^\alpha)d - (r^\alpha + \hat{q}^\alpha)f - r^\beta(b - e) - \hat{q}^\beta(b - c) \geq 0. \quad (8)$$

Note that  $r^\alpha + q^\alpha = k$ , since  $c < d$  and therefore the player who switches to  $\alpha$  will want to be linked (either passively or by supporting a link herself) to all other players choosing  $\alpha$ , i.e. to the total number  $k$  of  $\alpha$ -mutants. On the other hand, since  $c < \min\{f, b\}$ , we must also have that  $\hat{r}^\beta + \hat{q}^\beta = n - k - 1$  and  $r^\alpha + \hat{q}^\alpha = \hat{r}^\alpha + \hat{q}^\alpha = k$ , i.e. the player who chooses  $\beta$  must become linked to all other players, both those choosing  $\beta$  and those choosing  $\alpha$ .

We now ask: What is the lowest value of  $k$  consistent with (8)? Since  $c > e$ , the desired payoff advantage of action  $\alpha$  will occur for the lowest value of  $k$  when  $r^\beta = \hat{r}^\beta = 0$  and therefore  $\hat{q}^\beta = n - k - 1$ . That is, if the desired transition is to take place, the *necessary* condition (8) holds for the *minimum* number of required mutations when the arbitrary agent that must start the transition has *no* passive links from individuals choosing action  $\beta$ . Recall that  $m^{\beta, \alpha}$  stands for the minimum number of mutations required for the transition. From (8), it now follows that

$$m^{\beta, \alpha} \geq \frac{b - c}{(d - f) + (b - c)}(n - 1) \equiv H. \quad (9)$$

The above expression gives the minimum number of players choosing  $\alpha$  that are needed to induce some player to switch to action  $\alpha$  across all possible network structures. Next, we argue that this number of mutations is also *sufficient* to induce a transition from some  $s^\beta$  to some  $s^\alpha$ . The proof is constructive.

The main idea is to consider a particular state  $s^\beta$  where its corresponding (complete) network displays the maximal responsiveness to some suitably chosen mutations. Using the observations on the distribution of active and passive links,

this is seen to occur when there are some players who support links to all others – those are, of course, players with a “critical” role whose mutation would be most effective. Specifically, suppose that the network prevailing in  $s^\beta$  has every player  $i = 1, 2, \dots, n$  support active links to all  $j > i$ . (This means, for example, that player 1 supports links to every other player whereas player  $n$  only has passive links.) Then, recalling that  $\lceil z \rceil$  denotes the smallest integer no smaller than  $z$ , the most mutation-effective way of inducing the population to switch actions from  $\beta$  to  $\alpha$  is precisely by having the players  $\ell = 1, 2, \dots, \lceil H \rceil$  simultaneously mutate to action  $\alpha$  and maintain all their links. Thereafter, a transition to some state  $s^\alpha$  will occur if subsequent strategy revision opportunities are appropriately sequenced so that every player with index  $j = \lceil H \rceil + 1, \dots, n$  is given a revision opportunity in order. This, in effect, shows that the lower bound in (9) is tight and  $m^{\beta, \alpha} = \lceil H \rceil$ .

Step 2: Consider next the transition from some state  $s^\alpha$  to a state  $s^\beta$ . Using arguments from Step 1 above, it is easy to show that  $m^{\alpha, \beta}$ , must satisfy:

$$m^{\alpha, \beta} \geq \frac{d-f}{(d-f) + (b-e)}(n-1) \equiv H'. \quad (10)$$

Again, we can use previous arguments to show that  $\lceil H' \rceil$  is sufficient.

Step 3: Finally, we wish to study the difference  $m^{\beta, \alpha} - m^{\alpha, \beta}$  as a function of  $c$ . For low  $c$  (close to  $e$ ) and large  $n$ , this difference is clearly positive in view of the hypothesis that  $b-e > d-f$ . Next, to verify that it switches strict sign at most once in the range  $c \in (e, \min\{f, b\})$ , note that  $H - H'$  is strictly declining with respect to  $c$  in the interval  $(e, \min\{f, b\})$ .  $\triangle$

Lemma 3.3 applies both to the case where  $b < f$  and that where  $b > f$ . Suppose first that  $b < f$ . Then, since  $H - H' < 0$  for  $c = b$ , a direct combination of former considerations leads to the desired conclusion for the parameter range  $c \in (e, b]$ . We now take up the case  $f < b$  and focus on the range  $c \in (f, b)$ . We first derive the relative magnitude of the minimum mutation costs for  $s \in S^h$ , where  $h = \{\alpha, \beta\}$ .

**Lemma 3.4** *Suppose  $f < c < b$ .*

$$m^{\beta, \alpha} = \left\lceil \frac{b-c}{(d-f) + (b-c)}(n-1) \right\rceil; \quad m^{\alpha, \beta} = \left\lceil \frac{d-c}{(d-c) + (b-e)}(n-1) \right\rceil.$$

*Thus there is a threshold  $\check{c} \in [f, b)$  such that if  $c < \check{c}$  then  $m^{\beta, \alpha} - m^{\alpha, \beta} > 0$ , while if  $c > \check{c}$  then  $m^{\beta, \alpha} - m^{\alpha, \beta} < 0$ , for large enough  $n$ .*

The arguments needed to establish this result are very similar to those used in the proof of Lemma 3.3; we provide the computations in Appendix.

The principal complication in case  $c \in [f, b)$  is that the set of absorbing states is not restricted to  $S^\alpha \cup S^\beta$  but will generally include mixed states where the

population is segmented into two different action components (cf. Propositions 3.2 and 3.3). Let  $m^{h,\alpha\beta}$ , for  $h = \alpha, \beta$ , denote the minimum number of mutations needed to ensure a transition from some  $s \in S^h$  to some  $s \in S^{\alpha,\beta}$ . The first point to note is that by the construction used in Lemma 3.3,  $m^{\alpha,\alpha\beta} \geq m^{\alpha,\beta}$  and, similarly,  $m^{\beta,\alpha\beta} \geq m^{\beta,\alpha}$ . This implies that the transition from any state in some  $S^h$  towards a mixed equilibrium state in  $S^{\alpha,\beta}$  is *costlier* than a transition towards  $S^{h'}$  ( $h' \neq h$ ). Concerning now the converse transitions (i.e. from states in  $S^{\alpha,\beta}$  to either  $S^\alpha$  or  $S^\beta$ ), the following lemma indicates that it is relatively “easy” since it involves a suitable chain of *single* mutations..

**Lemma 3.5** *Let  $f < c < b$  and consider any equilibrium state  $s \in S^{\alpha,\beta}$  involving two non-degenerate ( $\alpha$  and  $\beta$ ) components,  $g^\alpha$  and  $g^\beta$ , with cardinalities  $|A(s)| > 0$  and  $|B(s)| > 0$ , respectively. Then, there is another equilibrium state  $s'$  with cardinality for the resulting  $\alpha$  component  $|A(s')| \geq |A(s)| + 1$  that can be reached from  $s$  by a suitable single mutation followed by the best-response dynamics. An identical conclusion applies to some equilibrium state  $s''$  with  $|A(s'')| \leq |A(s)| - 1$ .*

The proof of this Lemma is given in Appendix. We briefly sketch the argument here. Fix some mixed state, and suppose the strategy of some player  $i \in B(s)$  mutates as follows: she switches to action  $\alpha$ , while everything else remains as before. Now, have all the players in the  $\beta$  group move and suppose that they still wish to keep playing action  $\beta$ . Since  $c > f$ , their best response is to delete their links with player  $i$ . Next, have all the players in group  $\alpha$  move; their best response is to form a link with player  $i$ , since  $c < b$ . Finally, have player  $i$  choose a best response; since the original state was an equilibrium and  $c > f$ , her best response is to play action  $\alpha$  and delete all links with players in the  $\beta$  group. We have thus increased the number of  $\alpha$  players with a single mutation. This argument extends in a natural manner to prove the above result. We now have all the information to complete the proof of Theorem 3.1.

**Proof of Theorem 3.1:** Consider first the case where  $c < b$ . If  $f > b$ , the sets  $S^\alpha$  and  $S^\beta$  are the only candidates for stochastic stability and we simply need to compare  $m^{\alpha,\beta}$  versus  $m^{\beta,\alpha}$ . Then, the desired conclusion follows directly from Lemmas 3.2-3.4. The same applies if  $f < b$  but  $c < f$ . Thus, consider the case where  $f < c < b$ . Then, the states in  $S^\alpha$ ,  $S^\beta$ , and  $S^{\alpha,\beta}$  are possible candidates for stochastic stability. Take any state  $s \in S^h$  for some  $h = \alpha, \beta$ . With the help of Lemmas 3.1 and 3.5 we can infer that  $s$ -trees for any  $s \in S^\alpha$  will have the following minimum cost:  $m^{\beta,\alpha} + |S^\alpha| + |S^\beta| + |S^{\alpha,\beta}| - 2$ . For any  $s' \in S^\beta$ , the situation is symmetric, the minimum cost being equal to  $m^{\alpha,\beta} + |S^\alpha| + |S^\beta| + |S^{\alpha,\beta}| - 2$ . Next, concerning any  $s \in S^{\alpha,\beta}$ , we note that the corresponding  $s$ -tree would have to display a path joining some state in  $S^\alpha$  to  $s$  and some path joining some state in  $S^\beta$  to  $s$ . Thus, the cost of such an  $s$ -tree will be at least  $m^{\alpha,\beta} + m^{\beta,\alpha} + |S^\alpha| + |S^\beta| + |\tilde{S}^{\alpha\beta}| - 3$ . This expression is greater than the minimum  $s$ -tree costs for  $s \in S^h$  ( $h = \alpha, \beta$ ) since each  $m^{h,h'} > 1$  if the population is large. We therefore conclude that a state  $s \in S^{\alpha,\beta}$  cannot be

stochastically stable. Thus we only need to compare  $m^{\alpha,\beta}$  with  $m^{\beta,\alpha}$ ; the result follows.

Next, suppose that  $b < c < d$ . Then, the key point to observe is that the set of strict Nash equilibria and hence the set of absorbing states is simply  $\bar{S} = S^\alpha$ . This immediately establishes the result for this case. Finally, similar considerations apply to the case where  $c > d$ , in which case Propositions 3.2 and 3.3 indicate that  $\bar{S} = \{g^e\} \times A^n$ .  $\triangle$

In our analysis we have not placed any restrictions on the number of links a player can form and, in equilibrium, the nature of interaction is ‘global’. This has the implication that transitions from one strict Nash equilibrium to another require a number of mutations which is a proportion of the total number of players. As is well known, for large populations this implies that the rate of convergence will be slow. In section below we discuss the possibility that players might be limited in the number of links they can support, a modelling feature which would have a significant bearing on this issue.

## 4 Discussion of the assumptions

In this section, we discuss the main assumptions underlying the analysis. We do so by addressing in turn a number of variations of the basic model that highlight the role played by some of its key features. Due to space constraints, we will not present the different models in detail, nor therefore state formally and prove the results we have obtained. However, a full account of the proofs is available from the authors upon request.

### 4.1 The nature of link formation

An important aspect of our model is that link formation is *one-sided*. From a methodological point of view this formulation has the advantage that it allows us to study the social process of link formation and coordination as a non-cooperative game; from a substantive viewpoint this formulation is interesting since it allows for an explicit consideration of the role of active and passive links. To clarify this assumption, we explore in turn two variations of the model. In the first one, we maintain the feature that any link must be unilaterally decided by a particular player, but suppose that she alone derives the benefits of it. That is, the cost of links and the flow of benefits are both one-sided, which makes passive links payoff-irrelevant. In the second variation, the links are two-sided in the sense that both agents involved must express the desire to form it and bear an equal share of the linking cost. However, we keep the feature that such decisions are *independently* adopted by each player, and therefore the link-formation process can be modelled as a non-cooperative game.

*One-sided active links:* We have carried out a complete analysis of this model, whose main findings and implications may be briefly summarized as follows.

First, in the static setting, we have shown that the only (non-empty) network that can arise is the complete network. Moreover, in this network, everyone chooses the same action, thus social conformism obtains. These results demonstrate that coexistence of different conventions and the possibility of an unconnected society arises in our basic model solely due to the presence of (payoff-relevant) passive links. Next we note that a complete network in the present setting with only active links is one in which every player forms a link with every other player. Thus there is only one possible strategy configuration that can support the complete network. The only multiplicity that remains concerns the choice of action: both outcomes, everyone choosing action  $\alpha$  and everyone choosing action  $\beta$ , are possible in equilibrium.

Our analysis of the (stochastic) dynamics reveals that *there is a cut-off value for the cost of forming links  $\hat{c} \in (e, f)$ , such that for all  $c < \hat{c}$ , the risk-dominant action  $\beta$  prevails, while for all  $c > \hat{c}$ , the efficient action  $\alpha$  prevails*. It is worth noting that the cut-off level of costs  $\hat{c}$  with only active links is lower than the cut-off level of costs  $\bar{c}$  that arises in the presence of active and passive links. Thus passive links have the effect of making the risk-dominant action more likely. The intuition behind this finding is as follows. In our proof of Theorem 3.1 we showed that transition across equilibria is easiest, in terms of ‘the number of mutations required’, when the pattern of link formation is very asymmetric. This creates the maximum scope for passive links to act as a bridge to induce other players to switch actions. This construction also reveals why the risk-dominant strategy is favored in a setting with passive links. Passive links allow the connectedness of the network to be sustained, without costs being incurred by the recipients of the links. In a situation where players are choosing different actions, this has the effect of creating greater strategic uncertainty. And, as is well known, such strategic uncertainty acts in favor of the risk-dominant action, which in turn helps explain why the risk-dominant action prevails under a wider range of cost conditions once passive links are allowed.

*Two-sided links through independent decisions:* We next discuss the case where a link is formed if, and only if, both parties involved offer to form a link, in which case the linking cost is divided equally.<sup>21</sup> We have also analyzed this model fully. The findings are broadly in line with the results of the one-sided active links model. In the static setting, we show that the only (non-empty) strict Nash network is the complete network. Moreover, in this network, everyone chooses the same action, thus social conformism obtains. The only multiplicity that remains again concerns the choice of action: both outcomes, everyone choosing action  $\alpha$  and everyone choosing action  $\beta$ , are possible in equilibrium. Our analysis of the (stochastic) dynamics reveals that *there is a cut-off value for the cost of forming links  $\hat{c} \in (e, f)$  such that, for all  $c < \hat{c}$ , the risk-dominant action  $\beta$  prevails, while for all  $c > \hat{c}$  the efficient action  $\alpha$  prevails*. This cut-off value

---

<sup>21</sup>In this setting we need to make some assumption about the payoff implications of unreciprocated links. In our analysis, we assume that unreciprocated links involve no costs and yield no benefits either.

is, somewhat surprisingly, identical to the cut-off value in the one-sided active links model. It is worth noting, however, that in the two-sided setting the cost of forming a link is  $2c$  and therefore the precise value of the cut-off level should be interpreted with care here. What remains, however, as the most interesting observation is the similarity in the qualitative features of the result: for low costs the risk-dominant action prevails, while for high costs the efficient action prevails.

## 4.2 Negative payoffs and link refusal

Throughout the analysis of the basic model we have maintained the assumption that the (gross) payoffs to be earned by playing the bilateral game are all positive. This justifies the formulation that, at a zero linking cost, no player should refuse a unilateral proposal to play this game. But if this cost were positive (and higher than some of the payoffs in the game) or the game payoffs themselves were negative, such a one-sided approach could be hardly defended as a meaningful or plausible model of network formation.

To address this issue, we have studied an extension of the basic model where, after any player receives a unilateral link proposal, this player may object to forming the link. It is posited, in particular, that at the stage of possible refusal, the actions chosen by both players are already irreversibly fixed and therefore the agent in question can safely evaluate whether it is profitable or not to accept the proposal. As mentioned, this formulation may accommodate the case where passive links are costly and/or the payoffs of the bilateral game display some negative payoffs. And, of course, when passive links are costless and gross payoffs positive, the basic model studied in this paper follows from that general (one-sided) framework of network formation as a particular case.

In the context outlined, the main point of our analysis is that efficient state is selected for large enough costs of forming links. The exact value of the threshold is not generally the same as in Theorem 3.1. It always lies, however, between the lowest equilibrium payoff (i.e. payoff  $b$  in Table I) and the lowest payoff in the game (payoff  $e$ ).

## 4.3 Indirect links

In the real world, social networks are far from complete. This happens because there are a number of factors that limit the "linking capacity" of agents and also because indirect connections often facilitate transactions and make complete networks unnecessary. Motivated by these latter considerations, we have explored the role of *indirect linkages* that facilitate transactions between players. As before, the focus is on the architecture of stable networks and the influence of link formation on the behavior of players.

Consider a model in which two players can play a game if there is a path between them (recall Subsection 2.1). Given a network  $g$  and any two players,  $i$  and  $j$ , let us write  $i \xrightarrow{g} j$  when a path between them exist. Then, we may define the

*indirect* neighborhood of a player  $i$  by  $\hat{N}(i; g) \equiv \{j \in N : i \xleftrightarrow{g} j\}$ , i.e. the set of players with whom player  $i$  has a path in the network  $g$ . With this notation in place, the payoff to a player  $i$  from choosing some strategy  $s_i = (g_i, a_i)$  when other players choose  $s_{-i} = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$  is given by:

$$\hat{\Pi}_i(s_i, s_{-i}) = \sum_{j \in \hat{N}(i; \bar{g})} \pi(a_i, a_j) - \nu(i; g) \cdot c, \quad (11)$$

where recall that  $\nu(i; g) \equiv |N(i; g)|$  is the cardinality of the set of *direct* links established by player  $i$ .

In this setting, we find that *the unique stochastically stable architecture is the minimally-connected network that we call a center-sponsored star* (as in Figure 1b). We also find that there exists a critical cost of forming links, such that, for costs below this level, players coordinate on the risk-dominant action, while for linking costs above this level they coordinate on the efficient action. Thus, comparing matters with the basic model, we find a similar qualitative conclusion concerning the selection of efficiency versus risk dominance although, naturally, the specific network architecture that underlies players' interaction is very different. A detailed proof of these results can be found in our earlier working paper, Goyal and Vega-Redondo (1999).

#### 4.4 The costs of forming links

There are two types of assumptions we make on the costs of forming links: the first assumption is that everyone has the same costs of forming links, while the second assumption is that these costs are linear in the number of links. We briefly discuss each of these assumptions now.

Suppose some players have lower costs of forming links as compared to others. To fix ideas let there be two cost levels, high  $c_2$  and low  $c_1$  (i.e.  $c_2 > c_1$ ). Let  $\rho$  be the fraction of players that have high cost. Clearly, our results from the basic model carry over directly if  $c_2 < e$  or if  $c_1 > f$ . We therefore focus our attention on the intermediate values case,  $c \in (e, f)$ . Suppose  $b < f$  and let  $\bar{c} \in (e, b)$  be the threshold cost derived in Theorem 3.1. Define  $\lambda = \frac{b-e}{b-e+d-f} > 1/2$ . Then, for large  $n$ , it may be shown that  $\hat{S} = S^\alpha$  if  $\rho > 2\lambda - 1$ . It is worth noting that for  $\lambda$  close to  $1/2$  this requirement is weak and is satisfied for fairly small values of  $\rho$ . This result therefore suggests an interesting implication of cost-heterogeneity: consider a population with only low cost type players where  $c_1 < \bar{c}$  and suppose that  $\lambda$  is close to  $1/2$ . Then the stochastically stable action is  $\beta$ . Now introduce a relatively small number  $((2\lambda - 1)n)$  of high cost players in this population. Then, the aforementioned result indicates that the stochastically stable action changes to  $\alpha$ !

In our basic model it is assumed that costs of forming links are linearly increasing in number of links. This assumption has the implication that either a player

wishes to form no links or is willing to form up to  $n - 1$  links. In effect, therefore there is no constraint to link formation. It is certainly more realistic to assume that players are constrained in the number of links they can support due to time and resource constraints. Under this assumption, stable networks will generally be incomplete and possibly partially connected. We feel that a model with a limited number of individual links may also be more amenable to weaker assumptions concerning information on the network and the action profiles of players, issues which are of course interesting in their own right. For an analysis of the implications of limited links in the context of the two-sided links model, the reader is referred to Jackson and Watts (2002).

## 4.5 The model of mutations

In the basic model, we assume that players choose actions as well as their links at random with some small probability (in other words, there is a small probability of mutation affecting all the different strategy components). It can be argued, however, that links are more durable and substantial objects and perhaps less subject to this random choice. And from a technical point of view, it is also a matter of some interest whether our main insights are robust to this alternative specification of random choice. To explore these issues, in this section we discuss a model in which the probability of mutation on the links component of the strategy is set equal to zero.

We start by noting that the possibility of mutations in links is used in the proof of Theorem 4.1 only at one point: when we show that the sets  $S^\alpha$  and  $S^\beta$  are recurrent. In this context, the mutations in links play a crucial role as they allow for a one-step mutation across different complete networks. Our idea is to replicate this transition indirectly via the mutations in actions only. This idea is perhaps simplest to see if the costs of linking  $c \in (f, b)$ . Consider a complete network  $s = (g, a) \in S^\alpha$  and suppose that  $g_{i,j} = 1$ . We wish to transit to  $s' = (g', a) \in S^\alpha$  in which  $g'_{j,i} = 1$ . The ‘indirect’ transition works as follows: Let player  $j$ ’s action mutate from  $\alpha$  to  $\beta$ . Then let player  $i$  move and choose a best response. Her best-response is to persist with action  $\alpha$  (since  $n$  is large) and delete the link with  $j$  (because  $c > f$ ). Now get player  $j$  to move: her best response is to switch action from  $\beta$  back to  $\alpha$  (since  $n$  is large) and form a link with player  $i$  (since  $c < d$ ). This argument is general and also applies to  $s \in S$ .

We now note that this indirect route of transition between different complete networks in  $S^\beta$  does not really work if  $e < c < f$ . This is because, the second step in the above argument does not go through and player  $i$  (choosing action  $\beta$ ) would not find it optimal to delete the link with player  $j$  (with action  $\alpha$ ). One way out is to extend the myopic best-response decision rule to a better-response rule where a player places positive probability on all actions that yield a payoff as high as the payoff from the current strategy. This reflects the same spirit as best-response adjustment but is sufficient for our purposes as it allows

to construct the following indirect single step transition. First, suppose that player  $i$ 's action mutates from  $\beta$  to  $\alpha$ . Then, let player  $i$  move and choose an at-least-as-good-as-response in which she persists with action  $\alpha$  but deletes the link with  $j$ . This strategy increases her payoff as compared to the current strategy, since  $c > e$ . Now get player  $j$  to move: her best response is to persist with action  $\beta$  (since  $n$  is large), and form a link with player  $i$ , since  $c < f$ . Finally, allow player  $i$  to move and her best response is to switch back from action  $\alpha$  to  $\beta$ . This completes the one-step transition.

Thus, as explained above, suppose that mutations only affect the actions of players and that players choose with positive probability any strategy that yields (weakly) higher payoffs as compared to current strategy. Let  $\bar{c}$  be the cut-off value considered in Theorem 3.1. Then, using the above arguments, it is not difficult to show the following result: *If  $c < \bar{c}$  ( $\bar{c} < c$ ) then  $\beta$  ( $\alpha$ ) is the unique stochastically stable action, for large enough  $n$ .*

#### 4.6 Simultaneity of actions and links

We assume that players can form links and choose actions *simultaneously*. In independent work, Droste, Johnson and Gilles (1999), and Jackson and Watts (1999) also study a two-sided link model where directly connected agents play a bilateral game, but allow for players to choose links and actions in *separate* revision stages. They find that if the costs of link formation are high, all those states where players choose a common action are stochastically stable, i.e. *either* of the two actions may obtain. The contrast with our analysis arises out of the assumption that individuals choose links and actions at separate stages, i.e. players choose links taking actions as given while they choose actions taking the links as a given. Instead, in our setting, any individual undertaking a revision is allowed to modify *both* her action and her supported links. Our arguments pertaining to alternative models of link formation clarify that it is this simultaneity of actions and link formation decisions and not the one-sided nature of link formation that is critical for the difference in the results. This would motivate an examination of the effects of varying levels of flexibility in the two choice dimensions, links and actions – for example, one could allow for the possibility that link revision might be more rigid than action change.

## 5 Conclusion

We develop a simple model to study the interaction between partner choice and individual behavior in games of coordination. We suppose that two players can play a game only if they have a link between them. Our analysis shows that individual attempts to balance the costs and benefits of link formation yield a unique network. We also find, however, that the *dynamics* of network

evolution have a powerful impact on the nature of social coordination: at low costs of forming links, individuals coordinate on the risk-dominant action, while for high costs of forming links individuals coordinate on the efficient action. These findings are robust to modifications in the link formation process, different specifications of link formation costs, alternative models of mutations as well as the possibility of interaction among indirectly connected players.

## 6 Appendix

**Proof of Proposition 3.1:** The proof of part (a) follows directly from the fact that  $c < f$  and is omitted. We provide a proof of part (b). In this case  $f < c < b$ . We first show that  $a_i = a_j = a$ , if  $i, j$  belong to the same component. Suppose not. If  $\bar{g}_{ij} = 1$ , then it follows that the player forming a link can profitably deviate by deleting the link, since  $c > f$ . Similar arguments apply if  $i$  and  $j$  are indirectly connected. We next show that if  $i \in g'$  and  $j \in g''$ , where  $g'$  and  $g''$  are two components in an equilibrium network  $g$ , then  $a_i \neq a_j$ . If  $a_i = a_j$  then the minimum payoff to  $i$  from playing the coordination game with  $j$  is  $b$ . Since  $c < b$ , player  $i$  gains by forming a link, i.e. choosing  $g_{ij} = 1$ . Thus  $g$  is not an equilibrium network. The final step is to note that since there are only two actions in the coordination game, there can be at most two distinct components. We note that the completeness of each component follows from the assumption that  $c < b$ .

We next prove part (c). There are two subcases to consider:  $c > \max\{b, f\}$  or  $b < c < f$ . (Note, of course, that the former subcase is the only one possible if  $b > f$ .) Suppose first that  $c > \max\{b, f\}$ , and let  $g$  be an equilibrium network which is non-empty but also incomplete. From the above arguments in (b), it follows that if  $\bar{g}_{i,j} = 1$ , then  $a_i = a_j = \alpha$ . Moreover, if  $a_j = \beta$ , then player  $j$  can have no links in the network. (These observations follow directly from the hypothesis that  $c > \max\{b, f\}$ .) However, since  $g$  is assumed incomplete, there must exist a pair of agents,  $i$  and  $j$ , such that  $\bar{g}_{ij} = 0$ . First, suppose that  $a_i = a_j = \alpha$ . Then, since  $c < d$ , it is clearly profitable for either of the two players to deviate and form a link with the other player. Suppose next that  $a_i = a_j = \beta$ . Then, players  $i$  and  $j$  can have no links and, furthermore, since  $g$  is non-empty, there must be at least two other players  $k, l \in N$  such that  $a_k = a_l = \alpha$ . But then player  $i$  can increase her payoff by choosing action  $\alpha$  and linking to player  $k$ . Finally, consider the case where  $a_i \neq a_j$  and let player  $i$  choose  $\beta$ . Then, if this player deviates to action  $\alpha$  and forms a link with player  $j$  she increases her payoff strictly. We have thus shown that  $\bar{g}_{ij} = 0$  cannot be part of an equilibrium network. This proves that a non-empty but incomplete network cannot be an equilibrium network in the first sub-case considered.

Consider now the case  $b < c < f$  and suppose, for the sake of contradiction, that  $g$  is an equilibrium network which is non-empty but incomplete. Since  $b < c < d$ , it follows directly that not every player chooses action  $\alpha$  or  $\beta$ . Moreover, in the mixed configuration, all the players who choose  $\alpha$  are directly

linked (since  $c < d$ ), there is a link between every pair of players who choose dissimilar actions (since  $c < f$ ), but there are no links between players choosing  $\beta$  (since  $b < c$ ). But then it follows that every player choosing  $\beta$  can increase her payoff by switching to action  $\alpha$ . This contradicts the hypothesis that the mixed configuration is an equilibrium. This completes the argument for part (c).

Part (d) is immediate from the hypothesis that  $c > d$ .  $\triangle$

**Proof of Proposition 3.2:** We start proving Part (a). In view of Part (a) of Proposition 3.1 and the fact that the underlying game is of a coordination type, the inclusion  $S^\alpha \cup S^\beta \subset S^*$  is obvious. To show the converse inclusion, take any profile  $s$  such that the sets  $A(s) \equiv \{i \in N : a_i = \alpha\}$  and  $B(s) \equiv \{j \in N : a_j = \beta\}$  are both non-empty. We claim that such a  $s$  cannot be an equilibrium. Assume, for the sake of contradiction, that such a state  $s$  is a Nash equilibrium of the game and denote  $u \equiv |A(s)|$ ,  $0 < u < n$ . Recall from Proposition 3.1 that every Nash network in this parameter range is complete. This implies that for any player  $i \in A(s)$ , we must have:

$$(u-1)d + (n-u)e - \nu(i;g) \cdot c \geq (u-1)f + (n-u)b - \nu(i;g) \cdot c \quad (12)$$

and for players  $j \in B(s)$ :

$$(n-u-1)b + uf - \nu(j;g) \cdot c \geq (n-u-1)e + ud - \nu(j;g) \cdot c. \quad (13)$$

It is easily verified that (12) and (13) are incompatible.

Now, we turn to Part (b). The inclusion  $S^\alpha \cup S^\beta \subset S^*$  is trivial. To show that the inclusion is strict for large enough  $n$ , consider a state  $s$  where both  $A(s)$  and  $B(s)$ , defined as above, are both non-empty and complete components. Specifically, focus attention on those configurations that are symmetric within each component, so that every player in  $A(s)$  supports  $\frac{u-1}{2}$  links and every player in  $B(s)$  supports  $\frac{n-u-1}{2}$  links. (As before,  $u$  stands for the cardinality of  $A(s)$  and we implicitly assume, for simplicity, that  $u$  and  $n-u$  are odd numbers.) For this configuration to be a Nash equilibrium, we must have that the players in  $A(s)$  satisfy:

$$d(u-1) - \frac{u-1}{2}c \geq f\frac{u-1}{2} + b(n-u) - c(n-u) \quad (14)$$

where we use the fact that, in switching to action  $\beta$ , any player formerly in  $A(s)$  will have to support herself all links to players in  $B(s)$  and will no longer support any links to other players in  $A(s)$  – of course, she still anticipate playing with those players from  $A(s)$  who support links with him.

On the other hand, the counterpart condition for players in  $B(s)$  is:

$$(n-u-1)b - \frac{n-u-1}{2}c \geq du + e\frac{n-u-1}{2} - cu \quad (15)$$

where, in this case, we rely on considerations for players in  $B(s)$  that are analogous to those explained before for players in  $A(s)$ . Straightforward algebraic

manipulations show that (14) is equivalent to:

$$\frac{u}{n} \geq \frac{1}{n} \frac{2d - c - f}{2b + 2d - 3c - f} + \frac{2(b - c)}{2b + 2d - 3c - f} \quad (16)$$

and (15) is equivalent to:

$$\frac{u}{n} \leq \frac{1}{n} \frac{c + e - 2b}{2b + 2d - 3c - e} + \frac{2b - c - e}{2b + 2d - 3c - e}. \quad (17)$$

We now check that, under the present parameter conditions:

$$\frac{2b - c - e}{2b + 2d - 3c - e} > \frac{2(b - c)}{2b + 2d - 3c - f}. \quad (18)$$

Denote  $Y \equiv 2b - c$ ,  $Z \equiv 2b + 2d - 3c$ , and rewrite the above inequality as follows:

$$\frac{Y - e}{Y - c} > \frac{Z - e}{Z - f} \quad (19)$$

which is weaker than:

$$\frac{Y - e}{Y - f} > \frac{Z - e}{Z - f} \quad (20)$$

since  $c > f$ . The function  $\zeta(z) \equiv \frac{z - e}{z - f}$  is uniformly decreasing in  $z$  since  $b > f > e$ . Therefore, since  $Y < Z$ , (20) obtains, which implies (19). Hence it follows that, if  $n$  is large enough, one can find suitable values of  $u$  such that (16) and (17) jointly apply. This completes the proof of Part (b).

We now present the proof for part (c). We know from Proposition 3.1 that the complete and the empty network are the only two possible equilibrium networks. Since  $c > b > f > e$ , it is immediate that, in the complete network, every player must choose  $\alpha$  and this is a Nash equilibrium. Then note that, for the empty network to be an equilibrium, it should be the case that no player has an incentive to form a link. This implies that every player must choose  $\beta$ . On the other hand, it is easy to see that the empty network with everyone choosing  $\beta$  is a Nash equilibrium.

The proof of part (d) follows directly from the hypothesis  $c > \max\{d, b, f, e\}$ .  $\triangle$

**Proof of Proposition 3.3:** It is enough to show that, from any given state  $s^0$ , there is a finite chain of positive-probability events (bounded above zero, since the number of states is finite) that lead to a rest point of the best response dynamics.

Choose one of the two strategies, say  $\beta$ , and denote by  $B(0)$  the set of individuals adopting action  $\beta$  at  $s^0$ . Order these individuals in some pre-specified manner and starting with the first one suppose that they are given in turn the option to revise their choices (both concerning strategy and links). If at any given stage  $\tau$ , the player  $i$  in question does not want to change strategies, we set  $B(\tau + 1) = B(\tau)$  and proceed to the next player if some are still left. If none is

left, the first phase of the procedure stops. On the other hand, if the player  $i$  considered at stage  $\tau$  switches from  $\beta$  to  $\alpha$ , then we make  $B(\tau + 1) = B(\tau) \setminus \{i\}$  and, at stage  $\tau + 1$ , re-start the process with the first-ranked individual in  $B(\tau + 1)$ , i.e. *not* with the player following  $i$ . Clearly, this first phase of the procedure must eventually stop at some finite  $\tau_1$ .

Then, consider the players choosing strategy  $\alpha$  at  $\tau_1$  and denote this set by  $A(\tau_1) \equiv N \setminus B(\tau_1)$ . Proceed as above with a chain of unilateral revision opportunities given to players adopting  $\alpha$  in some pre-specified sequence, restarting the process when anyone switches from  $\alpha$  to  $\beta$ . Again, the second phase of the procedure ends at some finite  $\tau_2$ .

By construction, in this second phase, all strategy changes involve an increase in the number of players adopting  $\beta$ , i.e.  $B(\tau_2) \supseteq B(\tau_1)$ . Thus, if the network links affecting players in  $B(\tau_1)$  remain unchanged throughout, it is clear that no player in this set would like to switch to  $\alpha$  if given the opportunity at  $\tau_2 + 1$ . However, in general, their network links will also evolve in this second phase, because individual players in  $A(\tau_1)$  may form or delete links with players in  $B(\tau_1)$ . In principle, this could alter the situation of individual members of  $B(\tau_1)$  and provide them with incentives to switch from  $\beta$  to  $\alpha$ . It can be shown, however, that this is not the case. To show it formally, consider any given typical individual in  $B(\tau_1)$  and denote by  $\hat{r}^h$ ,  $h = \alpha, \beta$ , the number of links received (but not supported) by this player from players choosing action  $h$ . On the other hand, denote  $\hat{u} \equiv |A(\tau_1)|$ . Then, since the first phase of the procedure stops at  $\tau_1$ , one must have:

$$\begin{aligned} & \max_{q^\alpha, q^\beta} b(q^\beta + \hat{r}^\beta) + f(q^\alpha + \hat{r}^\alpha) - c(q^\alpha + q^\beta) \\ & \geq \max_{q^\alpha, q^\beta} e(q^\beta + \hat{r}^\beta) + d(q^\alpha + \hat{r}^\alpha) - c(q^\alpha + q^\beta) \end{aligned} \quad (21)$$

for all  $q^\alpha, q^\beta$  such that  $0 \leq q^\alpha \leq \hat{u} - \hat{r}^\alpha$ ,  $0 \leq q^\beta \leq n - \hat{u} - 1 - \hat{r}^\beta$ . Now denote by  $\tilde{r}^h$  and  $\tilde{u}$  the counterpart of the previous magnitudes ( $\hat{r}^h$  and  $\hat{u}$ ) prevailing at  $\tau_2$ . We now show that  $\tilde{u} \leq \hat{u}$ ,  $\tilde{r}^\alpha \leq \hat{r}^\alpha$ , and  $\tilde{r}^\beta \geq \hat{r}^\beta$ . First, we note that  $\tilde{u} \leq \hat{u}$  by construction of the process. Next note that if  $\tilde{r}^\alpha > \hat{r}^\alpha$  then this implies that some player who chooses action  $\alpha$  has formed an additional link with player  $i$  in the interval between  $\tau_1$  and  $\tau_2$ . This is only possible if  $c < e$ . It also implies that player  $i$  did not have a link with this player at  $\tau_1$ . This is only possible if  $c > f$ , a contradiction. Thus  $\tilde{r}^\alpha \leq \hat{r}^\alpha$ . Finally note that  $\tilde{r}^\beta \geq \hat{r}^\beta$  follows from the fact that all the players choosing  $\beta$  at  $\tau_1$  do not revise their decisions in the interval between  $\tau_1$  and  $\tau_2$ .

Therefore, (21) implies:

$$\begin{aligned} & \max_{q^\alpha, q^\beta} b(q^\beta + \tilde{r}^\beta) + f(q^\alpha + \tilde{r}^\alpha) - c(q^\alpha + q^\beta) \\ & \geq \max_{q^\alpha, q^\beta} e(q^\beta + \tilde{r}^\beta) + d(q^\alpha + \tilde{r}^\alpha) - c(q^\alpha + q^\beta) \end{aligned}$$

for all  $q^\alpha, q^\beta$  such that  $0 \leq q^\alpha \leq \tilde{u} - \tilde{r}^\alpha$ ,  $0 \leq q^\beta \leq n - \tilde{u} - 1 - \tilde{r}^\beta$ . This allows us to conclude that the concatenation of the two phases will lead the process to a rest point of the best response dynamics, as desired.  $\triangle$

**Proof of Lemma 3.1:** The proof is constructive. Let  $s \in S^h$ ,  $h = \alpha, \beta$ , and order in some arbitrary fashion all other states in  $S^h \setminus \{s\}$ . Also order in some discretionary manner all pairs  $(i, j) \in N \times N$  with  $i \neq j$ . For the first state in  $S^h \setminus \{s\}$ , say  $s_1$ , proceed in the pre-specified sequence across pairs  $(i, j)$  reversing the links of those of them whose links are different from what they are in  $s$  one at a time. This produces a well-defined path joining  $s_1$  to  $s$ , whose constituent states define a set denoted by  $Q_1$ . Next, consider the highest ranked state in  $S^h \setminus Q_1$ , say  $s_2$ . Proceed as before, until state  $s_2$  is joined to either state  $s$  or a state already included in  $Q_1$ . Denote the states included in the corresponding path by  $Q_2$ . Clearly, when a stage  $n$  is reached such that  $S^h \setminus (\cup_{\ell=1}^n Q_\ell) = \emptyset$ , the procedure described has fully constructed the desired  $s$ -tree restricted to  $S^h$ .  $\triangle$

**Proof of Lemma 3.2:** Let  $s^\alpha$  and  $s^\beta$  be generic states in  $S^\alpha$  and  $S^\beta$ , respectively. We want to determine the minimum number of mutations needed to transit across a pair of them in either direction.

(1). First, consider a transition from  $s^\beta$  to  $s^\alpha$  and let  $k$  be the number of mutations triggering it. If this transition is to take place *via* the best-response dynamics after those many mutations, there must be some player currently choosing  $\beta$  (i.e. who has *not* mutated) that may then voluntarily switch to  $\alpha$ . As before, denote by  $q^h$  the number of active links this player supports to players choosing  $h$  ( $h = \alpha, \beta$ ) and let  $r^h$  stand for the number of passive links she receives from players choosing  $h$  ( $h = \alpha, \beta$ ). The payoff from choosing  $\alpha$  for that player is given by:

$$\pi_\alpha = r^\alpha d + r^\beta e + q^\alpha(d - c) + q^\beta(e - c). \quad (22)$$

On the other hand, the payoff to choosing  $\beta$  is given by:

$$\pi_\beta = \hat{r}^\alpha f + \hat{r}^\beta b + \hat{q}^\alpha(f - c) + \hat{q}^\beta(b - c), \quad (23)$$

where  $\hat{q}^h$  and  $\hat{r}^h$  have the same interpretation of active and passive links as before, now associated to the possibility that the player chooses  $\beta$ . Clearly, we have  $\hat{q}^h = q^h$  and  $\hat{r}^h = r^h$  for each  $h = \alpha, \beta$ . Concerning the passive links, this is immediate; for active links, it follows from the fact that, since  $c < e$ , a player will want to create links to all unconnected players, independently of what they do. Analogous considerations also ensure that (i).  $r^\alpha + q^\alpha = k$  and (ii).  $r^\beta + q^\beta = n - k - 1$ . Thus, in sum, for a transition from some state in  $S^\beta$  to a state in  $S^\alpha$  to be triggered, one must have:

$$\begin{aligned} \pi_\alpha - \pi_\beta &= (r^\alpha + q^\alpha)(d - f) - (r^\beta + q^\beta)(b - e) \\ &= k(d - f) - (n - k - 1)(b - e) \geq 0 \end{aligned}$$

Let  $m^{\beta, \alpha}$  stand for the minimum number of mutations which lead to such a transition. The above considerations imply that

$$m^{\beta, \alpha} \geq \frac{b - e}{(d - f) + (b - e)}(n - 1), \quad (24)$$

which gives us the minimum number of mutations that are *necessary* for a transition from *any* state  $s^\beta$  to some  $s^\alpha$ . However, denoting by  $\lceil z \rceil$  the smallest integer no smaller than  $z$ , suppose that the *strategies* of  $\lceil \frac{b-e}{(d-f)+(b-e)}(n-1) \rceil$  players undergo a simultaneous mutation from any particular state  $s^\beta$  (i.e. these players maintain their links but switch from  $\beta$  to  $\alpha$ ). Thereafter, the repeated operation of the best-response dynamics is sufficient to induce a transition to a state  $s^\alpha$ . Thus the necessary number of mutations computed above is also sufficient to induce a transition from any  $s^\beta$  to some  $s^\alpha$ . That is, the inequality in (24) holds with equality.

(2). Consider on the other hand, the transition  $s^\alpha$  to  $s^\beta$ . Using the expressions (22) and (23), we can deduce that the minimum number of mutations  $m^{\alpha,\beta}$  needed to transit from some state in  $S^\alpha$  to a state in  $S^\beta$  satisfies:

$$m^{\alpha,\beta} \geq \frac{d-f}{(d-f)+(b-e)}(n-1). \quad (25)$$

As in the first case, this gives us the minimum number of mutations needed for a transition. However, consider any state  $s^\alpha$  and suppose that the strategies of  $\lceil \frac{d-f}{(d-f)+(b-e)}(n-1) \rceil$  players undergo a simultaneous mutation (i.e. they maintain their links but switch from  $\alpha$  to  $\beta$ ). It again follows that the operation of the best-response dynamics suffices to induce a transition to a state  $s^\beta$ . That is, (25) holds with equality.

To conclude, simply note that, if  $n$  is large enough,

$$\left\lceil \frac{b-e}{(d-f)+(b-e)}(n-1) \right\rceil < \left\lceil \frac{d-f}{(d-f)+(b-e)}(n-1) \right\rceil$$

since  $d-f < b-e$ . △

**Proof of Lemma 3.4 (Sketch):** The proof proceeds in the same way as the proof of Lemma 3.3. We therefore only spell out the main computations.

(1). First, consider transitions from state  $s^\beta$  to state  $s^\alpha$  and let  $k$  be the number of mutations triggering it. We focus on a player currently choosing  $\beta$  and aim at finding the most favorable (i.e. least mutation-costly) conditions that would induce him to switch to  $\alpha$ . Along the lines explained in the proof of Lemma 3.3, this leads to the following lower bound:

$$m^{\beta,\alpha} \geq \frac{b-c}{(d-f)+(b-c)}(n-1) \equiv H, \quad (26)$$

which again can be seen to be tight in the sense that, in fact,  $m^{\beta,\alpha} = \langle H \rangle$  – recall that  $\lceil z \rceil$  stands for the smallest integer no smaller than  $z$ .

(2). Analogous considerations for a transition from state  $s^\alpha$  to state  $s^\beta$  leads to the lower bound

$$m^{\alpha,\beta} \geq \frac{d-c}{(b-e)+(d-c)}(n-1) \equiv H', \quad (27)$$

which is also tight, i.e.  $m^{\beta,\alpha} = \lceil H' \rceil$ .

(3). Finally, to study how the sign of  $m^{\beta,\alpha} - m^{\alpha,\beta}$  changes for large  $n$  as a function of  $c$ , note that

$$H - H' \equiv \Delta(c) = \frac{(b-c)(b-e) - (d-f)(d-c)}{[(d-f) + (b-c)][(b-e) + (d-c)]} (n-1). \quad (28)$$

Observe that the denominator of  $\Delta(c)$  is always positive, the numerator is decreasing in  $c$ , and is moreover negative at  $c = b$ . This completes the proof.  $\triangle$

**Proof of Lemma 3.5:** Fix some  $s \in S^{\alpha,\beta}$ , with the players  $A(s)$  and  $B(s)$  of the  $\alpha$  and  $\beta$  components displaying respective cardinalities  $|A(s)| \equiv u > 0$  and  $|B(s)| \equiv n - u > 0$ , respectively. To address the first part of the lemma, suppose that a player  $i \in B(s)$  experiences a mutation, which has the effect of switching her action from  $\beta$  to  $\alpha$  and the deletion of all her links with players in  $B(s)$ . Now consider the players in the set  $B(s) \setminus \{i\}$ . There are two possibilities: either all of them wish to retain action  $\beta$ , or there is a player who wishes to switch actions.

In the former case, let all of them move and they will retain their earlier strategy except for one change: they will each delete their link with player  $i$ , since  $f < c < b$ . We now get players in  $A(s)$  to move and they all form a link with player  $i$ , since  $f < c < b < d$ . It may be checked that we have reached an equilibrium state  $s'$ , with  $A(s') \geq A(s) + 1$ .

Consider now the second possibility. Pick a player  $j \in B(s) \setminus \{i\}$ , who wishes to switch actions from  $\beta$  to  $\alpha$ . It follows that this player will delete all her links with players in  $B(s)$  and form links with all players in  $A(s)$  (since  $e < f < c < b < d$ ). We then examine the incentives of the players still choosing action  $\beta$ , i.e., players in the set  $B(s) \setminus \{i, j\}$ . If there are no players who would like to switch actions then we repeat step above and arrive at a new state with a larger  $\alpha$ -component. If there are players who wish to switch actions from  $\beta$  to  $\alpha$  then we get them to move one at a time. Eventually, we arrive at either a new state  $s' \in S^{\alpha,\beta}$ , or we arrive at a state  $s' \in S^\alpha$ .

In either case, we have shown that starting from a state  $s \in S^{\alpha,\beta}$ , we can move with a single mutation to a state  $s'$  such that  $A(s') \geq A(s) + 1$ . Since  $s \in S^{\alpha,\beta}$  was arbitrary, the proof is complete for the first part. The second conclusion concerning some new equilibrium state  $s''$  with  $|A(s'')| \leq |A(s)| - 1$  is analogous.  $\triangle$

## References

- [1] R. Aumann and R. Myerson (1989), Endogenous formation of links between players and coalitions: an application of the Shapley value, in A. Roth (eds) *The Shapley Value*. Cambridge University Press.

- [2] Albert, R., H. Jeong, and A.-L. Barabási (2000): “Error and attack tolerance of complex networks”, *Nature* **406**, 378-381.
- [3] Allen, F. and D. Gale (2000), Financial Contagion, *Journal of Political Economy*, **108**, 1, 1-33.
- [4] Bala, V. and S. Goyal (2000), A Non-Cooperative Model of Network Formation, *Econometrica*, **68**, 5, 1181-1231.
- [5] Bala V. and S. Goyal (1998), Learning from Neighbours, *Review of Economic Studies* **65**, 595-621.
- [6] Bhaskar, V. and F. Vega-Redondo (2003), Migration and the evolution of conventions, *Journal of Economic Behavior and Organization*, forthcoming.
- [7] Bliss, C. B. Nalebuff (1984): “Dragon-slaying and ballroom dancing: the private supply of a public good”, *Journal of Public Economics* **25**, 1-12.
- [8] Blume, L. (1993), Statistical Mechanics of Strategic Interaction, *Games and Economic Behavior* **4**, 387-424.
- [9] Blume, L. (1995): “Evolutionary equilibrium with forward-looking players”, working paper, Cornell University.
- [10] Canning, D. (1992), Average Learning in Learning Models, *Journal of Economic Theory*, **57**, 442-472.
- [11] Carlson, H. and E. van Damme (1993), Global Games and Equilibrium Selection, *Econometrica* **61**, 5, 989-1019.
- [12] Chwe, M. (2000), Communication and Coordination in Social Networks, *Review of Economic Studies*, **67**, 1, 1-17.
- [13] Coleman, J. (1966), *Medical Innovation: A Diffusion Study*, Second Edition, Bobbs-Merrill, New York.
- [14] Droste, E, R. Gilles, and K. Johnson (1999), Endogenous interaction and the evolution of conventions, mimeo, Tilburg University.
- [15] Dutta, B., A. van Den Nouweland and S. Tijs (1995), Link formation in cooperative situations, *International Journal of Game Theory*, **27**, 245-256.
- [16] Ellison, G. (1993), Learning, Local Interaction, and Coordination. *Econometrica* **61**, 1047–1071.
- [17] Ellison, G. and D. Fudenberg (1993), Rules of Thumb for Social Learning, *Journal of Political Economy* **101**, 612-644.
- [18] Ely, J. (2003), Local Conventions, *Berkeley Electronic Press Journals, Advances in Theoretical Economics*, 2,1,1.

- [19] Freidlin, M. I. and A. D. Wentzell (1984), *Random Perturbations of Dynamical Systems*. Translation of Freidlin and Venttsel' (1979) by J. Szücs. Springer-Verlag, New York.
- [20] Goyal, S. (1993), Sustainable Communication Networks, *Tinbergen Institute Discussion Paper*, 93/250.
- [21] Goyal, S. (1996), Interaction Structure and Social Change, *Journal of Institutional and Theoretical Economics*, **152**, 3, 472-495.
- [22] Goyal, S. and F. Vega-Redondo (1999), Learning, Network Formation, and Coordination, Working Paper, Erasmus University and University of Alicante.
- [23] Granovetter, M. (1974), *Getting a Job: A Study of Contacts and Careers*, Harvard University Press, Cambridge MA.
- [24] Granovetter, M (1985), Economic Action and Social Structure: The Problem of Embeddedness, *American Journal of Sociology* **3**, 481-510.
- [25] Haag, M. and R. Lagunoff (2000), Social Norms, Local Interaction, and Neighborhood Planning, *David Levine's Working Paper Archive 2049*, Department of Economics, UCLA.
- [26] Harsanyi, J.C. and R. Selten (1988), *A General Theory of Equilibrium Selection*, Cambridge, Mass., MIT Press.
- [27] Jackson, M. and A. Wolinsky (1996), A Strategic Model of Economic and Social Networks, *Journal of Economic Theory* **71**, 1, 44-74.
- [28] Jackson, M. and A. Watts (2002), On the formation of interaction networks in social coordination games, *Games and Economic Behavior*, 41, 2, 265-291.
- [29] Kandori, M. and G. J. Mailath and R. Rob (1993), Learning, Mutation, and Long Run Equilibria in Games. *Econometrica* **61**, 29-56.
- [30] Kirman, A. (1997), The Economy as an Evolving Network, *Journal of Evolutionary Economics* **7**, 339-353.
- [31] Lagunoff, R. and A. Matsui (1995): "Evolution in mechanisms for public projects", *Economic Theory* **6**, 195-223.
- [32] Lee, I. H. and A. Valentinyi (2000), Noisy Contagion without Mutation, *Review of Economic Studies*, **67**, 1, 17-47.
- [33] Mailath, G. Samuelson, L. and Shaked, A., (1994), Evolution and Endogenous Interactions, mimeo., Social Systems Research Institute, University of Wisconsin.
- [34] Morris, S. (2000), Contagion, *Review of Economic Studies*, **67**, 1, 57-79.

- [35] Oechssler, J (1997), Decentralization and the coordination problem, *Journal of Economic Behavior and Organization* **32**, 119-135.
- [36] Robson, A. and F. Vega-Redondo (1996), Efficient Equilibrium Selection in Evolutionary Games with Random Matching, *Journal of Economic Theory* **70**, 65-92.
- [37] Samuelson, L. (1994), Stochastic stability in games with alternative best replies, *Journal of Economic Theory* **64**, 35-65.
- [38] Skyrms, B. and R. Pemantle (2000), A dynamic model of social network formation, *Proceedings of the National Academy of Sciences* **97**, 16, 9340-9346.
- [39] Young, H. P. (1993), The Evolution of Conventions. *Econometrica* **61**, 57-84.